The Digital Identity Risk Universe - 2026

A Kantara Initiative Discussion Group Report

Version: 1.0

Document Date: 2026-03-23

Editors: John Fiske

Contributors: David Kelts, Jay Meier, Amit Sharma, Wendy Brown, Mohamed El Baih, Junmin (Eric) Kim, Joseph Payne, Samuel Steg, Cynetheia Brown, Shermaine Nedd, Anthony Moran

Produced by: Digital Identity Risk Universe Discussion Group (DIRU DG)

Status:

This document is an initial draft produced by the DIRU DG for approval by the Group. See the Kantara Initiative Operating Procedures for more information.

Abstract:

This document describes the motivation, methodology, output, and insights from the Digital Identity Risk Universe (DIRU) Discussion Group (DG). The Risk Universe itself is in a separate file (here) and is available as open source information. The DIRU is a compendium of major risks associated with digital identities, organized into five primary domains. This framework is intended to enable a dynamic ontology of digital identity-related risks, and to assist organizations in understanding and addressing diverse priorities and incentives (e.g. know your customer workflows, anti-fraud controls, cyber and IT security impacts, enterprise-wide risk, legal and regulatory compliance, etc.), whose equities are commonly impacted by identity risks.

We have restricted ourselves to three levels of risk identification, but expect that stakeholders interested in specific risk areas can and should develop more detailed risks.

Regarding the risks associated with the identification of AI agents, the DG decided to create two distinct (but overlapping) risk universes. AI Agents pose many novel risks so it would be confusing and counterproductive to combine them with the risk universe for human digital identities, though we do identify risks that are shared. We also note that this field is evolving very quickly, and as such, the risk universe should be regularly reviewed and amended to reflect ongoing technological or operational innovations. The risks highlighted in this initial framework are as of February 2026.

IPR Option:

This document is subject to the Kantara Initiative IPR Policy option Non-Assertion Covenant

Suggested Citation:

The Digital Identity Risk Universe - 2026 Kantara Initiative Discussion Group.

NOTICE

DEAR READER

Thank you for downloading this publication prepared by the international community of experts that comprise Kantara Initiative. Kantara is a global non-profit 'commons' dedicated to improving the trustworthy use of digital identity and personal data through innovation, standardization, and good practice.

Kantara is known around the world for incubating innovative concepts, operating Trust Frameworks to assure digital identity & privacy service providers, and developing community-led best practices and specifications. Its efforts are acknowledged by OECD ITAC, UNCITRAL, ISO SC27, other consortia, and governments around the world. 'Join, Innovate, Trust' captures the rhythm of Kantara in consolidating an inclusive, equitable digital economy offering value and benefit to all.

Every publication, in every domain, is capable of improvement. Kantara welcomes and values your contribution through membership, sponsorship, active participation in the Work Group that produced this, and participation in all our endeavors so that Kantara can reflect its value to you and your organization.

Table of Contents (Insert->Page Elements->Table of Contents->Dotted)

(finalized in Word prior to publication)**

1. Motivation

Digital identity is a cornerstone of online trust. Globally, billions of people now use some form of verifiable digital credential to support online banking, government services, and other digital services. Simultaneously, AI agents are exploding in usage, becoming ever more autonomous and capable, and quantum computing is fast becoming a reality. These trends create a complex set of digital identity-related risks including traditional risks around issuance, as well as new risks surrounding biometric binding, emerging wallet technologies, certificate management, ghost credentials, security, interoperability, cross-ecosystem usage, and ownership of AI decision-making to name a few.

Left unmanaged, these risks can lead to a wide range of harms including fraud, identity theft, misrepresentation, mass surveillance, social control and non-compliance to name a few. These risks are also shared across an ecosystem of identity stakeholders so managing them effectively will require organizations – both public and private – to coordinate their approach across each 'risk domain'. Mitigations must be implemented to protect individuals, businesses and governments as well as the entities they transact with – employees, clients, counterparties and others. Yet while identity-related risks may be a 'common throughline' across many facets of an organization, sadly the solutions developed are rarely holistic in nature given competing incentive structures and priorities. For example, anti-fraud teams striving to reduce identity compromise may have very different incentives than the know your customer (KYC) compliance teams who are focused on frictionless onboarding. We hope that organizations that recognize the core set of identity-related risks they face will approach the problem more holistically and create end to end solutions to address them. The purpose of creating this publicly available Digital Identity Risk Universe (DIRU) is to help businesses, regulators and policy makers better understand digital identity risks and manage them using common vocabulary and framing.

What is a Risk Universe?

A ‘risk universe’ is simply the superset of risks concerning a topic. Establishing a risk universe is typically the first step organizations or policy makers undertake when determining how to manage risk around a given topic. Many of the risks in a Risk Universe may not be relevant to a particular organization or function but it is nonetheless important to consider a full set of risks before narrowing scope to ensure one is considering all possible contingencies.

Once the broader risk universe is evaluated, organizations or teams will likely choose to manage a smaller set of risks (commonly called a risk register) that are relevant to their business or policy area. As an example of a comprehensive risk universe in the Artificial Intelligence space, please see MIT’s AI Risk Repository (https://airisk.mit.edu/).

This risk universe will hopefully serve as a starting point for organizations to begin their own risk assessment and management processes. As these risks constantly evolve, and technology and policies change, this Risk Universe represents major risk areas as of Q1 2026.

We welcome feedback to expand, deepen and maintain this risk universe as the space evolves. This authors of this paper can be reached by email at: dg-riskuniverse@kantarainitiative.org

2. Methodology and Scope

Over the course of assembling this risk universe, it became clear that the risks of Agentic AI Identity were fairly distinct from that of Verifiable Digital Credentials. The Discussion Group (DG) therefore decided to take different approaches to framing the two risk universes, though the methodology and general taxonomy were similar.

Methodology and Scope

Methodology

The DG was composed of experts from a number of specialties in digital identity. Over the course of early 2026, the DG developed the DIRU per the following method:

- Chair conducted preliminary research and proposed major risks
- DG reviewed these risks, and provided feedback and recommendations each week
- Chair adjusted risk universe to reflect consensus of DG
- DG provided final review and editorial input
- DG voted the project through Kantara approval process

Scope

The purpose of this DG is simply to define the set of potential risks in the risk universe. We do not seek to conduct any assessment of inherent risk, possible mitigations, or residual risk. Nor do we suggest any scoring of these risks. Scoring assessments are specific to each organization or stakeholder group and should be conducted per use case and in accordance with that organization's internal methodology.

Definitions and Taxonomy

As a working definition for 'Digital identity', we used NIST's definition of 'verifiable digital credentials' (VDC), namely all-digital credentials which are cryptographically authenticated between the Issuer, the Holder and the Verifier (regardless of the technology for verification). Most human-centered systems will leverage wallet technology and will enable some level of selective disclosure.

This risk universe is organized into five risk domains, based on the lifecycle/value chain of 'verifiable digital credentials'. The first three domains align roughly to the risks associated with action taken by an Issuer, Holder or Verifier, respectively. The last two domains concern end-to-end ecosystem risks, firstly risks associated with inadequate Governance, Accountability and Oversight, and secondly those focused on Systemic risks.[1]

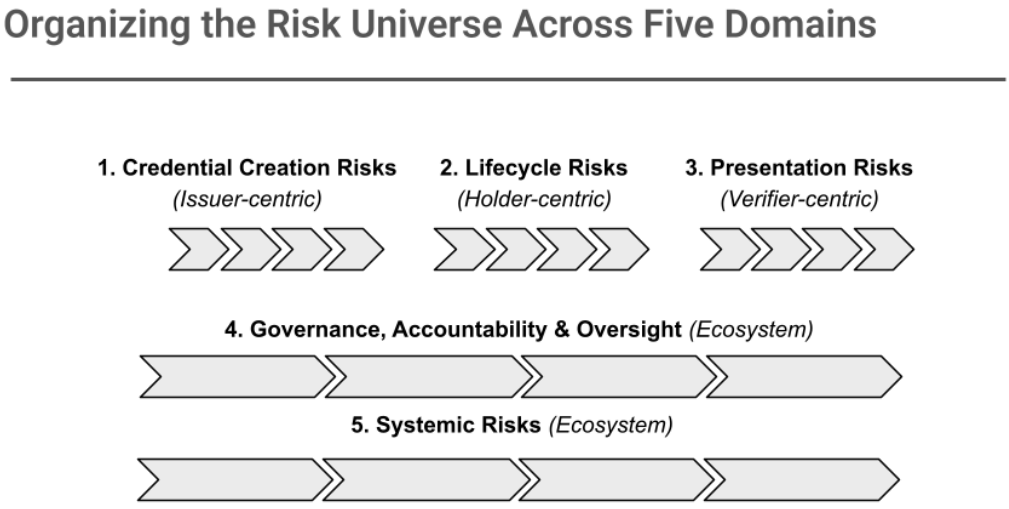

Figure 1. Overview of risk domain structure

[1] *It is important, however, for any user of this risk universe to consider risks from each of the five domains and not just focus on one. Issuers, for example, may still face risks when a credential is presented even if they are not a party to the transaction itself.*

In addition to identifying risks, this DG also flagged whether or not each of three risk themes was relevant to the risk:

1. Trustworthiness: Does this risk potentially impact trust in a credential?
2. Privacy: Does this risk potentially impact either an individual's personal privacy or an organization's compliance with privacy laws?
3. Agentic AI: Is this risk relevant to AI Agents?

For this Digital Identity Risk Universe, the DG adopted a four-tier risk taxonomy:

Level	***Description***
Domain	High level risk domains (grouping of major risk categories)
L1 Risks	Major risk categories
L2 Risks	Risk subcategories, and/or significant risks
L3 Risks	Risks and subcategories

Agentic Identity Risks

The DG used the same five-domain taxonomy to assess Agentic Identity Risks, but because we expect that AI Agents in practice will require interactions with multiple stakeholders, including layers of agents and subagents belonging to their stakeholders, the DG:

1. Added a new dimension of six 'stakeholders' to capture which identity risks are relevant to each group.
2. Increased emphasis on identifying and logging the ownership of decisions made as well as tracing authorization to access data and make those decisions

To be explicit: we are not attempting to define the broader set of risks surrounding AI agents (enterprise security, personal manipulation, dependencies, economic disruption etc) and are only considered identity-related risks, but given that authorization and ownership claims will likely be managed by identity credentials, they are in scope for the Agentic Identity risk universe.

Taxonomy

For Agentic Identity risks, we use the following stakeholder groups in our risk definitions:

Stakeholder	Description
Operators	Companies which deploy agents into their services or operations. They set policies, technical controls, business logic, and user experience.
Orchestrators	Companies which develop use-case specific agents and communication frameworks. They provide memory, planning and tool-use logic. They may be Model Providers, or may OEM models
Model Providers	Companies which build the underlying foundation models (LLMs, etc). They offer "intelligence" through APIs.
Beneficiary ('Principals')	End users or entities whose goals the agent aims to achieve. They provide the "prompt" or "mission" along with their personal background for agent decision-making.
End User	Entity (often, but not necessarily a human) who uses the agent; sometimes the same as the beneficiary
Counterparties	External systems, APIs, businesses or human parties with whom the agent interacts to complete a task.

3. Output

The 2026 Digital Identity Risk Universe is found in a spreadsheet here.

There are five worksheets for your consideration:

1. The ‘Digital Identity Risk Universe’ is the set of L1 to L3 risks that the DG associated with verifiable digital credentials and digital identities.
2. The ‘Agentic Identity Risk Universe’ is the set of L1 to L3 risks that the DG associated with AI agents. Note that these risks are mapped to the six primary stakeholders in columns D through I.
3. ‘Agentic AI Definitions / Use Cases’ includes more detailed descriptions of the stakeholders (and examples) used to define risks
4. ‘Taxonomy’ described the taxonomy used in the risk universe
5. ‘FAQs’ include answers to Frequently Asked Questions

4. Insights

The assembly of a risk universe is an exercise in compilation more than analysis. We are not evaluating policies or technologies, nor comparing national or corporate approaches to these risks. However, several larger themes emerged during the DG meeting series. In no particular order:

1. Risks associated with verifiable digital credentials primarily overlap with traditional physical ID risks (e.g. passports, drivers licenses, etc) in the issuance phase, where enrollment, identity proofing and revocation risks are similar for both physical and digital credentials. A major differentiator for digital credentials, however, are the risks introduced in the registration bridge from physical/analog to digital.
2. Issuers face similar risks in Governance and Oversight regardless which type of credential they issue (digital or physical). But once issued, the lifecycle risks of digital credentials and physical credentials are quite different.
3. Agentic Identity risks overlap with verifiable digital credential risk in the lifecycle and presentation phases (particularly around security risks). But AI Agent identities are fundamentally different from human-centered identities including
    - Agents can be instantiated rapidly, with huge numbers of agents (and their linked identities) potentially being spawned in short timeframes
    - An ephemeral lifecycle (sometimes measured in seconds) creates a new class of risks since tasks, permissions and authorizations must be created and revoked at a speed and scale that enterprise systems may be challenged to support
    - Agents themselves can be unidentified, forked, or duplicated which poses a range of questions about establishing accountability.
    - Agents may be subject to changes to underlying models, memory structures, tools, access rights, etc which may dynamically change the identity or trust levels of the agent

- Agents are quickly developing the ability to present as human, which creates a range of novel misrepresentation and trust risks for both humans and other agents in the loop
- Agents with enough autonomy may someday be able to defeat corporate safeguards to limit access rights, authorizations and other privileges based on identity or even perform social engineering
- Lack of industry standards for authentication and authorization of agents mean that identity claims between agents, humans, businesses, and other counterparties may not be secure or authentic, and may rely on blind or reputational trust.
- Numerous potential cyber risks and privacy risks are derived from agentic identities

We hope you find the DIRU useful, and would appreciate any feedback (positive or negative). The DG can by contacted via the Kantara LinkedIn page, or email to dg-riskuniverse@kantarainitiative.org

Thank you!