

General Assembly Topic

Data Security

Personal data often includes sensitive information, such as financial data, medical history, addresses, etc. In a world largely dominated by technology, the protection of this data has become more important over the past few decades. The U.N. has begun to pass measures to protect personal data, such as “The Principles on Personal Data Protection and Privacy,” which include three main aims:

1. Harmonize standards for the protection of personal data across the U.N.: all systems that protect privacy do not contradict each other
2. Facilitate accountable processing: controllers of data are being responsible and complying with data protection principles.
3. Ensure respect for the human rights and fundamental freedoms of individuals, in particular the right to privacy

However, the current implementation of this is far from comprehensive on a global scale. Many countries have begun to pass their own legislation for the protection of personal data and private information. Today, 71% of countries have some form of legislation in regards to Data Protection and Privacy, 9% have working drafts of legislation, and 15% have no legislation. One example of this is the European Union’s (EU) General Data Protection and Regulation (GDPR). Under this law, people are forced to comply with seven guidelines when handling data belonging to citizens of the EU: data holders must be transparent with their clients, specify the reason for processing their data, collect the absolute minimum amount of data necessary, keep that data accurate and up to date, only store data for as long as necessary, and ensure integrity and hold responsibility for the data. In violation of this law, people can be fined up to 20 million euros or 4% of global revenue, depending on which is higher. This type of legislation has also been passed in other parts of the world, such as Brazil’s General Data Protection Law and Australia’s Privacy Act of 1988; the latter was amended in 2014 and 2017 to accommodate more modern technologies.

Even with these principles in place, however, hackers have been able to bypass programs globally, a clear threat to the cybersecurity of the world. For example, in February 2024, a data breach occurred within French health insurance companies that affected 33 million French citizens, or nearly half of the country’s population. This compromise, according to the Center for Strategic and International Studies (CSIS), leaked “sensitive birth date, social security, and marital status information.” This attack was on a smaller scale when compared to a 2022 cyber attack in Australia, though. The health insurance company Medibank was hacked, and the private information of about 9.7 million people was leaked. Furthermore, the sensitive healthcare claims data of around 480,000 individuals were being published and extorted by criminals on the black market. This included the information mentioned above, as well as prior medical history that includes drug addiction treatments and abortion records. The Medibank attack exposed a serious data security problem: though Medibank was based in Australia, it was hacked by Russian hacker Aleksandr Ermakov using ransomware developed by a Russian

hacking group called REvil. Cybersecurity attacks are not limited to borders—they can traverse them. Oftentimes, a faceless entity is carrying out the attack, making catching the culprit exponentially harder. This highlights the importance of developing proactive frameworks focusing not only on solving issues arising in the aftermath of attacks, but also minimizing how often they occur.

With the advancement of technology has come the advancement of Artificial Intelligence, which comes with its own set of concerns in regards to data security. Hackers have already started utilizing this with the use of two main strategies: prompt injection attacks and data poisoning attacks. Through prompt injection attacks, attackers create an input for an AI designed to make the AI behave in an unintended way. Data poisoning attacks are when an attacker alters the data that trains an AI model, therefore training it to create outputs the AI is not supposed to make. This method is similar to another method of hacking databases called an SQL injection attack. These types of attacks are when attackers inject malicious data into a system or database to retrieve information. In August, 2007, the 7-Eleven corporate servers were attacked, and two months later in October, the database sent approximately 2 million different credit card numbers to the hackers. While there have been no documented large-scale prompt injection and data poisoning attacks, the potential of AI could make it more efficient and quicker for hackers to carry out these attacks, and as of now there is little regulation on artificial intelligence because of how new the technology is. It is crucial to recognize the dangerous potential of AI.

AI has also led to a rise in Deepfake technology, which could cause threats of financial and identity fraud. A Deepfake is a video in which AI creates believable and realistic photos, videos, and other media of people and events that never occurred in real life. These videos could be used as scams to get information out of either official businesses and corporations, or even friends and family of the victim. In some instances, attackers will use Deepfakes of certain individuals and contact their families, staging a fake kidnapping. The attackers will ask for ransom, and use the Deepfake to convince them that they are giving money or other sensitive information to save their family, not only giving their info to the hacker, but also allowing for it to circulate throughout the dark web. Another example is using Deepfakes to bypass voice authorization in certain institutions. For example, if a bank uses voice recognition to access an account, by using information from the dark web and Deepfake technology, hackers can access individuals' financial information, a serious personal data breach. Although Deepfakes have been around for quite some time, the introduction of AI has made making Deepfakes easier and has made the technology itself more seamless and untraceable. This creates more opportunity for hackers to find and steal sensitive information from individuals around the world.

Questions to consider:

- Does your country have any policies in place around cybersecurity? If so, in what areas? Are they more proactive, more focused on the aftermath (penalties for cyber criminals), both, or none?
- What does your country consider a cyberattack?

- Has your country experienced any large-scale cyber attacks? How often? How were they handled and managed? What was the exploited data used for?
- Were there any cases of cyber attacks in your country that involved demands to pay a ransom?
- What is the role of your country's government in cyber security? What is your country's stance and perception of it? Negative or positive?
- How has cybercrime affected your country's relationship with other nations?
- Are there any restrictions on AI within your country?
- How is AI production facilitated and/or authorized in your country? Is it government made and operated? Is it regulated? Is it made and managed by independent businesses?

Sources:

- <https://unsceb.org/privacy-principles>
- <https://unsceb.org/principles-personal-data-protection-and-privacy-listing>
- <https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents>
- <https://therecord.media/health-insurance-data-breach-affects-half-of-france-cnil>
- <https://krebsonsecurity.com/2024/01/who-is-alleged-medibank-hacker-aleksandr-ermakov/>
- <https://unctad.org/page/data-protection-and-privacy-legislation-worldwide>
- <https://gdpr.eu/what-is-gdpr/>
- <https://iapp.org/resources/article/brazilian-data-protection-law-lgpd-english-translation/>
- <https://www.ag.gov.au/rights-and-protections/privacy>
- <https://www.ncsc.gov.uk/guidance/ai-and-cyber-security-what-you-need-to-know>
- <https://www.justice.gov/iso/opa/resources/5182013725111217608630.pdf>
- https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf