VENDOR DUE-DILIGENCE PACKET

Vendor Due-Diligence Packet

Responsible vendor procurement evidence for AI security engagements

aisecurity.llc · hello@davidwolf.org · Generated June 26, 2026

Contents

1. Vendor Identity & Overview

aisecurity.llc is an AI product security engineering practice. We help product, security, legal, and finance teams ship AI features, copilots, RAG systems, agents, and workflows without letting customers, attackers, or enterprise security reviewers discover the weak points first.

Engagements follow the MADE methodology — Map what is launching, Attack what can be abused, Defend what must hold, and Evidence what is safe enough to ship — and are delivered under signed, scoped agreements.

Primary contact
David Wolf · hello@davidwolf.org

Flagship offer
AI Launch Security Review — first findings in 5 business days, launch-ready review in 5–10.

Engagement basis
No-Cost Scoping Retainer + Mutual NDA before any paid work; fixed-fee SOW or private offer thereafter.

Source of record: aisecurity.llc/about

2. Responsible Vendor Procurement Statement

This packet is assembled so a champion, procurement lead, or finance reviewer can document a responsible vendor-selection decision. It compiles our published security practices, AI-usage and data-handling commitments, subprocessor disclosures, contract and policy library, and the attestation artifacts we can produce after an engagement.

Nothing in this packet is a certification or a guarantee of a security outcome. It evidences how we operate, what we will and will not do with client material, and the controls and agreements that govern an engagement.

- No-cost scoping means confidentiality, access boundaries, and review planning before paid work — not free consulting.
- Paid work is governed by signed SOW, assessment terms, DPA (where applicable), and Rules of Engagement (for adversarial testing).
- Engagement-specific contract terms auto-populate from your scope intake to reduce review friction.

3. Security Practices

Client work is performed under minimum-necessary access. We do not request production credentials, secrets, or regulated/customer personal data unless an engagement explicitly requires it and the appropriate agreements are in place.

- Access-controlled, need-to-know handling of client material.
- Encryption of evidence stores where feasible; no unmanaged personal storage or consumer chat tools for client evidence.
- Authorized-testing boundaries only; no production exploitation without a signed Rules of Engagement.
- Emergency-stop procedure for any unexpected production impact, critical zero-day, or regulated-data exposure.

Source of record: aisecurity.llc/trust-center/security

4. Secure SDLC

Our delivery tooling and deliverables follow secure-SDLC practices: scoped change, review, and evidence capture appropriate to AI security engagements.

Source of record: aisecurity.llc/trust-center/secure-sdlc

5. AI Usage & Model-Training Commitments

We do not use client confidential engagement materials to train public models, publish examples, or improve unrelated offerings except as expressly permitted in an applicable SOW, DPA, or written approval.

- No client content authorized for provider model training.
- Human review of deliverables; AI-assisted analysis is reference, not authority.
- Redaction and minimization defaults on all evidence.
- Public claims only with client approval and appropriate caveats.
- No secrets or regulated data accepted through public forms.

Source of record: aisecurity.llc/ai-governance/customer-data-and-model-training

6. Evidence Handling, Retention & Redaction

Evidence is collected to the minimum necessary, tied to a work item, finding, or deliverable, and classified (public-safe, client-confidential, restricted-access, legal-hold, delete-on-close).

- Redaction of personal data, credentials, tokens, secrets, and sensitive operational details before sharing.
- Retention reviewed at engagement close; deletion requests honored unless a legal or contractual hold applies.
- Public-safe derivatives require client approval for reuse.

Source of record: aisecurity.llc/trust-center/contracts/evidence-handling-policy

7. Privacy & Data Processing

Where customer or personal data may enter scope, a Data Processing Addendum (DPA Lite) is executed before any such data is shared, allocating controller/processor responsibilities, processing purpose, security measures, deletion/return, breach notice, and an AI provider/model-training prohibition.

Privacy policy
Published at /legal/privacy.

Data Processing Addendum
Available at /legal/data-processing-addendum and as a packet contract.

Acceptable use
Published at /legal/acceptable-use.

Source of record: aisecurity.llc/legal/privacy

8. Subprocessors

Our subprocessor list documents the infrastructure and vendor services used to operate the practice. Material changes are communicated per the DPA notice terms.

Source of record: aisecurity.llc/legal/subprocessors

9. Contract & Policy Library

The following agreements and policies govern engagements. Engagement-specific terms auto-populate from your scope intake; signer-ready drafts are produced during scoping and executed through our document-signing flow.

No-Cost Scoping Retainer
Pre-engagement scoping: $0 fees, no obligation, NDA path, access boundaries, and a draft review plan before any paid work. Converts to a paid SOW only after approval.

AI Launch Security Review SOW
Scoped statement of work for the pre-release AI Launch Security Review — first findings in 5 business days, launch-ready review in 5–10. Auto-populated from your scope intake.

Scoped Services Framework
Master services framework for discovery, product review, red-team validation, governance evidence, and paid scopes without a standing retainer.

Sponsorship Agreement
Commercial sponsorship terms with explicit research-independence and disclosure boundaries.

Mutual NDA
Mutual confidentiality protections for pre-sales, delivery, and research collaboration contexts.

Commercial Services Addendum
Converts the services framework into scoped paid work with rate card, invoicing, and activation terms.

Data Processing Addendum
Controller/processor allocation, data protection obligations, subprocessing, security measures, AI provider boundaries, and customer-data handling for scoped services.

Assessment Terms Addendum
Scope, authorization, evidence use, testing boundaries, safe harbor, retesting, reporting limitations, and reliance limits for AI product security assessments.

Statement of Work Template
Mission-specific scope, deliverables, timeline, access, assumptions, and acceptance criteria for scoped AI security engagements.

AI Red Team Rules of Engagement
Rules of engagement for authorized AI red-team validation, including targets, test windows, allowed techniques, prohibited actions, safety controls, evidence handling, escalation paths, and stop conditions.

Consultant Mission Brief
Defines specialist role, client relationship model, confidentiality, deliverables, and independence boundary for consultant-led missions.

Sponsorship Launch Addendum
Campaign schedule, sponsor assets, labeling, approval process, and launch deliverables.

Security Operations Schedule
Operational control schedule for authorized AI security work, covering access, credentials, logging, AI/ML testing boundaries, incident handling, evidence retention, and client escalation.

Evidence Handling Policy

Evidence collection, classification, storage, redaction, retention, deletion, and publication boundaries for AI security assessments, red-team work, governance evidence, and public-safe deliverables.

Publication & Claim-Readiness Policy

Claim-readiness criteria for public research, trust pages, scorecards, attestations, sponsor materials, security review outputs, and buyer-facing evidence.

Data Retention & Redaction Policy

Retention, redaction, deletion, and post-engagement handling for client materials, research artifacts, assessment evidence, exports, and public-safe publication files.

Source of record: aisecurity.llc/trust-center/contracts

10. Attestation Artifacts

After a completed engagement we can produce attestation artifacts suitable for procurement and buyer assurance:

General Security Review Attestation

Independent technical AI security review artifact for procurement and buyer assurance.

RAG Authorization Review

Retrieval authorization, tenant-boundary, and leakage review attestation.

AI Red Team Completion

Confirmation of completed adversarial testing within an authorized scope.

Trust Surface Review

Public claim, trust-language, and evidence-surface review.

Governance Evidence

Operating-model, control-ownership, and evidence-cadence attestation.

Source of record: aisecurity.llc/trust-center/attestations

11. Buyer Assurances — What We Do Not Do

To keep the engagement scoped and defensible, the following are explicitly out of scope unless separately and expressly authorized in writing:

- No rubber-stamp approvals or certification claims.
- No open-ended governance program or platform migration.
- No production exploitation or adversarial testing without a signed Rules of Engagement.
- No use of client confidential materials to train public models or improve unrelated offerings.
- No acceptance of secrets or regulated data through public forms.

12. Contacts & Next Steps

To begin, request a No-Cost Scoping Retainer and Mutual NDA. Contact David Wolf at hello@davidwolf.org.

Start no-cost scoping
/marketplace/private-offers?track=procurement-fast-track

Scope a launch review
/scope?offer=ai-launch-security-review&pressure=launch

Trust center
/trust-center

Source of record: aisecurity.llc/scope?offer=ai-launch-security-review&pressure=launch

APPENDIX — CONTRACT & POLICY LIBRARY

No-Cost Scoping Retainer

Effective Date: the Effective Date

Provider: aisecurity.llc

Client: Client

Purpose: Pre-engagement scoping for the engagement agreed during scoping

Primary contact: the primary engagement contact

This No-Cost Scoping Retainer lets a champion move legal, finance, procurement, and technical scoping in parallel **before** any paid work begins. No-cost scoping does not mean free consulting — it means confidentiality, access boundaries, and review planning so the engagement can be defined cleanly. This is a business summary; the signer-ready draft is confirmed during scoping and is reviewable by counsel.

1. No fees, no obligation

- **$0 fees.** This scoping phase carries no charge.
- **No obligation to buy.** Entering this retainer does not commit either party to a paid engagement.
- Either party may stop the scoping phase at any time, for any reason, in writing.

2. Confidentiality

- Confidentiality is governed by the parties' Mutual NDA (executed alongside or before this retainer).
- Information exchanged during scoping is used solely to plan a potential engagement.

3. What this phase covers

- Preliminary scope planning only: target system, review window, deliverables, access boundaries, and a draft review plan.
- Vendor / procurement packet and an internal approval memo to help the champion get to yes internally.

4. What this phase does not authorize

- **No production testing** unless separately authorized in a signed Statement of Work and, where applicable, Rules of Engagement.
- **No adversarial testing** of any kind unless separately authorized in a signed AI Red Team Rules of Engagement addendum.

- No access to production credentials, secrets, or regulated/customer personal data.

5. Safe intake

- Do not submit secrets, credentials, or regulated/customer personal data through public forms or unencrypted channels.
- If customer or personal data may enter scope, a DPA Lite is executed before any such data is shared.
- Provider applies minimum-necessary access and the Evidence Handling Policy to anything shared during scoping.

6. Conversion to paid work

- This retainer converts to paid work **only** after the Client approves a scoped Statement of Work or private offer.
- Typical next step: the aisecurity.llc **AI Launch Security Review** SOW — first findings in 5 business days, launch-ready review in 5–10 business days.

7. Term

This retainer is effective from the Effective Date and continues until a Statement of Work is executed or either party ends the scoping phase in writing.

Provider: aisecurity.llc · Authorized signatory: David Wolf · hello@davidwolf.org

Client: Client · Authorized signatory: _____

Mutual Non-Disclosure Agreement

Effective Date: the Effective Date

Version: v1.1

Party A: aisecurity.llc

Party B: Client

Purpose: to be specified during scoping

Term: three (3) years from the Effective Date

1. Purpose

1.1 The parties wish to exchange Confidential Information to evaluate, negotiate, or perform a potential or existing business relationship involving AI security engineering, research, advisory, sponsorship, benchmarking, professional services, or related work (the "Purpose").

1.2 This Mutual Non-Disclosure Agreement sets out each party's obligations when receiving the other party's Confidential Information.

2. Definitions

2.1 "Confidential Information" means non-public information disclosed by or on behalf of a party (the "Disclosing Party") to the other party (the "Receiving Party") that is: (a) marked or identified as confidential at the time of disclosure; (b) identified orally as confidential and confirmed in writing within five (5) business days; or (c) of a nature that a reasonable person would understand it to be confidential given the circumstances of disclosure.

2.2 Confidential Information includes, without limitation: business and strategic plans, financial information, pricing, customer and prospect information, technical information, source code, security assessments, vulnerability information, incident information, research methods, benchmark structures, model weights and training data, prompt designs, credentials, system architectures, and unpublished materials.

2.3 "Representatives" means a party's employees, officers, directors, contractors, advisors, legal counsel, accountants, and agents who: (a) need to know the Confidential Information to advance the Purpose; and (b) are bound by confidentiality obligations at least as protective as this Agreement.

3. Obligations of Receiving Party

3.1 Receiving Party will use Confidential Information solely for the Purpose and for no other purpose.

3.2 Receiving Party will protect Confidential Information with at least reasonable care, but in no event with less care than it uses to protect its own confidential information of similar sensitivity.

3.3 Receiving Party will disclose Confidential Information only to its Representatives who need to know it for the Purpose and will ensure those Representatives comply with this Agreement.

3.4 Receiving Party is responsible for any unauthorized use or disclosure of Confidential Information by its Representatives.

3.5 Receiving Party will promptly notify Disclosing Party in writing upon discovering any unauthorized access, disclosure, or use of Disclosing Party's Confidential Information.

4. Exclusions

4.1 Obligations under this Agreement do not apply to information that Receiving Party can demonstrate, with contemporaneous written documentation where reasonably available:

1. is or becomes generally available to the public through no act or omission of Receiving Party or its Representatives;

2. was rightfully known to Receiving Party without restriction before disclosure by Disclosing Party;

3. is lawfully received from a third party without restriction on use or disclosure; or

4. is independently developed by Receiving Party without use of or reference to Disclosing Party's Confidential Information.

4.2 The fact that information falls within an exclusion does not authorize disclosure if a combination of excluded information and other Confidential Information would itself be confidential.

5. Compelled Disclosure

5.1 Receiving Party may disclose Confidential Information to the extent strictly required by applicable law, regulation, court order, subpoena, or governmental authority.

5.2 Before making any such disclosure, Receiving Party will, to the extent legally permitted: (a) give Disclosing Party prompt advance written notice of the requirement; (b) cooperate reasonably with Disclosing Party's efforts to obtain protective orders, confidential treatment, or other appropriate relief; and (c) disclose only the minimum portion of Confidential Information legally required.

5.3 Compelled disclosure under this Section does not authorize any use of the disclosed information beyond what is compelled.

6. Security Information

6.1 Security findings, vulnerability details, incident information, exploit information, credentials, infrastructure configuration, model weights, prompt designs, and remediation plans are Confidential Information regardless of whether they are marked as such.

6.2 Receiving Party will not publicly disclose any vulnerability or security incident information received from Disclosing Party without Disclosing Party's prior written authorization, except as required by applicable law.

6.3 Receiving Party will apply heightened care to security-related Confidential Information, including limiting access to security-cleared or need-to-know personnel and storing such information in access-controlled systems.

7. Return or Destruction

7.1 Upon written request by Disclosing Party, or upon termination of discussions or this Agreement, Receiving Party will, at Disclosing Party's election, promptly return or destroy all Confidential Information within thirty (30) days and certify completion in writing.

7.2 Receiving Party may retain archival or backup copies to the extent required by applicable law, court order, legal hold, or established records retention policies, subject to the ongoing obligations of this Agreement.

7.3 Confidentiality obligations continue to apply to any retained copies until they are destroyed or returned.

8. No License or Rights

8.1 Confidential Information remains the sole property of Disclosing Party.

8.2 No license, assignment, ownership interest, or other intellectual property right is granted by this Agreement or by any disclosure of Confidential Information, except the limited right to use Confidential Information for the Purpose.

8.3 Neither party acquires any right to use the other party's name, logo, or marks without separate written authorization.

9. No Obligation to Proceed

9.1 This Agreement does not obligate either party to enter into any transaction, engagement, investment, partnership, or commercial relationship.

9.2 Either party may discontinue discussions and negotiations at any time and for any reason without liability, unless a separate signed agreement expressly requires otherwise.

10. Remedies

10.1 Each party acknowledges that unauthorized disclosure or use of the other party's Confidential Information may cause irreparable harm for which monetary damages alone would be an inadequate remedy.

10.2 Disclosing Party is entitled to seek injunctive or other equitable relief to prevent or restrain any breach or threatened breach, without the necessity of proving actual damages and without the requirement to post any bond or other security.

10.3 The right to seek equitable relief is in addition to, and does not limit, any other remedies available at law or in equity.

11. Term and Survival

11.1 This Agreement begins on the Effective Date and continues for three (3) years, unless terminated earlier by either party on thirty (30) days' written notice.

11.2 Confidentiality obligations with respect to Confidential Information disclosed during the term survive for three (3) years after the date of each specific disclosure.

11.3 Trade secrets remain protected as long as they qualify as trade secrets under applicable law, regardless of the survival period above.

12. Representations and Warranties

12.1 Each party represents and warrants that: (a) it has the authority to enter into this Agreement; (b) this Agreement does not conflict with any other obligation; and (c) it will comply with applicable laws in performing its obligations.

12.2 Confidential Information is disclosed "as is." Disclosing Party makes no warranty as to the accuracy, completeness, or fitness for any particular purpose of Confidential Information, unless a separate signed agreement states otherwise.

13. Limitation of Liability

13.1 Neither party will be liable for indirect, incidental, special, consequential, exemplary, or punitive damages arising from or related to this Agreement, to the maximum extent permitted by applicable law.

13.2 This limitation does not apply to: (a) unauthorized disclosure or willful misuse of Confidential Information; or (b) liability that cannot be excluded under applicable law.

14. Governing Law and Disputes

14.1 This Agreement is governed by the laws of the laws of the State of Delaware, USA (excluding conflict-of-laws rules), without regard to conflict-of-law provisions.

14.2 The parties will first attempt to resolve any dispute through good-faith negotiation between senior representatives within thirty (30) days of written notice.

14.3 If negotiation fails, disputes will be resolved in the state and federal courts located in Delaware, USA.

14.4 Notwithstanding Section 14.3, either party may seek emergency injunctive or equitable relief in any court of competent jurisdiction to prevent irreparable harm.

15. General Provisions

15.1 **Relationship of Parties.** Nothing in this Agreement creates an employment, agency, partnership, or joint venture relationship.

15.2 **Assignment.** Neither party may assign this Agreement without prior written consent, except to an affiliate or successor in a merger, acquisition, or sale of substantially all assets. Any prohibited assignment is void.

15.3 **Waiver.** Failure to enforce any right is not a waiver of future enforcement. Waivers must be in writing and signed by the waiving party.

15.4 **Entire Agreement.** This Agreement is the entire agreement between the parties regarding confidentiality for the Purpose and supersedes all prior understandings on the same subject.

15.5 **Amendments.** Amendments require a written instrument signed by authorized representatives of both parties.

15.6 **Severability.** If any provision is found unenforceable, the remainder continues in force. The unenforceable provision will be modified to the minimum extent necessary to make it enforceable.

15.7 **Counterparts and Electronic Signatures.** This Agreement may be executed in one or more counterparts. Electronic signatures have the same legal effect as handwritten signatures.

15.8 **Notices.** Notices must be delivered by email with confirmed receipt, or by courier, to the contact identified in this Agreement or any updated address provided in writing.

16. Notices

16.1 Party A Notice Contact: hello@davidwolf.org

16.2 Party B Notice Contact: the Client notice email on file

17. Signature Blocks

Party A: aisecurity.llc

Signature: _____

Name: David Wolf

Title: Principal

Date: _____

Party B: Client

Signature: _____

Name: Client authorized signatory

Title: Authorized signatory

Date: _____

APPENDIX — CONTRACT & POLICY LIBRARY

AI Launch Security Review — Statement of Work

Effective Date: the Effective Date

Provider: aisecurity.llc

Client: Client

Engagement: the engagement agreed during scoping

Primary contact: the primary engagement contact

This Statement of Work ("SOW") describes a scoped, pre-release **AI Launch Security Review**. It is entered under the parties' Zero-Dollar Services Retainer / Scoped Services Framework (or as a standalone agreement) and the Assessment Terms Addendum. This is a business document; final signer-ready language is confirmed during scoping and is reviewable by counsel.

1. Target system

the AI system identified during scoping

In-scope AI surfaces: to be specified during scoping

2. Review window and timeline

- **First findings in 5 business days** from kickoff and access.
- **Launch-ready review in 5–10 business days.**
- Review window: the review window defined in the applicable SOW
- Detailed timeline: the timeline defined in the applicable SOW

A deeper 2–4 week AI Product Security Assessment is available as separate follow-on work.

3. Deliverables

the deliverables defined in the applicable SOW

The first deliverable is a decision package, not just a report: it states what must be fixed before launch and what evidence buyers can rely on.

4. Scope summary

the scope agreed during scoping

5. Assumptions

- Client provides timely access to the materials in the Technical Access Checklist (architecture, demo/staging, prompts, RAG sources, agent tools, authz, logs).
- Review reflects the system state during the review window and is point-in-time.
- Reference inputs: architecture diagrams, data-flow documentation, and prior assessment reports as available
- Participants: the Client engineering and product leads

6. Exclusions

the exclusions defined in the applicable SOW

This review is not a certification, not a guarantee of future security, not an open-ended program, and not a platform migration.

7. Access boundaries and authorization

- Environment: a staging or development environment unless production access is separately authorized in writing
- Data access: minimum-necessary access; no raw credentials or production customer data without explicit authorization
- Authorized testing: static analysis, architecture review, prompt-injection testing, and AI-specific risk modeling within the agreed scope
- Restrictions: no production exploitation unless separately authorized in a signed Rules of Engagement addendum

No production exploitation or adversarial testing occurs unless separately authorized in a signed AI Red Team Rules of Engagement addendum.

8. Acceptance criteria

Accepted when the agreed deliverables are delivered and the Client has had a reasonable factual-review window.

Accepted when the agreed deliverables are delivered and the Client has had a reasonable factual-review window.

9. Fees and payment

- Fee / private-offer path: as specified in the applicable Order Form or SOW
- Payment terms: Net 30 on invoice

Fees are scoped after triage, with fixed-fee options where possible. Typical budget categories: launch, AppSec, product security, red team, customer assurance, or security review.

10. Confidentiality and data handling

Confidentiality is governed by the Mutual NDA. Evidence handling, retention, and redaction follow the Evidence Handling Policy and Data Retention & Redaction Policy. Client confidential engagement materials are not used to train public models or improve unrelated offerings except as expressly permitted in writing.

Provider: aisecurity.llc · Authorized signatory: David Wolf · hello@davidwolf.org

Client: Client · Authorized signatory: _____

APPENDIX — CONTRACT & POLICY LIBRARY

Assessment Terms Addendum

Effective Date: the Effective Date

Version: v1.0

Client: Client

Provider: aisecurity.llc

Master Agreement: the Master Services Agreement or Scoped Services Framework

Assessment Type: to be specified during scoping

1. Purpose

1.1 This Assessment Terms Addendum ("Addendum") supplements the Master Agreement and governs any scoped AI product security assessment, agentic workflow review, Enterprise AI Security Evidence Pack, or related advisory engagement performed by Provider.

1.2 This Addendum is intended to keep assessment work explicit: what is authorized, what evidence may be collected, how findings are handled, and what the client may rely on. It does not create a certification, compliance guarantee, or audit opinion.

1.3 If this Addendum conflicts with a Statement of Work, the more specific assessment scope controls for the applicable engagement, provided the conflict does not reduce mandatory confidentiality, data handling, or safety protections in the Master Agreement.

2. Assessment Objectives

2.1 The assessment may include one or more of the following objectives:

- identify attack paths, abuse cases, and control gaps in LLM, RAG, agentic, or other AI-enabled systems;
- map data, prompt, retrieval, tool-use, and model boundaries;
- review evidence quality for enterprise scrutiny;
- produce a prioritized remediation backlog; and
- support buyer-facing security review, governance, or procurement readiness.

2.2 Provider will describe the exact objective set in the applicable Statement of Work or kickoff memo.

3. Authorization and Testing Boundaries

3.1 Client authorizes Provider to perform only the specific activities listed in the applicable Statement of Work, Rules of Engagement, or written amendment.

3.2 Unless expressly authorized in writing, Provider will not:

1. attempt destructive testing;

2. execute denial-of-service testing against production;

3. access data outside the defined scope;

4. exfiltrate production data except the minimum necessary to document a finding;

5. persist access, backdoors, or credentials; or

6. test third-party systems not expressly covered by written authorization.

3.3 If Provider encounters a condition that suggests the assessment should expand beyond the written scope, Provider will pause and request written approval before proceeding.

4. Evidence Use

4.1 Provider may collect only the minimum evidence reasonably necessary to document findings, support remediation, and preserve the integrity of the engagement record.

4.2 Evidence may include screenshots, request/response traces, logs, configuration snapshots, prompt traces, control maps, architecture notes, and short excerpts of relevant artifacts.

4.3 Provider will not publish raw client data, raw job-description corpora, raw survey answers, private keys, tokens, or other sensitive material in public deliverables.

4.4 If Provider needs to retain sensitive evidence longer than the engagement, Provider will follow the retention and redaction rules in the applicable Master Agreement, Schedule, or this Addendum.

5. Findings and Reporting

5.1 Findings are delivered as directional security engineering evidence, not as a guarantee that every issue has been identified.

5.2 Provider may assign severity or priority based on technical impact, exploitability, business context, and exposure. Those ratings are advisory, not a certification label.

5.3 Client may request factual corrections to the report. A factual correction request may not be used to suppress a valid finding, but it may be used to correct scope, terminology, architecture details, or evidence interpretation.

5.4 Provider will label any public-safe summary as claim-ready only where the supporting evidence and publication controls are sufficient for external use.

6. Retest and Verification

6.1 Unless otherwise specified in the Statement of Work, one follow-up verification pass may be included for the highest-priority findings if remediation evidence is provided within the agreed review window.

6.2 Retest is limited to validating whether the specific issue has been remediated. Retest is not a full re-assessment of the system unless separately scoped.

6.3 If the remediation changes system architecture, authentication boundaries, data flow, or authorization model, Provider may require an updated scope before retesting.

7. Client Responsibilities

7.1 Client will identify a primary technical contact and a security contact for rapid coordination during the assessment.

7.2 Client will provide timely access to the systems, documents, and environment(s) identified in the scope.

7.3 Client is responsible for ensuring that any third-party approvals required for the assessment have been obtained before testing begins.

7.4 Client will promptly notify Provider of any systems, data classes, or business periods that are off-limits.

8. No Certification or Warranty

8.1 Provider does not certify that Client is secure, compliant, or free of vulnerabilities.

8.2 Provider does not guarantee that all issues will be found, that all controls are effective, or that all findings will be exploitable in the same way in production.

8.3 Provider's assessment outputs are time-bound and reflect the state of the environment at the time of analysis.

9. Publication and Disclosure

9.1 Provider may reference the engagement in aggregate, anonymized, or public-safe form only in accordance with the publication and claim-readiness rules applicable to the engagement.

9.2 Client may not publicly quote or market the results of the assessment without prior written approval of the specific language and context.

9.3 Any sponsor or partner relationship is separate from the assessment and does not change methodology, findings, or conclusions.

10. Term

10.1 This Addendum remains in effect for the duration of the applicable assessment and survives as necessary for confidentiality, evidence handling, and dispute resolution.

11. Signature Blocks

Client: Client

Signature: _____

Name: Client authorized signatory

Title: Authorized signatory

Date: _____

Provider: aisecurity.llc

Signature: _____

Name: David Wolf

Title: Principal

Date: _____

APPENDIX — CONTRACT & POLICY LIBRARY

AI Red Team — Rules of Engagement

aisecurity.llc

Addendum to Statement of Work · Negotiation Draft

> **Required caveat:** This template is provided for transparency and scoping. It is not legal advice and does not replace a final reviewed and signed agreement. All bracketed placeholders must be completed and confirmed before testing begins. Testing must not start until this document is signed by both parties.

---

1. Engagement Summary

Field	Value
Client	Client
Provider	aisecurity.llc
Engagement Name	the engagement agreed during scoping
Related SOW	the applicable Statement of Work
Test Window	to be specified during scoping to to be specified during scoping
Client Security Contact	the Client security contact
Provider Test Lead	David Wolf · hello@davidwolf.org

---

2. Authorized Targets

Testing is authorized against the following systems, endpoints, and interfaces only:

- The specific AI systems, endpoints, models/providers, RAG corpora, and agent workflows enumerated and confirmed in writing during scoping.
- No system, endpoint, account, or data source outside that confirmed list is in scope.

Testing is explicitly limited to the above. Any system, endpoint, account, or data source not listed here is out of scope.

---

3. Authorized Attack Techniques

The following attack technique families are authorized:

- Prompt injection (direct and indirect)
- Jailbreak and policy-bypass attempts
- Context-window manipulation
- RAG corpus poisoning simulation (read-only unless separately authorized)
- Tool misuse and function-call abuse
- Unsafe action-path exploration
- Output-handling and data-exfiltration-via-output testing
- Agent goal hijacking
- Additional technique families only as confirmed in writing during scoping

Prohibited Actions

The following are explicitly prohibited regardless of technical capability:

- Accessing, modifying, exfiltrating, or destroying production data not in the authorized targets
- Attacking systems, endpoints, or accounts not in the authorized targets
- Social engineering of Client personnel unless separately authorized
- Denial of service or load testing unless separately authorized
- Introducing persistent artifacts (backdoors, modified prompts) into any environment
- Sharing test artifacts, prompts, or findings outside the named delivery contacts
- Publishing or referencing Client systems or findings without written consent

---

4. Test Windows and Scheduling

Window	Dates / Times
Primary test window	to be specified during scoping
Authorized testing hours	the agreed testing hours
Blackout dates	any blackout dates identified by the Client
Coordination cadence	an agreed coordination cadence

Provider will notify Client security contact at the start and end of each active testing session.

---

5. Evidence and Data Handling

- All test artifacts (prompts, outputs, screenshots, logs) are treated as confidential Client information.
- Evidence will be stored in a an access-controlled, encrypted store available only to named delivery contacts accessible only to named delivery contacts.
- Raw test outputs containing sensitive data will be redacted before inclusion in reports.
- Evidence will be retained for 30 days or until final delivery, whichever is later and then deleted or returned per Client instruction.
- Provider will not use test artifacts to train AI models or for any purpose beyond this engagement.

---

6. Emergency Stop Procedure

If testing causes unexpected production impact, discovery of a critical zero-day, or exposure of regulated data:

1. Provider immediately stops all test activity and notifies Client security contact via the agreed secure channel (phone, Signal, or secure email).

2. Client security contact confirms receipt within fifteen (15) minutes.

3. Both parties jointly assess impact before testing resumes.

4. If Client cannot be reached within fifteen (15) minutes, Provider halts all testing until contact is re-established.

Emergency contact (Client): the Client emergency contact

Emergency contact (Provider): David Wolf · hello@davidwolf.org

---

7. Reporting

Deliverable	Due Date
Daily or session status notes	at the end of each testing session
Preliminary findings (critical severity)	within 24 hours of discovery
Draft report	to be specified during scoping
Client review period	five (5) business days
Final report	to be specified during scoping
Readout session	to be specified during scoping

Reports will include: attack scenarios executed, findings with evidence, severity and exploitability notes, and mitigation guidance.

---

8. Severity and Escalation

Severity	Response
Critical	Immediate notification to Client security contact. Pause testing pending acknowledgment.
High	Notification within 24 hours. Continue testing unless Client requests pause.
Medium / Low / Info	Include in report. No pause required.

Severity is assessed by Provider based on exploitability and impact against the AI system and any downstream systems.

---

9. Caveats

- This engagement is time-bound and scenario-scoped. It does not guarantee exhaustive vulnerability discovery.
- Findings are limited to the authorized targets and attack techniques listed above.
- Test results reflect the state of the system at the time of testing. Changes after the engagement may affect the validity of findings.
- No certification, compliance approval, or formal audit opinion is provided.
- Public disclosure of findings requires written agreement from both parties.

---

10. Signatures

Testing must not begin until both parties have signed.

Client (authorizing test against listed targets):

Signature: _____

Name: _____

Title: _____

Date: _____

Provider (aisecurity.llc):

Signature: _____

Name: David Wolf

Title: Principal

Date: _____

These materials are provided for transparency and scoping. They are not legal advice and do not replace a final signed agreement. Consult qualified legal counsel before execution.

Data Processing Addendum (Lite)

Effective Date: the Effective Date

Version: v1.1

Controller: Client (Controller)

Processor: aisecurity.llc

Master Agreement: the Master Services Agreement or Scoped Services Framework dated the date of the Master Agreement

1. Purpose and Scope

1.1 This Data Processing Addendum ("Addendum") supplements the Master Services Agreement or Scoped Services Framework (the "Master Agreement") and governs aisecurity.llc's ("Processor") processing of personal data on behalf of Client (Controller) ("Controller") in connection with the services described in the Master Agreement.

1.2 The parties intend this Addendum to satisfy applicable data protection requirements, including where required by the EU General Data Protection Regulation (GDPR), UK GDPR, the California Consumer Privacy Act (CCPA), and comparable data protection laws.

1.3 This Addendum is incorporated into and forms part of the Master Agreement. In the event of conflict between this Addendum and the Master Agreement on data protection matters, this Addendum controls.

1.4 This Addendum applies only to Personal Data that Controller provides to Processor under the Master Agreement. It does not apply to Processor's processing of its own customer data or data collected independently.

2. Definitions

2.1 "Controller" means Client (Controller) as the party that determines the purposes and means of Personal Data processing.

2.2 "Processor" means aisecurity.llc as the party that processes Personal Data on behalf of and under the documented instructions of Controller.

2.3 "Personal Data" means any information relating to an identified or identifiable natural person that Controller provides to Processor under the Master Agreement.

2.4 "Processing" means any operation performed on Personal Data, including collection, storage, use, disclosure, transfer, or erasure.

2.5 "Security Incident" means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to Personal Data transmitted, stored, or

otherwise processed under this Addendum. For the avoidance of doubt, a Security Incident includes events that are suspected but not yet confirmed.

2.6 "Subprocessor" means any third party engaged by Processor to carry out Processing of Personal Data on Controller's behalf.

2.7 "Applicable Data Protection Law" means the data protection and privacy laws applicable to the Processing, which may include the GDPR, UK GDPR, CCPA, and comparable statutes.

3. Processing Details

3.1 The subject matter, nature, purpose, and categories of Personal Data processed under this Addendum are as follows:

1. Subject matter: to be specified during scoping

2. Nature and purpose: to be specified during scoping

3. Categories of data subjects: to be specified during scoping

4. Categories of personal data: to be specified during scoping

5. Special categories (if applicable): to be specified during scoping

6. Retention period: to be specified during scoping

3.2 This Addendum does not govern Processor's processing of publicly available job descriptions, public hiring signals, or aggregate benchmark data that does not constitute Personal Data under Applicable Data Protection Law.

4. Controller Instructions

4.1 Processor will process Personal Data only on Controller's documented instructions, which are set out in this Addendum, the Master Agreement, and any subsequent written instructions provided by Controller during the engagement.

4.2 Processor will promptly notify Controller if it reasonably believes an instruction violates Applicable Data Protection Law, unless prohibited by law from doing so.

4.3 Processor will not process Personal Data for any purpose beyond Controller's documented instructions, except to the extent required by applicable law, in which case Processor will notify Controller before processing (unless prohibited by law).

5. Processor Obligations

5.1 Processor will:

1. process Personal Data only as documented in Controller's instructions;

2. ensure that persons authorized to process Personal Data are committed to confidentiality or subject to statutory confidentiality obligations;

3. implement and maintain the technical and organizational security measures in Section 6;

4. engage Subprocessors only in accordance with Section 8;

5. assist Controller with data subject rights requests in accordance with Section 9;

6. notify Controller of Security Incidents in accordance with Section 10;

7. assist Controller with data protection impact assessments in accordance with Section 11;

8. return or delete Personal Data in accordance with Section 12; and

9. make available information reasonably necessary to demonstrate compliance with this Addendum and cooperate with Controller's audit rights in Section 13.

6. Technical and Organizational Security Measures

6.1 Processor will implement and maintain security measures appropriate to the nature, scope, context, and purposes of Processing and the risks posed to data subjects, including:

1. encryption of Personal Data in transit using TLS 1.2 or higher;

2. encryption of Personal Data at rest using AES-256 or equivalent;

3. access controls limiting Personal Data access to authorized personnel only, based on need-to-know;

4. logical and physical separation of Controller Personal Data from Processor's own data and other clients' data;

5. audit logging of access to and Processing of Personal Data;

6. vulnerability management, patch management, and security monitoring for systems that process Personal Data;

7. background verification practices for personnel with regular access to Personal Data; and

8. documented incident detection, classification, response, and notification procedures.

6.2 Processor's security measures will evolve with applicable technology standards and industry practice. Processor will maintain measures that provide at least equivalent protection to those described in Section 6.1.

6.3 Processor does not warrant that its security measures will prevent all unauthorized access, disclosure, or loss. Controller acknowledges that no system is completely secure.

7. Confidentiality of Processing

7.1 Processor will ensure that personnel authorized to process Personal Data are subject to appropriate confidentiality commitments, whether by contract or statutory obligation.

7.2 Processor will restrict access to Personal Data to personnel who require access to perform Processor's obligations under the Master Agreement.

8. Subprocessors

8.1 Controller grants general authorization for Processor to engage Subprocessors to carry out specific Processing activities on Controller's behalf, subject to the requirements of this Section.

8.2 Processor will maintain a current list of Subprocessors, which is available at https://aisecurity.llc/legal/subprocessors or will be provided upon request.

8.3 Processor will notify Controller of any intended addition, replacement, or removal of a Subprocessor at least thirty (30) days before the change takes effect, by email or by updating the Subprocessor list.

8.4 Controller may object to a new or replacement Subprocessor by providing written notice within fifteen (15) days of notification, stating legitimate grounds related to data protection. If the parties cannot agree on an alternative, Controller may terminate the affected services on written notice.

8.5 Processor will impose on each Subprocessor data protection obligations that are substantively equivalent to those in this Addendum and will remain fully responsible to Controller for the performance of those obligations by each Subprocessor.

9. Data Subject Rights

9.1 Processor will provide Controller with reasonable technical and organizational assistance to help Controller respond to data subject requests for access, rectification, erasure, restriction, portability, or objection under Applicable Data Protection Law.

9.2 Processor will forward data subject requests received directly by Processor to Controller without undue delay, and in any event within five (5) business days of receipt.

9.3 Controller remains responsible for all determinations about how to respond to data subject requests. Processor's assistance is limited to what is technically feasible within Processor's systems.

9.4 If Processor is legally prohibited from notifying Controller of a data subject request, Processor will inform the relevant authority and refer the data subject to Controller.

10. Security Incident Notification

10.1 Processor will notify Controller of a Security Incident without undue delay, and in any event within 72 hours of Processor becoming aware of the incident, regardless of whether the incident is confirmed or still under investigation.

10.2 Processor's initial notification will include, to the extent then available:

1. a description of the nature and likely cause of the Security Incident;

2. the categories and approximate number of data subjects and Personal Data records affected;

3. the name and contact details of Processor's data protection contact;

4. the likely consequences of the Security Incident for data subjects; and

5. the measures taken or proposed to address the Security Incident and mitigate its effects.

10.3 Where not all information is available at the time of initial notification, Processor may provide information in phases and will supplement its notice as additional information becomes available.

10.4 Notification to Controller does not constitute an acknowledgment of fault, causation, liability, or negligence on the part of Processor.

10.5 Controller is responsible for determining whether a Security Incident requires notification to a supervisory authority or affected data subjects under Applicable Data Protection Law. Processor will reasonably cooperate with Controller's notification obligations.

11. Data Protection Impact Assessments

11.1 To the extent required by Applicable Data Protection Law, Processor will provide reasonable assistance to Controller in conducting data protection impact assessments (DPIAs) and prior consultations with supervisory authorities relating to Processor's Processing of Personal Data under this Addendum.

11.2 Processor's assistance will be proportional to the nature and scope of its Processing activities and limited to information within Processor's knowledge and control.

12. Return and Deletion

12.1 Upon expiry or termination of the Master Agreement, or upon Controller's written request, Processor will, at Controller's election:

1. return all Personal Data to Controller in a commonly used and machine-readable format; or

2. securely delete or destroy all Personal Data, including copies held by Subprocessors.

12.2 Processor will complete return or deletion within thirty (30) days of written request and certify completion in writing.

12.3 Processor may retain Personal Data beyond the deletion deadline to the extent required by Applicable Data Protection Law, applicable regulatory obligation, legal hold, or dispute resolution requirement. Processor will notify Controller of any such retention and will continue to apply the protections in this Addendum to any retained Personal Data until it is deleted.

13. Audit Rights

13.1 Processor will make available to Controller, upon reasonable written request, information reasonably necessary to demonstrate compliance with this Addendum.

13.2 Controller may, no more than once per calendar year and with thirty (30) days' prior written notice, conduct or commission a third-party audit of Processor's data protection practices, limited to matters within the scope of this Addendum.

13.3 Audits must be conducted during normal business hours, with minimum disruption to Processor's operations. Auditors must execute a confidentiality agreement acceptable to Processor

before receiving access.

13.4 Where an industry-standard certification (such as ISO 27001, SOC 2 Type II) or qualified independent assessment covers the relevant controls, Processor may satisfy audit requests by providing the applicable report in lieu of an on-site audit.

13.5 Controller will bear audit costs and expenses unless the audit reveals a material breach of this Addendum by Processor, in which case Processor will bear the reasonable costs of that audit.

14. International Transfers

14.1 Processor will not transfer Personal Data to a country or territory outside the European Economic Area, the United Kingdom, or a jurisdiction recognized by the relevant supervisory authority as providing adequate protection, without:

1. Controller's prior written authorization; and

2. implementation of an appropriate transfer mechanism under Applicable Data Protection Law, such as Standard Contractual Clauses (SCCs) adopted or approved by the European Commission or UK Information Commissioner's Office, or binding corporate rules.

14.2 The parties will execute any required transfer mechanisms upon request, and Processor will update such mechanisms if required by changes in applicable law.

14.3 Where Processor uses SCCs or equivalent mechanisms, Processor will comply with all obligations imposed on data importers by those mechanisms.

15. Liability

15.1 Liability of each party for breaches of this Addendum is governed by the Master Agreement's limitation of liability provisions, to the extent permitted by Applicable Data Protection Law.

15.2 Where Applicable Data Protection Law allocates liability between controller and processor based on fault, each party will bear liability for the portion of harm, loss, or regulatory fine attributable to its own breach of its respective obligations.

16. Term and Survival

16.1 This Addendum is effective from the Effective Date and continues for as long as Processor processes Personal Data under the Master Agreement.

16.2 The following obligations survive termination of this Addendum: confidentiality; Section 10 (Security Incident Notification) for incidents discovered after termination relating to data processed during the term; Section 12 (Return and Deletion) until complete; Section 13 (Audit Rights) for the period permitted by applicable law; and Section 14 (International Transfers) for any retained Personal Data.

17. Signature Blocks

Controller: Client (Controller)

Signature: _____

Name: Client authorized signatory

Title: Authorized signatory

Date: _____

Processor: aisecurity.llc

Signature: _____

Name: David Wolf

Title: Principal

Date: _____

Evidence Handling Policy

Effective Date: the Effective Date

Version: v1.0

Owner: aisecurity.llc

Applies To: The State of AI Security Engineering 2026 and related client engagements

1. Purpose

1.1 This Policy defines how evidence is collected, labeled, stored, shared, redacted, retained, and deleted when Provider handles security evidence, assessment artifacts, research references, or client-supplied materials.

1.2 The goal is simple: preserve what is necessary to support the work, and remove what is not.

2. Scope of Evidence

2.1 Evidence may include:

- screenshots;
- request and response logs;
- configuration snapshots;
- architecture diagrams;
- short prompt traces;
- test notes;
- remediation artifacts;
- redacted exports;
- attestations; and
- supporting correspondence.

2.2 Evidence does not include credentials, secrets, or raw personal data unless the specific work requires temporary access and the material is protected accordingly.

3. Evidence Principles

3.1 Evidence collection must be minimum necessary, scope-bound, and tied to a work item, finding, or deliverable.

3.2 Provider will avoid collecting more data than needed to prove the condition, describe the risk, or support remediation.

3.3 Provider will prefer redacted, truncated, or synthetic representations when they convey the issue without exposing unnecessary detail.

3.4 Evidence used in public-safe publications must be reviewed for claim readiness before release.

4. Classification

4.1 Provider may classify evidence into the following practical categories:

1. public-safe;

2. client-confidential;

3. restricted-access;

4. legal-hold; and

5. delete-on-close.

4.2 Classification determines storage location, access controls, and retention requirements.

5. Storage and Access

5.1 Evidence will be stored in access-controlled systems appropriate to its sensitivity.

5.2 Access will be limited to personnel who need the material to perform the work.

5.3 Evidence with sensitive content will be encrypted at rest and, where feasible, in transit.

5.4 Provider will not store evidence in unmanaged personal storage, consumer chat tools, or unsecured shared folders.

6. Redaction Standards

6.1 Provider will redact or suppress:

- personal data;
- credentials and tokens;
- raw secrets;
- internal target lists;
- private contact details;
- privileged operational details; and
- any content that could cause unnecessary exposure if published.

6.2 Redaction must preserve enough context to make the evidence useful and understandable.

6.3 If redaction would remove the evidence value, Provider will use an explanatory summary instead.

7. Retention

7.1 Evidence will be retained only for the period required to complete the engagement, support retesting, satisfy legal or contractual obligations, or maintain the research record.

7.2 Unless a different schedule is specified in the applicable agreement, evidence should be reviewed for deletion or archival at engagement close.

7.3 Evidence subject to legal hold, dispute hold, or publication review may be retained until the hold is released.

8. Sharing

8.1 Evidence may be shared internally with personnel who need it for analysis, review, quality assurance, or delivery.

8.2 Client-facing sharing will use the least sensitive version that still supports the point being made.

8.3 Public sharing requires explicit review against claim-readiness and publication rules.

9. Deletion

9.1 When evidence is no longer needed, Provider will delete it using a commercially reasonable secure-deletion process.

9.2 If the platform does not support verifiable secure deletion, Provider will remove the material from active access and prevent further use to the extent reasonably possible.

9.3 Deletion requests from a client will be honored where they do not conflict with legal obligations, retention requirements, or active disputes.

10. Incident Response

10.1 If evidence is suspected to be exposed, altered, or mishandled, Provider will investigate, contain, and notify the appropriate internal and external contacts as required.

10.2 The incident record will include what was affected, what evidence was involved, who had access, and what corrective actions were taken.

11. Exceptions

11.1 Any exception to this Policy must be approved by an authorized manager and documented with the reason, scope, and expiration date.

12. Review

12.1 This Policy should be reviewed periodically and after any material change in evidence practices, delivery systems, or publication workflows.

Data Retention & Redaction Policy

Effective Date: the Effective Date

Version: v1.0

Owner: aisecurity.llc

Applies To: Client materials, research materials, evidence, exports, and publication files

1. Purpose

1.1 This Policy defines how long materials are retained, when they are redacted, and when they are deleted.

1.2 The policy applies across delivery, reporting, trust-center publication, and post-engagement cleanup.

2. Retention Principles

2.1 Retain only what is needed for the work, the record, or a legal obligation.

2.2 Prefer the shortest practical retention period that still supports delivery, review, dispute resolution, and publication controls.

2.3 If a file is no longer needed, it should not remain in active circulation.

3. Default Retention Classes

3.1 Provider may use the following working classes:

- working draft materials;
- active engagement materials;
- client delivery package;
- public-safe publication copy;
- archival research record; and
- delete-on-close material.

3.2 The retention period for each class should be documented in the project plan or applicable agreement when the standard default is not sufficient.

4. Redaction Rules

4.1 Before publication or wider sharing, Provider will redact:

- credentials;
- personal data;
- contact details;

- internal-only references;
- raw source payloads;
- private URLs;
- confidential pricing;
- non-public target lists; and
- other sensitive material that is not needed to understand the point.

4.2 Redaction should not remove the meaning of the document. If it would, the underlying claim should be rewritten or withheld.

5. Retention Schedule

5.1 Unless the applicable agreement states otherwise:

1. engagement work files may be retained through delivery and review;

2. client-sensitive evidence should be reviewed for deletion at engagement close;

3. public-safe publication files may be retained as part of the public record;

4. archived research records should be retained only for the period needed to support methodology continuity, claims review, or legal obligations; and

5. delete-on-close material should be securely removed when the relevant task is complete.

6. Legal Hold and Dispute Hold

6.1 If a legal hold, dispute hold, or regulatory obligation applies, the relevant materials must be preserved until the hold is released.

6.2 Materials on hold should remain access-controlled and clearly labeled.

7. Publication Copies

7.1 Publication copies must be checked for claim-readiness before release.

7.2 The publication copy should be the least sensitive version that still supports the published claim.

7.3 If a publication copy is later superseded, the older version should be archived or removed from active distribution as appropriate.

8. Deletion

8.1 Deletion should be performed using a secure, commercially reasonable process appropriate to the system storing the data.

8.2 If full cryptographic or verifiable deletion is not possible on a given system, Provider should remove access, prevent further use, and document the limitation.

8.3 Deletion records should note the material removed, the date, the method used, and any exceptions.

9. Backups and Replication

9.1 Backup systems may retain copies for operational continuity, but those copies should be governed by the same sensitivity and access principles as the primary record.

9.2 When a deletion request is honored in the primary store, backup removal should occur according to the normal backup lifecycle unless legal obligations require otherwise.

10. Exceptions

10.1 Exceptions require documented approval and a clear end date.

10.2 Exceptions should never be open-ended by default.

11. Review

11.1 Provider should review this Policy periodically and after any material change in data handling, publication workflow, or legal requirements.

APPENDIX — CONTRACT & POLICY LIBRARY

Publication & Claim-Readiness Policy

Effective Date: the Effective Date

Version: v1.0

Owner: aisecurity.llc

Applies To: Public report content, trust-center pages, sponsor materials, and related publications

1. Purpose

1.1 This Policy defines the criteria used to decide whether a statement, chart, benchmark, finding, or summary may be published as-is, published with a caveat, held internally, or withheld.

1.2 The policy protects the research, the reader, and the business from overclaiming.

2. Claim-Readiness Labels

2.1 Provider uses four practical claim-readiness values:

- `public_claim_ready`
- `public_claim_with_caveat`
- `internal_or_teaser_only`
- `do_not_claim`

2.2 These values are guidance for editorial and commercial review. They are not legal advice.

3. Public Claim Ready

3.1 A statement may be marked `public_claim_ready` when:

1. the underlying source is public-safe and can be cited or summarized without revealing restricted material;

2. the claim is supported by the project evidence or aggregate analysis;

3. the wording is accurate, balanced, and not misleading;

4. any sponsor relationship is clearly separated from the conclusion; and

5. the language can survive public scrutiny without requiring hidden context.

4. Public Claim With Caveat

4.1 A statement may be marked `public_claim_with_caveat` when the core point is sound, but the wording needs a visible limitation, source note, or method caveat to prevent misinterpretation.

4.2 Typical caveats include:

- based on analyzed job-description signals, not proof of internal maturity;

- aggregate benchmark, not company-level certification;
- directional signal, not universal truth;
- public-safe summary only;
- sponsor support does not affect findings.

5. Internal or Teaser Only

5.1 A statement may be marked `internal_or_teaser_only` when it is useful for sales, planning, or internal strategy, but not yet fit for public posting.

5.2 This category is appropriate when:

- evidence is incomplete;
- the sample is too small;
- the wording is too specific for public release;
- the artifact is still under review; or
- the statement depends on a future sample, appendix, or published asset that does not yet exist.

6. Do Not Claim

6.1 A statement must be marked `do_not_claim` when it would:

- overstate certainty;
- imply certification or endorsement;
- reveal private client details;
- expose raw evidence;
- claim maturity from weak signals; or
- create a misleading impression about a sponsor, partner, or client.

7. Review Criteria

7.1 Before publication, the reviewer should ask:

1. Is the source public-safe?

2. Is the evidence sufficient?

3. Is the wording precise?

4. Would the reader misunderstand the scope?

5. Does this create an endorsement, certification, or maturity claim we cannot support?

7.2 If the answer to any of the above is unclear, the item should be downgraded or held.

8. Sponsor Separation

8.1 Sponsor support does not influence methodology, scoring, findings, chart outputs, editorial conclusions, or publication timing.

8.2 Sponsor references must remain neutral and must not imply endorsement or validation.

8.3 Sponsor-provided copy may be used only if it is clearly labeled as sponsor perspective or sponsor material.

9. Research and Market Intelligence Claims

9.1 Job-description intelligence, public hiring signals, and aggregate benchmark outputs may be published only when the wording makes clear that they are directional and not proof of internal security maturity.

9.2 Psychometric outputs, if used, must be described as role-language evidence, not personality diagnosis or employee assessment.

10. Approval Workflow

10.1 Public claims should be reviewed by the content owner and at least one other reviewer familiar with the source evidence.

10.2 Any claim with legal, sponsor, or reputational risk should be escalated for additional review before publication.

11. Corrections

11.1 If a published claim is later found to be inaccurate, Provider will correct it promptly and, where appropriate, annotate the change.

12. Recordkeeping

12.1 Provider should retain the claim-readiness rationale for significant public claims, especially where the language is tied to research findings, sponsor materials, or buyer-facing proof.