JEARON WONG / AGENTIC LIFECYCLE GOVERNANCE INDUSTRY SERIES

Agentic AI Insurability & Risk Transfer White Paper 2026

A Lifecycle Evidence Guide for Underwriting, Claims, and Enterprise Risk Transfer

Jearon Wong - Protocol Architect for the Agent Era

Legal Subject	Risk Object	Agent Work	Evidence	Claim Review

DOCUMENT ID	VERSION	DATE
AIIRWP-2026-v1.0	v1.0 Public Research Edition	May 2026
STATUS	**SERIES**	**VISUAL SOURCE OF TRUTH**
Public research edition; repository governance seal complete with public terminology amendment	Agentic Lifecycle Governance Industry Series	HTML/PDF primary; manifest/checksum integrity

PUBLICATION BOUNDARY

Public Research Edition Status

Public research edition. Repository governance seal complete with public terminology amendment. Not legal advice, not insurance advice, not a coverage opinion, not underwriting guidance, not certification, not proof of insurability, not insurer endorsement, not regulator-approved, not a score, and not a standard. No public DOCX is authorized. The prior AIIRWP v0.2 candidate remains rejected and withdrawn; it is historical traceability context only and is not current source truth or citation source.

Publication Contents

HTML anchors are active in the public research edition. PDF rendering uses the shared whitepaper A4 print profile used by the compliance and auditability white paper artifacts.

Front Matter

Parts I-III

Parts IV-VI

Appendices

Public Research Edition Status: This public HTML artifact is the AIIRWP v1.0 public research edition. Public distribution is HTML/PDF plus manifest/checksum only; no public DOCX or source Markdown publication is authorized. Social announcement remains a separate owner decision.

Rejected v0.2 Note: The prior AIIRWP v0.2 public candidate remains rejected and withdrawn. It is historical traceability context only and is not current source truth or citation source.

Front Matter

Executive Thesis

The first insurance question for agentic AI is not whether "AI" is insured. That question is too large to answer and too vague to underwrite, broker, govern, or reconstruct after a loss.

The better question is more precise: when an AI agent participates in a loss, what exactly is being transferred? Is the insured subject the company that deployed the workflow, the professional who relied on it, the director or officer who oversaw the program, the vendor that supplied a model or platform, or another legal person named by a policy? Is the risk object the model, the output, the tool call, the business workflow, the delegated authority, the human approval event, the vendor dependency, the cyber event, or the completed work unit that produced the loss?

That distinction is the center of this paper. AI agents are usually not the insured legal subject. They are better understood, for insurance analysis, as loss-relevant parts of an agentic risk object: a bounded lifecycle work unit that can generate, shape, amplify, or obscure a loss. The insurable question is how that risk object maps back to a legal subject, a human responsibility role, a policy line, a coverage boundary, an underwriting evidence package, a claim reconstruction path, and a lifecycle evidence chain.

This is why agentic AI creates pressure that ordinary AI governance language does not resolve. A model can produce an output. A tool can execute an action. A human can approve a step. A vendor can host a service. A company can be the policyholder. A customer can suffer harm. An insurer can ask for evidence. A claim reviewer can ask what happened, who had authority, what policy line is implicated, what exclusion or sublimit may apply, and whether the event can be reconstructed. Those questions do not collapse into "AI was involved."

The public insurance market is already showing that split. Some AI-related risks are being packaged into affirmative products, including model-performance insurance, AI warranties, and AI-linked cyber coverages. Some AI risks are being bounded through form development, exclusions, endorsements, underwriting flexibility, limits, or sublimits. Some exposure remains silent inside existing lines such as cyber, E&O, Tech E&O, D&O, professional liability, media/IP, crime, and employment practices liability. Reinsurers and industry researchers are also watching accumulation, shared dependency, and systemic-loss questions around digital infrastructure and generative AI. The result is not a single market verdict. It is a mixed market that asks for sharper objects, better evidence, and cleaner boundaries. [1] [2] [3] [4] [5] [6] [7]

This paper uses that market split as the starting point. It does not begin with an internal model, a protocol claim, a maturity score, or a promise that lifecycle governance makes systems insurable. It begins with the practical problem faced by insurers, reinsurers, brokers, risk leaders, counsel, boards, finance leaders, and engineering teams: if an agentic workflow causes or contributes to a loss, can the organization identify the insured subject, the agentic risk object, the delegated authority, the human role, the dependency chain, the policy line, the relevant evidence, and the post-loss reconstruction path?

the compliance white paper and the auditability and assurance white paper matter because they help answer that evidence problem. The compliance white paper provides lifecycle governance objects: authority boundary, evidence partition, accepted outcome, substitution conformance, remediation closure, responsibility objects, and dependency visibility. The auditability and assurance white paper provides auditability concepts: an Audit Evidence Chain, auditability object, evidence request logic, and evidence sufficiency boundaries. But these papers do not make an AI system insurable. Auditability is necessary for risk transfer discussion because records matter. It is not sufficient because insurance also depends on legal subject, policy language, exclusions, limits, causation, notice, loss measurement, underwriting appetite, and claim review. [8] [9]

The role of this paper is therefore narrower and more useful than a claim that agentic AI is or is not insurable. It is a lifecycle evidence guide for reasoning about insurability. It explains what can be discussed as an insurable object, what cannot be assumed to be transferable, what evidence an insurer or risk reviewer would likely need, how compliance and auditability translate into insurance-facing reasoning, and why agentic AI needs claim reconstruction architecture before loss, not only incident response after loss.

Throughout the paper, the terms Agentic Insurability Object Model (AIO model) and Agentic Insurability Reasoning Model (AIRM) are analytical vocabulary only. They are not standards, scores, certifications, coverage triggers, underwriting requirements, pricing tools, claim approval methods, or evidence of insurer acceptance. Actual coverage, pricing, underwriting, liability, and claim outcomes remain external, policy-specific, jurisdiction-specific, and decision-specific.

The thesis can be stated simply:

An AI agent is usually not the insured legal subject. The risk-transfer problem is whether the agentic work that caused or shaped a loss can be bounded, evidenced, mapped to responsible legal and human actors, reviewed against a policy line, and reconstructed after the event.

Everything that follows is a way of making that sentence operational without overclaiming it.

Reader Map

This paper is written for readers who are facing the same problem from different sides of the table.

For insurers, the core need is object clarity. An underwriting or claim conversation cannot responsibly proceed if "AI system" means model, workflow, vendor service, customer-facing decision, human approval, API call, or all of those at once. Insurers need a way to separate the insured legal subject from the loss-generating work object, the policy line, the event trigger, the evidence set, and the reconstruction path. This paper does not provide an underwriting standard or pricing formula. It provides a vocabulary for asking cleaner questions.

For reinsurers, the central pressure is accumulation. Agentic AI can concentrate loss through shared models, cloud services, API dependencies, vendor platforms, common workflow patterns, or common failure modes. A single enterprise may see a workflow problem. A reinsurer may see a portfolio problem if many insureds depend on the same service layer or model ecosystem. This paper does not quantify capital, price systemic risk, or claim reinsurer acceptance. It explains why dependency visibility belongs in any serious insurability discussion. [7]

For brokers, the practical challenge is translating a confusing market into client language. A client may hear that AI coverage exists, that AI is excluded, that a cyber policy may respond to LLMjacking, that model warranties are available, or that D&O and E&O issues may arise. All of those statements can be partly true in different contexts. This paper gives brokers a way to explain the split without promising coverage or giving policy advice.

For CROs, the problem is risk inventory. Many enterprises can list AI tools, but cannot list the agentic work units that matter for risk transfer: what each workflow is authorized to do, what business value or customer impact it

touches, what data and tools it uses, what human role accepts the outcome, what dependencies it has, what incident history exists, and what evidence survives after loss.

For CFOs, the question is risk transfer boundary. Finance leaders may ask whether AI risk can be insured, what might affect premium, or whether evidence maturity changes the discussion. The paper answers carefully: variables such as authority scope, severity, frequency, dependency concentration, incident history, and evidence completeness may matter as analytical exposure inputs. They are not pricing formulas or premium recommendations.

For CTOs and CIOs, the paper is a warning that technical logs are not the same as insurance evidence. Logs, traces, prompts, tool calls, API records, and cloud bills can be essential. They still need to be connected to authority, responsibility, accepted outcome, dependency, remediation, and policy-line context. Engineering systems that cannot preserve those links before loss will struggle to reconstruct them after loss.

For counsel, the value is boundary discipline. The paper separates evidence architecture from legal conclusions. It does not determine liability, legal causation, coverage, exclusions, notice obligations, or claim outcomes. It provides a structure for preserving and organizing facts so that legal and coverage review can happen with fewer blind spots.

For boards, the issue is oversight. AI risk is no longer only a model-management question. It may become a governance, disclosure, cyber, professional liability, operational, customer harm, or vendor concentration question. Boards do not need another abstract AI taxonomy. They need to know what records would show who approved the agentic work, what authority was delegated, what incident escalation occurred, what remediation closed the loop, and whether the organization can explain the loss after the fact.

For AI governance leaders, this paper translates governance into risk-transfer reasoning. Policies, controls, model inventories, and review gates matter only if they can be tied to the work that created the risk. The paper keeps governance useful by asking whether each governance object survives underwriting discussion and claim reconstruction.

For engineering implementation leaders, the paper is concrete. It asks what must be recorded before something goes wrong: authority grants, tool permissions, model and vendor versions, human review criteria, accepted outcomes, incident timelines, substitution records, and remediation closure. Those records do not guarantee coverage. They make the work legible.

The shared reader problem is this: everyone can say "AI was involved." That is not enough. The paper is for readers who need to know what was involved, who owned it, what it was authorized to do, how it failed or contributed to loss, what evidence exists, and what kind of insurance question that evidence can support.

Boundary and Non-Claim Note

This paper is an analytical lifecycle evidence guide. It is not legal advice. It is not insurance advice. It is not a coverage opinion. It is not an underwriting standard. It is not actuarial pricing guidance. It is not a premium recommendation. It is not claims approval guidance. It does not determine legal liability, legal causation, policy interpretation, loss amount, claim payment, exclusion application, sublimit application, or insurer appetite.

The paper does not claim that AI is broadly insurable. It does not claim that AI is uninsurable. It does not claim that insurers accept agentic AI risk transfer as a category. It does not claim that any insurer, reinsurer, broker, regulator, standards body, or market participant has adopted the analytical vocabulary used here. It does not claim that the object model, the reasoning model, the compliance white paper, the auditability and assurance white paper, GAIC, AIAAWP, or any lifecycle evidence method makes a system coverage-ready, underwriting-ready, certified, approved, endorsed, or claim-ready.

The paper does make a narrower claim: public sources show a split insurance market around AI-related risk, and that split makes object clarity and evidence discipline more important. Some sources show affirmative AI-

related products. Some show form development and boundary tools. Some show cyber-linked AI coverage or cyber evidence needs. Some show silent exposure across traditional lines. Some show aggregation and dependency concerns. Together, those sources support a practical conclusion: agentic AI risk transfer cannot be discussed responsibly unless the insured subject, agentic risk object, event trigger, evidence chain, responsibility map, and policy line are kept separate. [1] [2] [3] [4] [5] [6] [7]

The paper uses source notes to distinguish public market evidence from analytical synthesis. Product pages are used as product examples, not as proof of market-wide acceptance. Broker and industry reports are used for market context, not for policy wording. Technical and incident-response sources are used for evidence architecture, not for insurance-market proof. the compliance white paper and the auditability and assurance white paper are used as internal source truth for lifecycle governance and auditability concepts, not as external insurance facts.

That discipline matters. Insurance is a contract and a market practice. Agentic AI is an operating pattern. A lifecycle evidence model can help the two meet, but it cannot replace policy terms, underwriting judgment, broker advice, legal review, actuarial analysis, or claim handling.

Relationship to the Compliance and Auditability White Papers

This is the third paper in an intended series, but it should not be read as a simple continuation of the first two. It changes the pressure test.

The Global AI Compliance White Paper 2026, asks how agentic and multi-agent systems become governable across lifecycle responsibility. It develops Missing Regulatory Objects, authority boundary, evidence partition, accepted outcome, substitution conformance, remediation closure, and lifecycle conformance reasoning. These concepts matter because insurance-facing evidence is rarely created at the moment of claim. It is created, or lost, during ordinary system operation. If the enterprise never records who authorized the work, which object was accepted, which dependency changed, or who closed remediation, it cannot reliably invent those facts after loss. [8]

The Agentic AI Auditability & Assurance White Paper 2026, asks how agentic systems become auditable. It develops the Audit Evidence Chain and related auditability concepts. That matters because logs alone are not enough. A log can show that an API was called. It may not show whether the call was authorized, whether the output was accepted, whether a human reviewer understood the boundary, whether a vendor dependency changed, whether an exception was escalated, or whether remediation closed the risk. [9]

this paper translates those ideas into insurability reasoning, but only after starting with the insurance market. That sequence is deliberate. If the paper began with governance objects, it would risk sounding like it is declaring insurance categories by theory. It does not. The market reality comes first: some AI risks are covered, some bounded, some silent, some cyber-linked, some warranty-like, some aggregation-sensitive, and some not yet cleanly named. Only after that reality is visible does compliance and auditability vocabulary become useful.

The relationship can be summarized this way:

Layer	What it asks	What it contributes to this paper	What it does not do
Compliance and life-cycle governance	Can the enterprise define authority, responsibility, evi-dence, accepted outcome, substitution, and remediation?	Gives agentic work a lifecycle structure that can later support risk review	Does not prove coverage or legal compliance
Auditability and assurance	Can a reviewer reconstruct evidence, sufficiency, scope, and boundary?	Gives logs and records an evidence chain rather than a pile of artifacts	Does not approve claims or make risk transferable
Insurability reasoning	Can an insurer or risk reviewer identify the insured subject, risk object, policy line, event, evidence, and reconstruction path?	Connects governance and au-ditability to underwriting and claims questions	Does not bind insurers or decide coverage

The key sentence for readers is this: auditability is necessary but not sufficient for risk transfer. It is necessary because a loss that cannot be reconstructed is difficult to discuss. It is not sufficient because even excellent evidence must still meet policy terms, legal standards, underwriting judgment, claim handling requirements, and coverage boundaries outside this paper.

How to Read This Paper

Part I establishes market reality. It asks what is being transferred, shows how the market is already splitting AI risk, and explains why model-performance or AI-linked cyber coverage is not the same as agentic lifecycle risk transfer. If you read only one part first, read Part I. It prevents the rest of the paper from becoming internal theory.

Part II defines the core object problem. It separates insured subject from agentic risk object, then shows why responsibility must be mapped across human roles, agent roles, corporate ownership, and vendor/platform dependencies. It also explains why "human in the loop" is not a responsibility structure by itself.

Part III translates compliance and auditability concepts into insurability reasoning. It shows how lifecycle governance objects and audit evidence chains can support underwriting discussion and claim reconstruction, while preserving the boundary that governance and auditability do not equal insurance.

Part IV turns the object problem into underwriting-facing evidence. It defines the evidence pack, exposure inventory, premium-relevant analytical variables, dependency concentration, and aggregation visibility needed for serious risk-transfer discussion. It does not price risk.

Part V turns the same architecture toward claims and disputes. It asks what evidence would be needed after loss: incident timeline, technical trace, authority trace, human/agent role trace, causality trace, policy boundary, and remediation closure. It does not approve claims or determine liability.

Part VI introduces the Agentic Insurability Object Model and the Agentic Insurability Reasoning Model only after the market, object, underwriting, and claim problems are established. These models are vocabulary for organizing evidence and reviewability. They are not scores, standards, or certifications.

Later appendices, if separately authorized, can make the body more usable through reference versions of the market signal register, compliance and auditability white paper mappings, reviewer evidence request structure, claim reconstruction package, premium variable dictionary, and boundary language table.

The best way to read the paper is from the outside in. Start with the market. Then define the object. Then ask what evidence survives. Then ask what underwriting and claim reconstruction can responsibly do with that evidence.

Part I: The Insurance Market Has Already Split AI Risk

Chapter 1: The New Insurance Question: What Exactly Is Being Transferred?

The meeting usually starts with the wrong question.

A business leader asks, "Are our AI agents insured?" A broker hears a version of it from a client. A CFO asks whether the new AI program changes premium. A CTO asks whether logs will be enough if something goes wrong. Counsel asks whether the vendor agreement or the company's own policy responds. A board member asks whether oversight records are adequate. Everyone wants a direct answer.

The direct answer is not available because the question is not precise enough.

Insurance does not usually begin with "the AI system." It begins with a named or described insured, a policy line, a covered event, conditions, exclusions, limits, sublimits, notice obligations, loss measurement, and claim evidence. Even before one reaches the policy text, there is a factual problem: what happened? Who did it? What was authorized? What object caused or shaped the loss? Was the event cyber, professional error, product failure, governance failure, employment practice, media/IP, crime, or something else? What evidence exists?

Agentic AI compresses those questions into one operational surface. A single workflow may involve a model, a prompt, a planning step, a tool call, an API, a vendor platform, a cloud service, a human reviewer, a business owner, a customer-impacting action, and a remediation record. After a loss, the enterprise may call that an "AI incident." The insurance file cannot stop there.

Consider a simple support workflow. A company deploys an AI agent to handle customer account requests. It can approve refunds under a threshold, update account records, trigger an external payment or CRM API, and send customer-facing notices. A customer asks for a refund after a disputed transaction. The agent reads the account history, selects a response, updates the account, triggers an external API, and sends a notice. Something goes wrong. The wrong customer receives the notice. The refund is processed twice. An account flag is changed incorrectly. A downstream partner receives an instruction that causes a service interruption. A complaint follows.

The first question is not "was AI involved?" Of course it was. That fact is too broad to settle anything. The useful questions are:

- Who is the insured subject?
- Which policy line could be implicated?
- What work object generated or shaped the loss?
- What authority was delegated to the workflow?
- Which human role reviewed, approved, supervised, or accepted the outcome?
- What model, tool, API, vendor, or cloud dependency participated?
- What event trigger matters for the policy line?
- What exclusion, limitation, or sublimit may need to be reviewed?
- What evidence exists to reconstruct the event?
- What remediation was performed and by whom?

Those questions reveal the core architecture of agentic AI insurability. The company may be the policyholder. A director or officer may be relevant if the loss turns into an oversight, disclosure, or governance question. A professional may be relevant under professional liability if the workflow supported advice or client service. A vendor or service provider may be relevant under contract, technology liability, or indemnity analysis. The AI agent may be central to the facts. But the AI agent is usually not the insured legal subject. It is not the policyholder. It is not usually the legal person whose liability is being insured. It is part of the work object that must be reconstructed.

This separation sounds simple until a loss occurs. In ordinary software incidents, the organization may identify the application, transaction, user, and failure mode. In agentic workflows, the failure may sit across a chain: a model output was plausible, a planner selected a tool, an API call executed, a human reviewer relied on a summary, a vendor service returned stale data, and the final business action caused loss. The "object" is not one component. It is the bounded work unit that connects authority, action, dependency, evidence, accepted outcome, and remediation.

That is why the phrase "AI risk transfer" can mislead. Risk transfer does not transfer a vibe. It transfers defined risks through policy language and market mechanisms. If the enterprise cannot say whether the risk object is model performance, cyber abuse, professional reliance, product underperformance, governance failure, customer harm, fraud, or vendor dependency, then it cannot have a serious conversation about transfer. It can only ask for a broad reassurance that the market is not in a position to give.

Public market sources already show this narrowing. Verisk/ISO has described form-development work around generative AI liability exposure in general liability filings, which signals that insurance infrastructure is trying to create language for boundaries rather than treating all AI exposure as one thing. [1] QBE has announced AI-focused cyber coverages and published LLMjacking guidance, which signals that AI-linked cyber risk can be addressed through a cyber lens while still requiring access, usage, and incident evidence. [2] [3] Munich Re describes AI performance insurance structures, which signals that some AI performance risk can be defined as a covered or warranty-like object. [4] These are different signals. They do not answer the same question.

If a customer refund agent causes a loss, a model-performance product might ask whether an AI model failed a defined performance promise. A cyber policy might ask whether there was unauthorized access, credential abuse, data compromise, or compute misuse. A professional liability policy might ask whether a professional service error occurred. A D&O policy might ask whether governance or disclosure failures are implicated. A general liability or product liability context might ask whether a third-party injury or property damage theory exists, subject to policy language. The agentic workflow sits across those questions, but none of them is identical to "the AI system."

This is the first principle of the paper:

The insured subject and the agentic risk object must be separated.

The insured subject is the legal person or organization whose risk may be addressed by a policy. The agentic risk object is the bounded work unit that generated, shaped, or obscured the loss. The two must connect, but they are not the same.

The second principle is that evidence must be separated from conclusion. Logs, traces, prompts, outputs, API records, and cloud usage records may be useful. They do not by themselves determine coverage, causation, liability, or claim outcome. They are inputs into reconstruction. For an agentic workflow, useful evidence must connect what happened technically to who had authority, who reviewed or accepted the outcome, which dependency participated, what policy line is implicated, and what remediation closed or failed to close the event. NIST and CISA incident response materials support the importance of structured timelines, response coordination, containment, recovery, and remediation records, but they do not turn those records into claim approval. [10] [11]

The third principle is that market signal is not market consensus. A product page, filing description, broker report, or industry research paper can show that a category is emerging or being bounded. It cannot be stretched into a universal statement that AI is covered, excluded, insurable, uninsurable, accepted, or rejected. The market is not making one decision about AI. It is sorting AI into narrower objects.

That sorting changes enterprise preparation. If a company wants to discuss risk transfer for an agentic workflow, it should not begin with a slide listing model names. It should begin with the insurance transfer question stack:

Question	What the reader must identify	Why it matters
Insured subject	Company, officer, professional, vendor, platform, or other legal subject	Insurance attaches to legal subjects and policy definitions, not loose technology labels
Risk object	Model, output, workflow, delegated action, API call, professional deliverable, governance decision, or bounded work unit	The object shapes the policy line, evidence, and causality review
Event trigger	Cyber event, professional error, product failure, governance failure, fraud, media/IP event, customer harm, or other loss type	Different lines ask different questions
Authority	What the agentic workflow was allowed to do and under whose control	Delegated authority changes exposure and responsibility analysis
Evidence	Logs, traces, approvals, prompts, tool calls, dependency records, incident timeline, remediation records	Without evidence, the event cannot be reconstructed reliably
Boundary	Exclusion, endorsement, limit, sublimit, condition, notice, or policy-specific review point	Transfer may be partial, bounded, or unavailable depending on policy terms

This table is not an underwriting standard. It is a thinking tool. Its purpose is to stop the enterprise from collapsing the whole loss into "AI." It gives the broker, insurer, counsel, risk team, and engineering team a common starting point.

The hardest part is cultural. Organizations like broad labels because they are easy to govern at the dashboard level. "AI system" is a broad label. "Agentic workflow" is a broad label. "LLM incident" is a broad label. But insurance is often forced to ask a narrower question after the fact. What was the action? Who authorized it? What line is implicated? What evidence survived? What changed after the event? What can be reconstructed?

Agentic AI makes that question urgent because the work is not only predictive. It can be delegated, tool-using, workflow-shaping, customer-facing, vendor-dependent, and partially autonomous. The final output may be less important than the path by which the work was authorized, executed, accepted, and remediated.

This is where the compliance and auditability white papers will later help. The compliance white paper gives language for lifecycle objects such as authority boundary, accepted outcome, evidence partition, substitution conformance, and remediation closure. The auditability and assurance white paper gives language for evidence chains and auditability. But those are not the opening argument. The opening argument is market-facing: if the insurance question is unclear, the enterprise has not yet defined the thing it is trying to transfer.

Once that becomes visible, the next chapter can ask what the market is doing with the ambiguity. It is not answering with one yes or no. It is covering some AI risks, excluding or bounding some, limiting some, leaving some silent, and packaging some model-performance risks in narrower products. That split is the market's way of telling enterprises that the object problem is real.

Chapter 2: What the Market Is Already Doing: Cover, Exclude, Sublimit, or Leave Silent

The AI insurance market is not a clean story. That is the point.

One buyer hears that AI insurance products exist. Another hears that insurers are writing AI exclusions. A cyber team hears that LLMjacking may be addressed through a cyber product. A professional-services firm hears that AI errors may create E&O exposure. A board hears that AI governance and cyber disclosure are becoming oversight concerns. A cloud-dependent engineering team hears that shared infrastructure creates accumulation risk. Everyone is hearing a fragment of the same larger pattern.

The market is splitting AI risk.

That split is more useful than either extreme. It is not accurate to say that AI is broadly covered. It is not accurate to say that AI is uninsurable. It is not accurate to say that all insurers are excluding AI. It is not accurate to say that affirmative AI coverage proves market acceptance of agentic AI lifecycle risk transfer. Public sources show a more practical market reality: AI risk is being sorted into cover, exclusion, endorsement, sublimit, silent exposure, cyber-linked coverage, professional liability ambiguity, governance exposure, claim reconstruction needs, aggregation concerns, and model-performance warranty structures. [1] [2] [3] [4] [5] [6] [7] [12]

That is why Chapter 2 matters. It proves that the paper is not inventing a theoretical problem. The problem is already visible in market behavior.

Imagine a broker receiving three questions in the same week.

The first client is an AI vendor that wants to offer buyers a performance warranty for its model. It asks whether underperformance against agreed metrics can be insured or backed by an insurance-linked product. Public sources from Munich Re, Armilla, and Chaucer show that product-specific AI performance and underperformance structures exist in defined contexts. [4] [5] [6]

The second client is a company whose cloud bill has spiked because stolen credentials were used to access LLM resources. It asks whether its cyber policy responds. QBE has published AI-focused cyber coverage announcements and LLMjacking materials that make this kind of AI-linked cyber risk concrete. [2] [3]

The third client is a company deploying generative AI into customer-facing and content-producing workflows. It asks whether general liability, professional liability, cyber, D&O, media/IP, or E&O policies will silently absorb the exposure. Verisk/ISO form-development activity and Aon market materials suggest that insurers and brokers are not treating these questions as settled. [1] [12] [13]

Those are not one question. They are three different insurance conversations. The fact that each contains "AI" does not make them interchangeable.

The following matrix is the evidence spine for Part I. It is intentionally framed as signal, not conclusion.

T-02-01 - AI Insurance Split-Market Signal Matrix

Market signal	What is happening	What it proves	What it does not prove	AIIRWP implication
Affirmative AI cover	Munich Re, Armilla, Chaucer/Armilla, QBE, and Beazley examples show defined AI-related products or coverage contexts for performance, warranty, liability, cyber, or cloud/AI service risk. [2] [4] [5] [6]	Some AI-related risks can be packaged for transfer or warranty-like remediation in defined settings.	Broad agentic AI insurability, claim payment certainty, policy applicability to all AI workflows, or insurer acceptance of this paper's vocabulary.	Name the covered object and event before saying "AI coverage."

Market signal	What is happening	What it proves	What it does not prove	AIIRWP implication
Exclusion / endorsement development	Verisk/ISO describes optional endorsements and rules around generative AI liability exposure; secondary reports point to named insurer exclusion activity that needs primary confirmation. [1] [14]	Insurance infrastructure is developing boundary tools for AI-related liability exposure.	Universal exclusion, AI uninsurability, final wording across all jurisdictions, or proof that every insurer is retreating.	Boundary formation shows the need to define the risk object and evidence boundary.
Sublimit / cap	Secondary reports describe AI-linked cyber caps or sublimit activity, while official QBE sources support AI-linked cyber coverage categories without giving all policy terms. [2] [15]	Some AI-linked transfer may be partial, limited, or specifically structured.	That cyber coverage equals agentic lifecycle coverage, or that exact sublimit wording is verified for all products.	Limit architecture is separate from coverage existence.
Silent AI exposure	Broker and insurer research discusses AI risk across existing lines such as cyber, E&O, Tech E&O, D&O, EPLI, crime, media/IP, and professional liability. [12] [13]	Existing insurance programs may face AI-linked ambiguity before dedicated AI products or exclusions resolve the issue.	That coverage applies, that coverage is denied, or that an insured is protected for every AI loss.	The insured subject may be covered while the agentic work object remains unclear.
Cyber-linked AI risk	QBE LLMjacking materials and AI-focused cyber announcements, plus cloud/AI service cyber examples, show compute abuse, API misuse, regulatory exposure, and cloud dependency as concrete risk channels. [2] [3] [16]	Cyber is one immediate channel for AI-linked loss and evidence reconstruction.	Complete AI insurance doctrine or lifecycle risk transfer.	Cyber records are useful, but they do not replace authority and responsibility evidence.
Professional liability ambiguity	Aon and product-specific AI liability sources point toward errors, professional reliance, technology services, and third-party liability issues. [6] [13]	AI can complicate professional service, technology, and liability analysis.	That professional liability coverage applies to every AI-assisted output.	The work unit and human acceptance path matter.
D&O / governance exposure	Aon AI risk materials, SEC cyber disclosure rules, and NAIC insurer AI governance context show governance, oversight, disclosure, and risk management relevance. [13] [17] [18]	AI risk can become a governance and oversight problem.	That D&O coverage applies or that any director/officer is liable.	Governance records must be separated from agent action and technical trace.
Claim reconstruction need	NIST, CISA, SEC, and QBE sources support timelines, incident response, coordination, containment, recovery, access logs, API records, and remediation evidence. [3] [10] [11] [17]	AI-linked losses require structured evidence after the event.	That incident response records guarantee claim approval.	Claim reconstruction requires more than final output or model trace.
Aggregation / reinsurance concern	Geneva Association and Swiss Re sources discuss cyber accumulation, generative AI risk, and cloud concentration. [7] [19]	Shared dependency and correlated loss matter to insurability analysis.	Actuarial pricing conclusions or reinsurer acceptance of agentic AI.	Agentic AI underwriting will need dependency visibility.
Model-performance warranty	Munich Re, Armilla, and Chaucer/Armilla examples focus on performance shortfall, KPI failure, model drift, hallucination, or underperformance in defined products. [4] [5] [6]	Model performance can be a defined object for transfer or warranty-like remediation.	That the entire agentic lifecycle is covered.	Model output/performance must be separated from delegated work.
Agentic lifecycle gap	The market examples mostly address model performance, cyber, product liability, professional liability, form boundaries, or silent exposure.	There is a gap between AI product coverage and end-to-end lifecycle work reconstruction.	That this paper's models are externally adopted or required.	this paper can provide analytical vocabulary for subject, object, evidence, responsibility, and reconstruction.

The first pattern in the matrix is affirmative coverage. It is real, and it matters. Munich Re's aiSure and related materials describe insurance approaches for AI performance risk and performance guarantees. Armilla describes AI insurance and warranty structures linked to performance, verification, and liability contexts. Chaucer's announced work with Armilla describes an AI third-party liability product for mechanical underperformance, hallucinations, model drift, and claims arising from underperformance. QBE has announced AI-focused cyber coverages. These examples show that market participants are not simply walking away from AI risk. [2] [4] [5] [6]

But the object matters. A model-performance warranty is not the same as a policy responding to every loss produced by an autonomous customer workflow. A cyber coverage extension is not the same as professional liability protection for AI-assisted advice. A cloud/AI service cyber product is not the same as coverage for every downstream business action triggered by an agent. The fact that a product exists tells us that a risk object can sometimes be defined. It does not tell us that all agentic risk has been made transferable.

The second pattern is exclusion and boundary formation. Verisk/ISO's general liability filing discussion is important because it shows that the market is developing optional endorsements and rules around generative AI liability exposures. [1] Whether a specific endorsement is adopted, filed, approved, or applied in a particular jurisdiction is a separate question. The safe conclusion is not that generative AI is excluded everywhere. The safe conclusion is that insurers and insurance infrastructure providers see a need for boundary tools.

That boundary formation is not a sideshow. It is the market saying: we need to know what kind of AI-related loss we are talking about. A generative AI content dispute, a product defect theory, a professional service error, a cyber incident, a customer refund error, and a board oversight failure are not interchangeable. If a policy form, endorsement, exclusion, or underwriting rule tries to draw a line, the enterprise needs evidence that can show on which side of the line its event belongs.

The third pattern is limits and sublimits. AI-linked cyber risk may be partially transferable while still being constrained. QBE's official materials support the existence of AI-focused cyber coverage categories and LLMjacking as a concrete cyber risk. Secondary reports about AI-linked cyber caps or sublimits are useful as market context, but not as primary policy wording. [2] [3] [15] The distinction matters because a sublimit is not the same as denial, and coverage existence is not the same as unlimited transfer. Limits are part of the architecture of risk transfer.

The fourth pattern is silent exposure. This may be the most common enterprise problem in the near term. Before a dedicated AI product or exclusion clearly addresses a loss, AI may enter through familiar lines: cyber, E&O, Tech E&O, D&O, EPLI, crime, professional liability, media/IP, or product liability. Aon materials frame AI as a cross-line risk issue for business leaders, not only a standalone product category. [12] [13] That does not mean coverage applies. It means the exposure may arrive through existing insurance language before the organization has built an AI-specific evidence file.

Silent exposure makes the insured subject vs risk object distinction sharper. A company may have a cyber policy. A professional firm may have professional liability coverage. A director or officer may be relevant under D&O. But the loss-generating object may be an agentic workflow, model output, tool call, API dependency, human approval event, or vendor platform. The policy might name the company. The evidence still has to reconstruct the work.

The fifth pattern is cyber-linked AI risk. LLMjacking is a useful example because it is concrete. It can involve stolen credentials, unauthorized use of LLM resources, abnormal token or compute consumption, API logs, identity records, cloud billing, containment, and remediation. QBE's LLMjacking material supports this evidence profile. [3] A cyber incident response file may include exactly the kind of artifacts that a claim reconstruction effort needs. But the same example also shows the limit of cyber framing. If the AI-linked loss is not only unauthorized access or compute abuse, but a delegated business action taken under valid credentials, cyber records may not answer the whole question.

The sixth pattern is professional liability and technology liability ambiguity. AI can sit inside advice, drafting, triage, code generation, underwriting support, claims support, financial analysis, engineering recommendations, medical workflow support, legal workflow support, customer service, or software-as-a-service delivery. If the loss is a professional error, product underperformance, technology service failure, or customer harm event, the policy-line question changes. Public product examples can show that AI liability products exist in defined settings, but they do not remove the need to ask what the professional or technology service actually did. [5] [6] [13]

The seventh pattern is governance exposure. AI risk can become a board, oversight, disclosure, control, or risk management issue. SEC cyber disclosure rules do not create insurance coverage, but they show why incident governance, materiality analysis, and board oversight records can matter for public companies. [17] NAIC insurer AI governance materials do not prove coverage for insureds, but they show that the insurance regulatory environment itself is paying attention to AI governance, risk management, controls, and third-party oversight. [18] For this paper, the implication is not that regulation equals insurance. It is that governance records can become part of the evidence environment.

The eighth pattern is claim reconstruction. NIST, CISA, QBE, and SEC sources point toward timelines, response procedures, access records, escalation, containment, remediation, and governance context. [3] [10] [11] [17] These sources are not insurance claim standards. They do not approve claims. They do show that post-loss review requires structured evidence. For agentic AI, that evidence must extend beyond ordinary incident response when the loss involves delegated authority, tool action, human acceptance, vendor dependency, or remediation closure.

The ninth pattern is aggregation. Reinsurers and insurance researchers are concerned with accumulation and correlated loss in cyber and digital infrastructure. Geneva Association work on cyber accumulation and generative AI risk, and Swiss Re work on cloud concentration, support the idea that shared dependencies matter. [7] [19] Agentic AI can intensify the same visibility problem. If many workflows depend on the same model provider, cloud service, API, orchestration tool, or vendor platform, then the enterprise and its insurers may need to understand correlation, not only isolated use cases.

The tenth pattern is model-performance warranty. This deserves its own category because it is often confused with broader AI insurance. A model-performance warranty can be a serious, useful product. It can define a performance promise, verification process, KPI, or underperformance remedy. It can make AI risk more legible in a narrow product context. But it does not automatically cover the lifecycle of agentic work. If a model performs within expected tolerance but the agent chooses the wrong tool, acts outside intended authority, relies on stale context, bypasses review, or triggers a harmful downstream action, the loss may not be a model-performance failure.

The final pattern is the agentic lifecycle gap. It is not a claim that no one covers AI. It is the gap between the objects the market is beginning to define and the work enterprises are beginning to delegate. The market has examples for model performance, warranty, cyber, technology liability, and form boundaries. Enterprises are building workflows that combine authority, planning, tool use, human review, vendor dependency, and customer impact. The gap is where this paper focuses.

What does this prove?

It proves that the insurance market is already asking narrower questions than "is AI insured?" It proves that AI-related risk can be covered in some settings, bounded in some settings, limited in some settings, and silent in some settings. It proves that evidence matters. It proves that the object being transferred must be named.

What does it not prove?

It does not prove broad agentic AI insurability. It does not prove universal exclusion. It does not prove claim payment. It does not prove that any insurer accepts the object model, the reasoning model, MPLP, or this paper's vocabulary. It does not prove that a company with good logs is coverage-ready. It does not prove that auditability equals insurability.

The responsible conclusion is more practical:

The market is splitting AI risk because AI is not one risk object. Enterprises that want risk-transfer conversations need to define the subject, object, event, evidence, boundary, and reconstruction path before the loss, not after it.

That conclusion leads directly to Chapter 3. If some AI products and cyber coverages exist, the next danger is false confidence. The reader may assume that "LLM insurance," "AI warranty," or "AI cyber coverage" solves the agentic lifecycle problem. It does not. Those products may be useful and important. They ask narrower questions.

Chapter 3: Why LLM Insurance Is Not Agentic Lifecycle Insurance

"AI insurance" is a label. It is not yet an answer.

The label can refer to several different things. It can mean model-performance insurance, where a model or AI solution fails to meet a defined performance target. It can mean a warranty or guarantee tied to AI product underperformance. It can mean third-party liability linked to model drift, hallucination, or underperformance in a defined product. It can mean AI-linked cyber coverage for LLMjacking, API misuse, cloud dependency, or regulatory costs. It can mean technology E&O, professional liability, or cyber lines encountering AI exposure without a dedicated AI product label.

Each of these categories matters. None should be dismissed. But none should be mistaken for end-to-end agentic lifecycle risk transfer.

The difference is easiest to see in a workflow example.

An enterprise uses a vendor model inside a customer-facing agent. The model produces a plausible output. The agent uses the output to select a next action. The agent calls a tool that updates a customer's account and triggers a downstream API. A human reviewer sees a short summary and approves the action. A vendor platform logs the tool call. The customer suffers a financial or service harm. The company later discovers that the model output was not obviously defective. The loss came from a combination of stale context, delegated authority, weak review criteria, and a tool action that was allowed but poorly bounded.

What question does model-performance insurance ask? It may ask whether the model or AI solution failed a defined performance promise, KPI, or agreed function. If the product is structured around underperformance, hallucination, model drift, or performance guarantee, the covered object is the model or AI product performance in that defined setting. Munich Re, Armilla, and Chaucer/Armilla sources support the existence of such product categories in public materials. [4] [5] [6]

What question does cyber coverage ask? It may ask whether there was unauthorized access, credential compromise, malicious activity, data exposure, compute abuse, business interruption, regulatory investigation, or another cyber-defined event. QBE's LLMjacking materials make the cyber version concrete: stolen access to LLM resources, abnormal usage, API logs, cloud or compute records, containment, and remediation. [2] [3]

What question does professional liability ask? It may ask whether a professional service or advice deliverable was defective, whether a human professional relied on AI inappropriately, whether a client relied on the output, and whether policy terms, exclusions, and professional standards are implicated. That is not the same as asking whether an AI model hit its KPI.

What question does agentic lifecycle risk ask? It asks how the work was authorized, planned, executed, reviewed, accepted, evidenced, changed, and remediated. It asks which legal subject owned the workflow, which human role accepted the outcome, which agent or tool performed the action, which vendor or cloud service participated, which authority boundary applied, which evidence survived, and whether the loss can be reconstructed. That is a broader operational question than model performance and a different question from cyber access.

This distinction is not a criticism of model-performance or AI cyber products. It is a boundary. A narrower product can be valuable precisely because it defines its object. If an AI vendor wants to give buyers confidence that a model will meet a specified performance promise, a model-performance warranty or insurance-backed product can address a real commercial problem. If an enterprise faces LLMjacking or AI-linked cyber abuse, cyber coverage and incident response services can address a real risk channel. The point is not that these products are weak. The point is that they are not the whole lifecycle.

Agentic lifecycle risk includes at least eight layers that model or cyber labels may not capture by themselves:

1. Authority: What was the workflow allowed to do?
2. Planning: How did the agent select a path or action?
3. Tool use: What external system, API, database, or service did it touch?
4. Human role: Who reviewed, approved, supervised, escalated, or accepted?
5. Dependency: Which model, vendor, cloud, API, or data source shaped the action?
6. Accepted outcome: What counted as done, correct, or approved?
7. Evidence: What logs, traces, approvals, records, and incident artifacts survived?
8. Remediation: What changed after the event, who owned the fix, and how was closure recorded?

These layers explain why a model can perform as expected while the agentic work still causes loss. A model can produce a plausible answer, but the workflow can apply it to the wrong account. A model can meet an accuracy benchmark, but the agent can call a tool outside the business context intended by the enterprise. A model can avoid hallucination, but a human reviewer can approve a summary without seeing the underlying evidence. A cyber policy can respond to credential theft, but not answer whether a validly authorized agent action created professional or operational loss. A warranty can address underperformance against a KPI, but not reconstruct authority, acceptance, dependency, and remediation across an enterprise workflow.

The same distinction applies to technical traces. Traces are useful. Logs are useful. Prompt records, output records, tool-call records, API logs, cloud bills, identity records, model versions, and vendor notices can be essential. It would be wrong to say they are useless. But they are not the same as claim evidence. A trace can show that a tool was called. It may not show whether the workflow had authority to call it, whether the human reviewer had the right context, whether the output was accepted under a defined criterion, whether the vendor dependency had changed, whether the policy line treats the event as cyber, E&O, Tech E&O, professional liability, product liability, crime, media/IP, or another category, or whether remediation closed the risk.

The auditability and assurance white paper's concepts become useful here, but with a boundary. An Audit Evidence Chain can help an enterprise move from raw logs to structured evidence. It can ask what object is being reviewed, what evidence was requested, what was sufficient, what was missing, and what boundary caveat applies. For claim reconstruction, that logic helps. But audit evidence is not claim approval. A technically complete evidence chain still needs policy context, legal review, coverage review, causality analysis, loss measurement, and claim handling outside this paper. [9] [10] [11]

Model-performance products also show why "covered object" is the right phrase. In those products, the object is narrowed: a model, AI solution, performance promise, KPI, underperformance event, hallucination, drift, or defined liability scenario. That narrowing makes transfer more plausible because the parties can identify what is being measured. Agentic lifecycle risk must do something similar, but with a different object: not just the model, but the bounded lifecycle work unit.

That work unit is not fully defined in Part I. Part II will do that work. For now, the key is to see why it is needed. Without a bounded work unit, agentic AI remains a blur of model, prompt, tool, vendor, human, and output. With a bounded work unit, the enterprise can begin to ask: what was delegated, what was authorized, what evidence was captured, who accepted the outcome, and what policy line might be implicated?

The market's current signals make this distinction unavoidable. Affirmative AI coverage examples show that some AI objects can be defined. Exclusion and endorsement signals show that insurers want boundary language. AI cyber examples show that some AI-linked losses can be handled through existing or extended cyber lines, but with event and evidence constraints. Silent exposure shows that existing policies may be pulled into AI losses before dedicated language catches up. Aggregation research shows that shared dependencies can turn isolated workflow risk into correlated exposure. [1] [2] [3] [4] [5] [6] [7] [12] [13] [19]

The lifecycle gap sits between those signals.

For an enterprise, the gap is practical. It may have a model inventory but no work-unit inventory. It may have logs but no authority records. It may have a human review step but no responsibility structure. It may have vendor contracts but no dependency map. It may have incident response procedures but no claim

reconstruction package. It may have strong governance language but no way to show which agentic work object caused the loss.

For an insurer or broker, the gap is equally practical. The conversation cannot stop at "AI." It must ask whether the loss is model performance, cyber, professional error, technology service failure, governance exposure, product liability, fraud, media/IP, employment practice, or something else. It must ask who is insured, what object is loss-relevant, what evidence exists, what policy language applies, and what remains unresolved.

For a reinsurer, the gap includes dependency concentration. If agentic workflows across many insureds use the same model, cloud service, API, orchestration framework, or vendor platform, then risk may accumulate through common infrastructure or common operating patterns. That is not a pricing conclusion. It is a visibility problem. [7] [19]

This chapter therefore makes a narrow but important claim:

LLM insurance, model-performance insurance, AI warranty products, and AI-linked cyber coverage are important market signals. They do not by themselves solve agentic lifecycle risk transfer.

They do not need to. Their value is that they teach the right lesson: define the object. Model-performance products define a performance object. Cyber products define a cyber event object. Agentic lifecycle risk transfer must define a work object that can connect authority, tool use, human role, dependency, accepted outcome, evidence, and remediation.

That is why Part II begins with the core object problem. Once the market reality is clear, the paper can ask the question that every later chapter depends on:

Is the insured object the company, the human, the officer, the professional, the vendor, the AI agent, or the bounded agentic work unit that created the loss?

The answer is not one object. It is a map. Part II builds that map.

Part II: The Insurable Agentic Risk Object

Part I ended with a map, not a slogan. The market does not need another broad claim that AI is insured, uninsurable, excluded, covered, risky, or revolutionary. It needs a cleaner object.

That object is not usually the AI agent itself. It is not simply the model output, the cloud event, the API trace, the workflow label, or the fact that a human clicked approve. Those records matter, but they do not by themselves say what was transferred, who owned it, what authority was delegated, what evidence survived, or what kind of policy question was created.

Part II defines the working object for the rest of the paper: the Insurable Agentic Risk Object. The phrase is an authored analytical construct. It is not an insurance standard, not a coverage trigger, not a policy definition, not a certification, and not proof that any insurer accepts the object. Its purpose is narrower and more useful. It gives insurers, brokers, reinsurers, risk leaders, counsel, boards, and engineering teams a way to describe the bounded work that might generate or shape loss.

The policyholder remains the legal insured subject. The agentic work can become the loss-relevant risk object. Part II is about keeping those layers separate long enough for underwriting discussion, incident response, and claim reconstruction to become possible.

Chapter 4: The Insurable Agentic Risk Object

The practical question after an agentic AI loss is rarely "which model was used?" That question matters, but it is not enough.

A company deploys an agent that can receive a customer request, classify the issue, approve a refund within a defined range, update the customer account, send a confirmation email, and trigger a downstream payment API. The model output looks plausible. The workflow completes. A human supervisor reviews only a summary queue. Days later, the company discovers that the agent applied the refund rule to the wrong class of customers and triggered a series of account changes that created financial loss, customer complaints, and regulatory concern.

The first insurance question is not whether "the AI" was involved. It is what object should be examined.

If the object is the model output, the analysis may miss the delegated refund authority. If the object is the cyber event, the analysis may miss that the action used valid credentials and ordinary system permissions. If the object is the runtime trace, the analysis may miss the human role and the accepted business outcome. If the object is the completed workflow, the analysis may miss the vendor dependency, substitution history, and exception record. If the object is only the company, the analysis may name the insured subject but fail to identify the work that produced the loss.

The Insurable Agentic Risk Object is the bounded lifecycle work unit that connects these fragments. It is the work as authorized, planned, executed, evidenced, handed off, accepted, excepted, remediated, and closed. It is a unit of analysis for insurance-facing review. It is not the insured legal subject.

This distinction is the hinge of this paper. The company may be the policyholder. A professional may be relevant under professional liability. An officer may matter in a governance or D&O context. A vendor or platform may matter in contract, service, technology, or liability analysis. But the agentic risk object is the bounded work that generated, shaped, amplified, or obscured the loss. It is where the insurance question becomes operational.

The object must be bounded because "AI system" is too large. It must be lifecycle-based because the loss may arise before or after the final output. It must be evidence-linked because post-loss reconstruction cannot depend on memory, dashboard summaries, or generic control claims. It must preserve responsibility because insurance conversations eventually ask who owned authority, review, acceptance, remediation, and external consequence.

The object has several minimum fields:

- initiating intent: what business purpose started the work;
- authority boundary: what the work was allowed to do, by whom, under what limit;
- agent role: what the agent drafted, recommended, selected, executed, monitored, or escalated;
- human role: who configured, supervised, approved, accepted, escalated, or remediated;
- tool action: which system, API, database, account, message, payment, code repository, or external service was touched;
- dependency chain: which model, runtime, vendor, cloud, data source, or orchestration layer shaped the work;
- evidence partition: where model, tool, human, vendor, data, project, and incident records are stored and how they can be joined;
- accepted outcome: what counted as done, correct, approved, or business-accepted;
- exception path: what happened when the work exceeded threshold, encountered uncertainty, or failed control checks;
- remediation closure: who fixed, rechecked, reauthorized, and closed the residual risk.

These fields borrow analytical vocabulary from the compliance white paper and the auditability and assurance white paper, but they are not insurance facts by themselves. The compliance white paper helps name lifecycle objects such as authority boundary, accepted outcome, evidence partition, substitution conformance, and remediation closure. The auditability and assurance white paper helps distinguish raw logs from evidence chains. In this paper, those ideas become inputs to insurability reasoning only when tied back to insured subject, policy line, loss event, and claim reconstruction. [20] [21]

This is why model output is insufficient as an insurable object. A model output can be correct while the work is wrong. The output may be applied to the wrong customer, used outside scope, executed by the wrong tool, accepted by a human without full context, or preserved without enough evidence to reconstruct the event. Model-performance products can define valuable and narrower objects, as Part I explained, but agentic lifecycle work requires a broader object.

AI cyber event framing is also insufficient by itself. Cyber coverage may be central when there is unauthorized access, credential abuse, compute theft, API misuse, business interruption, regulatory investigation, or data exposure. But an agentic loss may arise from a validly authorized business action performed through ordinary credentials. QBE's LLMjacking materials are useful because they show how access, usage, API records, containment, and remediation become evidence in a cyber-linked AI incident. They do not convert every agentic workflow loss into cyber risk. [22]

Runtime traces are insufficient for a different reason. A trace can show sequence. It can show a prompt, output, function call, API response, tool invocation, latency, token use, or model endpoint. But a trace often does not show why the work was authorized, which human role had authority, whether the action was business-accepted, whether a vendor substitution changed behavior, which evidence was withheld for privacy, or which policy line is implicated. A trace is an ingredient. It is not the meal.

Workflow completion is also insufficient. Enterprise software loves completed states: ticket closed, refund processed, email sent, pull request merged, alert resolved. Insurance review cares less about whether the workflow completed and more about whether the completed work can be reconstructed as a loss-relevant object. A workflow can complete successfully and still create the wrong legal, financial, professional, cyber, or customer consequence.

The Insurable Agentic Risk Object therefore works as a bridge. It sits between market categories and technical systems. It does not replace a policy. It helps the policy conversation ask better questions.

T-04-01 - Insurance Object Shift

Named insured, legal person, policyholder, officer, professional, vendor, or service provider	
Traditional insurance object	Named insured, legal person, policyholder, officer, professional, vendor, or service provider
AI governance object	AI system, model, control, policy, risk register, audit object
Agentic insufficiency	Names the actor or system but may not name the loss-generating work
Proposed risk object in this paper	Insurable Agentic Risk Object: bounded lifecycle work unit tied to insured subject and policy-line context
Boundary note	Analytical construct only; not a policy definition

Cyber incident, data breach, credential misuse, compute abuse, API misuse	
Traditional insurance object	Cyber incident, data breach, credential misuse, compute abuse, API misuse
AI governance object	Security event, access log, identity record, trace, cloud bill
Agentic insufficiency	May capture unauthorized activity but miss authorized harmful delegation
Proposed risk object in this paper	Authority-bounded tool-action work unit with identity, access, action, and remediation evidence
Boundary note	Does not determine cyber coverage

Professional service, advice, deliverable, technology service, product function	
Traditional insurance object	Professional service, advice, deliverable, technology service, product function
AI governance object	Human review, AI-assisted output, model response, workflow completion
Agentic insufficiency	May hide who accepted the output and how it became client- or customer-facing
Proposed risk object in this paper	Accepted-outcome work unit with human role, agent role, and external consequence
Boundary note	Does not determine professional liability

Governance event, board oversight record, disclosure issue, risk-management failure	
Traditional insurance object	Governance event, board oversight record, disclosure issue, risk-management failure
AI governance object	AI governance policy, model inventory, control attestation
Agentic insufficiency	May show governance posture without reconstructing the specific work
Proposed risk object in this paper	Responsibility-linked work unit connecting oversight, delegation, incident, and remediation records
Boundary note	Does not determine D&O liability

Claim file, incident report, notice, loss narrative	
Traditional insurance object	Claim file, incident report, notice, loss narrative
AI governance object	Log bundle, audit evidence chain, incident timeline
Agentic insufficiency	May collect artifacts without object boundaries
Proposed risk object in this paper	Claim-reconstructable work unit with source pointers and missing-evidence register
Boundary note	Does not approve claims

The table is intentionally modest. It does not say the risk object proposed here is insured. It says the object is what must be named before the insurance discussion becomes precise.

The boundary note for this chapter is simple: defining an Insurable Agentic Risk Object is not legal advice, not insurance advice, not underwriting guidance, not a coverage opinion, not certification, not proof of insurability, not insurer endorsement, and not a regulator-approved method. It is an analytical way to keep the insured subject, the work object, and the evidence path from collapsing into one vague phrase.

The next chapter asks what happens when that work object crosses hands. Agentic risk does not stay inside a single model call. It moves across humans, agents, tools, vendors, processors, projects, and remediation owners. The question becomes continuity.

Chapter 5: Agentic Risk Transfer and Responsibility Continuity

Risk does not transfer just because work moves.

In agentic systems, work moves constantly. A business user gives intent to an agent. The agent decomposes the task. A model generates text or selects a next step. A tool updates a record. A vendor service processes the request. A human approves a summary. A downstream workflow accepts the result. Another team reuses the same component in a different context. A later model substitution changes behavior. A support team remediates the incident. Each transition feels operationally normal.

For insurance analysis, every transition asks a harder question: did responsibility move with the work, or did only activity move?

Handoff is not risk transfer unless the responsibility transfer is evidenced. That does not mean every handoff needs legal ceremony. It means that post-loss reconstruction must be able to show who had authority, who understood the work boundary, who accepted the next state, what evidence moved with the task, and what remained with the prior actor.

Consider a professional services firm using an agent to draft client-facing recommendations. A consultant initiates the work. The agent retrieves prior engagement notes, uses a model to generate a recommendation, calls a document automation tool, and routes a summary for partner approval. The partner approves the summary but does not see the retrieval context or tool-action record. The document goes to the client. The client relies on it. A loss follows.

The firm may say the partner approved the work. The insurer or claims reviewer may ask a different question: approved what? The model output? The summary? The final deliverable? The retrieval context? The tool action? The client-specific suitability? The authority to send? The evidence that exceptions were absent?

That is the responsibility continuity problem.

Tool-action liability is where AI output becomes external consequence. A model can generate a recommendation without changing the world. A tool call can change an account, submit a filing, deploy code, send a payment instruction, issue a customer notice, post content, alter a medical workflow, or update an ERP record. Once the tool acts, the analysis shifts from information generation to external consequence. The risk object must preserve that shift.

Human approval is also not enough unless the human role authority is clear. A human can approve a queue item without owning the whole business decision. A reviewer can verify tone without verifying authority. A manager can approve deployment without seeing sensitive-data treatment. A professional can sign off on output without knowing that a model endpoint was substituted. The phrase "human in the loop" becomes meaningful only when the loop has role, authority, criteria, evidence, and accountability.

Vendor, model, and runtime substitution can break insurability reasoning when responsibility and evidence continuity are lost. A workflow that was reviewed under one model version may run under another. A tool connector may change. A vendor platform may modify logging behavior. A data processor may alter retention settings. A cloud or API dependency may become concentrated across many workflows. Swiss Re's cloud concentration work is useful here as an analogy: visibility into shared dependency matters because correlated infrastructure can affect many insured operations at once. [23]

Cross-project reuse creates another break. The same agentic component may be safe in one business context and risky in another. A refund classifier reused in collections, dispute handling, or compliance review now acts under a different authority scope, different customer impact, different policy line, and different evidence expectation. Reuse without reauthorization is not merely an engineering pattern. It is an exposure multiplier.

Responsibility continuity has five review questions:

1. What work moved?
2. Who or what received it?
3. What authority moved with it?
4. What evidence moved with it?
5. Who owned acceptance, exception, remediation, and closure after the move?

The compliance white paper's responsibility object and authority boundary concepts help name those states. the auditability and assurance white paper's audit evidence chain helps keep records from becoming disconnected artifacts. But again, these are analytical supports, not insurance standards. They help the risk reviewer ask whether continuity exists. They do not determine coverage, liability, claim outcome, or premium. [20] [21]

The responsibility map should include humans, agents, tools, vendors, processors, projects, and remediation owners. It should also include the absence of a responsible role. An empty cell is not a formatting problem. It is a risk signal.

T-05-01 - Responsibility Continuity Map

Human intent to agent task

Lifecycle transition	Human intent to agent task
Insurance-relevant question	Was the task authorized, bounded, and tied to an insured business activity?
Evidence needed	Initiating request, business owner, authority grant, scope limits, timestamp
Related MRO	Authority boundary; responsibility object
Failure if missing	The work cannot be tied to a responsible role or policy-line context

Agent plan to tool action

Lifecycle transition	Agent plan to tool action
Insurance-relevant question	Did the agent have permission to execute the external action it selected?
Evidence needed	Plan trace, tool permission, policy constraint, API call record, exception threshold
Related MRO	Tool-action evidence; evidence partition
Failure if missing	Output is visible but external consequence is not accountable

Tool action to business outcome

Lifecycle transition	Tool action to business outcome
Insurance-relevant question	What changed outside the model environment?
Evidence needed	Account update, payment instruction, email, filing, code deployment, customer notice, system state
Related MRO	Accepted outcome; external consequence record
Failure if missing	The loss event cannot be separated from ordinary workflow completion

Agent/human review to acceptance	
Lifecycle transition	Agent/human review to acceptance
Insurance-relevant question	What did the human approve, under what criteria, with what evidence?
Evidence needed	Review screen, approval criteria, reviewer role, visible evidence, override record
Related MRO	Accepted outcome; human responsibility role
Failure if missing	HITL becomes a comfort phrase rather than a responsibility record

Vendor/model/runtime substitution	
Lifecycle transition	Vendor/model/runtime substitution
Insurance-relevant question	Did a component change preserve authority, evidence, and expected behavior?
Evidence needed	Version history, vendor notice, conformance review, regression evidence, reauthorization
Related MRO	Substitution conformance; dependency visibility
Failure if missing	Later loss cannot be tied to the correct dependency or version

Project reuse to new context	
Lifecycle transition	Project reuse to new context
Insurance-relevant question	Was the work reauthorized for the new business scope and loss profile?
Evidence needed	Reuse approval, changed authority scope, data/context change, policy-line review marker
Related MRO	Cross-project lifecycle; authority boundary
Failure if missing	Safe behavior in one context is assumed transferable to another

Incident response to remediation closure	
Lifecycle transition	Incident response to remediation closure
Insurance-relevant question	Who owned containment, fix, recheck, residual risk, and closure?
Evidence needed	Incident timeline, containment action, fix record, retest, owner signoff, residual-risk note
Related MRO	Remediation closure; evidence partition
Failure if missing	The enterprise can describe the incident but not show closure

The map is not a liability allocation chart. It does not say which actor is legally responsible. It says which responsibility questions must be reconstructable if agentic work is to be discussed as a risk-transfer object.

The insurance significance is practical. If responsibility continuity is missing, the insurer, broker, counsel, or risk engineer may not be able to understand the event. The missing record may not be fatal in every policy context, and this paper does not decide that. But missing continuity makes the discussion harder because it turns a bounded work unit back into an undifferentiated AI system.

The boundary note for this chapter: responsibility continuity is not legal liability determination, not insurance advice, not coverage opinion, not underwriting guidance, not certification, not proof of insurability, and not insurer endorsement. It is an evidence architecture for making responsibility questions reviewable.

The next chapter turns from responsibility continuity to evidence. If the work object and responsibility handoffs can be named, what would an insurer, broker, reinsurer, or risk engineer need to review before treating the work as a more understandable risk?

Chapter 6: Underwriting Evidence Model for Agentic AI

Evidence does not guarantee coverage. Evidence does not guarantee insurability. Evidence does not guarantee a quote, a premium, an endorsement, a claim payment, or a favorable coverage position.

Evidence does something more basic: it makes the risk reviewable.

That distinction matters because enterprises often overestimate their evidence posture. A model inventory is not an underwriting evidence model. A SOC report is not a work-unit inventory. A log store is not a responsibility map. A dashboard showing agent runs is not a claim-reconstructable evidence chain. A policy document saying "human approval required" is not proof that the right human saw the right evidence at the right time.

For agentic AI, underwriting-facing evidence should be separated into four time horizons:

- pre-bind evidence: what the organization can show before coverage discussion or renewal;
- runtime evidence: what the system records while work occurs;
- post-incident evidence: what can be reconstructed after a loss or near miss;
- renewal evidence: what changed after incidents, substitutions, new workflows, or control improvements.

The categories overlap, but they should not be collapsed. Pre-bind evidence tells the reviewer what the enterprise intends and how it governs the work. Runtime evidence shows what actually happened. Post-incident evidence reconstructs event, authority, responsibility, consequence, and remediation. Renewal evidence shows whether the organization learned from experience or simply continued operating.

NIST AI RMF is useful as governance context because it frames AI risk management around governance, mapping, measuring, and managing risk. NAIC's model bulletin is useful because it shows insurance regulators thinking about insurer AI governance, risk management, controls, and third-party oversight. Those sources do not create an underwriting standard for enterprise insureds. They support the more cautious point that AI risk review increasingly requires governance, control, evidence, and third-party visibility. [24] [25]

Underwriting evidence for agentic work should include at least nine categories.

First is authority boundary. What can the agentic workflow do? What can it not do? What transaction value, customer impact, data class, tool permission, or external action requires escalation? Authority without boundaries is difficult to underwrite because the range of possible loss-generating behavior is undefined.

Second is role map. Which human role owns intent, configuration, review, acceptance, escalation, remediation, and closure? Which agent role drafts, recommends, executes, monitors, or escalates? Which vendor or platform role provides model, runtime, data, tool, storage, or logging service?

Third is tool-action record. The risk often becomes external when a tool acts. The evidence model should preserve not only model output, but the action taken: account update, API call, payment instruction, code deployment, customer message, content publication, database change, service ticket, or vendor handoff.

Fourth is evidence partition. The reviewer needs to know where records live and how they connect: model record, prompt/output record, tool record, human approval record, vendor record, privacy/redaction profile, incident record, remediation record, and missing-evidence note.

Fifth is privacy treatment. Agentic evidence can contain customer data, employee data, secrets, regulated data, trade secrets, prompts, source code, credentials, or privileged material. Evidence hoarding is not evidence maturity. A usable model must preserve what is needed while filtering or partitioning sensitive data responsibly.

Sixth is accepted outcome. The model output is not the business outcome. The evidence should show what counted as accepted, completed, delivered, filed, deployed, paid, sent, closed, or escalated.

Seventh is exception history. What thresholds were exceeded? What uncertainty was detected? What overrides occurred? What warnings were ignored? What false positives or false negatives happened? What near misses were recorded?

Eighth is remediation closure. Underwriting discussion may care not only that incidents occurred, but whether they were contained, investigated, fixed, rechecked, reauthorized, and closed. NIST and CISA incident-response sources support the value of preparation, response, recovery, remediation, tracking, and continuous improvement records. [26] [27]

Ninth is substitution conformance. If a model, tool, vendor, runtime, data source, or cloud dependency changes, the evidence should show whether the work remained within the same authority, evidence, privacy, and accepted-outcome boundaries. A silent substitution can make a previously reviewed workflow a different risk object.

Logs and traces appear in several categories, but they are not the category. A trace without authority is sequence without permission. A log without role mapping is activity without responsibility. A runtime dashboard without accepted outcome is motion without business consequence. A tool-call record without remediation history is action without closure.

T-06-01 - Underwriting Evidence Request Model

Evidence category	Insurance relevance	Example evidence artifact	Related MRO	Boundary note
Work-unit inventory	Shows what agentic work exists and what business process it touches	Work unit ID, process owner, business function, customer/financial impact flag	MRO object; cross-project lifecycle	Inventory supports review, not coverage readiness
Authority boundary	Defines delegated scope and escalation threshold	Permission matrix, transaction limits, tool scope, policy constraints	Authority boundary	Not legal delegation proof
Role map	Connects human, agent, corporate, vendor, and remediation roles	RACI, reviewer role, agent role, owner signoff, vendor support role	Responsibility object	Does not determine legal liability
Tool-action record	Shows where AI output became external consequence	API call, account update, email, payment, filing, code deployment	Tool-action evidence; accepted outcome	Does not decide policy line
Evidence partition	Shows how logs, traces, approvals, vendor records, and privacy-filtered data are joined	Evidence index, source pointers, retention map, redaction profile	Evidence partition	Not a mandatory schema
Privacy treatment	Shows evidence can be reviewed without uncontrolled sensitive-data retention	Redaction rules, data-class map, access controls, privilege flags	Evidence partition; privacy profile	Evidence hoarding is not maturity
Accepted outcome record	Shows what counted as complete, approved, or business-adopted	Approval criteria, final-state record, reviewer-visible evidence, delivery marker	Accepted outcome	Not legal acceptance or coverage proof
Exception history	Shows operational failure patterns, overrides, uncertainty, and near misses	Exception log, override register, incident/near-miss review	Exception path; remediation closure	Not a loss prediction formula
Remediation closure	Shows containment, fix, retest, reauthorization, and closure state	Incident timeline, fix record, retest evidence, owner signoff	Remediation closure	Does not prove no residual liability
Substitution conformance	Shows component changes preserved evidence and authority boundaries	Model/tool/vendor version change, conformance review, reauthorization	Substitution conformance	Not vendor certification
Dependency concentration	Shows shared model, cloud, API, vendor, or orchestration dependency	Dependency map, common service register, cross-workflow exposure list	Dependency visibility	Not actuarial pricing guidance

This model is useful because it turns "we have AI controls" into reviewable questions. What work? What authority? Which human? Which tool? Which vendor? Which evidence? Which accepted outcome? Which exception? Which closure? Which substitution?

For brokers, the model can help translate enterprise AI operations into risk-review language without promising a policy result. For insurers and reinsurers, it can help separate reviewable work from opaque automation. For CTOs and engineering leaders, it clarifies which records must be designed into the system before loss. For counsel, it keeps evidence architecture separate from legal conclusion.

The boundary note for this chapter: the Underwriting Evidence Model is not underwriting guidance, not an underwriting standard, not a premium recommendation, not actuarial pricing guidance, not coverage opinion, not insurance advice, not certification, not proof of insurability, and not insurer endorsement. It improves reviewability. It does not decide transfer.

The next chapter turns the same evidence model toward loss. When an incident occurs, the question changes from "what would a reviewer need to understand the risk?" to "what can the enterprise reconstruct now?"

Chapter 7: Claim Evidence Pack for Agentic Incidents

Incident notice is not enough.

An enterprise can notify an insurer, broker, vendor, regulator, customer, or internal executive that an AI-linked incident occurred. That notice may be important. It may be time-sensitive. It may be required by a policy, contract, regulation, or internal process. But the notice itself does not reconstruct the loss.

Agentic AI incidents require a Claim Evidence Pack: an authored analytical construct for organizing the records needed to understand the event. Like the other constructs in this paper, it is not a claims approval method, not a legal proof package, not an audit opinion, not a certification, and not a guarantee that a claim will be accepted.

Its job is to keep the event from dissolving into disconnected artifacts.

Take a small example. An agent incorrectly triggers a refund, sends a customer email, and changes a customer record. The business discovers the issue because a finance reconciliation flags an abnormal pattern. The engineering team retrieves logs. The support team sees the customer emails. The product team sees the workflow configuration. The AI team sees prompts and model outputs. The vendor has a tool-call trace. Counsel asks whether customer notice is required. The broker asks what policy line might be implicated.

Each group has a piece of the event. None has the whole object.

The Claim Evidence Pack should organize those pieces around the work unit:

- work unit ID;
- initiating intent;
- authority boundary;
- agent role;
- human role;
- tool action;
- external consequence;
- affected data;
- evidence chain;
- exception record;
- remediation action;
- closure state;
- privacy/redaction profile.

NIST and CISA sources support the value of structured incident-response processes, timelines, coordination, remediation, recovery, and tracking. QBE's LLMjacking materials show how AI-linked cyber incidents can require access, API, usage, containment, and remediation records. These sources do not define an insurance claim standard. They support the narrower point that reconstruction needs organized evidence, not only technical traces. [22] [26] [27]

The pack should answer six questions.

First: what was the work? This is the work unit ID and scope. A claim file cannot rely on "the AI made a mistake" as the object. It should identify the workflow, transaction, customer segment, project, repository, account, decision, deliverable, or tool action involved.

Second: what authority applied? The pack should show whether the agentic work was inside or outside delegated authority. Did it exceed transaction limits? Did it act on a protected data class? Did it bypass escalation? Did it use a tool permission that was technically available but not authorized for that context?

Third: who or what acted? The pack should separate agent role, human role, vendor role, model/tool role, and corporate owner. It should not assign legal fault. It should identify the roles needed for review.

Fourth: what external consequence occurred? A model output may have no loss by itself. The consequence may be account change, payment, customer message, professional deliverable, code deployment, privacy exposure, regulatory issue, service outage, publication, or other downstream effect.

Fifth: what evidence links the sequence? The pack should connect prompts, outputs, logs, tool calls, approvals, vendor records, data snapshots, exception history, incident timeline, and remediation records. It should also identify missing evidence.

Sixth: what privacy treatment applies? Claim evidence can include sensitive customer data, employee records, proprietary models, privileged communications, source code, security secrets, and vendor confidential information. A useful pack needs redaction, access controls, and source pointers, not uncontrolled data dumping.

T-07-01 - Claim Evidence Pack Components

Component	What it answers	Evidence pointer	Privacy treatment	Boundary note
Work unit ID	Which bounded agentic work created or shaped the event?	Workflow ID, run ID, ticket ID, transaction ID, repository or customer/account reference	Use scoped identifiers where possible	Analytical claim object only
Initiating intent	Why did the work begin?	User request, system trigger, scheduled job, business instruction	Redact personal data not needed for reconstruction	Does not prove authority
Authority boundary	What was the work allowed to do?	Permission matrix, delegated limit, escalation rule, policy constraint	Protect security-sensitive permissions	Not legal delegation proof
Agent role	What did the agent draft, select, execute, monitor, or escalate?	Prompt/output trace, plan record, tool call, agent state	Redact prompt content if sensitive while preserving pointer	Not insured legal subject by itself
Human role	Who reviewed, approved, accepted, overrode, or remediated?	Role record, approval record, reviewer-visible evidence, escalation note	Consider privilege and personnel data controls	Does not determine liability
Tool action	What external system or service changed?	API record, database update, email, payment, code deployment, filing	Mask secrets, credentials, and customer data	Does not determine policy line
External consequence	What loss-relevant effect occurred outside the model?	Customer impact, account change, service impact, financial record, notice record	Use data minimization and evidence index	Does not determine loss amount
Affected data	What data class was touched?	Data inventory pointer, class label, affected population, retention note	Apply redaction and access control	Not regulatory legal advice
Evidence chain	How do records connect across model, tool, human, vendor, and incident systems?	Evidence index, source pointers, timestamps, hash/checksum if used	Preserve source pointers over bulk copying when possible	Not audit opinion or claim approval
Exception record	What warning, threshold, override, or failure occurred?	Exception log, override note, near-miss record, alert history	Protect sensitive security/employee details	Does not prove causation
Remediation action	What was contained, fixed, rechecked, and reauthorized?	Incident ticket, patch, rollback, model/tool change, retest record	Separate operational secrets from review evidence	Does not prove no residual liability
Closure state	Who accepted closure and residual risk?	Owner signoff, closure note, residual-risk entry, follow-up task	Limit distribution of privileged or sensitive notes	Not settlement or claim outcome

The Claim Evidence Pack should include a missing-evidence register. Missing evidence is not a moral failure; it is a fact that must be visible. The pack should identify which records are absent, stale, overwritten, inaccessible, vendor-controlled, privileged, redacted, or outside retention. That register may be as important as the records that exist because it tells reviewers what cannot be reconstructed.

The pack should also distinguish technical causality from legal causation. A technical sequence can show that an agent called an API before a customer account changed. It cannot, by itself, decide proximate cause,

liability, coverage, exclusion, damages, or claim outcome. That boundary protects the paper from pretending that evidence architecture replaces legal and insurance review.

The Claim Evidence Pack is useful before the claim too. If an enterprise designs the pack only after an incident, the most important records may already be missing. The pack is therefore an incident-readiness design pattern as much as a post-loss organizing tool. It asks engineering, risk, legal, and business teams to preserve the records that future review will need.

The boundary note for this chapter: a Claim Evidence Pack is not claims approval guidance, not legal proof, not legal advice, not insurance advice, not a coverage opinion, not certification, not proof of insurability, and not insurer endorsement. It is a structured way to preserve and organize evidence for review.

The next chapter defines the negative space. If the risk object cannot be bounded, the responsibility cannot be traced, the evidence cannot be reconstructed, or remediation cannot be closed, then the agentic work may remain hard to discuss as a transferable risk.

Chapter 8: Uninsurable or Hard-to-Insure Agentic Risk Patterns

The most useful insurability architecture is honest about what it cannot carry.

This chapter does not declare final insurer positions. It does not say a system is legally uninsurable. It does not create underwriting rules. It identifies risk-readiness patterns that make agentic work difficult to review, difficult to reconstruct, or difficult to discuss as a transferable risk object.

The word "hard-to-insure" is safer than "uninsurable" for most of this discussion because insurance outcomes depend on policy language, market appetite, jurisdiction, insured profile, loss type, limits, exclusions, underwriting judgment, and claim review. Still, some patterns are so weak from an evidence and responsibility perspective that an enterprise should treat them as severe blockers for serious risk-transfer discussion.

The first pattern is fully opaque agentic execution. If the enterprise cannot identify what the agent did, which model or tool it used, which data it touched, or why it selected a path, the risk object is not reviewable. A black box may be commercially convenient, but insurance review needs some path from event to evidence.

The second pattern is no delegated authority boundary. If an agent can draft, decide, execute, message, transact, or deploy without a recorded scope, the range of possible consequences is undefined. The problem is not autonomy by itself. The problem is unbounded autonomy without authority evidence.

The third pattern is no human-role-to-agent-responsibility map. If the enterprise says "a human reviewed it" but cannot show who reviewed what, under what authority, using what criteria, with what visible evidence, the human step may not reduce ambiguity. It may add it.

The fourth pattern is no accepted outcome state. If the enterprise cannot show when an output became an accepted business action, it cannot separate generation from adoption. That matters for customer communications, professional deliverables, payments, filings, code deployments, account changes, and operational decisions.

The fifth pattern is no tool-action liability boundary. The agent may use tools that create external consequence, but the organization has not separated suggestion from action. It cannot tell when text became transaction, when recommendation became instruction, or when internal output became customer-facing effect.

The sixth pattern is broken or non-reconstructable evidence chain. Logs exist, but they are not joined. Traces exist, but they omit human acceptance. Vendor records exist, but are inaccessible. Retention is too short. Privacy controls erased needed pointers. Incident records do not connect to model or tool records. The result is evidence fog.

The seventh pattern is cross-project reuse without reauthorization. A component built for one workflow is reused in another business context with different authority, data, customer impact, policy-line exposure, or dependency concentration. The object looks familiar to engineers but different to risk reviewers.

The eighth pattern is vendor, runtime, or model substitution without conformance review. If the behavior of the work changes because a model endpoint, orchestration layer, tool connector, data source, or vendor platform changes, the enterprise needs evidence that the risk object remained within its approved boundary. Without that evidence, the prior review may not travel.

The ninth pattern is privacy evidence hoarding or uncontrolled sensitive-data trace retention. Insurability reasoning needs evidence, but evidence maturity is not the same as keeping every prompt, customer detail, credential, secret, privileged communication, or employee record forever. A serious evidence model must handle redaction, minimization, access control, privilege, and source pointers.

The tenth pattern is no dispute or remediation closure. The enterprise notices a failure, patches something, and moves on. Months later it cannot show who owned the fix, whether it was retested, whether residual risk was

accepted, whether the workflow was reauthorized, or whether the customer impact was closed. The loss may be over operationally, but not reconstructable.

These patterns often appear together. Fully opaque execution tends to create weak evidence chains. Weak evidence chains make human responsibility harder to show. Weak responsibility makes tool-action consequences harder to explain. Tool-action ambiguity makes claim reconstruction harder. Poor remediation closure makes renewal evidence weaker.

T-08-01 - Hard-to-Insure Agentic Risk Patterns

Fully opaque agentic execution

Pattern	Fully opaque agentic execution
Why it is hard to insure	The reviewer cannot identify what happened or why
Missing object	Bounded work unit
Evidence failure	Missing plan, action, model/tool, and dependency records
Possible remediation path	Add work-unit IDs, action traces, dependency pointers, and exception capture

No delegated authority boundary

Pattern	No delegated authority boundary
Why it is hard to insure	The possible loss perimeter is undefined
Missing object	Authority-bounded task
Evidence failure	No permission scope, escalation threshold, or transaction limit
Possible remediation path	Define authority grants, limits, expirations, and escalation rules

No human-role-to-agent-responsibility map

Pattern	No human-role-to-agent-responsibility map
Why it is hard to insure	HITL cannot be evaluated as responsibility
Missing object	Responsibility bridge
Evidence failure	Approval record lacks role, criteria, and visible evidence
Possible remediation path	Map intent, review, acceptance, escalation, remediation, and closure roles

No accepted outcome state

Pattern	No accepted outcome state
Why it is hard to insure	The enterprise cannot show when output became business action
Missing object	Accepted outcome
Evidence failure	Completion state is not tied to business acceptance
Possible remediation path	Record acceptance criteria, reviewer-visible evidence, and final-state marker

No tool-action liability boundary

Pattern	No tool-action liability boundary
Why it is hard to insure	External consequence is not separated from generated output
Missing object	Tool-action object
Evidence failure	API/payment/email/code/account action not linked to authority and outcome
Possible remediation path	Log tool actions with authority, role, and external consequence pointer

Broken evidence chain

Pattern	Broken evidence chain
Why it is hard to insure	Records exist but cannot reconstruct the event
Missing object	Evidence chain
Evidence failure	Logs, traces, approvals, vendor records, and incident records are disconnected
Possible remediation path	Build evidence index, source pointers, retention map, and missing-evidence register

Cross-project reuse without reauthorization

Pattern	Cross-project reuse without reauthorization
Why it is hard to insure	Prior review may not match new scope or loss profile
Missing object	Reauthorized work object
Evidence failure	New project context lacks authority and policy-line review marker
Possible remediation path	Require reuse review for authority, data, dependency, and external consequence changes

Vendor/runtime/model substitution without conformance review

Pattern	Vendor/runtime/model substitution without conformance review
Why it is hard to insure	The reviewed object may have changed silently
Missing object	Substitution-conformant dependency
Evidence failure	Version, endpoint, vendor, or tool change lacks reauthorization evidence
Possible remediation path	Record substitution, regression check, evidence continuity, and owner signoff

Privacy evidence hoarding

Pattern	Privacy evidence hoarding
Why it is hard to insure	Evidence creates separate privacy/security/legal exposure
Missing object	Privacy-filtered evidence object
Evidence failure	Excess sensitive trace retention or uncontrolled access
Possible remediation path	Use redaction, minimization, access control, privilege flags, and source pointers

No dispute/remediation closure	
Pattern	No dispute/remediation closure
Why it is hard to insure	The event cannot be closed as a reviewable lifecycle record
Missing object	Remediation closure
Evidence failure	Fix, retest, residual risk, and owner signoff are absent
Possible remediation path	Create closure checklist, retest record, residual-risk note, and reauthorization marker

The possible remediation paths in the table are not underwriting requirements. They are engineering, governance, and evidence design moves that can make the risk object more reviewable. Whether any insurer values them, requires them, discounts for them, excludes without them, or accepts them remains external.

This negative space is important because it prevents this paper from becoming a confidence machine. The paper is not here to say every agentic workflow can be insured if the enterprise fills out the right template. Some work may remain too opaque, too unbounded, too poorly evidenced, too privacy-invasive, too dependent on unreviewable vendors, or too hard to reconstruct.

The point is not pessimism. It is precision. A hard-to-insure pattern is often a design problem that can be surfaced earlier. If an enterprise identifies the missing object before deployment, it may be able to redesign authority, evidence, privacy, responsibility, substitution, and remediation. If it discovers the gap only after a loss, it may have fewer options.

The boundary note for this chapter: the hard-to-insure patterns are an analytical risk-readiness view, not final insurer position, not underwriting rule, not legal advice, not insurance advice, not coverage opinion, not certification, not proof of insurability, not proof of non-insurability, not insurer endorsement, and not regulator-approved method.

Part II has now built the object layer. Part I showed that the market is already splitting AI risk. Part II explains why the object has to be bounded, responsibility-linked, evidence-backed, claim-reconstructable, and honest about negative space. Part III can now return to the compliance white paper and the auditability and assurance white paper without confusing them for insurance facts: lifecycle governance and auditability matter because they help create reviewable objects, not because they make systems insurable by themselves.

Part III: From Lifecycle Governance and Auditability to Insurability Reasoning

Part I established the market reality: AI risk is not moving through insurance in one clean line. Some risk is affirmatively addressed. Some is bounded. Some is silent inside existing lines. Some belongs closer to model performance or cyber than to the full agentic lifecycle.

Part II then named the object problem. The policyholder remains the insured legal subject. The agentic work can become the loss-relevant risk object. That object must be bounded, responsibility-linked, evidence-backed, claim-reconstructable, and honest about negative space.

Part III translates the first two white papers into this insurance-facing layer.

The compliance white paper gives lifecycle governance vocabulary. It names missing regulatory objects, authority boundaries, responsibility states, accepted outcomes, evidence partitions, privacy constraints, substitution conformance, and remediation closure. The auditability and assurance white paper gives auditability vocabulary. It distinguishes raw logs from evidence chains, trace visibility from audit evidence, and audit readiness from professional assurance. Those foundations matter for this paper because they help describe the risk object. They do not create coverage, determine liability, approve claims, bind insurers, set underwriting rules, or make a system insurable.

This part is therefore a translation exercise. It asks how lifecycle governance and auditability become useful to insurability reasoning without being mistaken for insurance facts.

Chapter 9: From Lifecycle Governance Objects to Insurability Objects

Many enterprises now have AI governance language. They can describe models, controls, inventories, policies, risk committees, approval workflows, and sometimes even agentic use cases. After a loss, however, the insurance question often arrives in a different grammar.

What was the insured subject? What object generated or shaped the loss? What authority was delegated? What action crossed into the world? What evidence survived? Which policy line might be implicated? Which records show responsibility, acceptance, exception, and remediation?

The compliance white paper helps because it names lifecycle objects that ordinary governance language often leaves vague. Missing Regulatory Objects, or MROs, are not insurance objects in themselves. They are not coverage triggers, policy definitions, underwriting standards, legal proof, or insurer-endorsed categories. Their value in this paper is more precise: they help identify what must be bounded, evidenced, reviewed, transferred, or closed before agentic work can be discussed as a reviewable risk object. [28]

Consider a customer-operations agent that can classify requests, approve account adjustments, send customer messages, and trigger an external service ticket. The enterprise may have an AI policy, a model inventory, and a control statement saying that humans supervise the system. A risk reviewer will still need a sharper map. Who gave the agent authority to adjust accounts? Which human role accepted the outcome? Was the customer message treated as a draft, recommendation, or authorized communication? Did the external service ticket create a downstream obligation? If the model or runtime changed, did the prior review still apply? If customer data appears in traces, what evidence is retained, redacted, or minimized?

Those are MRO-shaped questions.

The first insurance-relevant MRO is the human-role-to-MAS-responsibility mapping. Agentic systems often make work look shared. A business owner initiates. A system designer configures. A model generates. An agent selects a tool. A reviewer approves. A vendor hosts. A support team remediates. The insurance problem is not that many roles exist. It is that the roles may not connect to responsibility. A role map does not determine legal liability, but without it the work unit can become difficult to reconstruct.

The second is the delegated authority boundary. Insurance analysis needs an exposure perimeter. If the agent can draft but not send, recommend but not execute, approve below a threshold but escalate above it, use one data class but not another, or call one API but not another, that scope matters. Authority boundary is not legal delegation proof. It is a record of what the agentic work was allowed to do and when escalation should have occurred.

The third is the distinction between agent role and human role. An agent role is not a human role wearing technical clothing. A human may own intent, professional judgment, business acceptance, escalation, remediation, or closure. An agent may draft, route, classify, retrieve, score, execute, monitor, or trigger. The same work unit can involve both. If the enterprise collapses those layers into "AI-assisted," the insurance object becomes blurry.

The fourth is accepted outcome compliance. Model output is not the same as business acceptance. A recommendation becomes risk-relevant when it is sent, filed, paid, deployed, relied on, published, merged, recorded, or otherwise adopted. Accepted outcome records help show when generated content became organizational action. They do not decide legal acceptance, policy compliance, or coverage.

The fifth is the tool-action liability boundary. This is where the agentic lifecycle often becomes external. A generated answer may be internal. A tool action can change a customer record, transmit funds, send an email, alter code, open a ticket, submit a form, or call a vendor API. The tool-action boundary helps separate suggestion from consequence.

The sixth is responsibility transfer across agents. Agentic work may pass from one agent to another, from an agent to a workflow platform, from a workflow to a vendor, or from a vendor to a human team. The insurance-relevant question is whether the responsibility state traveled with the work. Handoff is operational. Responsibility continuity is evidentiary.

The seventh is authority drift. A workflow that begins with a narrow authority can widen through reuse, configuration change, prompt update, tool permission expansion, or role confusion. Drift matters because the risk object originally reviewed may no longer be the object that caused loss.

The eighth is evidence partitioning. Agentic evidence is scattered by design: model records, prompts, tool calls, human approvals, vendor logs, cloud records, data labels, incident tickets, remediation notes, and privacy redactions may live in different systems. An evidence partition tells the reviewer where the pieces are and how they can be joined without pretending that a single log stream is enough.

The ninth is cross-project reuse compliance. Agentic components rarely stay in one place. A classifier built for support may be reused in billing. A code agent built for internal tooling may be reused in customer-facing deployment. A retrieval agent built for policy search may be reused in professional advice. Reuse changes exposure when authority, data, external consequence, or policy-line ambiguity changes.

The tenth through thirteenth MRO families sit around privacy: lifecycle privacy mapping, privacy-preserving third-party validation, evidence minimization and selective disclosure, and data subject rights versus evidence retention. These matter because insurance review needs evidence, while privacy and security disciplines limit what should be retained, copied, disclosed, or shared. Evidence maturity is not evidence hoarding.

The fourteenth is the third-party processor and subprocessor chain. Agentic systems often depend on vendors, cloud services, model providers, orchestration layers, data processors, and logging tools. The insured organization may not control every record needed for reconstruction. The processor chain is therefore not a procurement appendix. It is part of the evidence object.

The fifteenth is vendor, model, and runtime substitution conformance. A change in model endpoint, vendor platform, tool connector, runtime policy, embedding store, or API behavior can make a previously reviewed work unit behave differently. The insurance concern is not simply that change occurred. It is whether authority, evidence, privacy, and accepted-outcome boundaries remained continuous.

The sixteenth is incident, dispute, and remediation closure. A loss event does not end when a ticket is closed. Reviewers may need to know what was contained, what was fixed, who rechecked it, whether the workflow was reauthorized, what residual risk remained, and whether affected parties or business processes were closed out.

NIST AI RMF and NAIC insurer-governance materials are useful context here because they show the broader movement toward governance, mapping, risk management, documentation, controls, validation, and third-party oversight. They do not turn MROs into insurance standards, and they do not prove coverage or underwriting acceptance. [29] [30]

T-09-01 - MRO-to-Insurability Translation Map

Human Role to MAS Responsibility Mapping

MRO	Human Role to MAS Responsibility Mapping
Insurance-relevant question	Who owned intent, review, acceptance, escalation, remediation, and closure?
Evidence or control needed	Role map, RACI, approval record, owner signoff
Risk-transfer relevance	Makes responsibility questions reviewable
Boundary note	Does not determine legal liability

Delegated Authority Boundary

MRO	Delegated Authority Boundary
Insurance-relevant question	What was the agentic work allowed to do?
Evidence or control needed	Permission scope, transaction limits, data class limits, escalation rules
Risk-transfer relevance	Defines exposure perimeter
Boundary note	Not legal delegation proof

Agent Role is not Human Role

MRO	Agent Role is not Human Role
Insurance-relevant question	What did the agent do versus what the human owned?
Evidence or control needed	Agent task record, human approval criteria, reviewer-visible evidence
Risk-transfer relevance	Prevents HITL from becoming a vague control claim
Boundary note	Not an employment, agency, or liability conclusion

Accepted Outcome Compliance

MRO	Accepted Outcome Compliance
Insurance-relevant question	When did output become organizational action?
Evidence or control needed	Final-state marker, delivery record, approval criteria, acceptance timestamp
Risk-transfer relevance	Separates generation from business adoption
Boundary note	Not coverage proof or legal acceptance

Tool-Action Liability Boundary

MRO	Tool-Action Liability Boundary
Insurance-relevant question	Where did a recommendation become external consequence?
Evidence or control needed	API call, account update, email, filing, payment, deployment record
Risk-transfer relevance	Identifies loss-relevant action point
Boundary note	Does not decide policy line

Responsibility Transfer Across Agents

MRO	Responsibility Transfer Across Agents
Insurance-relevant question	Did responsibility move with the work or only activity?
Evidence or control needed	Handoff record, receiving role, evidence transfer, exception state
Risk-transfer relevance	Supports reconstruction across agent chains
Boundary note	Not risk transfer by itself

Authority Drift

MRO	Authority Drift
Insurance-relevant question	Did scope widen after review?
Evidence or control needed	Permission change log, reuse review, prompt/config change record
Risk-transfer relevance	Shows whether the reviewed object changed
Boundary note	Not a coverage trigger

MAS Evidence Partitioning

MRO	MAS Evidence Partitioning
Insurance-relevant question	Where are model, tool, human, vendor, incident, and privacy records?
Evidence or control needed	Evidence index, source pointers, retention map, missing-evidence note
Risk-transfer relevance	Converts scattered records into a reviewable object
Boundary note	Not a mandatory schema

Cross-Project Reuse Compliance

MRO	Cross-Project Reuse Compliance
Insurance-relevant question	Was the component reauthorized for the new context?
Evidence or control needed	Reuse approval, changed data/authority map, policy-line marker
Risk-transfer relevance	Controls exposure drift across business contexts
Boundary note	Not procurement guidance

Privacy / GDPR Lifecycle Mapping

MRO	Privacy / GDPR Lifecycle Mapping
Insurance-relevant question	What personal or regulated data appears in the work and evidence?
Evidence or control needed	Data-class labels, processing map, retention notes
Risk-transfer relevance	Keeps evidence review from creating unmanaged privacy exposure
Boundary note	Not GDPR or legal advice

Privacy-Preserving Third-Party Validation

MRO	Privacy-Preserving Third-Party Validation
Insurance-relevant question	Can evidence be reviewed without excessive disclosure?
Evidence or control needed	Selective disclosure, access control, redaction, trusted review channel
Risk-transfer relevance	Supports reviewability where vendor or personal data is sensitive
Boundary note	Not assurance certification

Evidence Minimization and Selective Disclosure

MRO	Evidence Minimization and Selective Disclosure
Insurance-relevant question	What is needed, what can be pointed to, and what should not be copied?
Evidence or control needed	Source pointer, redaction profile, privilege flag, minimization rule
Risk-transfer relevance	Balances reconstruction with privacy/security
Boundary note	Not a legal sufficiency rule

Data Subject Rights vs Evidence Retention	
MRO	Data Subject Rights vs Evidence Retention
Insurance-relevant question	Can retention support review without ignoring data rights?
Evidence or control needed	Retention schedule, deletion exception process, access log
Risk-transfer relevance	Flags tension before incidents occur
Boundary note	Not privacy compliance advice

Third-Party Processor / Subprocessor Chain	
MRO	Third-Party Processor / Subprocessor Chain
Insurance-relevant question	Which outside parties hold records or shape behavior?
Evidence or control needed	Vendor map, subprocessor list, contract pointer, logging access terms
Risk-transfer relevance	Shows dependency and evidence custody
Boundary note	Not vendor ranking

Vendor / Model / Runtime Substitution Conformance	
MRO	Vendor / Model / Runtime Substitution Conformance
Insurance-relevant question	Did component change preserve the reviewed boundary?
Evidence or control needed	Version change, conformance review, regression record, reauthorization
Risk-transfer relevance	Protects continuity of the risk object
Boundary note	Not vendor certification

Incident, Dispute, and Remediation Closure	
MRO	Incident, Dispute, and Remediation Closure
Insurance-relevant question	What ended, what remained, and who accepted closure?
Evidence or control needed	Incident timeline, fix record, retest, residual-risk note, closure signoff
Risk-transfer relevance	Supports post-loss and renewal review
Boundary note	Not claim approval or settlement guidance

The table should be read as a translation map, not a scoring model. It says that MROs can help describe the work object that insurance analysis needs to see. It does not say that the presence of an MRO creates coverage, satisfies an insurer, proves compliance, resolves causation, or makes the risk transferable.

The boundary for this chapter is deliberately narrow: The compliance white paper provides lifecycle governance object vocabulary for insurability reasoning. It does not provide legal advice, insurance advice,

underwriting guidance, a coverage opinion, certification, proof of insurability, insurer endorsement, or a regulator-approved method.

The next chapter moves from the compliance white paper's governance objects to the auditability and assurance white paper's auditability architecture. If MROs help name the object, audit evidence helps reconstruct what happened. But auditability still does not equal insurability.

Chapter 10: From Audit Evidence Chain to Claim Reconstruction

After an incident, teams often say the same thing: "We have logs."

Sometimes they do. They have model traces, API records, prompts, outputs, identity logs, cloud usage records, tool-call histories, approval timestamps, service tickets, alerts, and vendor dashboards. But a pile of logs is not a claim reconstruction. It may not even be an evidence chain.

the auditability and assurance white paper matters because it separated raw trace visibility from audit evidence. It introduced the idea that evidence must be linked to an object, a responsibility state, a request, a sufficiency boundary, and a review purpose. This paper uses that discipline for a different purpose: claim reconstruction. [31]

The difference is important. An audit evidence chain can help organize facts. It can show what work unit was reviewed, what evidence was requested, what records were available, and what gaps existed. A claim reconstruction effort may use the same records, but it asks additional questions: what loss occurred, which insured subject is involved, which policy line may be implicated, what event triggered notice, what evidence links action to consequence, what exclusions or limits may be relevant, what remediation occurred, and what remains unresolved?

Auditability is necessary for reconstruction because unreconstructable systems make post-loss review harder. It is insufficient for insurability because insurance decisions depend on policy language, underwriting appetite, line of coverage, loss facts, legal analysis, exclusions, limits, notice, causation, damages, and claim handling. The auditability and assurance white paper helps make the facts legible. It does not decide the insurance outcome.

Consider an agentic code-deployment workflow. A developer asks an agent to propose a patch. The agent retrieves context, writes code, runs tests, opens a pull request, and recommends deployment. A human reviewer approves a summary. The deployment tool pushes the change. A service outage follows.

The engineering logs may show the commit, deployment time, test output, and rollback. The agent traces may show the prompt, retrieved files, tool calls, and recommendation. The ticketing system may show the approval. The incident system may show detection, escalation, containment, recovery, and closure. A claim reconstruction needs all of that, but it also needs responsibility semantics. Was the agent authorized to modify the affected service? What did the reviewer see? Was the deployment inside an approved change window? Did a vendor or model substitution alter behavior? Was customer data exposed? What loss category is being asserted? What remediation closed the event?

NIST and CISA incident-response materials support the value of preparation, detection, coordination, containment, remediation, recovery, reporting, tracking, and continuous improvement. They are not insurance claim rules. They support the narrower point that reconstruction depends on organized, time-linked, role-aware records. [32] [33]

The translation from auditability evidence to risk-transfer analysis has several layers.

The work unit becomes the claim reconstruction object. The reviewer cannot reconstruct "AI failure" in the abstract. The event needs a bounded work unit: deployment run, customer-account update, refund approval, professional deliverable, payment instruction, email campaign, investigation workflow, or vendor handoff.

Authority becomes a reconstruction question. Was the work inside delegated scope? Did it exceed a threshold? Did it use a tool permission that existed technically but was not authorized for the context? Was escalation required?

Role becomes responsibility context. Who initiated, configured, reviewed, accepted, overrode, remediated, and closed the work? What did the agent do? What did the human do? What did the vendor do? What did the

organization own?

Tool action becomes the external consequence bridge. The claim file may need to show how an output became an account change, payment, code deployment, customer message, data transfer, filing, or service interruption.

Evidence pointer becomes the opposite of evidence dumping. The evidence chain should point to source records, retention status, redaction profile, access controls, and missing evidence. In many cases the reviewer needs proof that a record exists and can be accessed under appropriate controls, not a bulk export of every prompt and payload.

Accepted outcome becomes the moment of adoption. A model answer may be tentative. A business state may be final. Claim reconstruction needs to know when the organization accepted the action and what criteria were visible.

Exception becomes the early warning record. Thresholds, overrides, uncertainty, alerts, failed checks, near misses, and ignored warnings may explain why the event occurred or why it was not contained earlier.

Remediation closure becomes the post-loss state. The record should show what was contained, fixed, retested, reauthorized, monitored, or left open. It should also show who accepted residual risk.

Privacy treatment becomes part of evidence quality. A claim evidence chain that exposes unnecessary personal data, secrets, privileged materials, or vendor confidential information can create a second risk problem. Evidence must be reconstructable and controlled.

T-10-01 - Auditability-to-Claim-Reconstruction Crosswalk

the auditability and assurance white paper auditability concept	Claim reconstruction use	What it can show	What it cannot show	Boundary note
Agentic Audit Object	Identifies the bounded work that should be examined	Work unit, system boundary, evidence scope	Whether the object is covered or legally responsible	Analytical object only
Audit Evidence Chain	Organizes records across model, tool, human, vendor, and incident systems	Sequence, source pointers, gaps, reviewable record path	Claim approval, causation, damages, settlement, or coverage	Evidence chain is not claim outcome
Evidence Request	Helps define records needed after loss	Prompt/output, tool action, role, authority, incident, remediation records	Formal claim demand or insurer-required checklist	Not claims approval guidance
AARM Readiness Vocabulary	Helps discuss whether facts are observable, trace-linked, or evidence-structured	Evidence maturity and reconstruction gap	Underwriting score, certification, or insurer-adopted model	Not an underwriting standard
Trace vs Evidence Distinction	Prevents raw logs from being overclaimed	Technical sequence and available artifacts	Responsibility, authority, accepted outcome, policy fit	Logs are ingredients only
Evidence Sufficiency Boundary	Makes missing evidence visible	What is present, missing, stale, overwritten, vendor-held, or redacted	Legal sufficiency or proof burden	Not legal advice
Exception and Remediation Records	Links failure, containment, fix, retest, and closure	Operational response and post-loss state	No residual liability or guaranteed recovery	Closure does not decide liability
Selective Disclosure / Privacy Controls	Makes review possible without uncontrolled data exposure	Redaction, source pointers, access control, data labels	GDPR compliance or privilege decision	Not privacy legal advice

This crosswalk should prevent two mistakes.

The first mistake is technical overconfidence. A trace can show that a tool was called. It may not show whether the tool call was authorized, whether the human reviewer saw the relevant context, whether a vendor substitution changed behavior, whether the action became an accepted business outcome, or whether the affected data should have been retained or redacted.

The second mistake is insurance overconfidence. A strong evidence chain can make a claim file more coherent. It cannot decide causation, liability, coverage, exclusions, limits, damages, settlement, or claim payment. Those decisions sit outside this paper.

AI-linked cyber examples show why this matters. QBE's LLMjacking guidance points toward access, API usage, abnormal consumption, containment, and remediation records. Those records can be essential for reconstructing a cyber-linked AI event. They still do not answer every agentic lifecycle question: who authorized the workflow, what business action was accepted, what responsibility bridge existed, and what line ambiguity remains? [34]

The boundary for this chapter: The auditability and assurance white paper provides auditability and evidence-chain vocabulary that can support claim reconstruction. It does not provide legal advice, insurance advice, underwriting guidance, coverage opinion, certification, proof of insurability, claim approval guidance, legal liability determination, insurer endorsement, or a regulator-approved method.

The next chapter turns from evidence architecture to policy-line ambiguity. The same agentic event may be legible as cyber, professional liability, technology E&O, governance, crime, media, employment, product, or business interruption exposure. Evidence can help sort the questions, but it does not resolve coverage.

Chapter 11: Insurance Lines and Agentic Risk Ambiguity

An agentic incident rarely arrives wearing one policy label.

A customer-support agent triggers refunds to the wrong accounts. Is that a cyber event because credentials and APIs were involved? A technology E&O issue because a deployed software service failed? A professional liability issue because customer-facing advice or service was wrong? A crime or social-engineering issue if payment controls were manipulated? A regulatory or media issue if customer notices were misleading? A D&O or governance issue if the board had ignored known AI control gaps? The answer cannot be inferred from the word "AI."

Policy language controls. Facts matter. Jurisdiction, insured role, policy line, endorsements, exclusions, limits, sublimits, definitions, notice, causation, and loss category all matter. This chapter does not interpret policy wording or provide a coverage opinion. Its purpose is to show why agentic lifecycle evidence can help sort line ambiguity without resolving it.

Broker and market sources already frame AI as a cross-line exposure. Aon discusses AI risk across cyber, E&O/professional liability, employment, crime, D&O/governance, and related enterprise risk categories. That is useful context, not a coverage determination. [35] [36]

Cyber is often the first line people think of because AI systems touch identity, credentials, APIs, cloud services, data stores, prompts, and model endpoints. Cyber may be relevant when there is unauthorized access, credential misuse, data exposure, API abuse, service interruption, compute misuse, LLMjacking, or regulatory investigation tied to a cyber event. QBE's AI cyber and LLMjacking materials show why access records, usage logs, cloud bills, containment steps, and remediation records matter. They do not make every agentic incident a cyber claim. [34]

Technology E&O and professional liability ask different questions. Was there a client deliverable, service failure, implementation defect, professional advice, platform error, or AI product underperformance? Did the insured provide a technology service or professional service? Was the loss tied to a model output, product behavior, integration failure, workflow configuration, or human professional judgment? Model-performance and AI warranty products can be relevant examples, but they are narrower than agentic lifecycle risk transfer.

D&O and governance exposure sits at another layer. The loss may not be only the operational event. It may also involve oversight, disclosure, risk management, AI governance failure, cyber governance, or alleged misstatement. SEC cyber disclosure rules are relevant as governance and disclosure context, not as insurance coverage proof. [37]

General liability and product liability may enter where an AI-enabled product, physical system, public interaction, or generated content creates bodily injury, property damage, or other covered categories, depending on policy language. Verisk's ISO filing discussion is evidence that GenAI liability exposure is being addressed in form development, but exact wording and jurisdictional adoption need careful verification before any precise policy claim is made. In Part III, the point is narrower: agentic evidence must identify the product, action, outcome, and loss path.

Media and IP exposures may appear when agentic systems generate, publish, recommend, train on, or distribute content. The relevant object may be a generated article, image, ad, recommendation, takedown response, rights review, or publication workflow. Evidence should preserve prompt/output records, content provenance, approval, publication, and remediation. This does not decide media liability, copyright, defamation, or advertising coverage.

Employment practices liability may be implicated when agents support hiring, screening, performance review, scheduling, discipline, termination, or workplace investigations. Evidence may need to show human role,

decision criteria, data sources, notices, adverse action records, and exception handling. This paper does not provide employment-law analysis or coverage interpretation.

Crime and fraud lines may be implicated by deepfake fraud, synthetic identity, invoice manipulation, AI-enabled social engineering, payment instruction abuse, or internal control bypass. The claim ambiguity may turn on voluntary transfer, social-engineering wording, employee involvement, authentication controls, and cyber overlap. Agentic evidence helps show communications, approvals, identity checks, payment paths, and tool actions. It does not decide policy response.

Property and business interruption may appear when agentic systems affect operational continuity, cloud services, supply chain, code deployment, or physical systems. Beazley and cloud-related cyber examples show that AI/cloud services can enter product-specific cyber or business interruption contexts. Those examples should not be generalized beyond their terms.

The same event can straddle several of these lines. That is why this paper insists on the risk object. A line-of-business question cannot be answered if the enterprise cannot name the work unit, authority, tool action, accepted outcome, affected data, external consequence, dependency chain, incident timeline, remediation closure, and missing evidence.

T-11-01 - Insurance Line Ambiguity Map

Cyber

Potential insurance line	Cyber
Agentic trigger pattern	Unauthorized access, credential misuse, LLMjacking, API abuse, data exposure, service interruption, compute misuse
Evidence needed	Identity logs, API usage, cloud bills, forensics, containment, data impact, remediation
Ambiguity	Cyber event versus authorized operational misuse; sublimits/exclusions; causation
Boundary note	Does not determine cyber coverage

Tech E&O

Potential insurance line	Tech E&O
Agentic trigger pattern	AI product, platform, software, integration, API, or implementation failure
Evidence needed	Product version, model/tool logs, SLA/KPI context, customer impact, change history
Ambiguity	Product failure versus professional service versus cyber
Boundary note	Not a product coverage opinion

Professional Liability / E&O

Potential insurance line	Professional Liability / E&O
Agentic trigger pattern	AI-assisted advice, deliverable, client decision support, professional workflow error
Evidence needed	Engagement scope, AI use record, reviewer role, accepted output, client reliance, remediation
Ambiguity	Professional judgment versus tool failure; disclosure and standard-of-care issues
Boundary note	Not legal or professional-liability advice

D&O / Governance

Potential insurance line	D&O / Governance
Agentic trigger pattern	Board oversight, disclosure, risk-management failure, AI governance failure, cyber governance issue
Evidence needed	Board materials, risk register, AI inventory, incident escalation, disclosure timeline
Ambiguity	Governance claim versus operational loss; knowledge, materiality, conduct exclusions
Boundary note	Not D&O coverage analysis

General Liability / Product

Potential insurance line	General Liability / Product
Agentic trigger pattern	AI-enabled product or generated action causes bodily injury, property damage, or other third-party loss
Evidence needed	Product version, user interaction, tool action, physical-world effect, incident record
Ambiguity	Product defect versus software/service; emerging form wording
Boundary note	Policy terms control

Media / IP

Potential insurance line	Media / IP
Agentic trigger pattern	Generated or distributed content creates IP, defamation, advertising, or publication dispute
Evidence needed	Prompt/output, provenance, rights review, publication approval, takedown/remediation
Ambiguity	Media liability versus IP exclusions versus tech service
Boundary note	Not IP or media coverage advice

Employment Practices

Potential insurance line	Employment Practices
Agentic trigger pattern	AI-supported hiring, screening, discipline, scheduling, or performance decision
Evidence needed	Decision workflow, data sources, criteria, human review, adverse action record
Ambiguity	Employment practice versus technology error; discrimination causation
Boundary note	Not employment-law advice

Crime / Fraud

Potential insurance line	Crime / Fraud
Agentic trigger pattern	Deepfake, invoice manipulation, synthetic identity, payment instruction abuse, social engineering
Evidence needed	Communication record, payment approval, callback controls, identity verification, authority trace
Ambiguity	Crime versus cyber/social-engineering limit; voluntary transfer issues
Boundary note	Does not assume crime coverage

Property / Business Interruption	
Potential insurance line	Property / Business Interruption
Agentic trigger pattern	Agentic workflow, cloud, deployment, or automation disrupts operations
Evidence needed	Outage timeline, dependency map, affected systems, recovery record, financial impact
Ambiguity	Cyber/technology/property boundary; direct versus contingent BI
Boundary note	Not loss measurement or coverage opinion

Evidence helps sort these lines because it prevents the event from being described as "an AI problem." It shows whether the loss-relevant action was access, advice, deployment, payment, content publication, governance disclosure, data handling, service outage, or product behavior. But sorting is not deciding. The paper does not interpret policies, advise insureds, recommend procurement, or tell claims teams how to pay or deny a claim.

The boundary for this chapter: insurance line ambiguity analysis is not legal advice, insurance advice, underwriting guidance, coverage opinion, claim approval guidance, legal liability determination, certification, proof of insurability, insurer endorsement, or a regulator-approved method. It is an object-and-evidence map for understanding why the same agentic event can raise multiple insurance questions.

The next chapter moves from line ambiguity to portfolio ambiguity. Even when one event can be reconstructed, agentic AI can create aggregation and concentration risk through shared models, vendors, runtimes, cloud dependencies, reusable agents, and common workflows.

Chapter 12: Aggregation, Reinsurance, and Concentration Risk in Agentic AI

One agentic incident is difficult enough. Many similar incidents across the same dependency can become a different kind of insurance problem.

An enterprise may deploy a procurement agent in one division, a refund agent in another, a code agent in engineering, and a customer-response agent in support. Each workflow looks separate to its business owner. Under the surface, they may share the same model provider, orchestration framework, cloud region, authentication service, vector database, tool connector, logging vendor, or prompt library. A failure in that shared layer can affect many work units at once.

For insurers and reinsurers, the concern is not only individual incident severity. It is repeated, correlated, or systemic event shape.

Cyber insurance already gives the closest external analogy. Geneva Association cyber-accumulation work highlights the challenge of common technologies, shared vulnerabilities, providers, systemic loss, quantification limits, and capital capacity. Swiss Re's cloud concentration work emphasizes visibility into shared infrastructure dependencies. Geneva Association's GenAI insurance work adds a more direct AI-business risk context. These sources do not provide an actuarial model for agentic AI. They support the narrower conclusion that shared dependencies matter for insurability reasoning. [38] [39] [40]

Agentic AI adds a lifecycle layer to the cyber/cloud accumulation problem. The question is not only "which technology dependency failed?" It is also "which delegated work objects depended on it, under what authority, with what evidence, and with what accepted outcomes?"

A single reusable agent component can create correlated loss inside one organization. The same agent that classifies support tickets may be reused to classify refund requests, compliance inquiries, and escalation severity. If the component has a hidden failure mode, the losses may appear in different business units and policy lines. Without cross-project lifecycle records, the organization may not see the common object.

A single model or runtime change can create correlated behavior across many workflows. The model may become more permissive, less cautious, more verbose, more likely to call a tool, or less stable under certain prompts. If substitution conformance is not recorded, the enterprise may be unable to identify which workflows changed and which accepted outcomes were affected.

A single vendor or cloud dependency can create portfolio-level exposure. The same API, cloud region, identity provider, logging service, data processor, or orchestration platform can sit inside many insured operations. If it fails, is compromised, changes behavior, or becomes unavailable, the loss may not stay inside one policyholder's isolated workflow.

A single tool connector can create action concentration. Many agents may share the same permission to send emails, update accounts, create tickets, execute code, transfer files, approve payments, or access customer data. The concentration is not just the connector. It is the authority attached to the connector.

A single prompt, policy, or guardrail pattern can create governance concentration. If an enterprise copies one agent instruction template across many workflows, the same ambiguity can repeat across customer service, finance, operations, and engineering. A reused guardrail may be a control. It may also be a common failure point.

A single evidence architecture can create reconstruction concentration. If logs are retained too briefly, vendor records are inaccessible, privacy redaction removes source pointers, or tool-action records are not joined to approvals, then many agentic incidents may become unreconstructable at once.

Reinsurers care about aggregation because an apparently diversified book can contain hidden common dependencies. Enterprises should care for the same reason inside the firm. A risk leader may think ten

workflows create ten separate exposures. In practice, they may create one concentrated dependency repeated ten times.

Dependency mapping and substitution conformance make aggregation visible. They do not solve it automatically. A dependency map should identify shared models, runtime layers, cloud services, APIs, tool connectors, data sources, processors, subprocessors, prompt libraries, authentication services, and evidence repositories. Substitution records should show when those dependencies changed, which work units were affected, and whether authority, privacy, evidence, and accepted-outcome boundaries remained intact.

T-12-01 - Agentic Aggregation Risk Map

Shared model provider

Aggregation vector	Shared model provider
Example pattern	Many workflows use one model endpoint for customer, finance, and code tasks
Evidence needed	Model inventory, endpoint/version map, affected work units, substitution record
Reinsurance or portfolio concern	Correlated behavior change or outage across business lines
Boundary note	Not an actuarial pricing model

Shared runtime/orchestration layer

Aggregation vector	Shared runtime/orchestration layer
Example pattern	One agent framework routes plans, tools, memory, and approvals across teams
Evidence needed	Runtime version, policy config, tool permissions, incident logs
Reinsurance or portfolio concern	Common control failure across workflows
Boundary note	Not reinsurer acceptance proof

Shared cloud/API dependency

Aggregation vector	Shared cloud/API dependency
Example pattern	Agents depend on one cloud region, API, identity provider, or service account
Evidence needed	Cloud/API map, identity logs, usage records, outage/incident timeline
Reinsurance or portfolio concern	Accumulation through common infrastructure
Boundary note	Cloud analogy does not fully solve AI risk

Reusable agent component

Aggregation vector	Reusable agent component
Example pattern	A classifier or planner is reused across support, billing, compliance, and operations
Evidence needed	Component registry, reuse approvals, authority changes, context changes
Reinsurance or portfolio concern	Same defect repeated under different scopes
Boundary note	Not a procurement recommendation

Shared tool connector

Aggregation vector	Shared tool connector
Example pattern	Multiple agents can send email, approve refunds, update ERP, deploy code, or transfer files
Evidence needed	Tool permission map, action logs, escalation rules, external consequence records
Reinsurance or portfolio concern	Concentrated external-action authority
Boundary note	Does not determine policy line

Shared vendor/processor chain

Aggregation vector	Shared vendor/processor chain
Example pattern	Model, logging, storage, data labeling, or monitoring vendors hold key records
Evidence needed	Vendor/subprocessor map, record access terms, retention and notice records
Reinsurance or portfolio concern	Evidence custody and common vendor failure
Boundary note	Not vendor ranking

Shared prompt or policy template

Aggregation vector	Shared prompt or policy template
Example pattern	One instruction pattern is copied into many workflows
Evidence needed	Prompt library, template version, deployment map, exception history
Reinsurance or portfolio concern	Repeated governance ambiguity or guardrail failure
Boundary note	Not a compliance certification

Shared evidence repository	
Aggregation vector	Shared evidence repository
Example pattern	Many workflows rely on one logging or evidence system
Evidence needed	Evidence partition, retention rules, access controls, missing-evidence register
Reinsurance or portfolio concern	Reconstruction failure can affect many incidents
Boundary note	Not claim approval guidance

Cross-insured dependency	
Aggregation vector	Cross-insured dependency
Example pattern	Many insureds rely on the same cloud, model, API, or AI service vendor
Evidence needed	Portfolio dependency data, public/vendor incident records, concentration analysis
Reinsurance or portfolio concern	Correlated losses across a portfolio
Boundary note	Author analytical inference unless source-backed

The table deliberately stops before pricing. It does not offer rates, capital models, actuarial assumptions, accumulation thresholds, reinsurance attachment points, or portfolio management instructions. It says only that agentic AI can concentrate exposure through shared components and that the concentration must become visible before it can be discussed responsibly.

There is also a governance lesson. If the enterprise does not know which workflows depend on the same component, it cannot know which work units to review after a vendor change, model incident, cloud outage, prompt vulnerability, or tool permission error. The loss may look isolated because the evidence architecture is isolated.

The boundary for this chapter: aggregation analysis is not actuarial pricing guidance, not reinsurance underwriting guidance, not capital modeling, not legal advice, not insurance advice, not coverage opinion, not certification, not proof of insurability, not insurer endorsement, and not a regulator-approved method. Cyber and cloud accumulation sources are used as analogy and context unless the source directly addresses GenAI business risk.

The next chapter addresses a tension that runs through every evidence discussion in this paper. Insurance review needs records. Privacy, security, privilege, and data-minimization duties limit what records should be collected, retained, copied, or disclosed.

Chapter 13: Privacy, Evidence Minimization, and Insurance Review

Agentic AI creates a temptation: keep everything.

Keep every prompt. Keep every output. Keep every tool call. Keep every customer payload. Keep every memory entry. Keep every retrieved document. Keep every approval screen. Keep every vendor log. Keep every incident note. Keep every trace forever, just in case an insurer, auditor, regulator, lawyer, customer, or executive asks for it later.

That instinct is understandable. It is also dangerous.

Insurance review needs evidence, but uncontrolled evidence retention can create privacy, security, privilege, contractual, and legal exposure. Personal data may appear in prompts, memory, tool payloads, customer records, employee records, model outputs, retrieval context, incident tickets, vendor logs, screenshots, monitoring dashboards, and claim evidence packs. Secrets and credentials may appear in traces. Privileged communications may appear in remediation notes. Vendor confidential information may appear in dependency records.

Evidence maturity is not the same as hoarding. A mature evidence model preserves the ability to reconstruct the event while minimizing unnecessary sensitive-data exposure.

This is where privacy lifecycle concepts from the compliance white paper and selective-disclosure concepts from the auditability and assurance white paper become important. The compliance white paper helps name privacy lifecycle mapping, evidence minimization, selective disclosure, data subject rights tension, and third-party processor chains. The auditability and assurance white paper helps distinguish source pointers from bulk evidence dumps and reviewable evidence from unrestricted trace collection. These are analytical foundations, not GDPR advice, privacy legal advice, insurance requirements, or certification methods. [28] [31]

The core tension is simple. Underwriting and claim review may need to understand authority, role, tool action, accepted outcome, affected data class, external consequence, exception history, and remediation closure. But those records may contain more data than the reviewer needs to see. A good evidence architecture therefore separates existence, pointer, class, access, and disclosure.

Source pointers are often better than copies. A claim evidence pack can identify the system of record, timestamp, owner, retention period, hash or integrity marker where appropriate, and access pathway without duplicating sensitive payloads into a new uncontrolled file.

Redaction profiles should be designed before incidents. The organization should know which fields can be masked, tokenized, summarized, or disclosed under role-based access. If redaction is invented after a loss, the team may either over-disclose sensitive data or destroy context needed for reconstruction.

Access controls matter because evidence audiences differ. Engineering, counsel, brokers, insurers, vendors, auditors, regulators, and executives do not all need the same records. Some may need summaries. Some may need source pointers. Some may need full technical logs under controlled conditions. Some should not receive personal data, secrets, privileged materials, or vendor confidential information at all.

Data-class labels make evidence more useful. A prompt containing public information, personal data, payment data, health information, source code, security secret, regulated customer record, or privileged communication should not be treated as one generic trace. Labeling helps decide what can be retained, disclosed, minimized, or escalated.

Retention notes matter because evidence can disappear. Logs roll off. Vendor dashboards expire. Model records may be inaccessible after version changes. Data rights requests may require deletion or restriction. Contract terms may limit retention. Security policies may delete secrets. If evidence must be preserved for review, the retention basis and limits should be visible.

Privilege flags matter because incident response and legal review may create protected communications. this paper does not define privilege. It simply observes that evidence packs should not flatten privileged and non-privileged records into one uncontrolled bundle.

Third-party evidence creates a special problem. A vendor may hold model logs, runtime records, security events, cloud usage, monitoring data, or subprocessor details. The enterprise may need those records for reconstruction but may not have them by default. Processor and subprocessor mapping therefore belongs in the insurance evidence conversation. It is not only a privacy or procurement exercise.

Privacy-preserving third-party validation may become useful when direct disclosure is too sensitive. A reviewer may need assurance that evidence exists and supports a fact without receiving raw data. That can mean selective disclosure, controlled review rooms, redacted extracts, attestations limited to existence and integrity, or source-pointer verification. This paper does not define a validation standard. It identifies the evidence problem.

T-13-01 - Insurance Evidence vs Privacy Control Map

Evidence need	Privacy risk	Control pattern	Related MRO	Boundary note
Work unit ID and event scope	Identifier can expose customer, employee, project, or account information	Scoped ID, pseudonymous reference, system-of-record pointer	Evidence partitioning	Not legal sufficiency proof
Prompt/output record	May contain personal data, secrets, proprietary content, or privileged material	Redaction, field masking, source pointer, access tier	Privacy / GDPR lifecycle mapping	Not GDPR advice
Tool-action record	API payloads may include sensitive customer or operational data	Payload minimization, action metadata, protected source record	Tool-action boundary; evidence minimization	Does not determine policy line
Human review record	Personnel data and privileged notes may be exposed	Role-based summary, privilege flag, reviewer-visible evidence pointer	Human role mapping; accepted outcome	Not liability determination
Vendor/model/runtime record	Vendor confidential data or subprocessor information may be restricted	Contract pointer, vendor evidence request path, controlled disclosure	Processor/subprocessor chain; substitution conformance	Not vendor certification
Affected data class	Overbroad disclosure can expose unnecessary personal data	Data-class label, affected population range, redacted sample if needed	Privacy lifecycle mapping	Not regulatory advice
Incident timeline	Security-sensitive details may expose vulnerabilities or response playbooks	Need-to-know access, segmented timeline, sensitive-detail appendix	Incident and remediation closure	Not claims approval guidance
Exception and override history	Employee behavior or customer details may appear	Aggregated pattern plus source pointer for detailed review	Exception path; evidence partition	Not employment or legal advice
Remediation closure	Legal strategy, privileged analysis, or vendor defects may be included	Closure summary, residual-risk note, privilege segmentation	Remediation closure	Does not prove no residual liability
Claim Evidence Pack disclosure	Bulk evidence sharing can create secondary privacy/security exposure	Selective disclosure, redaction profile, access log, retention note	Evidence minimization and selective disclosure	Not insurance advice or coverage opinion

The most practical design move is an evidence index. The index should not be a dump of sensitive data. It should tell future reviewers what exists, where it lives, who owns it, what class of data it contains, what access controls apply, what retention limits exist, what redactions are available, what privilege concerns may exist, and what evidence is missing.

That index can support underwriting conversations, claim reconstruction, renewal review, and governance oversight without turning every agentic trace into a portable liability file.

The privacy chapter also reinforces the negative space from Chapter 8. A system can be hard to insure not only because it lacks evidence, but because it keeps evidence irresponsibly. Opaque execution is a problem. So is uncontrolled trace retention. The insurability question is not "more logs or fewer logs." It is whether the right evidence can be preserved, minimized, protected, and reconstructed for the right review purpose.

The boundary for this chapter: privacy and evidence-minimization analysis is not GDPR advice, privacy legal advice, legal advice, insurance advice, underwriting guidance, coverage opinion, certification, proof of insurability, insurer endorsement, regulator approval, or claim approval guidance. It is an evidence design lens for reconciling reviewability with data protection, security, privilege, and controlled disclosure.

Part III closes the translation layer. The compliance white paper contributes lifecycle governance objects. The auditability and assurance white paper contributes auditability and evidence-chain discipline. Insurance adds separate questions: insured subject, line ambiguity, covered object, aggregation, privacy-controlled evidence, claim reconstruction, and risk transfer boundary. Part IV can now move into underwriting-facing architecture with a clearer warning: evidence can make agentic risk more reviewable, but reviewability is still not coverage, pricing, or acceptance.

Part IV: Underwriting-Facing Architecture for Agentic AI Risk

Part IV moves from translation to architecture.

Part I showed that the insurance market is already splitting AI risk across affirmative cover, exclusions, sublimits, silent exposure, model-performance products, cyber-linked exposure, and unresolved lifecycle gaps. Part II defined the Insurable Agentic Risk Object as a bounded analytical object, not the insured legal subject. Part III translated compliance and auditability concepts into insurability reasoning without treating governance or auditability as insurance proof.

The next question is practical: what would an enterprise need to organize if it wanted agentic AI risk to be reviewable by a broker, risk engineer, underwriter, reinsurer, counsel, or internal risk team?

The answer is not a checklist that guarantees coverage. It is not an underwriting standard. It is not a rating model. It is not a certification path. It is not a premium-reduction playbook. It is an underwriting-facing architecture: a way to organize exposure, evidence, change, and review questions around bounded agentic work units.

The phrase "underwriting-facing" is deliberate. It means the architecture is useful for a risk-transfer discussion. It does not mean an insurer requires it, accepts it, discounts for it, or treats it as sufficient. Evidence can make the risk more legible. It cannot by itself make the risk covered, insurable, priced, or accepted.

Chapter 14: Underwriting Evidence Architecture

An enterprise can have a mature AI program and still be difficult to review from an insurance perspective.

It may have a model inventory, AI policy, model-risk committee, vendor list, security controls, incident process, privacy review, and logging system. Those are useful. But when a broker or risk engineer asks what the enterprise is actually trying to transfer, the answer may still be too broad: "We use AI in customer operations," "We use agents in engineering," or "We have governance controls for generative AI."

Underwriting-facing review needs something more exact. It needs evidence organized around bounded agentic work units.

Picture a mid-sized software company preparing for renewal. It has an AI coding assistant, a customer-support response agent, a finance reconciliation agent, and a sales proposal generator. The risk team can name the vendors and models, but it cannot say which work units can send external messages, which can update customer records, which can touch production code, which data classes are involved, which human role accepts the outcome, which tool actions are logged, which model/runtime changes occurred during the period, or which incidents were remediated and closed. The company has AI governance. It does not yet have underwriting-facing evidence architecture.

The architecture should begin with the work-unit inventory. The reviewer needs to know which bounded agentic work exists, not only which AI applications or models exist. A work unit may be a refund workflow, code-deployment assistant, claims triage process, contract review support workflow, customer message generator, invoice matching agent, procurement agent, or regulatory filing assistant. The work unit connects business function, authority, tool action, data, human role, vendor dependency, evidence, and consequence.

Authority boundaries come next. A draft-only agent creates a different review question from an agent that can approve, send, transact, deploy, delete, transfer, or update records. The boundary should show what the agentic work is allowed to do, what requires confirmation, what is prohibited, what thresholds apply, and where escalation occurs.

Role maps matter because agentic work rarely belongs to one actor. The architecture should identify the human role, agent role, vendor role, corporate owner, remediation owner, and any processor or subprocessor that holds relevant evidence. This is not a liability allocation. It is a review map.

Tool-action records are the point where output becomes consequence. Underwriting-facing evidence should identify the systems or services the agent can touch: email, CRM, ERP, payment system, code repository, deployment pipeline, ticketing system, database, cloud console, identity provider, third-party API, or customer record. The existence of tool records is not enough; the records should be linked to authority, role, outcome, exception, and closure.

Accepted outcome states help distinguish generated output from business action. A model response may be provisional. A sent message, approved payment, merged code change, updated customer account, or delivered professional recommendation may be accepted. Reviewers need to see how the enterprise knows the difference.

Exception history tells the reviewer where the system has already struggled. Overrides, threshold breaches, escalations, false positives, near misses, warning suppressions, vendor outages, prompt failures, and tool failures are not embarrassing side notes. They are evidence of how the risk behaves.

Privacy and redaction profiles make evidence usable. As Part III explained, a serious evidence model should not dump every prompt, payload, customer record, credential, secret, privileged note, or vendor log into a portable file. The architecture should show what can be disclosed, what must remain source-pointed, what is redacted, and who can access it.

Substitution and change records are central because agentic systems do not stay fixed. Model endpoints change. Vendor defaults change. Tool permissions widen. Data classes expand. Human review moves from required to sampled. A reviewed risk object can silently become a different risk object.

Remediation closure shows whether the organization can learn from events. A reviewer may need to see not only that incidents occurred, but whether they were contained, fixed, retested, reauthorized, and closed with residual risk visible. NIST and CISA incident-response sources support the importance of preparation, detection, response, recovery, remediation, coordination, and tracking, but they do not create insurance claim or underwriting standards. [41] [42]

Finally, the architecture should include a missing-evidence register. Missing evidence is part of the truth. It may be vendor-held, overwritten, outside retention, privileged, redacted, inaccessible, or never collected. A missing-evidence register helps reviewers understand uncertainty rather than pretend completeness.

Different users will use the architecture differently.

A broker may use it to translate enterprise AI operations into risk-transfer language. A risk engineer may use it to understand exposures, controls, dependencies, and reconstruction gaps. An underwriter may use it to ask better questions, subject to that insurer's own appetite, forms, guidelines, and judgment. A reinsurer may use it to see dependency concentration and aggregation shapes. Counsel may use it to preserve boundaries between evidence, privilege, legal advice, and coverage review. Enterprise risk teams may use it to make internal risk ownership visible.

None of those uses means the architecture is required, accepted, sufficient, or price-relevant. source research supports the general proposition that AI risk review may need exposure inventory, authority scope, dependencies, and evidence readiness, drawing on Aon, NAIC, NIST AI RMF, Geneva Association, Swiss Re, QBE, the compliance white paper, and the auditability and assurance white paper. The exact use by any insurer remains external. [43] [44] [45]

T-14-01 - Underwriting Evidence Architecture Components

Work-unit inventory

Component	Work-unit inventory
Review question	Which bounded agentic work creates exposure?
Evidence artifact	Work unit ID, business function, owner, external consequence tag
User of evidence	Broker, risk engineer, enterprise risk
Boundary note	Reviewability layer only

Authority boundaries

Component	Authority boundaries
Review question	What can the work do without additional approval?
Evidence artifact	Permission scope, threshold, escalation rule, prohibited action list
User of evidence	Underwriter, counsel, risk engineer
Boundary note	Not legal delegation proof

Human/agent/vendor role maps

Component	Human/agent/vendor role maps
Review question	Who initiates, executes, reviews, accepts, supports, remediates, and closes?
Evidence artifact	RACI, agent role record, vendor support role, processor map
User of evidence	Counsel, underwriter, enterprise risk
Boundary note	Does not determine legal liability

Tool-action records

Component	Tool-action records
Review question	Where does output become external consequence?
Evidence artifact	API calls, account updates, emails, payments, code deployments, filings
User of evidence	Risk engineer, claims/IR, underwriter
Boundary note	Does not determine policy line

Accepted outcome states

Component	Accepted outcome states
Review question	When does the enterprise treat the work as done or business-adopted?
Evidence artifact	Final-state marker, approval criteria, reviewer-visible evidence
User of evidence	Underwriter, counsel, business owner
Boundary note	Not coverage or legal acceptance proof

Exception history

Component	Exception history
Review question	What warnings, overrides, failures, or near misses occurred?
Evidence artifact	Exception log, override register, incident/near-miss review
User of evidence	Risk engineer, underwriter, renewal team
Boundary note	Not a loss prediction formula

Privacy/redaction profile

Component	Privacy/redaction profile
Review question	Can evidence be reviewed without uncontrolled sensitive-data exposure?
Evidence artifact	Data-class labels, source pointers, access controls, redaction rules
User of evidence	Counsel, broker, risk engineer
Boundary note	Not privacy legal advice

Substitution/change records	
Component	Substitution/change records
Review question	Did the reviewed object change over time?
Evidence artifact	Model/tool/vendor version history, reauthorization, conformance note
User of evidence	Underwriter, reinsurer, enterprise risk
Boundary note	Not vendor certification

Remediation closure	
Component	Remediation closure
Review question	What was fixed, retested, reauthorized, and closed?
Evidence artifact	Incident timeline, fix record, retest evidence, owner signoff
User of evidence	Renewal team, claims/IR, risk engineer
Boundary note	Does not prove no residual liability

Missing-evidence register	
Component	Missing-evidence register
Review question	What cannot be reconstructed?
Evidence artifact	Missing, stale, inaccessible, redacted, vendor-held, or overwritten record list
User of evidence	All reviewers
Boundary note	Not a claim denial or approval basis

The table is intentionally architectural. It does not say "submit these documents to get coverage." It says that a serious discussion of agentic AI risk needs a common evidence map before the parties can even understand the exposure.

The boundary for this chapter: underwriting evidence architecture is not legal advice, not insurance advice, not underwriting guidance, not an underwriting standard, not coverage opinion, not certification, not proof of insurability, not insurer endorsement, not a regulator-approved method, not actuarial pricing guidance, and not a premium recommendation.

The next chapter turns the architecture into an exposure inventory. Once the evidence components are known, the enterprise still needs to segment where agentic work creates different kinds of exposure.

Chapter 15: Agentic Exposure Inventory and Risk Segmentation

Many AI inventories are built for technology management, not risk transfer.

They list the application name, vendor, model family, business owner, data source, risk rating, and approval status. That may help procurement, security, or governance. It does not necessarily help a reviewer understand exposure. A single AI application may contain many agentic work units. A single model may support dozens of workflows. A single workflow may straddle cyber, professional liability, governance, crime, privacy, product, and business interruption questions.

An agentic exposure inventory should therefore sit at the level of work units, not only tools or models.

Consider a bank that lists "customer-service AI assistant" in its AI inventory. That label hides several different exposures. One work unit drafts customer responses. Another updates contact details. Another flags suspected fraud. Another recommends fee reversals. Another opens support tickets with a vendor. Another summarizes complaint history for a human reviewer. These work units differ in authority, data sensitivity, customer impact, external consequence, reversibility, dependency concentration, and line ambiguity.

Risk segmentation is the discipline of separating those work units before the insurance conversation starts.

The first segmentation dimension is business function. Agentic work in customer support, finance, engineering, legal, HR, procurement, sales, operations, claims, compliance, or security may create different exposure patterns. Function is not enough, but it orients the reviewer.

The second is external consequence. Does the work only draft internally, or can it send, file, deploy, pay, approve, deny, update, delete, publish, notify, or instruct an outside party? External consequence changes the risk conversation because the work leaves the model environment and affects someone or something else.

The third is authority level. A recommendation-only workflow is different from one that can transact. A sampled human review boundary is different from mandatory pre-action approval. An agent with emergency override permission is different from one confined to low-value routine activity.

The fourth is data sensitivity. Public content, internal business data, customer personal data, payment data, health data, employee data, source code, credentials, trade secrets, privileged material, and regulated records do not carry the same evidence and privacy implications.

The fifth is tool-action type. Emailing a customer, updating a CRM record, changing code, opening a vendor ticket, querying a database, initiating a payment, publishing content, or changing cloud infrastructure all produce different evidence and loss questions.

The sixth is customer or third-party impact. Some work units affect only internal productivity. Others affect customers, counterparties, vendors, employees, investors, regulators, patients, applicants, or the public. Impact surface matters even when the policy line is not yet known.

The seventh is dependency concentration. A work unit dependent on the same model, runtime, cloud region, identity service, tool connector, or vendor as many other work units carries aggregation relevance. Swiss Re and Geneva Association sources support the importance of dependency concentration and accumulation visibility as context and analogy. They do not provide agentic AI pricing conclusions. [46] [47]

The eighth is human confirmation boundary. "Human in the loop" is too vague. The inventory should record whether human confirmation is pre-action, post-action, sampled, threshold-based, exception-triggered, summary-only, or absent. It should also record what evidence the human sees.

The ninth is reversibility and remediation difficulty. A draft can be deleted. A sent customer notice may be corrected. A payment may or may not be reversible. A code deployment may be rolled back, but customer

impact may persist. A regulatory filing may create a different closure problem. Reversibility changes the remediation story.

The tenth is cross-project reuse. A component reused across projects may create hidden exposure drift. The same agentic component can be low-risk in one context and high-impact in another.

This inventory helps risk discussion because it separates exposure units before they are bundled into broad claims like "AI use," "GenAI," or "agent platform." It also connects back to the earlier insurable agentic risk object and insurance line ambiguity analysis. The inventory tells the reviewer which object is being discussed and what kinds of line questions may arise.

It still does not create insurability. A clean inventory can support discussion, but actual insurance outcomes depend on policy language, underwriting appetite, limits, exclusions, loss history, controls, insured profile, jurisdiction, and the facts of the event.

T-15-01 - Agentic Exposure Inventory Template

Exposure dimension	What to record	Why it matters	Related lifecycle object	Boundary note
Business function	Support, finance, engineering, HR, legal, sales, operations, compliance, security	Locates the exposure in the enterprise	Work-unit inventory	Not a coverage category by itself
External consequence	Draft, send, file, deploy, pay, approve, deny, update, delete, publish, notify	Shows whether work affects third parties or systems	Tool-action boundary; accepted outcome	Does not decide policy line
Authority level	Recommend, approve, transact, execute, escalate, override, block	Defines delegated action perimeter	Authority boundary	Not legal delegation proof
Data sensitivity	Public, internal, personal, payment, health, employee, source code, credential, privileged	Shapes evidence, privacy, and security controls	Privacy lifecycle mapping	Not privacy legal advice
Tool-action type	Email, CRM/ERP update, API call, payment, code deployment, filing, database update	Identifies the consequence channel	Tool-action record	Not claim causation proof
Customer/third-party impact	Internal only, customer-facing, vendor-facing, employee-facing, public-facing, regulator-facing	Frames potential harm surface	Accepted outcome; external consequence	Not liability determination
Dependency concentration	Shared model, runtime, cloud, API, identity, vendor, data processor, evidence repository	Shows common failure and aggregation pathways	Dependency visibility	Not actuarial pricing guidance
Human confirmation boundary	Pre-action, post-action, sampled, threshold-based, exception-only, summary-only, absent	Shows whether review is meaningful and evidenced	Human role mapping; accepted outcome	HITL is not proof of transferability
Reversibility/remediation difficulty	Easy rollback, delayed correction, irreversible transaction, public correction, regulatory closure	Helps frame containment and closure complexity	Remediation closure	Not loss estimate or claim outcome
Cross-project reuse	Original scope, reused scope, changed authority, changed data, changed consequence	Shows exposure drift across contexts	Cross-project lifecycle; substitution conformance	Not procurement recommendation

The strongest inventories will also show gaps. A work unit with unknown authority, unknown data class, unknown human confirmation, unknown vendor dependency, or unknown evidence location should not be forced into a false sense of completeness. Unknowns are risk information.

This chapter's boundary: an agentic exposure inventory is not legal advice, not insurance advice, not underwriting guidance, not an underwriting standard, not coverage opinion, not certification, not proof of insurability, not insurer endorsement, not a regulator-approved method, not actuarial pricing guidance, and not a premium recommendation. It organizes exposure units for review.

The next chapter addresses the question that CFOs and CROs often ask too quickly: which variables matter to premium? The safe answer is to discuss exposure variables without turning them into pricing.

Chapter 16: Premium and Exposure Variables Without Pricing Guidance

The premium question is understandable. It is also easy to mishandle.

A CFO asks whether better AI controls will lower premium. A CRO asks whether autonomous agents will cost more to insure. A broker asks what evidence to gather before renewal. A product leader asks whether a human approval gate changes the insurance conversation. An engineering leader asks whether dependency concentration matters. The pressure is to answer with a formula.

This chapter does not provide one.

It does not provide actuarial pricing guidance, rating methodology, premium recommendation, underwriting rule, insurer appetite statement, rate factor, surcharge, discount, credit, model score, band, or threshold. It identifies variables that may matter for risk review because they affect exposure, evidence, control, aggregation, or reconstruction.

The first variable is autonomy level. A system that drafts text for human use raises different review questions from one that can approve refunds, deploy code, transfer funds, change records, or send customer-impacting instructions. Autonomy matters because it changes timing, intervention, authority, and evidence expectations. It still cannot be converted into a pricing band.

The second is tool-action severity. The same model output can be harmless in a draft and serious when connected to payment, production infrastructure, regulated filing, customer notice, medical workflow, legal advice, or financial decision support. Tool action shows where consequence becomes concrete.

The third is transaction volume. A workflow that runs five times a month has a different opportunity surface from one that acts thousands of times per day. Volume may matter to frequency discussion, but frequency is not a rate formula.

The fourth is customer or third-party reach. Internal productivity tools, customer-facing tools, vendor-facing tools, employee-facing tools, investor-facing disclosures, and public content do not create the same external impact surface.

The fifth is data sensitivity. Agentic work involving personal data, payment data, credentials, source code, health data, employee data, privileged material, or regulated records may raise privacy, security, cyber, professional, and governance questions.

The sixth is reversibility. A low-value internal recommendation may be easy to correct. A payment, customer communication, production deployment, regulatory filing, or public statement may be harder to unwind. Reversibility shapes remediation discussion without deciding loss amount.

The seventh is human confirmation strength. A meaningful confirmation boundary includes role, authority, criteria, visible evidence, timing, and exception handling. A summary-only click may not have the same review value as pre-action approval with full context. This does not mean human review guarantees coverage or lowers premium.

The eighth is evidence maturity. Reviewers may ask whether records connect work unit, authority, role, tool action, accepted outcome, exception, privacy treatment, remediation, and change. Evidence maturity makes the risk more reviewable. It does not guarantee coverage, quote, or claim payment.

The ninth is vendor, model, and runtime concentration. Shared dependencies can create correlated exposure across workflows or insureds. Geneva Association and Swiss Re sources support the importance of accumulation and cloud concentration as context and analogy. QBE LLMjacking materials show concrete AI-linked cyber evidence needs around access, usage, containment, and remediation. None of these sources creates agentic AI pricing guidance. [46] [47] [48]

The tenth is prior incidents and near misses. Events matter because they show how the system behaves under stress and whether controls work. Near misses may be as informative as losses. The safe question is not "what premium change follows?" but "what happened, what evidence exists, and what changed after remediation?"

The eleventh is remediation maturity. Detection is only the beginning. The reviewer may ask how quickly the organization contained, fixed, retested, reauthorized, and closed the issue, and whether residual risk remained visible. NIST and CISA support structured incident and remediation records. They do not provide insurance pricing rules. [41] [42]

The twelfth is cross-project reuse. A component reused across many contexts can spread exposure. A model or tool that was reviewed for one workflow may not be reviewable in another without reauthorization. Reuse matters because it changes object boundaries.

These variables are still incomplete without policy language, line of coverage, insured profile, jurisdiction, market appetite, loss history, controls, exclusions, limits, sublimits, deductibles, retention, and the underwriter's own framework. A variable dictionary is therefore only a way to organize the conversation.

T-16-01 - Non-Pricing Exposure Variables

Autonomy level

Variable	Autonomy level
Why it may matter	Changes timing, intervention, and delegated authority
Evidence signal	Draft/recommend/approve/execute profile, escalation rule
Misuse to avoid	Turning autonomy into a rate band
Boundary note	Not pricing guidance

Tool-action severity

Variable	Tool-action severity
Why it may matter	Shows where output becomes consequential action
Evidence signal	Payment, code, account, filing, email, infrastructure, database action
Misuse to avoid	Assuming severity determines coverage
Boundary note	Not coverage opinion

Transaction volume

Variable	Transaction volume
Why it may matter	Indicates how often the work can create exposure
Evidence signal	Run count, transaction count, customer interactions, API usage
Misuse to avoid	Converting volume into rate
Boundary note	Not actuarial model

Customer/third-party reach

Variable	Customer/third-party reach
Why it may matter	Expands potential external impact surface
Evidence signal	Customer-facing, vendor-facing, employee-facing, regulator-facing marker
Misuse to avoid	Treating reach as liability determination
Boundary note	Not legal advice

Data sensitivity

Variable	Data sensitivity
Why it may matter	Shapes privacy, cyber, regulatory, and evidence controls
Evidence signal	Data-class label, affected data map, redaction profile
Misuse to avoid	Providing privacy compliance advice
Boundary note	Not GDPR advice

Reversibility

Variable	Reversibility
Why it may matter	Affects containment and remediation complexity
Evidence signal	Rollback path, correction process, irreversible action flag
Misuse to avoid	Estimating claim amount from reversibility
Boundary note	Not loss valuation

Human confirmation strength

Variable	Human confirmation strength
Why it may matter	Shows whether review is role-based and evidence-aware
Evidence signal	Reviewer role, visible evidence, approval criteria, timing
Misuse to avoid	Claiming HITL lowers premium or guarantees transfer
Boundary note	Not premium recommendation

Evidence maturity

Variable	Evidence maturity
Why it may matter	Makes authority, role, action, outcome, exception, and closure reviewable
Evidence signal	Evidence index, source pointers, missing-evidence register
Misuse to avoid	Treating completeness as claim approval
Boundary note	Not proof of insurability

Vendor/model/runtime concentration

Variable	Vendor/model/runtime concentration
Why it may matter	Reveals correlated dependency exposure
Evidence signal	Dependency map, version register, affected work-unit list
Misuse to avoid	Quantifying systemic premium effect
Boundary note	Not capital or pricing model

Prior incidents / near misses

Variable	Prior incidents / near misses
Why it may matter	Shows observed failure and response patterns
Evidence signal	Incident log, near-miss review, exception history
Misuse to avoid	Saying incident-free history earns a discount
Boundary note	Not insurer appetite statement

Remediation maturity

Variable	Remediation maturity
Why it may matter	Shows containment, fix, retest, reauthorization, and closure
Evidence signal	Timeline, fix record, retest, owner signoff, residual-risk note
Misuse to avoid	Creating closure thresholds
Boundary note	Not underwriting rule

Cross-project reuse

Variable	Cross-project reuse
Why it may matter	Shows exposure drift across contexts
Evidence signal	Reuse register, changed authority/data/consequence marker
Misuse to avoid	Treating reuse as procurement defect
Boundary note	Not procurement recommendation

The careful language is "may matter," "can shape exposure discussion," "risk reviewers may ask," and "evidence signal." The unsafe language is "will affect premium," "earns a discount," "creates a surcharge," "qualifies for coverage," "fails underwriting," or "meets a rating level."

This chapter's boundary: premium and exposure variables are analytical inputs for risk review. They are not actuarial pricing guidance, not rating methodology, not premium recommendation, not underwriting guidance, not underwriting standard, not coverage opinion, not legal advice, not insurance advice, not certification, not proof of insurability, not insurer endorsement, and not a regulator-approved method.

The next chapter adds time. Even a well-described exposure can change quickly when workflows, authority, tools, models, vendors, data classes, human review boundaries, incidents, or evidence gaps change.

Chapter 17: Renewal, Change, and Substitution Evidence

Agentic AI risk does not stand still between renewal cycles.

A workflow that was draft-only in January may send customer messages in April. A model endpoint may change in June. A tool connector may gain write permission in July. A vendor may change logging defaults in August. A team may reuse the same agent in a new business unit in September. A near miss may reveal that the human reviewer was seeing only a summary, not the full evidence. By renewal, the original risk object may exist only on paper.

Underwriting-facing evidence cannot be one-time only. It needs a change layer.

Renewal review should begin with new workflows. Which agentic work units were added, retired, expanded, merged, or reused? The reviewer should not have to infer change from a model inventory.

Changed authority is often the most important update. Did any workflow move from draft to send, recommend to approve, approve to transact, or human-required to sampled review? Did thresholds change? Were emergency overrides added? Was escalation weakened?

New tools matter because tool actions create external consequence. A workflow that gains access to email, payment, CRM, ERP, cloud infrastructure, code deployment, data export, vendor ticketing, or customer records has changed its exposure shape.

New models, vendors, and runtimes matter because substitution can alter behavior and evidence. A reviewed work unit may depend on a model version, orchestration layer, cloud service, logging tool, data processor, or API. If that dependency changes, the enterprise needs substitution conformance evidence.

Changed data classes matter because privacy, cyber, regulatory, and claim reconstruction implications change. A workflow that originally touched public content may now touch customer records, payment data, employee records, code, credentials, or privileged material.

Changed customer impact matters because an internal productivity tool can become customer-facing. A tool used for internal drafting can become a communication channel. A recommendation can become a decision. A decision can become an automated action.

Changed human review boundaries matter because a control may weaken without appearing to disappear. A reviewer may shift from full-context review to summary review, from pre-action review to post-action review, from every action to sampled review, or from named role to rotating queue.

Incidents and near misses matter because they are the evidence of how the system behaves. Renewal evidence should record what happened, which work units were affected, what was contained, what changed, what remains open, and what lessons were implemented. NIST and CISA support continuous improvement and remediation tracking as incident-response disciplines, not as insurance requirements. [41] [42]

Remediation actions matter only if closure is visible. A fix without retest, reauthorization, residual-risk note, owner signoff, or monitoring plan may not close the review question.

Unresolved evidence gaps should travel into renewal. If a vendor-held log was inaccessible, a retention period was too short, a prompt record was redacted without pointer, a human approval screen lacked context, or a tool action was not joined to authority, the gap should not disappear from the file.

The useful frame is before execution, during execution, and after execution.

Before execution, the enterprise defines work unit, authority, role, data, tool, dependency, and evidence expectations. During execution, it records action, confirmation, exception, and outcome. After execution, it

records incident, remediation, substitution, closure, and renewal change. This dynamic lifecycle record is what keeps the reviewed object from going stale.

T-17-01 - Renewal and Change Evidence Register

New workflows

Change category	New workflows
Evidence to review	Added work-unit list, business owner, external consequence marker
Why it matters	New work may create new exposure units
Trigger for escalation	Customer/third-party impact, regulated data, tool action
Boundary note	Not coverage-ready status

Changed authority

Change category	Changed authority
Evidence to review	Permission change, threshold update, escalation change, override addition
Why it matters	Authority defines the action perimeter
Trigger for escalation	Draft-to-send, recommend-to-approve, approve-to-transact
Boundary note	Not legal delegation proof

New tools

Change category	New tools
Evidence to review	Tool connector, API permission, system access, action log
Why it matters	Tools convert output into external consequence
Trigger for escalation	Write/delete/payment/deployment/customer-message access
Boundary note	Does not decide policy line

New models/vendors/runtimes

Change category	New models/vendors/runtimes
Evidence to review	Version change, vendor notice, runtime config, conformance review
Why it matters	Substitution may change behavior and evidence
Trigger for escalation	New provider, endpoint, logging default, model class
Boundary note	Not vendor certification

Changed data classes

Change category	Changed data classes
Evidence to review	Data map update, data-class label, privacy/redaction change
Why it matters	Data sensitivity changes privacy and cyber exposure
Trigger for escalation	Personal/payment/health/employee/credential/privileged data
Boundary note	Not privacy legal advice

Changed customer impact

Change category	Changed customer impact
Evidence to review	Customer-facing marker, notice channel, third-party dependency
Why it matters	Internal tools can become external-impact tools
Trigger for escalation	Public/customer/regulator-facing action
Boundary note	Not liability determination

Changed human review boundary

Change category	Changed human review boundary
Evidence to review	Reviewer role, evidence visible, pre/post/sampled review, exception trigger
Why it matters	HITL value depends on role, timing, criteria, and evidence
Trigger for escalation	Summary-only, sampled, post-action, queue-based review
Boundary note	Not proof of transferability

Incidents / near misses

Change category	Incidents / near misses
Evidence to review	Incident log, near-miss review, exception history, affected work units
Why it matters	Observed behavior informs future review
Trigger for escalation	Repeated pattern, severe action, missing evidence
Boundary note	Not claim approval guidance

Remediation actions

Change category	Remediation actions
Evidence to review	Fix, retest, reauthorization, residual-risk note, owner closure
Why it matters	Shows whether the event changed the risk object
Trigger for escalation	Unclosed fix, unresolved residual risk, no retest
Boundary note	Does not prove no residual liability

Unresolved evidence gaps

Change category	Unresolved evidence gaps
Evidence to review	Missing, stale, overwritten, vendor-held, inaccessible, privileged, redacted records
Why it matters	Gaps define uncertainty at renewal
Trigger for escalation	Gap affects high-impact work unit or claim reconstruction
Boundary note	Not claim denial or underwriting rule

Substitution conformance is the connective tissue. It asks whether a changed component preserves the work unit's authority, role, evidence, privacy, outcome, exception, and remediation boundaries. It is not vendor certification. It is a continuity question.

This chapter's boundary: renewal and change evidence is a continuity record for review. It is not legal or insurance advice, underwriting guidance, coverage opinion, certification, proof of insurability, insurer endorsement, regulator-approved method, actuarial pricing guidance, or premium recommendation.

The next chapter turns the architecture into a practical evidence request structure while preserving the most important boundary: optional evidence requests should not masquerade as standards.

Chapter 18: Reviewer-Facing Evidence Requests Without Creating a Standard

Evidence request lists are useful. They are also dangerous.

If written badly, they become fake standards. A template starts to sound like a requirement. A request list starts to sound like an underwriting checklist. A readiness package starts to sound like certification. A missing field starts to sound like automatic denial. A completed field starts to sound like coverage readiness.

That is not what this paper is doing.

This chapter offers a structured, optional evidence request model for enterprises, brokers, risk engineers, counsel, and reviewers who need a practical way to organize agentic AI risk discussions. It is not a claim demand, not an underwriting checklist, not a certification checklist, not a regulator-approved checklist, not a procurement requirement, and not proof that any insurer will accept, price, quote, renew, bind, endorse, or pay a claim.

The first request area is an inventory summary. It should name the population of agentic work units, not simply the AI tools. The summary should identify business function, owner, authority class, data class, external consequence, dependency concentration, and current status.

The second is a high-impact work-unit list. Reviewers do not need every low-risk experiment at the same depth. They may need to see work units that touch customers, money, regulated data, production systems, professional deliverables, employee decisions, public communications, or high-volume automated actions.

The third is an authority and role map. The request should ask who initiates, configures, approves, accepts, escalates, remediates, and closes, and what the agent, tool, vendor, and corporate owner do. This turns HITL into a reviewable responsibility structure.

The fourth is a tool-action sample. A useful sample should show how model output becomes external action, such as an API call, email, payment, account update, code deployment, database change, filing, or vendor ticket. It should also show authority and accepted outcome.

The fifth is an evidence-chain sample. The goal is not a full data dump. It is to show that core lifecycle records can be joined through source pointers across prompt, output, tool action, approval, exception, incident, remediation, and gap evidence.

The sixth is a privacy and redaction profile. Reviewers may need to know how sensitive data is protected without receiving uncontrolled raw traces. The request should ask for data-class labels, access controls, redaction rules, and source-pointer strategy.

The seventh is incident and near-miss history. The request should identify relevant events, affected work units, external consequence, evidence available, containment, fix, retest, reauthorization, and closure state. It should not ask the enterprise to pre-judge coverage or liability.

The eighth is remediation closure examples. A few examples can show whether the organization closes incidents as evidence objects or merely resolves tickets operationally.

The ninth is a substitution and change register. Reviewers may ask what changed since the last review or deployment: models, tools, vendors, runtimes, permissions, data classes, customer impact, and human review boundaries.

The tenth is a dependency map. This should show shared models, cloud services, APIs, identity providers, data processors, tool connectors, evidence systems, and vendors across work units. For reinsurers and portfolio-minded reviewers, dependency concentration can be as important as individual workflow design.

The eleventh is a missing-evidence register. This is where the enterprise shows intellectual honesty. It should say what is missing, why, who controls it, whether it is recoverable, whether it is redacted, whether retention expired, and whether remediation is planned.

T-18-01 - Optional Reviewer Evidence Request Structure

Inventory summary

Request area	Inventory summary
Example request	Provide a summary of agentic work units by function, owner, authority, data class, consequence, and dependency
Purpose	Establish review population
Sensitive boundary	Avoid exposing unnecessary low-level logs
Non-claim note	Optional structure only

High-impact work-unit list

Request area	High-impact work-unit list
Example request	Identify work units touching customers, money, regulated data, production systems, professional outputs, or public communications
Purpose	Focus attention on material exposure units
Sensitive boundary	Use scoped identifiers where possible
Non-claim note	Not coverage-ready evidence

Authority and role map

Request area	Authority and role map
Example request	Show who initiates, reviews, accepts, escalates, remediates, and closes, and what the agent/vendor/tool does
Purpose	Convert HITL into responsibility structure
Sensitive boundary	Protect personnel and privileged details
Non-claim note	Not liability determination

Tool-action sample

Request area	Tool-action sample
Example request	Provide a sample showing output, action, authority, and accepted outcome
Purpose	Show where external consequence occurs
Sensitive boundary	Mask secrets, credentials, and customer data
Non-claim note	Does not decide policy line

Evidence chain sample

Request area	Evidence chain sample
Example request	Show source pointers connecting prompts, outputs, tool calls, approvals, exceptions, incidents, remediation, and gaps
Purpose	Demonstrate reconstructability
Sensitive boundary	Prefer pointers/redaction over bulk export
Non-claim note	Not claim approval guidance

Privacy/redaction profile

Request area	Privacy/redaction profile
Example request	Describe data classes, redaction rules, access controls, privilege flags, and disclosure path
Purpose	Avoid evidence hoarding and uncontrolled sharing
Sensitive boundary	Preserve privacy, security, privilege, and vendor confidentiality
Non-claim note	Not legal or privacy advice

Incident/near-miss history

Request area	Incident/near-miss history
Example request	Summarize events, affected work units, consequence, containment, fix, retest, and closure
Purpose	Show observed risk and response maturity
Sensitive boundary	Separate factual timeline from legal analysis
Non-claim note	Not claim outcome prediction

Remediation closure examples

Request area	Remediation closure examples
Example request	Provide examples of closure records, owner signoff, residual risk, and reauthorization
Purpose	Show whether events become closed lifecycle records
Sensitive boundary	Segment privileged and sensitive notes
Non-claim note	Does not prove no residual liability

Substitution/change register

Request area	Substitution/change register
Example request	List model, tool, vendor, runtime, permission, data, customer-impact, and review-boundary changes
Purpose	Keep the reviewed object current
Sensitive boundary	Avoid vendor confidential over-disclosure
Non-claim note	Not vendor certification

Dependency map	
Request area	Dependency map
Example request	Show shared model, cloud, API, identity, vendor, data processor, tool connector, and evidence system dependencies
Purpose	Identify aggregation and concentration
Sensitive boundary	Limit disclosure of security architecture
Non-claim note	Not reinsurance pricing guidance

Missing-evidence register	
Request area	Missing-evidence register
Example request	Identify missing, stale, overwritten, inaccessible, vendor-held, privileged, or redacted records
Purpose	Make uncertainty visible
Sensitive boundary	Protect privileged and sensitive explanations
Non-claim note	Not claim denial or underwriting rule

The request model should be used with judgment. A small internal drafting assistant does not need the same depth as a high-volume payment workflow. A professional-services workflow may need different evidence from a customer-support bot. A cyber-linked LLMjacking event may require different access and usage records than a model-performance warranty scenario. QBE, Aon, Geneva, Swiss Re, NAIC, NIST, the compliance white paper, and the auditability and assurance white paper sources all support pieces of the evidence, governance, dependency, and reconstruction logic, but no cited source supports treating this request structure as a market-wide insurer requirement. [41] [42] [43] [44] [45] [46] [47] [48]

The final discipline is tone. A request should ask, not command. It should say "provide if available," "describe," "identify," "summarize," "show source pointers," and "note gaps." It should not say "must provide to qualify," "required for coverage," "certification evidence," "underwriting pass/fail," or "premium credit."

This chapter's boundary: reviewer-facing evidence requests are optional analytical structures. They are not legal or insurance advice, underwriting guidance or standards, coverage opinions, claim demands, certification checklists, proof of insurability, insurer endorsements, regulator-approved methods, actuarial pricing guidance, premium recommendations, procurement requirements, or claim approval guidance.

Part IV has built the underwriting-facing architecture: evidence components, exposure inventory, non-pricing variables, renewal/change evidence, and optional reviewer requests. It does not say the architecture is accepted by the market. It says the architecture is what an enterprise can use to make agentic AI risk more legible before underwriting, renewal, reinsurance, counsel, and risk-engineering discussions.

Part V can now move to the other side of the loss event: claims, disputes, responsibility and coverage boundaries, and post-loss remediation evidence.

Part V: Claims, Disputes, and Post-Loss Responsibility Evidence

Part IV organized agentic AI risk before loss: exposure inventory, underwriting-facing evidence, non-pricing variables, renewal change records, and optional reviewer evidence requests. Part V moves to the other side of the event.

The question changes after a loss. The enterprise is no longer asking only whether an agentic workflow is understandable enough for a risk discussion. It is asking what happened, who or what acted, what authority existed, what evidence survived, what cannot be reconstructed, what policy boundary questions may arise, what was fixed, and what the event should change before the next review.

That is not the same as claim approval. It is not coverage determination. It is not legal causation. It is not a liability finding. It is post-loss responsibility evidence: the disciplined reconstruction of a bounded agentic work unit after something has gone wrong.

Agentic AI makes this harder because the loss may not be located in a single model output. A model may have produced a plausible answer. An agent may have converted it into an action. A tool may have changed a customer record. A human may have approved a summary without seeing the full evidence. A vendor may hold the logs. A privacy rule may limit disclosure. A model or runtime may have changed before anyone investigated. The claim file may contain fragments, but fragments are not reconstruction.

Part V therefore builds the post-loss layer of the paper. It asks how claim reconstruction, dispute handling, coverage-boundary questions, remediation closure, and renewal feedback can be organized without pretending that evidence decides the legal or insurance outcome.

The compliance white paper supplies lifecycle vocabulary for authority, accepted outcome, substitution, and remediation closure. The auditability and assurance white paper supplies audit-evidence vocabulary for joining traces, records, and source pointers. In Part V, both are used as analytical scaffolding only, not insurance proof. [54] [55]

Chapter 19: Claim Reconstruction After Agentic AI Incidents

An incident notice is not claim reconstruction.

"The AI system made an error" is a useful alarm, but it is a weak claim narrative. It does not say which work unit failed, what the agent was authorized to do, what the human saw, what tool action occurred, which data changed, what loss followed, whether the event was exceptional or ordinary, what was remediated, or which evidence is missing. It may be enough to start an internal response. It is not enough to reconstruct an agentic incident for insurance-facing review.

Claim reconstruction begins with the bounded work unit. The unit may be a refund workflow, customer notice workflow, claims triage workflow, professional deliverable workflow, deployment assistant, payment exception process, or account-update agent. The phrase "AI failure" should be translated into "this work unit acted under this authority, through this role and tool path, producing this consequence, with this evidence and these gaps."

Consider a customer-support agent that incorrectly triggers a refund and sends a confirmation email. The customer receives money they were not owed. The account record changes. A finance reconciliation process flags the mismatch two days later. The team can produce a model trace and an email record, but it cannot immediately answer whether the agent had refund authority, whether a human reviewer approved the action or only reviewed the message, whether the CRM update and payment action were linked, whether the tool action exceeded a threshold, whether the model endpoint changed that week, or whether the exception was closed.

That is the difference between notice and reconstruction.

The first reconstruction layer is initiating intent. What task was the work unit supposed to perform? Was it answering a customer question, correcting an account, processing a refund, triaging a complaint, drafting professional advice, or changing production code? Without intent, the event is only a sequence of technical actions.

The second layer is authority boundary. Was the agent allowed to recommend, draft, approve, send, pay, update, delete, deploy, file, or escalate? Was the action inside a threshold, outside a threshold, or ambiguous? Authority does not decide legal liability, but it frames whether the work acted as designed, drifted, or exceeded the expected perimeter.

The third layer is agent role. Did the agent draft, classify, recommend, decide, execute, monitor, or remediate? A model output is not the same as an agentic role. The reconstruction has to identify whether the agent merely produced text or actually caused a tool-mediated external consequence.

The fourth layer is human role. Human-in-the-loop is not enough. The file should show who reviewed, what they could see, when they reviewed, what authority they held, what criteria applied, and whether they accepted the outcome. A human reviewer who sees only a summary is not in the same position as a reviewer who sees the work-unit evidence chain.

The fifth layer is tool action. This is where generated output becomes business consequence: email sent, refund issued, account changed, code deployed, vendor ticket opened, database updated, filing submitted, credential used, cloud resource consumed, or customer notice published. Tool action is often the hinge between "the model said something" and "the organization did something."

The sixth layer is external consequence. Who or what was affected? A customer, vendor, employee, applicant, patient, investor, regulator, public audience, cloud bill, data system, production service, or professional client may each require different evidence. The consequence also shapes which internal teams, policy lines, vendors, counsel, or response partners may become relevant.

The seventh layer is affected data. Personal data, payment data, employee records, customer records, credentials, source code, privileged material, health data, vendor data, and operational logs carry different

evidence and disclosure constraints. Claim reconstruction that ignores data class may either over-disclose sensitive material or fail to preserve source pointers.

The eighth layer is evidence chain. Logs and traces are ingredients. A reconstruction needs source pointers that connect intent, authority, role, prompt/output, tool action, human confirmation, external consequence, exception, remediation, and gap records. NIST and CISA incident-response sources support structured incident timelines, containment, response, recovery, remediation, and tracking. They do not determine claim approval, legal causation, or coverage. [49] [50]

The ninth layer is exception record. Was there a warning, override, failed validation, repeated near miss, manual correction, vendor outage, permission change, abnormal consumption pattern, or threshold breach? In AI-linked cyber settings such as LLMjacking, QBE materials point to access, API usage, abnormal consumption, containment, and remediation evidence as concrete reconstruction inputs. That support remains cyber-evidence support, not a promise of policy response. [51]

The tenth layer is remediation action and closure state. Was the workflow disabled, permission changed, account corrected, affected party notified, customer reimbursed, code rolled back, model endpoint reverted, prompt updated, human review strengthened, or vendor ticket resolved? Was the work unit retested and reauthorized? Was residual risk accepted? Closure is a lifecycle state, not a settlement.

The final layer is missing evidence. Missing evidence should not be hidden inside narrative. It should be named. Vendor-held logs, expired retention, redacted records without source pointers, missing approval screens, overwritten prompts, unjoined payment records, unavailable cloud usage data, and unclear version history all shape what can and cannot be reconstructed.

T-19-01 - Agentic Claim Reconstruction Map

Bounded work unit

Reconstruction layer	Bounded work unit
Question it answers	Which agentic work is being reconstructed?
Evidence needed	Work unit ID, business function, owner, scope, status
What it cannot determine	Whether coverage applies
Boundary note	Not coverage opinion

Initiating intent

Reconstruction layer	Initiating intent
Question it answers	What was the work supposed to do?
Evidence needed	Task request, trigger, workflow purpose, input context
What it cannot determine	Whether the intent was legally sufficient
Boundary note	Not legal advice

Authority boundary

Reconstruction layer	Authority boundary
Question it answers	What was the agent allowed to do?
Evidence needed	Permission scope, threshold, escalation rule, prohibited actions
What it cannot determine	Whether authority creates legal liability
Boundary note	Not liability determination

Agent role

Reconstruction layer	Agent role
Question it answers	What did the agent contribute?
Evidence needed	Role record, prompt/output, plan, decision or action trace
What it cannot determine	Whether the agent is an insured subject
Boundary note	Analytical object only

Human role

Reconstruction layer	Human role
Question it answers	What did the human see, approve, or accept?
Evidence needed	Reviewer role, evidence visible, approval record, timing
What it cannot determine	Whether the human is legally liable
Boundary note	Not legal causation

Tool action

Reconstruction layer	Tool action
Question it answers	Where did output become action?
Evidence needed	API call, payment, email, account update, deployment, filing
What it cannot determine	Which policy line responds
Boundary note	Not policy interpretation

External consequence

Reconstruction layer	External consequence
Question it answers	Who or what was affected?
Evidence needed	Customer/vendor/employee/system/data impact record
What it cannot determine	Damages or covered loss amount
Boundary note	Not loss valuation

Affected data

Reconstruction layer	Affected data
Question it answers	What data classes were involved?
Evidence needed	Data map, labels, source pointers, redaction profile
What it cannot determine	Privacy compliance conclusion
Boundary note	Not privacy legal advice

Evidence chain

Reconstruction layer	Evidence chain
Question it answers	Can the event be joined end to end?
Evidence needed	Source pointers, logs, approvals, tool records, exception links
What it cannot determine	Claim approval or denial
Boundary note	Not claim approval guidance

Exception record

Reconstruction layer	Exception record
Question it answers	What warnings or abnormal events occurred?
Evidence needed	Override, near miss, anomaly, vendor outage, threshold breach
What it cannot determine	Fault allocation
Boundary note	Not fault determination

Remediation and closure

Reconstruction layer	Remediation and closure
Question it answers	What was contained, fixed, retested, and closed?
Evidence needed	Containment, fix, retest, reauthorization, residual-risk note
What it cannot determine	No residual liability
Boundary note	Not settlement proof

Missing evidence

Reconstruction layer	Missing evidence
Question it answers	What cannot be reconstructed?
Evidence needed	Missing, stale, redacted, vendor-held, overwritten, inaccessible records
What it cannot determine	Claim denial or coverage result
Boundary note	Gap register only

The important distinction is between four different kinds of inquiry. The technical sequence asks what systems, prompts, tools, identities, and records moved in what order. The operational incident asks what business process failed and how it was contained. Legal causation asks questions that belong to counsel, courts, dispute forums, contracts, and applicable law. Insurance claim analysis asks policy-specific questions about notice, facts, wording, exclusions, limits, deductibles, causation, loss category, and claim handling authority.

Agentic claim reconstruction can support the first two and organize facts for the latter two. It cannot replace them.

This chapter's boundary: claim reconstruction after agentic AI incidents is not legal advice, not insurance advice, not underwriting guidance, not coverage opinion, not claim approval guidance, not legal liability determination, not certification, not proof of insurability, not insurer endorsement, not regulator-approved method, not actuarial pricing guidance, and not premium recommendation.

The next chapter turns to what happens when the reconstruction is incomplete or contested. Agentic incidents do not only create loss. They create disputes about responsibility, authority, acceptance, and evidence gaps.

Chapter 20: Dispute, Responsibility, and Evidence Gaps

Disputes often begin where the evidence stops.

An enterprise may tell a customer that an AI agent made an error. The customer may say the company made a promise. The enterprise may tell an insurer that the incident was a covered cyber event. The insurer may ask whether it was authorized operational misuse. The enterprise may tell a vendor that the platform failed. The vendor may point to configuration. A business team may say legal approved the workflow. Legal may say it never saw the tool-action authority. Security may have logs. Privacy may restrict disclosure. Engineering may know the model endpoint changed. No one may own the complete reconstruction.

This is why evidence gaps should be explicit, not hidden.

A dispute can arise between enterprise and customer. The customer experienced the external consequence: wrong notice, wrong refund, wrong account status, wrong professional deliverable, wrong access decision, wrong service interruption. The enterprise needs to explain what happened without overclaiming what the evidence proves.

A dispute can arise between insured and insurer. The question may not be "did AI cause a loss?" but whether the event fits the policy wording, notice requirements, covered peril, loss category, exclusion, sublimit, deductible, or condition. This paper does not answer that. It helps show which facts would need to be reconstructed.

A dispute can arise between enterprise and vendor. The enterprise may depend on a model provider, orchestration platform, cloud service, identity provider, data processor, tool connector, or evidence repository. Vendor-held logs, service notices, model versions, runtime defaults, and API records can be essential. If they are unavailable, the dispute becomes partly an evidence-access dispute.

A dispute can arise around human and agent responsibility. The human may have approved an output, but not the action. They may have seen a summary, but not the source. They may have been responsible for exceptions, but not routine executions. The agent may have acted within configured authority, but the authority may have been poorly designed. Responsibility evidence should not be collapsed into legal liability conclusions.

A dispute can arise inside the enterprise. Business, security, legal, privacy, compliance, engineering, procurement, finance, and risk teams often hold different pieces of the event. Each team may have a true fragment. The review problem is that none of the fragments alone reconstructs the lifecycle work.

The most common gaps are ordinary. There is no authority record. The approval context is missing. The tool-action record is separate from the model trace. The human approval screen did not show the evidence the reviewer would later need. The vendor has logs but the contract does not make them available. Retention expired. Redaction removed the useful source pointer. A model substitution occurred without conformance review. A remediation ticket closed operationally without lifecycle closure.

Those gaps are not just embarrassing documentation defects. They define the boundary of what the event can say.

An evidence gap register should therefore include the missing item, why it is missing, who may control it, whether it is recoverable, whether the absence affects reconstruction, whether a privacy or privilege constraint applies, and whether future remediation is planned. The register should not be written as an accusation. It should be written as a map of uncertainty.

NIST and CISA sources support the value of incident tracking, coordination, recovery, and remediation records. SEC cyber disclosure rules support governance and material incident disclosure context for public companies, but they do not turn an evidence gap into a securities, coverage, or liability conclusion. [49] [50] [52]

T-20-01 - Dispute and Evidence Gap Register

Enterprise/customer

Dispute type	Enterprise/customer
Typical trigger	Wrong notice, refund, account change, decision, advice, or service action
Evidence gap	Missing accepted outcome or customer-impact record
Review consequence	Harder to explain what was accepted and communicated
Boundary note	Not customer liability advice

Insured/insurer

Dispute type	Insured/insurer
Typical trigger	Policy line, notice, exclusion, sublimit, causation, or loss category question
Evidence gap	Missing timeline, authority, tool action, or loss category evidence
Review consequence	Harder to frame claim facts
Boundary note	Not coverage opinion

Enterprise/vendor

Dispute type	Enterprise/vendor
Typical trigger	Platform, model, runtime, logging, API, or service dependency question
Evidence gap	Vendor-held logs, inaccessible service record, missing version history
Review consequence	Dependency responsibility remains unclear
Boundary note	Not vendor liability determination

Human/agent responsibility

Dispute type	Human/agent responsibility
Typical trigger	Human approved summary; agent executed action
Evidence gap	Missing reviewer-visible evidence or role boundary
Review consequence	HITL cannot be evaluated as responsibility structure
Boundary note	Not legal causation

Business/security/legal teams

Dispute type	Business/security/legal teams
Typical trigger	Different teams hold different fragments
Evidence gap	No shared incident evidence index
Review consequence	Reconstruction becomes narrative-driven
Boundary note	Not claim handling guidance

Privacy/evidence conflict

Dispute type	Privacy/evidence conflict
Typical trigger	Sensitive data appears in prompts, logs, payloads, or records
Evidence gap	Redacted evidence lacks source pointer
Review consequence	Review may lose factual continuity
Boundary note	Not privacy legal advice

Substitution dispute

Dispute type	Substitution dispute
Typical trigger	Model, vendor, runtime, prompt, or tool changed before/after event
Evidence gap	Missing change or conformance record
Review consequence	Reviewed object may not match incident object
Boundary note	Not vendor certification

Remediation dispute

Dispute type	Remediation dispute
Typical trigger	Fix claimed but closure unclear
Evidence gap	No retest, reauthorization, owner signoff, or residual-risk note
Review consequence	Event remains open as lifecycle object
Boundary note	Not settlement proof

Retention dispute

Dispute type	Retention dispute
Typical trigger	Logs or prompts expired before review
Evidence gap	Expired, overwritten, or unavailable evidence
Review consequence	Technical sequence may be incomplete
Boundary note	Not claim denial basis

Privilege/confidentiality dispute	
Dispute type	Privilege/confidentiality dispute
Typical trigger	Counsel/vendor/security records cannot be shared directly
Evidence gap	No protected source-pointer strategy
Review consequence	Review may need scoped access or summary
Boundary note	Not disclosure advice

The value of a gap register is discipline. It prevents a team from filling uncertainty with confidence. It also prevents the opposite mistake: treating one missing record as proof that nothing can be reconstructed. Some events can be reconstructed from multiple partial sources. Others cannot. The register helps reviewers see the difference.

Responsibility disputes should remain responsibility disputes until the right forum decides otherwise. A role map can show who initiated, configured, approved, executed, accepted, remediated, and closed. It cannot decide legal liability. An evidence chain can show what happened. It cannot decide coverage. A missing-evidence register can show uncertainty. It cannot by itself approve or deny a claim.

This chapter's boundary: dispute and evidence-gap analysis is not legal advice, not insurance advice, not underwriting guidance, not coverage opinion, not claim approval guidance, not legal liability determination, not certification, not proof of insurability, not insurer endorsement, not regulator-approved method, not actuarial pricing guidance, and not premium recommendation.

The next chapter addresses the most sensitive post-loss question: where the coverage boundary might sit. The paper can help frame the question. It cannot answer it.

Chapter 21: Coverage Boundary Analysis Without Coverage Opinion

Agentic incidents often cross insurance lines before anyone agrees what kind of loss occurred.

A payment agent sends funds to the wrong party after a manipulated email. Is the review about cyber, crime, social engineering, authorized instruction, operational error, professional services, or internal control failure? A coding agent deploys a bad change that interrupts customer service. Is the question cyber, tech E&O, business interruption, professional liability, contract, or service failure? A generative content workflow publishes a misleading statement. Is the issue media liability, professional liability, D&O disclosure, advertising, IP, or operational governance?

The answer depends on policy wording, jurisdiction, facts, notice, exclusions, limits, sublimits, deductibles, causation, loss category, and claim handling control. This paper does not interpret policy language. It does not provide coverage opinion.

What it can do is identify why agentic evidence matters to coverage-boundary questions.

The first boundary is cyber versus authorized operational misuse. If credentials were stolen, an API abused, data exposed, cloud resources consumed, or service interrupted through unauthorized access, cyber evidence may be central. If a permitted workflow performed the wrong business action under configured authority, the question may look different. QBE's AI-linked cyber and LLMjacking materials support the need for access, API usage, abnormal consumption, containment, and remediation evidence, but they do not decide policy response. [51]

The second boundary is technology E&O versus professional liability. A vendor-provided AI service may fail as a product or platform. A professional firm may deliver advice, analysis, design, legal, financial, engineering, medical, or consulting work using agentic support. The same model output can appear in both settings. The evidence question is who promised what service, who accepted the output, what human judgment applied, and where the tool action affected the client.

The third boundary is product versus service. An AI-enabled device, software product, SaaS feature, or embedded agent may create different questions from an internal workflow that supports a service. The evidence needs include product version, use context, model/runtime record, tool action, user interaction, monitoring, and remediation. The paper can name those evidence needs. It cannot determine whether a product or service policy responds.

The fourth boundary is D&O or governance versus operational failure. A board or officer may be relevant where oversight, disclosure, public statements, securities claims, cyber governance, or risk-management failures are alleged. SEC cyber disclosure rules support the importance of governance and material incident context for public companies. They do not create D&O coverage proof or determine securities liability. [52]

The fifth boundary is crime/social engineering versus cyber. Deepfake fraud, invoice manipulation, synthetic identity, agent-assisted payment approval, and credential misuse can straddle policy concepts. The reconstruction should preserve communication records, authentication evidence, approval steps, payment authorization, identity proofing, and tool actions. It should not assume the line outcome.

The sixth boundary is media/IP versus generated content workflow. A workflow may create text, code, images, audio, marketing claims, legal summaries, or public statements. The coverage question may depend on publication, rights review, source material, approval, takedown, and policy wording. The agentic evidence helps show what was generated, reviewed, accepted, and published.

The seventh boundary is business interruption or property versus cloud/API/service interruption. A shared cloud, model, API, identity, data, or vendor dependency may affect operations. The evidence question is dependency map, service status, incident timeline, affected work units, restoration steps, and loss

measurement. That evidence may help frame a boundary question, but it is not a business interruption determination.

The eighth boundary is privacy or regulatory investigation versus operational event. A prompt, memory, log, tool payload, customer record, employee record, or evidence pack may contain personal or regulated data. The claim file may need source pointers and redaction controls, but privacy/legal conclusions remain outside this paper.

Aon materials support the general reality that AI risk can touch cyber, E&O, professional liability, crime, D&O, governance, employment, and other enterprise risk lines. source research on silent exposure uses that market context to explain why line ambiguity matters. Neither source permits a universal coverage conclusion. [53] [56]

T-21-01 - Coverage Boundary Question Map

Cyber vs authorized operational misuse

Boundary question	Cyber vs authorized operational misuse
Agentic fact pattern	Agent uses valid credentials or API to perform wrong action
Evidence needed	Access records, authority scope, API logs, incident timeline
Why ambiguous	Unauthorized access and authorized misuse may require different review
Boundary note	Not coverage opinion

Tech E&O vs professional liability

Boundary question	Tech E&O vs professional liability
Agentic fact pattern	AI platform fails or professional output causes client impact
Evidence needed	Service promise, deliverable, human review, model/tool record
Why ambiguous	Product/service failure and professional judgment may overlap
Boundary note	Policy wording controls

Product vs service

Boundary question	Product vs service
Agentic fact pattern	Embedded AI feature or internal service workflow causes harm
Evidence needed	Product version, use context, workflow record, monitoring, remediation
Why ambiguous	Product defect, software, service, and contract theories may mix
Boundary note	Not product liability advice

D&O/governance vs operational failure

Boundary question	D&O/governance vs operational failure
Agentic fact pattern	Incident raises oversight, disclosure, or risk-management allegations
Evidence needed	Board/risk records, disclosure timeline, incident escalation
Why ambiguous	Governance claims may arise from operational events
Boundary note	Not securities or D&O advice

Crime/social engineering vs cyber

Boundary question	Crime/social engineering vs cyber
Agentic fact pattern	Agent-assisted payment, impersonation, or manipulated instruction
Evidence needed	Payment approval, authentication, communication, tool-action record
Why ambiguous	Voluntary instruction, fraud, cyber, and crime concepts may straddle
Boundary note	No line determination

Media/IP vs generated content workflow

Boundary question	Media/IP vs generated content workflow
Agentic fact pattern	Generated or agent-published content creates dispute
Evidence needed	Prompt/output, source material, rights review, approval, publication, takedown
Why ambiguous	Content, technology, professional, and advertising issues may overlap
Boundary note	Not IP or media coverage opinion

BI/property vs cloud/API/service interruption

Boundary question	BI/property vs cloud/API/service interruption
Agentic fact pattern	Shared model, cloud, API, identity, or vendor outage affects operations
Evidence needed	Dependency map, service status, affected work units, restoration, loss records
Why ambiguous	Operational interruption and covered interruption differ by wording
Boundary note	Not loss valuation

Privacy/regulatory vs operational event

Boundary question	Privacy/regulatory vs operational event
Agentic fact pattern	Personal data appears in prompts, logs, payloads, or evidence packs
Evidence needed	Data map, affected data, source pointers, redaction, notification records
Why ambiguous	Operational error may trigger privacy, regulatory, or cyber questions
Boundary note	Not GDPR/legal advice

Exclusion/sublimit boundary	
Boundary question	Exclusion/sublimit boundary
Agentic fact pattern	AI, cyber, professional, privacy, or technology terms may limit response
Evidence needed	Actual policy text, event facts, source evidence, loss category
Why ambiguous	Public market signals cannot replace policy wording
Boundary note	No exclusion application

Notice/control boundary	
Boundary question	Notice/control boundary
Agentic fact pattern	Event timing, notice, consent, defense, vendor involvement, or remediation control is contested
Evidence needed	Notice timeline, claim communications, response authority, vendor coordination
Why ambiguous	Policy conditions and claim handling authority are case-specific
Boundary note	Not claim handling guidance

The table should be read as a question map, not an answer key. It helps a team avoid two errors: treating every AI-linked event as one line, and treating line ambiguity as hopeless. The right posture is narrower. Preserve the facts. Separate the work unit. Identify the action. Map the affected data and consequence. Preserve policy-specific questions for authorized review.

This chapter's boundary: coverage-boundary analysis is not legal advice, not insurance advice, not underwriting guidance, not coverage opinion, not claim approval guidance, not legal liability determination, not certification, not proof of insurability, not insurer endorsement, not regulator-approved method, not actuarial pricing guidance, and not premium recommendation. Policy wording, jurisdiction, facts, notice, exclusions, limits, sublimits, deductibles, causation, loss category, and claim handling control remain external and case-specific.

The next chapter moves from boundary questions to post-loss closure. Even when the coverage outcome is unresolved, the enterprise still needs to contain, fix, retest, reauthorize, and record residual risk.

Chapter 22: Post-Loss Remediation, Reauthorization, and Residual Risk

An operational fix is not necessarily remediation closure.

After an agentic incident, teams often move quickly. Disable the workflow. Revoke a permission. Patch the prompt. Change the model endpoint. Roll back code. Refund a customer. Send a correction. Update a playbook. Close the ticket. The organization feels the event is over because operations are stable again.

From a lifecycle evidence perspective, the event may still be open.

Post-loss remediation should show what was contained, what caused the event, which work units were affected, which customers or third parties were affected, which data was involved, which authority boundary failed, which tool permissions changed, whether a model/vendor/runtime substitution mattered, whether the human review boundary was updated, what was retested, who reauthorized the workflow, what residual risk remains, and whether the renewal/change register was updated.

Take a payment workflow incident. An agent incorrectly approves a refund above its intended limit. The team disables the refund tool, manually reverses what it can, updates the threshold, and apologizes to affected customers. That may be a real operational response. But remediation closure still needs to show root cause, affected work units, affected customers, data involved, authority correction, tool permission correction, evidence chain preservation, retest, reauthorization, residual-risk acceptance, and any unresolved gaps.

The first stage is containment. What was stopped, isolated, disabled, suspended, blocked, or escalated? Containment evidence should identify the time, owner, work units affected, tool permissions changed, and immediate customer/system impact. NIST and CISA sources support containment, response, recovery, coordination, and tracking as incident-response disciplines, not insurance requirements. [49] [50]

The second stage is root-cause analysis. Root cause in this paper does not mean legal causation. It means the operational explanation the organization can support: authority misconfiguration, tool permission drift, missing human context, prompt failure, model/runtime change, vendor outage, data-quality issue, identity/access problem, exception-handling failure, or monitoring gap.

The third stage is affected work-unit mapping. Agentic systems are reused. A fix for one work unit may not cover another. The remediation record should identify all workflows that used the same model, tool connector, prompt pattern, permission, vendor, runtime, identity path, or evidence repository.

The fourth stage is affected customer, third-party, and data mapping. Who or what experienced consequence? Which data classes were touched? Was any evidence redacted, privileged, vendor-held, or restricted? This mapping helps preserve reviewability without turning the remediation record into legal advice.

The fifth stage is authority correction. If the agent had too much authority, the record should show the corrected permission, threshold, escalation path, prohibited action, or human confirmation boundary. If the authority was correct but misunderstood, the record should show clarification and training context.

The sixth stage is tool permission correction. The most important post-loss changes may be outside the model: API permission, CRM write access, payment threshold, deployment right, data export, email send capability, ticketing integration, or cloud console access. Tool records connect remediation to consequence.

The seventh stage is substitution review. If the model, vendor, runtime, orchestration layer, logging setting, data processor, or tool connector changed, the organization should ask whether the new component preserves the core lifecycle object fields defined earlier: authority, role, evidence, privacy, outcome, exception, and remediation. This continues the the substitution conformance logic without turning it into vendor certification.

The eighth stage is human review boundary update. If the incident showed that the reviewer lacked context, saw only a summary, reviewed after action, reviewed too many items, or lacked authority to block, the

remediation record should say what changed.

The ninth stage is retest and reauthorization. A fix should not only be made. It should be tested against the relevant work-unit scenario and reauthorized by an accountable owner. Reauthorization is not regulatory approval or insurer approval. It is an internal lifecycle state.

The tenth stage is residual-risk acceptance. Some risk remains after remediation. The question is whether the organization names it, owns it, monitors it, and carries it into renewal and future review.

T-22-01 - Post-Loss Remediation Evidence Map

Containment

Remediation stage	Containment
Evidence artifact	Disablement record, permission block, escalation time, affected work-unit list
Responsible owner	Security, engineering, business owner
Renewal relevance	Shows immediate control of event
Boundary note	Not claim approval guidance

Root-cause analysis

Remediation stage	Root-cause analysis
Evidence artifact	Operational cause note, event timeline, contributing factor map
Responsible owner	Incident owner, engineering, risk
Renewal relevance	Shows what changed after event
Boundary note	Not legal causation

Affected work units

Remediation stage	Affected work units
Evidence artifact	Shared model/tool/vendor/runtime/dependency map
Responsible owner	Enterprise risk, product owner
Renewal relevance	Updates exposure inventory
Boundary note	Not coverage-ready evidence

Affected customers/third parties/data

Remediation stage	Affected customers/third parties/data
Evidence artifact	Impact record, data-class map, source pointers, notification status if applicable
Responsible owner	Business owner, privacy, counsel
Renewal relevance	Updates consequence and privacy profile
Boundary note	Not legal notice advice

Authority correction

Remediation stage	Authority correction
Evidence artifact	Revised authority scope, threshold, escalation, prohibited action list
Responsible owner	Business owner, governance owner
Renewal relevance	Updates authority map
Boundary note	Not legal delegation proof

Tool permission correction

Remediation stage	Tool permission correction
Evidence artifact	API, payment, CRM, deployment, email, database, cloud permission change
Responsible owner	Engineering, security, system owner
Renewal relevance	Updates tool-action risk
Boundary note	Does not decide policy line

Model/vendor/runtime substitution review

Remediation stage	Model/vendor/runtime substitution review
Evidence artifact	Version record, vendor notice, runtime setting, conformance note
Responsible owner	Procurement, engineering, risk
Renewal relevance	Updates dependency and substitution register
Boundary note	Not vendor certification

Human review boundary update

Remediation stage	Human review boundary update
Evidence artifact	Reviewer role, evidence visible, timing, criteria, approval screen change
Responsible owner	Business owner, governance, compliance
Renewal relevance	Updates confirmation boundary
Boundary note	HITL is not proof of transferability

Retest

Remediation stage	Retest
Evidence artifact	Scenario test, regression evidence, exception test, monitoring check
Responsible owner	Engineering, risk, business owner
Renewal relevance	Supports reauthorization
Boundary note	Not assurance opinion

Reauthorization	
Remediation stage	Reauthorization
Evidence artifact	Owner signoff, reactivation decision, monitoring plan
Responsible owner	Accountable business owner
Renewal relevance	Shows lifecycle restart state
Boundary note	Not insurer approval

Residual-risk acceptance	
Remediation stage	Residual-risk acceptance
Evidence artifact	Residual-risk note, owner, monitoring, open gaps
Responsible owner	CRO/risk owner, business owner
Renewal relevance	Carries risk into renewal/change review
Boundary note	Does not prove no residual liability

Renewal/change register update	
Remediation stage	Renewal/change register update
Evidence artifact	Incident, remediation, substitution, gap, and authority updates
Responsible owner	Risk, broker-facing owner
Renewal relevance	Feeds future review
Boundary note	Not premium recommendation

Remediation closure is the state in which the organization can say: the event was contained, the work unit was understood, affected parties and data were mapped, authority and tool boundaries were corrected where needed, substitutions were reviewed, human review was updated where needed, the workflow was retested, an accountable owner reauthorized or retired it, residual risk is visible, and future review records were updated.

That is not the same as legal settlement. It is not claim closure. It is not proof of no liability. It is not a renewal guarantee. It is the lifecycle discipline that prevents incidents from remaining unresolved objects inside the enterprise.

This chapter's boundary: post-loss remediation and reauthorization evidence is a lifecycle closure record. It is not legal or insurance advice, underwriting guidance, coverage opinion, claim approval guidance, legal liability determination, certification, proof of insurability, insurer endorsement, regulator-approved method, actuarial pricing guidance, or premium recommendation. Remediation closure does not eliminate liability or guarantee renewal, coverage, pricing, or claim outcome.

The next chapter closes Part V by sending post-loss learning back into the pre-loss architecture. A serious incident should change the future risk file.

Chapter 23: Claims-to-Renewal Feedback Loop

A serious agentic incident should not disappear after operational closure.

If the event revealed unclear authority, weak human review, missing tool-action evidence, vendor-held logs, privacy/redaction tension, dependency concentration, model substitution, or unclosed remediation, those lessons should feed the next review. Otherwise, the enterprise repeats the same story at renewal: "we have AI governance," while the actual risk object remains hard to understand.

The claims-to-renewal feedback loop is risk learning. It is not pricing guidance. It does not say an incident will increase premium, reduce premium, trigger a surcharge, earn a discount, change appetite, or determine renewal. It says that post-loss facts should update the evidence architecture that future reviewers use.

The loop begins with the incident record. The record should identify the bounded work unit, event timeline, action, consequence, affected data, parties involved, source pointers, and open questions. It should avoid unsupported conclusions about liability, coverage, or claim outcome.

The second input is the evidence gap register. Gaps should not be forgotten because the event was contained. If vendor logs were inaccessible, retention was too short, redaction lacked source pointers, or a tool action was not joined to authority, the gap belongs in future evidence design.

The third input is remediation closure. What was fixed, retested, reauthorized, retired, or left open? Who owns residual risk? What monitoring changed? Which unresolved items should be visible at renewal?

The fourth input is exposure inventory. If the event showed that a work unit had more authority, data sensitivity, customer impact, dependency concentration, or reversibility difficulty than expected, the inventory should change. If the same component appears in other work units, those work units should be reviewed.

The fifth input is the authority map. Did the event reveal unclear delegation, excessive permissions, weak thresholds, missing escalation, emergency override misuse, or an approval role without enough evidence? Authority changes should not live only in engineering tickets.

The sixth input is the dependency map. A model, cloud region, API, identity provider, vendor, processor, evidence repository, or runtime may have become a concentration point. That dependency should be visible before the next broker, risk-engineering, underwriting, reinsurance, or enterprise risk review.

The seventh input is the privacy and redaction profile. Post-loss evidence often exposes where the organization either hoarded sensitive traces or failed to preserve useful source pointers. The next review should improve selective disclosure, access control, data labels, privilege flags, and retention design.

The eighth input is the renewal/change register. Part IV argued that agentic risk changes over time. Part V adds that incidents are change events. A serious incident should update workflow status, authority, tools, data classes, dependencies, human review, remediation closure, missing evidence, and residual risk.

The ninth input is the reviewer-facing request package. The next time a broker, risk engineer, counsel, internal reviewer, or underwriter asks for evidence, the package should reflect what the organization learned. It should not be a static template that ignores observed failure.

Different users will use this feedback differently. Enterprise risk teams use it to own the internal risk story. Engineering teams use it to improve controls and evidence capture. Counsel uses it to preserve legal boundaries. Brokers and risk engineers may use it to frame future conversations. Underwriters may ask their own questions under their own forms, appetite, and judgment. None of those uses converts the loop into a market-wide standard.

T-23-01 - Claims-to-Renewal Feedback Loop

Incident record

Post-loss signal	Incident record
Update required	Work-unit event timeline, action, consequence, affected data, source pointers
Who uses it	Enterprise risk, claims/IR, counsel
Why it matters	Preserves factual starting point
Boundary note	Not claim approval guidance

Evidence gap register

Post-loss signal	Evidence gap register
Update required	Missing, stale, vendor-held, redacted, privileged, or overwritten evidence
Who uses it	Risk, engineering, broker-facing owner
Why it matters	Prevents uncertainty from disappearing
Boundary note	Not claim denial basis

Remediation closure

Post-loss signal	Remediation closure
Update required	Containment, fix, retest, reauthorization, residual-risk note
Who uses it	Business owner, security, renewal team
Why it matters	Shows whether event became closed lifecycle object
Boundary note	Not settlement proof

Exposure inventory

Post-loss signal	Exposure inventory
Update required	Update authority, data, consequence, dependency, reversibility, reuse
Who uses it	CRO, broker, risk engineer
Why it matters	Keeps exposure units current
Boundary note	Not coverage-ready status

Authority map

Post-loss signal	Authority map
Update required	Correct delegation, threshold, escalation, approval, override
Who uses it	Governance, business owner, counsel
Why it matters	Prevents repeated authority ambiguity
Boundary note	Not legal delegation proof

Dependency map

Post-loss signal	Dependency map
Update required	Update model, cloud, API, vendor, processor, evidence system dependencies
Who uses it	Reinsurer-facing review, risk, engineering
Why it matters	Shows concentration and correlation
Boundary note	Not reinsurance pricing guidance

Privacy/redaction profile

Post-loss signal	Privacy/redaction profile
Update required	Add source pointers, access controls, data labels, retention, privilege flags
Who uses it	Privacy, counsel, risk engineer
Why it matters	Makes evidence usable without uncontrolled disclosure
Boundary note	Not privacy legal advice

Renewal/change register

Post-loss signal	Renewal/change register
Update required	Record incident, changes, substitutions, gaps, residual risk
Who uses it	Renewal team, broker, enterprise risk
Why it matters	Carries post-loss facts into future review
Boundary note	Not premium recommendation

Reviewer-facing request package

Post-loss signal	Reviewer-facing request package
Update required	Update examples, samples, gaps, dependency and remediation records
Who uses it	Broker, risk engineer, underwriter, counsel
Why it matters	Aligns future evidence requests with observed risk
Boundary note	Not underwriting checklist

Governance and training record	
Post-loss signal	Governance and training record
Update required	Update role training, approval criteria, monitoring, escalation playbooks
Who uses it	Governance, business units, HR/compliance where relevant
Why it matters	Shows learning and ownership
Boundary note	Not certification

The feedback loop also clarifies what should remain separate. Enterprise internal learning is not insurer acceptance. Broker or risk-engineering review is not underwriting decision. Underwriting review is not coverage opinion. Legal and coverage claim handling remains case-specific. Claim reconstruction is not claim approval. Remediation closure is not no-liability proof. Renewal learning is not premium guidance.

The value of the loop is memory. Agentic systems change quickly. Without a feedback loop, a serious incident becomes a story, then a ticket, then an archived file. With a feedback loop, it becomes a source of better exposure segmentation, better authority design, better evidence capture, better privacy controls, better dependency visibility, better remediation closure, and a more honest future risk discussion.

This chapter's boundary: claims-to-renewal feedback is not legal advice, not insurance advice, not underwriting guidance, not coverage opinion, not claim approval guidance, not legal liability determination, not certification, not proof of insurability, not insurer endorsement, not regulator-approved method, not actuarial pricing guidance, and not premium recommendation. It does not provide pricing, discount, surcharge, insurer appetite, renewal, binding, endorsement, or claim outcome claims.

Part V has moved the paper from pre-loss reviewability to post-loss responsibility evidence. It has shown that agentic AI claims need bounded work-unit reconstruction, explicit evidence gaps, careful coverage-boundary questions, remediation closure, and a claims-to-renewal learning loop. The next part can turn the body into the paper's concluding analytical models and final architecture, while preserving the same boundary discipline: models can organize risk reasoning, but they cannot create coverage.

Part VI: Final Analytical Models and Insurability Architecture

The paper has now moved through the full arc of the insurability problem.

Part I established that the insurance market is not answering AI risk with a single yes or no. It is dividing risk across affirmative products, exclusions and endorsements, limits and sublimits, silent exposures, cyber-linked events, model-performance warranties, professional liability ambiguity, governance exposure, aggregation concern, and claim reconstruction needs. Part II separated the insured legal subject from the loss-relevant agentic work object. Part III translated lifecycle governance and auditability into insurability reasoning without treating either as insurance proof. Part IV organized underwriting-facing reviewability. Part V organized post-loss responsibility evidence.

Part VI now names the final analytical architecture.

The goal is not to invent a standard. It is not to create a score. It is not to declare that an agentic AI system is insurable, coverage-ready, underwriting-ready, claim-ready, certified, accepted by insurers, or approved by regulators. The goal is narrower and more useful: to give readers a disciplined way to discuss the object of risk, the evidence around that object, the maturity of reviewability, the boundaries of non-claim language, and the final conclusion of the paper.

The final synthesis is deliberately conservative. Agentic AI risk becomes more serious for insurance-facing discussion when the work can be bounded, evidenced, reconstructed, and updated. That is still not coverage. It is the precondition for a better conversation about risk transfer.

Chapter 24: Agentic Insurability Object Model

The central object of this paper is not "AI" in the abstract.

It is not the model alone. It is not the prompt. It is not a cyber event label. It is not a log bundle. It is not a governance policy. It is not an audit file. It is not a claim file. It is the loss-relevant agentic work object: the bounded work through which a legal subject, human role, agent role, tool action, evidence chain, dependency context, accepted outcome, exception, remediation state, and renewal feedback become intelligible.

The Agentic Insurability Object Model is an authored analytical model for describing that object. It does not replace policy wording. It does not create an underwriting checklist. It does not certify a system. It does not determine whether any loss is covered. Its purpose is to keep the insurance conversation from collapsing into a vague sentence such as "we use AI" or "the AI failed."

The first layer is the insured legal subject. Insurance begins with a policyholder, insured organization, officer, professional, vendor, additional insured, or other legal subject named or treated under policy language. The agent is usually not that subject. The agent is part of the risk object through which the subject's operations, decisions, services, duties, or dependencies create loss-relevant facts.

The second layer is the loss-relevant agentic work object. This is the bounded unit of work that can be described with scope, owner, business function, permitted actions, affected parties, data classes, dependencies, and lifecycle state. A customer refund workflow, code deployment assistant, claims triage process, professional deliverable workflow, customer notice agent, payment exception workflow, or account-update agent can be a candidate object for analysis. A brand name for a model or platform is too broad.

The third layer is authority boundary. The object has to show what the agent was allowed to recommend, draft, decide, send, transact, update, delete, deploy, escalate, or remediate. Authority does not prove legal authority. It establishes the operational boundary within which the loss-relevant work acted.

The fourth layer is responsibility continuity. Agentic work often moves across humans, agents, tools, vendors, processors, and projects. A handoff is not risk transfer unless the responsibility and evidence trail survives the handoff. Responsibility continuity asks who initiated, configured, approved, executed, accepted, remediated, reauthorized, or retired the work object.

The fifth layer is tool-action consequence. A model output becomes insurance-relevant when it crosses into consequence: payment sent, record changed, email published, code deployed, account locked, data exported, filing submitted, vendor ticket created, cloud resource consumed, or advice delivered. Tool action is where "AI said" becomes "the organization did."

The sixth layer is evidence chain. Logs and traces matter, but they are not enough. The evidence chain must join intent, authority, role, tool action, human review, accepted outcome, exception, remediation, dependency, privacy treatment, and gaps. The auditability and assurance white paper's evidence vocabulary helps organize this chain, while incident response sources support the need for structured records. Neither determines coverage or claim outcome. [57] [60]

The seventh layer is privacy and selective disclosure profile. Evidence that cannot be shared safely may become unusable. Evidence hoarded without data minimization can create new exposure. The object therefore needs source pointers, redaction profiles, access controls, retention notes, privilege flags, and data-class labels. the compliance and auditability white papers provide analytical vocabulary for evidence partitioning and selective disclosure; they do not provide legal advice. [57]

The eighth layer is dependency and substitution context. The object may depend on a model, runtime, cloud region, API, identity provider, tool connector, evidence repository, processor, vendor, or reusable agent component. If one of these changes, the reviewed object may no longer match the operating object.

Dependency visibility also matters to aggregation and reinsurance-facing concern, though it does not create a capital model or pricing method. [61]

The ninth layer is accepted outcome. A workflow that has no accepted state remains difficult to review. Did the work close, escalate, fail, require exception, or remain open? Accepted outcome is not legal acceptance. It is an operational lifecycle state.

The tenth layer is exception, dispute, and remediation state. A serious incident should not end as a ticket note. The object needs exception history, dispute posture, remediation evidence, reauthorization or retirement, residual-risk note, and missing-evidence register.

The final layer is renewal and change feedback. Agentic risk changes. The object should carry what has changed: new tools, new authority, changed model/vendor/runtime, changed data classes, changed human review, incidents, near misses, remediation, unresolved gaps, and dependency concentration. The feedback loop makes the object current enough for future review.

This model differs from model governance because it is not centered on a model asset. It differs from cyber event logging because it is not centered only on access, credential, network, or resource-use traces. It differs from raw traces because it gives those traces responsibility semantics. It differs from an audit evidence chain because it is oriented toward insurability reasoning, not audit sufficiency. It differs from an underwriting checklist because it is not a required submission model. It differs from a claim file because it does not decide notice, coverage, causation, damages, settlement, or claim payment.

T-24-01 - Agentic Insurability Object Model

Insured legal subject

Object layer	Insured legal subject
What it captures	Policyholder, organization, officer, professional, vendor, or other legal subject
Why it matters	Keeps the insured subject separate from the agentic object
What it does not prove	That the agent is an insured subject
Boundary note	Policy wording controls

Agentic work object

Object layer	Agentic work object
What it captures	Bounded lifecycle work unit, owner, scope, business function, status
Why it matters	Turns vague AI use into reviewable work
What it does not prove	That the work is covered
Boundary note	Analytical construct only

Authority boundary

Object layer	Authority boundary
What it captures	Permitted and prohibited actions, thresholds, escalation, delegated authority
Why it matters	Shows what the agent was expected to do
What it does not prove	Legal delegation or liability
Boundary note	Not legal advice

Responsibility continuity

Object layer	Responsibility continuity
What it captures	Human, agent, corporate, vendor, processor, and remediation roles
Why it matters	Preserves accountability across handoffs
What it does not prove	Fault allocation
Boundary note	Not liability determination

Tool-action consequence

Object layer	Tool-action consequence
What it captures	API call, payment, email, deployment, record change, filing, export, service action
Why it matters	Locates where output became consequence
What it does not prove	Which policy line responds
Boundary note	Not coverage opinion

Evidence chain

Object layer	Evidence chain
What it captures	Source pointers joining intent, role, action, outcome, exception, remediation, and gaps
Why it matters	Makes reconstruction possible
What it does not prove	Claim approval or denial
Boundary note	Not claim approval guidance

Privacy/selective disclosure profile

Object layer	Privacy/selective disclosure profile
What it captures	Data class, redaction, access control, retention, privilege, disclosure path
Why it matters	Keeps evidence usable without uncontrolled retention
What it does not prove	Privacy compliance conclusion
Boundary note	Not privacy legal advice

Dependency/substitution context	
Object layer	Dependency/substitution context
What it captures	Model, vendor, runtime, cloud, API, processor, tool connector, evidence repository
Why it matters	Shows concentration and object change
What it does not prove	Reinsurer acceptance or vendor certification
Boundary note	Not standard or certification

Accepted outcome	
Object layer	Accepted outcome
What it captures	Closure, escalation, rejection, exception, accepted state, unresolved state
Why it matters	Shows whether work reached an accountable endpoint
What it does not prove	Legal acceptance
Boundary note	Lifecycle state only

Exception/dispute/remediation state	
Object layer	Exception/dispute/remediation state
What it captures	Exception record, dispute gap, containment, fix, retest, reauthorization, residual risk
Why it matters	Prevents incidents from becoming unowned fragments
What it does not prove	Settlement or no-liability proof
Boundary note	Not legal proof

Renewal/change feedback	
Object layer	Renewal/change feedback
What it captures	Updated authority, exposure, dependency, privacy, evidence, and gap records
Why it matters	Keeps future review aligned with changed risk
What it does not prove	Renewal outcome or premium effect
Boundary note	Not premium recommendation

The model improves risk reasoning because it forces the reader to ask what object is being reviewed. It does not create coverage. It does not create insurability. It does not determine liability. It does not prove insurer acceptance. It gives the conversation a subject, object, evidence boundary, and update path.

This chapter's boundary: the Agentic Insurability Object Model is not legal advice, not insurance advice, not underwriting guidance, not coverage opinion, not claim approval guidance, not legal liability determination, not certification, not proof of insurability, not insurer endorsement, not regulator-approved method, not actuarial pricing guidance, not premium recommendation, not a score, not a public standard, and not readiness certification.

The next chapter turns from the object model to a non-scoring reasoning model. Once the object can be described, the next question is how reviewable it is without pretending reviewability is an insurance outcome.

Chapter 25: Agentic Insurability Reasoning Model

The word "readiness" is dangerous if it sounds like an outcome.

An enterprise can be more ready for a risk discussion without being coverage-ready. It can produce better evidence without being underwriting-ready. It can reconstruct an incident without being claim-ready. It can have stronger governance without being certified. It can move from vague AI use to bounded agentic work without proving insurability.

For that reason, this chapter uses a non-scoring Agentic Insurability Reasoning Model. It is an authored analytical model for discussing reviewability. It is not a benchmark, rating model, certification, underwriting rule, insurer-adopted method, score, grade, pass/fail test, market-acceptance label, or policy condition.

The model describes six reasoning states.

The first state is opaque and not reconstructable. The enterprise can say it used an AI system, but it cannot identify the bounded work unit, authority, tool action, evidence chain, accepted outcome, dependency path, privacy treatment, exception history, remediation closure, or owner. This state is not a declaration that the risk is uninsurable. It is a statement that the risk is too vague to review seriously.

The second state is log-visible but responsibility-poor. The enterprise has prompts, outputs, traces, API calls, or system logs. It can show activity, but not responsibility semantics. It cannot show what the agent was allowed to do, what the human saw, which role accepted the outcome, who owned the consequence, or how remediation closed. This state often creates false comfort: there is data, but not enough meaning.

The third state is evidence-linked but boundary-incomplete. The enterprise can connect some records across the work unit. It can show intent, action, and consequence in part. But authority boundaries, privacy controls, human-role clarity, accepted outcome, vendor-held evidence, substitution records, or remediation closure remain incomplete. The object is emerging, but the review still has material gaps.

The fourth state is reviewable work-unit architecture. The enterprise can describe bounded work units with authority, role, tool action, evidence, privacy, dependency, accepted outcome, exception, remediation, and missing-evidence fields. The review can now move from "AI system" to a concrete risk object. This still does not mean an insurer will accept, quote, bind, renew, cover, or pay a claim.

The fifth state is underwriting-facing evidence architecture. The enterprise can organize pre-bind, runtime, post-incident, renewal, dependency, privacy, and change evidence so that a broker, risk engineer, enterprise risk team, counsel, or underwriter can ask better questions. It remains optional and analytical. It is not an underwriting standard or mandatory request.

The sixth state is post-loss reconstructable and renewal-updatable. The enterprise can reconstruct incidents by bounded work unit, identify gaps, frame coverage-boundary questions without answering them, evidence remediation closure, and feed lessons into renewal/change review. This is the strongest reviewability state in the model, but it still does not decide coverage, liability, premium, renewal, or claim outcome.

The stages are intentionally non-numeric. A number would imply precision the sources do not support. A score would invite false use as a checklist, rating factor, certification threshold, or market signal. The model instead asks a sequence of questions: can the object be named, can authority be bounded, can responsibility survive handoff, can action be tied to consequence, can evidence be reconstructed, can privacy be preserved, can dependencies be seen, can remediation close, and can change feed the next review?

The compliance and auditability white papers provide analytical foundations for this reasoning. The compliance white paper helps describe lifecycle governance objects, authority, evidence partitioning, accepted outcome, substitution, and remediation closure. The auditability and assurance white paper helps distinguish raw traces

from evidence chains and auditability from sufficiency. Neither converts a reasoning state into insurance proof. [57]

External market and governance sources explain why this matters. Public insurance sources show a split market; governance and incident sources show the need for documentation, controls, and reconstruction; aggregation sources show why dependency visibility matters. They do not endorse this model or turn it into a market standard. [58] [59] [60] [61]

T-25-01 - Non-Scoring Insurability Reasoning Model

Reasoning state	What exists	What remains missing	Risk-review implication	Boundary note
Opaque / not reconstructable	General AI use description, vendor or model name, business claim	Bounded work unit, authority, role, action, evidence, outcome, owner	Risk is too vague for disciplined review	Not a denial or underwriting result
Log-visible but responsibility-poor	Prompts, outputs, traces, API logs, system events	Human/agent role semantics, authority, accepted outcome, remediation owner	Activity is visible but accountability is weak	Logs are not insurability
Evidence-linked but boundary-incomplete	Partial chain across intent, action, consequence, and records	Complete authority, privacy, dependency, substitution, exception, closure fields	Some review is possible, but major gaps remain	No coverage-ready claim
Reviewable work-unit architecture	Bounded work objects with authority, role, action, evidence, privacy, dependency, outcome, and gaps	External policy review, loss history, insurer appetite, line-specific requirements	Risk object can be discussed with more precision	Not insurer acceptance
Underwriting-facing evidence architecture	Pre-bind, runtime, post-incident, renewal, dependency, privacy, and change evidence organized for review	Insurer-specific forms, policy wording, pricing, appetite, underwriting judgment	Review questions become more concrete	Not underwriting standard
Post-loss reconstructable and renewal-updatable	Incident reconstruction, gap register, remediation closure, feedback loop, change register	Coverage, liability, damages, settlement, renewal, premium outcome	Strongest reviewability posture in this model	Not claim approval or premium guidance

A higher state means better reviewability, not guaranteed insurability. It means the enterprise can answer better questions and expose gaps more honestly. It does not mean the risk is acceptable, priced, covered, certified, or approved.

This chapter's boundary: the Agentic Insurability Reasoning Model is not legal advice, not insurance advice, not underwriting guidance, not coverage opinion, not claim approval guidance, not legal liability determination, not certification, not proof of insurability, not insurer endorsement, not regulator-approved method, not actuarial pricing guidance, not premium recommendation, not a score, not a standard, not readiness certification, and not an insurer-adopted method.

The next chapter translates the final architecture back to the paper's readers. Each audience can use the model differently, but none should mistake it for advice, policy interpretation, or obligation.

Chapter 26: What Enterprises, Brokers, Insurers, Reinsurers, and Counsel Should Take Away

The paper's final models are useful only if different readers know what to do with them without overclaiming them.

For enterprises, the takeaway is not "buy AI insurance" or "prove insurability." It is more practical: stop describing exposure only as AI adoption. Inventory bounded agentic work units. Name who owns them. Record authority boundaries. Capture tool-action consequences. Preserve evidence chains. Keep privacy and selective disclosure profiles. Map dependencies. Record substitutions. Close remediation. Carry incidents into renewal/change review.

For brokers and risk engineers, the takeaway is translation. A client may arrive with a slide that says "we use copilots," "we have an agent platform," or "we implemented AI governance." The useful move is to translate that into risk-object questions: which work units act, what authority they carry, what consequences they can produce, what evidence exists, what data is touched, what dependencies concentrate exposure, and what cannot be reconstructed.

For insurers, the takeaway is legibility. This paper does not tell insurers how to underwrite agentic AI. It does not define appetite. It does not recommend forms, exclusions, endorsements, sublimits, pricing, or claim handling. It says that agentic risk becomes more legible when the insured subject, work object, authority, role, tool action, evidence chain, dependency, privacy, accepted outcome, exception, remediation, and change feedback are separated. Policy language, appetite, portfolio judgment, and claim handling remain external.

For reinsurers, the takeaway is dependency visibility. Agentic AI can reuse the same model, cloud provider, runtime, API, identity path, data processor, evidence repository, or agent component across many business units or insureds. Aggregation concern is not only single-incident severity. It is the possibility that shared dependencies create correlated event shapes. Dependency visibility improves the conversation; it does not produce a capital model. [61]

For counsel, privacy, and governance teams, the takeaway is separation. Evidence reviewability is not legal conclusion. Governance is not liability proof. Selective disclosure is not legal advice. Retention is not automatically good if it hoards sensitive traces. Redaction is not automatically good if it destroys source pointers. The task is to preserve useful evidence while protecting legal, privacy, privilege, and confidentiality boundaries.

For boards, CROs, CIOs, CTOs, CCOs, and AI governance leaders, the takeaway is discipline. AI governance that cannot reconstruct lifecycle work is weak for insurance-facing review. A model inventory alone will not answer who acted, under what authority, through what tool, with what human role, producing what consequence, supported by what evidence, remediated by whom, and updated before the next review.

For claims and incident-response teams, the takeaway is continuity. A technical timeline, incident ticket, vendor notice, security alert, or model trace may be necessary. None is sufficient alone. Claim reconstruction needs the bounded work unit and its responsibility semantics. Incident response sources support the importance of structured response, recovery, remediation, and tracking, but they do not decide insurance outcomes. [60]

For implementation leaders, the takeaway is architecture. If agentic work is designed without authority boundaries, evidence pointers, privacy profiles, dependency records, accepted outcome states, exception paths, remediation closure, and change registers, the insurance-facing review problem is already being built into the system.

T-26-01 - Audience Takeaway Matrix

Audience	Practical takeaway	Evidence question to ask	Boundary to preserve	Next review focus
Enterprise risk / CRO	Inventory bounded agentic work units and gaps	Which work can cause loss, and what evidence describes it?	Inventory is not proof of insurability	Exposure and gap register
CFO / finance	Distinguish exposure variables from pricing outcomes	What transactions, values, volumes, and reversibility issues exist?	No premium recommendation	Non-pricing exposure view
CTO / CIO / engineering	Build evidence and authority into agentic workflows	Where do output, tool action, evidence, and dependency records join?	Technical trace is not coverage evidence by itself	Work-unit evidence design
AI governance leader	Connect governance to lifecycle work, not policy slogans	Which authority, role, outcome, exception, and closure records exist?	Governance is not claim approval	Lifecycle control architecture
Broker / risk engineer	Translate vague AI use into risk-object discussion	Which work objects, lines, consequences, and gaps need framing?	Not an underwriting checklist	Reviewer evidence package
Insurer / underwriter	Use object clarity to ask better questions where relevant	What evidence makes the risk legible under the insurer's own process?	No insurer acceptance implied	Policy/appetite-specific review
Reinsurer / portfolio reviewer	Look for shared dependency and correlated event shape	Which models, clouds, APIs, vendors, and reusable agents concentrate exposure?	Not capital or pricing model	Aggregation visibility
Counsel	Keep evidence review separate from legal conclusion	What facts can be preserved without over-disclosure or privilege loss?	Not legal advice	Boundary and privilege posture
Privacy / data governance	Make selective disclosure usable	What personal or sensitive data appears in traces and evidence packs?	Not privacy compliance opinion	Redaction and source-pointer design
Board / senior leadership	Ask whether AI governance can reconstruct work	Can the organization explain who or what acted, with what authority, and what changed?	Not fiduciary or D&O advice	Oversight evidence posture
Claims / incident response	Start reconstruction from the bounded work unit	What happened, what is missing, and what was remediated?	Not claim approval guidance	Post-loss evidence pack
Implementation leader	Design agentic work for future reviewability	What records will survive handoff, substitution, incident, and renewal?	Not certification path	Evidence-by-design backlog

These takeaways are not obligations. They are not legal requirements, insurance requirements, procurement requirements, regulator-approved requirements, or insurer requirements. They are practical implications of the paper's argument: if the work object cannot be bounded and evidenced, the risk conversation remains vague.

This chapter's boundary: the audience takeaways are not legal advice, not insurance advice, not underwriting guidance, not coverage opinion, not claim approval guidance, not legal liability determination, not certification, not proof of insurability, not insurer endorsement, not regulator-approved method, not actuarial pricing guidance, not premium recommendation, not a score, not a standard, and not readiness certification.

The next chapter gathers the paper's remaining caveats. The point is not to weaken the argument. The point is to protect it from becoming something the sources do not support.

Chapter 27: Residual Caveats and Non-Claim Discipline

Boundary discipline makes the paper stronger.

The temptation in a new risk field is to overstate the model. A vocabulary becomes a checklist. A checklist becomes a standard. A standard becomes a certification. A certification becomes an underwriting shortcut. A shortcut becomes a coverage promise. That sequence is exactly what this paper refuses.

AIIRWP v1.0 is a whitepaper for analytical synthesis. It is not a legal opinion. It is not insurance advice. It is not underwriting guidance. It is not actuarial pricing guidance. It is not a claim handling manual. It is not a coverage opinion. It is not a certification scheme. It is not a proof of insurability. It is not an insurer endorsement. It is not a regulator-approved method. It is not a procurement recommendation. It is not a vendor ranking. It is not a public standard.

The market sources should be read as context. Verisk, QBE, Munich Re, Armilla, Chaucer, Aon, Geneva Association, Swiss Re, Beazley, NIST, CISA, SEC, NAIC, and other sources support particular claims in particular ways. They do not collectively prove that agentic AI lifecycle risk is accepted by the insurance market. They show a split market, narrower products, cyber-linked channels, model-performance structures, governance and documentation expectations, incident response discipline, and aggregation concerns. [58] [59] [60] [61]

Cyber, cloud, and reinsurance sources are sometimes analogies. They are strong analogies where agentic AI depends on shared cloud, APIs, identity systems, models, vendors, processors, evidence repositories, and automated resource use. They are still analogies unless directly tied to AI or to a specific insurance product source. The paper should preserve that distinction in full-body assembly.

Policy language remains external. Exact policy wording, exclusions, endorsements, sublimits, deductibles, notice conditions, defense/control provisions, covered loss definitions, and claim forms require primary-source review. Public product pages and broker reports can support market context, not policy interpretation.

Exclusion and sublimit signals remain caveated where primary filings or policy wording are not available. the source research already records source gaps for AIG, WR Berkley, Great American, Beazley/QBE sublimit details, and exact Verisk form text. Later publication-stage work should recheck those sources before using precise language. [62]

Insurer claim documentation and AI underwriting questionnaire sources remain useful possible strengthening sources if later found. Their absence does not block the body draft, but it should keep the paper modest about claims handling and underwriting evidence. The paper can say what evidence would make a risk more reviewable. It should not say what any insurer will require, accept, price, bind, renew, or pay.

The compliance and auditability white papers are analytical foundations, not external insurance facts. The compliance white paper helps describe lifecycle governance objects. The auditability and assurance white paper helps describe audit evidence chains. This paper uses both to build insurability reasoning. It does not say the compliance white paper or the auditability and assurance white paper makes a system insurable, coverage-ready, underwriting-ready, claim-ready, certified, insurer accepted, or legally sufficient. [57]

The model names also require discipline. Agentic Insurability Object Model and Agentic Insurability Reasoning Model are authored analytical constructs. They can organize questions. They cannot score, certify, approve, validate, bind, insure, price, settle, or rank.

T-27-01 - Final Non-Claim Boundary Register

Restricted claim	Why it matters	Safer language	Where to preserve	Boundary note
Legal advice	Legal conclusions depend on law, facts, forum, and counsel	"legal questions remain external"	Front matter, Ch. 19-28, Appendix boundary	Not legal advice
Insurance advice	Coverage decisions depend on policies, facts, and authorized professionals	"insurance-facing review" or "risk discussion"	Whole paper	Not insurance advice
Coverage opinion	Policy wording, exclusions, limits, and facts control	"coverage-boundary question"	Ch. 11, 21, 27	Not coverage opinion
Underwriting guidance	Insurer appetite and process are external	"underwriting-facing evidence"	Ch. 6, 14, 18, 25	Not underwriting standard
Actuarial pricing guidance	Pricing requires actuarial data, filings, models, and insurer judgment	"non-pricing exposure variable"	Ch. 16, 23, 27	No pricing formula
Premium recommendation	Premium outcome is case-specific and insurer-controlled	"risk-review variable"	Ch. 16, 23, 26	No premium recommendation
Claim approval guidance	Claim outcomes depend on policy, facts, law, and claim handling	"claim reconstruction evidence"	Ch. 7, 10, 19, 23	Not claim approval
Legal liability determination	Liability depends on law, contracts, facts, and forum	"responsibility evidence"	Ch. 5, 20, 22	Not liability proof
Certification	Certification implies external validation authority	"authored analytical model"	Ch. 24-27	Not certification
Proof of insurability	Evidence improves reviewability, not outcome	"reviewability"	Whole paper	Not proof of insurability
Insurer endorsement	No insurer has endorsed these models as a method	"source-grounded synthesis"	Source notes, Ch. 24-27	No endorsement
Regulator-approved method	Regulator sources support context, not this model	"regulatory context"	Ch. 9, 14, 27	Not regulator approved
Procurement recommendation	Procurement choices require separate review	"implementation or vendor decisions remain external"	Ch. 26-27	No procurement recommendation
Vendor ranking	The paper is not evaluating vendors	"dependency context"	Ch. 12, 15, 24, 26	No vendor ranking
Public standard	A whitepaper model is not a public standard	"analytical construct"	Ch. 24-27	Not a standard
Score or grade	Numeric labels imply unsupported precision	"reasoning state"	Ch. 25	Not a score
Coverage-ready / underwriting-ready / claim-ready	Readiness labels imply outcome or acceptance	"more reviewable"	Ch. 25-28	No readiness certification
External adoption	No adoption evidence is claimed	"proposed by this paper"	Whole paper	No adoption claim
Certification, external approval, or publication overclaim	Separate owner authorization is required	"draft body" or "rewrite in progress" in governance files	Governance only, not body	No certification or external approval claim

This register is not defensive padding. It is the condition for credible synthesis. The paper asks insurance readers to take agentic AI risk seriously. That requires not pretending that evidence is insurance, governance is coverage, auditability is insurability, or a model name is a market outcome.

This chapter's boundary: residual caveats and non-claim discipline are not legal advice, not insurance advice, not underwriting guidance, not coverage opinion, not claim approval guidance, not legal liability determination, not certification, not proof of insurability, not insurer endorsement, not regulator-approved method, not actuarial pricing guidance, not premium recommendation, not a score, not a standard, and not readiness certification.

The final chapter returns to the thesis. Agentic AI risk is not a single market verdict. It is an object problem, an evidence problem, and a responsibility problem.

Chapter 28: Conclusion: From AI Risk Noise to Agentic Risk Objects

"Is AI insurable?" is the wrong opening question.

Asked in the abstract, it produces noise. One reader thinks about hallucinated professional advice. Another thinks about stolen LLM credentials. Another thinks about model-performance warranty. Another thinks about a customer refund agent. Another thinks about board oversight. Another thinks about cloud concentration. Another thinks about a policy exclusion. Another thinks about a claim file with missing logs. All of them are talking about AI risk. They are not talking about the same risk object.

The better question is: what exactly is being transferred, by whom, through what object, under what authority, with what evidence, across what boundary, and after what event?

This paper's answer is that AI risk is not insurable or uninsurable in the abstract. It becomes more reviewable when the loss-relevant agentic work object can be bounded, evidenced, reconstructed, and updated. The insured legal subject remains the person or organization under policy language. The agentic risk object is the work through which loss-relevant action, consequence, evidence, and responsibility become visible.

Part I showed why the market forces this question. Public sources do not show a single AI insurance answer. They show affirmative AI products, model-performance warranties, AI-linked cyber coverages, exclusion and endorsement development, sublimit and cap signals, silent exposures across existing lines, governance and disclosure context, claim reconstruction needs, and aggregation concern. That split means the paper cannot begin with "AI covered" or "AI excluded." It has to begin with object, evidence, boundary, and line.

Part II defined the insurable agentic risk object. The policyholder remains the legal insured subject. The work object is the loss-relevant lifecycle unit that can carry authority, role, tool action, evidence, accepted outcome, exception, remediation, and closure. Human-in-the-loop alone is not a responsibility structure. A tool action matters because it turns output into consequence. Hard-to-insure patterns emerge when authority, responsibility, evidence, privacy, substitution, or closure cannot be reconstructed.

Part III translated compliance and auditability concepts into risk-transfer analysis without letting them dominate it. The compliance white paper contributes lifecycle governance vocabulary. The auditability and assurance white paper contributes auditability and evidence-chain vocabulary. Auditability is necessary for reconstruction but not sufficient for insurability. Compliance is not auditability. Auditability is not coverage. Evidence chains are not claim approval.

Part IV built the underwriting-facing architecture: work-unit inventory, exposure segmentation, non-pricing variables, renewal/change/substitution evidence, and optional reviewer-facing evidence requests. It did not turn those requests into a standard, turn variables into pricing guidance, or claim that any insurer will accept or reward the architecture.

Part V moved to the post-loss side. Incident notice is not reconstruction. Coverage-boundary analysis is not coverage opinion. Dispute mapping is not liability determination. Remediation closure is not settlement. Claims-to-renewal feedback is not premium guidance. Post-loss evidence matters because agentic AI incidents can leave fragments unless responsibility, evidence, and lifecycle state are joined.

Part VI names the final analytical architecture. The Agentic Insurability Object Model describes what the risk object must contain. The Agentic Insurability Reasoning Model describes reviewability states without scoring them. The audience matrix translates the argument into practical questions. The boundary register keeps the paper from becoming advice, standard, certification, or market claim.

The final rule is simple:

Reviewability is not coverage.

Evidence is not insurance.

Governance is not claim approval.

Auditability is not insurability.

Without bounded lifecycle evidence, agentic AI risk remains too vague to discuss seriously.

That is the contribution of the paper. It does not promise that agentic AI risk can always be transferred. It does not declare that agentic AI risk cannot be transferred. It says the serious conversation begins only when the insured legal subject and the agentic risk object are separated, the object is bounded, authority is visible, responsibility survives handoff, tool action is tied to consequence, evidence is reconstructable, privacy is preserved, dependencies are mapped, remediation closes, and change feeds future review.

This public research edition exposes bounded public HTML, PDF, manifest, and checksum artifacts. It does not create certification, external approval, DOCX distribution, or source Markdown publication.

This conclusion does not claim certification, external approval, insurer acceptance, coverage-ready status, underwriting-ready status, claim-ready status, certification, endorsement, legal proof, insurance advice, legal advice, coverage opinion, underwriting standard, actuarial pricing guidance, premium recommendation, claims approval guidance, external adoption, indexing, SEO-GEO outcome, answer-engine recognition, or future implementation.

Appendices

Integrated Appendix Status: Appendix A-H are included as reference material after Chapter 28. They do not revise the body, do not create standards, checklists, certifications, insurer-adopted methods, coverage opinions, claim guidance, pricing guidance, or external approval status.

Appendix A - Agentic Insurability Object Model Reference

Purpose

Appendix A provides a compact reference version of the Agentic Insurability Object Model introduced in Chapter 24 and supported by Chapters 4, 5, 6, 7, 9, 14, 19, 22, and 23. It is intended to help a reader see the loss-relevant agentic work object as a bounded analytical unit rather than as a general AI system, a raw log set, or a governance program.

This appendix references body tables T-04-01, T-05-01, T-06-01, T-07-01, T-09-01, T-14-01, T-19-01, T-22-01, T-23-01, and T-24-01.

Reference Object Layers

Insured legal subject

Object layer	Insured legal subject
Field or evidence focus	Named entity, business unit, contracting party, or responsible legal person associated with the activity.
Body relationship	Chapters 4 and 5; T-04-01 and T-05-01.
Reviewability function	Separates the person or entity whose risk may be reviewed from the technical artifact performing work.
Boundary note	Does not determine legal liability, coverage, or insurer acceptance.

Loss-relevant agentic work object

Object layer	Loss-relevant agentic work object
Field or evidence focus	The bounded task, workflow, decision support step, transaction, or operational action in which an agent participates.
Body relationship	Chapters 4, 6, 14, and 24; T-04-01, T-06-01, T-14-01, T-24-01.
Reviewability function	Provides the unit around which authority, evidence, consequence, and responsibility can be organized.
Boundary note	Not a policy definition, coverage trigger, or proof of insurability.

Initiating intent

Object layer	Initiating intent
Field or evidence focus	Business purpose, user instruction, system objective, or workflow trigger.
Body relationship	Chapters 6, 7, 19; T-06-01, T-07-01, T-19-01.
Reviewability function	Helps reconstruct why the agentic work object began and whether later actions stayed inside the intended work unit.
Boundary note	Does not prove authorization, compliance, or proper use by itself.

Delegated authority boundary

Object layer	Delegated authority boundary
Field or evidence focus	Permitted action range, approval threshold, escalation rule, tool permission, or blocked action.
Body relationship	Chapters 5, 6, 14, 19; T-05-01, T-06-01, T-14-01, T-19-01.
Reviewability function	Shows whether the agentic action can be compared with an explicit boundary rather than inferred after the fact.
Boundary note	Not underwriting guidance, legal advice, or claim approval guidance.

Agent role

Object layer	Agent role
Field or evidence focus	The agent's assigned function, autonomy level, tool access, and workflow position.
Body relationship	Chapters 4, 6, 14, 24; T-04-01, T-06-01, T-14-01, T-24-01.
Reviewability function	Identifies what the agent was expected to do and which actions were within the modeled work object.
Boundary note	Does not classify the agent as insurable or uninsurable.

Human role

Object layer	Human role
Field or evidence focus	Human initiator, reviewer, approver, override actor, escalation recipient, or post-loss reviewer.
Body relationship	Chapters 5, 7, 19; T-05-01, T-07-01, T-19-01.
Reviewability function	Preserves responsibility continuity between automated action and human governance.
Boundary note	Does not decide negligence, liability, employment responsibility, or claims outcome.

Vendor/runtime/model context

Object layer	Vendor/runtime/model context
Field or evidence focus	Model, tool, agent framework, cloud service, API, version, runtime configuration, or vendor-held dependency relevant to the work object.
Body relationship	Chapters 12, 14, 17, 23; T-12-01, T-14-01, T-17-01, T-23-01.
Reviewability function	Makes dependency, substitution, concentration, and version-change questions visible.
Boundary note	Does not endorse or rank any vendor, model, or platform.

Tool-action consequence

Object layer	Tool-action consequence
Field or evidence focus	External command, transaction, data movement, message, system change, or operational effect.
Body relationship	Chapters 6, 7, 19, 24; T-06-01, T-07-01, T-19-01, T-24-01.
Reviewability function	Connects agentic behavior to potential loss, dispute, remediation, or renewal review.
Boundary note	Does not prove causation, damages, coverage, or claim payment entitlement.

Evidence chain

Object layer	Evidence chain
Field or evidence focus	Pointers to logs, approvals, prompts, tool-call records, exception records, human review, and remediation evidence.
Body relationship	Chapters 7, 10, 14, 19; T-07-01, T-10-01, T-14-01, T-19-01.
Reviewability function	Supports reconstruction of the work unit without treating raw traces as sufficient evidence.
Boundary note	Evidence improves reviewability; it is not insurance or audit proof by itself.

Privacy/selective disclosure profile

Object layer	Privacy/selective disclosure profile
Field or evidence focus	Redaction, minimization, retention, privileged data, sensitive data, and disclosure-control posture.
Body relationship	Chapters 13, 14, 18; T-13-01, T-14-01, T-18-01.
Reviewability function	Allows evidence review to be framed without assuming unlimited disclosure or over-retention.
Boundary note	Not privacy legal advice, disclosure advice, or regulator-approved method.

Dependency/substitution context

Object layer	Dependency/substitution context
Field or evidence focus	Third-party service, model replacement, API dependency, fallback path, or substitution record.
Body relationship	Chapters 12, 17, 23; T-12-01, T-17-01, T-23-01.
Reviewability function	Helps explain whether a change in the operating environment changed the reviewed risk object.
Boundary note	Does not establish aggregation pricing, reinsurance treatment, or policy response.

Accepted outcome

Object layer	Accepted outcome
Field or evidence focus	Expected output, completion criterion, authorized result, or business acceptance condition.
Body relationship	Chapters 4, 5, 9, 24; T-04-01, T-05-01, T-09-01, T-24-01.
Reviewability function	Provides a reference point for deciding whether an incident involved deviation, incomplete work, or disputed completion.
Boundary note	Does not certify quality, compliance, or claim validity.

Exception/dispute/remediation state

Object layer	Exception/dispute/remediation state
Field or evidence focus	Incident status, exception type, dispute posture, remediation step, closure evidence, or residual gap.
Body relationship	Chapters 19, 20, 22, 23; T-19-01, T-20-01, T-22-01, T-23-01.
Reviewability function	Connects post-loss evidence to the original work object and later review.
Boundary note	Does not settle a dispute, prove no residual liability, or decide coverage.

Renewal/change feedback

Object layer	Renewal/change feedback
Field or evidence focus	Evidence of control changes, reauthorization, substitution, recurring gaps, or revised operating assumptions.
Body relationship	Chapters 17 and 23; T-17-01 and T-23-01.
Reviewability function	Makes claims-to-renewal learning visible without converting it into pricing or acceptance.
Boundary note	Not actuarial pricing guidance, premium recommendation, or renewal instruction.

Minimal Object Card

An appendix-level object card may later summarize the object model in a shorter form:

Card field	Prompt for internal use	Boundary note
Work object	What bounded agentic work unit is being reviewed?	Not a coverage trigger.
Responsible subject	Which legal subject or business owner is connected to the work unit?	Not a liability determination.
Authority	What was the agent allowed and not allowed to do?	Not underwriting guidance.
Evidence	Which evidence pointers reconstruct intent, action, review, and consequence?	Not proof of claim approval.
Consequence	What external action, loss, dispute, or remediation issue made the work object loss-relevant?	Not causation proof.
Boundary	Which privacy, dependency, substitution, or dispute limits affect review?	Not legal or insurance advice.
Feedback	What should be visible for later renewal/change review?	Not pricing or renewal guidance.

Boundary Note

Appendix A is an authored analytical model only. It is not a policy definition, coverage trigger, underwriting standard, certification, proof of insurability, insurer endorsement, regulator approval, or insurer-adopted method.

Appendix B - Non-Scoring Agentic Insurability Reasoning Model

Purpose

Appendix B provides a compact reference version of the non-scoring Agentic Insurability Reasoning Model from Chapter 25. It supports Chapters 8, 10, 13, 14, 18, 21, 23, and 24 by describing reviewability states without ranking, scoring, certifying, or predicting an insurance outcome.

This appendix references body tables T-08-01, T-10-01, T-13-01, T-14-01, T-18-01, T-21-01, T-23-01, T-24-01, and T-25-01.

Non-Scoring Reasoning States

Reasoning state	What exists	What remains missing	Reviewability implication	What it does not prove
Opaque / not reconstructable	AI use is known or suspected, but the work unit, authority, action, evidence, and consequence are not reconstructable.	Bounded work object, role map, tool-action record, evidence chain, exception state, and privacy-aware disclosure plan.	Risk discussion remains vague because the reviewer cannot locate what acted, who was responsible, or what evidence exists.	Does not prove uninsurability, policy exclusion, legal liability, or claim denial.
Log-visible but responsibility-poor	Logs, traces, prompts, tickets, or system records exist.	Responsibility continuity, authority boundary, accepted outcome, human role, and consequence mapping.	Technical visibility improves but still does not explain who owned the work or whether action stayed inside authority.	Logs alone do not prove auditability, insurability, liability, coverage, or claim approval.
Evidence-linked but boundary-incomplete	Evidence pointers can connect intent, action, and outcome for selected work units.	Complete authority boundaries, privacy treatment, dependency context, substitution record, and dispute/remediation state.	A reviewer can begin reconstruction, but the risk object may still be too boundary-poor for serious insurance-facing discussion.	Does not create a coverage path, underwriting rule, or readiness certification.
Reviewable work-unit architecture	Work units, responsible subjects, authority, evidence, privacy controls, and accepted outcomes are organized.	Insurance-facing aggregation, renewal/change feedback, post-loss evidence handling, and external source caveats.	The enterprise can discuss agentic risk as bounded work rather than generic AI use.	Does not guarantee insurability, premium effect, insurer acceptance, or claim outcome.
Underwriting-facing evidence architecture	Evidence requests, exposure inventory, authority maps, privacy/redaction profiles, dependency maps, and change registers can be presented for review.	Actual insurer appetite, policy terms, exclusions, sublimits, endorsements, jurisdictional facts, and final underwriting judgment.	Risk is more legible for a reviewer without turning the paper into an underwriting standard.	Does not bind an insurer, create underwriting guidance, or imply coverage-ready status.
Post-loss reconstructable and renewal-updatable	Claim reconstruction, dispute gaps, remediation closure, and claims-to-renewal feedback can be tied back to the work object.	Legal determination, coverage determination, claim handling judgment, settlement posture, actuarial pricing, and future market response.	The evidence environment can support post-loss analysis and future change review.	Does not approve claims, prove liability, determine coverage, or set renewal terms.

Use Notes

- The states are descriptive, not hierarchical grades.
- Movement between states is analytical, not a pass/fail path.
- A higher reviewability state means more reconstructable evidence, not guaranteed insurability.

- Source refs `[1]` through `[62]` remain preserved in the body; this appendix does not add new external claims.

Boundary Note

Appendix B is non-scoring. The reasoning states describe reviewability, not insurance outcomes. They are not a score, grade, benchmark, rating model, maturity score, pass/fail certification, readiness certification, insurer appetite statement, underwriting rule, coverage path, or insurer-adopted method.

Appendix C - Underwriting-Facing Evidence Request Structure

Purpose

Appendix C provides an optional analytical structure for evidence a reviewer might ask to see when discussing agentic AI risk. It supports Chapters 6, 14, 15, 16, 17, and 18 and references T-06-01, T-14-01, T-15-01, T-16-01, T-17-01, and T-18-01.

This appendix is deliberately framed as an optional request structure, not as a checklist or condition for insurance.

Optional Evidence Request Categories

Category	Example request	Purpose	Sensitive boundary	Non-claim note
Work-unit inventory	Identify the recurring agentic work units that can create external consequences.	Separates general AI use from loss-relevant work.	Avoid collecting more operational detail than review requires.	Not a procurement requirement or underwriting checklist.
Authority and role map	Show the permitted action range, escalation rule, human role, and agent role for selected work units.	Connects delegated authority to responsibility continuity.	Role evidence may include privileged, employment, or sensitive governance context.	Not a legal liability determination or insurer acceptance condition.
High-impact work-unit list	Identify work units that affect payments, customer communications, regulated decisions, production systems, data movement, or external commitments.	Focuses review on work likely to matter for loss, dispute, or dependency analysis.	Avoid implying that unlisted work units are risk-free.	Not a coverage-ready or underwriting-ready label.
Tool-action sample	Provide selected examples of tool calls, system changes, messages, transactions, or external actions.	Shows whether the agent could create consequences outside model output.	Samples may contain customer, employee, security, or confidential business data.	Not proof of causation, claim approval, or claim payment support.
Evidence chain sample	Provide pointers to intent, approval, tool action, exception, human review, remediation, and closure evidence.	Tests whether the work unit can be reconstructed.	Use pointers and redaction rather than unnecessary bulk disclosure.	Evidence improves reviewability only.
Privacy/redaction profile	Explain how sensitive evidence is minimized, redacted, retained, or selectively disclosed.	Keeps reviewability separate from over-collection.	Privacy, privilege, retention, and disclosure issues remain external.	Not privacy legal advice or regulator-approved treatment.
Incident/near-miss history	Summarize relevant incidents, near misses, exceptions, escalations, and disputed outcomes.	Helps distinguish theoretical AI risk from observed operational patterns.	Incident records may require legal, privacy, HR, customer, or security review.	Not a claim demand or admission.
Remediation closure examples	Show how selected incidents were remediated, reauthorized, or closed with residual gaps noted.	Connects post-loss learning to future review.	Closure language should avoid overclaiming no residual risk or no liability.	Not settlement guidance or proof of no residual liability.
Substitution/change register	Identify model, tool, vendor, API, runtime, prompt, workflow, or permission changes that affect reviewed work units.	Makes renewal/change review possible.	Vendor and security details may require minimization.	Not renewal guidance, pricing guidance, or binding condition.
Dependency map	Map cloud, API, model, data, vendor, or runtime dependencies for selected work units.	Supports aggregation, concentration, and operational dependency review.	Dependency maps may expose security or commercial sensitivities.	Not reinsurance guidance or vendor ranking.

Category	Example request	Purpose	Sensitive boundary	Non-claim note
Missing evidence register	Identify known gaps, unavailable records, vendor-held logs, expired retention, or redacted evidence without pointers.	Makes uncertainty explicit rather than hiding it.	Some gaps may reflect legitimate minimization or privilege decisions.	Not claim denial support or underwriting rejection guidance.

Boundary Note

Appendix C is an optional analytical request structure only. It is not an underwriting checklist, claim demand, procurement requirement, certification checklist, premium-credit path, binding condition, insurer acceptance statement, mandatory coverage requirement, or insurer-adopted method.

Appendix D - Claim Reconstruction and Evidence Gap Register

Purpose

Appendix D provides a reference structure for organizing post-loss reconstruction and evidence gaps. It supports Chapters 7, 10, 19, 20, 22, and 23 and references T-07-01, T-10-01, T-19-01, T-20-01, T-22-01, and T-23-01.

The register is designed to make evidence questions visible after an agentic AI incident. It does not decide legal liability, coverage, settlement, claim payment, or claim approval.

Claim Reconstruction Fields

Work unit ID	
Field	Work unit ID
What to capture	Identifier for the bounded agentic work object involved in the incident.
Why it matters	Prevents the claim discussion from collapsing into generic AI use.
Boundary note	Not a policy definition or coverage trigger.

Initiating intent	
Field	Initiating intent
What to capture	User instruction, system objective, workflow trigger, or business purpose.
Why it matters	Helps reconstruct why the agentic work started.
Boundary note	Does not prove authorization or proper use.

Authority boundary	
Field	Authority boundary
What to capture	Permission, threshold, escalation rule, blocked action, or delegated authority scope.
Why it matters	Shows what the agent was allowed to do.
Boundary note	Not legal liability proof or underwriting guidance.

Agent role

Field	Agent role
What to capture	Assigned agent function, autonomy level, and tool access.
Why it matters	Connects the agent's role to later action.
Boundary note	Does not prove insurability or claim outcome.

Human role

Field	Human role
What to capture	Initiator, reviewer, approver, override actor, escalation recipient, or post-loss reviewer.
Why it matters	Preserves responsibility continuity.
Boundary note	Does not determine negligence or liability.

Tool action

Field	Tool action
What to capture	Tool call, message, transaction, system change, data access, or external command.
Why it matters	Connects the agentic work to possible consequence.
Boundary note	Does not prove causation or damages by itself.

External consequence

Field	External consequence
What to capture	Customer, system, financial, operational, regulatory, security, or third-party effect.
Why it matters	Identifies why the work unit became loss-relevant.
Boundary note	Does not decide coverage or claim value.

Affected data

Field	Affected data
What to capture	Data, records, credentials, customer information, intellectual property, or business process affected.
Why it matters	Supports privacy, evidence, and impact review.
Boundary note	Not privacy legal advice.

Evidence chain pointer

Field	Evidence chain pointer
What to capture	Location of relevant approvals, logs, prompts, tool records, reviews, exceptions, and remediation records.
Why it matters	Supports reconstruction without requiring all evidence in one place.
Boundary note	Evidence pointers are not claim approval guidance.

Exception record

Field	Exception record
What to capture	Escalation, anomaly, override, control break, or disputed action.
Why it matters	Shows where the normal work-unit path changed.
Boundary note	Does not prove fault.

Remediation action

Field	Remediation action
What to capture	Containment, correction, reauthorization, rollback, notice, user communication, or control change.
Why it matters	Connects post-loss response to the original work object.
Boundary note	Not settlement guidance or proof of no residual liability.

Closure state

Field	Closure state
What to capture	Open, remediated, reauthorized, monitored, disputed, incomplete, or unresolved.
Why it matters	Makes residual uncertainty explicit.
Boundary note	Does not close legal, claim, or coverage questions.

Missing evidence

Field	Missing evidence
What to capture	Unavailable, expired, vendor-held, privileged, redacted, or not collected evidence.
Why it matters	Prevents silent gaps from being mistaken for proof.
Boundary note	Missing evidence is not automatic claim denial support.

Privacy/redaction treatment

Field	Privacy/redaction treatment
What to capture	Redaction basis, minimization method, disclosure pointer, or protected category.
Why it matters	Balances reviewability with sensitive evidence handling.
Boundary note	Not privacy legal advice or disclosure instruction.

Evidence Gap Register

Evidence gap	Typical issue	Reviewability effect	Boundary note
Missing authority record	No clear record of what the agent was allowed to do.	Weakens responsibility and boundary reconstruction.	Does not prove liability or coverage position.
Missing approval context	Human approval exists but lacks purpose, timing, or scope.	Makes it difficult to connect human review to agent action.	Not claim approval guidance.
Missing tool-action record	External action is known, but tool-call detail is absent.	Breaks the link between agentic work and consequence.	Does not itself decide causation.
Vendor-held logs	Relevant records are controlled by a provider or platform.	Creates dependency and disclosure friction.	Not vendor ranking or procurement advice.
Expired retention	Records existed but were overwritten or aged out.	Creates a reconstruction gap.	Not retention legal advice.
Redacted evidence without pointer	Information is withheld or redacted but no substitute pointer exists.	Preserves privacy but may weaken reviewability.	Not a privacy compliance conclusion.
Missing version/substitution record	Model, prompt, tool, API, or workflow changed without traceable record.	Makes recurrence and renewal review harder.	Not renewal or pricing guidance.
Unclear remediation closure	Response occurred but closure state or residual gap is unclear.	Leaves post-loss responsibility and future review unresolved.	Not settlement or no-liability proof.

Boundary Note

Appendix D organizes reconstruction evidence. It does not provide claim approval guidance, legal liability proof, settlement guidance, coverage determination, legal causation determination, claim payment support, insurer endorsement, or insurer-adopted method.

Appendix E - Coverage Boundary Question Map

Purpose

Appendix E maps recurring coverage-boundary questions that agentic AI incidents can raise. It supports Chapters 3, 7, 11, 19, 20, and 21 and references T-07-01, T-11-01, T-19-01, T-20-01, and T-21-01.

The purpose is to preserve questions, not answer them. Policy wording, facts, jurisdiction, notice, exclusions, limits, sublimits, deductibles, causation, loss category, and claim handling remain external to this paper.

Coverage-Boundary Question Categories

Cyber vs authorized operational misuse

Category	Cyber vs authorized operational misuse
Agentic fact pattern	An agent uses valid credentials or permitted tools in a harmful but not obviously unauthorized way.
Evidence needed	Authority boundary, credential context, tool-action record, human role, external consequence.
Why ambiguous	The event may look operational, cyber, insider-like, or control-failure-like depending on policy wording and facts.
External decision note	Coverage analysis belongs to policy, facts, claim handling, and applicable law.

Tech E&O vs professional liability

Category	Tech E&O vs professional liability
Agentic fact pattern	Agentic output or workflow action harms a client, customer, or third party in a service context.
Evidence needed	Work unit, service obligation, accepted outcome, review record, human role, customer effect.
Why ambiguous	The boundary between technology service failure and professional service error may be fact-sensitive.
External decision note	This map does not interpret policy forms or assign coverage.

Product vs service

Category	Product vs service
Agentic fact pattern	Agentic capability is embedded in a product, platform, API, or managed service.
Evidence needed	Product/service description, deployment context, user control, vendor/runtime context, consequence.
Why ambiguous	The loss may be framed as product defect, service failure, operational negligence, or contractual performance issue.
External decision note	Product and service characterization remains external.

D&O/governance vs operational failure

Category	D&O/governance vs operational failure
Agentic fact pattern	Board or executive governance decisions are challenged after an agentic AI incident.
Evidence needed	Governance record, authority delegation, risk reporting, oversight evidence, incident record.
Why ambiguous	Governance allegations may overlap with operational control failures without being the same risk object.
External decision note	This appendix does not determine director/officer liability or policy response.

Crime/social engineering vs cyber

Category	Crime/social engineering vs cyber
Agentic fact pattern	Agentic workflow is used in deceptive payment, identity, invoice, or communication events.
Evidence needed	Instruction path, human approval, payment flow, tool-action record, fraud indicators, exception record.
Why ambiguous	Loss may involve deception, authorized transfer, unauthorized access, or social engineering depending on policy terms.
External decision note	Claim handling and policy wording remain external.

Media/IP vs generated content workflow

Category	Media/IP vs generated content workflow
Agentic fact pattern	Generated content or agent-assisted publication triggers IP, media, defamation, or content dispute.
Evidence needed	Prompt/objective record, source material, human review, publication path, takedown/remediation record.
Why ambiguous	The event may involve content creation, publication, vendor tool use, human review, and rights questions.
External decision note	This appendix does not provide IP or media-liability advice.

Business interruption/property vs cloud/API/service interruption

Category	Business interruption/property vs cloud/API/service interruption
Agentic fact pattern	Agentic workflow fails because a cloud, model, API, or service dependency is unavailable or degraded.
Evidence needed	Dependency map, outage record, affected work unit, fallback path, business process impact.
Why ambiguous	The boundary between technology dependency, business interruption, cyber event, and operational outage can be policy-specific.
External decision note	Loss category and policy response remain external.

Privacy/regulatory investigation vs operational event

Category	Privacy/regulatory investigation vs operational event
Agentic fact pattern	Agentic system accesses, discloses, transforms, or retains sensitive data in a disputed way.
Evidence needed	Affected data, authority, tool-action record, privacy/redaction profile, notice/investigation record.
Why ambiguous	The event may be framed as privacy incident, regulatory issue, cyber event, professional failure, or operational control problem.
External decision note	Privacy, regulatory, and coverage conclusions remain external.

Boundary Note

Appendix E frames coverage-boundary questions only. It is not a coverage opinion, legal advice, policy interpretation, exclusion application, sublimit application, claim handling instruction, or insurer-adopted method.

Appendix F - Source and Claim Boundary Notes

Purpose

Appendix F consolidates source-use and claim-boundary notes for the appendices and the revised body. It supports Front Matter and Chapters 1-28.

The revised body preserves source refs `[1]` through `[62]`. This appendix does not create new source IDs, add unsupported external factual claims, or use rejected v0.2 as source truth.

Source Group Notes

Source group	How it is used	Caveat to preserve
Market/context sources	Used to ground the observation that AI risk transfer is already being discussed through market products, exclusions, sublimits, warranties, governance expectations, and silent exposure concerns.	Market/context sources are not proof of market-wide acceptance, insurer endorsement, policy availability, or universal coverage treatment.
Product examples	Used to illustrate that AI-related cover, warranty, or risk-transfer signals exist in the market.	Product examples do not prove agentic lifecycle coverage, policy response, claim outcome, or adoption of the paper's models.
Governance and incident-response sources	Used to support evidence architecture, lifecycle controls, remediation, and reconstruction concepts.	Governance and incident-response sources do not become insurance advice, underwriting standards, or claim handling guidance.
Cyber/cloud/reinsurance analogies	Used to reason by analogy about aggregation, concentration, dependency, incident evidence, and reviewability.	Analogies remain analogies unless directly tied to AI or agentic AI by the cited source.
Regulatory/legal context	Used to understand the external environment in which AI governance, privacy, liability, and insurance questions may arise.	These sources do not create legal advice, coverage opinion, legal authority, or regulator approval of the paper.
Compliance and auditability foundations	Used as internal analytical foundations for lifecycle governance, Missing Regulatory Objects, auditability, Audit Evidence Chain, Agentic Audit Object, and related concepts.	The compliance and auditability white papers are analytical foundations only; they do not prove insurability, coverage, or claim approval.
Author synthesis	Used when the paper combines source-supported context with authored analytical models such as AIO/AIRM-style constructs.	Author synthesis must remain labeled as analytical construction, not external standard, score, certification, insurer-adopted method, or market fact.

Rejected v0.2 Exclusion

The rejected v0.2 candidate remains archived for traceability only. It is not current source truth and is not a citation source for this appendix set.

Unresolved Source Gaps

The following gaps remain strengthening opportunities, not blockers for the appendix draft:

- exact exclusion/sublimit primary wording;
- insurer claim documentation;
- AI underwriting questionnaires.

If a later wave adds insurer-specific, claims-handling, regulatory, reinsurance, broker, or vendor-specific factual claims beyond the current body, those claims should be source-backed or explicitly marked as author inference.

Boundary Note

Appendix F supports traceability and caveat discipline. It does not convert sources into legal, insurance, underwriting, actuarial, claim, regulator, or market conclusions. It does not imply legal authority, insurer endorsement, regulator approval, market-wide acceptance, policy wording proof, or source support beyond actual source scope.

Appendix G - Final Non-Claim Language Register

Purpose

Appendix G provides a consolidated non-claim language register for future QA and drafting control. It supports Front Matter and Chapters 6, 7, 16, 18, 19, 21, 24, 25, 27, and 28, and it references T-27-01.

This appendix is a publication guardrail only. It is not legal advice.

Restricted Claim Register

Restricted claim	Why it matters	Safer language	Where to preserve
Legal advice	Legal conclusions depend on facts, jurisdiction, counsel, and applicable law.	legal context; legal question; counsel-reviewed issue; external legal determination	Front Matter, Chapters 21 and 27, Appendix E.
Insurance advice	Insurance decisions depend on broker, insurer, policy wording, facts, and appetite.	insurance-facing reasoning; risk-transfer discussion; reviewer question	Front Matter, Chapters 1, 6, 18, 27.
Underwriting guidance	Underwriting decisions remain external to the paper.	underwriting-facing evidence architecture; reviewer-facing evidence; risk review question	Chapters 6, 14, 18, Appendix C.
Coverage opinion	Coverage depends on policy wording, facts, law, claim handling, exclusions, limits, sublimits, and notice.	coverage-boundary question; coverage ambiguity; external coverage analysis	Chapters 11 and 21, Appendix E.
Actuarial pricing guidance	Pricing requires actuarial data, assumptions, insurer practice, and regulatory context.	non-pricing exposure variable; pricing-relevant question; exposure signal	Chapter 16, Appendix C.
Premium recommendation	Premiums are insurer and market decisions.	premium-sensitive factor; pricing-relevant exposure; non-pricing discussion	Chapter 16.
Claim approval guidance	Claim handling is external and fact-specific.	claim reconstruction; evidence gap; post-loss reviewability	Chapters 7, 19, 20, Appendix D.
Legal liability determination	Liability depends on facts, law, contracts, duties, causation, and adjudication or settlement.	responsibility continuity; liability question; legal determination remains external	Chapters 5, 19, 20, 22.
Certification	Certification implies an external attestation not created by the paper.	internal reference; authored analytical model; reviewability construct	Chapters 24 and 25, Appendices A and B.
Proof of insurability	Insurability depends on insurer judgment, policy terms, market appetite, and facts.	improves reviewability; makes risk more legible; supports discussion	Chapters 24, 25, 28.
Insurer endorsement	The paper's models are not adopted or endorsed by insurers.	insurer-facing; reviewer-facing; market context	Front Matter, Appendices A, F.
Regulator-approved method	No regulator approval is claimed.	regulatory context; governance context; external regulatory question	Chapters 13 and 27, Appendix F.
Procurement recommendation	Vendor and procurement choices are outside scope.	dependency context; vendor/runtime context; source caveat	Chapters 12, 17, Appendices A, F.
Vendor ranking	The paper does not compare vendors or tools.	vendor/runtime context; dependency visibility	Chapters 12, 17, Appendix A.
Score	Scores imply numeric evaluation or ranking.	non-scoring reasoning state; reviewability state	Chapter 25, Appendix B.
Standard	Standards imply normative external requirements.	reference model; analytical construct; optional structure	Chapters 18, 24, 25, Appendices A-C.

Restricted claim	Why it matters	Safer language	Where to preserve
Readiness certification	Readiness language can imply external acceptance.	reviewability; internal preparedness for discussion; evidence architecture	Chapter 25, Appendix B.
Insurer-adopted method	Adoption requires evidence outside the paper.	authored analytical model; not insurer-adopted	Chapters 24, 25, Appendices A and B.
Public release	Release requires later authorization.	public research edition available; no external approval status	Front Matter, README/governance files.
Public edition	Candidate status requires later authorization.	internal appendix draft; internal revision package	README/governance files.
Certification or external approval	Certification or external approval is not claimed.	internal draft; subject to QA	README/governance files.
Release-ready status	Public HTML/PDF artifacts are available for research-edition access; external approval status requires separate owner decision.	not certification; social announcement not executed	README/governance files.

Boundary Note

Appendix G is a drafting and publication guardrail only. It is not legal advice, insurance advice, underwriting guidance, coverage opinion, claim handling guidance, regulator-approved language, certification language, or release authorization.

Appendix H - Table Inventory and Layout Risk Register

Purpose

Appendix H inventories body tables and flags future layout risk before any artifact planning. It supports Chapters 2 and 4-27 and references T-02-01 and T-04-01 through T-27-01.

This appendix is a source-level layout-risk register. The public research edition package exposes HTML, PDF, manifest, and checksum artifacts, while DOCX distribution and external approval status remain unauthorized.

Table Inventory

Table ID	Chapter	Title	Appendix posture	Wide-table risk flag	Future artifact note
T-02-01	2	AI Insurance Split-Market Signal Matrix	Appendix-summary candidate; source caveat support.	Yes	May need responsive/abbreviated treatment later.
T-04-01	4	Insurance Object Shift	Body-only; Appendix A reference support.	No	Keep body version unless later layout QA requires change.
T-05-01	5	Responsibility Continuity Map	Body-only; Appendix A reference support.	No	Keep body version.
T-06-01	6	Underwriting Evidence Request Model	Body-only; Appendix C reference support.	No	Keep body version; avoid checklist framing.
T-07-01	7	Claim Evidence Pack Components	Body-only; Appendix D reference support.	No	Keep body version; preserve claim-approval boundary.
T-08-01	8	Hard-to-Insure Agentic Risk Patterns	Body-only; Appendix B reference support.	No	Keep body version; avoid uninsurability overclaim.
T-09-01	9	MRO-to-Insurability Translation Map	Appendix-summary candidate; Appendix A/B reference support.	Yes	May need shortened body version or appendix parity QA later.
T-10-01	10	Auditability-to-Claim-Reconstruction Crosswalk	Body-only; Appendix D reference support.	No	Preserve auditability/insurability distinction.
T-11-01	11	Insurance Line Ambiguity Map	Body-only; Appendix E reference support.	No	Preserve coverage-opinion boundary.
T-12-01	12	Agentic Aggregation Risk Map	Body-only; Appendix F source caveat support.	No	Keep analogy caveats visible.
T-13-01	13	Insurance Evidence vs Privacy Control Map	Body-only; Appendix C/F reference support.	No	Preserve privacy/selective disclosure boundary.
T-14-01	14	Underwriting Evidence Architecture Components	Appendix-summary candidate; Appendix C/H reference support.	Yes	May need responsive or appendix-detailed treatment later.
T-15-01	15	Agentic Exposure Inventory Template	Body-only; Appendix C reference support.	No	Preserve non-certification posture.
T-16-01	16	Non-Pricing Exposure Variables	Body-only; Appendix C/G boundary support.	No	Preserve no-pricing/no-premium boundary.
T-17-01	17	Renewal and Change Evidence Register	Body-only; Appendix C reference support.	No	Preserve no-renewal-instruction boundary.
T-18-01	18	Optional Reviewer Evidence Request Structure	Appendix-summary candidate; Appendix C/H reference support.	Yes	May need body/appendix split later; avoid checklist framing.

Table ID	Chapter	Title	Appendix posture	Wide-table risk flag	Future artifact note
T-19-01	19	Agentic Claim Reconstruction Map	Body-only; Appendix D reference support.	No	Preserve no-claim-approval boundary.
T-20-01	20	Dispute and Evidence Gap Register	Body-only; Appendix D reference support.	No	Keep dispute/evidence-gap framing.
T-21-01	21	Coverage Boundary Question Map	Appendix-summary candidate; Appendix E/H reference support.	Yes	May need careful PDF/HTML treatment later; no coverage opinion.
T-22-01	22	Post-Loss Remediation Evidence Map	Body-only; Appendix D reference support.	No	Preserve no-settlement/no-liability-proof boundary.
T-23-01	23	Claims-to-Renewal Feedback Loop	Body-only; Appendix A-D reference support.	No	Preserve no-pricing/no-renewal-outcome boundary.
T-24-01	24	Agentic Insurability Object Model	Body-only; Appendix A reference support.	No	Preserve authored analytical model framing.
T-25-01	25	Non-Scoring Insurability Reasoning Model	Body-only; Appendix B reference support.	No	Preserve non-scoring framing.
T-26-01	26	Audience Takeaway Matrix	Body-only; Appendix F/G support.	No	Keep audience implications non-advisory.
T-27-01	27	Final Non-Claim Boundary Register	Appendix-summary candidate; Appendix G/H reference support.	Yes	May need future layout treatment; do not weaken caveats.

Wide-Table Risk List

The tables currently carried forward as wide-table risks are:

- T-02-01;
- T-09-01;
- T-14-01;
- T-18-01;
- T-21-01;
- T-27-01.

Future Artifact Note

Future artifact planning may consider responsive HTML wrappers, shorter body tables with appendix detail, PDF landscape treatment, table footnote compression, or body-to-appendix parity QA. None of those treatments is implemented in this appendix draft.

Boundary Note

Appendix H is a layout-risk register only. Public HTML, PDF, manifest, and checksum access remains bounded by public research edition status; Appendix H does not create final PDF readiness, visual acceptance, DOCX distribution, certification, external approval, or social announcement authorization.

Deferred Appendix Note

Appendix I - Assembly and Revision Change Log is not included in this integrated draft. It remains optional and deferred because assembly/revision traceability already exists in governance packages and may be too process-heavy for the public white paper unless later authorized by the owner.

This deferred note does not create Appendix I content, certification, external approval, social announcement, DOCX distribution, or source Markdown publication.

Boundary: This public research edition is available with bounded HTML/PDF, manifest, and checksum access, but it does not authorize social announcement, certification, external approval, DOCX distribution, insurer acceptance, coverage readiness, underwriting readiness, claim readiness, certification, score, standard, or regulator approval.