A Tweet Consumer's Look At Twitter Through Linked Data Goggles Via Google Analytics

Thomas Steiner and Arnaud Brousseau

Google Germany GmbH, ABC-Str. 19, 20354 Hamburg, Germany
{tomac|arnaudb}@google.com

Abstract.

1 Introduction

1.1 Google Chrome Extensions

Google Chrome extensions[1] are small software programs that can be installed to enrich the browsing experience with the Google Chrome browser. They are written using a combination of standard Web technologies, such as HTML, JavaScript, and CSS. Chrome extensions bundle all their files into a single file that gets usually (but not necessarily) distributed through the Chrome Web Store. There are several types of extensions, for this paper we focus on extensions based on so-called content scripts. Content scripts are JavaScript programs that run in the context of Web pages, similar to the Firefox Greasemonkey extension[2]. By using the standard Document Object Model (DOM), they can read or modify details of the Web pages a user visits. Examples of such modifications are, e.g., changing hyperlinks to remove potential @target="_blank" attributes, or increasing the font size.

1.2 Google Analytics

2 Twitter Swarm NLP Extension

With our Twitter Swarm NLP extension[3], we inject JavaScript code via a content script into the Twitter.com homepage. The extension first checks if the user is logged in, and if so, retrieves the tweets of the logged-in user's timeline one-by-one, and performs NLP analysis via a remote NLP Web service on each of the tweets. The extracted entities are then displayed on the righthand-pane of the Twitter.com homepage, and sent to Google Analytics for further processing.

[1] Google Chrome Extensions: http://code.google.com/chrome/extensions/index.html. Text adapted from the description to be found there.
[2] Firefox Greasemonkey extension: http://www.greasespot.net/
[3] https://chrome.google.com/webstore/detail/dpbphenfafkflfmdlanimlemacankjol

2.1 Twitter Swarm NLP Web Service

We have created a wrapper NLP Web service that merges results from existing third-party NLP Web services, namely from OpenCalais[4], Zemanta[5], AlchemyAPI[6], and DBpedia Spotlight[7].

2.2 Dealing With Extracted Entites On the Client Side

2.3 Dealing With Extracted Entites On the Google Analytics Side

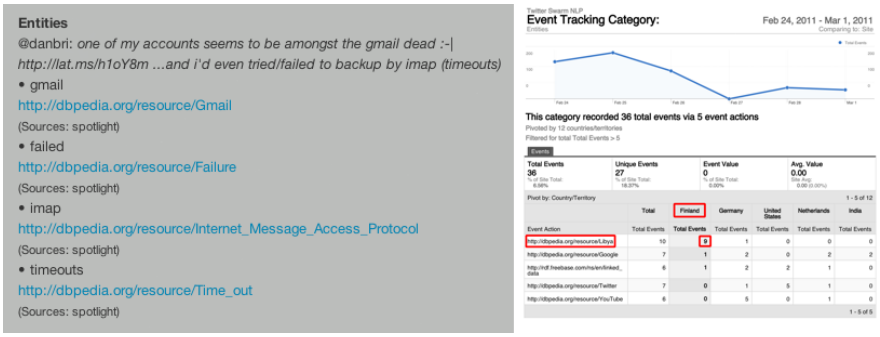

Fig. 1. Left: Screenshot of the extracted entites of a particular tweet as displayed by the Twitter Swarm NLP Extension. Right: Test.

3 Related Work

3.1 Linked Open Social Signals (TWARQL)

@ToDo Arnaud `http://knoesis.wright.edu/library/download/paper-wi10-MPKS.pdf` `http://wiki.knoesis.org/index.php/Twarql`

3.2 Twopular

@ToDo Arnaud `http://twopular.com/tag#about`

4 Conclusion

References

[4] `http://www.opencalais.com/`
[5] `http://www.opencalais.com/`
[6] `http://www.alchemyapi.com/`
[7] `http://dbpedia.org/spotlight`