From Deepfakes to Research Ethics

Today I'm going to focus on the ethics of science. Imagine that you notice a video on social media. In this video, president of the United States declares war on China. You do not believe it and replay the video time and again, but everything there is so real: the voice, the facial expression, the gestures, and even the motion of his mouth: flawless, this must be real. You collapse in terror, fearing that your fatherland will soon be razed to the ground.

Welcome to the age of disinformation. Nowadays malicious production of fake videos has been made much easier than before, all because of a small piece of codes made public by three computer scientists in 2017. And today, I am going to condemn this kind of scientists who publish their work without consideration of consequences, and I shall argue that **scientific research should not be published without limit, but be subject to evaluation of its social impact**.

Everyone knows that any kind of technology has its own virtue and vice. But for many people in the academia, this universal two-sidedness of technology simply becomes a justification for any kind of research. For many, the existence of potential hazard of a future technology is only considered as something of minor importance. They are never deterred by the negative impact of their research, because they always believe there will be a workaround. This line of thinking might be valid in the pre-Internet age. However, as we are entering an age dominated by the Internet, the seriousness of this problem escalates to an unprecedented level.

Let's go back to our opening example. The so-called deep fake video technology is originally developed for creating cinematic effect in movie industry. However, what the researchers didn't take into consideration is that once their work is published, everyone with a personal computer can utilize this invention and make fake videos that will corrupt the social media. The openness and anonymity of the Internet offers great opportunity of weaponizing technologies that rely on computers. What's more disturbing is that this kind of computer-based technologies are oftentimes difficult to counteract once they are turned into malicious use. In this example, due to the peculiarity of AI, any effort to detect fake videos will in return make them increasingly resistant to detection, just like antibiotics will not kill some bacteria but make them stronger. Once the Pandora's Box is opened, there is no returning.

This is the curse of the Internet. Misuse of technology has never been so easy and so hazardous. It's time, therefore, for researchers and academic reviewers to realize the gravity of this issue. Do not allow any paper to be published unless the negative impact is estimable and acceptable. Otherwise, a technological dystopia might be drawing near.