

# Finding ground truth in social media

December 13 2019, by David Bradley

---



Credit: CC0 Public Domain

Is it possible to extract "journalistic" as opposed to "general" commentary from social media? Writing in the International Journal of Grid and Utility Computing, a team from Portugal describes an approach to human and automatic extraction of updates and reports one might

describe as coming from citizen journalists. Their algorithm is trained on automatically annotated and human-annotated data sets and shows that the wholly automated approach homes in on "ground truth" data much more efficiently and effectively than when the data has had the human touch.

Nuno Guimarães, Filipe Miranda, and Álvaro Figueira of the University of Porto explain how social networks and [social media](#) have provided the means for constant connectivity and fast information dissemination. Eyewitnesses to events and happenings the world over can share information and insight in [real-time](#) from a sports event or other entertainment, from the scene of a disaster, crime, or other happening, in a way that is impossible for members of the conventional media unless they happen to be at the scene themselves. Moreover, [citizen journalists](#) can add a personal perspective that is precluded from independent journalist inquiry.

The automated algorithm builds up an internal trust that is not possible when human, subjective, classification is carried out. After a sufficient number of training data points, the system can unambiguously discern which updates are based on personal beliefs and which are ground truth.

**More information:** Nuno Guimarães et al. Identifying journalistically relevant social media texts using human and automatic methodologies, *International Journal of Grid and Utility Computing* (2019). [DOI: 10.1504/IJGUC.2020.103971](#)

Provided by Inderscience

Citation: Finding ground truth in social media (2019, December 13) retrieved 9 April 2024 from <https://techxplore.com/news/2019-12-ground-truth-social-media.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.