Identity Salon June 2026 — Final Report

Topic: Identity Systems Under Pressure: What AI Vendors and Standards Efforts Reveal About How Security Decisions Break Down

Table of Contents

Executive Summary

The June 2026 Identity Salon examined a problem that is becoming increasingly visible across identity, security, and AI systems: organizations have many of the technical components needed to support complex access, trust, and risk decisions, but those components do not necessarily produce a coherent decision system when connected.

The first discussion tested a functional model introduced in the pre-reading:

Identity + Action Context + Signals → Policy → Decision → Runtime Enforcement → Execution → Audit

The model describes how an outcome may be assembled across identity providers, risk systems, policy engines, gateways, applications, infrastructure controls, automation platforms, and audit systems. Each component may work correctly within its own scope, while the complete decision remains difficult to understand, predict, or reconstruct.

The second discussion tested a proposed agentic trust architecture built largely from mechanisms that already exist. The proposal combined workload identity, an agent registry, a gateway, a trust broker, authorization and relationship data, transaction-specific credentials, Shared Signals, and an observability layer. Rather than beginning with a new agent identity protocol, it asked how current components might be assembled into a usable trust fabric.

To make the implications concrete, consider a simple scenario used throughout this report:

> *An employee asks an AI travel agent to book a particular hotel within a stated budget. The agent identifies itself, discovers the relevant booking service, obtains limited authority for the transaction, and completes the booking.*

Even this apparently straightforward task raises several questions.

Can the service distinguish the employee from the agent? Does the agent receive all of the employee's travel-system access or only what this booking requires? How is the instruction to book a particular hotel represented? What happens if the agent selects a different hotel while remaining within budget? Which system decides whether that action is permitted? What evidence remains afterward? The Salon discussion showed that these questions cannot be reduced to "AI is non-deterministic" or "we need better authorization."

Machine-learning systems have contributed probabilistic risk and fraud assessments for years. Core policy and enforcement controls remain deliberately deterministic. The more useful question is where uncertainty enters the decision path: signal generation, interpretation of context, task planning, tool selection, execution, or downstream interpretation.

Authorization also has limits. It can constrain which APIs, resources, operations, or data an agent may access. It cannot always determine whether the agent interpreted the task correctly. An agent may be technically authorized to book travel while still violating the instruction to book one particular hotel.

Several conclusions followed:

- Probabilistic analysis should be separated from deterministic authorization and enforcement.
- An agent should receive its own restricted authority rather than impersonating a user.
- The initiating principal, acting workload, delegated task, and resulting action should remain distinguishable.
- Hard controls are more credible when enforced around an AI model rather than expressed only within its instructions.
- Trust-domain boundaries are important points for re-evaluating identity, authority, and policy.
- Registries need to connect agent instances to governed classes, ownership, lifecycle, mission, and provenance.

- Observability must link decisions to actions without requiring organizations to retain every prompt or internal model event.
- Concrete reference architectures are valuable because they show where context, authority, and assumptions are lost.

The proposed trust architecture provided a plausible starting point. Participants largely tested it through clarification rather than rejecting its foundations. At the same time, the discussion showed that the problem is not merely one of assembling existing components. The architecture addresses lifecycle and transaction evidence at a conceptual level, but important questions remain about how those mechanisms should work in practice, alongside broader gaps in semantics, delegation, signal interpretation, and cross-domain authority.

The industry has many of the pieces, and participants broadly agreed that an acting workload needs an identity distinct from the person or organization behind it. It does not yet have sufficient agreement about which identity properties, relationships, context, authority, and decision evidence must remain meaningful as a transaction moves between systems and trust domains.

How to Read This Report

This report is a companion to two documents distributed prior to the Las Vegas Salon:

- *Identity Systems Under Pressure*, the pre-reading that introduced the functional decision model; and
- *You Already Have the Pieces. Now Build It.*, which proposed an agentic trust architecture assembled from existing identity and security mechanisms.

It does not reproduce either document—we recommend you read all three in conjunction—and instead focuses on what the discussion added:

- where participants challenged or refined the original arguments;
- what the proposed architecture looks like when applied to a concrete transaction;
- where identity and authorization controls reach their limits;

- and what implementers, vendors, and standards communities should do next.

Why This Matters Operationally

The issues discussed in Las Vegas are not theoretical problems that would arise only in highly autonomous systems. They appear today in 'classical' systems whenever decisions are assembled across multiple identity, security, risk, and automation products. Automation (through AI or otherwise) accelerates, amplifies, and extends the issues, but does not directly give rise to them.

The discussion and pre-reading point to four recurring failure patterns.

1. **A valid credential can be used outside the intended task**

   The employee may have broad permission to arrange travel. If the travel 'agent' receives equivalent authority, it may be able to book any hotel, change existing reservations, or access unrelated travel records. The credential is valid. The action may be within the broad entitlement. It may nevertheless exceed the bounds of the task the employee was assigned.

2. **A derived assessment can be mistaken for a fact**

   A fraud or anomaly system may classify the booking as high risk. Another system may receive that classification without knowing:

   - which model produced it;
   - which data informed it;
   - how fresh it is;
   - how confident the model was;
   - or whether it was intended as advice or as a policy outcome.

   The receiving system may then treat a probabilistic assessment as an authoritative statement.

3. **A downstream service may not know who is acting**

   If the agent simply uses the employee's token, the booking service may 'see' only the employee.

   The service cannot distinguish:

   - the person actually responsible for the request;
   - the workload that interpreted it;
   - the particular instance that acted;
   - or the authority delegated for this transaction.

   The logs may show that the employee booked a hotel even though an agent selected and executed the transaction.

4. **Logs may show what happened without showing why**

   The booking platform may record the reservation. The gateway may record a request. The policy engine may record an allow decision. The agent platform may retain a conversation.

   Unless those records are connected, the organization may be unable to determine:

   - which task was active;
   - what authority the agent had;
   - which signals affected the decision;
   - which policy was evaluated;
   - or why the selected hotel was considered acceptable.

   Each system may have worked as designed. The resulting decision may still be difficult to explain.

[Subscribe to receive the full report]

Conclusion

The identity industry has spent decades building mechanisms for authentication, federation, provisioning, authorization, workload identity, signaling, attestation, and auditing. Agentic systems do not make that work obsolete, but they do reveal where the connections remain weak.

The central challenge is not merely identifying an agent or creating a more advanced policy engine. It is preserving enough meaning as identity, context, authority, decisions, and actions move among systems.

The June Salon sharpened several points.

Probabilistic systems do not occupy one fixed place in the architecture. Uncertainty may arise in signal generation, interpretation, task planning, or execution. Deterministic controls can still govern access around those points.

Authorization remains essential, but it cannot guarantee that an agent understands a task correctly. It must be combined with distinct identity, restricted authority, task-specific constraints, external enforcement, and linked observability.

Delegation must preserve the difference between the responsible principal and the acting workload. An agent should receive authority appropriate to a transaction, not simply inherit a person's identity and permissions.

Trust-domain boundaries offer the strongest points for interoperable control. They are where systems can require explicit identity, provenance, authority, transaction context, and local policy evaluation.

Concrete architectures are valuable even when they do not define every interface. The proposed trust model showed that many components can be assembled now in a modular architecture while also exposing unresolved issues around semantics, relationships, registries, lifecycle, credential chaining, and evidence.

The industry already has many of the pieces. The work now is to determine what must remain true as a decision moves from identity to action—and to preserve enough evidence to know whether it did.

Acknowledgments

The Identity Salon thrives because of the generosity and openness of its participants. Each month, technologists, policymakers, researchers, and implementers step into a Chatham House Rule space to test ideas, challenge assumptions, and share experience. Their candor and willingness to explore difficult questions make the Salon work.

The discussion summarized in this report was entirely human. The conversations, insights, and debates captured here came directly from the participants in the room. ChatGPT was used as a drafting aid to help organize and synthesize the meeting notes into a readable report. As always, the interpretation and final editorial decisions remain human.

About The Identity Salon

The Identity Salon™ provides a unique, exclusive environment where seasoned digital identity architects, technical standards experts, and researchers can engage in meaningful, protected conversations. Limited in size to foster genuine connections, this gathering allows experienced professionals to dive into complex, long-term challenges with peers who understand the depth and breadth of identity's impact.

We host the Identity Salon under the Chatham House Rule, facilitating candid dialogue that often isn't possible in larger, more public settings. Participants have the rare opportunity to explore the '5-year problems' in identity, share leading practices, and discuss emerging approaches with like-minded experts. Our aim is to bridge the gap between academic and industry research and real-world practice, connecting public and private sectors to advance knowledge and drive practical solutions.

Heather Flanagan, is the Principal at Spherical Cow Consulting, where she helps organizations navigate the fast-moving world of digital identity and Internet standards. With more than 15 years of experience translating complex technical concepts into clear, actionable strategy, Heather is known for her ability to bridge communities, guide collaborative work, and make standards development a little less intimidating. Named to the 2025 Okta Identity 25 as one of the top thought leaders in digital identity, Heather is also a regular speaker and writer, focusing on standards, governance, and the real-world challenges of identity implementation.

Andrew Hindle, is an independent consultant focusing on digital identity, privacy, cyber security, and corporate governance. A co-founder of The Identity Salon, Andrew is also Conference Chair for both the Identiverse and Authenticate conferences, serves as a non-executive member of the board at Curity, and chairs the UK Advisory Board at the Kantara Initiative. Andrew has over 25 years' experience in the software industry in a range of technical sales, pre-sales, product marketing, business development and corporate governance roles. He maintains CIPP/E, CIPM and CIPT privacy certifications with the IAPP, a CIDPRO certification with IDPro; and holds a BA in Oriental Studies (Japanese) from Oxford University and an advanced professional diploma in corporate governance. Outside of the world of identity, Andrew holds several governance roles with the Scouts at both local and county level; rides regularly with a local road cycling group; and plays keyboard, guitar and bassoon (not at the same time) with more enthusiasm than skill, and for an audience of one. Andrew is based in the UK.

Thanks to Our 2026 Supporters!

Hindle Consulting

Spherical Cow Consulting

Weave Identity