

How to protect UK general elections from AI generated disinformation

Professor Steven Wheatley

Policy Context

Election interference is not new, but techniques are emerging that require the United Kingdom to clarify its position on the international law of state cyber operations that target elections.

The [UK's National Cyber Security Centre](#) says that emerging technologies, such as AI generated disinformation, are cheaper, easier to target millions of voters, and have a lower risk of detection. It is important that the UK has a clear legal framework regarding techniques that target elections.

Key research findings

AI generated disinformation poses a danger for UK elections. There is no specific international law rule against election interference, but the non-intervention rule of international law provides a firm basis to establish that the use of AI-generated disinformation is unlawful under international law.

- AI generated disinformation can weaken social cohesion, undermine trust in democracy and influence the outcome of elections. However, the international law on AI generated disinformation is unclear.
- One way of understanding if AI generated disinformation is unlawful is by using the non-intervention rule of international law.
- The non- intervention rule forbids states from intervening in the internal affairs of other states.
- The non-intervention rule states that intervention is unlawful when it uses methods of coercion. This means getting the target to do something that they would not otherwise do, leaving them no choice in the matter.
- The non- intervention rule can be applied to AI generated disinformation by asking two questions: was the information communicated with the intention of deceiving the audience into believing that something false is actually true? Would a reasonable observer conclude that the communication was intended to influence the audience's decision making to such an extent they would be left without a meaningful choice about what to think and therefore what to do?
- If the answer to both questions is 'yes' then the AI generated disinformation is a violation of the non-intervention rule because it is intended to get the target audience to do something they would not otherwise do and leave them with no choice in the matter.

Key Research Findings

Applying the non intervention rule to AI generated disinformation reveals that:

- State cyber influence operations that use AI generated disinformation are generally unlawful as a matter of international law.
- State cyber influence operations that promote racial or religious hatred are always unlawful as a matter of international law.
- But State cyber influence operations that place true information into the public domain are not unlawful as a matter of international law.

Policy recommendations

- The UK Government should support a public information campaign to explain the problem and illegality of State-sponsored disinformation campaigns
- The UK Government should clarify its position on international law in relation to AI generated disinformation targeting elections, making clear that AI disinformation is illegal under international law.
- Clarification includes establishing the rules of international law in this area, including how the non-intervention rule applies to AI generated disinformation.

Work with me

Professor Steven Wheatley is Professor of International Law at Lancaster University. Professor Wheatley has two principal areas of research interest: human rights theory and democracy.

Contact Professor Wheatley at s.wheatley@lancaster.ac.uk if you would like to learn more about their research, invite them to speak at your event, or collaborate with them.