digital

Is AI a New Threat to Digital Safety?

Is AI a New Threat to Digital Safety?

Let me walk you through a scenario: Imagine waking up one day to an overwhelming number of notifications on your phone, and the incoming calls are becoming unmanageable. You start thinking that something terrible may have happened, perhaps someone has died. You decide to answer one of the calls, and the first thing the caller asks is, "Are you okay?" Confused by such a question, you wonder what could be going on. Then they continue, "I saw the picture how did it get out?" Now you're even more bewildered. What picture are they talking about? They then send you what they've seen circulating on social media, and to your shock, the image shows a naked person with your face. You don't recall ever taking such a picture, and the body in the photo isn't yours. And then you recall someone mentioned AI-generated content.

Over the past year, artificial intelligence (AI) has become a key topic of conversation in social media, workshops, etc. The discussion revolves only around its potential to revolutionize various sectors and often leaves out how it's being misused. While AI has brought about many advancements, it is increasingly being used to infringe on digital users' rights. From generating fake music mimicking popular artists to the deeply troubling rise of AI-generated pornographic content featuring the faces of real individuals, AI's misuse is raising serious questions about digital safety.

The Dark Side of AI Content Creation

AI was the talk of the town, celebrated for its creative capabilities and how it simplifies one's work, for example in this article, the author was excited to use AI to enhance their work and flow of the content, what a brilliant technology. Music lovers were excited by AI models that could generate songs in the style of their favorite artists, the likes of Drake, Rihanna, and Dua Lipa. What began as an exciting technological breakthrough soon revealed its dark side. People began using AI-generated music to imitate artists without consent, raising copyright infringement issues and damaging the reputation of original creators.

Not only the music industry has been affected by the misuse of AI. In a more alarming trend, AI has been weaponized to create deepfake content. Deepfakes are highly realistic videos or images where a person's face is superimposed onto someone else's body. Increasingly, AI is

being used to generate pornographic material featuring the faces of real people often without their knowledge or consent. This is a clear violation of privacy and personal safety, raising ethical, legal, and digital safety concerns.

Violations of Digital Rights and Privacy

Now, let's return to the scenario. Imagine that the picture has been seen by millions of people. Suddenly, you feel an overwhelming urge to delete all your social media accounts, missing out on the opportunities that often come with these platforms. You lose valuable connections, and perhaps that one project you were looking forward to collaborating on with a major company has now fallen through all because of a fake picture that has circulated.

Every digital citizen has a right to use digital platforms without feeling unsafe. The implications of AI misuse are far-reaching. When AI is used to create deepfakes or other harmful content, it directly infringes upon an individual's digital rights. The right to privacy, the right to control one's own image, and the right to safety in digital spaces are all compromised. In some cases, AI-generated content has been used to harass and exploit individuals, leading to significant emotional, psychological, and reputational damage.

Maybe the challenge has not yet emerged in Tanzania but surely it will. Celebrities and public figures are frequent victims of this type of abuse. Once someone's image is manipulated in such a way, it is difficult to undo the harm. The content spreads quickly online, and laws have struggled to keep pace with the rapid evolution of AI technology.

Nevertheless, let's address the issue of misinformation. AI-generated content has increasingly been used to spread misinformation, particularly in the political arena. As we approach local and general elections, AI is often misused to create false information aimed at fueling propaganda. These AI-driven manipulations can take the form of fabricated speeches, altered images, or deepfake videos, all designed to mislead the public and distort the truth. The spread of such content can have serious consequences, influencing voter opinions and undermining the democratic process.

What do the laws and policies say?

It is alarming that Tanzania has not yet introduced AI-specific laws and policies, or at least incorporated them into the existing laws. The fast-growing technology with numerous numbers of users is yet to be regulated. However, several existing laws and policies including the Cybercrime Act 2015, do not specifically address the issue of AI, but it has covered a broad range of offenses such as unauthorized access to computer systems, cyberbullying, online harassment, and data breaches, these offenses are relevant to the provisions of the cybercrime act.

The case of Electronic and Postal Communication Act (EPOCA) 2010 governs telecommunications services and could apply to AI-driven communication platforms or services. Any misuse of AI for illegal content distribution, online harassment, or privacy violations could fall under this law's provisions. When an AI system processes personal data without consent or in violation of privacy rights, this law can address the offense. Although AI is not specifically addressed in Tanzania's National ICT Policy 2006, it stresses the importance of innovation while advocating for digital security, privacy, and the responsible use of technology. This policy provides a framework for future AI-related regulations in sectors such as healthcare, agriculture, and education.

Despite our laws seeming sufficient to address offenses, public mistrust in AI technologies may grow, hindering adoption and acceptance and complicating enforcement efforts. The absence of a comprehensive legal framework for AI poses risks to individuals and society, undermining the technology's potential benefits while exposing users to harm.

What other countries are doing

How do other countries address the issue? The U.S. has created various sector-specific agencies and organizations dedicated to tackling some of the challenges associated with the advancement of AI. For example, California passed a law making it illegal to use deepfake technology for malicious purposes, such as defaming or harassing individuals.

European Union has proposed the AI Act which aims to establish strict guidelines for the collection, use, and retention of personal data. This act is designed to enhance the EU's control over AI development and usage while prioritizing transparency, accountability, and ethical principles to address user concerns. On the other hand, The Chinese Cybersecurity Law and the New Generation AI Development Plan outline strategies for data protection and cybersecurity in AI, highlighting the importance of compliance and proactive risk management.

What LP Digital has been doing

Over the past few years, LP Digital has remained committed to raising awareness about digital rights and digital safety. Through our platforms, we have consistently communicated the laws and policies governing the use of digital platforms, as well as various aspects of digital rights. In collaboration with the Tanzania Digital Rights Coalition (TDRC), which was formed to create a unified voice for digital rights advocacy in Tanzania and to promote safe, inclusive, and equitable digital spaces, we have worked to advocate for changes in laws and policies that reflect current technological developments posing threats to digital safety, including data protection.

Additionally, we have been actively educating digital platforms users,law enforcement, including the police, on emerging digital trends to help them stay informed and equipped to handle these challenges. We have also built a strong relationship with the Tanzania Communications Regulatory Authority (TCRA), working together to ensure a safer digital environment for all.

Understanding the magnitude of the problem, LP Digital has established a cyber helpline to provide immediate support and resources for individuals facing digital safety issues. This helpline aims to assist users in navigating challenges related to online threats, misinformation, and privacy concerns, ensuring they have access to the necessary guidance and protection.

Recommendations on What Needs to be Done?

- It is important for government agencies to work closely with digital rights activists, civil society organizations, and academic institutions to address AI-related challenges collectively.
- Engage in international dialogues to create standardied regulations for new emerging technologies like Artificial Intelligence.
- Tanzania should participate in global efforst to establish ethical quidelines for AI development.
- Establish clear mechanisms for individuals to report violations of digital rights and seek redress for harms caused by AI technologies.
- There should be calls for stricter regulations on the use of AI, particularly in the areas of deepfakes and content manipulation. In some countries, laws are being drafted to criminalize the creation and distribution of non-consensual AI-generated content.
- Tech companies have a responsibility to implement safeguards to prevent the misuse of their AI tools. This could include stricter guidelines on content creation, better detection algorithms for deepfakes, and policies that hold users accountable for malicious uses of AI.
- At the individual level, awareness is key. Users must be cautious about the content they share online and stay informed about how their digital rights can be infringed. As AI continues to develop, the boundaries between ethical use and abuse will need to be constantly monitored and reevaluated.

Conclusion

While AI has enormous potential to benefit society, its misuse is causing a significant threat to digital safety. The use of AI to infringe on users' digital rights whether through generating fake music or creating harmful deepfakes demands serious attention. Without proper regulations, safeguards, and awareness, the unchecked development of AI could lead to further digital rights violations, making it crucial for all stakeholders to take immediate action to address this emerging threat.