Human-Centric Cybersecurity Partnership

Artificial Intelligence and Trust

SECURING DIGITAL IDENTITIES

YARA ALSIYAT,
MOHAMMADHOSSIEN AMOUIE,
ARIEL BURGESS,
NATALIE CHU,
PAOLA MARMORATO,
PULKIT MOGRA

Human-Centric Cybersecurity Partnership

HUMAN-CENTRIC CYBERSECURITY REPORT PROJECT

The 2024 Human-Centric Cybersecurity Report Project brought together postgraduate students from across Canada and the United Kingdom to work with our partners from both private industry and the public sector to produce reports looking at the problem of ransomware through a transdisciplinary lens.

ABOUT HC2P

The Human-Centric Cybersecurity Partnership (HC2P) is a transdisciplinary group of scholars, government, industry and not-for-profit partners that generate research and mobilize knowledge that will help create a safer, more secure, more democratic and more inclusive digital society.

ACKNOWLEDGEMENTS

We would like to thank Accenture, Bell Canada, The Canadian Centre for Cyber Security (CCCS), Canadian Cyber Threat Exchange (CCTX), Desjardins, Field Effects, GoSecure, Innovation, Science and Economic Development Canada, The National Bank of Canada, The National Cybercrime Coordination Centre (NC3), Public Safety Canada, The Royal Canadian Mounted Police, Statistics Canada, The University of Montreal, and The University of Ottawa for their efforts in supporting this project.

Cover Art by Michael Joyce

Cite as:

Alsiyat, Y., Amouie, M., Burgess, A., Chu, N., Marmorato, P., & Mogra, P. (2024). Artificial intelligence and trust: Securing digital identities. Human-Centric Cybersecurity Partnership (HC2P).Dépôt légal,

ISBN: 978-1-7387249-9-4

The Human-Centric Cybersecurity Partnership is supported in part by funding from the Social Sciences and Humanities Research Council.

Social Sciences and Humanities Research Council of Canada

Conseil de recherches en sciences humaines du Canada

Canada

Contents

Artificial Intelligence and Trust

SECURING DIGITAL IDENTITIES

Executive Summary

This report examines the critical role of Artificial Intelligence (AI) in digital identity management and highlights the need to address trust issues to enable its broader adoption in cybersecurity. AI offers significant potential to enhance the accuracy, speed, and scalability of digital identity systems by detecting and predicting unauthorized access more effectively than traditional methods. However, several barriers hinder its widespread adoption. These challenges include the opacity of AI decision-making processes, concerns over data privacy, biases within AI systems, security vulnerabilities, and the high costs of implementation. Additionally, the lack of clear regulatory frameworks exacerbates the trust deficit, causing users and organizations to hesitate to fully embrace AI solutions.

To address these challenges, the report proposes four key mechanisms to build trust in AI systems: the development of trustworthy AI technologies, human oversight, comprehensive regulation, and incentives. Trustworthy AI systems must prioritize transparency, fairness, and security. This transparency, combined with accountability

frameworks, can ensure that AI systems treat all users equitably. Human oversight through the establishment of review boards to validate AI models before deployment and giving users a say in how their digital identity data is used and shared will also be important. This oversight helps ensure that AI systems are not left unsupervised, reducing the risk of biases or security flaws going unnoticed.

Regulation plays a vital role in ensuring ethical and accountable AI deployment. The report emphasizes the need for cooperative efforts among public, private, and academic sectors to develop effective legal frameworks. These frameworks should promote accountability, protect against undue influence from external interests, and foster a competitive environment that encourages innovation. Furthermore, regulatory standards should prioritize the ethical sharing of knowledge and technological compatibility across sectors to ensure best practices in AI deployment. Finally, incentives are essential to stimulate AI development and adoption. Government funding, especially for small and medium-sized businesses, can help foster decentralized innovations in AI. Investments in law enforcement training on AI-driven cybercrime intelligence will improve the protection of individuals and organizations, reducing the risks associated with identity fraud and cyberattacks.

The report concludes that addressing these challenges requires a multifaceted approach that incorporates technological advancements, human governance, regulatory oversight, and financial support. By implementing these trust-building mechanisms, AI can become a reliable tool for enhancing digital identity security and improving cybersecurity practices across various industries.

“POOR DIGITAL IDENTITY MANAGEMENT POSES A SERIOUS RISK TO CYBERSECURITY...”

1 Introduction

Digital identity is crucial to securing access to online data, computer network systems, and online platforms. Storing and sharing sensitive information online, such as an individual’s personal and financial information, increases the risk of the illegitimate use of a digital identity for unauthorized access. Digital identity management is therefore a top priority to ensure that access to certain information and systems is only accessible by approved parties. Identification, authentication, authorization, and accountability processes are critical components of digital identity management in cybersecurity.

Digital identity management also poses significant challenges, as it relies upon large amounts of information, constant information flows, and individual user adaptability. This makes prioritization, outlier identification, and response difficult to achieve efficiently given the limitations of human cognition and currently employed technologies. Artificial intelligence (AI) can revolutionize digital identity management, as it can process vast amounts of data, recognize patterns, prioritize risks, and detect unusual activities. However, large-scale employment of AI technologies for digital identity management, and specifically for recognizing and predicting outliers of digital identity, is currently limited.

This report seeks to investigate both the benefits and barriers of AI adoption in the context of digital identity management, and specifically explores the relationship between AI and trust, guided by the questions:

- How can trustworthy AI minimize risks and address possible threats caused by misuse of digital identity?
- How do the challenges and potential

vulnerabilities associated with using AI for preventing unauthorized access inform trust building mechanisms for AI?

The report will first define key terms related to digital identity management and AI before reviewing what has currently been studied in this area to identify any gaps. The report then discusses the importance of identity management in cybersecurity, delving into the risks associated with outliers of digital identity, including digital identity theft. AI-driven solutions, and their barriers to adoption, are then explored. Mechanisms for trust building, including model requirements, human oversight, cooperative efforts, and funding, are then discussed, and specific recommendations for deploying each mechanism are provided.

The report seeks to provide an understanding of the relationship between AI adoption and trust for policymakers, industry leaders, researchers, and individuals, to enhance and improve usable cybersecurity in a variety of sectors.

1.1 Key Terminology

This section introduces and defines key terminology used throughout the report, including digital identity, authentication, outliers of digital identity, digital identity theft, artificial intelligence (AI), trust in AI, and human-oriented cybersecurity practices.

- Digital identity refers to the personal information, biometric data, and digital footprints associated with an individual, often authenticated through credentials such as passwords.
- Authentication is the process of verifying an individual's digital identity, often via credentials or a unique attribute like biometric data.
- Digital identity theft is the illegitimate access and/or use of a digital identity.
- For the purposes of this report, Artificial Intelligence (AI) is a machine trained on large amounts of input data to produce content, predictions, recommendations, or decisions.
- Trustworthy AI refers to the confidence that an individual user has in AI's ability to produce reliable, fair, and safe outcomes. This confidence is informed by an individual user's set of beliefs about AI, which may change over time.
- Human-Oriented Cybersecurity Practices are a set of measures that individuals, organizations, or cyber ecosystems adopt to maintain their cybersecurity, which focus on user behaviours and interactions with digital technologies and spaces.

1.2 Current Research on Digital Identity and AI

The authentication of an individual's digital identity, informed by personal information,

biometric data, and digital footprints, is essential to securing their online data and granting access to computer network systems and online platforms. The process of digital identity authentication involves the comparison of an artifact presented by an individual against a known unique record. This often involves verifying something that an individual uniquely knows via mechanisms such as usernames and passwords in login credentials, and/or by verifying an attribute unique to that individual, such as personal information or biometric data (Sule et al., 2021).

Problems with the authentication process may cause an individual to lose access to their online data and computer network systems. Alternatively, an individual's digital identity may be authenticated without the individual's permission, leading to unauthorized access to data and systems (Conklin et al., 2004). An illegitimate use of identity may be a result of successful digital identity theft made possible by digital trace data from online activities, or from a failure to adopt appropriate and effective human-oriented cybersecurity practices. This may include maintaining software and operating systems (via updates), securing login credentials, recognizing phishing attacks, using tools like antivirus software and firewalls, and monitoring data breaches (Gupta & Furnell, 2022).

Unfortunately, preventing misuse of digital identity is difficult, even with proper practices and protocols in place. Given the complexities involved in identifying and reacting to unauthorized use of computer systems and the heavy burden placed on individual users to maintain their own cybersecurity, many common mechanisms for managing this issue can be ineffective (Cremer et al., 2022).

Artificial Intelligence technologies offer new potential for data processing that could be applied to the challenge of digital identity authentication. Consequently, the use of AI presents opportunities to supplement human-oriented cybersecurity practices. For example, threat intelligence and hunting, which involves the identification and system-wide search for bad actors and threats, are mechanisms through which AI can be used to recognize and predict the unauthorized use of digital identities (IBM, 2021).

The implementation of these technologies is not without its challenges, as AI systems are subject to technology-related barriers such as data quality, interoperability, and the lack of standardized data formats, which can hinder the development and deployment of AI (Alami et al., 2024). Poor implementations of technology could lead to it being considered to be untrustworthy, which may in turn limit its deployment and the accompanying real-world testing and revision, which are important for the development of universally robust technologies. It is clear then that trust must be developed in AI systems implemented in a cybersecurity setting.

Building trust is dependent upon the characteristics and behaviours of an individual user that affect how they interact with a technology, the design and development of a technology, and external factors that impact how a technology is deployed (Adediji et al., 2022). A lack of trust in AI may be evident in beliefs about reliability, performance, and fairness of outcomes of AI models themselves (Mennella et al., 2024), or through distrust of a particular developer or owner of AI technologies.

The literature as reviewed suggests that AI offers substantial potential to improve digital identity

authentication by addressing the limitations of human-centred cybersecurity practices, but its effectiveness hinges on establishing trust. This trust must be built through technological reliability, fairness, and transparency, while also addressing issues such as data quality and interoperability. Without broad user confidence, AI deployment may be limited, constraining the real-world testing and iteration necessary for its refinement. Developing robust and trustworthy AI systems is vital to unlocking their full potential in securing digital identities in the cybersecurity landscape.

2 The Importance of Digital Identity Management

2.1 Situating Digital identity within the Cybersecurity Context

Maintaining the availability of computer systems to users is a core element of information and cybersecurity (Althonayan & Andronache, 2018). However, these systems are regularly attacked by malicious actors for crimes including scams, fraud, password cracking, and digital identity theft (Asselin & Bilodeau, 2023). A secure system must consequently have robust mechanisms allowing legitimate users or entities to access digital resources in a reasonable time frame and consistent manner while still protecting said digital resources from malicious attacks, illegitimate or unauthorized access, and illegal disclosure (Sule et al., 2021).

Digital identities are essential in the core functions of cybersecurity, including Identification, Authentication, Authorization, and Accountability (IAAA). These identities serve as the unique markers that allow systems to recognize and differentiate between users, forming the basis for **Identification**. Through an **Authentication** process, systems verify that the claimed identity matches stored credentials or attributes, ensuring that the entity accessing the system is legitimate (Sule et al., 2021). Once authenticated, a predetermined **Authorization** policy determines what actions or resources the individual can access, based on their digital identity (Pimenidis, 2010). Importantly, a process of **accountability** ensures that these actions are traceable to an identity, which is important for responding to unauthorized or malicious activity (Asselin & Bilodeau, 2023). Digital identities and their management play a critical role in managing secure access to digital ecosystems and protecting systems from unauthorized breaches or misuse, serving as the foundation for modern cybersecurity protocols.

2.2 Risks Related to Digital Identity Management

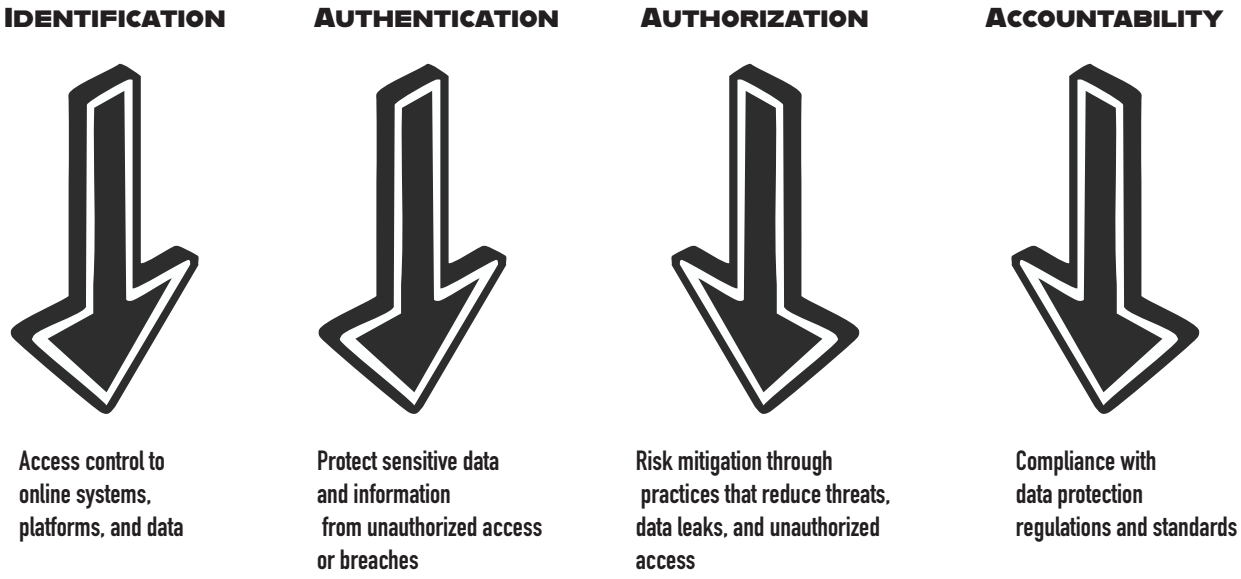
Digital identity management can also create vulnerabilities (Anderson et al., 2008). For example, gaps in authentication processes can be leveraged to gain access to sensitive information, which may lead to financial losses, reputational damage, and fraud. Additionally, privacy concerns can arise for users when their personal data is collected for verification purposes and is misused or exposed by an attacker (Cassim, 2015).

Common risks related to digital identity management include:

- System vulnerabilities, which may be exploited to gain access to sensitive information.
- Inaccurate identification, resulting in false positives and false negatives, may either deny legitimate users access or grant unauthorized users' entry.
- Single points of failure, arising from reliance on a single form of digital identity verification and potentially leading to widespread security breaches.
- User acceptance and compliance with digital identity management mechanisms also pose a challenge, as stringent verification measures may pose an inconvenience or be resisted (Cassim, 2015).

2.3 User Compliance & Digital Identity Management

There is often a mismatch between cybersecurity measures that may prevent risks to digital identity management and the ability of users to consistently meet the requirements of these measures. Users are often burdened with complex tasks that may prevent access to important data in a timely manner, impact productivity, or work against their cognition (Vishwanath et al., 2020). As a result, they may find workarounds that are less secure, or potentially refuse to adopt certain security measures altogether. For example, security requirements for credential management often include passwords with a complex variety of characters and a long length. Given that cybersecurity education often stresses not writing down or reusing passwords, the user is put in a position where they are expected to memorize it. This task is unrealistic and incompatible with human cognition (Sasse et al., 2001). Therefore, users create coping strategies that are less secure (Stobert & Biddle, 2018), creating risks for the individual user or the cyber ecosystems they belong to. With that in mind, it's crucial to design digital identity authentication systems that are efficient and secure, without burdening the system users.



The IAAA Framework

The IAAA framework, adapted from Stewart et al. (2012) and Danturthi (2024)

3 Poor Digital Identity Management: Impacts & Implications

3.1 Digital identity Theft

Poor digital identity management poses a serious risk to cybersecurity, most notably through digital identity theft. In 2022–2023, 91% of Canadians reported concerns about digital identity theft, with nearly half being extremely concerned, a trend that has been consistent since 2018 (Canada, 2023). This concern is not unfounded; in 2022, police-reported fraud rates (which include digital identity theft and digital identity fraud) were 78% higher than in 2012. Since 2011, fraud incidents have increased nearly every year (Canada, 2023). Additionally, crimes linked with digital identity fraud, such as extortion, have increased fivefold since 2012 (Canada, 2023). A significant portion of these crimes are conducted online, with 23% of fraud and 48% of extortion incidents reported as cybercrimes (Canada, 2023). Combined, these offences account for more than half of all cybercrimes reported in 2022 (Asselin & Bilodeau, 2023). More challenging is that digital identity theft and digital identity fraud perpetrated online may go unreported or undetected. Further compounding the potential harm from these thefts is the low availability of mechanisms through which an individual or entity can track whether their digital identity has been flagged or collected, or a database in which it is stored compromised, by cybercriminals. The rising trend of digital identity theft and fraud underscores the urgent need for robust digital identity management practices in cybersecurity. With the risks posed to individuals and organizations escalating and the lack of mechanisms to mitigate the threat, victims are in clear need of tools that reduce the impact of poor digital identity management on society.

3.2 Consequences of Digital Identity Theft

3.2.1 Impacts on Individuals

Digital identity theft has far-reaching impacts on individuals, affecting multiple aspects of their lives. The harms can include an impact on the financial, health, and legal well-being of affected individuals and can negatively influence their freedoms and employment.

Victims of digital identity theft often face significant financial losses, which extend beyond unauthorized transactions to include a variety of out-of-pocket costs. These can encompass fraudulent charges on credit cards, drained bank accounts, and loans or credit taken out in the victim's name (Irvin-Erickson, 2024). The process of resolving these issues is time-consuming and costly, often involving lengthy disputes with financial institutions and legal actions. Furthermore, victims may see long-term damage to their credit scores, complicating future financial opportunities (Sule et al., 2021). In severe cases, legal fees and identity recovery services further burden the victims, amplifying the financial toll (Cassim, 2015). These wide-ranging financial consequences underscore the profound and lasting impact that identity theft can have on individuals' financial stability.

Health consequences of digital identity theft can be severe, often manifesting as emotional distress and physical ailments. Over 80% of digital identity theft victims report emotional distress, and more than 21% experience physical health issues lasting a month or more (DeLiema et al., 2021). Victims of digital identity theft often experience emotional distress, including depression, anxiety, feelings of violation, anger, and a sense of powerlessness. These repercussions can lead to sleep disturbances, suicidal thoughts, headaches, high blood pressure, muscle tension, fatigue, and upset stomach (Golladay & Holtfreter, 2017). Medical fraud is another concern, where stolen identities are used to receive medical services, resulting in inaccurate medical records that can pose health risks besides the disputes over billing (Seh et al., 2020). The extensive health impacts underscore the need for robust measures to prevent digital identity theft and protect individuals from its harmful effects.

Legal issues may stem from both criminal and civil proceedings related to identity theft. Victims may face wrongful criminal charges and long legal battles if their stolen identities are used to commit crimes. Civil litigation can arise if digital identity thieves use stolen information to defraud businesses or individuals, necessitating lengthy legal processes to clear the victim's name and avoid liability (Cassim, 2015). Complications can also occur if stolen identities are used for immigration purposes, potentially resulting in wrongful deportation proceedings or difficulties in obtaining visas and other legal documents (Irvin-Erickson, 2024). Tax and social security fraud are also common outcomes of digital identity theft (Ngugi et al., 2021). This can lead to victims facing audits, penalties, and the challenge of correcting their tax records (Ngugi et al., 2021).

The employability or employment may be impacted not only by these legal issues but also by impacts on the victim's credit history, social security contributions, and employment records, particularly if they use stolen identities to gain employment for themselves (Anderson et al., 2008). Victims may also face audits, penalties, and tax issues related to tax and social security fraud (Ngugi et al., 2021). This can result in the potential restrictions on the freedoms of victims, as their residence and movements may be restricted while they correct their relationship with the various institutions and individuals who are also victimized in this form of crime.

3.2.2 Impacts on Businesses

Businesses impacted by digital identity theft often suffer significant financial losses. These losses may arise from fraudulent transactions resulting from the impersonation of employees or from compromised customer accounts, where companies may be held responsible for the cost to customers. In addition to these direct costs, businesses must also bear the burden of legal fees and the expenses associated with addressing the fraud, such as conducting investigations and enhancing security measures. Furthermore, regulatory fines may be imposed if the company is found to have inadequate protection, adding to the financial strain (Golladay & Holtfreter, 2017). The cumulative effect of these costs can severely affect a company's financial health.

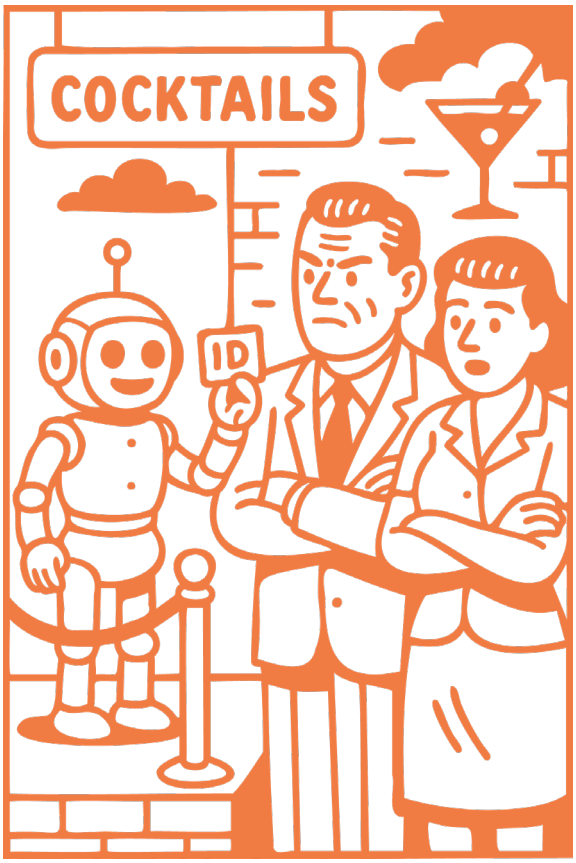
In addition to financial losses, businesses face significant reputational risks when digital identity theft occurs. Customers who feel that their personal information is not adequately protected may lose trust in the company, potentially leading to a loss of business. This reputational damage can be long-lasting and, in many cases, more harmful than the immediate financial impact, as trust is difficult to rebuild once it is lost (Golladay & Holtfreter, 2017). Dissatisfied customers may share their negative experiences, further harming the company's image and causing a ripple effect that erodes brand loyalty (Irvin-Erickson, 2024). While the total cost of a loss of reputation is difficult to calculate and may vary for each type of crime and differ across industries and business types, it is difficult to argue that reputation isn't an important aspect of the crimes resulting from identity theft.

Legal and regulatory challenges further compound the risks businesses face when digital identity theft occurs. Companies are required to comply with various data protection regulations, such as the Privacy Act, PIPEDA, and the GDPR in the European Union, to ensure that personal data is protected. Failure to meet these standards can result in hefty fines and legal action, further damaging the company's financial and operational standing (Sirur et al., 2018). Additionally, the legal repercussions of data breaches can involve lengthy lawsuits and settlements, which can divert resources away from business growth and innovation. As regulatory environments become stricter, businesses ultimately are served by prioritizing compliance to avoid these legal pitfalls and ensure their long-term viability.

3.2.3 Impacts on Government

Digital identity theft can have profound effects on election results, particularly through the manipulation of voter registrations. A notable example occurred during the 2016 Republican primary in Riverside County, California, where attackers altered voter registrations by changing addresses, requesting absentee ballots, or modifying party affiliations without the voter's knowledge (Sweeney et al., n.d.). Such actions can disenfranchise voters, resulting in them not casting their ballots and skewing election outcomes. This is particularly concerning in closely contested elections, where even a small number of altered votes can shift the result. It is arguable that disenfranchisement does not require the direct compromise of any specific election system, as the public perception of the possibility could result in the same problematic outcomes. The potential for voter disenfranchisement due to digital identity theft presents serious concerns, as it can potentially damage public perceptions about the integrity of the electoral process, leading to diminished voter participation and increasing public distrust in election outcomes (Biggers et al., 2023).

Digital identity theft also poses a significant financial threat to governments. Fraud rings often exploit stolen identities to file fraudulent tax returns and claim government benefits, resulting in substantial financial losses. One such case involved a fraud ring that stole over $6 million in government funds by using stolen identities (Internal Revenue Service, n.d.). These types of scams can have a far-reaching impact, draining public resources and undermining trust in government services. As financial losses accumulate, governments are



forced to allocate additional funds to investigate and combat these fraudulent activities, further stretching already limited budgets. Thus, the financial consequences of digital identity theft for governments are extensive and damaging.

In addition to financial losses, digital identity theft places an increased strain on government services. Government agencies, already tasked with protecting sensitive information, must implement additional security measures to combat identity theft (Internal Revenue Service, n.d.). This added workload diverts resources from other essential services, making it more difficult for agencies to operate efficiently. The need for continuous updates to security protocols and the implementation of new safeguards requires significant time, effort, and financial resources. As a result, digital identity theft not only compromises the integrity of governmental operations but also stretches the capacity of government services, ultimately impacting their ability to serve the public effectively.

4 Mitigating Digital Identity Theft

Risk can be understood as a combination of the probability of an event occurring and the harm that results from that event. Consequently, risk management involves either reducing the likelihood of the event or minimizing the harm it causes. While risk reduction activities generally need to occur before the event, risk mitigation efforts can be implemented both before and after the event to reduce its impact. Prevention mechanisms related to identity theft focus primarily on authentication systems, which help reduce the risk of unauthorized access. On the other hand, harm mitigation efforts, often referred to as resilience measures, are typically employed after identity theft has occurred.

In the context of digital identity theft, the harmful action is the misuse of a stolen identity. Reducing the possibility of this misuse after the theft has occurred is critical to minimizing long-term damage. Initially, incident response measures, such as locking compromised accounts and changing passwords, are necessary to halt immediate misuse. These should be combined with ongoing resilience measures that can extend long after the initial incident. These can include cred-

it monitoring, digital identity theft insurance, and legal protections that continue to shield victims from further financial or reputational harm (Chawki & Abdel Wahab, 2006).

Artificial intelligence (AI) offers significant potential in both preventing digital identity theft and supporting the initial stages of resilience. AI-driven systems can detect suspicious behaviour patterns, enhance fraud detection, and automate responses to breaches, providing an extra layer of protection. Additionally, AI can assist in real-time monitoring and swift mitigation efforts, making it a promising tool in both identity theft prevention and early resilience measures.

This section of the report will discuss the current commonly employed authentication systems before introducing how AI technologies could be introduced to reinforce digital identity theft mitigation strategies.

4.1 Common Authentication Systems

The earliest and most widely used method of authentication is the username and password system, which requires users to input a unique identifier (username) and a secret code (password). This method remains popular due to its simplicity, convenience and ease of implementation (Zviran & Haga, 1999). A key benefit of passwords is their memorability, as users can often choose passwords based on personal information. However, this ease of use comes at a cost: passwords are vulnerable to being guessed, stolen, or shared. Poor password management, such as using weak or easily guessed passwords, can lead to significant security breaches. Despite these vulnerabilities, passwords continue to be the first line of defence in many systems, though their security largely depends on user behaviour and system enforcement of strong password policies.

One advancement beyond traditional password systems is the use of biometric authentication, which identifies users based on physiological or behavioural traits, such as fingerprints, facial recognition, or voice patterns. Biometric systems offer a higher level of security than passwords because they rely on unique biological attributes that are difficult to replicate or steal (Idrus et al., 2013). The major advantage of biometrics is that they eliminate the need for users to remember complex passwords, offering both convenience and enhanced security. However, if biometric data is compromised, the consequences are severe, as unlike passwords, biometric traits cannot be easily changed. The cost and complexity of implementing biometric systems also pose challenges, particularly in ensuring accuracy and protecting biometric data.

Alternatively, some form of token-based authentication may be used. This method requires users to possess a physical or digital token to verify their identity. These tokens can come in various forms, such as hardware devices (e.g., security key fobs) or software-based tokens (e.g., mobile authentication apps). Unlike traditional username and password systems, token-based authentication ensures that only individuals in possession of the token can access secure systems, significantly reducing the risk of unauthorized access (Idrus et al., 2013). Furthermore, as they are able to be easily changed, the compromise of a token can be resolved relatively easily when compared

with biometric systems. A major advantage of token-based systems is that they generate dynamic, one-time credentials, making them resistant to common attacks such as phishing or replay attacks. However, token-based systems come with their own challenges. Managing physical tokens may impose logistical challenges, particularly for organizations with many users. The physical tokens can be lost, damaged, or stolen, leaving users temporarily locked out of systems or compromising their security. Software tokens, while more convenient, may still be vulnerable to malware or device theft. Despite these potential drawbacks, token-based authentication remains a powerful security tool, especially when combined with other authentication methods, such as passwords or biometrics, to create multi-factor authentication systems.

Another important mechanism is multi-factor authentication (MFA), which combines the approaches discussed in a layered approach, greatly increasing the complexity for attackers (Idrus et al., 2013). The system requires users to provide two or more verification factors before accessing a system. These factors typically include something the user knows (password), something the user has (security token), and/or something the user is (biometrics). However, while MFA significantly improves security, it is not infallible and it can introduce friction in the user experience, as users must interact with multiple systems. Additionally, the costs associated with implementing and maintaining MFA systems can be a barrier for smaller organizations.

An approach to measuring the performance of authentication systems so that their effectiveness may be compared involves the recording of False Rejection Rates (FRR) and False Acceptance Rates (FAR). A high FRR can prevent legitimate users from accessing critical systems, causing inconvenience and potentially hampering productivity. Conversely, a high FAR increases the risk of unauthorized access, as the system may incorrectly accept a fraudulent user. It is important to note that there is not an inherent performance difference between the system types outlined above, as their implementation greatly affects their performance. For example, a password system using four characters may show a better FRR than a fingerprint reader, and a simplistic token implementation may have a worse FAR than a complex password. Balancing these rates is a critical challenge in developing reliable authentication systems.

An additional concern for the implementation of these systems is that the implementation of the system should be more resistant to attack than the authentication method. System vulnerabilities remain a major concern for all types of authentication technologies. Weak security measures or outdated technology make systems particularly susceptible to exploitation by cybercriminals. Cyberattacks can take advantage of these weaknesses, allowing unauthorized access and data breaches, which calls for continuous updates to security protocols to mitigate risks.

Data security is another key concern, as authentication systems create a need for the storage and processing of additional sensitive information, that is, they store usernames and passwords, biometric-related data, etc. These systems require robust security measures to prevent unauthorized access, data breaches, and misuse. The leak of this information can

create serious issues for the security of the system, as it may put privileged accounts (e.g., system administrator), and privileged identities (e.g., CEO) at risk of being abused by malicious actors. The sensitivity of the data created by authentication systems makes strong encryption and secure data storage practices essential in any modern identification system.

5 The Potential for AI Deployment

AI-driven authentication, involving continuous verification and behavioural biometrics, presents excellent potential for identity theft detection (Al-Shehari et al., 2024). AI-driven systems, by continuously learning and adapting to user behaviours, provide a proactive and continuous defence against unauthorized access attempts. These systems increase security and can improve user experience by reducing the need for frequent authentication disruptions.

Continuous verification includes monitoring user behaviour in real time to validate actions. Unlike traditional systems, which rely on one-time authentication (e.g. “logging in”), continuous verification ensures that users are authenticated throughout their session. This method utilizes real-time monitoring of behavioural patterns, such as how users interact with their devices, including their typing speed, mouse movements, and touchscreen gestures. Behavioural biometrics allow systems to recognize users based on subtle, unique traits that are difficult for attackers to replicate (Folino et al., 2023). This real time, continuous monitoring not only enhances security but also offers a seamless experience, as legitimate users are not interrupted during their sessions.

The data gathering and processing power of modern networked computer systems when combined with machine learning algorithms allow for anomalies to be detected in real time (Martín et al., 2021). Machine learning allows AI systems to process vast amounts of behavioural data and identify deviations from normal patterns. By learning what constitutes typical behaviour for a particular user, these systems can flag unusual activity, such as accessing data from an unfamiliar location or at odd hours. As part of a multi-factor approach, this behaviour represents the passive representation of a factor for identification (i.e., it presents data on “something you are”), which strengthens security by reducing reliance on other authentication systems. These machine learning-driven systems make it harder for attackers, as they will have to mimic legitimate users while also achieving their intent, which is likely to involve actions unusual to most users, thereby creating a more secure and efficient authentication process (Verma et al., 2022). The continuous updating of the behavioural patterns of each user, as would be required in order to reduce false rejection rates, would result in an authentication process that evolves with user behaviour over time, staying ahead of potential threats such as those related to data breaches.

These forms of technology also offer promise in assisting the detection of deepfake video processing technologies for identity theft. Deepfake technology, which can be used to impersonate individuals in digital spaces, presents a growing threat to cybersecurity. Continuous verification makes it significantly harder for an attacker to maintain access using a deepfake without being detected due to the

mismatch in behavioural patterns over time (Hoque et al., 2021). For instance, an attacker using a deepfake might successfully bypass initial verification, but continuous monitoring would eventually reveal inconsistencies in typing cadence, device interaction, or other behavioural markers.

Verification requirement levels can also be adjusted dynamically, informed by real-time risk assessments (Hoque et al., 2021). AI systems can assess the risk level of any given user session by analyzing factors such as location, device, and behaviour. When suspicious activity is detected, the system can adjust the security protocols, such as requiring additional authentication steps, like multi-factor authentication or limiting access to certain features. This adaptability ensures that security measures remain proportional to the risk, balancing the need for strong security with a smooth user experience. Dynamic verification makes it more difficult for attackers to predict and bypass security measures, keeping the system secure without being overly burdensome to legitimate users.

By leveraging vast amounts of data and prioritizing it, AI has the capacity for nuance and proactive detection and reaction that would not otherwise be possible (Yang et al., 2024). AI systems are capable of processing far more data than traditional authentication methods, including data from multiple sources, such as location, behaviour, and device metrics. This enables the system to make nuanced decisions based on a comprehensive view of the user's activity, catching threats that may have gone unnoticed in more static systems. By proactively responding to potential risks and continuously learning from new data, AI-driven authentication systems offer a level of security that is both advanced and adaptable. These capabilities make AI indispensable in the future of digital identity protection and threat detection.

6 Challenges to AI Deployment

While AI offers significant potential for improving cybersecurity practices, especially in outlier identification, it also introduces barriers that impact user trust. Trust is a crucial factor in the adoption of AI technologies, as it shapes how users perceive the reliability, fairness, and security of these systems. For AI to be widely accepted, users must trust that the technology will function as expected and in alignment with their values. Mayer et al. (1995) define trust as the willingness to be vulnerable to the actions of another party based on the expectation that they will perform a particular action important to the trustor. This concept applies directly to AI, where users place their trust in the system's ability to make accurate and fair decisions without fully understanding how these decisions are made. Trust becomes essential, particularly because AI systems often operate with a level of complexity and opacity that makes it difficult for users to directly verify their performance. Many users express concerns about data privacy, algorithmic bias, the complexity of AI decision-making, and a lack of transparency in how AI systems operate. These factors contribute to a significant trust deficit, which must be addressed if AI is to be widely adopted in critical areas such as cybersecurity.

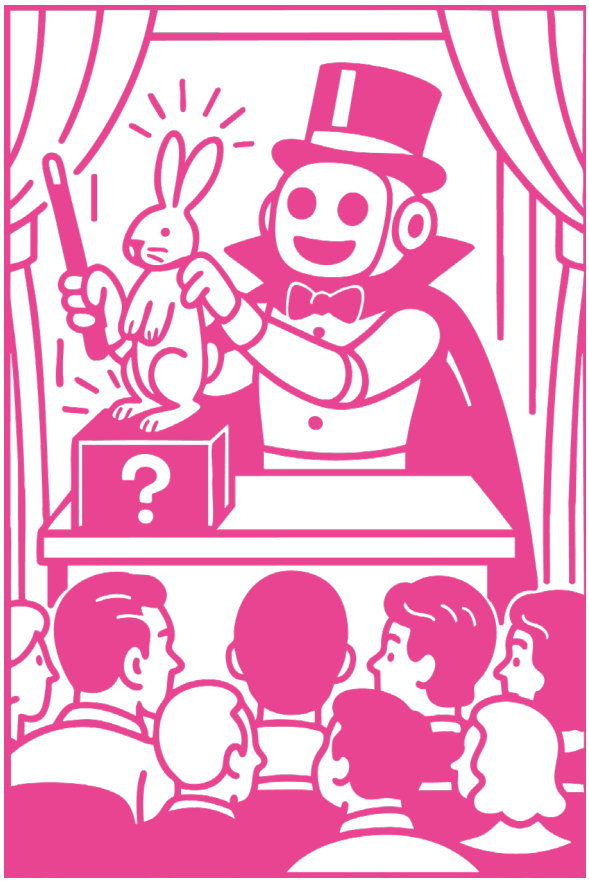
Moreover, Wu et al. (2011) highlight that trust significantly influences the acceptance of technology through the Technology Acceptance Model (TAM). Their meta-analysis shows that trust moderates users' perceptions of the usefulness and ease of use of new technologies, which in turn affects their willingness to adopt them. In the context of AI, if users believe that AI systems are trustworthy, they are more likely to perceive the technology as beneficial and less concerned about potential risks. This emphasizes the importance of trust-building strategies, such as improving transparency, ensuring fairness in decision-making, and protecting user data, to encourage the adoption of AI technologies across various industries.

This section outlines the various technical and non-technical barriers that hinder the widespread adoption of AI in cybersecurity. Technical barriers include the opaque nature of AI, data quality and interoperability issues, concerns about privacy and security, biases in AI models, and scalability limitations. On the other hand, non-technical barriers such as high costs, public perception, lack of regulatory frameworks, socio-economic disparities, and geopolitical tensions also present significant challenges to AI adoption. Addressing these barriers is critical to improving trust and ensuring the successful integration of AI in identity management and other critical applications.

6.1 Technical Barriers

6.1.1 Opaque Nature of AI

The inner workings of AI systems are often complex, which may reduce trust among users who are unable to understand how AI makes decisions and arrives at conclusions (Emaminejad & Akhavian, 2024). Trust in AI technologies is heavily influenced by users' perception of the system's reliability, fairness, and transparency. According to Radhakrishnan and Chattopadhyay (2020), one of the key barriers to AI adoption is the lack of transparency, as users struggle to understand how AI algorithms process data and arrive at conclusions. This opacity can create a trust deficit, especially in high-stakes applications like healthcare or finance, where errors or biases can have serious consequences. For users to adopt AI technologies, developers must prioritize creating systems that not only perform well but also provide clear explanations of their processes, thus fostering greater user trust.



6.1.2 Privacy & Security Concerns

As AI systems deployed for outlier identification must process sensitive information, raising concerns about data privacy and security. Users may be resistant or non-compliant to the use of AI as a result (Emaminejad & Akhavian, 2024). Privacy and security concerns can create significant barriers to trust in AI systems. Users may fear that their personally identifiable sensitive information, such as financial data or personal health records, could be exposed to unauthorized parties or misused. This fear can be exacerbated by past incidents of data breaches and scandals, which have heightened public awareness of the risks associated with data privacy. Furthermore, Radhakrishnan and Chattopadhyay (2020) emphasize that security breaches involving AI data processing could lead to long-term damage to trust, which is crucial for the system's adoption. The secure storage and transfer of sensitive data are key concerns that AI developers must address to maintain user trust and compliance with regulatory frameworks.

6.1.3 Issues of Bias & Fairness

AI systems may perpetuate biases from their training data, leading to discrimination, unfair treatment, and distrust (Alami et al., 2024). Ensuring fairness and limiting bias requires careful data selection, algorithm design, and continuous monitoring. Without standardization and regulation, these efforts may be applied unevenly. For example, biased AI algorithms in hiring processes may unfairly disadvantage certain demographic groups. This awareness can erode trust in AI systems and deter users from relying on them for critical decisions. As Wu et al. (2011) highlight, trust in AI is often undermined when users perceive a lack of fairness, especially in systems with high impact on life decisions like employment, financial services, and legal determinations. Addressing bias and ensuring fairness are essential for improving trust in AI technologies, as trust is a primary driver of AI adoption.

6.1.4 Scalability & Performance Limitations

As AI systems become increasingly complex and are adopted more frequently in large-scale settings, challenges such as failure to manage data volume and variety, or an inability to deliver timely and accurate outcomes could decrease trust in technology (Alami et al., 2024). Any delays, inaccuracies, or failures in AI performance can lead to frustration and a lack of confidence in the technology. Scalability issues often arise in environments where data is complex and voluminous, such as in real-time trading or large healthcare systems (Radhakrishnan & Chattopadhyay, 2020). If AI systems cannot perform consistently under these conditions, users may lose faith in the technology's reliability.

6.2 Non-Technical Barriers

6.2.1 High Costs

Adoption and maintenance of AI systems can be costly, particularly for small businesses or organizations in developing regions. The Canadian Federation of Independent Business (CFIB) reports that numerous small businesses in Canada find it challenging to use advanced technologies such as AI due to its high cost and their financial constraints (Can-

adian Federation of Independent Business, 2023). High upfront costs, alongside infrastructure, software, and talent acquisition expenses, present significant barriers to AI adoption (Radhakrishnan & Chattopadhyay, 2020). These financial burdens are particularly felt by small and medium-sized enterprises (SMEs) and businesses in developing regions, where the necessary technological infrastructure may not exist. Moreover, ongoing costs such as system maintenance, upgrades, and integrating AI into existing processes add to the challenge. Potential solutions, such as leveraging cloud-based AI services or open-source tools, can help lower the cost barriers and make AI more accessible.

6.2.2 Public Perception & Trust

The public's trust in AI affects its adoption. Negative perceptions, such as fears over job loss, can hinder AI implementation. Radhakrishnan and Chattopadhyay (2020) highlight that public concerns over job displacement, privacy, and transparency are significant barriers to the widespread use of AI technologies. Additionally, individuals may be uncomfortable with continuous data monitoring, as required for systems that use AI for verification. These concerns are compounded by the "black box" nature of many AI systems, where users cannot easily understand how decisions are made. This lack of transparency can lead to distrust. Public perception is also influenced by fears over algorithmic bias and fairness, which can negatively impact marginalized groups in critical areas such as hiring and criminal justice. Building trust through transparency, accountability, and communication is crucial for AI's success (Ryan, 2020). Explaining how AI systems work and giving users more control over their data could help reduce public skepticism.

6.2.3 Lack of Regulatory and Legal Frameworks

The absence of clear regulatory and legal frameworks in AI development can hinder adoption (Ryan, 2020). The lack of well-defined legal structures around liability, data protection, and compliance creates uncertainty for businesses, which may be reluctant to invest in AI without clear guidelines (Radhakrishnan & Chattopadhyay, 2020). In sectors such as healthcare and finance, where the stakes are high, the absence of robust AI regulations may slow adoption further. Specific regulatory challenges include managing cross-border data flows, ensuring algorithmic transparency, and defining accountability for AI-driven decisions. Creating and enforcing comprehensive legal frameworks will be essential for ensuring that AI systems are used ethically and safely, while also fostering innovation and adoption.

6.2.4 Socio-economic Disparities

Unequal access to AI technologies can worsen socio-economic disparities. It is crucial to develop and deploy AI inclusively to prevent social and economic gaps from widening (Ryan, 2020). Businesses and regions with limited digital infrastructure face additional challenges in adopting AI, as they may lack access to high-quality data, skilled personnel, and the necessary hardware (Radhakrishnan & Chattopadhyay, 2020). The digital divide between wealthier and poorer regions, as well as between large corporations and small businesses,

can exacerbate existing inequalities. Furthermore, education and workforce development are key factors in addressing these disparities. Without access to AI education and training, certain regions and groups may be excluded from AI-driven economic growth. Ensuring equitable access to AI technologies and training programs is essential for minimizing these gaps and ensuring that the benefits of AI are shared widely.

6.2.5 Geopolitical Tensions

The integration and utilization of AI technologies are increasingly intertwined with the geopolitical environment. The pursuit of dominance in AI can result in the prioritization of national agendas over international cooperation and the utilization of AI for military and surveillance objectives. Geopolitical competition, particularly between AI leaders such as the US and China, often stifles international collaboration and creates fragmented technological advances (Radhakrishnan & Chattopadhyay, 2020). For example, export restrictions on AI technologies or data sovereignty laws can prevent the free flow of AI innovations across borders. This geopolitical rivalry also raises ethical concerns over the use of AI in warfare and surveillance, further complicating international AI development. Facilitating global partnerships, establishing universal benchmarks, and ensuring the peaceful and constructive utilization of AI present significant challenges that necessitate continuous diplomatic efforts (Ryan, 2020). Without international cooperation, the risk of AI development exacerbating geopolitical conflicts increases.

7 Building Trust in AI

This report proposes four key mechanisms to address the factors contributing to the trust deficit: the development of trustworthy AI technologies, human oversight in model decision-making, regulation, and incentives. Although AI and authentication present specific challenges within cybersecurity, these issues—such as transparency, fairness, and security—are increasingly central to the field. These mechanisms are designed to be adaptable and flexible, allowing them to address trust concerns in AI and authentication and across a wide range of cybersecurity applications. By focusing on long-term effectiveness, these approaches aim to ensure that AI technologies can be trusted and widely adopted.

7.1 Trustworthy AI Technologies

The development of trust in IT technologies for security will be greatly aided by them being developed in such a manner that they are trustworthy. As such, AI authentication technologies and their underlying models should be developed in such a way that features that manifest their trustworthiness are prioritized. These features would include mechanisms for transparency and fairness; equitable standards; security and data protection; and ongoing monitoring and updates.

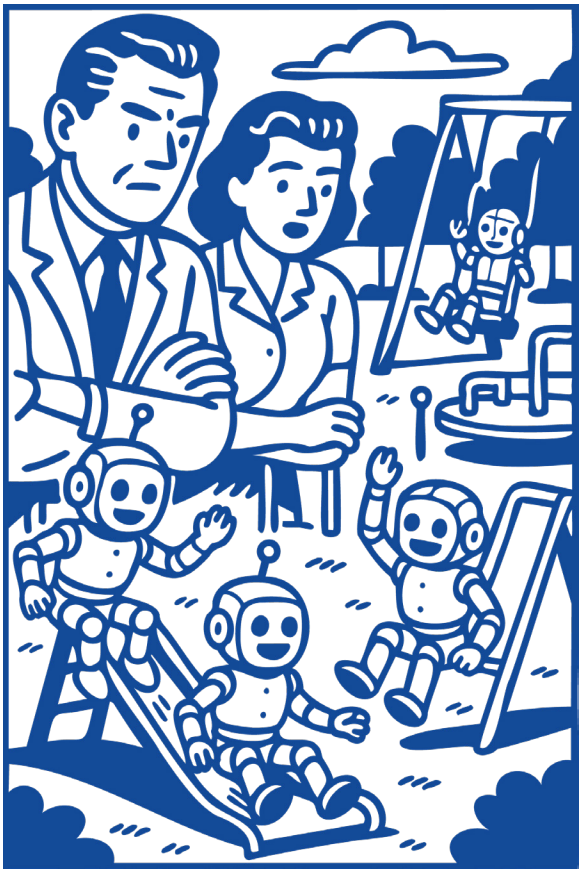
7.1.1 Transparency and Fairness

AI systems should be transparent, clearly explaining their decision-making processes so that users clearly understand how the data is used, what criteria are considered, and what

the process itself encompasses (Doshi-Velez & Kim, 2017). Transparent AI systems enable users to trust the technology by providing clear, understandable explanations of how decisions are made. While machine learning technologies are often characterized as being opaque, there are clear pathways towards this outcome, such as the incorporation of saliency maps, class activation maps, and gradient-based methods to help explain AI decision-making processes. By making the decision-making process visible, AI developers can ensure accountability, reduce the risk of misuse, and foster trust among users. Furthermore, this transparency should be leveraged with accountability frameworks directed to the end of fairness in AI systems, ensuring that decisions are equitable in that they treat all users equally regardless of their background, which is critical in minimizing bias and discrimination.

7.1.2 Equitable Standards in AI Development

AI models should be developed with equitable standards to minimize biases and ensure equal treatment for all users (IBM, 2021). Ensuring fairness in AI is critical because biased models can perpetuate existing inequalities, especially in systems used for hiring, lending, or criminal justice. Developers must carefully curate training datasets, avoiding inherent biases that could disproportionately affect certain demographic groups. Additionally, fairness algorithms and regular audits of AI systems can help detect and correct biased outcomes, ensuring that all users receive equitable treatment. By promoting fairness from the outset, AI developers can reduce the risks of harm and discrimination, while fostering trust in the system's reliability and ethical integrity.



7.1.3 Security and Data Protection

Both AI systems and the data they are protecting must have robust security measures that prioritize protection of sensitive data from breaches and attacks (Yang et al., 2024). As AI systems increasingly handle large volumes of sensitive information, particularly data used for authentication, they become prime targets for cyberattacks. To mitigate these risks, AI systems must incorporate advanced encryption methods, regular vulnerability assessments, and real-time threat detection mechanisms. By prioritizing security, AI systems can better encourage user trust by protecting the integrity of the information they process.

7.1.4 Ongoing Monitoring and Updates

AI models must be constantly monitored and updated to ensure the system remains effective and reliable, demonstrating high precision and ongoing dependability (Rosenberg et al., 2021). Regular updates allow AI models to integrate new information, correct errors, and refine their decision-making processes. This is particularly important where they are implemented to perform a security function. Continuous monitoring helps detect any emerging biases or performance degradation, ensuring that the system maintains high standards over time. As organizational and personal needs evolve, the models behind the security system must adapt. By investing in ongoing model maintenance, organizations can ensure that their AI systems stay capable of meeting the demands of dynamic real-world environments.

7.2 Human Oversight in Model Decision-Making

Ethical and responsible deployment of AI systems requires active human oversight and governance mechanisms. AI technologies developed for cybersecurity purposes should integrate human oversight into their processes, ensuring that their design and operation conform to expectations. Review boards and approval processes should validate AI systems before deployment (Mennella et al., 2024), while leadership should be interdisciplinary and prioritize the navigation of dilemmas and issues as they arise (Hagendorff, 2020). At a systems level, trained employees should engage in proactive and ongoing sampling and oversight activities to ensure that AI models remain effective and do not become vulnerable to exploitation (via data poisoning or constant return of false positives, for example). At the individual level, users should have a say in how their digital identity data is used and shared (Golladay & Holtfreter, 2017). Balancing this oversight with the use of AI to relieve the burden of cybersecurity practices related to digital identity will likely differ depending on context and environment; the priority should be to ensure that humans and AI work symbiotically. Giving humans a role in oversight builds trust by granting humans control over systems in critical ways and ensures that the inherent vulnerabilities of AI are not left unsupervised. By maintaining a human in the loop approach, there is a greater chance that new problems will be able to be identified before they result in catastrophic outcomes.

7.3 Regulation and Standards

While the introduction of regulation for the use of AI appears inevitable, with efforts such as the European Union regulatory framework for AI (European Parliament, 2023), the proposed Artificial Intelligence and Data Act (AIDA) representing the desire for clear rules (ISED, 2023). In order for this regulation to be effective in such a dynamic technological environment, it is important that it is developed by means of cooperative efforts, meaning the establishment of collaborative undertakings and collective understandings across public, private, and academic sectors.

Regulatory processes must incorporate a wide

variety of actors to ensure robust accountability mechanisms and clear guidelines for ethical conduct (Novelli et al., 2023). Regulatory oversight should be adequately funded and insulated from both public and private interests that may seek to influence legal or political frameworks. The regulation developed should promote a competitive environment in the industry to ensure diversification of development and opportunity, and avoid centralization of expertise.

To complement regulation, procedural and technological standardization must be adopted to facilitate compatibility between individual users, controllers, and technologies to ensure best practices are both known and adopted (Ryan, 2020). Prioritizing AI interoperability requires certification and educational opportunities to be made available to those with sufficient knowledge and experience. In prioritizing these features, standards can identify and rectify unique vulnerabilities to certain models, contribute to the development of expertise that is not limited by model-specific knowledge, and ensure accessibility of expertise. Further, ethical knowledge sharing will be a key component of both regulatory development and standardization and works as a mechanism of trust building in and of itself (Hagendorff, 2020). This involves the development of frameworks and forums for sharing information that does not contribute to or create new risks. This would be best served by a cross-sectoral effort to ensure comprehensive understanding of threats, actors, vulnerabilities, and best practices and should be pursued through both data/technology sharing and experiential or research-based sharing. Developing standard frameworks in each is essential to protecting data privacy and security, as well as preventing malicious actors from exploiting shared knowledge (Ryan, 2020)

By developing standards and regulations that ensure accountability, responsible use of AI, and effective management of potential threats or vulnerabilities, and by fostering expertise through ethical knowledge-sharing frameworks, trust in the field of AI and cybersecurity can be enhanced (Humphreys et al., 2024).

7.4 Incentives

Investments and funding into AI development, education, and training are necessary for stimulating expertise, increasing accessibility, and ensuring growth in available technologies and opportunities. Increasing government funding sources available to small and medium-sized businesses for the development, design, and adoption of AI would help promote decentralized innovations in addressing problems with authenticating identities. Additionally, enforcing standards in AI deployment will require investments in law enforcement to build the skills and knowledge required for the use of AI for cybercrime intelligence (Walter, 2024). Providing training and education on investigating cybercrimes involving misuse of digital identity will allow law enforcement to better protect individuals and ultimately lower the costs that result from cyberattacks against individuals and businesses (Walter, 2024). Investing in education and training programs is also important for equipping individuals with skills and knowledge that enhance their trust in how AI is designed, developed, and deployed.

8 Summary of Recommendations

The following recommendations present key considerations for advancing the responsible development and deployment of artificial intelligence.

- Promote Transparent and Fair AI Systems: AI systems should be transparent, clearly explaining their decision-making processes, the data used, criteria considered, and the process itself, to foster user trust and accountability.
- Prioritize Equitable Standards in AI Development: AI models should be developed with equitable standards to minimize biases and ensure equal treatment for all users, particularly in sensitive fields like hiring, lending, and criminal justice.
- Prioritize Security and Data Protection in AI Systems: AI systems and the data they process must incorporate robust security measures, such as encryption, vulnerability assessments, and real-time threat detection, to protect sensitive data and prevent breaches.
- Implement Ongoing Monitoring and Updates: AI models must be constantly monitored and updated to maintain effectiveness, detect emerging biases, correct errors, and ensure reliability in dynamic environments.
- Ensure Human Oversight in AI Systems: Ethical and responsible AI deployment requires active human oversight, review boards, and governance mechanisms to validate systems, prevent misuse, and ensure accountability at both system and individual levels.
- Establish Comprehensive AI Regulations and Standards: Regulation and standardization efforts should involve collaboration between public, private, and academic sectors, promoting ethical conduct, accountability, and competition while ensuring compatibility across AI technologies and practices.
- Foster Ethical Knowledge Sharing and Interoperability: Cross-sector efforts should focus on ethical knowledge sharing and interoperability, providing forums for information exchange while preventing exploitation of shared data and maintaining security and privacy.
- Provide Incentives for AI Development and Adoption: Government funding, particularly for small and medium-sized businesses, should be provided to stimulate decentralized investment in AI innovations.
- Invest in AI Education and Training: Governments and organizations should invest in AI education and training programs to equip individuals with the skills needed to develop, understand, and trust AI systems, fostering greater adoption and innovation. Further, education and training for law enforcement to improve AI-related skills, not only to protect against cybercrime but also to enforce AI-related regulations.

9 Looking Forward: Report Conclusions

Addressing digital identity-related cybersecurity issues protects individuals, organizations and systems. It mitigates the risks and harms associated with unauthorized access, fraud, data breaches, and digital identity theft. By integrating AI as a mechanism to better authenticate and protect digital identities, it may be possible to both increase security through continuous monitoring and reduce

the reliance on human users and human information security personnel, alleviating them of unreasonable burdens.

However, to use AI effectively to supplement identity management practices, trust-building mechanisms for AI are necessary. This can be accomplished through combinations of appropriately featured technologies, human oversight in AI decision-making, regulation, and incentives. What these combinations and the realizations of these mechanisms will look like in practice will depend on the environment and context. Recognizing this, the recommendations in this report are broad in scope and actionable in a variety of ways. Once trust in AI is built sufficiently, and AI given the legitimacy required to be a viable solution, the robustness of cybersecurity related to identification can be improved. This, in turn, will result in further trust, as the benefits derived from AI for outlier identification are demonstrated and experienced.

10 Bibliography

Adediji, D., Aldridge, M., Ouellet, M., Puopolo, A., Thompson, D., & Frank, R. (2022). Challenges of Virtual Trust: A Matter of Cooperation, Education, and Cybersecurity. Human-Centric Cybersecurity Partnership HC2P.

Adham, M., Azodi, A., Desmedt, Y., & Karaolis, I. (2013). How to Attack Two-Factor Authentication Internet Banking. In A.-R. Sadeghi (Ed.), Financial Cryptography and Data Security (pp. 322–328). Springer.

Alami, H., Lehoux, P., Papoutsi, C., Shaw, S. E., Fleet, R., & Fortin, J.-P. (2024). Understanding the integration of artificial intelligence in healthcare organisations and systems through the NASSS framework: A qualitative study in a leading Canadian academic centre. BMC Health Services Research, 24(1), 701. https://doi.org/10.1186/s12913-024-11112-x

Alawida, M., Omolara, A. E., Abiodun, O. I., & Al-Rajab, M. (2022). A deeper look into cybersecurity issues in the wake of Covid-19: A survey. Journal of King Saud University - Computer and Information Sciences, 34(10), 8176–8206. https://doi.org/10.1016/j.jksuci.2022.08.003

Al-Shehari, T. A., Rosaci, D., Al-Razgan, M., Alfakih, T., Kadrie, M., Afzal, H., & Nawaz, R. (2024). Enhancing Insider Threat Detection in Imbalanced Cybersecurity Settings Using the Density-Based Local Outlier Factor Algorithm. IEEE Access, 12, 34820–34834. https://doi.org/10.1109/ACCESS.2024.3373694

Althonayan, A., & Andronache, A. (2018). Shifting from information security towards a cybersecurity (pp. 68–79).

Anderson, K. B., Durbin, E., & Salinger, M. A. (2008). Digital identity Theft. Journal of Economic Perspectives, 22(2), 171–192. https://doi.org/10.1257/jep.22.2.171

Asselin, G., & Bilodeau, H. (2023a, July 11). The changing landscape of cyber security following the COVID-19 pandemic. https://www150.statcan.gc.ca/n1/pub/22-20-0001/222000012023001-eng.htm

Asselin, G., & Bilodeau, H. (2023b, July 11). The changing landscape of cyber security following the COVID-19 pandemic. https://www150.statcan.gc.ca/n1/pub/22-20-0001/222000012023001-eng.htm

Biggers, D. R., Elder, E. M., Hill, S. J., Kousser, T., Lenz, G. S., & Lockhart, M. (2023). Can Addressing Integrity Concerns about Mail Balloting Increase Turnout? Results from a Large-Scale Field Experiment in the 2020 Presidential Election. Journal of Experimental Political Science, 10(3), 413–425. https://doi.org/10.1017/XPS.2022.31

Canada, O. of the P. C. of. (2023a, June 14). Survey of Canadians on Privacy-Related Issues. https://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2023/por_ca_2022-23/#fig13)

Canada, O. of the P. C. of. (2023b, June 14). Survey of Canadians on Privacy-Related Issues. https://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2023/por_ca_2022-23/#fig13)

Cassim, F. (2015). Protecting personal information in the era of digital identity theft: Just how safe is our personal information from digital identity thieves? Potchefstroom Electronic Law Journal, 18(2), 68–110. https://doi.org/10.4314/pelj.v18i2.02

Chawki, M., & Abdel Wahab, M. S. (2006). Digital identity Theft in Cyberspace: Issues and Solutions Edition Speciale Immigration et Securite. Lex Electronica, 11(1), 1–41.

Chen, S., Hao, M., Ding, F., Jiang, D., Dong, J., Zhang, S., Guo, Q., & Gao, C. (2023). Exploring the global geography of cybercrime and its driving forces. Humanities & Social Sciences Communications, 10(1), 71. https://doi.org/10.1057/s41599-023-01560-x

Conklin, A., Dietrich, G., & Walz, D. (2004). Password-based authentication: A system perspective. 37th Annual Hawaii Inter-

national Conference on System Sciences, 2004. Proceedings of The, 10-. https://doi.org/10.1109/HICSS.2004.1265412

Cremer, F., Sheehan, B., Fortmann, M., Kia, A. N., Mullins, M., Murphy, F., & Materne, S. (2022). Cyber risk and cybersecurity: A systematic review of data availability. The Geneva Papers on Risk and Insurance. Issues and Practice, 47(3), 698–736. https://doi.org/10.1057/s41288-022-00266-6

Danturthi, R. S. (2024). Database and Application Security: A Practitioner's Guide (1st ed.). Addison-Wesley Professional. ISBN-13: 978-0-13-807373-2. Published March 12, 2024

DeLiema, M., Burnes, D., & Langton, L. (2021). The Financial and Psychological Impact of Digital identity Theft Among Older Adults. Innovation in Aging, 5(4), 043. https://doi.org/10.1093/geroni/igab043

Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. https://www.semanticscholar.org/paper/Towards-A-Rigorous-Science-of-Interpretable-Machine-Doshi-Velez-Kim/5c39e37022661f81f79e481240ed9b175dec6513

Dutson, J., Allen, D., Eggett, D., & Seamons, K. (2019). Don't Punish all of us: Measuring User Attitudes about Two-Factor Authentication. 2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW, 119–128. https://doi.org/10.1109/EuroSPW.2019.00020

Emaminejad, N., & Akhavian, R. (2024). Trust in Construction AI-Powered Collaborative Robots: A Qualitative Empirical Analysis (pp. 513–521). https://doi.org/10.1061/9780784485224.062

Folino, G., Otranto Godano, C., & Pisani, F. S. (2023). An ensemble-based framework for user behaviour anomaly detection and classification for cybersecurity. The Journal of Supercomputing, 79(11), 11660–11683. https://doi.org/10.1007/s11227-023-05049-x

Golladay, K., & Holtfreter, K. (2017). The Consequences of Digital identity Theft Victimization: An Examination of Emotional and Physical Health Outcomes. Victims & Offenders, 12(5), 741–760. https://doi.org/10.1080/15564886.2016.1177766

Gupta, S., & Furnell, S. (2022). From Cybersecurity Hygiene to Cyber Well-Being (pp. 124–134). https://doi.org/10.1007/978-3-031-05563-8_9

Hagendorff, T. (2020). The Ethics of AI Ethics: An

Evaluation of Guidelines. Minds and Machines, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hoque, M. A., Ferdous, M. S., Khan, M., & Tarkoma, S. (2021). Real, Forged or Deep Fake? Enabling the Ground Truth on the Internet. IEEE Access, 9, 160471–160484. https://doi.org/10.1109/ACCESS.2021.3131517

Humphreys, D., Koay, A., Desmond, D., & Mealy, E. (2024). AI hype as a cyber security risk: The moral responsibility of implementing generative AI in business. AI and Ethics. https://doi.org/10.1007/s43681-024-00443-4

I.B.M. (2021, July 19). What Is Threat Hunting? https://www.ibm.com/topics/threat-hunting

Idrus, S. Z. S., Cherrier, E., Rosenberger, C., & Schwartzmann, J.-J. (2013). A review on authentication methods. Australian Journal of Basic and Applied Sciences, 7(5), 95–107.

Independent Business, C. F. (2023). The challenges of competing in a digital world: The experiences of Canadian small businesses with Amazon. Canadian Federation of Independent Business. https://www.cfib-fcei.ca/hubfs/research/reports/2023/2023-09-challenges-competing-digital-world-amazon-en.pdf

Innovation, Science and Economic Development Canada [ISED]. (2023, September 27). Artificial Intelligence and Data Act. Innovation, Science and Economic Development Canada. https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act

Internal Revenue Service [IRS]. (2024a, January 12). Three individuals sentenced for roles in fraud and digital identity theft ring that stole over $6 million in government funds. https://www.irs.gov/compliance/criminal-investigation/three-individuals-sentenced-for-roles-in-fraud-and-identity-theft-ring-that-stole-over-6-million-in-government-funds

Internal Revenue Service [IRS]. (2024b, May 8). To protect against digital identity theft, IRS adds additional protections to Centralized Authorization File, Transcript Delivery System; changes designed to protect sensitive tax pro, taxpayer information. https://www.irs.gov/newsroom/to-protect-against-identity-theft-irs-adds-additional-protections-to-centralized-authoriz-

ation-file-transcript-delivery-system-changes-designed-to-protect-sensitive-tax-pro-taxpayer-information

Irvin-Erickson, Y. (2024). Digital identity fraud victimization: A critical review of the literature of the past two decades. Crime Science, 13(1), 3. https://doi.org/10.1186/s40163-024-00202-0

Martín, A. G., Beltrán, M., Fernández-Isabel, A., & Diego, I. (2021). An approach to detect user behaviour anomalies within digital identity federations. Computers & Security, 108, 102356. https://doi.org/10.1016/j.cose.2021.102356

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. The Academy of Management Review, 20(3), 709–734. https://doi.org/10.2307/258792

Mennella, C., Maniscalco, U., Pietro, G., & Esposito, M. (2024). Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. Heliyon, 10(4), 26297. https://doi.org/10.1016/j.heliyon.2024.e26297

Ngugi, B. K., Hung, K.-T., & Li, Y. J. (2021). Reducing tax digital identity theft by identifying vulnerability points in the electronic tax filing process. Information & Computer Security, 30(2), 173–189. https://doi.org/10.1108/ICS-05-2021-0056

Novelli, C., Taddeo, M., & Floridi, L. (2023). Accountability in artificial intelligence: What it is and how it works. AI & SOCIETY. https://doi.org/10.1007/s00146-023-01635-y

Parliament, E. (2023, June 8). EU AI Act: First regulation on artificial intelligence. Topics | European Parliament. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Pimenidis, E. (2010). Digital identity management. In Handbook of Electronic Security and Digital Forensics (pp. 279–294).

Radhakrishnan, J., & Chattopadhyay, M. (2020). Determinants and barriers of artificial intelligence adoption–A literature review (pp. 89–99).

Rosenberg, I., Shabtai, A., Elovici, Y., & Rokach, L. (2021). Adversarial Machine Learning Attacks and Defense Methods in the Cyber Security Domain. http://arxiv.org/abs/2007.02407

Ryan, M. (2020). In AI We Trust: Ethics, Artificial Intelligence, and Reliability. Science and

Engineering Ethics, 26(5), 2749–2767. https://doi.org/10.1007/s11948-020-00228-y

Sasse, M. A., Brostoff, S., & Weirich, D. (2001). Transforming the 'Weakest Link'—A Human/Computer Interaction Approach to Usable and Effective Security. BT Technology Journal, 19(3), 122–131. https://doi.org/10.1023/A:1011902718709

Seh, A. H., Zarour, M., Alenezi, M., Sarkar, A. K., Agrawal, A., Kumar, R., & Ahmad Khan, R. (2020). Healthcare Data Breaches: Insights and Implications. Healthcare, 8(2), 133. https://doi.org/10.3390/healthcare8020133

Sirur, S., Nurse, J. R. C., & Webb, H. (2018). Are We There Yet? Understanding the Challenges Faced in Complying with the General Data Protection Regulation (GDPR. Proceedings of the 2nd International Workshop on Multimedia Privacy and Security, 88–95. https://doi.org/10.1145/3267357.3267368

Stewart, J. M., Chapple, M., Gibson, D., & Chapple (2012). Cissp : Certified information systems security professional study guide : certified information systems security professional study guide. John Wiley & Sons, Incorporated.

Stobert, E., & Biddle, R. (2018). The Password Life Cycle. ACM Transactions on Privacy and Security, 21(3). https://doi.org/10.1145/3183341

Stuart, A., Bandara, A. K., & Levine, M. (2019). The psychology of privacy in the digital age. Social and Personality Psychology Compass, 13(11), 12507. https://doi.org/10.1111/spc3.12507

Sule, M.-J., Zennaro, M., & Thomas, G. (2021). Cybersecurity through the lens of Digital Digital identity and Data Protection: Issues and Trends. Technology in Society, 67, 101734. https://doi.org/10.1016/j.techsoc.2021.101734

Sweeney, L., Yoo, J. S., & Zang, J. (XXXX). Voter Digital identity Theft: Submitting Changes to Voter Registrations Online to Disrupt Elections. Technology Science. https://techscience.org/a/2017090601/

Verma, A., Moghaddam, V., & Anwar, A. (2022). Data-Driven Behavioural Biometrics for Continuous and Adaptive User Verification Using Smartphone and Smartwatch. Sustainability, 14(12), 12. https://doi.org/10.3390/su14127362

Vishwanath, A., Neo, L. S., Goh, P., Lee, S., Khader, M., Ong, G., & Chin, J. (2020). Cyber hygiene: The concept, its measure, and its

initial tests. Decision Support Systems, 128, 113160. https://doi.org/10.1016/j.dss.2019.113160

Walter, Y. (2024). Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences. Discover Artificial Intelligence, 4(1), 14. https://doi.org/10.1007/s44163-024-00109-4

Wu, K., Zhao, Y., Zhu, Q., Tan, X., & Zheng, H. (2011). A meta-analysis of the impact of trust on technology acceptance model: Investigation of moderating influence of subject and context type. International Journal of Information Management, 31(6), 572–581.

Yang, L., Tian, M., Xin, D., Cheng, Q., & Zheng, J. (2024). AI-Driven Anonymization: Protecting Personal Data Privacy While Leveraging Machine Learning. https://doi.org/10.48550/arXiv.2402.17191

Zviran, M., & Haga, W. J. (1999). Password security: An empirical study. Journal of Management Information Systems, 15(4), 161–185.

Human-Centric
Cybersecurity
Partnership