

NewScientist

UK police are using AI to spot spikes in Brexit-related hate crimes

TECHNOLOGY 28 August 2019

By **Donna Lu**



Reported hate crimes spiked after the UK voted to leave the EU in 2016
REUTERS/Yves Herman

THE UK police are monitoring hundreds of thousands of Twitter posts related to Brexit every day. It is part of a pilot project to predict spikes in hate crimes in the run up to 31 October, when the UK is due to [leave the European Union](#).

The Online Hate Speech Dashboard is being used by analysts at the National Police Chiefs' Council's online hate crime hub, which was established by the Home Office in 2017 to "tackle the emerging threat of online hate crime".

It gathers Twitter posts from across the UK and uses [artificially intelligent algorithms to detect speech](#) that is, for example, Islamophobic, anti-Semitic or directed against people from certain countries or with disabilities or from LGBT+ groups.

[Matthew Williams](#) at Cardiff University, UK, and his colleagues created the dashboard so that government organisations could monitor hate speech.



Ads by Teads

The dashboard flags between 500,000 and 800,000 tweets per day related to Brexit, of which between 0.2 per cent and 0.5 per cent are classified as hateful. About 0.2 per cent of these are from users tagged with city locations within the UK, which the dashboard presents as a map of hate hotspots. If there is a spike, the information can be passed by analysts to the relevant local police forces, says Williams. Previously, such monitoring had to be done manually.

Read more: [Graphene inventor Andre Geim: No-deal Brexit would destroy UK science](#)

The main aim of the project is to identify patterns of hate speech in the lead up to 31 October to warn police and support organisations of any potential issues.

The team recently established for the first time that an increase in hate speech on Twitter leads to a corresponding increase in crimes against minorities on London streets (*British Journal of Criminology*, [doi.org/c9qh](https://doi.org/10.1017/S0007417819000094)). The pattern is similar to what happens with domestic violence, which often escalates from verbal to physical abuse, says Williams.

The team found that as the number of tweets that were antagonistic about race, ethnicity or religion increased, so did the incidence of aggravated crimes, including violence, harassment and criminal damage. A similar study in 2018 found a link between the number of anti-refugee statements on Facebook and violent crimes against refugees in Germany.

Relevant government authorities such as police forces and councils may use the information from the hub for [counter-messaging on social media](#). These include awareness campaigns, reiterating zero tolerance for hate crimes and encouraging people to report incidents to True Vision, a national crime reporting hub.

“The dashboard flags between 500,000 and 800,000 Brexit-related tweets per day”

Last year, the UK government launched a nationwide hate crime awareness campaign, which included adverts on social media.

The hope is that the dashboard will lead to a reduction in online hate speech. It includes information about the trends in hate speech against each group over time, and commonly used words and hashtags in hateful tweets. In addition, it shows networks of tweeters who interact with each other, although their identities are anonymised. These clusters can provide information about how much of the hate speech results from coordinated efforts, says Williams.

Williams and his colleagues measure the performance of the dashboard using an F1 score, a statistical measure of accuracy

that takes into account the rate of true and false positives. “Usually, our algorithms come in between 85 and 95 per cent,” says Williams.

Less than half of hate crimes are reported to the police. According to the Crime Survey for England and Wales, racially and religiously motivated crimes in the two nations spiked after the [Brexit vote in 2016](#), with 5605 crimes reported in July that year, up 44 per cent from the same period in 2015.

People with racist views feel emboldened to target others by events like the vote, says Imran Awan at Birmingham City University, UK.

The police are often slow in reacting, he says. Awan attributes this to scepticism about the link between online and offline abuse. “The perception is: ‘Do I really need to come out and speak to somebody because they’ve posted a tweet?’.”

Hate-speech detection tools that analyse aggregated data may not be able to prevent individual acts of violence, says Timothy Quinn at Hatebase, a firm that provides hate speech resources to law enforcement agencies. Such tools are more useful for governments to identify overall rises in hate speech across a region, giving opportunities to prevent it escalating into violence in the form of riots, for example, he says.

Article amended on 29 August 2019

We have corrected the nature of the tweets flagged by the dashboard and how many of these are geotagged. We have also clarified the involvement of the National Police Chiefs' Council.

More on these topics: [crime](#) [social media](#) [Brexit](#) [government](#) [Twitter](#)



TRENDING

LATEST

VIDEO

FREE

Net zero emissions target in peril as tropical forests absorb less CO2

1

Major science journal retracts study blaming climate change on the sun

2

Silver uses a surprising trick to stop the spread of bacteria

3

Have we really found an alien protein inside a meteorite?

4

Can an N95 face mask protect you from catching the new coronavirus?

5

Magazine issue 3245 , published 31 August 2019



SUBSCRIBE

VIEW IN THE APP

BUY IN PRINT

Previous article

Record Amazon rainforest fires spark row between Brazil and France

Next article

Dorian batters Bahamas with strong hurricane winds for record time

[Contact us](#) | [Help](#) | [About us](#) | [Privacy & cookies](#) | [Terms & conditions](#)

[Advertise](#) | [Write for us](#) | [Events](#) | [Science jobs](#) | [Syndication](#) | [RSS feeds](#)

[Gift subscriptions](#) | [Student subscriptions](#) | [Educational subscriptions](#) | [Corporate subscriptions](#)

GET THE APP



FOLLOW US

© Copyright New Scientist Ltd.

Back to top

