Digital Literacy Education in the Era of Artificial Intelligence (AI): Learn to Discern AI-Generated

Deepfakes

Project Number: 6000856

Principal Investigator: Prof. Guanxiong HUANG

Grant Type: TDG

Abstract:

Misinformation, disinformation, and fake news has been rampant on the Internet. With the aid of AI-driven technologies, videos can be easily manipulated and have become a major media format for misinformation distribution on social media. For example, a deepfake video portrays Meta CEO Mark Zuckerberg saying, "Imagine this for a second: One man, with total control of billions of people's stolen data, all their secrets, their lives, their futures," was widely circulated on social media. But in fact, this video was created by using a footage of Zuckerberg made in 2017 and manipulating Zuckerberg's face with AI technology to make it appear he said something he did not say. Deepfake videos like this may accelerate the spread of disinformation and propaganda, posing a significant threat to citizens and society. It is thus essential to equip citizens with digital literacy knowledge and skills that enable them to discern AI enabled deepfakes.

To address this urgent issue, this project aims to understand college students' perceptions and knowledge about AI-generated deepfakes and develop and evaluate digital literacy education material that can be used to improve college students' ability to discern AI enabled deepfakes. This project will be implemented in three phases: (1) Phase 1: focus groups will be conducted to understand how college students perceive AI-generated deepfakes, how accurately they are able to judge the authenticity of AI-generated videos, and what method would be the most effective to improve their discernment; (2) Phase 2: based on the findings of the focus groups,

multimedia instructional material that teach students how to discern AI-enabled deepfakes will be created and pretested; (3) Phase 3: an experiment will be conducted in undergraduate classes to examine the effectiveness of the multimedia instructional material in terms of improving students' digital literacy and promoting deepfake discernment.

As such, the proposed project will generate education material for improving AI deepfake discernment, which can be incorporated into the digital and information literacy curriculum for general education. The digital literacy training program will build students capabilities to recognize and resist the incoming misinformation and better navigate in today's information environments.