Twitter related academic papers

A dataset from 2007-2011

# Overview

This dataset consists of a bibliography of papers relating to Twitter studied in [Williams, Terras, and Warwick (to appear](#_ENREF_3)). The aim of the study was to locate and classify academic papers on Twitter. Papers were identified by searching two databases:

* Scopus (http://www.info.sciverse.com/scopus)
* Web of Science (http://wok.mimas.ac.uk/

Papers were identified by using the search word “Twitter”, limiting the searches to the abstracts, keywords and titles of journal and conference papers, published between 2007 (the first year there were any papers related to Twitter) and 2011 (the last full year before the study). The searches were all conducted on 12th January 2012, Web of Science returned 384 items and Scopus 1132. Data cleansing was performed to remove obvious duplicates, and items with missing data, leaving a total of 1161 items, subsequently another nine were found to be duplicates, leaving a corpus of 1152 papers.

Each paper’s title and abstract was read and re-read and classified according to the paper’s focus:

* Focussed. The paper is focussed on Twitter or another microblogging system.
* Mentions. The paper mentions Twitter or microblogging but it is primarily about something else.
* Not. The paper is not about microblogging, the term twitter is used in another sense, for example the noise made by birds.

This dataset lists these papers with 575 identified as Twitter-focussed; 550 as mentioning Twitter and 27 not related to microblogging.

# Limitations

The classification of papers is based on the opinions of the researchers formed after reading the title and abstract. A reader of the full paper may make a different classification.

The dataset was collected on 12th January 2012, since that time the journals and conferences indexed by the databases have increased and so there are papers in earlier years that would appear in the dataset the would be included if it were collected today. For example the paper by [Chang (2010](#_ENREF_1)) was not in the original set as the Proceedings it is in was not at the time indexed. Some editions of periodicals with 2011 publications dates were not indexed until later in 2012 for example [Eysenbach G (2011](#_ENREF_2)) is not included as the final 2011 issue of the journal was not apparently indexed at the time.

# References

Chang, H.-C. (2010). *A new perspective on Twitter hashtag use: Diffusion of innovation theory*. Paper presented at the Proceedings of the American Society for Information Science and Technology.

Eysenbach G. (2011). Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. *journal of Medical Internet Research, 13*(4), e123.

Williams, S., Terras, M., & Warwick, C. (to appear). What people study when they study Twitter: Classifying Twitter related academic papers. *Journal of Documentation*.