Nerding Out on Twitter: Fun, Patriotism and #Curiosity

Victoria Uren
Aston University
Aston Triangle
Birmingham, B4 7ET, UK
v.uren@aston.ac.uk

Aba-Sah Dadzie
The OAK Group, DCS,
The University of Sheffield.
Sheffield, S1 4DP
a.dadzie@dcs.shef.ac.uk

ABSTRACT
This paper presents an analysis of tweets collected over six days before, during and after the landing of the Mars Science Laboratory, known as Curiosity, in the Gale Crater on the 6th of August 2012. A sociological application of web science is demonstrated by use of parallel coordinate visualization as part of a mixed methods study. The results show strong, predominantly positive, international interest in the event. Scientific details dominated the stream, but, following the successful landing, other themes emerged such as fun, and national pride.

Categories and Subject Descriptors
J.4 [**Social and Behavioural Sciences**], K.4 [**Computers and Society**].

General Terms
Experimentation.

Keywords
Public Engagement with Science, Web Science, Parallel Coordinates Visualization, High-dimensional Visualization.

1. INTRODUCTION
Social media, including microposts, provide spaces for public engagement with science, which are open to researchers, professional communicators and the public. We are investigating methods and metrics for characterizing this growing public dialogue in order to understand what people think about science. In previous work, we have studied the occurrence of scientific terms from the UNESCO thesaurus [1] and public interest in meteor showers [2]. In this paper, we present a case of large-scale social media engagement, specifically, the Twitter stream surrounding the landing of the Curiosity Mars rover in August 2012. At this time the hashtags #curiosity and #MSL were trending on Twitter; this is unusual for scientific discussion which is often relatively low key [2]. The availability of samples on the scale of thousands offers us opportunities for analysis that is not possible with only hundreds.

Studying Curiosity allows us to look at a research area which has been in the focus of public attention for decades. Since the launch of Sputnik in 1957, space research has fired the public's imagination, been a source of national pride, and spun off a range of now familiar technologies from Velcro to Teflon. Public engagement during the space race era of the 1960s and 1970s was not necessarily founded on understanding of the technology or appreciation of the scientific agenda, but often on respect for science and perceptions of the place of space technology in the wider context of the Cold War [3].

WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

From a social science perspective the issue of message framing is important, i.e., the interpretation of storylines with the aim of facilitating communication by appeals to social concerns and interests. The case of climate science illustrates the power of positive use of framing. Nisbet reports how climate science had been debated on the grounds of scientific validity and, in the USA, people had become entrenched in partisan divisions [4]. He then shows how reframing the discussion in moral (particularly biblical) terms, helped bridge some of these divides and win new advocates for responsible behaviour to protect the environment.

Web science has much to offer social science, particularly for the analysis of social media. Yet, at the time of writing, the Social Science Research Network[1] eLibrary contains just 13 papers containing the phrase "web science" compared with 555 for "social media". Meyer and Schroeder [5] speak of the "silos" within social science occupied by fields such as Webometrics, which use methods closest to what would be considered "traditional" methods by researchers with a computer science background. Web scientists can meet social scientists halfway by demonstrating how web science methods can contribute to social science studies. The work reported here is a mixed methods study, in which we deploy text analysis for initial exploration of the data, content analysis to gain rich qualitative interpretation and visual analysis using parallel coordinates to give a more nuanced view of trends in language usage. The triangulation of these methods strengthens the validity of the results.

2. CURIOSITY
Curiosity, also known as MSL (Mars Science Laboratory), is a robot the size of a small car[2]. It has been designed to conduct geological survey work on the surface of Mars under direction by scientists working at NASA's Jet Propulsion Laboratory (JPL). It was launched on the 26th of November 2011 and landed on the surface of Mars on the 6th of August 2012.

Considerable publicity surrounded the landing, heightened by the history of failed landings on Mars. In the run up to the landing, the phrase "seven minutes of terror" emerged to describe the period of the spacecraft's descent through the Martian atmosphere, during which it had to slow down from 13,000 mph to zero. There was also speculation about the likely success of the Sky Crane, the extraordinary, jet powered apparatus that would gently lower Curiosity the last few meters of its descent. Following the successful landing there was celebration and excitement as the first images came back to earth.

With hindsight the landing was a success, but it was not obvious beforehand that public reaction would be so overwhelmingly

[1] http://papers.ssrn.com

[2] http://www.nasa.gov/mission_pages/msl

positive. The Curiosity rover is nuclear powered, containing 4.8kg of plutonium. The launch was not marked by protests like those that accompanied the Cassini and New Horizon probes, due perhaps to a better understanding of the public health risks [6], but the ethics of sending nuclear materials to other planets is still debatable[3]. Comments about government spending priorities might also be expected as Curiosity cost about $2.5bn at a time when the NASA budget has faced spending cuts [7].

3. RESEARCH QUESTIONS

In the study reported here, a mixed methods approach was used to address the following exploratory research questions:

1. How did tweeting and retweeting activity vary over the course of the days and hours around the Curiosity landing?
2. What message frames were commonly used in tweets?
3. What was the prevailing sentiment?

The remainder of the paper describes the three methods used with their individual results: text analysis, content analysis and multi-dimensional visualization using parallel coordinates. The conclusions are derived from a triangulated view of the three analytical approaches applied to the Curiosity Twitter stream.

The dataset for the work was harvested, using the Twitter API, from the public data stream, from 03-09 Aug 2012. The analysis reported in this paper focuses on the 241,748 tweets containing the filter term "#curiosity", which was trending on 06 Aug 2012. Tweets were processed and stored in a mySQL database to facilitate the selection of subsets of microposts for analysis.

4. TEXT ANALYSIS

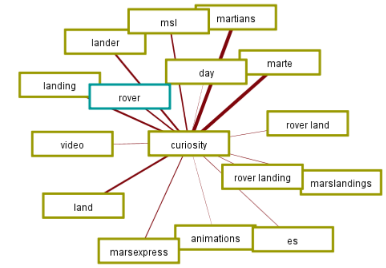
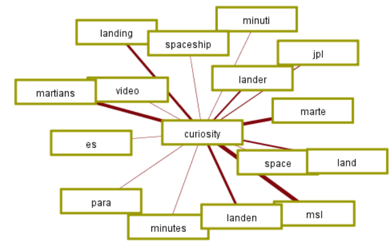
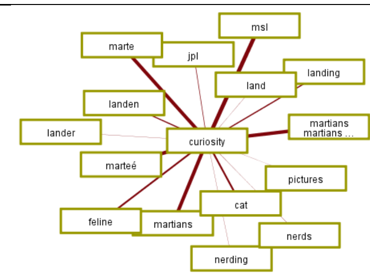
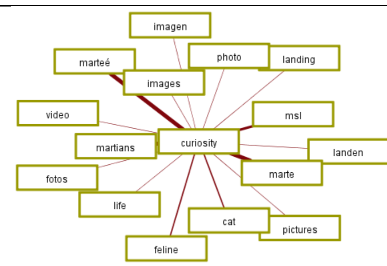
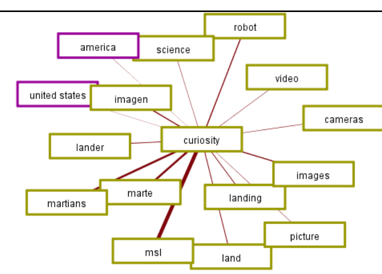
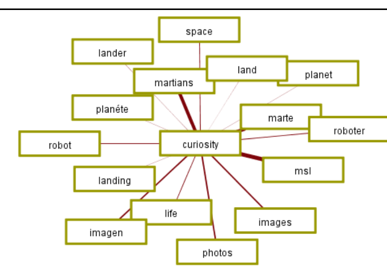
Text analysis was exploratory, with deeper analysis planned for future work. IBM's SPSS Modeller was used. Samples comprised all the tweets collected on each full GMT calendar day, from 04-09 Aug 2012, which contained the term #curiosity. One concept map was produced per day, centered on "curiosity".

4.1 Results

Table 1 shows how the topics discussed evolved over the days sampled. Some concepts, like "landing" and "lander" are present over the whole period, but other concepts change as time passes. Before the landing the term "animations" is presented, referring to a video which showed how the Sky Crane landing gear was intended to work, as well as "minutes", as in the "seven minutes of terror". On 06 Aug 2012 we see the emergence of the slang concept "nerding", sometimes as part of the phrase "nerding out". This expressed the excitement people felt and the pleasure derived from the event, with an emphasis on this being unusual because it was due to scientific activity. After the landing concepts such as "images", "photos" and "camera" emerged, as the first pictures taken by the rover started to come back to earth.

The concept maps also highlight frame-related concepts (see Figure 1). The science frame appears strong, with concepts such as "rover", "science", "life" and "robot". Politics emerges on 08 Aug, two days after the landing, with "america" and "usa", linked to expressions of patriotic feeling. Fun is also well represented, especially at the height of activity 06-07 Aug with jokes featuring the reaction of "martians" and the demise of a "cat".

Table 1: SPSS Concept Maps – 04-09 Aug 2012

<table>
<tr><td>04 Aug 2012
Frames:
Science</td><td>
</td></tr>
<tr><td>05 Aug 2012
Frames:
Science</td><td>
</td></tr>
<tr><td>06 Aug 2012
Frames:
Science
Fun</td><td>
</td></tr>
<tr><td>07 Aug 2012
Frames:
Science
Fun</td><td>
</td></tr>
<tr><td>08 Aug 2012
Frames:
Science
Politics</td><td>
</td></tr>
<tr><td>09Aug 2012
Frames:
Science</td><td>
</td></tr>
</table>

[3] http://www.space4peace.org/

5. CONTENT ANALYSIS

Content analysis can be defined as a qualitative "research technique for making replicable and valid references from texts" [8]. It comprises a sequence of procedures: selection of the texts to be coded, defining the unit of analysis and the codes to be used to address research questions, building a detailed coding manual providing definitions of codes and examples of their use, coding itself, and assessment of agreement between coders.

Samples of 200 random tweets were selected for the content analysis (randomisation was achieved using the SQL "order by rand()" statement). Six samples were taken, one for each day 04-09 Aug. We define the unit of analysis as a single tweet. Each tweet had to be assigned at least one code and could have a maximum of two (one sentiment and one frame).

The codes (Figure 1) were selected to address research questions 2 and 3. For the analysis of frames the schema proposed by Schäfer [9] was used, providing four codes: Scientific, Political, Economic, ELSI (Ethical, Legal and Social). To these a fifth code was added to reflect the light-hearted nature of many posts on Twitter, which was labeled "Fun". For the analysis of sentiment three standard codes were used: Positive, Negative and Neutral. Neutral was required in order to be certain that at least one code could be applied to any tweet that concerned Curiosity. Finally, two categories were created to cover all tweets that could not be categorized using a frame or sentiment code. The code Off Topic was defined for tweets that used the term #curiosity but were not about the landing (the hashtag is also used in its more general sense). Other Languages was added because, although most of the tweets containing #curiosity in languages other than English were clearly about Curiosity the lander, interpreting the subtleties of framing in 140 character tweets that had been passed through automatic translation was too error prone. A benefit of this code was that it indicated international interest.

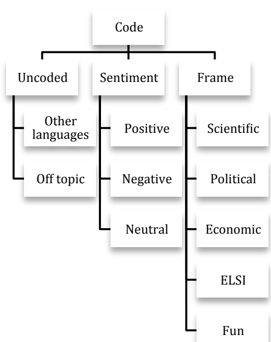


Figure 1: Taxonomy of Codes

The definitions in the coding manual were refined during an initial round of coding and finalized by agreement between the two coders. The tweets were then coded. Coder A coded the whole of each sample. Coder B coded a block of 50 tweets from each sample, a checking rate of 25% (over the total sample size of 1200 manually annotated tweets). Agreement between the two coders was calculated using Hooper's measure [10]:

$$\boldsymbol{H} = \frac{\boldsymbol{C}}{\boldsymbol{A} + \boldsymbol{B} - \boldsymbol{C}}$$

where C is the number of codes on which both coders agree, A is the number of codes assigned by coder A and B is the number of codes assigned by coder B. Agreement, reported in Table 2, was measured datewise (Hdate) and codewise (Hcode). Note that Cohen's Kappa [11] is inappropriate for this study as 1) the coding units, i.e. tweets, are not independent of each other because of retweeting, and 2) the codes are not mutually exclusive because both a frame and a sentiment code may be applied.

5.1 Results

The content analysis indicates that discussion of the Curiosity landing typically had positive sentiment, and that the predominant frames were Scientific (for example reporting landing times or the release of images), and Fun (there were many Martian jokes and references to parties taking place). Tweets in the Fun frame were lowest the day before the landing when anxiety about the "seven minutes of terror" was observed. Political tweets (most often expressing national pride, from both the USA and other nations with links to the Curiosity project) started to emerge after the successful landing. Very few Economic tweets were identified as comments on the cost of Curiosity were often classified as Political because they addressed other issues, e.g. "*Worth pointing out: both the #Olympics & #Curiosity are still cheaper than a month fighting in Iraq & Afghanistan*". Up to 50% of the tweets in the samples were in languages other than English, indicating strong international interest in the event.

Table 2: Number of codes applied by Coder A with Hooper agreement compared to Coder B

Date	4th	5th	6th	7th	8th	9th	Hcode
Positive	29	52	39	28	38	29	0.63
Negative	0	1	2	3	4	6	0.00
Neutral	1	1	13	3	2	1	0.00
Scientific	49	56	34	42	68	60	0.74
Political	1	4	13	11	6	5	0.50
Economic	1	0	0	1	0	0	0.00
ELSI	0	1	0	9	4	3	0.43
Fun	24	16	35	32	24	36	0.51
Off Topic	17	6	44	27	6	12	0.74
Other L.	100	95	44	70	79	80	0.98
Hdate	0.81	0.75	0.61	0.80	0.77	0.73	

Perfect agreement between the coders was not expected due to the inherent ambiguity of 140 character tweets and the subtlety of judgment required for some frames. Nonetheless, agreement was good for the datewise samples (Hdate). For individual codes (Hcode) it was good for the Scientific and Positive codes, but only moderate for Fun, possibly because humour is very personal (the coders had an eight year age difference and different cultural backgrounds). These initial results indicate that future content analysis with larger samples would be worthwhile. For this study, the validity of the results was assessed by triangulation with the other analytical approaches.

6. VISUALIZATION

To complement the text and content analysis, we carried out visual analysis of the data. Information visualization augments analytical capability [[12],[13],[14],[15]], by harnessing advanced human perception, making it easier to "see" information hidden within data. First high-level, exploratory overviews of data structure and content are examined [[12],[16],[17]]. Detailed analysis of regions of interest (ROIs) is then carried out using visualization methods selected based on data type and task, and the target audience. The merits of visualization-based analysis are, however, often limited by screen real estate and resolution, rendering even the most intuitive representations increasingly difficult to interpret with increase in data size, density, dimensionality and (other factors contributing to) complexity [18]. Interactive visualization is especially useful then, as it aids the recognition of (emerging) patterns and trends [[13],[15]], by allowing the analyst additional perspectives on the data, as well as functionality for suppressing or highlighting data of low or high relevance respectively [17].

Our analysis aims to detect dynamicity in term usage, meaning and frequency in tweets. We therefore consider each term as a facet or dimension in the data, in addition to the temporal dimension. With 241,748 tweets containing an undetermined amount of noise, the complete dataset is sufficiently complex to warrant support for visual, exploratory analytics, for the human analyst to draw confident conclusions about its content. Because we are investigating public engagement with science, we aim in the long term to contribute to support for more effective analysis of such data. We therefore focus on accessible visualization approaches able to support exploration across all facets in high-dimensional (high-D) data, with a temporal element across all others. Dimensionality reduction techniques [18] are typically employed during pre-processing of high-D data. However, the open, shared data collected, even in this atypically trending case, even with noise, cannot be considered as very large scale. Therefore the benefits in dimensionality reduction are outweighed by the potential sacrifice of the key features in the data we wish to make explicit [15].

Multi-dimensional data visualization typically employs multiple, co-ordinated views [[16],[19],[17],[20]], allowing different perspectives on selected data types, e.g., cartography for the geographical dimension, and timelines for temporal data [[13],[19]], in addition to custom implementations such as tag or topic clouds [[19],[21],[17],[20]]. The simple, 2D scatterplot is one of the best-known examples of a visual analysis technique for comparing any two dimensions in a dataset [[19],[15],[18]]. Scatterplots (and other techniques) may incorporate visual cues such as colour, node size and shape [[19],[22],[15],[18],[20]], or be extended to 3D or matrices [[15],[17]], in order to compare across additional dimensions. Chan et al. [15] describe an extension of scatter plots to illustrate sensitivity due to the influence of additional attributes in a dataset as a flow or stream.

Small multiples, as described by Tufte [14], are particularly useful for displaying both discrete and continuous changes in data that lends itself to a pictorial representation. Parallel co-ordinates (//-coords) [23] plot each dimension (co-ordinate) in high-D data on a separate vertical axis in 2D space, providing an overview that reveals clusters of related information across multiple dimensions, while isolating outliers [18]. //-coords are often used as timelines [[13],[20]], with the temporal dimension on the horizontal axis. Additional benefits include the ability to reorder and cluster axes, flip poles, and extract each co-ordinate to a scatter plot that focuses on each in isolation, or a collection of dimensions, as in Tufte's [14] small multiples approach. *ParallelTopics* [19] demonstrates the use of //-coords, linked to topic clouds, a scatter plot and a theme river, to support topical analysis of large-scale text corpora. Riehmann et al. [22] describe customisation of //-coords to support use by the lay public. The authors report successful evaluation, demonstrating effective, intuitive support for exploring and filtering on product attributes in online stores.

Krstajic et al. [24] describe the use of CloudLines, to explore entities and events in temporal data. Kim et al. [21] illustrate an extension of tag clouds to augment content analysis. They generate overview document summaries, as a node-link network of key terms, bridged by the relationships between entities. Collins et al. [20] demonstrate a combination of tag clouds with //-coords for large scale, exploratory, faceted text analysis. Faceted views as illustrated also in [17], are a not unusual option on websites for navigating through and filtering on categories (facets) in document or item collections. A challenge highlighted in each of these papers is the need to enable more intuitive identification and interpretation of key terms or themes – facets or dimensions – in textual data, as amount and dimensionality increase.

//-coords are well suited to the exploratory analysis we report, as they allow us to obtain a quick overview of trends across all terms (facets) in the data, as well as the evolution of each facet and the entire dataset with time. We aim, in follow-up work, to build on the small multiples technique [23], to foster more detailed anlaysis of evolution across facets extracted from the //-coords plot to representations that map to different user types and analytical tasks. Table 1, for instance, uses a series of time-based concept maps to illustrate trends in terminology during the evolving event of the Curiosity landing. Tag clouds, already in common use for visualizing term frequency and (perceived) importance on the (readable and semantic) Web and in social media data, may provide an alternative perspective, to support further text and content analysis for the ordinary end user.

6.1 Data Overview: Filters & Aging Factor

We use a trial version of Macrofocus's ||-*coords*[4] to analyse the multiple dimensions hidden within the tweets, each of which is a specification of a code in Figure 1. Initial detailed analysis, to identify trends and ROIs in the data, was restricted to the samples of 200 tweets for each day (from 04-09 Aug) used in the content analysis. The smaller number also allowed manual inspection of the data content necessary for (content) coding.

Datasets for the visualization were filtered first using four key Scientific terms: curiosity, MSL, #curiosity, @marscuriosity (the rover's Twitter username) and one multi-term search ((7 OR seven) AND terror). The overview, in Figure 3, plots the dataset, split into six hour periods from 00:00 GMT on August 4th to 23:59 GMT on August 9th, showing, for all tweets collected over the full period, the trends for each term. For this and each of the following //-coords plots, each polyline, cutting across axes for each filter term, represents a temporal unit of measure, one of:

- **date range**: 00:00-23:59, e.g., date_range_0 = 04 Aug
- **quarter of a day**: successive counts from q0 - 00:00-05:59 (dateA00, e.g., 6A00 – 1st quarter of 06 Aug)

Tweet counts peak on 06 Aug, with the highest for all filter terms falling on **6A06** – 0600-11:59 – the period during which Curiosity successfully touched down on Mars.

To measure retweeting activity we used the Ageing Factor metric, AF, defined in [2]:

$$\boldsymbol{AF} = \sqrt[i]{\frac{\boldsymbol{k}}{\boldsymbol{k}+\boldsymbol{l}}}$$

where i is the cut-off time in hours since the originating tweet, k is the number of retweets originating at least i hours ago and l is the number of retweets originating less than i hours ago.

[4] http://ll-coords.com (now available as *ProfilePlot*)

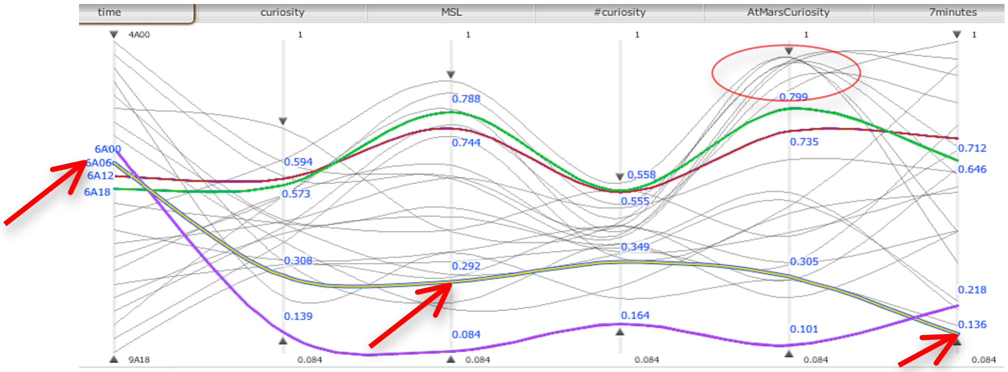


Figure 2: 1hAF values for the tweets in Figure 3. 6A06 (highlighted) shows relatively low AF values. The preceding quarter, 6A00 records the lowest AF values.

This metric allows a synchronous view of retweeting behaviour in which the originating tweet does not need to be present in the sample. Values of AF range from 0 to 1. Low AF is assumed to indicate rapid retweeting, such as might occur around a currently unfolding event of interest. We found previously [2] that 1hAF (a 1 hour cutoff) provides useful information for evolving events on Twitter, and subsequent breakdown into quarters in each day provides distinct trends. Figure 2 plots 1hAF for the data in Figure 3; **6A00** records the lowest AF for all but the term "7 Minutes of Terror" or *7minutes*. **6A06** (highlighted) is at its largest (the filter #curiosity) the 4th lowest AF value, and the lowest overall recorded for 7minutes.

High AF is assumed to indicate enduring interest in a topic. 1hAF for the second half of 06 Aug is much higher, nearer the top of the range. "7 minutes of terror" is overall, the least frequently tweeted term, and the only term to record maximum 1hAF (1), during **9A00**. We also note high 1hAF for AtMarsCuriosity, forming a distinct cluster (area encircled), starting before the landing and continuing after it. This indicates enduring interest from people sufficiently engaged to follow the rover's official persona.

6.2 Detailed Visual Analysis in ROIs

Visualization is useful for obtaining an overview of data structure, followed by more detailed analysis of ROIs [[12],[13],[21]]. To determine if the trends revealed in the smaller sample were representative of the larger dataset we increased the sample ten-fold, to 2000 tweets each. The plots (of 2000) showed persistence in overall data structure as well as in the trends and ROIs revealed in the smaller samples. While the content analysis reported in Section 5 is restricted to the first set of 200, the remainder of the discussion in this section is based on the larger random sample. We focus on activity on 06 Aug, which contains the two quarters of particular interest: the peak in Figure 3, touch down at 6A06, and the preceding quarter, 6A00, the lowest AF polyline in Figure 2 and second highest (tweet count) in Figure 3.

Visual analysis based on the larger random sample (2000 tweets each) strengthens the case for the particular visualization approach for intuitive analysis as data size, density and dimensionality increase. Analysis using //-coords is often restricted to visualization experts, as the approach is sometimes seen to have a steep learning curve. However, more recent research [[13],[22]] has shown the approach to be particularly useful for high-D data, even for non-experts. Used with alternative, co-ordinated visualization options, such as scatter plots and other layouts specific to selected dimensions (e.g., maps for cartographic data), and with the simple filtering obtained through range sliders as illustrated here and in early seminal work in [25], //-coords may serve as a powerful analytical tool. Examples in Tufte's "narratives of space and time" [14] illustrate a parallel with commonly used bus and train timetables and other pictorial representations of temporal data in nature; this may provide an avenue for additional co-ordinated visualizations and that may help to minimize any initial challenges in use. We illustrate here the potential of //-coords to support domain experts and the ordinary end user interacting with moderate to large amounts of high-dimensional data, the norm in today's information-rich society. Two factors may contribute to the success of //-coords in this case, as reported in our previous work [1], the use of scientific terms on Twitter is relatively low, with frequency further reduced when filtered to remove noise (use in non-scientific contexts). Our dataset in this case, even for the atypically trending scientific terms surrounding Curiosity, while still large enough for the human user to require support for effective analysis (ranging from 100s, to100s of 1000s, to 1.25M), is much smaller than typical counts (100s of 1000s to several millions) extracted from the Twitter firehose stream for trending data. //-coords are also particularly useful for highlighting clusters

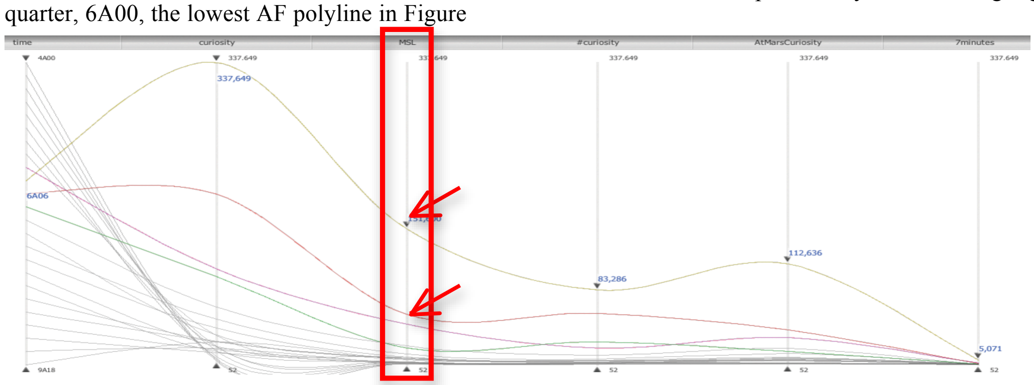


Figure 3: Tweet counts for the filter terms for the complete dataset, for each 6 hour period from 04-09 Aug 2012. The peak, 06:00-11:59 on 06 Aug – the period during which Curiosity landed on Mars, is highlighted in successive plots.

along and between dimensions and for isolating outliers. For the atypically trending scientific terms – #Curiosity and #MSL – peaks are found for all terms in 6A06, and while falling during the other three periods on 06 Aug (see Figure 3), clearly removed from the much lower clusters for the other 5 days surrounding Curiosity's landing. The filter term #curiosity is disregarded for the detailed analysis as it occurs in every tweet in the samples. The resulting co-ordinates, normalized across the remaining four terms, are plotted for each day in Figure 4.

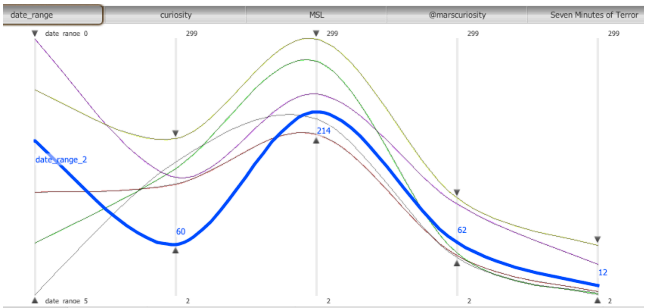


Figure 4: Frequency of terms per day, for a random sample of 2000 tweets each – data_range_0 maps to 00:00-23:59, 04 Aug 2012. The filter term #curiosity is hidden in this plot.

The shape of each polygon is fairly consistent across all 6 days. The relative totals for 06 Aug, however, drop, compared to the other days, with the skew largely attributable to the last two quarters in the day. Evidence for the variance is seen in both tweet counts and 1hAF values. Manual inspection of tweet content indicates this could be due to a lull after the main event was concluded, with further interest sustained by what may be more avid enthusiasts within the general public, in addition to the scientific community. We aim to carry out a more detailed study as we explore further visual analytics options for detailed analysis.

The term #MSL shows a clear peak throughout, consistent with it trending over this period. The compound term 7minutes records the lowest count over the same period, peaking the day before the landing, which also records maximum values across all terms. The peaks match anticipation of the landing. The counts for 7minutes continue to drop quickly over the subsequent days (but remains above 0), after this period elapses and Curiosity lands safely.

Considering the low 1hAF values for the first half of 06 Aug in Figure 2 and the peak in tweet count for the whole day for the complete dataset (Figure 3), we take a more detailed look, in Figure 5, at each of the four quarters in this day. The plot reflects more closely the results seen in the overview. 6A00 shows a clear peak for all terms, followed by 6A06. The drop in #MSL and 7minutes from the peaks at 6A00 is pronounced (halved).

In addition to plotting the data, Macrofocus's ||-*coords* implementation supports content inspection for each data point. Combined with interactive sliders (see [25]), it enables visual, dynamic filtering, which we used in the more detailed analysis of the tweet content for 6A00 and 6A06. The codes defined in Section 5 guided this analysis. An interesting point to note is that non-English posts (code *Other languages*), including those using a non-Latin alphabet, typically made use of English hashtags. So while these tweets are not considered during the content analysis (as explained in Section 5), we include all posts in the overview plots for the exploratory analysis we report here, as relevance using manual inspection is easily inferred.

As in previous work [[1],[2]], a significant proportion of posts containing scientific terms use a colloquial interpretation, with reference to known brands, or in jokes. We found a large proportion of tweets that would be coded as *Fun* in 6A06, albeit with clear reference to Curiosity, largely puns on the name and about martians; people tweeting about what they or others were doing while watching the landing, e.g., "*just watched #missioncontrol #passthepeanuts. Feel pretty lucky to get to witness this. #curiosity #nasa #mars*"; congratulatory messages to the Curiosity team and tweets about parties in the lead up to the landing or to celebrate its success. Figure 6 plots additional terms extracted during the text analysis (see Section 4); some of the more common terms found for the frame *Fun* were "curiosity killed the cat | cat(s)/feline | martian(s) | Mars Bar | Peanuts". Other terms of interest were "Congratulations | Celebration | party", and in a more *Scientific* frame, "Nerd | Science | Landing | Images/Photos | JPL | NASA | Rover".

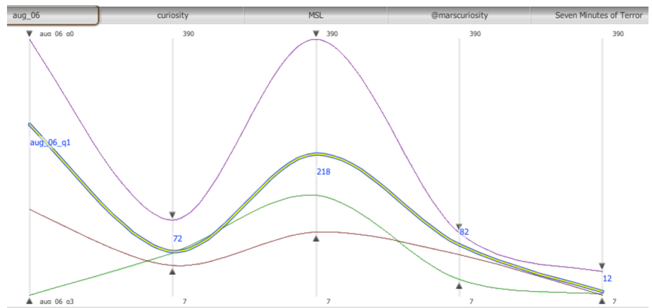


Figure 5: Frequency of terms per quarter, for the random sample of 2000 tweets each for 06 Aug 2012. The period from 06:00-11:59 is highlighted. (The filter #curiosity is hidden.)

6A00 on the other hand contained a much larger proportion of tweets coded as *Scientific*. These are more likely to be from users with a more research-oriented basis for their interest, including tweets pointing to sources of more information about Curiosity and the landing process. We posit that in 6A06 the relief at the successful landing attracted more tweets from the general public with an interest in science, heightened by the publicity surrounding the landing, in addition to "high fives" from and within the scientific community. For instance, while Landing, (#)NASA and #MSL show clear peaks for 6A00, there is only one tweet containing the term congrat(ulation)s. The first three terms drop significantly during 6A06; however congrats peaks at 79, mixed with other celebratory messages and comments about partying at NASA, and tweeters' own Curiosity party plans. Jokes of the type "*How ironic would it be if #CURIOSITY landed on a Martian cat?*" and/or curiosity killing the cat rise in 6A06 to over 140, three times that for 6A00. Tweets of the kind "*Also it is not made of Mars Bars.*" and about eating Mars Bars and peanuts during the "seven minutes of terror" and to celebrate the successful landing, while being in the lowest trough overall, doubled from q0 (6A00) to q1 (6A06). "*Let s all dance for 7 minutes. #Curiosity*" may be considered to fall in the *Fun* frame, but due to the more pertinent terms 7minutes and #curiosity, could be weighted toward *Scientific*.

Considering sentiment, we find a large number of positive tweets in the lead up to the landing, expressing anticipation, excitement and pride, and in 6A06, increasingly showing also relief and celebration post landing. A smaller number of tweets fall in the frames *Political*, *Economic* and *ELSI* (not shown in Figure 6). *ELSI* particularly tended to range from neutral to negative.

Overall, the trends for all four quarters of 06 Aug, and also for the full period, from 04-09 Aug. (Figure 7), are fairly consistent. One significant difference from our previous work is, a much lower

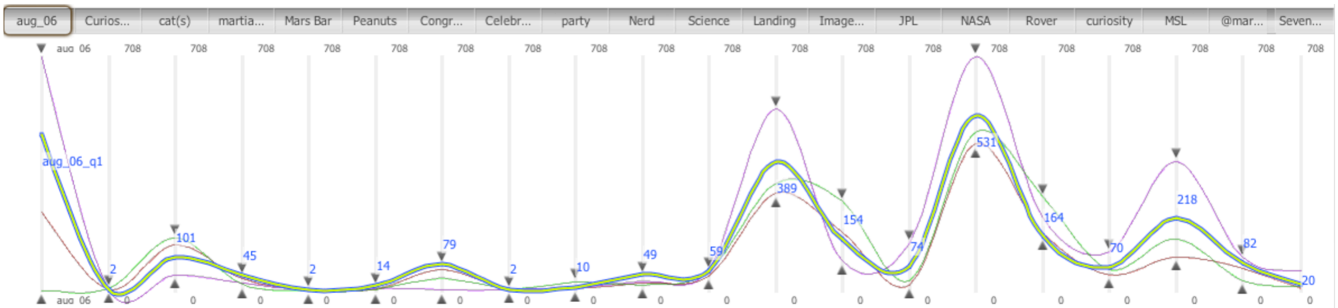


Figure 6: Search on additional terms extracted during the text analysis for the four quarters in 06 Aug. (6A06 highlighted). 5 peaks are found, in descending order, for NASA, Landing, #MSL, and much lower, cat/feline and congratulatory messages. A tiny peak is seen for nerd in 6A06.

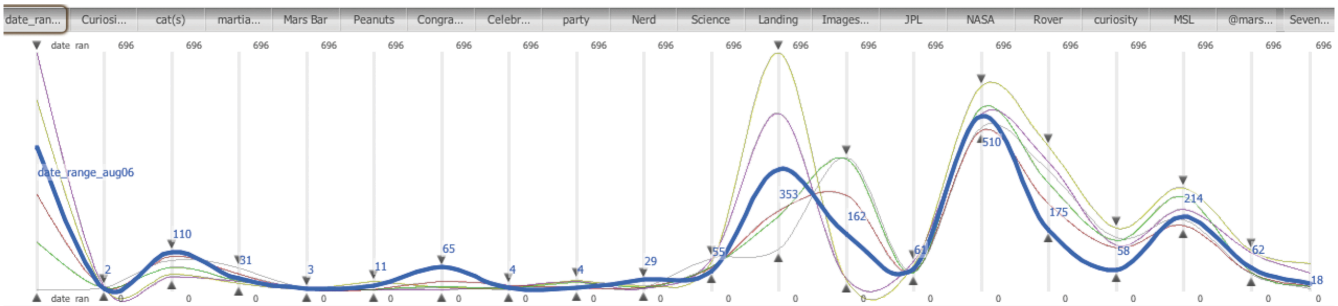


Figure 7: Trends for the additional terms (as in Figure 6), for the full period (04-09 Aug 2012). 06 Aug. is highlighted. The patterns for term frequency largely follow those in Figure 6, but with an additional peak for Images/Photos after the landing (07-09 Aug).

proportion of noise. It must be noted however that the hashtag #MSL, which was found to be trending at the time (along with the random sample filter term #curiosity), was present in a large number of tweets judged to be *Fun*, rather than *Scientific* or one of the other three frames. However, these were still found, for the most part, to be in reference to Curiosity/MSL.

7. CONCLUSIONS

The analysis reported here is initial and exploratory. The visualization component of the study served as independent validation of text analysis and content analysis. The latter was also found to serve as a springboard – with the richer qualitative understanding it yielded suggesting routes to explore the data further with the visualization. Thus the methods were found to be complementary and indicate that this kind of mixed methods study could be used successfully to characterize public engagement with science expressed in microposts. In order to further assess the efficacy of the approach and the kinds of insights it can yield for studies of public engagement with science, the approach must be applied to different kinds of events and to topics that are not event driven.

The advantage in the visual overviews was easily recognizable variation in trends as the event unfolded, with very clear peaks and troughs at distinct periods. These were found, on more detailed investigation, to map to key points – just before and after the Curiosity landing. Looking at isolated points in the overviews, we found more evidence to support our initial theories on how aging factor maps to intense retweeting during evolving events, as opposed to longer term interest in topics of interest.

Triangulation of the mixed methods results for the research questions stated in Section 3 is as follows:

1) *How did tweeting and retweeting activity vary over the course of the days and hours around the Curiosity landing?* The //-coords visualizations show that both the number of tweets and the speed of retweeting (expressed as 1hAF) were most intense in the two quarter days either side of the landing event (6A00 and 6A06). Further, while the scientific frame was stronger prior to the landing, the fun was at its strongest immediately after the landing.

2) *What message frames were commonly used in tweets?* Evidence from all three of the analytical methods used indicates that the Scientific and Fun frames were the strongest for this event. The result for Fun is validated by evidence for jokes about "martians" and "cats" from both the text analysis and the visualization. The result for the Scientific frame is also validated with evidence for interest in the "landing", "MSL" and "images", the latter particularly after the landing event. The concept map for 08 Aug includes associations with "united states" and "america" both associated with the Political frame.

3) *What was the prevailing sentiment?* The content analysis indicated sentiment was largely positive. Evidence from the visualization, which identified a peak for "congratulations" on 06 Aug supports this conclusion.

An additional outcome, which emerged from recording the code Other Languages, was that international interest was high.

From a methodological perspective, the study used parallel co-ordinates visualization, a method which has previously been seen as more suitable for expert users. As found in more recent work, however, the technique may be used effectively for exploratory analysis by the non-expert, especially when combined with simple filtering techniques, such as query sliders, and custom, co-ordinated views that allow a focus on selected dimensions in ROIs. It was found that it gave intuitive, interpretable results for this dataset, both for obtaining overviews of the data and for more detailed trend analysis for selected ROIs in data subsets. We consider it to have promise for web science studies.

8. ACKNOWLEDGMENTS

Aba-Sah Dadzie is currently funded by the UK MRC project Time to Change (129941).

9. REFERENCES

[1] V. Uren and A.-S. Dadzie, "Relative Trends in Scientific Terms on Twitter," in *altmetrics11: Tracking scholarly interest on the Web, Workshop at ACM WebSci'11*, 2011.

[2] V. Uren and A.-S. Dadzie, "Ageing Factor: a Potential Altmetric for Observing Events and Attention Spans in Microblogs," in *1st International Workshop on Knowledge Extraction and Consolidation from Social Media (KECSM)*, 2012.

[3] M. C. Nisbet and D. A. Scheufele, "What's next for science communication? Promising directions and lingering distractions.," *American Journal of Botany*, vol. 96, no. 10, pp. 1767–1778, 2009.

[4] M. C. Nisbet, "Commmunicating climate change: why frames matter for public engagement," *Environment*, vol. 51, no. 2, pp. 12–23, 2009.

[5] E. T. Meyer and R. Schroeder, "Untangling the Web of e-Research: Towards a Sociology of Online Knowledge," *Journal of Informetrics*, vol. 3, no. 3, pp. 246–260, 2009.

[6] W. E. Kastenberg and R. Wilson, "Risk of nuclear powered space probes," *Reliability Engineering & System Safety*, vol. 86, no. 1, pp. 53–59, 2004.

[7] T. Dinerman, "NASAs continuing problems," *The Space Review*, 18-Apr-2011.

[8] K. Krippendorp, *Content Analysis: An Introduction to its Methodology*. Sage, 2004.

[9] M. S. Schäfer, "From Public Understanding to Public Engagement : An Empirical Assessment of Changes in Science Coverage," *Science Communication*, vol. 30, p. 475, 2008.

[10] R. S. Hooper, "Indexer consistency tests - Origin, measurements, results and utilization," Bethesda, 1965.

[11] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, pp. 37–46, 1960.

[12] S. K. Card, J. D. Mackinlay, and B. Shneiderman, *Readings in information visualization - using vision to think*. Academic Press, 1999.

[13] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, "Mastering The Information Age – Solving Problems with Visual Analytics," 2010.

[14] E. Tufte, *Envisioning information*. Cheshire, CT, USA: Graphics Press, 1990.

[15] Y.-H. Chan, C. Correa, and K.-L. Ma, "Flow-based scatterplots for sensitivity analysis," in *Proc., IEEE Symposium on Visual Analytics Science and Technology*, 2010, pp. 43–50.

[16] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, "Visual Analytics: Definition, Process, and Challenges," in *Information Visualization: Human-Centered Issues and Perspectives*, A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, Eds. Berlin/Heidelberg: Springer-Verlag, 2008, pp. 154–175.

[17] V. Thai, P.-Y. Rouille, and S. Handschuh, "Visual Abstraction and Ordering in Faceted Browsing of Text Collections," *ACM Transactions on Intelligent Systems and Technology*, vol. 3, no. 2, pp. 21:1–21:24, Feb. 2012.

[18] N. Elmqvist and J.-D. Fekete, "Hierarchical Aggregation for Information Visualization: Overview, Techniques, and Design Guidelines," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 3, pp. 439–454, 2010.

[19] W. Dou, X. Wang, R. Chang, and W. Ribarsky, "ParallelTopics : A Probabilistic Approach to Exploring Document Collections," *System*, vol. 07, pp. 231–240, 2011.

[20] C. Collins, F. B. Viegas, and M. Wattenberg, "Parallel Tag Clouds to explore and analyze faceted text corpora," in *Proc., IEEE Symposium On Visual Analytics Science And Technology*, 2009, pp. 91–98.

[21] K. Kim, S. Ko, N. Elmqvist, and D. S. Ebert, "WordBridge: Using Composite Tag Clouds in Node-Link Diagrams for Visualizing Content and Relations in Text Corpora," in *Proc., 44rd Annual Hawaii International Conference on System Sciences*, 2011, pp. 1–8.

[22] P. Riehmann, J. Opolka, and B. Froehlich, "The product explorer: decision making with ease.," in *International Working Conference on Advanced Visual interfaces (AVI'12)*, 2012, pp. 423–432.

[23] A. Inselberg, *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*. Springer-Verlag New York, Inc., 2009.

[24] M. Krstajic, E. Bertini, and D. Keim, "CloudLines: Compact Display of Event Episodes in Multiple Time-Series," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2432–2439, 2011.

[25] C. Ahlberg and B. Shneiderman, "The alphaslider: a compact and rapid selector," *Conference companion on Human factors in computing systems CHI 94*, pp. 365–371, 1994.