Reality Defender

\ Case Study

How Reality Defender Exposed Political Misinformation in Canada

\ Case Study

How Reality Defender Exposed Political Misinformation in Canada

Note: All names and PII have been removed to protect confidentiality.

As deepfake and generative content technology increases in complexity and prominence, so too are the bad actors using them for misinformation. Utilizing everything from off-the-shelf technology (to render “cheapfakes,” or low-to-mid-quality deepfakes) to resource-intensive yet high-quality video manipulations, politically-motivated bad actors are increasingly disrupting the news cycle.

Armed with convincing-enough deepfaked videos and photos, news media and social platforms are now rife with content purportedly showing politicians and key notable figures speaking and engaging in disparaging ways that never actually happened. Such is the case of a senior Canadian government official who was made a target of a deepfake-oriented disinformation campaign.

\ 01

The Challenges

A senior Canadian government official in Canada made the news rounds in late 2022 after a series of short videos appeared on social media depicting them and an associate mocking indigenous Canadians. The government official initially apologized for the media, while simultaneously denying that such incidents ever occurred. Having made the rounds for a few days, the videos painted the victim, now in the private sector, in an unflattering light, one that left little room for interpretation past first glance.

The immediate blowback was felt across all tier-one social channels, which dealt a serious blow to the senior official's personal and professional reputation.

> **With their name left in tatters overnight, the government official considered the idea that a bad actor had deepfaked their likeness and voice (as well as that of their associate) to create the videos.**

As they were a public figure with many recorded speeches and videos, using just a minute of their voice and a small sample of video from their speeches would be enough to generate convincing-enough deepfakes like the supposed deepfaked videos making the rounds.

This sounded like the most likely hypothesis — the creation of falsified videos using advanced technologies — leading the government official to engage with Reality Defender to run said videos through our ensemble of detection models and accept whatever results they generated.

\ 02

The Solutions

After a quick call with the client, Reality Defender gave them access to our web application. This afforded the corresponding Canadian governing branch the ability to upload the videos to the Reality Defender platform and detect the supposed video and audio deepfakes. Upon doing this, the government official discovered the four videos to be probabilistically fake.

Instead of providing a definitive “yes” or “no” answer as to whether media contains deepfaked material, Reality Defender’s models each give probability scores (ranging from 1 to 99) on the validity of uploaded media. This allows clients to export Reality Defender’s findings in CSV and PDF form and weigh the results.

Our models use the latest existing deepfake models, along with theoretical models, to determine the authenticity of media. This allows for Reality Defender to not only provide up-to-date detection on the newest and most current methods of deepfaking; it also allows us to stay one step ahead of new and upcoming models that arise every week.

\ 03

The Advantages

In this specific instance, all four videos and their respective audio tracks scored high, meaning their probability of being deepfaked video and audio was extremely likely.

> **This gave sufficient evidence to the senior government official's claims that the media depicting them and their associate never existed in the first place, allowing them to provide a strong retort to the social media and press coverage given to the videos.**

After conferring with Reality Defender over the findings, the senior government official then released the exported PDFs to tier-one media corporations in Canada, who contacted Reality Defender to further verify our platform's findings. Our CEO, Ben Colman, was featured in the press developments that followed to discuss how our platform works, how we grade suspected deepfakes, and whether or not the videos submitted by the senior government official and their team were falsified.

\ 04

Conclusions

In the months since, the senior government official was targeted by bad actors seeking to blackmail them with further misinformation, lending further credence to the results generated by the Reality Defender platform. Due to the prevalence of the videos before the senior government official contacted our team, the findings of the videos as deepfakes were reported less than the news of the initial videos.

> **This incident is not unique with deepfake-centered stories in news media, as the damage done by deepfakes is far more controversial than the revelation of deepfakes being involved.**

This highlights the dire need for media institutions, social media platforms, and government agencies to scan deepfakes at the onset before they can cause any potential harm. Such an implementation would give extra power to a media outlet's research and fact-checking abilities, while allowing social platforms and government agencies to pinpoint bad actors before they cause any irreparable harm and sow discord.

Should the government official see similarly targeted campaigns against them in the future, they and their team can continue to rely on Reality Defender to disprove (or verify) any media used to portray them in an unfavorable light.

As we continue to address newer deepfake models and bolster our detection algorithms against future ones, we are ready to assist politicians in combating any misinformation and work by bad actors targeting them.

Reality Defender

realitydefender.com
ask@realitydefender.com