



The Alan Turing Institute

**Cyber Threat Observatory
for National Identity Systems**

**Quarterly Report: Generative
AI and the Rise of Credential
Fraud in Digital Public
Infrastructure**

November 2025

Authors

- Dr Salim Awudu, Research Associate
- Dr Mark Hooper, Technical Development Manager
- Professor Carsten Maple, Professor, and Director of the NCSC-EPSC, Academic Centre of Excellence in Cyber Security Research at Institute partner, University of Warwick (WMG)
- Professor Jon Crowcroft, Researcher-at-Large, Marconi Professor of Communication Systems at Institute partner University of Cambridge

Cyber Threat Observatory for national identity systems

The programme's flagship observatory focuses on offering timely insight and analysis of cyber threats that have potential to negatively impact identity systems globally. For more information visit: <https://www.turing.ac.uk/TDI/Cyobs>

Acknowledgements

This work is supported, in whole or in part, by the Gates Foundation [INV-057591]. Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic Licence has already been assigned to the Author Accepted Manuscript.

The Alan Turing Institute
British Library
96 Euston Road
London
NW1 2DB

Executive Summary

This report provides a comprehensive and strategic analysis of the escalating threats posed by synthetic identities and deepfake-enabled credential fraud within Digital Public Infrastructure (DPI). As DPI systems, particularly digital identity platforms, become foundational to public service delivery, their risk exposure also increases.

Drawing on data from the National Vulnerability Database (NVD), the report highlights a 300% increase in identity-related Common Vulnerability Exposures (CVEs) and Common Weakness Enumeration (CWE) between 2020 and 2025, with frequent exploitation of weaknesses such as CWE-287 (Improper Authentication), CWE-863 (Incorrect Authorisation), and CWE-200 (Information Exposure).

Key findings as highlighted in the report include:

- Experience from advanced digital economies shows that synthetic identity activity can surge dramatically, with over 500% increase in just three years for the UK. This underscores the need for early safeguards as Global South countries expand their Digital Public Infrastructure.
- Documented deepfake-enabled breaches have occurred across sectors, including banking, healthcare, and national ID programmes. They demonstrate that AI-driven credential fraud is now a transnational and cross-sectoral threat, not confined to high-income countries.
- Synthetic identities are increasingly being weaponised for cross-border financial crime, social-benefit fraud, and coordinated disinformation campaigns, posing complex governance and cybersecurity challenges for countries expanding their Digital Public Infrastructure.

The report integrates insights from the Universal DPI Safeguards Framework, the Cambridge Centre for Alternative Finance (CCAF) NIST SP 800-63 (USA), eIDAS 2.0 (EU), African Union Digital ID Framework, UNDP DPI Safeguards, OECD Digital Identity Governance, and the UK Government Cyber Security Strategy to contextualise these threats within broader governance, resilience and regulatory challenges. It emphasises that synthetic identity fraud is not merely a technical anomaly but a systemic risk to digital trust, democratic integrity, and inclusive development. To address these risks, the report recommends a multi-layered mitigation strategy grounded in:

- Zero Trust Architecture and biometric anti-spoofing.
- DPI safeguards principles such as “Do no harm,” “Do not exclude,” and “Ensure effective redress.”
- Cross-sectoral governance, including regulatory harmonisation and public-private partnerships.
- Lifecycle-based risk mitigation, aligning with DPI development stages from conception to operations.
- Adoption of the Cyber Assessment Framework (CAF) to ensure consistent and proportionate cyber security assurance across identity systems.
- Establishment of coordinated threat intelligence sharing platforms, inspired by the UK Government Cyber Coordination Centre (GCCC), to enhance detection and response capabilities.

The report calls for proactive governance, continuous monitoring, and international collaboration to secure identity systems and uphold the integrity of DPI. As DPI scales globally especially in low and middle-income countries (LMICs), embedding safeguards and resilience into identity infrastructure will also be essential to counter the evolving threat landscape shaped by generative AI.

Table of Contents

1. Background & Motivation	6
2. Introduction	6
2.1 Research Objectives and Questions	7
2.2 Research Structure	7
3. Methodology	8
3.1 Data Sources	8
3.1 CVE/CWE Filtering Criteria.....	9
3.2 Sectoral Mapping & Attack Vector Classification.....	9
3.3 Case Study Selection	9
4. The New Threat Landscape	10
4.1 Deepfake Evolution and Detection Challenges.....	10
4.2 Face Morphing as a Biometric Spoofing Technique	11
4.3 GANs as Dual-Use Tools.....	11
4.4 Global Scale and Urgency of the Threat	12
4.5 Exploitation of Systemic Vulnerabilities	12
4.6 Sectoral Impact and Attack Vectors	13
4.6.1 Finance Sector: Deepfake-Enabled Money Laundering in Fintech.....	13
4.6.2 Healthcare Sector: Biometric Spoofing in Patient Portals	14
4.6.3 Government Sector: Synthetic Identity Fraud in National ID Systems	14
4.6.4 Cross-Sectoral Observations	15
4.7 Emerging Trends	15
5. CVE/CWE Analysis	16
5.1 Overview of CVE Trends (2020–2025)	16
5.2 Dominant CWEs in Identity Systems.	21
5.3 Attack Vector Classification.....	22
5.4 DPI Governance Stakeholder Map.....	23
5.5 DPI Maturity vs. Financial Inclusion Outcomes	24
5.6 Emerging Technologies and Regulatory Challenges	25
5.6.1 Designing Inclusive AI-Enabled Public Service Delivery.....	26
5.6.2 Safeguarding Privacy in AI-Powered Accessibility Tools.....	27
5.7 Implications for DPI Security.....	28
6. International Standards and Policy Alignment	28
6.1 Foundational Standards for Identity Security	28

6.2 DPI Safeguards and Ethical Governance	28
6.3 Regulatory Convergence and Global Cooperation	29
7. Deep Dive: The Digital ID Safety Pack (Minimum Security Baseline)	29
7.1 Technical Controls	29
7.2 Governance and Policy Measures	30
7.3 Cross-Sector Collaboration	31
7.4 Future-Proofing Identity Systems.....	31
8. Implementation Pathways for Countries (Quick, Medium, and Long-Term).....	31
8.1 QShort-Term (0–6 Months)	32
8.2 Medium-Term (6–18 Months)	32
8.3 Long-Term (18+ Months)	32
9. Conclusion: Protecting Trust in Digital Public Infrastructure	33
References	34

1. Background & Motivation

The emergence and impacts of digital identity systems on the economic, social and political sectors of a country's Digital Public Infrastructure (DPI) cannot be underestimated. This has led to their increased adoption particularly in low and middle-income countries (LMICs). Coupled with the increased adoption is the requirement for the maintenance of the systems' integrity, security and overall trustworthiness due to the rise in credential fraud enabled by generative AI technologies using synthetic identities and deep fake media. Synthetic identities and deepfake media exploit systemic weaknesses in federated identity architectures, API-based authentication flows, and biometric data storage.

In this report, generative AI refers to Generative Adversarial Networks (GANs) used to create synthetic media such as deepfake videos, face morphs, and other synthetic biometric data. While this report focuses primarily on facial biometrics, it also considers other modalities such as voice and fingerprint spoofing where relevant.

This report is motivated by the urgent need to understand and address emerging trends related to synthetic identities and deepfake media in DPI. The report provides a strategic, evidence-based analysis of the emerging threats by drawing on CVE and CWE data from the National Vulnerability Database (2020–2025), sectoral case studies, and international frameworks such as the Universal DPI Safeguards, NIST SP 800-63, and the UK Cyber Assessment Framework (CAF). The goal is to inform policymakers, technologists, and development partners on how to secure identity systems and uphold digital trust in an era of AI-driven threats.

2. Introduction

Digital identity systems have modernised Digital Public Infrastructure (DPI), and are facilitating access to essential services such as healthcare, financial inclusion, voting, and government benefits. As DPI use increases globally especially in low and middle-income countries, the security, integrity and trustworthiness of identity systems have become central to national resilience, economic stability, and democratic governance.

Despite these positive transformations, these systems have facilitated the introduction of some complex cybersecurity and governance challenges. Among them are the rise of synthetic identity fraud and deepfake-enabled credential spoofing. These threats are increasingly weaponised by adversaries leveraging generative AI technologies, such as Generative Adversarial Networks (GANs), to mimic biometric traits and fabricate digital personas. These synthetic identities, are capable of bypassing Know Your Customer (KYC) protocols, biometric authentication, and remote onboarding systems, undermining the security of DPI. These threats, combined with existing vulnerabilities in federated identity systems, API-based authentication, and biometric data storage, have created a perfect storm for identity systems. According to the U.S. Federal Reserve, synthetic identity fraud is the fastest-growing type of financial crime, accounting for up to 20% of credit losses in some institutions (Federal Reserve Bank of Boston, 2019). Additionally, LexisNexis Risk Solutions reports a 527% increase in high-risk synthetic identities in the UK between 2020 and 2023 (LexisNexis Risk Solutions, 2023). Generative AI enabled crimes are not only common in the developed economies, the threat is especially severe for low and middle-income countries (LMICs), where DPI systems have been recently adopted to facilitate development goals. Challenges that are prevalent in the LMICs such as weak cybersecurity controls, limited regulatory oversight, and infrastructural gaps make such identity systems more vulnerable to exploitation by these cyber criminals. Such exploits or breaches have severe impacts like exclusion from vital services, financial fraud, and erosion of public trust. Further, cybersecurity researchers have also demonstrated that deepfake videos can bypass facial recognition systems with

success rates exceeding 80% in the absence of robust liveness detection (Vyshegorodtsev et al., 2024).

These risks are also compounded by the increasing commoditisation of offensive cyber tools and services, which lower the barrier for malicious actors to exploit identity systems. The UK Government Cyber Security Strategy highlights that approximately 40% of cyber incidents managed by the National Cyber Security Centre (NCSC) between 2020 and 2021 targeted the public sector (UK Cabinet Office, 2022), underscoring the urgency of strengthening cyber resilience across DPI.

This report therefore, provides a comprehensive analysis of the emerging generative AI-enabled threats and credential fraud within the context of DPI. It also provides a strategic analysis of the evolving threat landscape, presents evidence of scale and impact, and outlines practical strategies for securing DPI globally. The study draws upon multiple data sources to examine emerging vulnerabilities and mitigation strategies in digital identity systems. It analyses CVE and CWE records from the National Vulnerability Database (2020–2025) to identify systemic weaknesses, including those linked to deepfake-enabled attacks. Real-world case studies from the finance, healthcare, and government sectors illustrate how synthetic identities and biometric spoofing are being exploited in practice. The analysis is situated within the broader landscape of cybersecurity and digital trust, drawing on academic and industry literature (ENISA, NIST, CyberArk, LexisNexis) and aligning with leading strategic frameworks such as the Universal DPI Safeguards Framework, Zero Trust Architecture, NIST SP 800-63, OWASP guidelines, and the UK Cyber Assessment Framework.

2.1 Research Objectives and Questions

This study seeks to inform policymakers, cybersecurity professionals, and Digital Public Infrastructure (DPI) architects of the urgent need to adapt security and governance models to a rapidly evolving threat environment shaped by generative AI. The overarching goal is to strengthen digital trust and resilience by securing identity systems against emerging AI-driven threats.

Specifically, the research will:

- Investigate how generative AI technologies exploit vulnerabilities in digital identity systems and verification processes.
- Assess the impact of synthetic and credential-based identity fraud across critical sectors, including finance, healthcare, and government services.
- Evaluate the effectiveness of policy instruments, cybersecurity standards, and architectural principles in mitigating these threats and supporting trustworthy DPI.

To achieve these objectives, the study addresses three guiding questions:

1. How are generative AI technologies exploiting vulnerabilities in digital identity systems across DPI ecosystems?
2. What are the impacts of credential and synthetic identity fraud in finance, healthcare, and government services?
3. What strategies are needed to mitigate these threats and ensure resilience in AI-enabled DPI systems?

2.2 Research Structure

The structure of the report is as follows:

- Section 3 provides the methodology by detailing how this work analysed the existing generative AI technologies used to compromise digital identity systems.
- Section 4 explores the evolving threat landscape, detailing how generative AI technologies such as deepfakes and synthetic identities are used to compromise digital identity systems. It also demonstrates the scale and impact of sector-specific case studies across finance, healthcare, and government sectors, highlighting systemic vulnerabilities and cross-sectoral risks.
- Section 5 details the Common Vulnerabilities and Exposures (CVEs) and their associated Common Weakness Enumerations (CWEs) trends from 2020 to 2025, DPI stakeholders' challenges and their impact on the security of Digital Public Infrastructure (DPI)
- Section 6 outlines international standards and policy frameworks such as NIST, ISO, GDPR, eIDAS, and the UNDP DPI Safeguards to provide guidance for securing identity systems.
- Section 7 provides the Digital ID Safety Pack, a minimum-security baseline comprising technical controls and governance measures to protect credentials and biometric data.
- Section 8 provides implementation pathways for countries at different stages of DPI maturity, offering quick, medium, and long-term strategies tailored to national contexts.
- Section 9 concludes with a call to action for proactive governance, international and cross- sectoral collaboration, and secure-by-design identity systems to uphold trust in Digital Public Infrastructure.

3. Methodology

This section details the methodology used for the study. The study employs a multi-layered approach to assess the systemic vulnerabilities exploited in generative AI-enabled credential fraud across Digital Public Infrastructure (DPI). The approach integrates CVE/CWE analysis, sectoral case studies, and threat classification to provide a comprehensive understanding of the evolving risk landscape.

The study adopted the following key stages:

- CWE Frequency Analysis: Identified dominant weakness types across sectors.
- Cross-Sectoral Comparison: Tracked propagation of vulnerabilities between domains.
- Strategic Mapping: Linked CVEs to mitigation strategies using Zero Trust Architecture, DPI Safeguards, and CAF profiles.

3.1 Data Sources

To achieve the research objective and the research questions, the research work focused on the following data sources:

- National Vulnerability Database (NVD): CVE and CWE data from 2020–2025, focusing on identity-related vulnerabilities.
- National Vulnerability Database (NVD): CVE and CWE data from 2020–2025, focusing on deepfake-related vulnerabilities.
- Sectoral Reports: Case studies from finance, healthcare, and government sectors, including data from TransUnion, LexisNexis, and UK Finance assessed between July 2025 and September 2025.

- Academic & Industry Literature: Insights from ENISA, NIST, CyberArk, and the Cambridge Centre for Alternative Finance assessed between July 2025 and September 2025.
- Policy Frameworks: DPI Safeguards Framework, NIST SP 800-63, OWASP API Security Top 10, and the UK Cyber Assessment Framework (CAF) assessed between July 2025 and September 2025.

The subsequent sub sections provide a more detailed description of the data sources used in this research.

3.1 CVE/CWE Filtering Criteria

The research adopted the inclusion and exclusion criteria to filter the CVE/CWE data as detailed below:

- Inclusion:
 - CVEs referencing identity verification, biometric modules, federated login APIs, and authentication tokens.
 - CVEs with mapped CWEs relevant to deepfakes, credential fraud and biometric spoofing such as content/script injection, UI spoofing, forced actions, and internal fetches between 2020 and 2025.
- Exclusion:
 - CVEs unrelated to identity systems (e.g., IoT, industrial control systems).
 - CVEs lacking descriptive metadata or CWE classification.

Regarding data cleaning and normalisation, the data were deduplicated to ensure unique CVE IDs. This means that, CVEs were included if they referenced identity verification, biometric modules, federated login APIs, or authentication tokens. CVEs without descriptive metadata or lacking CWE classification were excluded. Further, where multiple CWEs existed, the study counted the record once per CWE.

3.2 Sectoral Mapping & Attack Vector Classification

The research identified and mapped the attacks to the relevant sectors using the following processes:

- CVEs were mapped to sectors based on affected systems and vendors.
- Attack vectors were classified using CVSS metrics:
 - Network: Remote exploitation via APIs and login endpoints.
 - Local: Device-level access or compromised credentials.
 - Adjacent Network: Shared infrastructure (e.g., federated kiosks).
 - Physical: Biometric spoofing and hardware tampering.

3.3 Case Study Selection

Real-world incidents were selected based on:

- Public disclosures and regulatory investigations.
- Documented use of AI-driven identity fraud or deepfake media.
- Exploited CVEs/CWEs and systemic impact on DPI services.

4. The New Threat Landscape

The threat landscape surrounding digital identity systems has been revolutionised by generative AI. Credential fraud for instance, which was once limited to stolen passwords or forged documents, now includes sophisticated attacks using synthetic identities, deepfake media, and biometric spoofing. These threats exploit systemic vulnerabilities in identity architectures, biometric verification systems, and onboarding workflows to pose significant risks to the integrity and trustworthiness of Digital Public Infrastructure (DPI). This is illustrated in Figure 1:

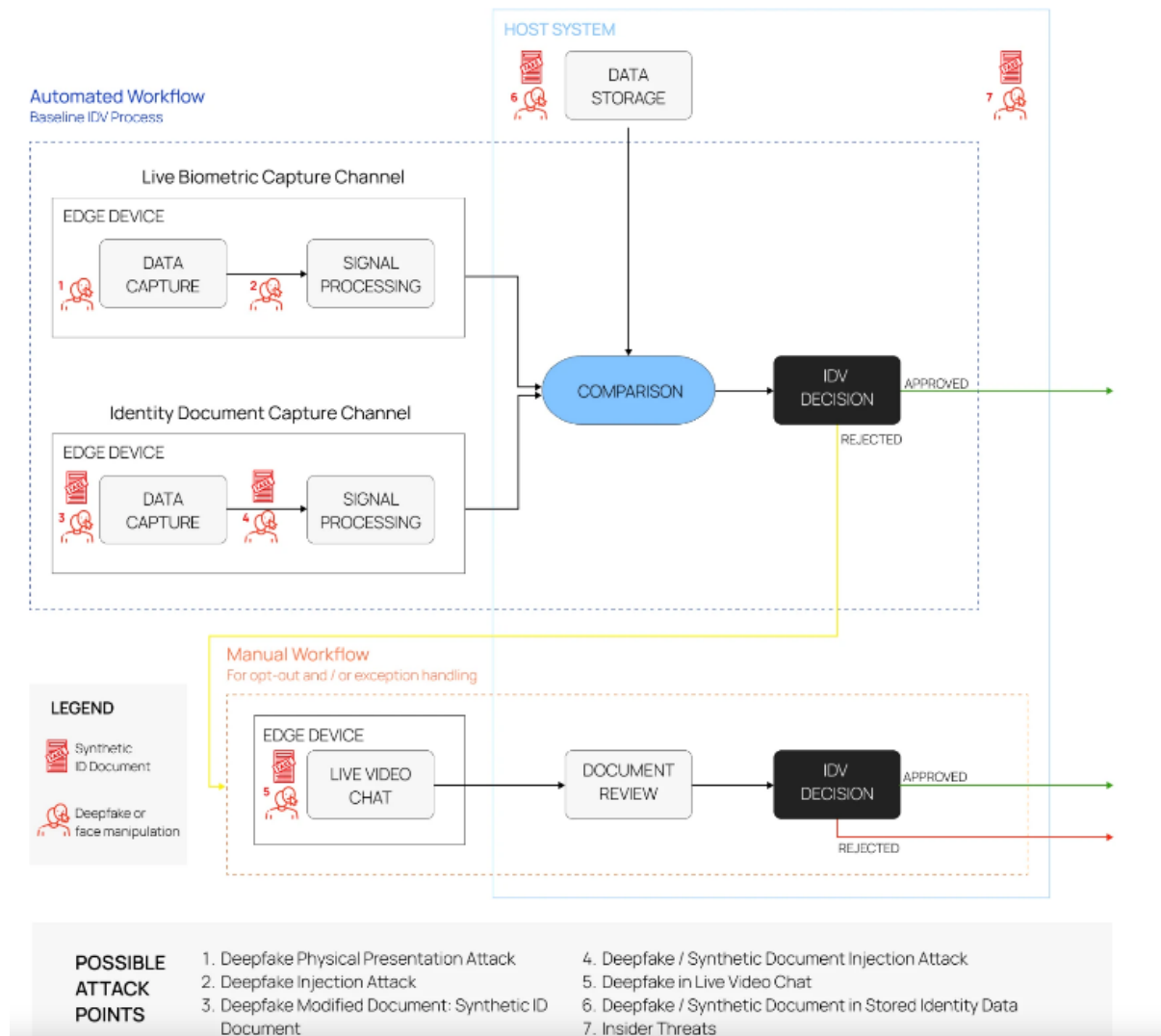


Figure 1: Deepfake threats and attack vectors in remote identity verification

Source: (Paravision, 2024).

4.1 Deepfake Evolution and Detection Challenges

Advances in generative adversarial networks (GANs) have introduced the concept of deepfakes, where fraudulent media is essentially indistinguishable from ones that are authentic. These GANs are AI models capable of generating synthetic biometric content, including facial images, voice recordings, and other identity-related media. While the primary focus is on facial biometrics, this report acknowledges the continued threat from voice cloning and the use of synthetic fingerprints. Detection methods such as facial landmark inconsistencies and GAN artifact analysis remain limited in real-world deployment. Without

robust liveness detection and multi-modal biometric safeguards, identity verification systems are highly vulnerable.

In LMICs, DPI systems often rely on low-cost biometric solutions, and any resulting absence of robust anti-spoofing measures is a major concern. This is particularly an issue since researchers have found that deepfake videos could bypass facial verification with success rates exceeding 80% (Vyshegorodtsev et al., 2024). The challenge in combatting the use of AI for identity fraud. For example, in Nigeria, over 6000 AI-assisted image morphing and finger-blending cases have been reported in examination fraud by the Special Committee on Examination Infractions set up by the Joint Admissions and Matriculation Board (The Guardian, 2025). The committee recommended that the Joint Admissions and Matriculation Board “deploy AI-powered biometric anomaly detection, dual verification systems, real-time monitoring, and a National Examination Security Operations Centre”.

4.2 Face Morphing as a Biometric Spoofing Technique

Face morphing involves blending features from two or more individuals into a single composite image that can deceive both human inspectors and automated facial recognition systems. When such morphed images are used during enrolment, they can generate biometric templates that falsely match multiple people, allowing shared access to identity-linked services. Research by Bello and Thomas (2023) highlights that similar manipulation can exploit weak internal controls, leading to inflated false acceptance rates, ghost-worker records, and other forms of credential fraud. As identity systems increasingly underpin payroll, welfare, and border-control processes, morphing represents a critical integrity risk, especially where verification pipelines rely solely on facial recognition.

The growing accessibility of morph-generation tools and AI-based image editing compounds this threat. Countermeasures include integrating morph-detection algorithms, trusted or in-person enrolment processes, and multi-modal biometric verification that combines face, fingerprint, or iris data. Equally important are governance and policy responses that include certification of biometric vendors, auditable capture provenance, and capacity-building for enrolment operators, to ensure resilience across Digital Public Infrastructure (DPI). For low- and middle-income countries, where cybersecurity and oversight capacities may be limited, addressing morphing risks early is vital to maintaining public trust and preventing systemic abuse of digital identity systems.

4.3 GANs as Dual-Use Tools

Generative Adversarial Networks (GANs) have become quintessential dual-use technologies that are capable of both enabling and undermining digital identity systems. They are used to generate synthetic facial identities and deepfakes that can facilitate impersonation or credential fraud, yet the same architectures are increasingly employed to strengthen fraud detection systems by creating synthetic training datasets and simulating attack scenarios. This co-evolution blurs the boundary between attacker and defender innovation. As GANs improve, detection models trained on earlier data rapidly lose efficacy, shortening the defensive lead time.

In the context of Digital Public Infrastructure, especially in low- and middle-income countries where regulatory and cybersecurity capacities may be limited, this duality raises urgent governance challenges. It underscores the need for robust AI assurance frameworks, auditable data provenance, and secure model deployment practices to ensure that the benefits of synthetic data and AI-driven fraud detection are not outweighed by escalating adversarial risks.

4.4 Global Scale and Urgency of the Threat

Credential and identity fraud have escalated worldwide, fuelled by AI-generated content, weak verification ecosystems, and expanding digital service coverage. In the UK alone, total fraud losses reached £1.17 billion in 2023 across 2.97 million cases (UK Finance, 2024). Synthetic identity fraud and deepfake-enabled impersonation now constitute some of the fastest-growing attack vectors in digital banking and fintech. The UK Government Cyber Security Strategy (Cabinet Office, 2022) notes that 40% of incidents handled by the NCSC targeted the public sector, underscoring systemic vulnerability in identity-dependent services. Key statistics highlight the accelerating scale:

- Nearly 3 million AI-driven synthetic identities circulate within UK systems.
- Businesses lose an estimated £15,000 per confirmed synthetic identity case.
- National losses could exceed £4.2 billion by 2027.

Comparable patterns are emerging globally. In India, Aadhaar-linked benefit systems have faced deepfake-enabled impersonation attempts, prompting investment in biometric anti-spoofing and AI-based anomaly detection (Protean eGov Technologies, 2025). In South Africa, the Department of Home Affairs reported a surge in synthetic identity registrations linked to fraudulent passports (Phandle, 2024). Meanwhile, Australia has committed over \$20 billion to strengthen cyber resilience, including enhanced biometric verification and counter-identity operations (Australian Information Security Association, 2024).

These trends illustrate that AI-driven credential fraud is no longer confined to high-income economies. Low- and middle-income countries (LMICs) face disproportionate exposure, as limited cyber resilience, fragmented governance, and infrastructural constraints magnify the risks to Digital Public Infrastructure (DPI)—particularly where social protection, health, and financial systems rely on digital identity rails without robust fraud analytics or cross-sector coordination.

4.5 Exploitation of Systemic Vulnerabilities

Synthetic identity and deepfake threats often exploit known vulnerabilities catalogued in the National Vulnerability Database (NVD). According to The Alan Turing Institute (2025), CVEs linked to identity systems increased by over 300%, with a concentration in the following Common Weakness Enumerations (CWEs) between 2020 and 2024:

- CWE-287: Improper Authentication
- CWE-863: Incorrect Authorisation
- CWE-200: Information Exposure
- CWE-306: Missing Authentication for Critical Function
- CWE-798: Use of Hardcoded Credentials

These weaknesses are frequently found in federated identity systems, single sign-on (SSO) platforms, and API-based authentication flows. For example:

- CVE-2023-23397 (*Microsoft, 2023*): Exploited authentication tokens.
- CVE-2024-20328 (*Cisco, 2024*): Exposed flaws in multi-factor authentication logic.

Further to the above is the presence of specific deepfake relevant CWEs which have been increasing over time. The identified CWEs which include CWE-79, CWE-601, CWE-918, CWE-352, CWE-20, CWE-22 and CWE-611 highlight systemic vulnerabilities that need addressing especially in DPI systems.

In LMICs, many DPI systems rely on open-source identity platforms or third-party providers, increasing exposure to these vulnerabilities. Without regular patching, secure API design, and encryption-at-rest policies, these systems remain attractive targets.

4.6 Sectoral Impact and Attack Vectors

Credential fraud enabled by generative AI has become global and impacts various sectors of an economy. To understand the scale, level of sophistication, the systemic impact of synthetic identity threats and deepfake-enabled credential fraud, this section analysed real-world case studies across three critical sectors: finance, healthcare, and government. These sectors are highly dependent on digital identity systems and have experienced significant disruptions due to identity-related vulnerabilities. Each case study highlighted attack vectors, exploited vulnerabilities (CVEs/CWEs), and systemic implications, while integrating insights from DPI governance models, safeguards and national cyber resilience strategies including the LMICs.

4.6.1 Finance Sector: Deepfake-Enabled Money Laundering in Fintech

In early 2024, a Southeast Asian fintech startup was targeted by a transnational fraud ring using AI-driven identity fraud to launder illicit funds. Attackers created hundreds of fake user profiles using deepfake videos, AI-generated documents, and stolen personal data purchased from dark web marketplaces. The startup's onboarding process relied on asynchronous video KYC verification, which was bypassed using pre-recorded deepfake clips. This development in Southeast Asia resulted in fintech platforms recording a 116% rise in synthetic identity fraud, with deepfake-enabled laundering attacks targeting onboarding APIs (Fintech News Singapore, 2025).

A similar incident occurred in the UK, where a major digital bank reported a coordinated fraud campaign involving AI-driven identity fraud and deepfake facial videos. Attackers exploited weaknesses in the bank's KYC process, which lacked robust liveness detection and behavioural biometrics (Burt, 2024).

Attack Vector: Network-based (API abuse during onboarding).

Exploited CWEs:

- CWE-287: Improper Authentication
- CWE-200: Information Exposure
- CWE-639: Insecure Direct Object Reference (IDOR)

Systemic Implications:

- Demonstrated the scalability of synthetic identity fraud in digital lending ecosystems.
- Triggered regulatory audits mandating liveness detection and behavioural analytics.
- Increase awareness of the dual-use nature of GANs in both fraud and fraud detection.
- Reinforced the need for sector-wide threat intelligence sharing, as promoted by the UK's Government Cyber Coordination Centre (GCCC).

Impact:

- Over £10 million in fraudulent loans and credit lines were issued before detection.
- The UK Financial Conduct Authority (FCA) launched a regulatory review recommending mandatory liveness detection for remote identity verification (Jumio, 2020).

4.6.2 Healthcare Sector: Biometric Spoofing in Patient Portals

In 2023, a major hospital network in the United States experienced a breach involving AI-driven deepfake-based impersonation using AI-generated facial images and forged documents. Attackers exploited weaknesses in the hospital's remote onboarding system, which relied on facial recognition but lacked liveness detection and behavioural biometrics (TechRadar, 2025).

Face morphing techniques were used to register synthetic patients whose biometric profiles could match multiple real individuals, complicating detection and remediation. In South Africa, similar vulnerabilities were found in provincial health systems where biometric authentication was used for patient registration (Techish Kenya, 2025). The lack of multi-modal verification and encryption-at-rest policies exposed sensitive health data to manipulation and fraud.

Attack Vector: Local and network-based

Exploited CWEs:

- CWE-306: Missing Authentication for Critical Function
- CWE-284: Improper Access Control
- CWE-798: Use of Hardcoded Credentials

Systemic Implications:

- Exposed the fragility of biometric-only verification systems.
- Highlighted the need for morph detection algorithms and multi-modal authentication.
- Prompted updated guidance from the U.S. Department of Health and Human Services (HHS) on biometric data protection under HIPAA (HHS, 2024).
- Echoed the UK strategy's emphasis on secure-by-design principles and data classification policies.

Impact:

- Over 1.2 million patient records were compromised.
- The breach led to a class-action lawsuit and new federal guidance on biometric encryption and access control.

4.6.3 Government Sector: Synthetic Identity Fraud in National ID Systems

In late 2024, El Salvador experienced a massive data breach involving its national digital identity platform. Over 5.1 million citizens' national ID numbers (DUIs), facial images and other biometric data were leaked on the dark web (Resecurity, 2024). Investigations revealed a misconfigured API endpoint and inadequate encryption of biometric data. The exposure of such facial images and biometric identifiers significantly increases the risk of synthetic identity creation. These datasets can be used to train generative adversarial networks (GANs) to produce deepfakes or synthetic biometric personas, potentially undermining future authentication systems and video-based KYC processes.

Between 2022 and 2024, Germany recorded significant rise (1500%) in forced verifications after attackers used high-resolution deepfake videos and morphing techniques to bypass video-based KYC verification systems (Sumsub, 2023). Over 3,000 fraudulent digital IDs were issued, some linked to cross-border SIM card fraud and benefit theft.

In Nigeria, the National Identity Management Commission (NIMC) reported a surge in synthetic identity registrations linked to fraudulent SIM card activations and mobile money scams (News Digest, 2025). These incidents prompted calls for biometric anti-spoofing and encryption-at-rest policies.

Attack Vector: Network and adjacent network (API spoofing and federated identity compromise)

Exploited CWEs:

- CWE-862: Missing Authorisation
- CWE-200: Information Exposure
- CWE-918: Server-Side Request Forgery (SSRF)

Systemic Implications:

- Exposed the risks of relying on third-party identity providers without rigorous biometric anti-spoofing.
- Led to national reviews of digital identity governance and the adoption of biometric encryption and tokenised identity attributes.
- Prompted the European Data Protection Board (EDPB) to recommend stronger anti-spoofing standards under eIDAS 2.0.
- Reinforced the UK strategy's call for cross-government cyber assurance and secure data sharing.

Impact:

- Leaked data was used to create deepfakes for cross-border fraud.
- The breach triggered calls for biometric data localisation and encryption-at-rest policies.

4.6.4 Cross-Sectoral Observations

This study identified and focused on risks to three key sectors (Finance, Health and Government). Across all three sectors, common patterns emerge:

- Federated identity systems are often targeted, especially when relying on third-party providers without rigorous auditing.
- Biometric data, once compromised, are difficult to be “reset” like passwords, making its protection paramount.
- AI-driven identity fraud is increasingly used not just for financial gain, but also for espionage, disinformation, and disruption of public services.
- Cyber resilience must be embedded across DPI systems, with coordinated governance, threat intelligence sharing, and secure-by-design architecture.

The above case studies reinforce the need for a unified, cross-sectoral approach to security strategies where the system ensures the integration of technical controls, regulatory oversight, and international cooperation as outlined in the UK Government Cyber Security Strategy. These approaches enhance the DPI systems' security and to mitigate the evolving risks posed by generative AI-powered credential fraud.

4.7 Emerging Trends

Recent evidence points to several converging trends that are rapidly reshaping the threat landscape for digital identity and public infrastructure:

- **AI-as-a-Service:** Commercial and open platforms now offer on-demand deepfake and morph-generation capabilities, dramatically reducing the skill and cost barriers for attackers.

- **Synthetic Media Marketplaces:** Dark web forums actively trade synthetic biometric profiles and deepfake kits tailored to bypass specific national identity or KYC systems.
- **Cross-Sector Propagation:** Breaches in one sector (e.g., healthcare) increasingly expose authentication credentials used in another (e.g., finance or government), reflecting the growing interdependence of digital identity ecosystems.
- **Commoditised Cybercrime:** As highlighted in the *UK Government Cyber Security Strategy* (UK Cabinet Office, 2022), the rise of Ransomware-as-a-Service (RaaS) and similar models has industrialised attack capabilities, fuelling a surge in identity-based cyber incidents.

Collectively, these developments underscore the urgency of cross-sector, proactive cybersecurity strategies that integrate threat intelligence, coordinated response mechanisms, and shared defensive capabilities. The UK’s “*Defend as One*” approach exemplifies such alignment, promoting unified resilience across government systems.

The following section analyses illustrative cases from the financial, health, and government sectors to examine how these emerging dynamics manifest in practice.

5. CVE/CWE Analysis

Understanding the technical foundations of synthetic identity and deepfake-enabled credential fraud requires a close examination of Common Vulnerabilities and Exposures (CVEs) and their associated Common Weakness Enumerations (CWEs). This section analyses trends in CVE disclosures from 2020 to 2025, focusing on vulnerabilities that directly impact digital identity systems, biometric authentication, and federated login architectures, which serve as the core components of Digital Public Infrastructure (DPI).

5.1 Overview of CVE Trends (2020–2025)

Between 2020 and 2024, the National Vulnerability Database (NVD) recorded a significant increase in CVEs affecting identity-related systems. The number of CVEs referencing identity verification, authentication tokens, biometric modules, and federated login APIs rose by over 300%, reflecting both the growing adoption of DPI platforms and the expanding attack surface (The Alan Turing Institute, 2025). The diagrams in Figure 2 below presents the CVE distribution from the NVD Database.

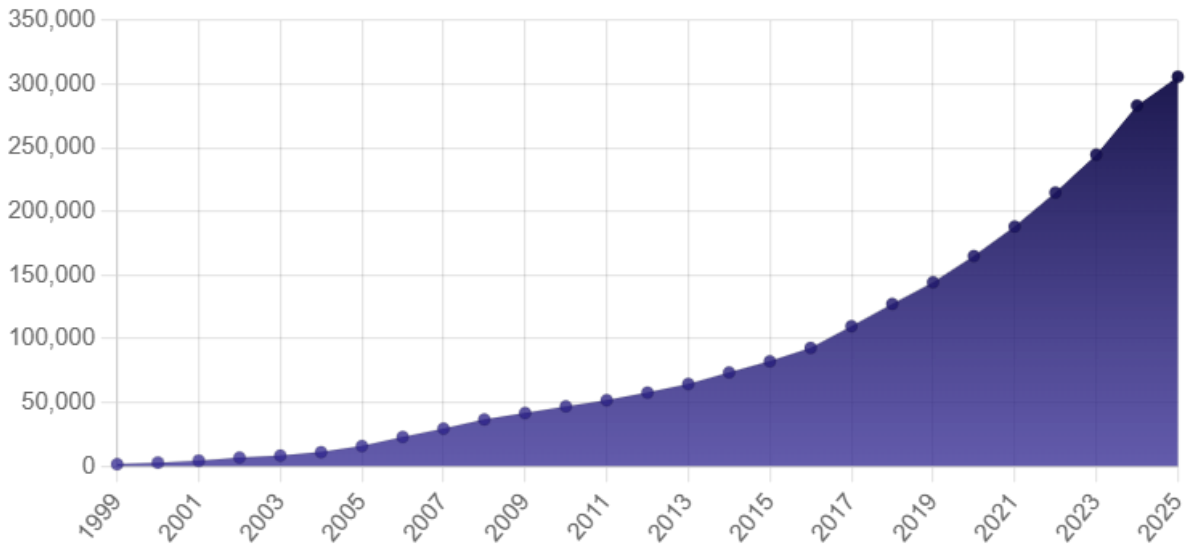


Figure 2: Yearly Cumulative count of CVEs Distribution

Emerging threats such as face morphing and deepfake media injection are not yet fully captured in CVE taxonomies. For this research, the study analysed from the National Vulnerability Database (NVD) CVE and CWE data from 2020–2025 and focused on deepfake-related vulnerabilities. The details show an increasing rate of vulnerabilities over time as shown below in Figure 3.

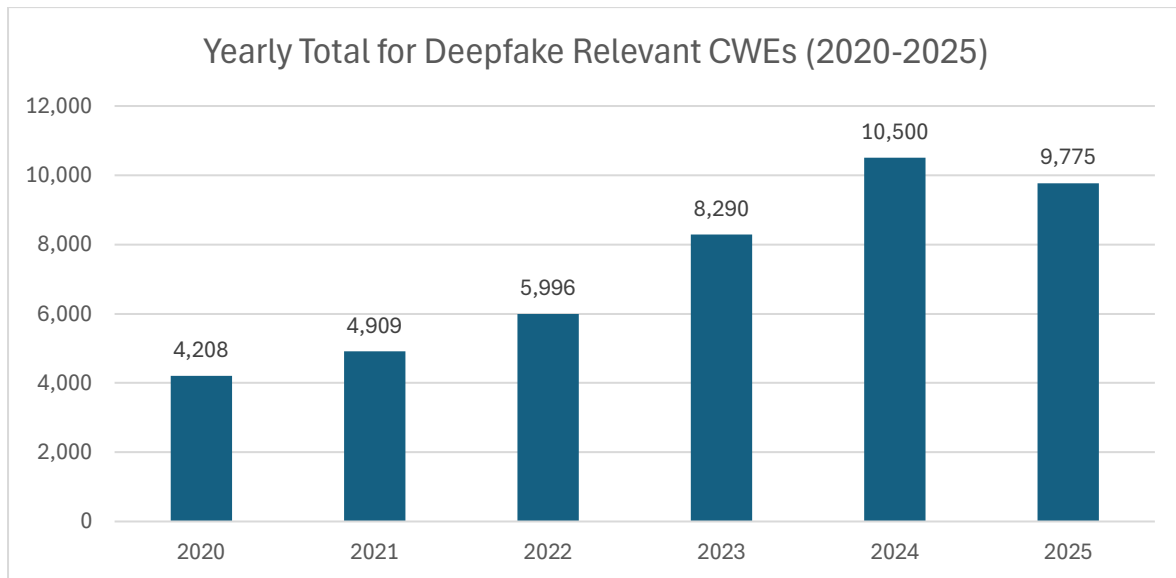


Figure 3: Yearly Deepfake Relevant CWE Distribution between 2020 to 2025 (September)

It can be observed that deepfake-relevant CWEs increasing on year-to-year basis with a recorded figure of 9775 as at September 2025 where the year has not ended at the time of writing this report.

Table 1: Deepfake CWE Yearly Distribution to September 2025

Year	CWE-79 (XSS)	CWE-601 (Open Redirect)	CWE-918 (SSRF)	CWE-352 (CSRF)	CWE-20 (Improper Input Validation)	CWE-22 (Dir. Traversal)	CWE-611 (XXE)	Yearly Total
2020	2,201	100	130	414	808	436	119	4,208
2021	2,724	133	188	520	671	547	126	4,909
2022	3,378	137	230	766	672	690	123	5,996
2023	5,102	168	240	1,392	522	742	124	8,290
2024	7,443	113	372	1,434	103	925	110	10,500
2025	6,778	122	422	1,574	0	790	89	9,775

Analysing the CWEs, we can identify those that can serve as practical paths that attackers could use to deliver or legitimise deepfake content:

- XSS (CWE-79) for Cross-Site Scripting could be an avenue for adversaries to embed or auto-play deepfake media or manipulate page chrome to impersonate sources. The NVD defined Improper Neutralization of Input during Web Page Generation as “Cross-site Scripting”. The NVD Slice (2025) describes the Cross-site Scripting (XSS) as where “The software does not neutralize or incorrectly neutralises user-controllable input before it is placed in output that is used as a web page that is served to other users”. Mitre, 2025 also adds that “Some cross-site scripting vulnerabilities can be exploited to manipulate or steal cookies, create requests that can be mistaken for those of a valid user, compromise confidential information, or execute malicious code on the end user systems for a variety of nefarious purposes. Mitre (2025) provides varied ways that attackers could use to inject deep fakes by stating that “The attack methods for XSS can vary depending on the type of XSS and the attacker's goal”. The following three process were outlined by mitre (2025):
 - “Reflected XSS exploits (Type 1) occur when an attacker causes a victim to supply dangerous content to a vulnerable web application, which is then reflected back to the victim and executed by the web browser. The most common mechanism for delivering malicious content is to include it as a parameter in a URL that is posted publicly or e-mailed directly to the victim. URLs constructed in this manner constitute the core of many phishing schemes, whereby an attacker convinces a victim to visit a URL that refers to a vulnerable site. After the site reflects the attacker's content back to the victim, the content is executed by the victim's browser”.
 - “In a Stored XSS exploit (Type 2), the optimal place to inject malicious content is in an area that is displayed to either many users or particularly interesting users. Interesting users typically have elevated privileges in the application or interact with sensitive data that is valuable to the attacker. If one of these users executes malicious content, the attacker may be able to perform privileged operations on behalf of the user or gain access to sensitive data belonging to the user. For example, the attacker might inject XSS into a log message, which might not be handled properly when an administrator views the logs”.

- “DOM-based XSS (Type 0) generally involves server-controlled, trusted script that is sent to the client, such as JavaScript that performs sanity checks on a form before the user submits it. If the server-supplied script processes user-supplied data and then injects it back into the web page (such as with dynamic HTML), then DOM-based XSS is possible”.

The above according to Mitre (2025) helps attackers to “encode the malicious portion of the attack, such as URL encoding or Unicode, so the request looks less suspicious” and “emulate trusted web sites and trick the victim into entering a password, allowing the attacker to compromise the victim's account on that web site”.

- Open Redirect (CWE-601) is commonly used in phishing to route users through trusted domains to attacker-controlled deepfake sites. Mitre (2025) states “The phishing attack may point to an attacker-controlled web page that appears to be a trusted web site” and “the phishers may then steal the user's credentials and then use these credentials to access the legitimate web site”. Metre (2025) further adds that “the user will then be redirected to the attacker's web site (attacker.example.net) which the attacker may have made to appear very similar to bank.example.com” and “the user may then unwittingly enter credentials into the attacker's web page and compromise their bank account”.

Additionally, the NVD Slice (2025) describes the Open Redirect as where “A web application accepts a user-controlled input that specifies a link to an external site and uses that link in a Redirect”.

- SSRF (CWE-918) for Server-Side Request Forgery is used to fetch internal or cloud-hosted content (e.g., pulling staged media or exfiltrating artifacts to support convincing fakes). Regarding CWE-918, the NVD Slice (2025) described it as where “the web server receives a URL or similar request from an upstream component and retrieves the contents of this URL, but it does not sufficiently ensure that the request is being sent to the expected destination”. Mitre (2025) also adds that “By providing URLs to unexpected hosts or ports, attackers can make it appear that the server is sending the request, possibly bypassing access controls such as firewalls that prevent the attackers from accessing the URLs directly”. Mitre (2025) further adds that “The server can be used as a proxy to conduct port scanning of hosts in internal networks, use other URLs such as that can access documents on the system (using file://), or use other protocols such as gopher:// or tftp://, which may provide greater control over the contents of requests”.
- CSRF (CWE-352) for Cross-Site Request Forgery can be used by adversaries to force unauthorised actions such as posting or sharing manipulated media from legitimate accounts). The NVD Slice (2024) also describes it as where “the web application does not, or cannot, sufficiently verify whether a well-formed, valid, consistent request was intentionally provided by the user who submitted the request”. Mitre (2025) also clarified that “an attacker could trick a client into making an unintentional request to the web server via a URL, image load, XMLHttpRequest, etc., which would then be treated as an authentic request from the client - effectively performing any operations as the victim, leading to an exposure of data, unintended code execution, etc” and “If the victim is an administrator or privileged user, the consequences may include obtaining complete control over the web application - deleting or stealing data, uninstalling the product, or using it to launch other attacks against all of the product's users”. Mitre (2025) further adds “Because the attacker has the identity of the victim, the scope of CSRF is limited only by the victim's privileges”.
- Improper Input Validation (CWE-20) can potentially lead to unexpected values and cause a programme to crash or lead to data compromise. The NVD Slice (2025) describes CWE-20 as where “the product receives input or data, but it does not validate

or incorrectly validates that the input has the properties that are required to process the data safely and correctly". Mitre (2025) explains that "an attacker could use malicious input to modify data or possibly alter control flow in unexpected ways, including arbitrary command execution" and "an attacker could provide unexpected values and cause a program crash or arbitrary control of resource allocation, leading to excessive consumption of resources such as memory and CPU".

- Path Traversal (CWE-22) can lead to filesystem abuse for hosting/injection to enhance fake contents sharing. Regarding the CWE-22, the NVD Slice (2025) describes it as where "The software uses external input to construct a pathname that is intended to identify a file or directory that is located underneath a restricted parent directory, but the software does not properly neutralize special elements within the pathname that can cause the pathname to resolve to a location that is outside of the restricted directory". The Mitre (2025) explains that "By using special elements such as ".." and "/" separators, attackers can escape outside of the restricted location to access files or directories that are elsewhere on the system". Mitre (2025) further adds that "The attacker may be able to create or overwrite critical files that are used to execute code, such as programs or libraries"; "The attacker may be able to overwrite or create critical files, such as programs, libraries, or important data. If the targeted file is used for a security mechanism, then the attacker may be able to bypass that mechanism. For example, appending a new account at the end of a password file may allow an attacker to bypass authentication"; and "The attacker may be able read the contents of unexpected files and expose sensitive data. If the targeted file is used for a security mechanism, then the attacker may be able to bypass that mechanism. For example, by reading a password file, the attacker could conduct brute force password guessing attacks in order to break into an account on the system". Mitre (2025) adds that "The attacker may be able to overwrite, delete, or corrupt unexpected critical files such as programs, libraries, or important data. This may prevent the product from working at all and in the case of protection mechanisms such as authentication, it has the potential to lock out product users".
- XXE (CWE-611) could be used for internal fetch/data exfil or Improper Restriction of XML External Entity Reference) to "prove" authenticity for deepfakes. The NVD Slice (2025) describes CWE-611 as where "The software processes an XML document that can contain XML entities with URIs that resolve to documents outside of the intended sphere of control, causing the product to embed incorrect documents into its output". Mitre 2025 explains that "If the attacker is able to include a crafted DTD and a default entity resolver is enabled, the attacker may be able to access arbitrary files on the system. By submitting an XML file that defines an external entity with a file:// URI, an attacker can cause the processing application to read the contents of a local file". Mitre (2025) further adds that "An attacker may supply a crafted DTD using URIs with schemes such as http://, forcing the application to make outgoing HTTP requests to servers that the attacker cannot reach directly, which can be used to bypass firewall restrictions; hide the source of attacks such as port scanning; or otherwise leverage the server's trust relationship with other entities".

The above highlight that as biometric spoofing techniques evolve, CVEs will need to expand to include vulnerabilities in morph detection algorithms and AI-generated media validation.

Trend highlights from both the identity-related and deepfake vulnerabilities analysis show that:

- 2020–2021: Legacy authentication flaws and misconfigured access controls in government and healthcare systems.
- 2022–2023: Surge in biometric-related vulnerabilities, especially in mobile SDKs and facial recognition APIs.

- 2024: Emergence of some deepfake-specific CVEs targeting liveness detection bypasses and spoofing-resistant biometric protocols.

These trends align with the UK Government Cyber Security Strategy’s emphasis on proactive vulnerability management and the need for a cross-government vulnerability reporting service to accelerate detection and remediation.

5.2 Dominant CWEs in Identity Systems.

Recurring classes of Common Weakness Enumerations (CWEs) are increasingly responsible for critical vulnerabilities within digital identity ecosystems. Data from the National Vulnerability Database (2024) indicates that these systemic software weaknesses continue to shape the overall risk profile of identity platforms.

Figure 4 below illustrates the distribution of the most prevalent CWE categories. In addition to those shown, the following CWEs were most frequently associated with CVEs affecting digital identity systems (Table 2).

These weaknesses are particularly dangerous in federated identity systems, where a single compromised token or misconfigured API can grant access across multiple services. The Government Cyber Security Strategy underscores the importance of secure-by-design architecture and the adoption of the Cyber Assessment Framework (CAF) to mitigate such systemic risks.

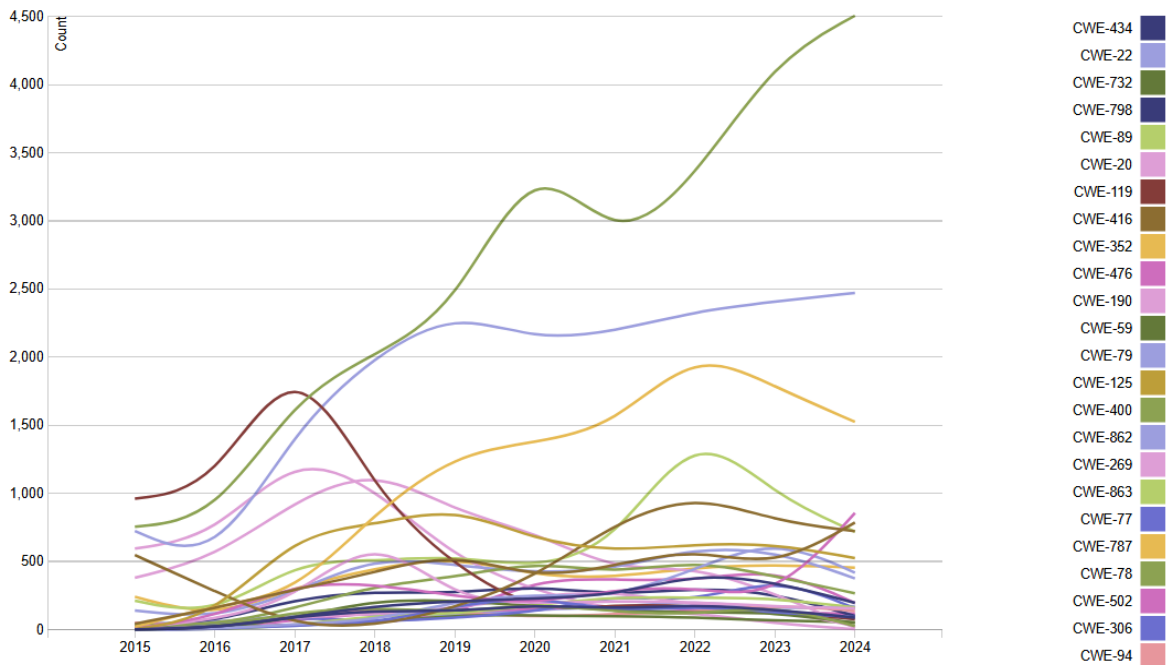


Figure 4: Yearly Dominant CWE Distribution

Source: National Vulnerability Database, 2024

Table 2: CWEs and their CVEs impact

CWE ID	Name	Description	Frequency (2020–2025)
--------	------	-------------	-----------------------

CWE-287	Improper Authentication	Failure to verify identity correctly	High (160+ CVEs)
CWE-863	Incorrect Authorisation	Inadequate enforcement of access controls	High (140+ CVEs)
CWE-862	Missing Authorisation	Absence of access control checks	Moderate (90+ CVEs)
CWE-200	Information Exposure	Leakage of sensitive identity data	High (120+ CVEs)
CWE-306	Missing Authentication for Critical Function	Unprotected access to sensitive operations	Moderate (80+ CVEs)
CWE-798	Use of Hardcoded Credentials	Static credentials embedded in code	Moderate (60+ CVEs)
CWE-918	Server-Side Request Forgery (SSRF)	Exploitation of internal services via crafted requests	Emerging (30+ CVEs)
CWE-639	Insecure Direct Object Reference (IDOR)	Manipulation of user-controllable identifiers	High (75+ CVEs)

5.3 Attack Vector Classification

The study assessed and analysed the publicly available CVSS data from the National Vulnerability Database. Using CVSS metrics, CVEs were classified by attack vector as detailed below in Table 3.

Table 3: CVEs Attack Vector Classification Results using CVSS Metrics

Attack Vector	Prevalence	Primary Targets	Key CWEs
Network	High	APIs, federated login, cloud services	CWE-287, CWE-863, CWE-918
Local	Moderate	Mobile SDKs, client-side apps	CWE-798, CWE-306
Physical	Emerging	Biometric sensors, kiosks	CWE-287, CWE-306
Adjacent Network	Moderate	Shared infrastructure, federated systems	CWE-862, CWE-200

This classification highlights the need for layered defences that address both remote and proximity-based threats. Bello and Thomas (2023) reinforce the importance of recognising morphing attacks as a physical and local threat vector, particularly in biometric onboarding systems. The UK strategy also recommends embedding security controls throughout the technology lifecycle and ensuring that detection capabilities are proportionate to risk (UK Cabinet Office, 2022).

5.4 DPI Governance Stakeholder Map

Digital Public Infrastructure (DPI) governance requires effective strategy that includes a whole-of-society approach. The Universal DPI Safeguards Framework and the CCAF DPI–DFS Convergence report both emphasise a multi-stakeholder approach to DPI governance, involving government, regulators, technology providers, donors, and civil society (DPI Safeguards Initiative, 2024; Cambridge Centre for Alternative Finance, 2025). The UN Universal DPI Safeguards Framework report for instance, identifies stakeholder map consisting of five key groups as detailed in Figure 5 below. Additionally, the report detailed the roles of the stakeholders as summarised in Table 4.

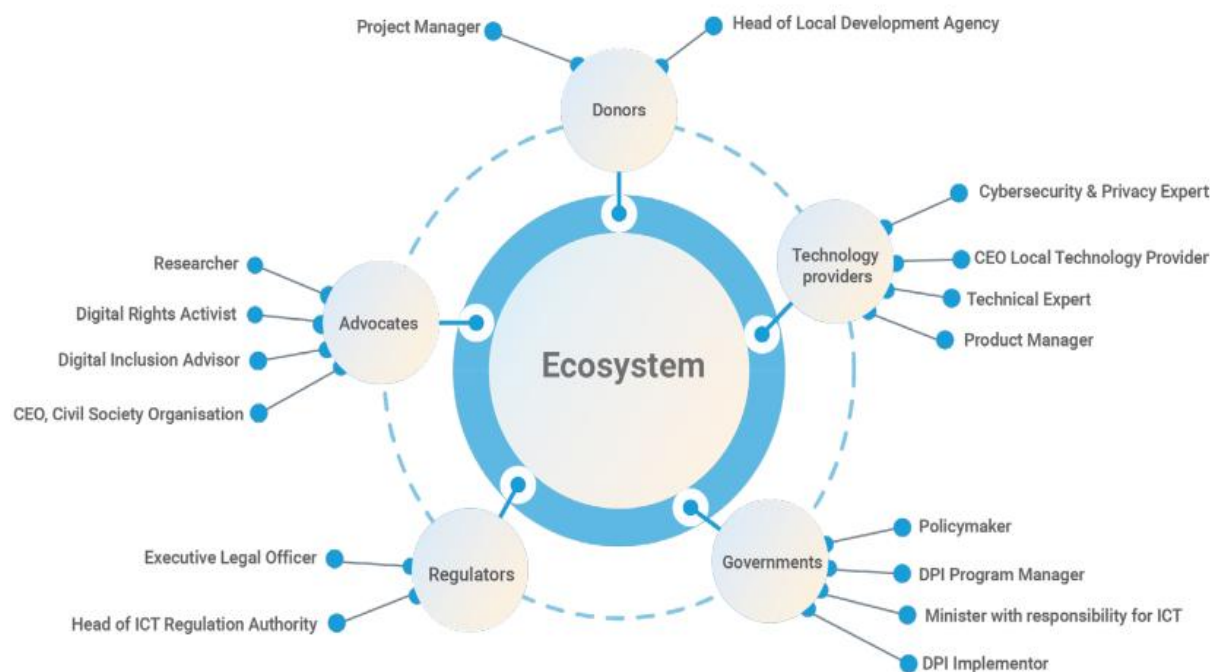


Figure 5: Mapping of DPI Ecosystem

Source: Universal DPI Safeguards Framework, Figure 2.1 (pp. 16)

Table 4: Stakeholders' Role and Responsibilities in DPI

Source: Universal DPI Safeguards Framework, Table 2.1 (pp. 17–19)

Stakeholder Group	Example Roles	DPI Lifecycle Responsibilities
Governments	Policymakers, ICT Ministers	Governance, funding, legal frameworks
Regulators	Legal officers, ICT regulators	Standards, enforcement, oversight
Donors	Development agencies	Funding, progress tracking
Technology Providers	Cybersecurity experts, product managers	Implementation, risk mitigation, technical support
Advocates	Civil society, digital rights activists	Inclusion, redress, community feedback

Drawing on the Universal DPI Safeguards Framework, the CCAF DPI–DFS Convergence report and the UK Government Cyber Security Strategy, this research proposes that the key actors, roles, and lifecycle responsibilities that apply to DPI governance in the LMICs, as shown in Table 5.

Table 5: Stakeholders' Roles and Responsibilities in DPI Governance

Stakeholder Group	Example Roles	Lifecycle Responsibilities
Governments	ICT ministries, central banks, treasury departments, User government department and agencies	Strategic direction, legal frameworks, funding, public service integration
Regulators	Financial authorities, data protection agencies, competition bodies	Standards setting, enforcement, oversight, risk mitigation
Technology Providers	Cybersecurity firms, Fintechs, platform developers	Infrastructure design, implementation, technical support, innovation
Donors and Development Partners	Multilateral institutions, aid agencies	Funding, capacity building, monitoring and evaluation
Civil Society and Advocates	Digital rights groups, consumer protection bodies	Inclusion, transparency, redress mechanisms, community engagement

These groups operate across three governance tiers:

- Policy: Strategic mandates, legal harmonisation, DPI principles integration.
- Technical: Secure-by-design architecture, interoperability standards, privacy safeguards.
- Operational: Monitoring, feedback loops, incident response, user support.

The CCAF report emphasises that fragmented oversight across these groups can undermine DPI effectiveness and recommends the need to establish cross-regulatory coordination mechanisms, such as national DPI councils or inter-agency forums, to ensure coherence and accountability.

5.5 DPI Maturity vs. Financial Inclusion Outcomes

The maturity of Digital Public Infrastructure (DPI) systems with regards to the integration of digital identity, real-time payments, and consent-based data sharing are having a considerable impact on financial-services outcomes, i.e. inclusion, credit, payments, transfers, and resilience, as well as regulatory/governance aspects, i.e. market integrity, competition, and privacy/data governance (Cambridge Centre for Alternative Finance, 2025). For instance, the CCAF DPI–DFS Convergence report presents empirical evidence linking DPI maturity to improvements in access to financial services by assessing 113 jurisdictions based on the presence of three core DPI components:

- Digital Identity
- Real-Time Payments
- Consent-Based Data Sharing

Each jurisdiction received a DPI score from 1 to 3:

- Score 1: One component implemented
- Score 2: Two components implemented
- Score 3: All three components implemented

From the CCAF report we observe trends shown in table 6 which illustrate jurisdictions with higher DPI maturity consistently report better financial inclusion outcomes. The figures also show a strong relationship between DPI maturity and improved financial access:

- Credit access triples from 25% to 77% as DPI components increase.
- Digital payment usage rises from 45% to 83%.
- Government transfer receipt doubles from 14% to 28%.

Table 6: DPI Financial Maturity Assessment
Source: Cambridge Centre for Alternative Finance (2025)

DPI Score	Credit Access (%)	Digital Payments (%)	Gov Transfers (%)
1	25	45	14
2	39	56	17
3	77	83	28

Interpretation and Strategic Implications

The data presented suggests that jurisdictions with more mature DPI ecosystems are better positioned to:

- Expand access to formal financial services and social protection services.
- Deliver targeted government support efficiently, with higher efficiency and reduced leakage.
- Lower entry barriers by addressing documentation gaps and streamlining digital onboarding.

This reinforces the strategic importance of investing in secure, interoperable, and inclusive DPI systems. It also highlights the need for cross-sectoral coordination to ensure that DPI components are not developed in isolation but as part of a cohesive digital infrastructure strategy.

5.6 Emerging Technologies and Regulatory Challenges

This report highlights that deepfakes, synthetic identities, and AI-enabled credential manipulation present a significant and fast-evolving threat to the integrity of digital identity

systems. These risks undermine user trust, authentication reliability, and the credibility of digital public infrastructure (DPI) as a whole.

Technical solutions, including distributed ledger technologies (DLT), artificial intelligence (AI), and digital wallets, can help mitigate these vulnerabilities while balancing innovation, security, and inclusion. Distributed Ledger Technology (DLT) refers to a decentralised database system in which records (or “ledgers”) are shared and synchronised across multiple nodes, reducing reliance on a single authority. By design, DLT offers immutability, transparency, and traceability, which can strengthen auditability and data integrity in identity ecosystems.

Despite their benefits to DPI, technologies such as DLT, AI, and digital wallets embed trade-offs that directly affect system design, regulatory strategy, and user outcomes. While DLT may improve transparency and traceability, it can also increase energy consumption or create new risks of decentralised accountability. AI can automate compliance and personalise services, but it also raises concerns about bias, opacity, and control over decision-making. Digital wallets have enabled real gains in payment inclusion, yet they often come bundled with surveillance risks and limited data protection.

This section examines key trends shaping the future of DPI, categorised across technological, political-regulatory, and socio-economic domains. It explores both the promise, and the risks associated with each development, supported by global case studies and expert insights. The goal is to inform a more holistic and anticipatory approach to DPI policy, helping decisionmakers design infrastructure that is not only innovative but also inclusive, resilient, and fit for purpose.

5.6.1 Designing Inclusive AI-Enabled Public Service Delivery

AI-enabled public service delivery systems have increasingly relied on digital identity verification as a prerequisite for access. While this can streamline operations and reduce fraud, it also risks excluding individuals who lack formal identification or experience authentication failures due to technical, infrastructural, cultural or socio-economic barriers.

The World Bank’s ID4D initiative reports that over 850 million people globally lack any form of official ID, disproportionately affecting women, rural populations, and displaced persons (World Bank, 2023). In AI-driven systems, this exclusion is compounded when identity verification is automated and rigid, leaving no room for alternative pathways to access essential services.

To ensure that essential benefits are not denied to those most in need, systems should be designed with the following safeguards:

- Tiered Access Models
 - Implement progressive onboarding that allows users to access basic services with minimal verification.
 - Example: India’s Aadhaar-linked DBT system uses offline verification and assisted modes for rural populations (Economic Times, 2025).
- Fallback Mechanisms
 - Introduce human-in-the-loop verification for cases where biometric or digital authentication fails.
 - Use community attestations, local government records, or mobile-based identity proxies as interim solutions.
- Inclusive Identity Design
 - Support multi-modal authentication (e.g., PINs, QR codes, voice recognition).
 - Ensure systems are language-localised and accessible via low-tech devices.

- Regulatory Safeguards
 - Mandate non-discrimination clauses in digital ID systems.
 - Adopt frameworks like the UNDP DPI Safeguards (DPI Safeguards Initiative, 2024) and OECD’s Recommendation on Digital Identity Governance (OECD, 2023).
- Monitoring and Redress:
 - Establish real-time grievance redressal systems and audit trails.
 - Publish disaggregated access data to monitor equity in service delivery.

5.6.2 Safeguarding Privacy in AI-Powered Accessibility Tools

AI-powered accessibility tools such as sign-language translation, emotion-aware interfaces, and gesture recognition are enhancing and transforming public service delivery for persons with disabilities. Despite these developments, the same tools often require continuous video capture in public or semi-public spaces, raising significant privacy concerns for both intended users and bystanders (AAAS, 2025).

Privacy Risks

- Incidental Data Collection: Individuals not actively engaging with the system may be recorded, leading to unintended biometric data capture.
- Facial Recognition and Surveillance: Continuous video feeds may be repurposed for surveillance or profiling.
- Consent and Transparency: Users and bystanders may be unaware of data collection, processing, or retention practices.

Design and Governance Solutions:

- Privacy-by-Design Architecture
 - Implement sensitive data localisation and avoid cloud transmission.
 - Use anonymisation techniques during model training and inference.
 - Apply data minimisation principles—only collect what is strictly necessary.
- Contextual Consent Mechanisms
 - Deploy visual indicators (e.g., signage, light signals) to inform users and bystanders of active data capture.
 - Persons could be given the option to opt-out or be given privacy shields in public spaces.
- Regulatory and Ethical Safeguards
 - Align with frameworks such as the EU AI Act (European Commission, 2025), OECD Privacy Guidelines, and UNDP DPI Safeguards.
 - Require impact assessments and community consultations before deployment.
- Inclusive Governance
 - Engage disability rights groups and affected communities in the design and oversight of accessibility systems.
 - Ensure that accessibility does not come at the cost of dignity, autonomy, or surveillance.

5.7 Implications for DPI Security

The CVE/CWE analysis reveals that identity systems are not only vulnerable in isolation but also act as amplifiers of risk across DPI. A single CVE in an identity provider can cascade into financial fraud, healthcare data breaches, and disruption of government services.

To mitigate these risks, DPI architects must:

- Prioritise CWE-informed secure design practices.
- Conduct regular CVE monitoring and patch management related to AI-generated media and biometric spoofing.
- Implement threat modelling frameworks that account for cross-sector propagation.
- Adopt secure-by-design principles for biometric and federated identity systems.

6. International Standards and Policy Alignment

Securing digital identity systems against generative AI-powered credential fraud go beyond technical controls. It requires alignment with managerial, operational, and socio-cultural considerations which are in line with international standards, ethical frameworks, and inclusive governance models. This section outlines key global instruments that provide guidance for designing secure, rights-respecting Digital Public Infrastructure (DPI).

6.1 Foundational Standards for Identity Security

Several international standards offer robust frameworks for securing identity systems:

- NIST SP 800-63 (U.S.): Provides guidelines for digital identity proofing, authentication, and federation. It emphasises multi-factor authentication, risk-based assurance levels, and biometric anti-spoofing (NIST, 2025).
- ISO/IEC 24745: Focuses on biometric information protection, including template security, encryption, and unlinkability (ISO, 2022). It is widely adopted in biometric systems across Europe, Asia, and Africa to meet confidentiality, integrity and privacy in biometric systems requirements.
- GDPR Article 9 (EU): Regulates the processing of special categories of personal data, including biometric identifiers. It mandates explicit consent, data minimisation, and purpose limitation (GDPR, 2016/679).
- eIDAS 2.0: The updated EU regulation on electronic identification and trust services introduces stronger anti-spoofing requirements and interoperability standards for cross-border identity systems (European Commission, 2024).

These standards are increasingly referenced in national cybersecurity strategies, including the UK Government Cyber Security Strategy and India's National Digital Communications Policy.

6.2 DPI Safeguards and Ethical Governance

The DPI Safeguards Initiative (2024) provides a holistic approach to DPI governance, emphasising inclusion, transparency, and accountability. It outlines six core principles:

1. Do no harm
2. Do not exclude
3. Ensure effective redress

4. Promote autonomy and agency
5. Uphold the rule of law
6. Reinforce transparency and accountability

These principles among others are particularly relevant in LMICs, where DPI systems are often deployed rapidly and at scale. For example, in Ethiopia’s digital ID pilot, safeguards were embedded to ensure that users could access services through assisted onboarding and community attestations (UNDP, 2025).

The OECD’s article, Recommendation on Digital Identity Governance, complements these safeguards by promoting ethical use of AI, privacy-by-design, and inclusive identity verification (OECD, 2023). It encourages member states to adopt interoperable standards and conduct impact assessments before deploying biometric systems.

6.3 Regulatory Convergence and Global Cooperation

Fragmented regulatory oversight can undermine DPI effectiveness. The Cambridge Centre for Alternative Finance (CCAF) recommends establishing cross-regulatory coordination mechanisms, such as national DPI councils or inter-agency forums, to ensure coherence and accountability (Cambridge Centre for Alternative Finance, 2025).

In the Global South, regional cooperation is emerging. The African Union’s Digital ID Framework promotes harmonised standards across member states (African Union, 2023), while ASEAN countries are exploring shared biometric verification protocols for cross-border financial services (Mahusin & Prilliadi, 2025).

International development partners including the World Bank, and UNDP, are supporting DPI deployments with technical assistance, funding, and policy guidance. These partnerships are essential for building cyber resilience and ensuring that identity systems are secure, inclusive, and interoperable.

7. Deep Dive: The Digital ID Safety Pack (Minimum Security Baseline)

The evolving threat landscape of synthetic identities and deepfake-enabled credential fraud require multi-layered, cross-sectoral approaches and strategies. DPI systems require a minimum-security baseline, a “Digital ID Safety Pack” that integrates technical safeguards, governance principles, and operational best practices. This section therefore outlines the strategic recommendations for mitigating the risks in accordance with cybersecurity and governance frameworks. These controls when embedded across the identity lifecycle, from system design to deployment and monitoring enhance the resilience of Digital Public Infrastructure (DPI), their success and assurance at both architectural and operational levels and generally enhance the resilience, transparency and trustworthiness of such identity systems and the DPI in general.

7.1 Technical Controls

Technical controls serve as an effective safeguard in mitigating emerging threats in digital identity systems particularly those amplified by generative AI and synthetic media. The following measures are recommended to enhance system resilience and trustworthiness:

- **Multi-Modal Biometric Verification** - Implement biometric systems that combine multiple modalities such as facial recognition, voice authentication, and behavioural

biometrics to reduce the risks of spoofing. Multi-modal systems are significantly more resistant to deepfake attacks than single-modality systems.

- **Liveness Detection and Anti-Spoofing Algorithms** - Integrate liveness detection into biometric workflows to distinguish between real users and synthetic media. Techniques include challenge-response prompts, infrared imaging, and micro-expression analysis. Face morph detection algorithms should also be deployed.
- **Secure API Design and Rate Limiting** - Identity APIs should enforce strict validation, authentication checks, and rate limiting to prevent enumeration and abuse especially in remote onboarding systems.
- **Encryption of Biometric Templates** - Biometric data must be encrypted both in transit and at rest. Storage should be isolated from personally identifiable information (PII) to prevent correlation attacks.
- **AI-Based Deepfake Detection** - Deploy machine learning models to detect anomalies in identity verification workflows, such as synthetic media artifacts or unusual behavioural patterns.

7.2 Governance and Policy Measures

Digital identity systems are supposed to be not only technically robust but also ethically grounded and legally compliant. This requires governance and policy frameworks to be embedded into Digital Identity Systems design and operation. To achieve these, the following measures are recommended:

- **DPI Safeguards Integration** - Adopt the foundational DPI principles from the Universal DPI Safeguards Framework: Do no harm, do not exclude, ensure effective redress, promote autonomy and agency, Uphold Rule of Law, and Reinforce transparency and accountability. These principles should guide the design, deployment, and operation of identity systems.
- **Regulatory Alignment and Standards Adoption** - Governments should enforce standards such as:
 - NIST SP 800-63 (Digital Identity Guidelines)
 - ISO/IEC 24745 (Biometric Information Protection)
 - GDPR Article 9 (Special Categories of Personal Data)
 - eIDAS 2.0 (EU Digital Identity Regulation)
- **Cyber Assessment Framework (CAF) Adoption** - As recommended in the UK strategy, DPI operators should adopt the CAF to ensure consistent and proportionate cyber security assurance. Tiered CAF profiles should be used to align security outcomes with threat levels and system criticality.
- **Mandatory Disclosure and Incident Reporting** - Establish mandatory vulnerability disclosure policies (VDPs) and breach notification frameworks for identity providers and DPI operators.

- **Synthetic Identity Simulation Exercises** - Conduct regular red-teaming and synthetic identity simulation exercises to test system resilience against deepfake and spoofing attacks.

7.3 Cross-Sector Collaboration

Addressing synthetic identity threats and other biometric vulnerabilities effectively requires action and coordination across sectors. Data and skills sharing help stakeholder to collaborate and enhance the resilience of digital identity systems. The following collaborative measures are recommended for building resilient digital identity ecosystems:

- **Threat Intelligence Sharing** - Create national and regional platforms for sharing indicators of compromise (IOCs), attack patterns, and mitigation strategies related to synthetic identity fraud.
- **Public-Private Partnerships** - Encourage collaboration between governments, academia, and industry to co-develop secure identity technologies and conduct joint research on biometric spoofing.
- **Cyber Workforce Development** - Invest in cyber security skills and leadership across DPI stakeholders. Establish learning academies, career pathways, and accreditation frameworks to build a sustainable cyber workforce.

7.4 Future-Proofing Identity Systems

Digital identity systems must be designed to withstand current and emerging forms of threats. This means, such systems would require proactive investment in resilient architectures and forward-looking standards especially when the threats relate to cryptography, Artificial Intelligence or other technological shifts. The following are measures recommended for future-proofing digital system:

- **AI Use and Governance Programme** - Establish a structured framework to manage AI models used in identity systems, ensuring fairness, robustness, and accountability across their lifecycle, in line with NIST AI RMF (2023).
- **Post-Quantum Cryptography Readiness** - Prepare identity systems for future cryptographic threats by exploring post-quantum algorithms for biometric template protection and digital signatures (NIST, 2023; ETSI, 2024).
- **Secure-by-Design DPI Architecture** - Embed security principles into every layer of identity infrastructure from biometric sensors to federated login APIs ensuring resilience against evolving threats in line with recommendation of the UK Cyber Assessment Framework (NCSC, 2025).

8. Implementation Pathways for Countries (Quick, Medium, and Long-Term)

The implementation strategies to protect Digital Public Infrastructure (DPI) against generative AI-enabled credential fraud require distinct measures. Such measures are generally tailored to each country's technological capacity, regulatory environment, and risk exposure. This section outlines practical implementation strategies for governments and DPI institutions to

strengthen their identity systems across three horizons: quick or short-term, medium-term, and long-term.

8.1 QShort-Term (0–6 Months)

These actions focus on immediate risk mitigation and foundational security upgrades:

- **Deploy Liveness Detection in Biometric Systems:** Integrate basic liveness detection into facial recognition workflows to prevent deepfake spoofing during onboarding and authentication.
- **Establish Threat Intelligence Sharing Platforms:** Create national or regional platforms for sharing indicators of compromise (IOCs), attack vectors, and mitigation strategies. LMICs can leverage existing networks such as the African Union’s Cybersecurity Collaboration Hub for threat intelligence sharing (African Union, 2024).
- **Conduct Synthetic Identity Simulation Exercises:** Run red-teaming exercises to test system resilience against deepfake-enabled identities, morphing attacks, and credential spoofing (OECD, 2023).
- **Enforce Secure API Practices:** Apply rate limiting, input validation, and authentication checks to identity APIs, especially in mobile and federated environments.
- **Launch Public Awareness Campaigns:** Educate citizens and service providers about the risks of AI-driven identity fraud and the importance of secure credential handling.

8.2 Medium-Term (6–18 Months)

These measures aim to institutionalise security and governance practices:

- **Align with International Standards:** Adopt frameworks such as NIST SP 800-63, ISO/IEC 24745, GDPR Article 9, and eIDAS 2.0 to guide identity system design and operation.
- **Implement the DPI Safeguards Framework:** Embed principles like “Do no harm,” “Do not exclude,” and “Ensure effective redress” into procurement, system architecture, and user engagement.
- **Adopt the Cyber Assessment Framework (CAF):** Use tiered CAF profiles to assess and improve cybersecurity maturity across DPI components.
- **Develop Inclusive Identity Design:** Introduce multi-modal authentication (e.g., PINs, QR codes, voice recognition) and fallback mechanisms for users facing biometric or digital barriers in line with World Bank recommendations (World Bank, 2023).
- **Build Cybersecurity Workforce Capacity:** Establish training programmes, career pathways, and accreditation schemes to develop a skilled cyber workforce across government and industry.

8.3 Long-Term (18+ Months)

These strategies focus on future-proofing DPI systems and embedding resilience:

- **Invest in Post-Quantum Cryptography Readiness:** Prepare for emerging cryptographic threats by exploring quantum-resistant algorithms for biometric template protection and digital signatures.
- **Design Secure-by-Default DPI Architectures:** Embed security principles into every layer of identity infrastructure — from biometric sensors to federated login APIs — ensuring resilience against evolving threats.

- **Adopt an AI Use and Governance Programme:** Establish a structured framework to manage AI models used in identity systems, ensuring fairness, robustness, and accountability across their lifecycle.
- **Localise Biometric Data and Enforce Encryption-at-Rest:** Store biometric data within national jurisdictions and apply strong encryption to prevent unauthorised access and cross-border misuse.
- **Establish National DPI Councils:** Create cross-sectoral governance bodies to coordinate policy, technical standards, and incident response across identity systems as recommended by the Cambridge Centre for Alternative Finance, (2025).
- **Integrate DPI with Broader Digital Ecosystems:** Ensure that identity systems are interoperable with digital payments, consent-based data sharing, and public service platforms to maximise inclusion and efficiency.

9. Conclusion: Protecting Trust in Digital Public Infrastructure

Synthetic identity threats and deepfake-enabled credential fraud represent a paradigm shift in the cybersecurity and governance landscape. These challenge the foundational assumptions of digital identity systems and Digital Public Infrastructure (DPI). The threats discussed above highlight that they are no longer speculative but are endemic and increasingly weaponised across sectors.

From the CVE and CWE data between 2020 to 2025, this report has documented a marked increase in vulnerabilities targeting identity verification systems, biometric authentication modules, and federated login architectures. Weaknesses such as CWE-287 (Improper Authentication), CWE-863 (Incorrect Authorisation), and CWE-200 (Information Exposure) are being actively exploited by threat actors using generative AI to create deepfake-based impersonation capable of bypassing traditional security controls.

The sectoral case studies presented also illustrate the real-world consequences of these threats: financial fraud, data breaches, service disruption, and erosion of public trust. In each case, attackers leveraged a combination of technical vulnerabilities and procedural oversights to compromise identity systems and exploit them at scale.

These findings reinforce the urgency of implementing robust detection, governance, and cross-sector collaboration. Synthetic identities and deepfake-enabled fraud are not merely technical challenges. They are systemic threats to digital trust, financial stability, health care and public service delivery.

To counter these threats, stakeholders must:

- Integrate biometric anti-spoofing, synthetic identity screening, and AI-driven anomaly detection into identity systems.
- Align with international standards such as NIST SP 800-63, ISO/IEC 24745, and eIDAS 2.0.
- Adopt Cyber Assessment Frameworks (CAF) that ensure consistent and proportionate cyber security assurance across DPI systems.
- Embed DPI safeguards principles—such as “Do no harm,” “Do not exclude,” and “Ensure effective redress” into every stage of the identity lifecycle.
- Foster public-private partnerships and threat intelligence sharing platforms.

- Prepare for future threats through post-quantum cryptography readiness and secure-by-design identity architectures.

As DPI continues to expand globally, particularly in low and middle-income countries, the risks associated with AI-driven identity fraud and deepfakes will only grow more complex and consequential. To meet these challenges, stakeholders must be more proactive than reactive in their security models and approaches. This means embedding security-by-design principles into every layer of identity infrastructure from biometric sensors to cloud-based identity providers. It also means fostering a culture of transparency, accountability, and continuous improvement across the public and private sectors.

Ultimately, the security of digital identity systems is not just a technical issue, but rather, a matter of national resilience, economic continuity, and democratic integrity.

References

AAAS. (2025). *Key considerations when using artificial intelligence in the public sector*. American Association for the Advancement of Science. <https://www.aaas.org/sites/default/files/2025-01/Key%20Considerations%20AI%20for%20Public%20Sector.pdf>

African Union. (2023). *AU interoperability framework for digital ID*. <https://au.int/en/documents/20231211/au-interoperability-framework-digital-id>

African Union. (2024). *AU Cybersecurity Collaboration Hub*. <https://au.int/en/cybersecurity-collaboration-hub>

Australian Information Security Association. (2024). *Australian Federal Government's 2024–25 Budget*. https://www.aisa.org.au/Public/Public/News_and_Media/News/2024/Australian-Federal-Government-s-2024-25-Budget.aspx

Bello, M. I., & Thomas, D. R. (2023). Curbing ghost worker fraud in developing countries using consortium blockchain. In *Proceedings of the 9th ACM International Workshop on Security and Privacy Analytics (IWSPA '23)* (pp. 1–7). https://pure.strath.ac.uk/ws/portalfiles/portal/161303395/Bello_Thomas_IWSPA_2023_Curbing_ghost_worker_fraud_in_developing_countries.pdf

Burt, N. (2024, October 14). How bank-based digital ID can neutralise the threat of deepfake fraud. *TechUK*. <https://www.techuk.org/resource/how-bank-based-digital-id-can-neutralise-the-threat-of-deepfake-fraud.html>

Cambridge Centre for Alternative Finance. (2025). *Digital Public Infrastructure and Digital Financial Services: Convergence, landscape and regulatory considerations*. <https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/publications/digital-public-infrastructure-and-digital-financial-services/>

Cisco. (2024). *Cisco Duo Authentication for Windows Logon and RDP Authentication Bypass Vulnerability*. <https://sec.cloudapps.cisco.com/security/center/content/CiscoSecurityAdvisory/cisco-sa-duo-win-bypass-pn42KKBm>

Deepstrike.io. (2025). *Deepfake statistics 2025: Trends, threats, and insights*. Retrieved October 8, 2025, from <https://deepstrike.io/blog/deepfake-statistics-2025>

DPI Safeguards Initiative. (2024). *Universal DPI Safeguards Framework*. <https://www.dpi-safeguards.org/framework>

Economic Times. (2025, June 12). RBI simplifies KYC rules to allow face-to-face, video and OTP-based onboarding for customers. *Economic Times*. <https://economictimes.indiatimes.com/news/economy/policy/rbi-know-your-customer-kyc-rules-customer-onboarding-aadhaar-biometric-norms/articleshow/121797850.cms>

ETSI. (2024). *Quantum-safe cryptography: Technical report*. European Telecommunications Standards Institute. https://www.etsi.org/deliver/etsi_tr/103600_103699/103619/01.01.01_60/tr_103619v010101p.pdf

European Commission. (2024). *eIDAS 2.0 compliance: Unlocking EU-wide interoperability, convenience, and security*. <https://community.infineon.com/t5/Blogs/Solutions-to-fulfil-the-main-eIDAS-2-0-requirements-of-an-EU-wide/ba-p/1029751>

European Commission. (2025, February 4). *Guidelines on prohibited artificial intelligence practices under the AI Act*. <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act>

Federal Reserve Bank of Boston. (2019). *Synthetic identity fraud in the U.S. payment system: A review of causes and contributing factors*. <https://fedpaymentsimprovement.org/wp-content/uploads/frs-synthetic-identity-payments-fraud-white-paper-july-2019.pdf>

Fintech News Singapore. (2025, June 12). APAC fintech faces 116% spike in fraud linked to deepfakes, synthetic IDs. *Fintech News Singapore*. <https://fintechnews.sg/112593/security/sumsub-fintech-fraud/>

FIRST.Org, Inc. (2023). *CVSS v4.0 specification document*. <https://www.first.org/cvss/specification-document>

GDPR. (2016/679). *Art. 9 GDPR – Processing of special categories of personal data*. <https://gdpr-info.eu/art-9-gdpr/>

HHS. (2024, December 27). HIPAA Security Rule Notice of Proposed Rulemaking to strengthen cybersecurity for electronic protected health information. U.S. Department of Health and Human Services. <https://www.hhs.gov/hipaa/for-professionals/security/hipaa-security-rule-nprm/factsheet/index.html>

ISO. (2022). *ISO/IEC 24745:2022 – Information security, cybersecurity and privacy protection – Biometric information protection*. <https://www.iso.org/standard/75302.html>

Jumio. (2020, April 9). Liveness detection a must for FCA recommended customer selfies. *Jumio*. <https://www.jumio.com/fca-selfies-liveness-detection/>

Lemos, R. (2025, September 19). Synthetic identities plague finance & lending sector. *Dark Reading*. <https://www.darkreading.com/cybersecurity-operations/synthetic-identities-finance-lending-sector>

LexisNexis Risk Solutions. (2023). *Uncovering the hidden threat: Synthetic identity fraud in the UK*. <https://risk.lexisnexis.co.uk/insights-resources/white-paper/synthetic-identity-fraud-in-the-uk>

Mahusin, M., & Prilliadi, H. (2025). *Integrating digital payments in ASEAN: Harmonising regulations and strengthening security for inclusive growth*. Economic Research Institute for ASEAN and East Asia. <https://www.eria.org/uploads/Integrating-Digital-Payments-in-ASEAN.pdf>

Microsoft. (2023, March 24). Guidance for investigating attacks using CVE-2023-23397. *Microsoft Security Blog*. <https://www.microsoft.com/en-us/security/blog/2023/03/24/guidance-for-investigating-attacks-using-cve-2023-23397/>

MITRE. (2025). *CWE-79: Improper Neutralization of Input During Web Page Generation (“Cross-site Scripting”)*. Retrieved October 1, 2025, from <https://cwe.mitre.org/data/definitions/79.html>

MITRE. (2025). *CWE-601: URL Redirection to Untrusted Site (“Open Redirect”)*. Retrieved October 1, 2025, from <https://cwe.mitre.org/data/definitions/601.html>

MITRE. (2025). *CWE-918: Server-Side Request Forgery (SSRF)*. Retrieved October 1, 2025, from <https://cwe.mitre.org/data/definitions/918.html>

MITRE. (2025). *CWE-352: Cross-Site Request Forgery (CSRF)*. Retrieved October 1, 2025, from <https://cwe.mitre.org/data/definitions/352.html>

MITRE. (2025). *CWE-20: Improper Input Validation*. Retrieved October 1, 2025, from <https://cwe.mitre.org/data/definitions/20.html>

MITRE. (2025). *CWE-22: Improper Limitation of a Pathname to a Restricted Directory (“Path Traversal”)*. Retrieved October 1, 2025, from <https://cwe.mitre.org/data/definitions/22.html>

MITRE. (2025). *CWE-611: Improper Restriction of XML External Entity Reference*. Retrieved October 1, 2025, from <https://cwe.mitre.org/data/definitions/611.html>

National Cyber Security Centre. (2025). *Cyber Assessment Framework (CAF)*. <https://www.ncsc.gov.uk/collection/cyber-assessment-framework>

National Vulnerability Database (NVD). (2024). *CWE over time*. <https://nvd.nist.gov/general/visualizations/vulnerability-visualizations/cwe-over-time>

National Vulnerability Database (NVD). (2024). *Vulnerability metrics*. <https://nvd.nist.gov/vuln-metrics/cvss>

National Vulnerability Database (NVD). (2025). *NVD CWE Slice*. <https://nvd.nist.gov/vuln/categories>

News Digest. (2025, September 16). *NIN-SIM linkage cuts fraud, strengthens Nigeria's national security says Pres. Tinubu*. *News Digest*. <https://newsdigest.ng/nin-sim-linkage-cuts-fraud-strengthens-nigerias-national-security-says-pres-tinubu/>

NIST. (2023). *Getting ready for post-quantum cryptography*. <https://csrc.nist.gov/publications/detail/white-paper/2023/07/05/getting-ready-for-post-quantum-cryptography/final>

NIST. (2025). *NIST Special Publication 800-63 Digital Identity Guidelines*. <https://www.nist.gov/identity-access-management/projects/nist-special-publication-800-63-digital-identity-guidelines>

NIST AI RMF (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf> accessed October 15, 2025

OECD. (2023). *Recommendation of the Council on the Governance of Digital Identity*. <https://digitalgovernmenthub.org/library/recommendation-of-the-council-on-the-governance-of-digital-identity/>

OWASP Foundation. (2024). *Zero Trust Architecture Cheat Sheet*. https://cheatsheetseries.owasp.org/cheatsheets/Zero_Trust_Architecture_Cheat_Sheet.html

Paravision. (2024). *Deepfake threats and attack vectors in remote identity verification*. <https://www.paravision.ai/news/deepfake-threats-and-attack-vectors-in-remote-identity-verification/>

Phandle, G. (2024, August 2). *Corrupt home affairs official sentenced to 35 years for passport scheme*. *SowetanLIVE*. <https://www.sowetanlive.co.za/news/south-africa/2024-08-02-corrupt-home-affairs-official-sentenced-to-35-years-for-passport-scheme/>

Protean eGov Technologies. (2025). *How UIDAI and Protean protect Aadhaar KYC from deepfake fraud*. <https://proteantech.in/articles/secure-aadhaar-kyc-02-06-2025/>

Resecurity. (2024, May 6). *Massive dump of hacked Salvadorean headshots and PII highlights growing threat-actor interest in biometric data*. *Resecurity*. <https://www.resecurity.com/blog/article/massive-dump-of-hacked-salvadorean-headshots-and-pii-highlights-growing-threat-actor-interest-in-biometric-data>

Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture (NIST SP 800-207)*. National Institute of Standards and Technology. <https://www.nist.gov/publications/zero-trust-architecture>

Sumsub. (2023, June 9). *New digital fraud statistics in the UK and continental Europe: Forced verification and deepfake cases multiply at alarming rates*. *Sumsub*. <https://sumsub.com/newsroom/new-digital-fraud-statistics-in-the-uk-and-continental-europe-forced-verification-and-deepfake-cases-multiply-at-alarming-rates/>

Techish Kenya. (2025, January 31). *Identity fraud in Africa reaches crisis levels: How AI is fueling the problem*. *Techish Kenya*. <https://tech-ish.com/2025/01/31/generative-ai-identity-fraud-africa/>

TechRadar. (2025, February 10). *Top US health provider tells 882,000 patients they were hit in August 2023 breach*. *TechRadar*. <https://www.techradar.com/pro/security/top-us-health-provider-tells-882-000-patients-they-were-hit-in-august-2023-breach>

The Alan Turing Institute. (2025). *Cyber Threat Observatory for National Identity Systems Quarterly Report*. https://www.turing.ac.uk/sites/default/files/2025-05/threat_report_cve_april-4.pdf

The Guardian. (2025, September 9). *AI-assisted image morphing, finger blending among 6,000 exam fraud cases — JAMB Panel*. *The Guardian*. <https://guardian.ng/news/nigeria/national/ai-assisted-image-morphing-finger-blending-among-6000-exam-fraud-cases-jamb-panel/>

TransUnion. (2025). *2025 identity fraud trends impacting government agencies*. <https://www.transunion.com/report/fraud-trends-impacting-government-agencies>

UK Cabinet Office. (2022). *Government cyber security strategy: 2022 to 2030*. <https://www.gov.uk/government/publications/government-cyber-security-strategy-2022-to-2030>

UK Finance. (2024). *Annual fraud report 2024*. <https://www.ukfinance.org.uk/policy-and-guidance/reports-and-publications/annual-fraud-report-2024>

United Nations Development Programme. (2025). *Showcasing Global Progress Towards Building Safe and Inclusive DPI for Societies*. UNDP. Retrieved October 10, 2025, from <https://www.dpi-safeguards.org/video-dpi/ethiopia-digital-ids-enable-essential-services-for-refugees>

Vyshegorodtsev, K., Kudiyarov, D., Balashov, A., & Kuzmin, A. (2024). Deepfake detection in videos with multiple faces using geometric-fakeness features. *arXiv*. <https://arxiv.org/html/2410.07888v1>

World Bank. (2023). *ID4D global dataset*. Identification for Development (ID4D). <https://id4d.worldbank.org/global-dataset>