CMSGu2024-01

Computer Emergency Response Team of Mauritius

Enhancing Cyber Security in Mauritius

Guideline on Cybersecurity Ethics and Artificial Intelligence (AI)

CERT-MU

Mauritius

Version 1.0

January 2024

Issue No. 1

Table of Contents

1.0 Introduction....4
1.1 Purpose and Scope....4
1.2 Audience....4
1.3 Document Structure....4
2.0 Background....5
3.0 Cybersecurity Ethics....6
4.0 Responsible AI....7
5.0 Some of the most important ethical challenges of AI....8
6.0 How to get started using generative AI responsibly....10
7.0 Conclusion....12
8.0 References....13

DISCLAIMER: *This guideline is provided "as is" for informational purposes only. Information in this guideline, including references, is subject to change without notice. The products mentioned herein are the trademarks of their respective owners.*

1.0 Introduction

1.1 Purpose and Scope

The purpose of this guideline is to give readers an overview of how they can make use of generative AI responsibly and ethically.

1.2 Audience

The target audience for this guideline is the general public.

1.3 Document Structure

This document is organised into the following sections:

Section 1 gives an outline of the document's content, the targeted audience and the document's structure.

Section 2 presents a background on AI the its responsible use.

Section 3 talks about cybersecurity ethics.

Section 4 elucidates the responsible use of AI.

Section 5 describes some of the key ethical challenges of AI.

Section 6 provides a few tips on how to get started using generative AI responsibly.

Section 7 concludes the document.

2.0 Background

Artificial Intelligence (AI) has become a vital tool in cybersecurity due to its ability to detect and prevent cyber-attacks more efficiently than traditional methods. However, the use of AI in cybersecurity has raised various ethical concerns that must be addressed. Ethical AI refers to AI that follows clear ethical guidelines. These guidelines are based on important values including individual rights, privacy, fairness, and avoiding manipulation. When organizations use ethical AI, they have well-defined policies and review processes to make sure they are following these guidelines. Ethical AI goes beyond just what is allowed by law. While laws set the minimum acceptable standards for AI use, it sets even higher standards to respect fundamental human values. Responsible AI is the practice of designing, developing, and deploying AI with good intention to empower employees and businesses, and fairly impact customers and society, allowing companies to engender trust and scale AI with confidence.

3.0 Cybersecurity Ethics

- **Privacy and Data Protection**

  The cornerstone of cybersecurity ethics lies in the protection of privacy and personal data. As digital interactions become increasingly pervasive, individuals are entrusting a vast array of sensitive information to online platforms. Ethical cybersecurity practices dictate that this data be collected, stored, and processed in a manner that safeguards the privacy and security of individuals.

- **Transparency and Accountability**

  Transparency and accountability are fundamental principles that underpin ethical cybersecurity. Organizations are expected to be forthright about their cybersecurity measures and practices. In the event of a breach or incident, there is a moral imperative to promptly notify affected parties and take responsibility for the breach.

- **Equity and Fairness**

  Cybersecurity should be applied equitably, providing equal protection to all individuals and groups. This extends to vulnerable populations who may be at greater risk due to socio-economic, geographical, or demographic factors. Ethical cybersecurity practices work to bridge gaps in security vulnerabilities.

- **Ethical Hacking and Testing**

  Ethical hacking, or penetration testing, is an integral aspect of cybersecurity ethics. It involves systematically identifying vulnerabilities in systems, networks, and applications. This is conducted with explicit permission and within legal and ethical boundaries, contributing to the fortification of digital security.

- **Education and Awareness**

  Promoting cybersecurity awareness and education is an ethical responsibility that extends to organizations, educational institutions, and the broader public. By empowering individuals with the knowledge and skills to protect themselves online, the collective digital community becomes more resilient against cyber threats.

4.0 Responsible AI

- **Transparency and Explainability**

  In the realm of AI, transparency and explainability are paramount. Responsible AI practices require that users have insight into the decision-making processes of AI systems. This not only fosters trust but also allows for meaningful human oversight.

- **Accountability and Oversight**

  Clear lines of accountability must be established for AI systems. Developers, operators, and organizations are responsible for the outcomes of their AI applications. This accountability framework ensures that AI technologies are deployed responsibly and ethically.

- **Sustainability**

  Responsible AI practices extend to the environmental impact of AI systems. Efforts are made to ensure that AI technologies are energy-efficient and do not contribute excessively to resource consumption.

- **Human Oversight and Intervention**

  While AI systems can automate numerous tasks, human oversight and intervention remain critical, especially in high-stakes situations. This ensures that AI operates within ethical boundaries and that humans retain ultimate decision-making authority.

- **Security and Privacy**

  Responsible AI includes robust security measures to protect against malicious attacks or misuse of AI systems. Furthermore, it adheres to data protection regulations, respecting user privacy and confidentiality.

5.0 Some of the most important ethical challenges of AI

- **AI bias**

  This refers to the potential tendency of algorithms to produce results that reflect and replicate human biases. Not only do they reflect human biases, but they also give them an objective status in a way that can be perceived by others as scientific and credible. AI bias risks the reinforcement of existing biases and stereotypes in the population. It can be hurtful to the already marginalized, vulnerable, and disadvantaged group of people.

- **AI's impact on jobs**

  AI is affecting the workforce and many people are concerned that it will lead to an increased rate of unemployment as some job positions are being replaced by machines. However, this concern should be well analyzed as the transformation in the job market is mostly taking a shift from specific roles to another, rather than job loss. The fact that many new jobs and tasks will emerge from AI should also be taken into consideration. AI has still limited capabilities, therefore, it is unlikely to replace many jobs. Furthermore, it requires humans input and it depends on them.

- **AI and privacy**

  One of the most concerning things about AI is privacy which is a fundamental human right. Many technologies like cameras, smartphones, the internet, facial recognition, and other digital services have made it easier to collect personal data which can be used for unintended or malicious purposes. In order to prevent such issues, organizations are implementing privacy information management systems and governments have developed many regulations and laws, such as GDPR, CCPA, etc.

- **AI and humanity**

  AI is shifting the way people behave and interact with each other. Many AI bots are being used to model human conversations and build relationships. However, it is debatable whether this is a positive or negative thing as many people have started building relationships with machines, even though they cannot replace many human features.

- **AI and security**

  AI has the potential to face many security threats, including model manipulation and poisoning, data privacy, data tampering, insider threats, deliberate attacks, mass adoption, online manipulation, and vulnerability to attacks. relationships with machines, even though they cannot replace many human features.

6.0 How to get started using generative AI responsibly

- **Set risk-based priorities**

  Some generative AI risks are more important to your stakeholders than others. Adjust or establish escalation frameworks so that governance, compliance, risk, internal audit and AI teams give the greatest attention to the greatest risks.

- **Revamp cyber, data and privacy protections**

  Update cybersecurity, data governance and privacy protocols to help mitigate the risks of malicious actors' generative AI inferring private data, unravelling identities or conducting cyberattacks.

- **Address opacity risk**

  With some generative AI systems, explainability is not an option. It is impossible to unravel "why" a certain system produced a certain output. Identify these systems, consider what practices can help support their fairness, accuracy and compliance, and tread carefully when oversight is impossible or impractical.

- **Equip stakeholders for responsible use and oversight**

  Teach employees who may need to use generative AI the basics of how it works, when and how to use it, and when and how to verify or modify outputs. Provide compliance and legal teams with skills and software to identify intellectual property violations and other related risks.

- **Monitor third parties**

  Know which of your vendors provide content or services that use generative AI, how they manage the related risks and what your possible exposure may be.

- **Watch the regulatory landscape**

  Policymakers around the world are issuing more and more guidance on AI development and usage. This guidance is still a patchwork, not a complete regulatory framework, but new rules are continually emerging, especially regarding AI's impact on privacy, AI bias and how AI should be governed.

- **Add automated oversight**

  With generative AI-created content ever more common, consider emerging software tools to identify AI-generated content, verify its output, assess it for bias or privacy violations and add warnings as needed.

7.0 Conclusion

The use of AI in cybersecurity presents numerous ethical considerations that must be addressed. Key considerations include avoiding bias and discrimination in AI algorithms, ensuring transparency, protecting data privacy and security, mitigating the impact on jobs, promoting responsibility and accountability, and developing legal and regulatory frameworks that protect individual rights and freedoms. Prioritizing fairness, accountability, transparency, and responsibility in the use of AI can improve cybersecurity while also protecting the rights and dignity of individuals. By doing so, we can ensure that AI is used to improve cybersecurity while also protecting the rights and dignity of individuals.

8.0 References

- https://medium.com
- https://www.accenture.com/
- https://builtin.com/artificial-intelligence
- https://pecb.com