Corrections table — first public release, text version 1 (25 September 2026) to version 2 (29 September 2026)

What this is. Every change made to the text, in order, in three columns a reader can check against the two versions: the sentence as it was, the sentence as it is now, and the class of the change. Stage 1 lists every change between the text posted on appendonly.ai on 25 September 2026 and the first corrected text (version 2) of 29 September. Later stages list the changes made within the release text before deposit, each dated.

How to read the classes. WRONG = the record contradicted the sentence · OVERCLAIM = the record supports something weaker · CITATION = an outside number or quote was misstated, mis-sourced or unverifiable · OMISSION = something the record holds that the text left out · DECISION = a choice by the owner, not a factual correction · ADDED / WORDING = new or clearer text without a change of claim.

What is omitted. The internal table has a sixth column naming the internal readers, files and the rows of the record each change traces to; it is omitted here. One cell is withheld and marked in place: it carried internal pricing. The retracted sentences in the 'old' column are printed as they stood; the point of the table is to show what was corrected and how.

Counts. The count is high because the text was checked sentence by sentence against its sources and the record, and every change is listed, not only the ones that changed a conclusion; most rows tighten one sentence by one qualifier. 130 rows: 39 WRONG · 36 OVERCLAIM · 25 CITATION · 17 WORDING · 9 ADDED · 3 DECISION · 1 OMISSION.

#	Where	Class	v1 (old)	v2 (new)
1	Title line	WORDING	**AppendOnly by 5R Industries** · Adrian Outlaw · first public release, 28 September 2026	**AppendOnly by 5R Industries** · Adrian Outlaw · first public release (text version 2), September 2026
2	Abstract	CITATION	more than 40 percent of agent projects are forecast to be cancelled by 2027,	more than 40 percent of agentic AI projects are forecast to be cancelled by 2027,
3	Abstract	OVERCLAIM	This report describes such a record as deployed since March 2026 on a small engineering company, run by one person with an AI chief of staff.	This report describes such a record, running in its first form on a small engineering company since March 2026 and in its present form since 9 September 2026, when signing, public timestamps, and action capture began; most of the numbers below come from the period since July, when the independent verifier started. The company is run by one person with an AI chief of staff.
4	Abstract	WRONG	and a claim-labelling precision measured so far on six human-graded rows.	and a claim-labelling precision checked so far on five decided rows, all of them rows the verifier had flagged.
5	About this release	ADDED	the argument (Part A) and the measured results (sections 22 to 24 and 26).	the argument (Part A), the measured results (sections 22 to 24), a short related-work note (section 25), and the limitations (section 26).
6	About this release	OVERCLAIM	Every number below was measured on 17 September 2026 from the live system and is reproducible from files in the record; the filing that covers the mechanisms was made on 24 September 2026. This report was written on the record it describes: the agent that helped write it is the agent verified in it, every draft is a signed commit in that record, and the human signs the release.	Every number of ours below was measured on 17 September 2026 from the live system and is reproducible by us from files in the record; what a stranger can check today is the integrity of the record (its signatures and public timestamps), not yet its contents. The filing that covers the mechanisms was made on 24 September 2026. This report was written on the record it describes: the agent that helped write it is the agent verified in it, and every draft is a signed commit in that record. This text is version 2; version 1 (SHA-256 f5900deb…) was posted on appendonly.ai on 25 September 2026 and is superseded by this one, with every change listed in a corrections table in the record. This release is licensed CC BY 4.0.
7	§1 opening	WRONG	The numbers below are other people's, gathered on 17 September 2026 and cited in full in the companion research note.	The numbers below are other people's, gathered on 17 September 2026 and checked against their sources on 26 September 2026; every source is listed with its link in the Sources section at the end of this report.
8	§1 spend	CITATION	Per-enterprise AI spend is forecast to rise 65 percent in 2026, and the share paid from innovation budgets fell from 25 to 7 percent: AI is now a core operating line, not an experiment (a16z).	Enterprises expect their spend on large language models to rise about 65 percent in 2026 (a16z, January 2026), and by mid-2025 the share of that spend paid from innovation budgets had already fallen from a quarter to 7 percent: AI is now a core operating line, not an experiment (a16z, June 2025).
9	§1 Gartner	CITATION	Gartner forecasts that more than 40 percent of agent projects will be cancelled by the end of 2027	Gartner forecasts that more than 40 percent of agentic AI projects will be cancelled by the end of 2027
10	§1 Gartner quote	CITATION	"due to governance gaps identified only after production incidents."	"due to governance gaps identified only after production incidents occur."
11	§1 KPMG	CITATION	Only 26 percent of companies deploying agents have real-time visibility into what the AI is costing to operate (KPMG, mid-2026).	Only 26 percent of large U.S. companies report full, real-time visibility into what their AI systems cost to operate (KPMG, Q2 2026).

#	Where	Class	v1 (old)	v2 (new)
12	§1 BambooHR	CITATION	Workers spend about 47 days a year on AI and about 20 of them fixing its output (BambooHR, September 2026).	Workers spend about 47 days a year on AI and about 20 of them troubleshooting errors and re-prompting (BambooHR, September 2026).
13	§1 RegLab	CITATION	Leading legal research tools fabricate 17 to 33 percent of the time (Stanford RegLab, 2025).	Leading legal research tools gave incorrect or mis-sourced answers 17 to 33 percent of the time in Stanford RegLab's 2024 tests (published 2025).
14	§1 Deloitte	CITATION	A consulting firm refunded part of a $440,000 government report after its citations were found to be fabricated (October 2025).	A consulting firm refunded part of an A$440,000 (about US$290,000) government report after its citations were found to be fabricated (October 2025).
15	§1 sanctions	CITATION	Courts have sanctioned lawyers in 1,668 catalogued cases for AI-fabricated citations, at about eight new cases a day (July 2026).	Courts have addressed AI-hallucinated citations or content in 2,079 catalogued decisions, 827 of them involving lawyers, at about five new decisions a day since July (Charlotin database, 25 September 2026).
16	§1 IBM	CITATION	Breaches involving unsanctioned AI cost $670,000 more than the average and appeared in 43 percent of incidents in 2026; the average breach in the United States now costs $11.5 million and takes 241 days to identify and contain (IBM, 2025 and 2026).	In 2026, unsanctioned AI was present in 43 percent of security incidents, and breaches involving it cost $5.39 million on average against $4.99 million overall; the average breach in the United States now costs $11.5 million and takes about 247 days to identify and contain (IBM, 2026).
17	§1 Splunk	CITATION	An incident investigation runs 50 to 150 staff hours and about $58,000 in direct cost (Splunk, 2025).	An incident investigation runs 50 to 150 staff hours and about $58,000 in direct cost (Ponemon research, as cited by Splunk, 2025).
18	§1 post-mortems	OVERCLAIM	And the post-mortems keep hitting the same wall: they cannot answer "what did the agent touch," because the agent had no record of its own.	And the post-mortems keep finding the same gap: the records existed, but nothing checked them while the agent worked, and the agent's own harness could alter them.
19	§1 EU AI Act	CITATION	The European Union's AI Act requires high-risk systems to log events automatically over their lifetime and providers to keep the logs, with fines up to 7 percent of global turnover.	The European Union's AI Act requires high-risk systems to log events automatically over their lifetime and providers to keep those logs; breaches of the high-risk obligations carry fines up to 3 percent of global turnover, and for most high-risk systems those obligations now apply from December 2027.
20	§1 SEC	WRONG	United States financial regulators have said existing books-and-records and supervision rules apply unchanged to AI output, and the SEC's 2026 examination priorities ask whether "changes to system logic are documented."	FINRA has said existing books-and-records and supervision rules apply unchanged to generative-AI output, and the SEC's 2026 examination priorities say examiners will assess whether firms "have implemented adequate policies and procedures to monitor and/or supervise their use of AI technologies."
21	§1 FDA	CITATION	and data-integrity findings appear in most drug-manufacturing warning letters.	and data-integrity findings, once cited in about four of five drug warning letters, still appear in about one in seven (2025), and in 60 percent of those issued to sites in India.
22	§1 NAIC	CITATION	Twenty-four states have adopted the insurance regulators' model bulletin on AI record retention.	More than half of U.S. states have adopted the insurance regulators' model bulletin on insurers' use of AI, which expects a written AI program, board accountability, and documentation available on examination (NAIC, December 2025).
23	§1 House bill	CITATION	A bill introduced in the House in September 2026 would task NIST with standards for continuous verification of agent actions and tamper-proof action logs, mandatory for federal agencies and their contractors.	A bipartisan House bill introduced in September 2026, the Stop Rogue AI Act (H.R. 10362), would have NIST set standards for continuous monitoring and verification of AI agents and tamper-evident logs of their material actions, which federal contracts entered after enactment would then have to require.
24	§2 capture	OVERCLAIM	A record. Every action an AI agent takes inside your company is written down by the system at the moment it happens, below the level of language, so the agent cannot describe its work more favourably than the facts. Every sentence the agent later says about its work is compared to that record, and a claim that exceeds its evidence is marked as such, automatically, without the agent's cooperation.	A record. Every tool call an AI agent makes inside your company is written down by a hook the agent does not control, at the moment the call returns, as fingerprints of what went in and what came out, below the level of language, so the agent cannot describe its work more favourably than the facts. Every statement the agent then commits to the record about its work is compared to that evidence, and a claim that exceeds it is marked as such, automatically, without the agent's cooperation.

#	Where	Class	v1 (old)	v2 (new)
25	§2 integrity	WRONG	The record cannot be erased or edited after the fact without the change showing by the next morning: each entry carries the fingerprint of the one before it, every entry is signed, every morning the newest fingerprint is stamped into a public timestamp service, and a second machine that the agent cannot write to checks the whole thing every minute. All of it runs on your own hardware. Nothing leaves your building. Three small computers on a desk are enough.	The record cannot be quietly changed after the fact: each entry carries the fingerprint of the one before it, every commit since 9 September 2026 is signed, every morning since then the newest fingerprint has been stamped into a public timestamp service (one morning missed), and a separate machine that accepts only appends from the agent, never a rewrite, checks every minute that no committed history has changed and grades each new commit. A change to anything already stamped shows by the next morning. The record, the verifier, and the checks run on your own hardware; what leaves is what your agent already sends to the model vendor you chose, plus a 32-byte fingerprint to the timestamp service. Three small computers on a desk are enough.
26	§2 person floor	WRONG	The record is of agents and evidence. It is never used to evaluate an individual person. Every person sees every row the record holds about them, and every read of those rows is itself a row. That is a property of the system, written into the pilot terms and enforced in the product, not a policy promise.	The record is of agents and evidence. It is not used to evaluate employees; the one person it has graded is its operator, by his own choice. A person the record is about can read what it holds about them, and every read is itself a row, except where a legal duty of confidentiality named in the contract says otherwise, and then the withheld rows are recorded as withheld. We commit to this in the pilot terms; the read path is being built before the first pilot.
27	§3 table	OVERCLAIM	| Total model spend at list price, 55 days | $620.65 |	| Seat sessions' model spend at list price, 55 days (subagents and the scribe add roughly a sixth) | $620.65 |
28	§3 table	OVERCLAIM	| Heaviest build days (9 to 17 September) | $17 to $111 per day, average $46 |	| Build days (9 to 17 September) | $0 to $111 per day, average $46 |
29	§3 table	OVERCLAIM	| Context tokens served from cache | 98.0 percent (811 million cached, 247 thousand at full price) |	| Context tokens served from cache | 98.0 percent: 811 million read from cache, about 17 million written to cache, 247 thousand sent uncached |
30	§3 table	WRONG	| What the same context would have cost uncached | about $8,100, roughly forty times the $203 paid |	| What the same context would have cost uncached | about $8,100, roughly twenty times the about $415 paid for context (cache reads $203, cache writes about $207, uncached about $3) |
31	§3 table	WRONG	| One seat left running three days without rotation (12 September, I-39) | $168, peak context 997,000 tokens |	| One seat left running about two days without rotation (10 to 12 September, I-39) | $168, about $76 a day; peak context 997,000 tokens; ended stalled |
32	§3 table	WORDING	| The six rotated seats of the following day | $64 in total |	| The six rotated seats of 13 September | $64 for the day, no stall |
33	§3 table	OVERCLAIM	| Cold start: a fresh instance orienting from the record alone | $1.04 |	| Cold start: a fresh instance orienting from the record alone (answer not yet graded) | $1.04 |
34	§3 table	WORDING	| Handoff: warm relief to its first useful commit | 3 minutes, about 110,000 tokens |	| Handoff: a pre-warmed relief to its first useful commit | 3 minutes, about 110,000 tokens |
35	§3 record cost	WRONG	Under two dollars a day, under five percent of the seat's spend.	About forty cents a day in model spend for the scribe; the hooks' injected context is paid inside the seat figure above and was not separated out. Well under two dollars a day by our estimate.
36	§3 argument	WRONG	Three things in that table are the argument. Caching is what makes a long-running agent affordable at all, and caching only works when the context is append-only and stable, which is exactly the shape a tamper-evident record imposes. The one day a seat was allowed to run without rotation cost more than the whole following day of six rotated seats, and it was also the day the seat stalled: the record's discipline is the cost discipline. And a fresh instance can pick up the company from the record alone for a dollar, which is what makes the work survive a model change, a vendor change, or a person's absence.	Three things in that table are the argument. Caching is what makes a long-running agent affordable at all, and caching works when the context is append-only and stable, the same discipline a stable record encourages; a 98 percent cache share is typical of a well-run harness, not unique to ours. The one seat allowed to run without rotation cost more per day, and about 45 percent more per model call, than the following day's six rotated seats, and it ended in a stall: the record's discipline is also the cost discipline. And a fresh instance produced a next-actions answer from the record alone for a dollar, its grade still open, which is what should let the work survive a model change, a vendor change, or a person's absence.
37	§3 industry	CITATION	Measured studies find agents spending half their tokens on input and up to 62 percent re-reading context they have already seen; failed runs burn most of their tokens after the first sign of failure (2026).	A 2026 study of agentic software engineering found input tokens were 54 percent of consumption; in failed multi-agent runs that raised a warning, 58 percent of tokens were spent after the first warning (2026 preprints).

#	Where	Class	v1 (old)	v2 (new)
38	§4 Line 1	WRONG	The record labels every agent statement as verified, claimed, partial, or contradicted, at the moment it is made. On our record, across 557 verifications, half the agent's statements are `claimed`: asserted, not established by evidence. The other half needed no second look. Review effort goes where the label says, not everywhere.	The record labels every claim the agent commits to the record as verified, claimed, partial, or contradicted, within a minute of the commit. On our record, across 557 verifications, about a third (197) were verified against evidence; half (280) were `claimed`, asserted but not established by the evidence in the commit; the rest were partial (78) or contradicted (2). Review goes to the two-thirds the label marks. How often the `verified` label is itself wrong is not yet measured.
39	§4 Line 2	OVERCLAIM	**Line 2. Incidents are caught by tomorrow, not at the audit.** Our record holds 43 incidents in six months; the ones that mattered were caught by the record before they were visible in the agent's own account, most within a day. Industry figure: 241 days to identify and contain;	**Line 2. Incidents are caught by the record, not at the audit.** Our incident register holds 43 entries in six months (as of 17 September): 38 of our own failures, near-misses and design holes, and 5 from the field, each dated with the mechanism that caught it or failed to. Several were caught by the record within minutes: the agent editing its own gate, a forked sequence. Others were found by red-team review or by the operator. The worst, a verifier silently degraded for five weeks, was caught late, and the fix was a liveness check that pages. Industry figure: about 247 days to identify and contain;
40	§4 Line 3	OVERCLAIM	Every day's work is already signed, stamped, and independently verified. A compliance package that today costs weeks of evidence collection is a query. Industry figure: a first SOC 2 runs $10,000 to $80,000; a small defense contractor's CMMC assessment $30,000 to $70,000; bank model validation twenty weeks per model; data-integrity findings in most FDA warning letters.	Since 9 September, every day's commits are signed, stamped each morning, and graded by an independent machine. A compliance package that today costs weeks of evidence collection could be produced by query; building and timing that query is part of the pilot. Industry figure: a first SOC 2 runs $10,000 to $80,000; DoD's own estimate of a small contractor's CMMC Level 2 certification assessment is about $102,000, before the cost of implementing the controls; bank model validation runs about twelve weeks per top-tier model in the United States and twenty in Europe; data-integrity findings are still cited in about one FDA warning letter in seven.
41	§4 Line 4	WRONG	Ours: 98 percent cache share, $1.48 per commit, and the one un-rotated day that cost more than the next six seats combined.	Ours: 98 percent cache share, $1.48 per commit, and the one un-rotated seat that cost about 45 percent more per model call than the next day's six rotated seats, and stalled.
42	§4 Line 5	OVERCLAIM	A replacement instance, a different model, or a different vendor picks up from the record in minutes, because the record is the memory. Ours: three minutes and one dollar. Industry figure: knowledge workers spend about a fifth of the week finding information; 5.3 hours a week waiting for or recreating knowledge (dated surveys, cited as such).	A replacement instance picks up from the record in minutes, because the record is the memory; a different model or a different vendor is what the design is for and is not yet measured. Ours: a pre-warmed relief to its first useful commit in three minutes; a cold instance's orientation for about a dollar, its answer not yet graded. Industry figure: knowledge workers spend about a fifth of the week finding information (McKinsey, 2012); 5.3 hours a week waiting for or recreating knowledge (Panopto, 2018).
43	§4 closing	OVERCLAIM	The single-company numbers above are real, dated, and reproducible from files in the record.	The single-company numbers above are real, dated, and reproducible by us from files in the record.
44	§5 Federal	CITATION	Proposal cost runs 0.2 to 3 percent of contract value; 1,688 bid protests were filed in fiscal 2025; DCAA examined $788 billion and found $18.8 billion in exceptions. What is coming: a House bill requiring continuous verification and tamper-proof logs of agent actions for contractors, and a defense authorization act folding AI security into CMMC, whose certification already costs a small contractor $75,000 to $150,000 in year one.	Proposal cost runs 0.2 to 3 percent of contract value by consultants' estimates; 1,688 bid-protest cases were filed with GAO in fiscal 2025; DCAA examined $788 billion and found $18.8 billion in exceptions. What is coming: a House bill (H.R. 10362) that would require tamper-evident logs of agent actions in new federal contracts, and a defense authorization act folding AI security into CMMC, whose Level 2 certification assessment DoD itself estimates at about $102,000 for a small contractor, before the cost of implementing the controls.
45	§5 Engineering	CITATION	and a single inspector-general review disallowed $15 million in indirect costs.	and a single 2009 inspector-general audit found $15.7 million in unallowable indirect costs claimed by design firms.
46	§5 Financial	CITATION	Regulators have said books-and-records and model-risk rules apply unchanged to AI output; the SEC's 2026 examination priorities ask whether changes to system logic are documented; initial validation of a top-tier model takes about twenty weeks; compliance consumes more than a tenth of operating cost.	FINRA has said books-and-records and supervision rules apply unchanged to AI output; the SEC's 2026 examination priorities test whether firms' AI controls match their disclosures; initial validation of a top-tier model takes about twelve weeks at U.S. banks and twenty in Europe; compliance is commonly put at a tenth or more of a bank's operating cost.

#	Where	Class	v1 (old)	v2 (new)
47	§5 Pharma	CITATION	Data-integrity findings appear in most drug-manufacturing warning letters; warning letters rose 59 percent in fiscal 2025; remediation under consent decrees has cost individual companies hundreds of millions.	Data-integrity findings still appear in about one in seven drug warning letters, and in 60 percent of those issued to sites in India; warning letters to drug and biologics makers rose 59 percent in fiscal 2025; remediation under consent decrees has cost individual companies hundreds of millions (Schering-Plough paid $500 million under its 2002 decree).
48	§5 Legal	CITATION	where the leading AI tools fabricate 17 to 33 percent of the time and courts have sanctioned lawyers in 1,668 catalogued cases.	where the leading AI research tools gave incorrect or mis-sourced answers 17 to 33 percent of the time in Stanford's tests and courts have addressed AI-hallucinated citations in more than 2,000 catalogued decisions.
49	§5 Insurance	CITATION	under the regulators' model bulletin, adopted in 24 states, which requires a written AI program, board accountability, record retention, and production of documents on examination.	under the regulators' model bulletin, adopted in more than half the states, which expects a written AI program, board accountability, record retention, and production of documents on examination.
50	§6 hardware	OVERCLAIM	a third holding the observer and the local verifier, which accepts no writes from the agent. The frontier model is whatever you already pay for. The verifier is an open model on a $2,000 desktop machine.	a third holding the observer and the local verifier, which accepts only appends from the agent and refuses any rewrite. The frontier model is whatever you already pay for; custody of the record stays home, the model's compute is rented wherever you already allow it, and the record outlives whichever vendor supplied it. The verifier is an open-weights 32-billion-parameter model on a desktop machine with 24 gigabytes of memory.
51	§6 pricing	DECISION	[v1 sentence withheld: internal pricing]	Pilot pricing is quoted per workflow, priced like compliance work and not like software seats.
52	§6 comparables	CITATION	AI governance platforms run $30,000 to $150,000 a year and produce policy documents. Compliance-automation tools run $10,000 to $25,000 a year to collect evidence. The category Gartner created in June 2026, AI governance platforms, is forecast at $492 million this year and more than a billion by 2030.	AI governance platforms, at reported prices in the tens of thousands of dollars a year, manage AI inventories, policies and approvals. Compliance-automation tools run about $10,000 to $40,000 a year at typical purchase prices to collect evidence. Gartner, which published its first Magic Quadrant for AI governance platforms in June 2026, forecasts spending on them at $492 million this year and more than $1 billion by 2030.
53	§7	OVERCLAIM	the value is that it is real, dated, continuous, and cannot be manufactured after the fact.	the value is that it is real, dated, continuous, and cannot be changed after its first public timestamp (9 September 2026) without the change showing; what came before that date is attested from that date, not from when it was written.
54	§7	WRONG	Precision of the claim-labelling is measured on six human-graded rows and is the start of a measurement, not a result. The system caught its own mistakes because it was built to, and section 24 below lists the ones it did not catch in time.	Precision of the claim-labelling has been checked on five decided rows, all of them rows the verifier had flagged, each labelled by the agent and confirmed by the operator; recall has not been measured. That is the start of a measurement, not a result. The system caught many of its own mistakes because it was built to; section 24 below lists five, including the one it caught five weeks late and the one its operator saw first.
55	§22 table	WORDING	| Actor-ledger rows since 9 Sep | 46,534 |	| Actor-ledger rows since 9 Sep (mostly the observer's own polling) | 46,534 |
56	§22 table	OVERCLAIM	| Decisions recorded | 92 |	| Decision entries recorded (80 decided, the rest open or tabled) | 92 |
57	§22 table	WRONG	| Rows where the semantic verifier changed the deterministic verdict | 248 of 557 (tightened 233, loosened 15) || Human adjudications | 6 (3 true positives, 2 false positives, 1 unsure) || Flagging precision / recall on labelled rows | 60 % / 100 % (n = 6) |	| Rows where the semantic verifier was available | 404 of 557 || Rows where it disagreed with the deterministic floor | 248 of 404; because the stricter verdict wins, it tightened the final label on 233 and loosened none || Human adjudications | 6, all on rows the verifier flagged: 3 correct, 2 false positives, 1 unsure || Flagging precision | 3 of 5 decided rows; too few to state as a rate || Recall | not yet measurable: no unflagged row has been adjudicated |
58	§22 table	OVERCLAIM	| Incidents recorded | 43 |	| Incident register entries (as of 17 September: 38 ours, 5 from the field) | 43 |
59	§22 text	OVERCLAIM	Half of the observer's rows are `claimed`: the commit message asserted something its own evidence did not establish. Most are benign in content and exactly the class the protocol exists to label. Six labels against 557 rows is the start of a precision measurement, not one.	Half of the observer's rows are `claimed`: a 32-billion-parameter checker judged that the commit asserted more than its own evidence shows (the deterministic floor alone says so for about a fifth). Most are benign in content and exactly the class the protocol exists to label. Both `contradicted` rows were adjudicated as false positives, so the verifier's only two contradictions were its own errors. Five decided labels against 557 rows is the start of a precision measurement, not one.

#	Where	Class	v1 (old)	v2 (new)
60	§23	OVERCLAIM	The verdict mix is stable at roughly half `claimed` across two months and three seat models: the label measures a property of agent language, not of one model.	Since late July the cumulative mix has held at roughly half `claimed`, across two versions of the seat model; a per-model breakdown is not yet measured. We read that as a property of how agents describe work a diff cannot show (tests run, services restarted), which is what the action capture added on 10 September now records; published benchmarks find the same behaviour under controlled conditions (section 25).
61	§23	OVERCLAIM	A cold-start drill on 5 September produced a correct next-actions answer from the record alone for $1.04; its grading table is still empty, and we record that rather than hide it.	A cold-start drill on 5 September produced a next-actions answer from the record alone for $1.04; it has not yet been graded, and we record that rather than hide it.
62	§24 opening	WORDING	Forty-three incidents, each with a date, a class, the mechanism that caught it or failed to, and the measure that followed. Five, briefly.	Forty-three register entries as of 17 September, each with a date, a class, the mechanism that caught it or failed to, and the measure that followed. Five, briefly.
63	§24 I-32	WRONG	At the edge of the context window the harness replaced working memory with a model-written summary placed in the operator's slot; a gate read its quoted words as typed. The record showed the summary was model text. Fix: every gate keys to the prompt hook, so a summary's quoted words no longer pass as typed; summaries are claims; the seat rotates before the edge.	At the edge of the context window the harness replaced working memory with a model-written summary placed in the operator's slot; a gate read its quoted words as typed. The record did not catch it: the summary became working memory with no row anywhere, and the verbatim original existed only in the local session store. The operator noticed first; the session store, once examined, showed the summary was model text, and the original was captured to the record after the fact. Fix: every gate keys to the operator's session hook, so a summary's quoted words no longer pass as typed; summaries are claims; the seat rotates before the edge.
64	§24 closing	OVERCLAIM	The pattern across all five: the failure was visible in the record before it was visible in the agent's account of itself.	In three of the five, the failure was visible in the record before it was visible in the agent's account of itself. In the other two it was the operator, or hindsight, that found it: visible in the record, if not always seen. Each became a fix.
65	§25 (new)	ADDED	## 26. Limitations	(new section; see v2)
66	§26	WRONG	One operator, one system, one agent, six labelled rows. Signing began in September; 427 earlier commits are unsigned and graded rather than re-signed. An early SMB share on the observer machine was a second door, closed in September. The capture plane began on 9 September; earlier agent messages are checked only at commit granularity. One semantic verifier model; a second architecture designed, not deployed. Precision unmeasured until the adjudication corpus grows through ordinary use. No customer deployment; the pilot is the next measurement.	One operator, one system, one agent, five decided rows. Signing began on 9 September; 427 earlier commits are unsigned and graded rather than re-signed, and the first public timestamp is 9 September, so what came before is attested from that date. One signing key, held on the working machine the agent uses: a signature proves custody by that machine, not a human act. An early file share on the observer machine was a second door, closed in September for the record's clone; the observer's home folder is still reachable from the working machine, an open item recorded as such. Capture began on 10 September and runs inside the agent's own harness as one hook per tool call: a shell command is one row and what happens inside it is not broken out, nothing outside the harness is captured, and no completeness figure is yet in the record; earlier agent messages are checked only at commit granularity. What the agent cannot do is change the record afterward without the witnesses seeing it by morning. The 557 verifications and the 6 adjudications live on the verifier machine, outside the git record, in a folder the working machine can still reach, and are not yet externally stamped; a hash of the 557-row cut is committed beside this release, and the record's daily stamp covers it from then on. One semantic verifier model; a second architecture designed, not deployed. Precision is a first reading on five rows; recall is unmeasured until unflagged rows are adjudicated. No customer deployment; the pilot is the next measurement.
67	About the author	WORDING	He is a retired Army officer, Lieutenant Colonel (Ret.), a graduate of West Point in mechanical engineering and of the Darden School in business,	He retired from the Army as a lieutenant colonel, and is a graduate of West Point in mechanical engineering and of the Darden School in business,
68	Sources (new)	ADDED	*—First public release, 28 September 2026. Cut from working draft v5 of 19 September 2026 (Parts A and B unchanged since v4; Part C interim). Outside figures are cited to their sources in the companion research note.*	(new section; see v2)

Stage 2 — an internal hostile read of v2 (27 September), applied to the v2 text

#	Where	Class	v2 (old)	v2.1 (new)
S1	Abstract	WRONG	kept where the agent cannot edit it, checked by something the agent does not control	kept where the agent cannot quietly change it, checked by something the agent does not control
S2	Abstract	WRONG	and in its present form since 9 September 2026, when signing, public timestamps, and action capture began; most of the numbers below	and in its present form since 9 September 2026, when signing and public timestamps began (action capture followed the next day); most of the numbers below
S3	About this release	OVERCLAIM	version 1 (SHA-256 f5900deb…) was posted on appendonly.ai on 25 September 2026 and is superseded by this one, with every change listed in a corrections table in the record.	version 1 (SHA-256 f5900debd1c191cc1ce288edadefd05fb05ddaad13a2c67846008b6547c142ec) was posted on appendonly.ai on 25 September 2026 and is superseded by this one, with every change listed in a corrections table kept in the company's record.
S4	§2 capture	WRONG	written down by a hook the agent does not control, at the moment the call returns, as fingerprints of what went in and what came out, below the level of language,	written down by a hook that runs beside the agent, not by the agent's own account of itself, at the moment the call returns, as fingerprints rather than text (a shell command is recorded as one row, by its length, with what happened inside it not broken out), below the level of language,
S5	§2 capture	ADDED	and a claim that exceeds it is marked as such, automatically, without the agent's cooperation. The record cannot be quietly changed	and a claim that exceeds it is marked as such, automatically, without the agent's cooperation. Any change to the hook itself is reported by the next check. The record cannot be quietly changed
S6	§2 data that leaves	OVERCLAIM	plus a 32-byte fingerprint to the timestamp service. Three small computers	plus a 32-byte fingerprint to the timestamp service and the operator's own pager messages. Three small computers
S7	§3 caching	WORDING	Caching is what makes a long-running agent affordable at all, and caching works when the context is append-only and stable, the same discipline a stable record encourages; a 98 percent cache share is typical of a well-run harness, not unique to ours.	Caching is what makes a long-running agent affordable at all, and it works only when what the agent re-reads is stable. The record is that stable context: the orientation digest and the hooks' injected material are paid once and read from cache after. Our 98 percent share is what a well-run harness should show; the point is that the record rides inside it for almost nothing.
S8	§4 Line 2	ADDED	and 5 from the field, each dated with the mechanism that caught it or failed to. Several were caught by the record within minutes: the agent editing its own gate, a forked sequence. Others were found by red-team review or by the operator. The worst, a verifier silently degraded for five weeks, was caught late, and the fix was a liveness check that pages. Industry figure:	and 5 recorded from outside our own operation, each dated with the mechanism that caught it or failed to. Several were caught by the record within minutes: the agent editing its own gate, a forked sequence. Others were found by red-team review or by the operator. The worst, a verifier silently degraded for five weeks, was caught late, and the fix was a liveness check that pages. Each late catch became a check the record now runs: the verifier's silence now pages, and a summary's words no longer pass as typed. Industry figure:
S9	§6 pricing	WORDING	Pilot pricing is quoted per workflow, priced like compliance work and not like software seats.	Pilot pricing is quoted per workflow, priced like compliance software and not by the seat.
S10	§22 table	WORDING	| Incident register entries (as of 17 September: 38 ours, 5 from the field) | 43 |	| Incident register entries (as of 17 September: 38 ours, 5 recorded from outside our own operation) | 43 |
S11	§25 closing	OVERCLAIM	Not first and not largest; measured, dated, and checkable.	Not first and not largest; measured, dated, and, from the first evidence bundle on, checkable.
S12	§26	OVERCLAIM	The 557 verifications and the 6 adjudications live on the verifier machine, outside the git record, and are not yet externally stamped.	The 557 verifications and the 6 adjudications live on the verifier machine, outside the git record, in a folder the working machine can still reach, and are not yet externally stamped; a hash of the 557-row cut is committed beside this release, and the record's daily stamp covers it from then on.

Stage 3 — the dated measurement file (27 September)

#	Where	Class	v2.1 (old)	v2.2 (new)
M1	About this release (measurement file)	ADDED	and is reproducible by us from files in the record; what a stranger	and is reproducible by us from files in the record, with the commands and their output committed in a dated measurement file beside this release; what a stranger

#	Where	Class	v2.1 (old)	v2.2 (new)
M2	§3 intro (rates per model, cutoff)	WRONG	taken from the session records themselves. Prices are Anthropic's published list rates, used as a conversion so that a reader can compare; the account actually runs on a subscription.	taken from the session records themselves (cutoff 17 September 2026, 13:53 EDT). Prices are Anthropic's published list rates for the model each message actually ran on (Fable 5 until 1 September, Fable 5.1 after, a few Opus 4.8 turns), as listed on the vendor's price page on 26 September 2026 and assumed unchanged since July; they are a conversion so that a reader can compare, and the account actually runs on a subscription.
M3	§3 table (total)	WRONG	| Seat sessions' model spend at list price, 55 days (subagents and the scribe add roughly a sixth) | $620.65 |	| Seat sessions' model spend at list price, 55 days (subagents and the scribe add roughly a sixth) | $705.51 |
M4	§3 table (per day)	WRONG	| Average per day | about $11 |	| Average per day | about $13 |
M5	§3 table (per commit)	WRONG	| Cost per commit to the record (419 commits) | about $1.48 |	| Cost per commit to the record (419 commits) | about $1.68 |
M6	§3 table (uncached ratio)	WRONG	| What the same context would have cost uncached | about $8,100, roughly twenty times the about $415 paid for context (cache reads $203, cache writes about $207, uncached about $3) |	| What the same context would have cost uncached | about $8,100, roughly sixteen times the about $500 paid for context (cache reads $293, cache writes $206, uncached $2) |
M7	§3 table (output share)	WRONG	| Output tokens, all sessions | 4.1 million, about one third of the bill |	| Output tokens, all sessions | 4.1 million, about 29 percent of the bill |
M8	§3 table (handoff tokens)	OVERCLAIM	| Handoff: a pre-warmed relief to its first useful commit | 3 minutes, about 110,000 tokens |	| Handoff: a pre-warmed relief to its first useful commit | 3 minutes |
M9	§3 scribe	OVERCLAIM	the scribe that digests each session runs on a small model at about four cents per digest, nine times a day;	the scribe that digests each session runs on a small model about eight times a day, at a few cents per digest by our estimate;
M10	§3 scribe (daily)	OVERCLAIM	About forty cents a day in model spend for the scribe;	A few tens of cents a day in model spend for the scribe, by that estimate;
M11	§4 Line 4	WRONG	Ours: 98 percent cache share, $1.48 per commit,	Ours: 98 percent cache share, $1.68 per commit,
M12	§22 table (actor ledger)	OVERCLAIM	| Actor-ledger rows since 9 Sep (mostly the observer's own polling) | 46,534 |	| Actor-ledger rows since 9 Sep (mostly the observer's own polling) | about 46,500 |
M13	§22 table (vault notes)	WRONG	| Vault notes | 1,113 |	| Vault notes | 1,112 |
M14	§22 verb list	OVERCLAIM	retrievals 40; analyses 23; messages sent 22; answers 16; failures 1.	retrievals 40; analyses 23; messages sent 22; answers 16; tool calls 14; failures 1.

Stage 4 — cache writes were all one-hour writes (27 September)

#	Where	Class	v2.2 (old)	v2.3 (new)
W1	§3 intro (write TTL)	WRONG	as listed on the vendor's price page on 26 September 2026 and assumed unchanged since July;	as listed on the vendor's price page on 26 September 2026 and assumed unchanged since July; every cache write the seat made was a one-hour write and is priced at that rate;
W2	§3 table (total)	WRONG	| Seat sessions' model spend at list price, 55 days (subagents and the scribe add roughly a sixth) | $705.51 |	| Seat sessions' model spend at list price, 55 days (subagents and the scribe add roughly a sixth) | $829.12 |
W3	§3 table (per day)	WRONG	| Average per day | about $13 |	| Average per day | about $15 |
W4	§3 table (build days)	WRONG	| Build days (9 to 17 September) | $0 to $111 per day, average $46 |	| Build days (9 to 17 September) | $0 to $133 per day, average $56 |
W5	§3 table (per commit)	WRONG	| Cost per commit to the record (419 commits) | about $1.68 |	| Cost per commit to the record (419 commits) | about $1.98 |
W6	§3 table (uncached ratio)	WRONG	| What the same context would have cost uncached | about $8,100, roughly sixteen times the about $500 paid for context (cache reads $293, cache writes $206, uncached $2) |	| What the same context would have cost uncached | about $8,300, roughly thirteen times the about $625 paid for context (cache reads $293, cache writes $330, uncached $2) |
W7	§3 table (output share)	WRONG	| Output tokens, all sessions | 4.1 million, about 29 percent of the bill |	| Output tokens, all sessions | 4.1 million, about a quarter of the bill |
W8	§3 table (I-39 seat)	WRONG	| One seat left running about two days without rotation (10 to 12 September, I-39) | $168, about $76 a day; peak context 997,000 tokens; ended stalled |	| One seat left running about two days without rotation (10 to 12 September, I-39) | $201, about $91 a day; peak context 997,000 tokens; ended stalled |
W9	§3 table (six seats)	WRONG	| The six rotated seats of 13 September | $64 for the day, no stall |	| The six rotated seats of 13 September | $74 for the day, no stall |

#	Where	Class	v2.2 (old)	v2.3 (new)
W10	§3 argument (per call)	WRONG	cost more per day, and about 45 percent more per model call, than the following day's six rotated seats,	cost more per day, and about 50 percent more per model call, than the following day's six rotated seats,
W11	§4 Line 4	WRONG	Ours: 98 percent cache share, $1.68 per commit, and the one un-rotated seat that cost about 45 percent more per model call	Ours: 98 percent cache share, $1.98 per commit, and the one un-rotated seat that cost about 50 percent more per model call

Stage 5 — release decisions of 27 September, corrected by a second hostile read

#	Where	Class	v2.3 (old)	v2.4 (new)
D1	About	WRONG	The second release, the system as deployed, is scheduled for 12 October 2026. The third, a case study of one federal proposal run on the record, follows the captured run.	The full report, the system as deployed, and Part C, a case study of one afternoon's work run on the record, follow this release once the filing that covers their additions is on the record.
D2	About	OVERCLAIM	with every change listed in a corrections table kept in the company's record.	with every change listed in a corrections table published beside it.
D3	§5, after the federal paragraph	OMISSION	—	We sell for acquisition, administrative, engineering, and document-review decisions.
D4	§22 table	WORDING	Observer verifications (Node 3, since 14 Jul)	Observer verifications (the observer machine, since 14 Jul)
D5	§24	WORDING	**The genesis case (I-01, 28 March).**	**The first case (I-01, 28 March).**
D6	§26	OVERCLAIM	An early file share on the observer machine was a second door, closed in September for the record's clone; the observer's home folder is still reachable from the working machine, an open item recorded as such.	An early file share on the observer machine was a second door; it was closed for the record's clone in September and shut entirely on 27 September, so the working machine now reaches the observer only through a restricted channel that accepts commits and answers read requests, never a write to the observer's own store.
D7	§26	OVERCLAIM	in a folder the working machine can still reach, and are not yet externally stamped;	in a folder the working machine can no longer reach, and are not yet externally stamped;
D8	Sources	CITATION	SEC Division of Examinations, Fiscal Year 2026 Examination Priorities.	SEC Division of Examinations, Fiscal Year 2026 Examination Priorities, p. 17.
D9	§26	WRONG	in a folder the working machine can no longer reach, and are not yet externally stamped;	in a folder the working machine can read over the restricted channel but can no longer write to, and are not yet externally stamped;
D10	§26	OVERCLAIM	so the working machine now reaches the observer only through a restricted channel that accepts commits and answers read requests, never a write to the observer's own store.	so the working machine now reaches the observer only through a restricted channel: it can add commits to the record's clone there and read the observer's ledgers, and it cannot write to those ledgers.
D11	About	WORDING	The full report, the system as deployed, and Part C, a case study of one afternoon's work run on the record, follow this release once the filing that covers their additions is on the record.	The full report (the system as deployed) and Part C (a case study of one afternoon's work run on the record) follow this release once the filing that covers what they add is on the record.
D12	Sources	CITATION	SEC Division of Examinations, Fiscal Year 2026 Examination Priorities, p. 17.	SEC Division of Examinations, Fiscal Year 2026 Examination Priorities, p. 13 (page 17 of the PDF).

Stage 6 — the release day moved to 29 September, and the owner's review notes of 29 September (v2.4 → v2.5)

#	Where	Class	v2.4 (old)	v2.5 (new)
D13	header (PDF running header; corrections-table title and preamble)	DECISION	first public release · 28 September 2026	first public release · 29 September 2026
E1	About	DECISION	The full report (the system as deployed) and Part C (a case study of one afternoon's work run on the record) follow this release once the filing that covers what they add is on the record.	The full report (the system as deployed) and Part C (a case study of one afternoon's work run on the record: the rebuild of our own website) are planned for Thursday 1 October 2026 and follow this release once the filing that covers what they add is on the record.

#	Where	Class	v2.4 (old)	v2.5 (new)
E2	§2	ADDED	with what happened inside it not broken out), below the level of language, so the agent cannot describe its work more favourably than the facts.	with what happened inside it not broken out). A language model works in language: left to itself, it can only tell you what it did. The record does not take the telling. Each row names the act from a fixed list of verbs (searched, fetched, read a file, ran a command, committed), with what it acted on and when, so the path the work took is the evidence, and an agent cannot change that path with words. This is the method in our patent filings, the Action-Evidence Protocol (AEP).
E3	§2	ADDED	A change to anything already stamped shows by the next morning.	A change to committed history shows at the observer's next check, and a change to anything already stamped shows against the public stamp; on our system the first runs every minute and the second every morning, and a customer sets the schedule.
E4	§4 Line 2	ADDED	each dated with the mechanism that caught it or failed to. Several were caught	each dated with the mechanism that caught it or failed to. How fast a catch comes depends on the check and the schedule it is given: on our system the observer checks every minute and the public stamp is taken every morning, and a customer sets the schedule. Several were caught
E5	§6	WORDING	**A pilot with a number at the end.**	**A pilot with a measured number at the end.**
E6	§7	WORDING	a change to the past shows by the next morning;	a change to the past shows at the next check, on our system every minute at the observer and every morning at the public stamp;
E7	§26	WORDING	without the witnesses seeing it by morning.	without the witnesses seeing it at their next check.
E8	§24	WORDING	At the edge of the context window the harness replaced working memory with a model-written summary placed in the operator's slot; a gate read its quoted words as typed.	At the edge of the context window the harness compacted the session automatically, with no command from the operator: it replaced the agent's working memory with a model-written summary and placed that summary in the operator's slot, and a gate read its quoted words as typed. A compaction is an unverified rewrite of memory.

Stage 7 — an internal hostile read of stage 6, 29 September (v2.5 → v2.6)

#	Where	Class	v2.5 (old)	v2.6 (new)
G1	§2	OVERCLAIM	Each row names the act from a fixed list of verbs (searched, fetched, read a file, ran a command, committed), with what it acted on and when, so the path the work took is the evidence, and an agent cannot change that path with words. This is the method in our patent filings, the Action-Evidence Protocol (AEP).	Each row names the act from a fixed list of verbs (searched, fetched, read a file, wrote a file, ran a command), with the tool, the time, and fingerprints of what went in and what came out; a file row names the file and a fetch names the site. The path the work took is the evidence, and nothing the agent says afterward changes a row. This is the Action-Evidence Protocol (AEP), first described in our provisional patent application of April 2026.
G2	§2	OVERCLAIM	and a change to anything already stamped shows against the public stamp; on our system the first runs every minute and the second every morning, and a customer sets the schedule.	and a change to anything already stamped shows at the next check against the public stamp; on our system the first runs every minute and the second every morning. In a pilot the schedule is set with the customer.
G3	§4 Line 2	OVERCLAIM	on our system the observer checks every minute and the public stamp is taken every morning, and a customer sets the schedule.	on our system the observer checks every minute and the public stamp is taken every morning. In a pilot the schedule is set with the customer.
G4	About	WORDING	are planned for Thursday 1 October 2026 and follow this release once the filing that covers what they add is on the record.	are planned for 1 October 2026. They will not be released before the filing that covers what they add is on the record.