



AI-generated news & misinformation during elections

Paige Morrow, ARTICLE 19

ARTICLE 19

Global organisation promoting freedom of expression.





Uses

→ Investigative journalism

[ProPublica](#) maps Senators' pet issues.
NY Times does image recognition of Congress members.

→ Automated reporting

Geo-targeting, WP's Heliograf election day coverage 2016.

→ Back-end

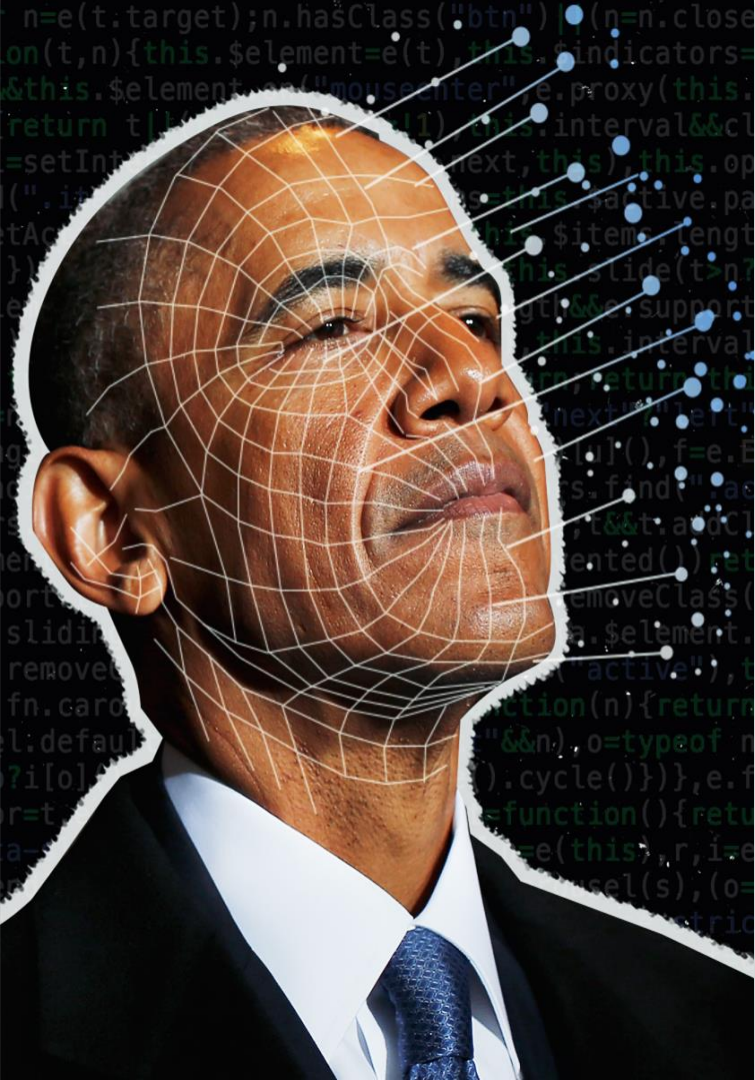
Content moderation, ads (inclu. RTB),
personalising content, content
distribution via social media

→ Fact-checking

What human rights issues are raised?

(focusing on freedom of expression)

1. **Bias.** Difficult to detect & mitigate bias within training data, e.g. data set lacking gender/race parity. (Qualitative evaluation + model card helps.)
 2. **Uncertainty.** Journalists need to verify outcomes. Algorithm may overlook important info or return false positives.
-



Misinformation

'Fake news'

- GPT-2 model generates (more) convincing text. Credibility score 6.91/10.
- Detection is challenging. OpenAI ~95% detection.

Deep/shallow fakes

- (Rare) political videos, e.g. Nancy Pelosi slurring
- Bot social media accounts with fake photos.
- **Response:** California banned deepfakes pre-elections (AB 730).

What are people saying?

Disinfo erodes trust in institutions & media & harms democracy through less informed decisions.

European Commission

Democracy is at risk from malicious & relentless targeting with disinformation and personalised 'dark adverts' from unidentifiable sources.

UK DCMS Committee on Disinfo

Americans rate misinfo as a bigger problem than racism, climate change or terrorism.

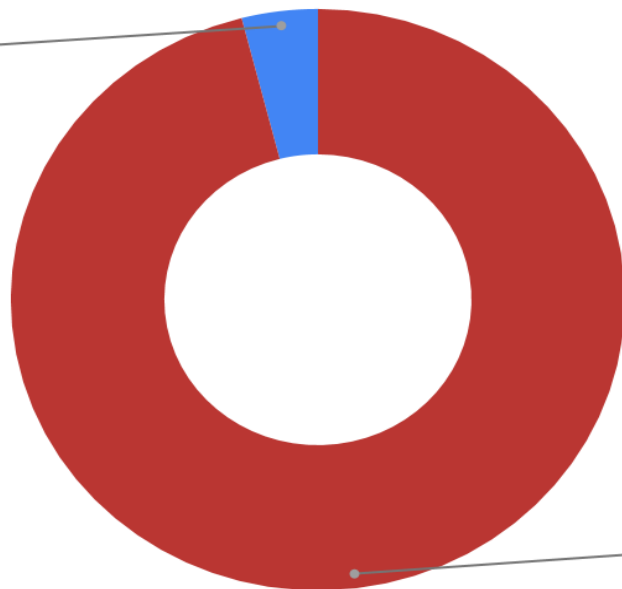
Pew Research Center Study 2019

Time to panic?

14,678

Total number of deepfake videos online

Political
4.0%



Pornography
96.0%

—

**What percentage of UK adults
are trapped in a
filter bubble?**



—

Just eight.

(According to a representative sample of 2,000 UK adults)*

*Elizabeth Dubois & Grant Blank (2018) The echo chamber is overstated: the moderating effect of political interest and diverse media, *Information, Communication & Society*, 21:5, 729-745, DOI: [10.1080/1369118X.2018.1428656](https://doi.org/10.1080/1369118X.2018.1428656)



Challenges

- Reliance on ad (RTB) revenue
- Epistemic closure
- Social media as secondary gatekeeper

(Quasi-)regulatory responses

Germany:

- NetzDG
- Proposals on media pluralism

France:

- Law on combatting the manipulation of info

United Kingdom

- White Paper on Online Harms 'duty of care'

EU Code of Practice on disinfo

International Grand Committee on disinfo & fake news

Pluralism in social media markets: potential solutions

1. Regulated pluralism

Impose *must-see* obligations on dominant platforms

1. (Unbundled) access to content moderation

Functionally separate hosting from content moderation

—

Thank you

paige@article19.org

