Blockchain Governance Initiative Network

Workshop: Vulnerability Handling in the Agentic AI Era

BGIN – Julien Bringer & Mitchell Travers

Global Digital Collaboration | Geneva | 2–3 September 2026

The Challenge — From Alert Floods to Actionable Intelligence

The Problem Is Getting Worse, Fast. Agentic AI is dramatically increasing the volume and complexity of vulnerability signals:

- Automated scanning generates thousands of findings per day
- Chain-visible issues expose vulnerabilities to defenders **and adversaries simultaneously**
- Third-party reports arrive across jurisdictions, vendors, and regulated institutions
- Triage speed expectations are outpacing human team capacity

The **stakes** are uniquely high for:
Digital wallets · Verifiable credentials · Public ledgers · Custody infrastructure 🔗 **Takeaway: The Core Tension.** How do you share vulnerability intelligence fast enough to matter — without a central authority, across borders, and without feeding your adversaries? **Co-organized by:** BGIN · Blockchain Bundesverband · Ethereum Climate Platform · LNet OpenWallet Foundation · Trust Over IP · LF Decentralized Trust

Why This Problem Is Different in the Decentralized Stack

« Traditional »	Decentralized / Web3 Security
Centralized patch authority	No single coordinating owner
Private vulnerability disclosure	On-chain data is inherently public
Single-jurisdiction compliance	Multi-jurisdictional by design
Vendor-managed CVE lifecycle	Ecosystem-wide, multi-vendor impact
Alert fatigue is manageable	Alert floods are structurally built-in

Three Compounding Factors

- **Transparency paradox** — Public ledgers mean threats are visible to everyone, including attackers
- **Multi-stakeholder complexity** — Regulated banks, DeFi protocols, wallet vendors, and national agencies must cooperate without hierarchy
- **AI acceleration** — Agentic systems generate findings faster than any governance model was designed to process

🔗 **Takeaway:** The Question Is Not "Should We Share?" — It's "How Do We Share Trustworthily?"

Building on Real Foundations — What BGIN Has Already Started

We Are Not Starting From Zero. BGIN's existing work provides the scaffolding for today's session:

BGIN Cyber Security Information Sharing Framework

- Developed with ISO & ISO/IEC JTC 1 liaisons in view
- Establishes principles for **trustworthy, structured** vulnerability sharing across decentralized ecosystems
- Informs ISAC-style and cross-border cooperation models

Security Agentic AI Proof of Concept

Addresses the **triage bottleneck** directly. Three key capabilities already under development:

- **Prioritization** — Signal-to-noise filtering at scale
- **Abstraction** — Translating technical findings for multi-audience use
- **Quality controls** — Governing AI-assisted analysis outputs

NIST Workshop Input (14 May 2026) (Minimal high-value data elements for sharing, AI-assisted analysis governance, Public–private sharing model design) & **preparation of a collaboration with several US stakeholders** 🔗 **Takeaway:** Today's session builds on these foundations — it does not duplicate them

Who Should Be in the Room

Sector	Role in This Problem
Security ops & vuln management leads	Live with the alert flood every day
ISAC / CSIRT / SOC teams	Coordinate response across organizations
Wallet & custody operators	High-value targets, complex disclosure obligations
Financial & digital-asset regulators	Define compliance expectations for sharing
National cyber agencies	Cross-border coordination authority
NIST / NCCoE & standards developers	Translate practice into durable frameworks
BGIN Cyber Security & Agentic AI WG	Building the tools and frameworks & supporting a neutral coordination

What We Will Shape Together in 50 Minutes

Three Key Discussion Items — Your Input Shapes Real Outputs

1. **Key Discussion Item 1: AI-Era Triage Models**
   How should teams handle AI-generated vulnerability signals? What prioritization logic is defensible and auditable?
2. **Key Discussion Item 2: V0 Minimal Data Elements**
   What is the **minimum viable information set** for trustworthy cross-org sharing? Informed by NIST workshop outputs — refined here by practitioners
3. **Key Discussion Item 3: Multistakeholder Sharing Pilots**
   What cooperation structures actually work across borders and sectors? ISAC-style, bilateral, or something new?

Outputs from This Room Feed Directly Into: BGIN Block #15 — October 2026 · NIST & MITRE follow-up dialogue · ISO standardization track 🔗 **Takeaway:** Discussion will continue with BGIN Block #15 (in person or remote) – YOU ARE INVITED !!!

From Alert Floods to Actionable Sharing

A working frame for this breakout

Blockchain Governance Initiative Network

Shin'ichiro Matsuo, Ph.D.

BGIN Co-Chair | Breakout: Vulnerability Handling in the Agentic AI Era

Global Digital Collaboration | Geneva | 2–3 September 2026

Discussion starter

The bottleneck is no longer discovery

Vulnerability handling, incident response, and threat-intelligence sharing were built for **human-speed** disclosure. Frontier systems now generate, rewrite, and chain findings faster than those pipelines can absorb.

Those pipelines remain necessary — a CVE-shaped identifier, an ISAC ticket, a 911-style intake queue. They do not scale when discovery is AI-accelerated *and* the objects themselves are mixed.

Protocol, not CVE-shaped

A bridge or protocol bug that will not fit a classical software CVE.

Already handling-labeled

A stolen-key pattern that operators already share under TLP-style labels.

Between two desks

A social-engineering campaign that sits between disclosure and response.

🔗 **Takeaway:** Volume is now AI-speed. Priority is still local: which findings affect *this* organization's assets and duties?

What already exists — and what it is not

A public tool, already in use

- The **Community Security Agent** is public under the MIT License.
- At Block 13 (Georgetown, October 2025) a sharing-framework discussion became running software the same day. Tokyo (Block 14) repeated the pattern.
- It assists draft, triage, and matching under TLP and Codex, with multilingual briefing support.
- Matching turns AI-speed findings into organization-specific priorities.

Not a second clearinghouse

- National clearinghouses are being rebuilt for this volume. That is evidence of the problem, not a model to copy in this room.
- Do not replace the TI stacks operators already run.
- Do not invent a third private language beside existing matrices.
- Handling-labeled tickets stay in operator stacks.

The **public layer** is source, schemas, taxonomies, and lessons. The BGIN Information Sharing Framework (ISF) is the related sharing-format standard — not a second product — with an ISO/TC 307 pathway in view.

Why an ecosystem around AI agent systems matters

AI agent systems already assist draft, triage, and matching. That does not by itself produce trustworthy sharing. The missing work is a community-run ecosystem whose **working surface** is the agent.

A lone agent is a demonstration

- Each organization can run a private agent and still drown in AI-speed findings.
- Dialects proliferate, so matching cannot be checked across communities.
- Handling-labeled content never leaves, or it leaks into a public tracker.

The ecosystem is the effort

- The agent is where volume becomes organization-specific priority.
- Schemas, provenance, validation, and review rules have to travel with that surface.
- The public layer can grow; handling-labeled tickets stay in operator stacks.

🔗 **Takeaway:** Faster models raise volume. Only an ecosystem keyed on the agent turns that volume into shareable, locally actionable practice.

Why this has to be global collaboration

Neither **volume** nor **quality** can keep up inside one country or one stakeholder. A single queue is overrun; a single reviewer cannot make matching trustworthy at that speed.

Why one actor fails

- Volume: AI-speed, mixed-object findings already outrun any one country's queue.
- Quality: triage, matching, and handling degrade when one actor tries to go faster alone.
- A second national copy restores neither volume nor quality.

Where collaboration has to land

- A public layer that can travel: source, schemas, taxonomies, and lessons.
- Handling that stays local: TLP/Codex tickets remain in operator stacks.
- A shared working surface, without a central authority, so triage is comparable across communities.

🔗 **Takeaway:** Collaboration is required because volume and quality have to be pursued together — not as a courtesy after one actor builds a system.

Four conditions for a reliable sharing ecosystem

Matching is the step that matters. Reliability at that step depends on more than a faster model. A community-run open-source ecosystem has to connect public information and expert knowledge to **local context**.

1. Shared schemas and semantics

So a finding can move between communities without a private dialect.

2. Provenance

So a receiver can see where a finding came from and what was transformed.

3. Validation evidence

So assisted triage can be checked, not merely trusted.

4. Review rules

TLP/Codex, sensitivity gates, and human quality assurance for high-impact outputs.

🔗 **Takeaway:** The next job is to *plan* that ecosystem around an existing public tool. Scaling the agent is not the first job.

Questions for this room

This breakout is for shaping triage practice, cooperation structures, and the first data elements — not for locking a single architecture.

1. Which finding classes belong in the **public layer**, in **national identifier** pipelines, and in **handling-labeled** operator stacks?
2. What are the **minimal high-value data elements** worth sharing first, before full harmonization?
3. When do existing vocabularies fit, and when is **honest misfit** better than a new private language?
4. What cooperation model — ISAC-style, cross-border, private-sector-first — can operate **without a central authority**?

Follow-on: BGIN Block #15, Washington, October 2026. Public tool:
`https://github.com/smitgu/community-security-agent-public`

Join us at BGIN Block 15

15–16 October 2026 | Georgetown University, Washington, D.C. | In person (remote available). This discussion continues in the Block 15 working sessions.

Related sessions

- 16 Oct, 09:50–11:20 (Room B). Cyber + IKP: Security AI Agent — Information Sharing & Vulnerability Handling
- 15 Oct, 09:50–11:20 (Open Space). Agent Hack / Local Pi Private Knowledge Network
- 15 Oct, 11:30–13:00 (Room A). Cyber: Governance of the Security Supply Chain

Related session
bgin-global.org/.../sessions/2-4

Block 15 site
bgin-global.org/events/20261015-block15

Register
eventbrite.com/.../bgin-block15-tickets

Takeaway: Hybrid. Connection details are sent after registration. You are invited.