Foundations and Positioning of the Zemi Method

Version 1.4

Kevin V. Watson

August 10, 2026

Final publication revision August 25, 2026

Document information

Companion to: The Zemi Method, Version 1.2

Version: 1.4

Status: Final publication version. Author revision; not peer reviewed or empirically validated. This revision carries forward the Version 1.3 literature and positioning record and adds the Version 1.2 rationale and modern AI-forensics interface.

Zemi Method Version 1.2 DOI: https://doi.org/10.5281/zenodo.22084060

Versioning note: This companion is versioned independently because it records positioning, literature, acknowledgements, and pressure-test history rather than normative controls alone.

Prepared: August 10, 2026

Final publication revision: August 25, 2026

Canonical home: https://zemimethod.com

Author: Kevin V. Watson

Web: https://kevinvwatson.com

Contact: kevinvwatson@gmail.com

Material changes from v1.3

Explains the Version 1.2 clarity revision; adds trace-centric decision-making; defines Zemi's interface with AI governance; adds synthetic-media detector-output promotion as an application example; records the agentic-AI pressure tests supporting AI Action Provenance and AI-role classification; defines a proportionate Control Provenance requirement and its relationship to Human–Agent Attribution Discontinuity; and explains representation completeness for sources that may differ between human-visible and machine-readable forms.

Contents

Purpose and claim

The Zemi Method is deliberately compact and normative. This companion identifies the traditions on which it relies, records the pressure tests that shaped it, and explains its position beside forensic standards, attribution methods, evaluative-reporting frameworks, AI governance, and emerging AI-forensics practice.

The method does not claim to have invented hypothesis testing, independent verification, bounded conclusions, evidence mapping, competing explanations, or human accountability. Its contribution is the integration of claim-as-hypothesis reasoning, source completeness, independent corroboration, human adjudication, bounded classification, AI-role control, and a standing defensibility check into one repeatable discipline for investigation and assurance.

That integration operates across two related subjects. Zemi can govern work in which AI assists an investigation, and it can govern work in which an AI system, its outputs, controls, or actions are themselves under examination. The same boundary holds in both cases: machine-generated records and outputs are evidence to be tested, while findings remain human-adjudicated.

Foundations carried forward from Version 1.3

The Version 1.3 companion's complete literature and positioning record remains part of the development history. Version 1.4 carries forward these foundations:

- Popper's falsificationist account of testing and the need for a stated breaking condition.
- Toulmin's separation of data, claim, warrant, and qualifier.
- Wigmore's insistence that inferential paths be explicit and inspectable.
- Heuer's testing of competing hypotheses against the same evidence.
- Bayesian and ENFSI evaluative-reporting traditions where formal strength-of-evidence assessment is required.
- GRADE as a precedent for a spare, published, versioned classification vocabulary.
- NIST, ISO, and SWGDE process standards governing the handling and examination of digital evidence beneath Zemi's reasoning layer.
- Zero Trust Digital Forensics as the nearest published predecessor to verification-before-trust at the artifact level.
- DERDS, Digital Evidence Certainty Descriptors and their critique, preliminary evaluative opinions, SOLVE-IT, digital-forensic bias and reliability research, peer-review research, and automated-system evaluation as the method's immediate scholarly neighborhood.
- FACT as a complementary attribution framework, particularly where a Zemi-governed record must support a person-level attribution.

Version 1.4 does not replace that record with a claim of novelty. It narrows the claim further by showing how newer material informs the method without turning every modern problem into a new phase or control.

Why Version 1.2 is a clarity revision

Version 1.1 established the seven-phase lifecycle and the principal controls for authority, source completeness, independent corroboration, custody, classification, automation, AI-role reporting, and defensible release. Subsequent review found that several ideas central to the method were present but not stated with enough directness.

Version 1.2 therefore defines a finding, states that evidence is not itself a finding, makes the evidence-to-finding pathway visible, strengthens the independence test, clarifies that Zemi supports decisions rather than making them, connects *Artifact-to-Finding Promotion*, Working Paper Version 1.1, to the wider method, and states the method's integrative contribution without overclaiming originality. The working paper is aligned to these Version 1.2 concepts rather than treated as a dependency that the method requires in every engagement.

The agentic-AI pressure tests identified related operational requirements rather than a new lifecycle phase. Version 1.1 required preservation of automated inputs, outputs, tools, and versions, but consequential agents create a longer evidentiary object: an action trajectory operating through changing tools, permissions, state, environmental responses, and human interventions. Version 1.2 names this requirement AI Action Provenance. Where attribution is material, it also requires proportionate Control Provenance so evidence about the action is not collapsed into evidence about who controlled, authorized, or bears responsibility for it.

Trace-centric decision-making

Divoy, Lemans, and Bitzer's trace-centric utility-based case study treats crime-scene investigation as a sequence of decisions about attendance, search, collection, analysis, and the utility of traces. Its relevance to Zemi lies less in any particular crime-scene procedure than in the timing of defensibility.

Defensibility begins before a report is written. Decisions about what sources to seek, what traces to collect, which exclusions to make, and what analysis to prioritize shape what can later be claimed. This supports Zemi's placement of Authority and Scope before Frame, its source-completeness requirement, and its obligation to disclose evidentiary gaps created by earlier collection or selection decisions.

Trace-centric reasoning also sharpens a practical reporting duty: where material, the practitioner should be able to explain why a source, trace, artifact, or investigative step was pursued, excluded, or escalated. This does not require a new core control in Version 1.2; it provides foundation and context for controls already present.

Interface with AI governance

Zemi is not an AI governance framework. It does not select an organization's risk appetite, assign governance roles, create legal authority, or certify that a governance program is adequate. In AI assurance work, it can operate as the evidentiary discipline beneath a chosen governance framework by testing whether governance claims are supported in practice.

The governance basis of an AI assurance claim may include a law, regulation, contract, policy, standard, risk framework, control, approval, or organizational commitment. Relevant questions include:

- What rule, policy, control, or framework authorized and governed the AI use?
- What legal, regulatory, contractual, or organizational obligation applied?
- Who owned the system, risk, control, and resulting decision?
- Was the use approved, experimental, shadow, vendor-provided, or embedded in another workflow?
- What controls were supposed to exist, and what evidence shows whether they operated?
- Was the organization's understanding of the system accurate enough to support the decision made?

Lawrence and Cripps' *Governed to Grow* is relevant because it emphasizes the gap between governance documentation and an operational understanding of how AI systems function, what risks they create, and how those risks are controlled. Zemi addresses that gap narrowly: a documented control is a claim until evidence shows its design, implementation, operation, exception handling, and limits in the context being examined.

Synthetic media and detector-output promotion

Synthetic-media examination illustrates Artifact-to-Finding Promotion. A deepfake detector score, frame-level classifier output, or AI-generated authenticity assessment is a tool-derived artifact and investigative lead. It is not a finding by itself.

Iuliani's deepfake-video workflow supports a multi-domain examination involving visual artifacts, metadata and file structure, AI-based detection, geometric consistency, and spectral analysis. Under Zemi, these domains are not automatically independent merely because different tools display them. Their provenance and failure modes must still be examined. Where materially separate traces converge, they may support a human-adjudicated finding with stated processing and uncertainty limits.

The reporting question should not be forced into a binary "deepfake or authentic" conclusion where the evidence supports only narrower propositions. A defensible report may establish manipulation indicators, AI-assistance indicators, provenance limits, or unresolved alternatives involving compression, editing, CGI, transcoding, or metadata variation. Negative detector output does not establish authenticity.

This example confirms existing controls. It does not justify a new synthetic-media phase or a detector-specific rule in the core method.

Representation completeness and document-borne instructions

The Connecticut court-filing prompt-injection scenario exposed a narrow source-completeness issue. A document can present one account to a person viewing the rendered page and another to software extracting its text, OCR, metadata, annotations, or embedded content. The visible page and machine-

readable layers are not automatically independent sources. They are representations or transformations of the same artifact, but differences among them may be material evidence.

Version 1.2 therefore requires proportionate comparison where a material source may differ between human and machine readers. The practitioner preserves the relevant visible, machine-readable, and transformed representations and documents any material difference. This applies beyond prompt injection to hidden text, OCR-layer manipulation, concealed metadata instructions, poisoned retrieval content, and similar document-borne attacks.

The requirement does not turn Zemi into a prompt-injection security framework. It prevents an examiner from treating ordinary visual rendering or one extraction path as the complete source when another material representation exists.

Agentic AI, AI forensics, and action provenance

Agentic systems shift the forensic object from a prompt-and-response pair to an observable sequence of consequential actions. A useful abstraction is:

> ***Objective → Context and state → Tool selection → Invocation → Environmental response → State change → Subsequent action → Outcome***

The sequence may span model providers, agent frameworks, memory stores, plugins or skills, credentials, sandboxes, browsers, APIs, cloud systems, and human approval points. Configuration is therefore capable of being evidence. Internet access, credential exposure, tool permissions, guardrail state, orchestration logic, and evaluator intervention may explain an outcome more accurately than a model-generated narrative about what occurred.

AgentCanary independently reinforces the analytical value of full-trajectory evaluation rather than judging an agent solely from its final response or a single tool call. Its purpose is security evaluation rather than post-incident forensic doctrine, so it supports the direction of the requirement without defining Zemi's evidentiary standard.

Reported 2026 cyber-evaluation incidents involving agentic systems provide a useful structural pressure test. Public accounts describe materially different mechanisms and levels of evidence, including external connectivity made available through evaluation configuration and more complex third-party compromise sequences. Zemi should resist collapsing those mechanisms into the narrative label "rogue AI."

An investigator should distinguish:

- the action that was technically executed;
- the model, orchestrator, tool, or deterministic component that selected, transmitted, or performed it;
- the objective assigned to the system and whether the evidence supports independent objective formation;
- the permissions, credentials, connectivity, and controls that enabled the action;

- human instructions, configuration, approvals, omissions, interventions, or overrides;
- causal contribution to the outcome;
- whether the action was authorized;
- whether intent, awareness, deception, escape, or autonomy can be supported; and
- where organizational, operational, legal, and governance accountability remains.

The central distinction is that autonomous execution does not establish autonomous escape, intent, independent objective formation, or organizational accountability. AI-generated accounts and orchestration logs must be corroborated where they are relied on to establish consequential actions or explanations.

This pressure test supports AI Action Provenance as a core Version 1.2 requirement. It does not support rewriting the seven-phase lifecycle or requiring hidden chain-of-thought. The relevant evidence is the observable or recorded action trajectory, execution context, control state, human involvement, and resulting state change.

Control provenance and the attribution boundary

AI Action Provenance and Control Provenance answer different questions. Action provenance concerns what the system did, through which components and conditions, and with what result. Control provenance concerns the evidence of who or what created, configured, authorized, deployed, operated, modified, supervised, or intervened in the system or its material capabilities.

The distinction prevents a common promotion error. An account registration, credential event, device association, administrative role, infrastructure record, or approval entry may be relevant to control, but none uniquely establishes a natural person's material control, authorship, knowledge, intent, or responsibility. Delegated or shared access, credential compromise, inherited permissions, service accounts, scheduled processes, and other automation remain plausible alternatives where the record permits them.

For clarity, an execution record or path is the observable technical record of events captured by systems. An action trajectory is the investigator's evidence-bounded reconstruction of the material sequence. Neither should be treated as a complete human-attribution account merely because a person, account, or organization appears somewhere in the record.

The practical sequence is:

> ***Technical execution → Control attribution → Authorization or intent → Responsibility or accountability***

Each movement requires additional evidence. Establishing one level does not establish the next. Zemi requires that the relevant evidence be preserved, tested, bounded, and reported, but it does not turn this sequence into a new lifecycle or decide legal, organizational, or moral responsibility.

Where the engagement requires detailed attribution of a consequential AI-agent action to a natural person, the related *Human–Agent Attribution Discontinuity (HAAD)* methodology provides the

specialized proposition-level analysis. HAAD examines identity, credential, session, control, authorship, authorization, voluntariness, knowledge, intent, causation, and attributional discontinuities in greater depth. Zemi remains the governing evidence-to-finding discipline; HAAD is applied proportionately where that specialized attribution question is within authority and scope.

Premature AI attribution and causal closure

AI forensics must work in both directions. It must be capable of establishing consequential AI action where the evidence supports it, and of preventing AI from becoming a convenient causal label where configuration, conventional automation, permissions, environment design, or human action better explains the outcome.

The 2025 Replit/SaaStr incident and Amazon's later public account of an AWS Cost Explorer interruption provide a useful contrast. The AI Incident Database and related reporting recorded allegations that a Replit agent deleted production data, produced fabricated data or test results, and incorrectly described rollback availability. Replit subsequently stated that its applications had used one database for both development and live customer data and announced development/production separation. Amazon, by contrast, publicly attributed its limited interruption to a misconfigured role and access controls rather than AI, while adding mandatory peer review for production access.

The public sources do not independently establish identical root causes. They demonstrate that similar classes of evidence can be publicly framed toward or away from AI. Zemi therefore treats both "the AI did it" and "it was user error" as claims requiring component-level testing.

The visible AI action may be established and still not be the complete causal finding. A defensible analysis asks what the AI selected or executed, which deterministic components performed the action, what credentials and permissions made it effective, what environment and approval controls existed, what humans configured or authorized, and whether the outcome would have followed through the same pathway without the AI component. An agent's confession is an output artifact, not an adjudication of cause.

This inverse pressure test confirms the neutrality of Version 1.2. The evidence-to-finding pathway, AI-role classification, AI Action Provenance, proportionate Control Provenance, independent corroboration, and separation of execution, control, causation, authorization, intent, autonomy, and accountability prevent premature causal closure when properly applied.

Additional Version 1.2 pressure tests

Six further public-incident tests examined whether the Version 1.2 controls remained usable across different forms of agentic and AI-mediated activity. The tests covered the OpenAI and Hugging Face evaluation incident, Anthropic's three reported cybersecurity-evaluation incidents, the UK AI Security Institute's unsanctioned agent behaviour, an Australian gym-booking incident involving another member's reservation, the Connecticut court-filing prompt injection, and the human-directed AI-agent campaign against Taiwanese government systems.

The tests did not establish the underlying incidents independently. They assessed whether Zemi required the relevant evidentiary questions and prevented public labels from becoming findings. Across the cases, the method separated assigned objective from selected method, stated authority from technical capability, sandbox description from instantiated environment, action execution from independent objective formation, human direction from agent orchestration, and linguistic or account artifacts from person or state attribution.

Three explanatory formulations emerged without becoming new finding classifications:

- **Authorized-scope overrun**, where an agent acts beyond the authorized task boundary through capabilities the environment made technically available.
- **Delegated-authority overrun**, where an agent uses technical capability beyond the authority delegated by its operator, including authority the operator did not possess and could not delegate.
- **Human-directed, AI-orchestrated cyberattack**, where people establish the target and objective while agents coordinate or execute material attack steps under that direction.

The first two remain companion language. The third applies the existing AI-role classification rather than adding a category. The only core refinement produced by these tests was representation completeness. The other scenarios confirmed the existing Authority and Scope gate, AI Action Provenance, Control Provenance, independent-corroboration rule, human adjudication boundary, and separation of execution, authorization, causation, intent, autonomy, attribution, and accountability.

What remains outside the method

Zemi remains a reasoning and reporting methodology. It does not replace:

- lawful authority, legal analysis, disclosure rules, or professional duties;
- forensic acquisition, examination, incident-response, malware-analysis, or media-authentication procedures;
- discipline-specific evaluative reporting or formal probabilistic frameworks;
- attribution methods where person-level attribution is the question;
- AI governance, safety engineering, red-team, model-evaluation, or security-control frameworks;
- tool validation, peer review, or competent human judgment.

Its value is to make the path from source material to decision-supporting finding explicit, testable, bounded, and reviewable across those domains.

Validation boundary

Version 1.2 is an author revision shaped by literature review and pressure testing. Public incident reporting can show whether the method asks necessary questions, but it cannot demonstrate performance against raw evidence. AgentCanary and the cited forensic literature provide external intellectual support; they do not validate Zemi as a methodology.

The next credibility steps remain external practitioner review, engagement-based application where appropriate, inter-rater testing of finding classifications, reader-comprehension testing, assessment of the burden and completeness of AI Action Provenance and proportionate Control Provenance across architectures, and eventual peer-reviewed study.

References

AI Incident Database. (2025). *Incident 1152: LLM-driven Replit agent reportedly executed unauthorized destructive commands during code freeze, leading to loss of production data*. https://incidentdatabase.ai/cite/1152/

Anderson, T., Schum, D., & Twining, W. (2005). *Analysis of Evidence* (2nd ed.). Cambridge University Press.

Anthropic. (2026, July 30; updated August 3). Investigating three real-world incidents in our cybersecurity evaluations. https://www.anthropic.com/news/investigating-incidents-cybersecurity-evals

Biedermann, A., & Kotsoglou, K. (2020). Digital evidence exceptionalism? A review and discussion of conceptual hurdles in digital evidence transformation. *Forensic Science International: Synergy, 2*, 262-274. https://doi.org/10.1016/j.fsisyn.2020.08.004

Bollé, T., Casey, E., & Jacquet, M. (2020). The role of evaluations in reaching decisions using automated systems supporting forensic analysis. *Forensic Science International: Digital Investigation, 34*, 301016. https://doi.org/10.1016/j.fsidi.2020.301016

Casey, E. (2020). Standardization of forming and expressing preliminary evaluative opinions on digital evidence. *Forensic Science International: Digital Investigation, 32*, 200888. https://doi.org/10.1016/j.fsidi.2019.200888

Champod, C., Biedermann, A., Vuille, J., Willis, S., & De Kinder, J. (2016). *ENFSI Guideline for Evaluative Reporting in Forensic Science: A Primer for Legal Practitioners*.

Chernyshev, M., Baig, Z., Syed, N., Doss, R., & Shore, M. (2026). Large language models in digital forensics: Capabilities, challenges and future directions. *Forensic Science International: Digital Investigation, 56*, 302043. https://doi.org/10.1016/j.fsidi.2025.302043

Divoy, J., Lemans, I., & Bitzer, S. (2026). The decision-making process of crime scene investigation: A trace-centric utility-based case study. *Forensic Science International, 383*, 112840. https://doi.org/10.1016/j.forsciint.2026.112840

Flemyng, E., Noel-Storr, A., Macura, B., Gartlehner, G., Thomas, J., Meerpohl, J. J., Jordan, Z., Minx, J., Eisele-Metzger, A., Hamel, C., Jemiolo, P., Porritt, K., & Grainger, M. (2025). Position statement on artificial intelligence use in evidence synthesis across Cochrane, the Campbell Collaboration, JBI and

the Collaboration for Environmental Evidence 2025. *Environmental Evidence, 14*, 20. https://doi.org/10.1186/s13750-025-00374-5

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schunemann, H. J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ, 336*(7650), 924-926.

Hargreaves, C., van Beek, H., & Casey, E. (2025). SOLVE-IT: A proposed digital forensic knowledge base inspired by MITRE ATT&CK. *Forensic Science International: Digital Investigation, 52*, 301864. https://doi.org/10.1016/j.fsidi.2025.301864

Heuer, R. J. (1999). *Psychology of Intelligence Analysis*. Center for the Study of Intelligence, Central Intelligence Agency.

Horsman, G. (2019). Formalising investigative decision making in digital forensics: Proposing the Digital Evidence Reporting and Decision Support framework. *Digital Investigation, 28*, 146-151. https://doi.org/10.1016/j.diin.2019.01.007

Horsman, G. (2020). Digital Evidence Certainty Descriptors. *Forensic Science International: Digital Investigation, 32*, 200896. https://doi.org/10.1016/j.fsidi.2019.200896

Horsman, G., & Sunde, N. (2020). Part 1: The need for peer review in digital forensics. *Forensic Science International: Digital Investigation, 35*, 301062. https://doi.org/10.1016/j.fsidi.2020.301062

International Organization for Standardization. (2012). *ISO/IEC 27037:2012: Information technology - Security techniques - Guidelines for identification, collection, acquisition and preservation of digital evidence*.

International Organization for Standardization. (2015). *ISO/IEC 27043:2015: Information technology - Security techniques - Incident investigation principles and processes*.

Iuliani, M. (2026, July 28). Deepfake forensics workflow for video analysis. *Amped Software Blog*. https://blog.ampedsoftware.com/2026/07/28/deepfake-forensics-workflow-for-video-analysis

Koebler, J. (2026, August 13). Person hides prompt injection in legal filing telling AI to side with them. *404 Media*. https://www.404media.co/person-hides-prompt-injection-in-legal-filing-telling-ai-to-side-with-them/

Lawrence, A., & Cripps, A. (2026). *Governed to Grow: AI Governance as a Foundation for Innovation*. LexisNexis Australia. https://www.lexisnexis.com/community/au-resources/b/whitepapers/posts/a-practical-guide-for-legal-and-technology-leaders-navigating-accountability-innovation-and-opportunity-in-australia-s-evolving-ai-landscape

Li, P., Wang, S., Huang, Y., Shi, Y., Zhang, C., Li, Q., Lyu, Y., Shan, C., Li, F., Feng, C., Zhu, C., & Chen, L. (2026). AgentCanary: A security evaluation framework for autonomous AI agents in real executable environments. *arXiv*. https://doi.org/10.48550/arXiv.2606.10484

OpenAI. (2026, July 21; updated July 28 and 29). OpenAI and Hugging Face partner to address security incident during model evaluation. https://openai.com/index/hugging-face-model-evaluation-security-incident/

Taiwan Ministry of Digital Affairs, Administration for Cyber Security. (2026, August 13). Overseas hackers launch AI Agent attacks on government agencies; Ministry of Digital Affairs initiates response and joint defence to strengthen cybersecurity. https://moda.gov.tw/ACS/press/news/press/20394

UK AI Security Institute. (2026, August 4). Incident report: Unsanctioned agent behaviour during cyber testing. https://www.aisi.gov.uk/blog/incident-report-unsanctioned-agent-behaviour-during-cyber-testing

Wilson, C., & Hobbins, R. (2026, August 10). AI assistant hacks gym website in first known Australian autonomous cyber attack. *ABC News Australia*. https://www.abc.net.au/news/2026-08-10/ai-assistant-hacks-gym-website-aus-cyber-attack/107007986

Replit Team. (2025, July 21). Introducing a safer way to Vibe Code with Replit Databases. *Replit*. https://replit.com/blog/introducing-a-safer-way-to-vibe-code-with-replit-databases

Splunk. (2026, June 23). The AI paradox and the hidden costs of downtime. https://www.splunk.com/en_us/blog/cio-office/governing-ai-agents-for-resilience.html

Amazon Staff. (2026). AI coding bot didn't take down AWS, Amazon confirms. *Amazon*. https://www.aboutamazon.com/news/aws/aws-service-outage-ai-bot-kiro

Kent, K., Chevalier, S., Grance, T., & Dang, H. (2006). *Guide to integrating forensic techniques into incident response* (NIST Special Publication 800-86). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-86

Neale, C. (2023). Fool me once: A systematic review of techniques to authenticate digital artefacts. *Forensic Science International: Digital Investigation, 45*, 301516. https://doi.org/10.1016/j.fsidi.2023.301516

Neale, C., Kennedy, I., Price, B., Yu, Y., & Nuseibeh, B. (2022). The case for Zero Trust Digital Forensics. *Forensic Science International: Digital Investigation, 40*, 301352. https://doi.org/10.1016/j.fsidi.2022.301352

Popper, K. R. (1959). *The Logic of Scientific Discovery*. Hutchinson.

Shavers, B. (2025). *The FACT Attribution Framework* (v1.1). Zenodo. https://doi.org/10.5281/zenodo.18005597

Scientific Working Group on Digital Evidence. (2025). *SWGDE best practices for digital evidence collection* (Version 2.0, SWGDE 18-F-002-2.0). https://www.swgde.org/documents/published-complete-listing/18-f-002-2-0/

Sunde, N. (2021). What does a digital forensics opinion look like? A comparative study of digital forensics and forensic science reporting practices. *Science & Justice, 61*(5), 586-596. https://doi.org/10.1016/j.scijus.2021.06.010

Sunde, N., & Dror, I. E. (2021). A hierarchy of expert performance applied to digital forensics: Reliability and biasability in digital forensics decision making. *Forensic Science International: Digital Investigation, 37*, 301175. https://doi.org/10.1016/j.fsidi.2021.301175

Toulmin, S. E. (1958). *The Uses of Argument*. Cambridge University Press.

Wigmore, J. H. (1913). *The Principles of Judicial Proof*. Little, Brown and Company.

Watson, K. V. (2026). *Artifact-to-finding promotion: Evidentiary requirements for harm from correctly functioning AI systems* (Working Paper Version 1.1).

Willis, S. M., McKenna, L., McDermott, S., O'Donnell, G., Barrett, A., Rasmusson, B., Hoglund, T., Nordgaard, A., Berger, C. E. H., Sjerps, M. J., Molina, J. J. L., Zadora, G., Aitken, C. G. G., Lovelock, T., Lunt, L., Champod, C., Biedermann, A., Hicks, T. N., & Taroni, F. (2015). *ENFSI Guideline for Evaluative Reporting in Forensic Science*. European Network of Forensic Science Institutes.

This companion is maintained alongside the method and revised as the surrounding field develops. Its purpose is not to argue that the method is unique. It is to keep the method's foundations, debts, boundaries, and reasons for change on the record.