



Obfuscating Transparency:

The Promise and Risks of
Defusing Deepfake Disinformation through Blockchain

Annette Lee

What are deepfakes?

- **Generally, refer to realistic videos created with machine-learning algorithms** (Westerlund 2019; Vaccari and Chadwick 2020; Wilkerson 2021; Paradis 2022)
- **May also include audio, images, text** (Chesney and Citron 2018; Kirchengast 2020; Venema 2023; Byman et al. 2023, Anand and Bianco 2021)
- **Encompass several methods and algorithms** (Seow et al. 2022)



Deepfakes and Challenges to Peacebuilding

- Deepfakes disinformation: vivid, emotional, realistic, hard to debunk
 - Powerful people
 - Incitement of violence, support of violence, especially building on existing tensions (Zelenskyy, Putin)
 - Allegations that something is a deepfake (Gabon)
- Indirect: propaganda, sabotage, blackmail, weakening diplomatic ties, trust, polarization, autocratization, liar's dividend
- Largely speculative research





Deepfake Disinformation and Peacebuilding: Conditions

- Rate of technological advancement
- Access
- Messaging
- Private sector voluntary action
- Legislation and regulation
- Education and training
- Political polarization

What is blockchain?



- **Record data on an unchangeable, shared, and decentralized shared database in a fixed order**
- **Proof of Authenticity:** provenance, metadata, creator control
- **Smart Contracts:** execute following agreements in the code

Blockchain and Deepfake Disinformation

- Prove media is fake vs prove media is authentic: origin/provenance
 - Tracing media over time and across different sources
 - Creator control: authorized modification using smart contracts
 - Combine with watermarks for embedding identifying info
 - Example: Fox Corp. and Polygon Labs; Reuters and Canon; Truepic
- Potential limitations: societal buy-in, technical effectiveness
 - Critical mass needs to adopt; access; social norms around truth
 - Energy cost; ownership risks; storage limits; security risks; confidentiality

Blockchain Could Make Deepfake Disinformation More Harmful!

- Greater confusion and distrust about information: skepticism about real information, could cover up actual disinformation
- Further marginalization of the most at-risk: persecution of activists and media who can't risk such full transparency
- Further institutionalization of harmful power dynamics, norms, constraints

Decentralization doesn't mean equality.

Recommendations

- Plan for the long-term and for sustainability
- Don't treat it as a silver bullet, but rather as one tool in critical thinking
- Adopt high standards of transparency, which are currently lacking
- Carefully design to accommodate authentic additions and modifications
- Allow anonymization of certain information
- Increase public awareness of blockchain's function and its vulnerabilities and risks
- Consider power dynamics and coercion, reject claims of neutrality, make values explicit