DEMO: Automated Analysis Of Charities’ Communication Styles On Twitter
Christopher Phethean, Thanassis Tiropanis, Lisa Harris
University of Southampton
England, UK
{C.J.Phethean, tt2, L.J.Harris}@soton.ac.uk

ABSTRACT
This extended abstract presents an overview of a technology demonstration that would consist of a novel piece of software for automatically analysing charitable organisations’ posts on the micro-blogging service Twitter. The main contribution is an inbuilt taxonomy of communication styles, which is used to classify each charity based on the results of the analysis undertaken. The demonstration will showcase this automated process, as well as explaining the restrictive properties for each category, and why certain charities are classified into certain styles.

Categories and Subject Descriptors
H.3.5 [Online Information Services]: Web Based Services

General Terms
Measurement, Performance

Keywords
Social media; Twitter; charities; marketing; web science; classification

1. INTRODUCTION
Measuring charities’ performance of using social media is difficult because of the variety of ways in which these services can be utilised. Each way of using each different social media tool can produce vastly different outcomes, which are frequently subjected to generic measurements to establish how “well” the charity is performing. However these often do not take into account specific strategies of the charities themselves, which may be aiming to achieve something different. Previous work by the authors has examined how and why charities use social media, finding that the aims range from relationship building and mobilisation of supporters to generation of referral traffic to their websites [3]. In order to establish the performance of any of these aims, numerous variables relating to the engagement and interaction the charity receives back on its social media posts could be measured; however it is believed that the first step should not be to assess the responses that are collected, but to determine whether the organisation’s own communication style is conducive to produce the desired output.

2. DETAILS OF APPLICATION
SCENARIO
An automated classification system has been developed which assesses the communication style of organisations on Twitter. Previous work into analysing communicator roles on social media has largely focused on individuals, looking for example at the conversation role based on how frequently a user retweets another user, and how often they are retweeted themselves [4]. Alternatively, work has been carried out to represent different groups of customer in a model based on the extent to which they are involved with social media, with each “step” requiring increased levels of engagement and activity [2]. Relating to organisations, there is literature to suggest how US governmental agencies utilise Twitter, focusing mainly on distinguishing between one-way or two-way communication, although in this case the analysis was carried out as an aggregation of the entire sector, rather than for a single agency [5]. This demonstration will therefore showcase an original approach to analysing communication styles, by focusing the analysis on individual organisational accounts, rather than users. The classification approach builds on work carried out on user roles on Twitter but instead tailors the variables that are measured towards those that would indicate a particular type of communication from a charity or similar organisation. These include:

- New Tweets – a broadcast mechanism indicative of information dissemination.
- Replies – directed messages that suggest a willingness to engage in conversation with supporters, and respond to queries.
- Retweets – message posting mechanism that reveals the organisation is willing to take notice of other users’ content, and share it with their own followers.
- URLs – links to other content which suggest a desire to refer users to another website for more information.
- Hashtags – descriptive words or phrases that indicate a categorisation system to the organisation’s posting, where for example, all tweets related to a particular campaign could be grouped.

3. DEMONSTRATOR AND INSTALLATION
The installation will showcase the running of the automated classification system, which takes the form of a lightweight piece of software. Taking a charity’s Twitter username as a search parameter, the software collects the previous 7 days’ worth of tweets from the organisation and calculates the frequencies of the
five variables listed above. Based on these values, further calculations are made to assess the favoured communication style of the charity from a stored taxonomy of categories that describe different types of communication (Figure 1). From the outputs of these calculations, the software then assigns the charity to the particular category that is most suitable from the evidence of the data collected.

The demonstrator is the developer of the software and will explain what analysis the software is performing at each stage, along with providing a visual display of the internal taxonomy that the classifier uses as to make the classifications. They will also be able to discuss why certain charities are assigned into particular categories, by explaining the restrictive properties on each category that were used to develop the taxonomy.

4. TECHNOLOGY USED

The development of this software began with the intention of making an automated system that could form part of the Web Observatory1 and analyse organisational communication styles on Twitter. The software is currently comprised of a Python program that accesses the Twitter Search API to obtain the latest tweet data from each organisation, and then performs calculations to classify the charity based on the internally coded taxonomy. This innovates on previous Twitter analysis that simply analyses Tweets and may produce an output based on how many of each kind of message is sent, but does not then go on to assess the frequencies and proportions of theses. The software presented here does this, and as such is able to compare the observed use to a standardised taxonomical model, and ultimately create a communication profile of the organisation.

5. LESSONS LEARNED

Numerous lessons were learned from producing this application. Firstly, while an initial classification was made based on 7 different categories from a pilot sample of charities, a test run of the classifier revealed some initial problems whereby certain charities weren’t classified. To resolve this, 2 new categories were added to the taxonomy, to ensure that regardless of the type of use, a category will always be assigned.

Furthermore, it was found in the initial tests that there was a vast difference between charities who had, in other reports ([1]), been ranked as some of the top social brands in the country, compared to those that had not. It was therefore important to contain sufficient categories to cover all charities, regardless of how extensively they adopt social media.

6. ACKNOWLEDGMENTS

This research was funded by the Research Councils UK Digital Economy Programme, Web Science Doctoral Training Centre, University of Southampton. EP/G036926/1.

7. REFERENCES


---

1 http://webscience.org/web-observatory/