

# Separating fact from fiction using a 'fake news' algorithm

February 16 2017, by Adela Talbot

---



Victoria Rubin, a professor in the Faculty of Information and Media Studies, has been working on deception detection since 2010, more recently focusing on developing an algorithm to detect fake news. Credit: Adela Talbot // Western News

The impetus behind Victoria Rubin's research is a tip from Ernest

Hemingway: "Develop a built-in bullshit detector."

Working with a team of graduate students in the Faculty of Information and Media Studies (FIMS), Rubin has been studying deception detection since 2010, more recently focusing on developing an algorithm to detect [fake news](#).

The project, funded by the Social Sciences and Humanities Research Council, looks to identify deliberate misinformation using text analytics in an online environment. Rubin isn't looking to identify the kind of 'fake news' U.S. President Donald Trump might refer to; her research team has honed in on satirical news, specifically, and is working to develop an algorithm that will help hone the skill of [deception detection](#) online.

While deliberate journalistic fraud and online hoaxes that spread quickly via social media equally qualify as 'fake news,' satirical news pieces from outlets such as The Onion and The Beaverton presented Rubin and her doctoral students with a concrete opportunity to contrast and compare, to put fiction and truth side by side, in order to determine what separates the real from the false.

"(Satire) was something we could really understand very clearly in terms of linguistic properties. We collected a large data set of The Onion as opposed to The New York Times, and Toronto Star, as opposed to The Beaverton, and we tried to find the differences," Rubin explained.

Identifying the differences is crucial to developing an algorithm to test the accuracy of a 'news' piece, she noted. Every brand of 'fake' – whether it is journalistic deception, a hoax or a satirical news piece – would need its own algorithm.

The art of developing an algorithm to weed out fake news is part

linguistics, part information science, part technology, part data management and part natural language processing, which Rubin teaches in FIMS.

Combining these elements when comparing real news and fake news, her students (Niall Conroy, Yimin Chen, Sarah Cornwell and Toluwase Asubiaro) identified five features that serve as a litmus test in determining whether or not something is legitimate. A high prevalence of absurdity, humour elements, long sentences, negative affect and punctuation indicates a likelihood something is fake.

And if you think assessing the legitimacy of a satirical news piece is trivial, consider the story of Prime Minister Justin Trudeau and last year's 'elbowgate' scandal. In May, Trudeau was accused of "manhandling" two opposition members of parliament. Following the incident, Canada's satirical news outlet, The Beaverton, ran a piece that claimed members of parliament, in solidarity with the two Trudeau reportedly "manhandled," showed up to work in neck braces. The fake was picked up, re-reported and later retracted by the Hamilton Spectator.

"Even a qualified journalist wasn't able to tell that was a satirical piece," Rubin said, stressing the importance of reporting accurate information and individual discretion – both of which her algorithm and resultant satire detector are meant to bolster.

"Automation is not enough; education is crucial. It's a matter of activating critical thinking and providing as much assistance as we can to humans in distinguishing types of fakes," she explained.

"We've been going with the idea everybody should be able to identify (fake news), and if they can't, perhaps they can rely on some assistance, or having the kind of tools that would allow a person to be more aware. It's about increased awareness, decision making, how you proceed with

your daily life, how you proceed with your life as a citizen. Having misleading information changes your thinking; it affects elections and many aspects of life."

Rubin's satire detector, which will soon be live on a website hosted by FIMS, will allow anyone to upload a story to assess its accuracy using the algorithm her students developed. The detector will take into account the five linguistic elements and indicate to the user whether or not the story is likely to be a fake. The detector has a relatively high accuracy, Rubin noted, with a likelihood of being right 80-85 per cent of the time. She hopes to see engagement with the detector to improve its accuracy, she said.

"We're not collecting information; were just interested if we got it right or not. We're considering making it more of a game. If people can come up with different ways of being satirical (in news), we can modify and improve the [algorithm](#)."

Provided by University of Western Ontario

Citation: Separating fact from fiction using a 'fake news' algorithm (2017, February 16) retrieved 18 April 2024 from <https://phys.org/news/2017-02-fact-fiction-fake-news-algorithm.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.