

YouTube's Notre Dame-9/11 flub highlights AI's blind spots

16 April 2019, by Barbara Ortutay



In this image made available on Tuesday April 16, 2019 flames and smoke rise from the blaze as the spire starts to topple on Notre Dame cathedral in Paris, Monday, April 15, 2019. An inferno that raged through Notre Dame Cathedral for more than 12 hours destroyed its spire and its roof but spared its twin medieval bell towers, and a frantic rescue effort saved the monument's "most precious treasures," including the Crown of Thorns purportedly worn by Jesus, officials said Tuesday. (AP Photo/Thierry Mallet)

YouTube might need a few more humans. The machines whose job is to tamp down conspiracy theories are not cutting it just yet.

As people around the world Monday turned to YouTube to watch Notre Dame Cathedral burn in Paris, an automated system attached background information about the Sept. 11 terror attacks in New York to livestream videos of the fire.

The cause of the blaze has not been determined, but authorities said it appeared to be accidental, not arson or terrorism.

The background note was posted by a system

YouTube recently put in place to combat well-known conspiracies about such events as the moon landing or 9/11. In this case, the algorithm might have had the opposite effect, fueling speculation about the cause of the fire and who might be behind it.

It's the latest example of [artificial intelligence](#) misfiring—and a sign that we have a long way to go before AI becomes smart enough to understand nuance and context.

In a statement, YouTube explained that the background information—an entry from the Encyclopedia Britannica—was mistakenly placed there by algorithms intended to protect users from fake material that spreads in the wake of some news events.

YouTube's algorithms have a history of misfiring and labeling videos inappropriately. Joshua Benton, director of the Nieman Journalism Lab at Harvard University, noted several in a blog post Monday.

Last fall, for instance, YouTube labeled a [video](#) of a professor's retirement from Michigan State University with the Encyclopedia Britannica entry for "Jew," along with a Star of David placed under the image. The professor, Ken Waltzer, had been head of the university's Jewish studies program, but Benton noted that nothing in the video's title or description mentioned anything Jewish.

YouTube's algorithm, which is presumably primed to bat down anti-Semitic conspiracies, somehow did that on its own.

When YouTube announced its anti-conspiracy efforts last summer, it said it would counter bogus information with sources people generally trusted, such as Wikipedia and Encyclopedia Britannica. It said it would add background from these sources to videos that feature common conspiracy subjects (for example, vaccinations, school shootings or the

1995 Oklahoma City bombing), regardless of whether the videos supported a conspiracy theory.

Videos of the Notre Dame fire were shown by large, trusted news organizations. YouTube's artificial intelligence, however, made no exceptions.

On Monday, the company quickly fixed the Notre Dame error and said its systems "sometimes make the wrong call." It turned off the information panels for the videos of the fire but did not say whether it was looking at the practice more broadly.

"I think they are sort of back and forth about how much good this is doing," Benton said. "It does get at the core question that we see with Facebook and YouTube and any other tech platform that aspires to global scale. There is just too much content to monitor and you can't have human beings monitor every video."

Instead, we have machines that are clearly still learning on the job.

"It's one thing to get something wrong when the stakes are low," Benton said. "When it's the biggest news story of the world, it seems like they could have more people looking at it."

© 2019 The Associated Press. All rights reserved.

APA citation: YouTube's Notre Dame-9/11 flub highlights AI's blind spots (2019, April 16) retrieved 23 September 2020 from <https://techxplore.com/news/2019-04-youtube-notre-dame-flub-highlights.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.