



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/87991/>

Conference or Workshop Item:

Derczynski, Leon (2014) Leveraging the Power of Social Media. In: USES 2014 - The University of Sheffield Engineering Symposium, 24 June 2014, The Octagon Centre, University of Sheffield.

<https://doi.org/10.15445/01022014.16>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Leveraging the Power of Social Media

Leon Derczynski

Department of Computer Science, University of Sheffield

Abstract

From a business and government point of view there is an increasing need to interpret and act upon information from large-volume, social media streams, such as Twitter, Facebook, and forum posts. While natural language processing from newswire has been very well studied in the past two decades, understanding social media content has only recently been addressed in NLP research.

Social media poses three major computational challenges, dubbed by Gartner the 3Vs of big data: volume, velocity, and variety. NLP methods, in particular, face further difficulties arising from the short, noisy, and strongly contextualised nature of social media. To address the 3Vs of social media, novel language technologies have emerged, e.g. using locality sensitive hashing to detect new stories in media streams (volume), predicting stock market movements from tweet sentiment (velocity), and recommending blogs and news articles based on users' own comments (variety).