Artificial Intelligence & Disinformation Campaigns

Dylan W. Wheeler
Departments of Computer Science & Philosophy, University of New Hampshire

University of New Hampshire

Introduction

Disinformation is false information that is **deliberately intended to mislead** people.

Computer-generated imagery (CGI) has traditionally been **very expensive** and only used in high-budget Hollywood movies.

New advancements in AI allow anyone to **generate videos of anyone saying anything.**

Influencers

People:

Nick Bostrom, *Swedish philosopher*
Yuval Noah Harari, *Israeli historian*
Demis Hassabis, *British video game designer*
Ray Kurzweil, *American inventor*
Yann LeCun, *French computer scientist*
Richard Sutton, *Canadian computer scientist*

Organizations:

DARPA, *Government agency*
Google, *Technology company*

Contact Me!

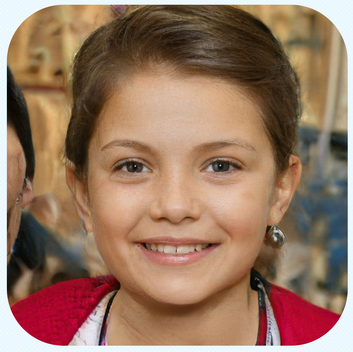
We're all in this together. Please reach out if you'd like to discuss more!

dww1002@wildcats.unh.edu
linkedin.com/in/dylan-wheeler

These People Do Not Exist

thispersondoesnotexist.com



Research Questions

- To what extent are algorithms being developed to **combat the spread** of disinformation and what is their effectiveness?
- What threats does disinformation pose to our **democracy** and the global economy?
- What can **policymakers** do to ensure AI is not used to spread disinformation?
- Can **market forces** adequately address these issues, or does it require governmental—and even global—regulation?

Discussion

- Does social media play a role in the way you receive and digest news?
- Do you feel like you can **trust social media**?
- Why would people **want to deceive** you by faking news media?
- How might we **identify factual information** when anyone with a smartphone can fake articles, audio, images, and videos?
- Is it possible to **ban or regulate** this technology?
- Can we trust ourselves to vote accurately in elections when **we've all been exposed** to disinformation?

The Algorithm

Generative adversarial network (GAN) - a class of machine learning systems and a form of unsupervised learning.

Two neural networks (a disinformation generator and a detector) contest in a zero-sum game framework causing both to **improve rapidly**.

Future Research

I will conduct this research through a paid UNH Summer Undergraduate Research Fellowship.

The technological and philosophical components of this phenomenon will become the basis for my senior thesis.