> **About this rendering.** Built 2026-09-19 by the data-room pipeline from `DIGITAL_PILOT_WHITE_PAPER.md` (source last edited 2026-09-19) against the Nightingale OS repository at commit `f3e28b3c5` (2026-09-19). The company's Canonical Facts Register (Edition 1 · compiled 2026-09-12) is the authority on every count; where the text was typed before the register's current measurement, the measured value is stated here and the text is left as written pending its correction.
>
> This paper carries both a superseded and the measured bundle count; at this build's commit: 1,347 (905 surfaced). Every dollar figure and percentage is modelled on a synthetic facility and reads as modelled. The refusal-code count in the text is one of several live values the register has not reconciled. Its retraction list is the register's F-046.
>
> The corporation is Nightingale OS Inc., Delaware — certificate of incorporation executed 12 August 2026 and held in the company record (register F-002, RESOLVED by the instrument). The Delaware filing itself is founder-attested (F-003, 2026-09-14) and is not yet evidenced by a file-stamped copy. An entity line elsewhere in this document may carry an older styling from its source; the certificate governs.
>
> Statements about data handling, PHI or security in this document describe design intent at the cited commit. Nothing here states a conclusion about the company's PHI handling or security posture, in either direction, and no certification of any kind is held.

The Nightingale OS Digital Pilot

A control-facility rerun of the 274-bed pilot, with the July 2026 model retracted and replaced

Kristen Cline, BSN, RN, CEN, CPEN, TCRN, CFRN, CTRN, CCRN Founder & Architect, Nightingale OS
Nightingale OS Inc., a Delaware C-Corp · certificate of incorporation executed 12 August 2026 and held in the company record; the Delaware filing itself is founder-attested (2026-09-14) and is not yet evidenced by a file-stamped copy · Pre-Seed Correspondence: kristen@nightingaleos.com

September 19, 2026 — Revision 3.1 (current control rerun). *Supersedes the July 1, 2026 edition (Revision 2), which is retained unaltered at `superseded/DIGITAL_PILOT_WHITE_PAPER_REV2_2026-07-01.md`. Revision 3 is a retraction-and-replacement edition: the July model was audited line by line against the code and the literature, thirty-six of its published statements did not survive, and the replacements are produced by two reproducible machine runs rather than by a hand-maintained metrics file. Revision 3.1 retains Revision 3's September 2 run as a historical artifact and makes the September 19, 2026 control rerun the current artifact. The legacy July Word and DocSend renders are frozen records; current PDFs are rendered through the data-room pipeline.*

STATUS: APPROVED by Kristen Cline, 2026-09-19. Revision 3.1 updates the current control-run figures; the September 2 artifacts remain historical records and must not be represented as current evidence. **Corrected after approval, 2026-09-15, on her instruction — three factual corrections, no argument changed.** The masthead still carried "founder attestation 2026-09-13; certificate to follow", contradicting this document's own footer, and is now aligned to it — the 2026-09-13 date was unsourced. And the patent was described as filed on a founder statement, when the USPTO Electronic Acknowledgement Receipt for application 64/139,773, received 2026-08-24, is in the company record.

Founder review — four items that need a decision before this ships

1. **The risk-purchase sentence (§5.1).** Revision 3 prints, in the company's own voice: *a

substitution-only ROI does not clear the Owl subscription price, and the purchase is a risk purchase rather than a savings purchase.* That sentence is true and defensible, and it hands a CFO the weakest form of the commercial argument in writing. It stays or goes on a founder call.

1. **The Base-tier inversion (bead `nightingale-eaave`).** Holding modeled value constant per bed

while pricing sub-150-bed facilities at Base $300/bed/yr made the smallest hospitals show the highest return in Revision 2's scaling table — 25× at Base against 12.6× at the pilot's Standard tier. Revision 3 withdraws the per-bed value figure that table rested on, so the inversion does not reach print here, and it is escalated rather than repaired.

1. **Patent wording.** Charlie Nuthatch is described as **filed** — USPTO application 64/139,773,

received by the USPTO 2026-08-24 (Electronic Acknowledgement Receipt, in the company record). Revision 3 claims nothing further for it — no grant, no allowance, no scope.

1. **The never-event category falls from $264,870 to $112,556.50.** The July model applied an

all-falls baseline (3.5 per 1,000 patient-days) and then priced and labeled the result as falls *with injury*, which over-counts avoided injuries by about 3.8×; the injurious-fall baseline is 0.93 per 1,000. The CAUTI cost basis moves from $25,000 to AHRQ's pooled $13,793 for the same reason of source discipline. Both corrections are shown beside the July arithmetic in §4.2.

Abstract

Nightingale OS is a bedside-first hospital operating system. It captures structured clinical event data once at the bedside, composes FHIR R4 transaction bundles for the legal chart automatically, and enforces patient-safety and staffing guardrails as software constraints rather than as policies a busy unit can quietly bypass. Revision 2 of this paper projected $2.06M of annual value at a synthetic 274-bed community hospital against a $164,400 platform cost. That projection came from a hand-maintained JSON record that no code path read, and this revision withdraws it.

Revision 3.1 reports what the platform's own engines produce when they are run against the same 274-bed control facility, seeded, DB-free, and reproducible from a clean clone by two commands. Two current machine artifacts carry the current-run summary printed here: an engines run (`docs/owl/pilot/control-facility-engines/run-2026-09-19/`, run hash `76eedcef…`) and an Owl gap assessment and remediation roadmap (`docs/owl/pilot/control-facility/run-2026-09-19/`, run hash `27bee202…`). The engines run reproduces the July never-event arithmetic to the cent, then prints the corrected figure beside it — $264,870.00 as published against $112,556.50 corrected — and identifies eleven of the July model's value lines that no code in the repository computes at all. The Owl run assesses nine accreditation programs over 110 evaluable requirement elements — six more could not be read — bands the facility at-risk on evidence coverage, returns 45 findings, a 51-step roadmap with a 28-day critical path, and prices the work to close at $31,285 … $35,857 … $44,565 including one $25,000 external mock survey.

Revision boundary. Sections and tables below that expressly name the September 2, 2026 artifacts remain the Revision 3 historical analysis. They are not current evidence. The current artifacts are the September 19 runs named above and in Appendix A; `latest` resolves to those artifacts and both official `--verify` commands pass against it.

The paper reports the model, its provenance, its pre-registered falsifiable claims audited against current code, and its limitations — including that no patient is real, that no readiness score has ever been compared against an actual survey outcome, and that six named refusals in the Owl run are the machine declining to price things for which no published basis exists. Nothing here is a claim of effect. It is an auditable projection in which every number traces to a machine artifact, a published source, or a labeled assumption. The machine itself was red-teamed under three adversarial lenses before this edition shipped — twenty-two findings, twenty-one of them repaired in the code this revision runs on — and §9 names every finding still open, from this review and from the reviews before it.

1. What changed since Revision 2

Revision 2 was audited in full against the repository at commit `6afb979ee` and against the primary literature, over eight read-only passes and one adversarial completeness review. The findings below are the ones that changed a printed statement. Each names what Revision 2 said, what Revision 3 says instead, and where the correction comes from. Nothing on this list is a softening; each item is a retraction of a sentence the company published.

1.1 The value model

1. **The metrics record is retired.** Revision 2's Appendix A was reproduced from

`data/demo/pilot_metrics.json`. A grep across `server/src`, `client/src`, `landing`, `scripts` and `docs` returns no reader of that file: it is hand-maintained JSON, not a generated artifact, and no code path connects the stated method to the published numbers. Revision 3 replaces it with two committed machine artifacts and their run hashes.

1. **Never events: $264,870.00 as published, $112,556.50 corrected.** Both computations sit side

by side in §4.2 with their denominators, baselines and cost bases.

1. **The falls line falls from $176,400 to $39,841.20.** Revision 2 applied 3.5 falls per 1,000

patient-days — an all-falls rate — and priced the outcome at a fall-with-injury cost. The injurious-fall baseline is 0.93 per 1,000 (26.1% of 3.56, Bouldin 2013), and the defensible reduction is 34% rather than 40% (Dykes 2020).

1. **The never-event denominators are now published, and so is their scale.** Falls run on 8,400

patient-days, pressure injuries on 600 admissions, CAUTI on 2,100 catheter-days, CLABSI on 1,260 line-days. That patient-day base is an annualized cohort snapshot with an average daily census near 23, against 100,010 bed-days at 274-bed occupancy — a factor of 11.9. Revision 2 printed none of this. The cohort base is kept as the floor rather than rescaled upward.

1. **Average length of stay.** Revision 2 printed 21.0 days. The fixture supports "20.97-day mean

expected length of stay across the 15 patients carrying a projected discharge date"; 98 of the 100 patients share one admission date, so the cohort is a point-in-time census snapshot.

1. "**85% census-forecast accuracy**" **is dropped.** Nothing in the repository computes it; the

register entry is Appendix B, line 10.

1. **$50,000 per ambulance diversion is relabeled an unsourced assumption.** The literature prices

diversion per hour; at McConnell 2006's $1,086/hour the figure implies a 46-hour episode.

1. **$2,000 per agency shift is relabeled an unsourced assumption,** and the reuse behind it is

disclosed: the agency line and the overtime line are the same unrounded quantity, 67.392 shifts, credited to two levers carrying two different reduction rates.

1. **The $55/hour fully-loaded RN rate gets its first anchor.** BLS OEWS May 2025 puts the RN mean

at $48.76/hour (SOC 29-1141, series `OEUN0000000000000029114103`), so $55 implies a 1.13× loading — conservative against the usual 1.3–1.4×, and now stated rather than asserted.

1.2 Sources and citations

1. **$25,000 per CAUTI becomes $13,793.** AHRQ's pooled additional-cost estimate is $13,793 (95%

CI $5,019–$22,568); $25,000 sits above its upper bound and roughly 28× Zimlichman's $896.

1. "**2.5% of admissions (NPIAP)**" **becomes 2.58% HAPI prevalence (IPUP).** NPIAP's fact sheet says

2.5 *million cases per year* and gives incidence of 2.1 per 1,000 (0.21%). The 2.5% figure is a transcription; the defensible source is VanGilder 2021's IPUP prevalence survey at 2.58%, which is a point prevalence rather than an incidence per admission, and the paper says so.

1. **The falls baseline was mis-attributed.** 3.5 per 1,000 patient-days is the NDNQI figure

(3.56) from Bouldin 2013, not an AHRQ patient-safety benchmark.

1. **None of Appendix B's four reduction-rate citations resolves as printed.** The CLABSI entry

("BMJ Qual Saf, 2020") resolves to a paper reporting a *non-significant* result (IRR 0.75, p=0.13). The 35% rate survives against Berenholtz 2014 (43%); the citation does not. Falls, CAUTI and pressure injuries are re-anchored or relabeled as bounded assumptions in §4.2.

1. **Reference 2's title word.** Synthea's published title reads "for *generating* synthetic

patients," not "creating."

1. **References 1, 5, 9 and 11 were not citable as printed** — no locator, version, DOI or

statute number. All four are resolved or explicitly labeled unresolved in the References.

1.3 Product and capability claims

1. **California's medical/surgical ratio is 1:5, not 1:4.** Title 22 §70217 sets 1:2 critical

care, 1:5 medical/surgical, 1:3 step-down and 1:4 telemetry. Revision 2 printed 1:4 for med-surg *and claimed it was verified against the authoritative text*, which is the worst place in the document to have been wrong. The repository's own registry had it right.

1. **Oregon and Washington are reframed as portability evidence.** Revision 2 cited "Oregon's

1:4 acute-care ratio effective 2026"; ORS §441.765 sets 1:5 for medical-surgical, with amendments operative 1 July 2026, and Washington's statute establishes staffing committees rather than fixed ratios. The rule engine genuinely carries both, which is evidence that it is state-portable. The commercial beachhead is California and Oklahoma; the earlier Oregon-and-Washington framing was residue from a superseded ideal-customer profile.

1. **Eagle: two of eight domains, and the service is named.** The eight-domain chassis is

`server/src/services/currentStateTwinModeler.js`; two domains are data-grounded, and `server/tests/currentStateTwinDataBacked.test.js:114` asserts `groundedCount` is 2. Revision 2's "four of its eight domains (up from two)" is withdrawn.

1. **Self-learning reachability is 22% (17 of 76), not 92% (57 of 62).** The figure was false on

Revision 2's own as-of date; the family was re-scoped on 2026-07-04 and surfacing has not caught up.

1. **Charlie does not ship a local speech binary.** The Whisper.cpp seam is real, loopback-

enforced and covered by 194 passing tests (counted 2026-09-03 in the Whisper seam suite; a per-seam figure with no derived fact behind it), and no binary ships: `whisperBinaryExists()` returns false and the route falls through to a stub that returns an empty transcript. Revision 2's "transcribes locally via Whisper.cpp on an edge device" is withdrawn.

1. **Charlie's SOAP, MAR and medication-name claims are withdrawn.** No SOAP composer exists on

the ambient path (the pipeline emits flowsheet rows); no ambient service writes to the medication administration log; and the medication-name refusal is syntactic rather than dictionary-backed — a live run accepted the invented drug name "Zzzblargin 10 mg po daily" at its maximum confidence of 0.8.

1. **Owl's readiness scoring is a built composer with an empty read side.** The demo database

holds zero `owl.posture.snapshot` Decision rows (verified 2026-09-02), so Revision 2's "with live survey-readiness scoring" becomes "composer built; no snapshot has been written."

1.4 Platform inventory

1. **Server lines of code: 1,059,721 excluding `_generated`** (4,042 files), against Revision

2's ~959,000 under an August stamp. The published figure was a late-July value.

1. **Client lines of code: 348,877 including client tests, 247,199 excluding them,** against

~302,000.

1. **Prisma models: 944,** against 943.
2. **Feature bundles: 1,347 total, 905 surfaced, 418 unsurfaced, 24 unwired, 27 families,**

against 1,343 / 902 / 417 / 24 / 27.

1. **Server route patterns: 6,455,** against "on the order of 6,250."
2. **Tests: 2,885 files, 72,122 passed, 0 failed, 17 skipped** on a local batched run of

2026-09-03 — never CI, and never "exercised on every push." The results artifact is regenerated on every local batched run, so the 2026-08-27 figures Revision 3 first drafted (2,871 / 71,867 / 2 / 17) are no longer reproducible by anyone, including this company. Revision 2's "64,088 passing tests exercised on every push" fails on both clauses and is withdrawn; §7 names every check the most recent push-triggered run failed on.

1. **Refusal codes: 3,651 distinct across 4,147 declarations** in 1,347 manifests, against

"3,600+ declared." The published phrasing read as the distinct count and understated it.

1. **Citations resolved: 4,489 of 8,754 (51%),** against 3,189 of 7,957. The published pair was

seven weeks stale on its own stamp and understated the platform.

1. **Owl's two denominators are separated.** 50 programs are modeled end-to-end against **583**

seed requirement elements; **611** is the criterion count of the 58-pack certifier tier. Revision 2 welded the two into one sentence.

1.5 Pricing, scaling and governance

1. **The price card is a model, not an exact figure.** Revision 2 called the platform cost

"exact and stated on the company's published price card." Every quotation of the card in this revision carries the qualification instead (§6).

1. **The Enterprise return was wrong twice, and the scaling table goes with it.** The published

7.5× was computed at $1,000/bed, which is neither the card's $1,500 ceiling nor the $1,150 midpoint of its $800–$1,500 band; recomputing it at the ceiling gives a different and smaller figure, which Revision 3 does not print, because it would be a return multiple derived from a per-bed value this revision withdraws. A second defect ran alongside it: the column headed "Return" divided *gross* value by cost while sitting beside a

net-value column. Both are moot in Revision 3, because the $7,532-per-bed figure every row multiplied is withdrawn with the total it came from, and §6 prints no return multiple.

1. **The 50%-discount sentence is deleted.** The pilot-cohort offer is a flat $95,000 all-in

annual package that becomes the facility's flat annual fee; a written invitation to negotiate against it contradicted the one term the business model treats as non-negotiable.

1. **The national ceiling is restated, split and re-sourced.** Revision 2's "roughly $550M"

became one blended number covering two differently priced products, and its bed count came from an American Hospital Association edition that has since been superseded. The corrected arithmetic and both editions' figures are in §6.

1. **The pre-registration gained a scope clause and an erratum.** Revision 2 cited a document

nobody had checked. Pre-registered claim 8 is narrower in code than in prose, and the document named a rights holder the founder retired on 2026-08-09. A dated erratum carrying both corrections is now appended to that document (commit `9419e2ca9`); §8.1 states the scope clause, and Appendix C audits all eight claims against the repository.

2. The control facility

The control facility is the committed fixture set at `data/demo/` — a hand-authored 274-bed community hospital that every engine in this paper is run against. It is one hospital, named once, and nothing else is mixed into it. The repository also carries a Prisma database seed (`server/prisma/seed.js`) describing a *different* hospital of 11 units and 217 beds; no figure in this paper comes from that seed, and the two are never combined.

Parameter	Value	Basis
Licensed beds	274	computed from `data/demo/units.json`
Clinical units	18	computed
Patients (census snapshot)	100	computed from `data/demo/patients.json`
Total staff	50	computed from `data/demo/employees.json`
Registered nurses	22, including 4 charge RNs	computed
High acuity (`acuityScore` ≥ 4)	56 of 100	computed at a stated threshold
ICU occupancy	22 patients in 20 beds = 110%	computed; an authored over-census
ED visits, Q1 2026	150	computed from `data/demo/ed-visits.json`
ED admission rate	44 of 150 = 29.3%	computed
Mean expected length of stay	20.97 days, n = 15, range 16.17–61.17	computed
Mean staff seniority	6 years	computed

The length-of-stay figure carries its own denominator for a reason. Only 15 of the 100 patients hold a projected discharge date, so 20.97 days is a mean over those 15 and not over the cohort; 98 of the 100 share a single admission date, which makes the fixture a point-in-time census snapshot rather than a longitudinal stay record. Any sentence quoting an average length of stay for this facility has to carry n = 15.

Unit	Kind	Beds	Census	Occupancy	Target HPPD
ICU	intensive-care	20	22	110%	22.5
Med-Surg 6N	med-surg	32	1	3.1%	6.8
Med-Surg 7N	med-surg	32	6	18.8%	6.8
ED	emergency	40	15	37.5%	4.2
NICU	neonatal-icu	24	10	41.7%	18
L&D	labor-delivery	12	0	0%	10
Postpartum	postpartum	18	8	44.4%	7.5
OR-1 … OR-6	or	1 each	4 of 6 occupied	—	6
PACU	pacu	12	1	8.3%	8
BH 3W	behavioral-health	18	0	0%	5.5
Geriatric 4S	geriatric	24	18	75%	7
Oncology 5N	oncology	28	11	39.3%	9.5
Dialysis	dialysis	8	4	50%	4

3. The machine

Revision 2 described a method and published numbers that no code produced. Revision 3 publishes a machine and the numbers that machine emits. It runs in four stages.

Intake. A delimited evidence inventory is read through a named intake profile, attested by a named human, and hashed. The control run accepted 105 rows from `evidence-inventory.csv` under attester `emp-049`, preview hash `df912990f9…`. Evidence may only be filed against a requirement that already exists: a malformed companion file was dry-run in the same pass and all three of its rows rejected with `OWL_INTAKE_016`, so the commit would have refused whole.

Gap assessment. Each program's requirement elements are classified against the attested evidence at a threaded as-of date, producing met, partial, not-met or unevaluable. A program whose evidence source cannot be read returns unevaluable rather than zero, and its elements are never counted as findings. Caller-supplied input is refused before any read if it carries structure, repeats itself, or names nothing the platform has registered — refusal code `PGA_009`, cited to OWASP A03:2021 and AHIMA data-quality management — so a caller's string never becomes a row, a cell boundary or a citation in a surveyor-facing document. An imported survey finding that cites no registered element is reported unmatched rather than attached to every element whose free text happens to contain the string; this control run produced no unmatched finding. Declared intake receipts are resolved inside one facility: `facilityId` is a required predicate on the lookup, and an absent receipt, a receipt belonging to another facility and a receipt whose preview hash differs all return the same sentence, so the refusal cannot be used to confirm that a row exists.

Roadmap. Every finding becomes one or more steps carrying an owner role, a dependency, minimum calendar days, an hours band and a labor cost band priced from BLS OEWS May 2025 national wages. Where an occupation has no published national wage, the hours are still reported and only the dollars are refused. Where a step is flagged as requiring capital, the labor half prices and the capital dollars are refused outright, because no public per-requirement capital benchmark exists.

Readiness. Per-program readiness is composed from evidence-coverage signals and rendered with its evidence floor. A domain with no freshness envelope rides out labeled unevaluated rather than as a verified zero.

Both runs are DB-free, seeded and byte-reproducible from a clean clone:

```
node server/scripts/run-digital-pilot.mjs --as-of 2026-09-19T00:00:00Z --seed 42 \
  --out docs/owl/pilot/control-facility-engines/run-2026-09-19
node server/scripts/run-owl-pilot-gap-assessment.mjs --facility SYN0274 \
  --as-of 2026-09-19T00:00:00Z --accept-partial \
  --out docs/owl/pilot/control-facility/run-2026-09-19
```

Each command has a `--verify` twin that byte-compares the committed artifact and exits 0.

Receipts. The Owl run wrote 154 Decision receipts through the platform's real `recordDecision` writer into a file-backed store rather than production Postgres: 105 `accreditation.recorded`, 29 `owl-designation-ledger`, 18 `owl-regulatory-lifecycle`, 1 `owl-intake` and 1 `owl-pilot-gap-assessment`. Forty-nine of them sit on hash-chained kinds and every one verifies end to end — 29 of 29, 18 of 18, and 1 of 1 on each of the two single-row kinds. The 105 evidence-capture rows carry a null content hash because `accreditation.recorded` is not a high-stakes kind, which is the platform's existing contract rather than a property of this run — the same scope limit §8 records against pre-registered claim 8.

The run hash binds the run's inputs and its observed outputs, not the bytes of the source that produced them; `codeVersion` is a label. Editing a fixture, a seed, an as-of or an emitted figure moves the hash, and editing an engine's source without changing any output does not.

4. Control-run results

4.1 What the engines computed

The September 2 detail that follows is retained as Revision 3's historical analysis. The current engines artifact is `docs/owl/pilot/control-facility-engines/run-2026-09-19/` (run hash `76eedcef9979374366c4b9658c5c7b3b88455a1a516fd7c44ffe0b56cfa8be7b`), verified at the source revision that publishes this edition.

All simulation figures below come from `docs/owl/pilot/control-facility-engines/run-2026-09-02/` (run hash `869a7401534e…`), seeded with mulberry32 at seed 42, 168-hour horizon, 20 replications, 95% intervals by normal approximation across replications. Two runs at the same as-of and seed produce byte-identical output.

Scope caveats, stated before any number is read. `hospitalFlowSim` simulates emergency, intensive-care and medical-surgical only — 40 + 20 + 64 = 124 of the facility's 274 beds — while echoing all 18 unit names back to the caller; passing 18 units returns metrics byte-identical to passing 3. The remaining 150 beds have no home in the model. The engine's `ED.beds` input is validated and then never used: emergency capacity is gated on providers and triage nurses. The metric the engine names `lwbs` is the queue backlog at the horizon rather than an abandonment count; nobody abandons in this model.

Three volume scenarios. D1 runs the fixture exactly as it ships, at 1.701 emergency arrivals per day. D2 scales only the arrival level, holding the hour-of-day shape, unit configuration, service times, horizon, replications and seed identical: 155,400,000 US emergency-department visits (CDC/NCHS, NHAMCS 2022) ÷ 907,216 US staffed beds (AHA Fast Facts 2026) = 171.2933 visits per staffed bed per year × 274 beds ÷ 365 = **128.5873 arrivals per day**, a scale factor of 75.594321. Falvo 2007 observed 139.1 emergency visits per bed per year at one 450-bed community teaching hospital, 19% below that intensity. D2a holds staffing at D1 and saturates by construction; D2b sizes emergency providers at 16, which is what the repository's own Erlang-C sizing returns for the benchmark rate at an 85% target service level.

Metric	D1 — fixture rate	D2a — saturated	D2b — Erlang-sized
Daily ED arrivals	1.701	128.5873	128.5873
Throughput, patients	9.25	316.55	826.8
Triage wait, min	0	15.21	14.86
ED-bed wait, min	none observed	3,096.15	18.85
Admission wait, min	none observed	none observed	577.34
Peak ED-bed queue	0	554.1	7
Backlog at horizon	0	554.45	10.45
ED provider utilization	0.0292	0.9943	0.6516
ICU utilization	0.0089	0.3211	0.7662
Med-surg utilization	0.0049	0.1367	0.337

Metric	D1 95% CI	D2b 95% CI
Throughput, patients	7.98–10.52	813.71–839.89
ED-bed wait, min	none observed	15.21–22.50
Admission wait, min	none observed	418.08–736.60
Peak ED-bed queue	0–0	5.72–8.28
ED provider utilization	0.0234–0.0351	0.6376–0.6656
ICU utilization	0.0041–0.0136	0.7251–0.8074

The shipped fixture cannot load a queueing model. At its own arrival rate, every queueing metric is zero or near zero: no patient ever waits for a bed, peak queues are 0, and emergency provider utilization is 2.92%. That is the honest reading of the fixture, reported as a result.

The fixture disagrees with itself about emergency volume by a factor of 105.8. An emergency census of 15 patients at a two-hour mean treatment time implies 180 arrivals per day by Little's law, against the 1.701 arrivals per day the visit log produces. The census is consistent with a real 274-bed hospital; the visit log is not. Any downstream model has to say which one it read.

Steady state against discrete-event simulation. `discreteEventTwin` cross-checks the emergency station. At D1 the two agree to four decimal places (largest hourly gap −0.0002). At D2 they diverge by up to 0.9607 in the probability of waiting at hour 21, then 0.8673 at hour 8 and 0.7854 at hour 15. Erlang-C treats each hour as an independent equilibrium; the simulation carries queue state across hours. The gap is the finding, and it is the reason the platform runs both.

Chokepoint detection has no lookback bound. Over a synthesized 88-day admission-discharge- transfer stream, `throughputAnalyzer` reports 90 critical chokepoints with elapsed times up to 126,704 minutes, for patients who left the emergency department in January. Bounded to 24 hours it reports 1 chokepoint from 4 events and 1 patient. Every detector in the service measures an *open* interval rather than a completed one, so it cannot structurally emit a statement of the form "median boarding time fell." Revision 2's "130 boarding hours cut" is not a quantity this engine can produce.

Staffing and ratios

Every ratio finding below rests on an assumed on-duty count. The fixture carries a roster and no shift schedule, so a unit's on-duty registered nurses are taken as the day-shift staff RNs assigned to it, with charge nurses excluded by default because §70217 counts a charge nurse only when carrying a direct patient-care assignment. No finding here observes a real staffing event.

Unit	Provision	Max pts/RN	Census	RN on duty	Required	Compliant
ICU	`ca-title22-ratio-icu`	2	22	2	11	no
ED	`ca-title22-ratio-ed`	4	15	1	4	no
Med-Surg 6N	`ca-title22-ratio-med-surg`	5	1	2	1	yes
Med-Surg 7N	`ca-title22-ratio-med-surg`	5	6	1	2	no

Three of the four evaluated units are out of ratio, and the registry returns a total penalty exposure of **$0 on every one of them** — because the California entry carries no `penaltyPerViolationUsd` field at all, not because the units comply. A reader who took that zero at face value would conclude the opposite of what the findings say. The telemetry tier (`ca-title22-ratio-tele` , 1:4) was not exercised: no fixture unit is designated telemetry, and that tier is the source of the 1:4 figure Revision 2 attached to medical-surgical. Fourteen of the 18 units — 150 of 274 beds — have no ratio tier in the registry at all, and the engine emits no refusal for them; the caller has to notice.

`ratioEngineValidation` scores the facility's hours-per-patient-day floors at **31.3 out of 100 with three violations**: registered nurses pass at 1.2571 against a 0.85 floor, licensed practical nurses are critical at 0.0 against 0.55 (the fixture has no LPNs at all, so the floor is missed by construction), nursing assistants are critical at 0.9714 against 2.2, and total nursing is major at 2.2285 against 3.6. The hours conversion and the assignment of all 17 fixture job titles to payroll-based-journal buckets are both stated mapping decisions rather than facts in the data.

The eight ratio violations Revision 2 priced remain an assumption. Nothing in the repository counts ratio violations over a period; the registry evaluates one unit at one moment. The *price* is now sourced: California Health & Safety Code §1280.3 as amended by SB 596 (2025, ch. 773, approved 2025-10-13) sets $15,000 for a first violation and $30,000 for each subsequent one, and directs that violations on separate days are separate violations. Eight events price at **$120,000** if every one is a first violation, or **$225,000** on the escalating schedule. The paper keeps $120,000 and states plainly that it is the conservative reading of a count nothing measures, and that a per-day statutory unit and an event count are not the same dimension.

The dollar engine

`edLwbsRoi` is the only service in the repository that emits a dollar return from queueing inputs, and it is a per-lever engine rather than a rollup. It embeds a stale wage: `edLwbsRoi.js:47` hard-codes $47.32 as "BLS OEWS May 2024" against the verified May 2025 figure of $48.76, so every added-staff cost it prints is 2.95% low. At the Erlang-C-sized staffing of 16 providers the baseline abandonment rate is already 0.93%, and the ladder is **negative at every rung** — −$307,376 for the first added provider, −$1,441,824 by the fourth. That is a real finding about the lever, and it is printed rather than dropped. The 4-provider ladder sits outside any operating regime (offered load 10.72 against 4 servers, 62.75% predicted abandonment) and its multi-million- dollar rungs are that catastrophic baseline being walked back; they are not a business case.

Cohort scoring

`losPredictor` scored all 100 patients: mean projected length of stay 8.271 days, range 1.6–39.2, `isCalibrated` false on every patient with a heuristic confidence source. Three of its inputs (comorbidity index, social complexity, prior admissions) are absent from the fixture and were left undefined rather than invented, and 42 of 100 admit units have no counterpart in the engine's multiplier table. Nothing in this family models a length-of-stay *reduction*.

`dischargeReadiness` returned "ready" for all 100 patients, and that is a null result printed as one: the fixture carries none of the 17 gate inputs the service reads (counted 2026-09-03 from the named destructure in `dischargeReadiness.js`; no predicate derives it), and the engine does not penalize unknown inputs by design, so every gate scores 100 by absence. It must never be quoted as a readiness finding.

`fallRisk` and `pressureInjury` were run on **assumed answers**. The fixture holds no Morse and no Braden assessment, so each instrument's items were derived from acuity, age band and a ventilator flag by a stated rule table. The Morse distribution (40 high, 56 low, 4 moderate, mean 29.1) and the Braden distribution (56 very-high, 31 high, 12 mild, 1 no-risk, mean 9.22) are outputs of that assumption table. A

different table gives different scores, and no number from either instrument is an observation of the control facility.

4.2 Never events, three steps and two answers

The method is unchanged from Revision 2: a published baseline rate applied to a stated denominator, reduced by a published reduction rate, multiplied by a published cost per event, with no rounding until display. What changed is that the denominators are now printed, and that the inputs were checked against the sources they were attributed to.

Denominators. 8,400 patient-days = 400 annualized patients × a 21-day length of stay, an average daily census near 23. 600 admissions = 150 first-quarter emergency visits × 4. 2,100 urinary-catheter days and 1,260 central-line days are the model bases as originally decoded. The safety category is sized on that annualized cohort snapshot and **not** on 274-bed occupancy, which would be 100,010 bed-days — roughly 11.9× larger. The cohort base is kept as the floor.

As the July 2026 pilot computed it, reproduced exactly:

Line	Denominator	Baseline	Expected	Cut	Prevented	$/event	Value
Falls	8,400 patient-days	3.5/1,000	29.4	40%	11.76	$15,000	$176,400.00
CAUTI	2,100 catheter-days	1.5/1,000	3.15	30%	0.945	$25,000	$23,625.00
CLABSI	1,260 line-days	1.0/1,000	1.26	35%	0.441	$45,000	$19,845.00
Pressure injury	600 admissions	2.5%	15	25%	3.75	$12,000	$45,000.00
Total							**$264,870.00**

Corrected against the resolved sources:

Line	Denominator	Baseline	Expected	Cut	Prevented	$/event	Value
Falls	8,400 patient-days	0.93/1,000	7.812	34%	2.65608	$15,000	$39,841.20
CAUTI	2,100 catheter-days	0.74/1,000	1.554	30%	0.4662	$13,793	$6,430.30
CLABSI	1,260 line-days	1.0/1,000	1.26	35%	0.441	$45,000	$19,845.00
Pressure injury	600 admissions	2.58%	15.48	25%	3.87	$12,000	$46,440.00
Total							**$112,556.50**

At the July baseline of 1.5 per 1,000 catheter-days with the corrected AHRQ cost, CAUTI prevents 0.945 events worth $13,034.39 and the category totals $119,160.59. Both sensitivities are printed because the CAUTI baseline is the one input in this table that could not be resolved to any NHSN publication; CDC's 2023 acute-care reporting implies a national pooled rate near 0.74.

The event column is where a reviewer's arithmetic breaks against Revision 2. The dollars were always computed on unrounded expected values while the counts were rounded for display, so dividing $176,400

by "12 falls" returns $14,700 against a stated $15,000 basis. We print the unrounded expected values instead. Two of them — CAUTI at 0.945 and CLABSI at 0.441 — are below one event per year, and the fractional-event convention now covers both rather than only CLABSI. Falls, CAUTI and pressure injuries all round up in Revision 2's display and CLABSI rounds down, which is the inconsistency that produced the appearance of an error.

4.3 What stays an assumption

Four of the July model's value lines have no computational counterpart anywhere in the repository — CAUTI events prevented, CLABSI events prevented, ambulance diversions avoided, and the ambient-documentation hours returned. Three more are measured by nothing that models a reduction: the 0.8-day length-of-stay cut, the 130 boarding hours, and the agency and overtime shift counts. Appendix B is the register — thirteen lines, each labeled and sourced or explicitly unsourced. Nothing in it is presented anywhere in this paper as an engine output.

5. The Owl control run — gap assessment, roadmap, and cost to close

The current September 19 control run is `docs/owl/pilot/control-facility/run-2026-09-19/` (run hash `27bee20243095644903318d7af5f0b9ea3846cbc6e2f6c86d6034c13b76a77cd`). It reports 45 findings, 51 roadmap steps, a 28-day critical path, labor cost of $6,285 … $10,857 … $19,565, and facility total of $31,285 … $35,857 … $44,565. The detailed September 2 table below is retained as the Revision 3 historical record, not as the current rerun summary.

Owl's bottom line is the one an executive already understands: helping hospitals pass inspections by attaining continuous survey readiness. This section is what that produces on the control facility. Every figure comes from `docs/owl/pilot/control-facility/run-2026-09-02/` (run hash `7a0ac762fd3b…`, as-of 2026-09-02, provenance mode `control`, in-memory fixture store).

Metric	Value	Basis
Programs in scope	9	counted programs
Programs unevaluable	1	counted `PGA_003` refusals
Requirement elements	110 evaluable (+6 unevaluable = 116 in scope)	counted seed elements
Findings, not met or partial	44 — 22 not met + 22 partial	counted elements
Labor hours, point	228.26 h, before suppression	modeled assumption
Labor hours reported but not priced	6.80 h	modeled assumption
Labor cost band	$6,247 … $10,792 … $19,448	BLS OEWS wage × stated hours band
External cost, remediation lane	$25,000	public procurement
Facility total	$31,247 … $35,792 … $44,448	modeled assumption
Findings refused a price	1	counted refusals
Roadmap steps	50	counted steps
Critical path	28 days	stated planning convention

Those bands are summed independently across lines. They are **not** confidence intervals: there is no correlation model and no Monte Carlo behind them, and the top of every dollar band comes from the hours convention rather than from wage data, because published wages give a median and a mean and no upper bound. The hours figure is the whole facility before small-cell suppression, and it already includes the 6.80 hours whose dollars were refused rather than sitting above them.

Forty-four is the count of elements that are not met **or** partial. The run was executed with `--accept-partial`, which is what makes a partial element a finding with a remediation step attached; the scoreboard's Not-met column carries the 22 that are outright not met. A reader who adds that column and expects the headline finds half of it, which is why both halves are printed here.

The facility-wide readiness band is at-risk. That band is the artifact's own §1 headline and it rests on the evidence-coverage signals tabulated below. The run publishes no facility-wide score beside it, because scoring in this machine is per program; the score column below carries the eight programs that could be scored at all.

Program	Readiness	Evaluable	Met	Partial	Not met	Unevaluable	Evidence floor
joint-commission	76.0%	measured	14	5	6	0	below floor (20)
cms-deemed	88.0%	measured	18	4	3	0	at or above floor (12)
state-doh	not evaluable	unknown	—	—	—	6	at or above floor (5)
magnet	77.8%	measured	5	2	2	0	below floor (9)
aacn-beacon	83.3%	measured	4	1	1	0	below floor (6)
ena-lantern	80.0%	measured	3	1	1	0	below floor (5)
acs-trauma-level-2	87.5%	measured	10	4	2	0	at or above floor (14)
primary-stroke-center	76.5%	measured	9	4	4	0	at or above floor (12)
acep-geda-level-3	57.1%	measured	3	1	3	0	below floor (6)

Program	Score band	Score	Days to survey
joint-commission	at-risk	12	120
cms-deemed	at-risk	56	400
state-doh	not evaluable	—	210
magnet	at-risk	66	900
aacn-beacon	watch	77	640
ena-lantern	at-risk	74	700
acs-trauma-level-2	at-risk	72	560
primary-stroke-center	at-risk	53	300
acep-geda-level-3	at-risk	43	480

Two band vocabularies sit across those two tables, and the machine keeps them apart. *Evaluable* is `accreditationWorkflow.composeReadinessStatus` answering whether a denominator could be read at all — `measured` or `unknown`. *Score band* and *Score* are `owlContinuousReadinessScore`'s `ready` / `watch` / `at-risk` over a 0–100 composite of a program's own domain signals: evidence coverage less 6 points per open finding, 4 per gap and 2 per stale artifact, clamped. It is a stated scoring convention rather than a measurement, and the coverage input is never a mock-survey pass rate. A program can be `measured` and still `at-risk`, and an unevaluable program has no score at all. *Days to survey* is a counted date difference against the run's as-of.

Five of the nine programs sit below their evidence floor, which means their readiness percentage rests on too little evidence to carry weight even where the percentage looks healthy. `state-doh` returns *not evaluable* rather than 0%: its evidence source was made unreadable as a declared fault injection, and its six elements are unevaluable rather than not-met, because the denominator could not be read. A program whose evidence cannot be read never renders as a confident zero.

One defect rides on that same row, and this paper prints it rather than the reassuring reading. `state-doh` shows *at or above floor (5)* in the evidence-floor column, because the unevaluable branch leaves the below-floor flag false. The column is reporting a floor cleared over a numerator the run never read. It is open, and §9 carries it.

When the surveys actually land. The same run carries a survey schedule for all nine designations, composed by `owlRegulatoryLifecycle` from the facility's own registered renewal receipts, never from the attestation counts above. A designation whose lifecycle was never registered prints that in every cell of its row — never a sibling's dates, a catalogue default or a guess.

Designation	Cycle	Last survey	Next due	Days remaining	Schedule band
Joint Commission	730 d	2024-12-31	2026-12-31	120	current
CMS Medicare Certification	1,095 d	2024-10-07	2027-10-07	400	current
State Department of Health	365 d	2026-03-31	2027-03-31	210	current
ANCC Magnet	1,461 d	2025-02-18	2029-02-18	900	current
AACN Beacon Award	1,096 d	2025-06-03	2028-06-03	640	current
ENA Lantern Award	1,096 d	2025-08-02	2028-08-02	700	current
ACS Trauma Center Level II	1,096 d	2025-03-15	2028-03-15	560	current
Primary Stroke Center	730 d	2025-06-29	2027-06-29	300	current
ACEP GEDA Level 3 (Bronze)	1,095 d	2024-12-26	2027-12-26	480	current

All nine rows band `current`, and the nearest survey is 120 days out, so nothing on this control facility sits inside a renewal window. `Days remaining` is the same counted difference the scoreboard above prints as `Days to survey`. `Schedule band` is a third vocabulary, separate from both columns above it: it tracks the calendar rather than the evidence. The cycle is a rolling cadence counted in days, so a due date reproduces to the day instead of drifting on a 30-day-month approximation. Every cadence here is an in-repo catalogue constant for a synthetic facility, and a real hospital is governed by its designating body's own published cycle. The owner role on all nine rows is the same seeded quality coordinator.

`designationLedger.revenue` is `null` for this run, and the ledger says why in its own words: `OWL_LEDGER_004` — revenue-at-risk requires an aggregate payer mix, and absent one the ledger reports readiness without dollars rather than fabricating a figure. This paper prints no revenue-at-risk figure for the same reason.

Every finding is submittable as printed. Each of the 44 finding cards carries the seven plan-of-correction required elements read from `owlPlanOfCorrection.REQUIRED_ELEMENTS` — deficiency, standard cited, root cause, corrective actions, responsible role, due date, measure of success — with the two human-judgment fields left explicitly unauthored rather than filled in by a machine. The responsible party is a role, never a named individual, and the underlying root-cause analysis stays peer-review privileged.

A specimen, printed in full in the artifact: joint-commission NPG.05.03.01, hand hygiene, status not-met, 0 evidence artifacts on file against a 3-month cadence, owner role nursing-education, severity high, scope large, matched to imported survey finding `sf-0002`, 17.46 … 29.10 … 52.38 labor hours at $819 … $1,418 … $2,554, 28 minimum calendar days, target date 2026-10-02.

The roadmap. Fifty steps — 44 evidence-update steps and 6 staff-retraining steps, each retraining step dependent on its own evidence step. Forty-five steps carry a 7-day minimum and five carry 14 days; the 28-day critical path is the topological longest path over those dependencies. Target dates cluster at 2026-10-02 (8 steps), 2026-11-01 (20) and 2026-12-01 (22).

Program	Steps	Labor cost (low … point … high)
joint-commission	15	$2,742 … $4,743 … $8,543
cms-deemed	8	$724 … $1,249 … $2,251
primary-stroke-center	8	$859 … $1,484 … $2,676
acs-trauma-level-2	6	$302 … $518 … $934
acep-geda-level-3	5	$872 … $1,507 … $2,716
magnet	4	$337 … $581 … $1,048
aacn-beacon	2	$224 … $387 … $698
ena-lantern	2	$187 … $323 … $582

Severity	Findings	Hours (point)
high	4	79.35
medium	18	102.45
low	22	46.46

The by-role rollup is mostly suppressed, and that is the correct behavior rather than a gap in the data. Under the CMS cell-suppression policy and 45 CFR 164.514(b), a rollup whose denominator sits below the small-cell floor of 11 is withheld with its reason printed. Suppression binds the machine-readable dossier and the run's own Decision row as well as the rendered report, so a withheld role's hours are absent from every consumer rather than blacked out in one of them. Quality-coordinator and nursing-education (n=0) publish no findings count, no hours and no dollars anywhere in this run; the control roster carries no nurse-educator job profile at all, so that headcount is genuinely zero. Registered nurses clear the floor and publish: 22 findings, 55.00 hours, $2,682. The emergency-department medical director row is withheld for

a different reason — the headcount was never supplied, so the denominator is unknown and the cell is withheld rather than published over an unproven n.

The report's own footnote counts three suppressed cells: two withheld by the assessment before the report was rendered, one by the report's pass over the rollups above. Two things follow, and both are printed here rather than smoothed over. The 228.26-hour headline is a pre-suppression total, so it still counts the role cells the rollup withholds — which is why the by-role rows do not add up to it, and why this paper prints no by-role hour arithmetic. And `dossier.json` reports one suppressed cell where the report reports three, so a consumer reading the machine-readable file alone undercounts the suppression. That defect is open too; §9 lists both.

Roles are mapped to published occupations, and the mapping is printed so a reviewer can disagree with it. Quality-coordinator, nursing-education, clinical nurse leader and registered nurse price at SOC 29-1141 (median $46.90, mean $48.76); chief nursing officer, leader and manager at 11-9111 ($59.55 / $67.77); compliance officer, operations-compliance and policy owner at 13-1041 ($38.81 / $42.50); clinical informatics at 29-9021 ($32.70 / $36.04). Four roles are refused a price rather than substituted: chief medical officer, emergency-department medical director, pharmacy director and privacy officer have no row in the fetched national wage table, and pricing a physician as a health-services manager would be an unstated substitution that materially understates the wage. Their hours ride out in full. The quality-coordinator row moves the largest cost line in this run and is mapped to a nurse occupation because the control facility's own quality coordinator holds registered-nurse licensure; a reviewer who disagrees can re-price that single row.

External anchors sit in three lanes, and the lane type is what prevents double counting.

Anchor	Lane	Dollar year	Amount
mock-survey-single-site	remediation (summed)	2009–2011	$25,000
standards-manual-annual	displaced incumbent (never summed)	2023–2026	$67,000
mock-survey-scope-pair	context benchmark	2009–2011	$25,000 … $151,000
mock-survey-network-wide	context benchmark	2009–2011	$151,000
readiness-consulting-engagement	context benchmark	2008–2015	$177,000 … $285,000
survey-tracking-cockpit-annual	context benchmark	2017–2021	$234,000
trauma-level-2-readiness-annual	context benchmark	2013 survey	$2,330,000

Only the single-site mock survey is ever a cost line, capped at one per facility per run. The standards-manual subscription is what the platform displaces, and summing a tool a facility already pays for into the price of fixing a finding would be double counting. The remaining anchors are denominators printed for scale. Two of them have no single point value because their two endpoints describe different scopes, and averaging them would invent a number. Every one is a published federal purchase-order price in the dollar year shown, and none is a quote for this facility.

Six refusals. A refusal here means *the machine cannot evaluate this*, never *the facility is non-compliant*.

#	Code	Where	What the machine declined
1	`PGA_003`	program `state-doh`	reading an evidence store it could not open
2	`EST_008`	step `p1-e12-step-2`	pricing per-learner delivery with no headcount
3	`EST_004`	step `p2-e23-step-1`	pricing an occupation with no published wage
4	`EST_003`	step `p9-e07-step-1`	modeling capital with no public benchmark
5	`PGA_008`	designation roster	resolving an unknown program key
6	`OWL_INTAKE_016`	intake dry run	filing evidence against unregistered requirements

Each refusal names the authority it cites and what would unblock it; both columns are in Appendix A. A supplied head count of zero is an evaluable zero rather than a refusal, which is the distinction the second code exists to preserve. The fifth refusal is a pass-through: the designation ledger's own `OWL_LEDGER_005` is carried verbatim rather than re-derived, and it is never softened into "no designations found."

5.1 What Owl is worth — a risk-and-substitution disclosure

Owl does not get a four-category value table in this paper, and the reason is arithmetic. Built honestly from sourced substitution credits alone it does not clear its own price.

Substitution credit (shown, not added). The defensible incumbent basket for a single 274-bed community hospital is the standards-manual subscription at $67,000/year plus one mock survey at the $25,000 floor amortized over a three-year accreditation cycle, or $75,333/year. Crediting half of that gives **$37,667/year** against a $60,000/year software-only Owl price — a ratio of 0.63. The credit excludes the $234,000/year survey-tracking cockpit and the $177,000–$285,000 consulting engagement outright, because both are enterprise-scope federal obligations rather than single-facility prices, and crediting either to one hospital would be the largest single overstatement available. It uses the floor of the mock-survey range, it does not inflation-adjust 2009–2011 dollars, and it credits only half the remaining basket because a hospital may rationally keep buying a human mock survey alongside the software. **A hospital that buys none of this basket today saves none of it, and this line is zero for that hospital.**

Registrar abstraction time (a measurement instrument, not a dollar line). Evans and Samanta 2020 give a regression intercept of 38.95 minutes per trauma-registry chart; at 1,200 cases a year that is 779 registrar-hours, which independently corroborates the 800–1,000-hour estimate in the company's own research. The intercept is the predicted time with every acuity covariate at zero and every coefficient in the model positive, so it understates the burden at any real case mix. No published reduction rate exists for governed-AI registry abstraction, so we apply none. At 10% to 30% displacement and $40–$50 per hour the band runs $3,116 to $11,685 — under 0.6% of the July pilot's total. That is a credibility line rather than a dollar line, and a facility holding no trauma designation scores zero on it.

Designation revenue at risk (an exposure, not a value). MS-DRG 023 pays $41,698 per case and DRG 024 $28,466, against a median comprehensive-stroke-center volume of 76 cases a year — about $2–3M of annual revenue tied to a designation, on a mix the underlying research does not state. Converting exposure

into value requires a probability of designation loss, and no study since 2015 isolates revenue lost specifically from losing a trauma or stroke designation, so we apply no probability to it.

The readiness denominator (scale, never a credit). Ashley 2017 measured Georgia trauma centers' annual readiness cost at an average of $2,333,113 for Level II centers (n = 9, 2013 survey dollars, stratified by trauma-center level rather than by bed count). Most of that total is clinical medical-staff standby and a held-open operating room — capacity costs software cannot reduce by a dollar — and the four-category split needed to isolate the addressable share is not in the abstract, so we credit no fraction of it. It appears as what it is: a five-figure subscription set against a seven-figure annual readiness investment.

FOUNDER REVIEW. The honest summary of the four disclosures above is that *a substitution-only ROI does not clear the Owl subscription price, and the purchase is a risk purchase rather than a savings purchase.* The commercial case rests on the exposure lines and on the readiness-spend denominator, which is the argument the company's public material already makes.

6. Two purchase scenarios

Nightingale prices on a hybrid model, so this pilot carries two cost lines rather than one. They price different products and are not alternatives to the same deployment.

Scenario A — Owl-first, zero integration

What is deployed	Owl on delimited exports the quality office already produces
Integration required	none — no EHR feed, no bedside capture
Price	$95,000/yr all-in, or $60,000/yr software-only
Priced against	the facility's own annual readiness spend and incumbent basket
Budget share	$95,000 = 4.1% of a $2,330,000 readiness spend; $60,000 = 2.6%
Against the incumbent cockpit	$234,000 − $95,000 = **$139,000/yr less**, portfolio-wide
Modeled annual value	not modeled — see §5.1
Payback	not computed, because no value model is claimed

The $95,000 figure is an all-in annual package covering implementation, year one and live survey support, and it becomes the facility's flat annual fee rather than a first-year rate that renegotiates. Both prices sit under the $100,000 board-approval line so a quality office can buy directly. The readiness-spend denominator is in scope for *this* facility specifically: the control run's designation roster includes ACS Trauma Level II at 87.5% readiness, which is the center type Ashley 2017 measured. A facility holding no such designation carries a different denominator, and the budget-share line has to be recomputed against it rather than transferred.

Scenario A states no return multiple. We have not measured how much readiness labor Owl displaces, and we will not estimate it before a design partner does. What we can state is a price against a spend the facility

already carries.

Scenario B — the integrated platform

What is deployed	live HL7/FHIR feeds, bedside capture, ambient documentation, analytics
Price	$164,400/yr — 274 staffed beds × $600/bed/yr
Tier	Standard ($400–$800/bed/yr) on a $200–$1,500 card, modeled at the midpoint
Modeled annual value	**withdrawn** — see below
Net, return, payback	not computed in this revision

The price card is a model, not an exact figure: the per-bed tiers recalibrate at the first design partner, and Revision 2's word "exact" is withdrawn. The 274 beds are licensed beds priced as though all were staffed, which overstates the cost line and understates any return. Facilities under 50 beds and over 500 beds are quoted a negotiated per-facility flat rate rather than the per-bed card, so any extrapolation to those sizes is a card arithmetic exercise rather than an offered price.

Scenario B carries no total annual value in this revision. Of the thirteen lines in Appendix B's register, four have no computational counterpart in the repository and three more are measured by nothing that models a reduction (§4.3). One category was recomputed end to end by the engines, and it fell from $264,870.00 to $112,556.50. Restating a headline total by subtracting that one correction from a rollup whose other nine lines are unsourced assumptions would repeat the error this revision exists to correct. We withdraw the July figures — $2,063,632 of annual value, $1,899,232 net, 1,155% first-year return, roughly 29 days to payback — and return the integrated scenario to a price line plus an assumption register until a design partner supplies measured inputs.

Revenue ceilings. Price multiplied by facility count is arithmetic that survives the withdrawal of the value model, because it depends on no reduction rate. At full penetration the integrated tier reaches roughly **$548M** in annual recurring revenue (913,136 staffed beds × $600/bed/yr blended = $547,881,600). Owl's per-facility line is separate and additive: **$366M** at 6,093 hospitals × $60,000/yr, or **$579M** at $95,000/yr. Both are ceilings at full penetration rather than forecasts, and the blend assumption matters — a single $600 blend across every US staffed bed prices the large majority of hospitals above the Base tier their bed count would place them in. The bed and hospital counts are the American Hospital Association's 2025 PDF edition; its canonical page now serves the 2026 edition at 907,216 staffed beds across 6,100 hospitals, and the engines run in §4.1 uses the 2026 figures.

7. Engineering inventory

The value projections were estimates and this revision withdraws most of them. The engineering counted below is not an estimate: every row is a command a reader can run against the repository, and the exact commands are Appendix D, keyed by the reference in the last column. Every row was re-run at `888975d8c`. That commit carries the Owl gap-assessment dossier exactly as printed here (run hash `7a0ac762fd3b…`), so the table and that artifact describe the same tree. **The engines artifact is one commit later:** at `888975d8c`

it hashes `4890ee933d91…`, and the `869a7401534e…` printed in §3, §4 and Appendix A first exists at `bd0aa24cd` (2026-09-03 15:56, *fix(ci): make the Owl pilot engines artifact environment-independent*), which is the only commit carrying both hashes as printed. Verified 2026-09-15 by reading both artifacts at both commits. The same commands on `main` return smaller figures in five rows, by the bundle and the routes this branch adds.

Metric	Value	As of	Cmd
Prisma data models	944	`888975d8c`	D1
Feature bundles	1,347	`888975d8c` catalogue	D2
Bundles surfaced to a clinician	905	`888975d8c` catalogue	D2
Bundles unsurfaced	418	`888975d8c` catalogue	D2
Bundles unwired	24	`888975d8c` catalogue	D2
Catalogue families	27	`888975d8c` catalogue	D3
Server route patterns	6,455	`888975d8c`	D4
Server lines of code	1,059,721 in 4,042 files, excl. `_generated`	`888975d8c`	D5
Client lines of code	348,877 incl. tests; 247,199 excl.	`888975d8c`	D6
Server test files	2,885	2026-09-03 local run	D7
Tests passed / failed / skipped	72,122 / 0 / 17	2026-09-03 local run	D7
Refusal codes	3,651 distinct across 4,147 declarations	`888975d8c`	D8
Citations resolved to a URL	4,489 of 8,754 (51%)	`888975d8c` catalogue	D2
Owl programs catalogued	148	`888975d8c`	D9
Owl programs modeled end to end	50	`888975d8c`	D10
Owl seed requirement elements	583	`888975d8c`	D11
Owl certifier packs / pack criteria	58 / 611	`888975d8c`	D12
Eagle domains data-grounded	2 of 8 (`currentStateTwinModeler`)	`888975d8c`	D13
Self-learning reachability	17 of 76 (22%)	`888975d8c` catalogue	D14
Commits, all time	4,314 since 2026-05-28	`888975d8c`	D15
Commits since 2026-07-01	3,650 (85%)	`888975d8c`	D15

The test figures come from a local batched run, recorded in a git-ignored results artifact that the next local run overwrites. It is the best source available, a third party cannot audit it, and the first party cannot re-derive an earlier run either: the 2026-08-27 figures this revision first drafted — 2,871 files, 71,867 passed, 2 failed — are gone, and the row above is the dated snapshot that replaced them. It is not continuous integration, and the platform makes no claim that its tests pass on every push. The most recent push-

triggered run on `main` (`33729948364`, 2026-09-03) failed twelve checks. Nine are steps of the lint-gates job: the deprecation-manifest overdue check, the ESLint silent-catch and inline-role-predicate gates, the Nuthatch engine suite, the Nuthatch README fresh-default truth gate, the Nuthatch sync scrub-gate and lockstep-drift check, the Owl citation gate, the replat freeze gate, the success-without-persistence gate and the time-bomb probe. Three more are whole jobs — client tests, the Semgrep SAST scan, and server tests, whose three shards were cancelled. Three of those nine steps are the three red Nuthatch gates §8.2 names.

The branch that carries this paper's two artifacts also carries a CI step that re-runs both commands in `--verify` mode and byte-compares the committed files. That step has never executed: no CI run has ever been triggered on this branch.

Two counting conventions are stated because they have been conflated before. The 148 catalogued programs and the 158 keys in the classification table (counted 2026-09-03 from `PROGRAM_CLASSIFICATION`) are different populations — ten legacy or alias keys the classifier still answers for are not in the roster — so any external count has to say which it means. And the 50 programs modeled end to end carry **583** seed requirement elements, while **611** is the criterion count of the separate 58-pack certifier tier; the two cannot appear in one sentence.

Built since the 8 May 2026 pre-registration: HL7 v2 filing appliers (ORU lab results numeric and text-coded, plus MDM, medication administration and ADT), a FHIR R4 Subscriptions implementation, real-time clinical alerting over server-sent events, risk-adjusted observed-to-expected outcome analytics, and the Owl accreditation-evidence layer this paper exercises. Owl's continuous readiness composer is built and pure; its read side is empty, because no posture snapshot has ever been written to the demonstration database.

The Joint Commission corpus is the one behind. Twenty-five crosswalk members remain proposed and un-attestable and fifteen pre-2026 keys are held on old numbering, so any Joint Commission standard number in this paper's Owl output may be a superseded identifier — and that is precisely the certifier a buyer tests first.

8. Governance

8.1 The pre-registration, audited against current code

The platform and its architectural claims were pre-registered on 8 May 2026, committing to eight measurable, falsifiable claims. Revision 2 cited that document and no one had checked it. It has now been audited claim by claim against the repository at commit `6afb979ee`; the full table with file-and-line evidence is Appendix C. Seven of the eight name a mechanism that exists in code at the cited location, with two qualifications worth stating in the body of the paper.

Claim 1's timing half is untested. A closed clinical event does build a FHIR R4 transaction bundle and does refuse to post unless the event is resolved. No live EHR has ever accepted one, because there is no live deployment, so the "within seconds" half of the claim has never been measured. The mechanism is built; the timing claim is untested.

Claim 8 is narrower in code than in prose, and needs a scope clause. It reads that the hash-chained audit log is verifiable end to end. In code the Decision chain covers high-stakes kinds only: both chain

columns are nullable, non-high-stakes rows skip the chain by design, and authentication and system events live on a separate chain. The defensible statement is that the Decision log is hash-chained for high-stakes kinds and verifiable end to end *within that scope*. The control run in §3 is a worked example of exactly this — 49 of 154 receipts are on chained kinds and all 49 verify; the other 105 carry a null content hash by contract.

The erratum the pre-registration needed is appended. `PREREGISTRATION.md` named Regulation Loop LLC as rights holder; that entity was retired on 2026-08-09 in favor of Nightingale OS Inc., with no d/b/a, so a diligence reader following the citation would have landed on a dead company name. A dated erratum (2026-09-03, commit `9419e2ca9`) now sits at the top of that document and carries both the entity correction and claim 8's scope clause; the 8 May body is unchanged. What remains outstanding is Reference 1's stable locator, a commit hash or a public URL, before the claims audited here can be checked by anyone outside the company.

8.2 Charlie Nuthatch — a build disclosure

Charlie Nuthatch governs how Nightingale's own software is built, not how a hospital runs. The platform is largely written by AI coding agents, and Nuthatch is a witness installed at the boundary where those agents ask a runtime to write a file or run a command. Every such request is recorded before it happens as a tamper-evident receipt in an append-only, hash-chained ledger. A receipt records the shape of a request and never its contents: proposed file content becomes a SHA-256 digest and a byte count, and a shell command becomes a structural sketch in which only words from a pre-registered vocabulary appear literally and everything else becomes an integer-only placeholder. That is why the ledger can be shown to an auditor without becoming a new disclosure surface. Nuthatch runs nowhere inside a hospital, sees no patient data, and makes no clinical decision.

Disclosure	Value
Posture band	observing
Enforcement mode declared	ask
Build receipts recorded	155,968 — 30,649 live plus 125,319 archived; last sequence 155,967
Signed transparency-log heads	2,393
Overflow rows outside the live ledger	3,204
Controls in the pack	26, of which 26 declare `ask` and 0 block
Receipt-chain state	not verified in this run
Control pack signature	not verified in this run

Those are file counts read from the vendor's own build ledger on 2026-09-03, not a composer verdict. The governance composer does not return within ten minutes against a ledger this size (bead `nightingale-oi7os`), so the control artifact carries a lead-verified snapshot with `composerRun: false` and that bead printed beside it. No linkage claim is made: the chain was not walked for this artifact.

The harness is not in force, and three independent reasons stack. A hook-side process cannot prove which other hooks the host resolved for the same tool call, so composition is unverified and every matcher reports *observing — unverified*. The control pack's preflight binding no longer matches the loaded pack. A competing handler claims the same shell-command surface. Three of six Nuthatch continuous-integration gates are red, including its own documentation truth gate, so this paper quotes nothing from that documentation while that gate is red.

The wording discipline is the point. When the disclosure says a control *would have* escalated or refused, it means **observed**, never **prevented**. We do not describe it as blocking unsafe AI actions, because no control declares a blocking verdict and the pack does not currently load. We do not describe it as securing the repository: its own self-protection document calls it a flight recorder with a brake rather than a security boundary against a process running with the same identity. The patent is **filed** — USPTO application 64/139,773, received by the USPTO 2026-08-24 (Electronic Acknowledgement Receipt, in the company record); nothing further is claimed for it.

9. Limitations and honest disclosures

Stamped 2026-09-02, the as-of date of both control runs.

1. **The hospital is synthetic.** No real facility was assessed, and no readiness score in this

paper has been compared against an actual accreditation decision — by this company or by anyone. No correlation between the two has been measured.

1. **There is still no live hospital.** HL7 v2 and FHIR R4 paths are built and tested against

fixtures; no production feed has flowed, and vendor adapters remain sandbox-bound.

1. **Hours are modeled, not measured.** Exactly one effort line in the Owl run has a published

measurement behind it — 0.65 hours per evidence artifact, from Evans and Samanta 2020 — and one more has a regulatory count with a declared rate. Everything else is a planning assumption, and the plausible range is a planning convention rather than a confidence interval.

1. **The dollar bands are asymmetric in provenance.** Published wage data supplies a median and a

mean; the top of every band comes from the hours convention, not from wage evidence.

1. **The external anchors are federal procurement obligations in their original dollar years**

(2008–2015, 2009–2011, 2017–2021, 2023–2026), not quotes for this facility and not inflation-adjusted. Only the single-site mock survey is ever a cost line.

1. **The wage table is national, wage-only, and incomplete.** BLS OEWS May 2025 means and medians

carry no benefits or overhead multiplier, and every physician and pharmacist role in the roadmap is refused a price rather than substituted.

1. **The simulation covers 124 of 274 beds.** Three unit types enter the flow model; the other

fifteen units have no home in it, and the engine reports no refusal for them.

1. **The synthetic generator's operational layer is unvalidated.** Seven distributional checks

against a held-out 30% Synthea split (n = 3,426 patients, 201,738 encounters) all pass: six score a Kullback-Leibler divergence at or below 0.005 in $\log_2$, and medication frequency — the hardest vocabulary-mismatch case — scores 0.031. Synthea validates the clinical layer only. Four nursing-operational tables (triage, nursing assessments, safety events, bed assignments) have no open reference corpus and are unvalidated. That validation is a point-in-time artifact from 2026-06-18, reproduced on 2026-09-02, and it is not a standing gate on any build.

1. **MIMIC-IV credentialing has not started.** PhysioNet credentialing is tracked as bead

`nightingale-aqp` and is not complete, so patient-level validation of the synthetic layer is future work. A clean clone cannot reproduce the calibrated path, because the reference cohort is not re-downloadable.

1. **The policy lane was not evaluated.** No policy rows are seeded for the control facility. The

lane was called and returned nothing, which is an unknown rather than a finding of zero policy gaps.

1. **Six fault injections are deliberate and declared,** so a reader never mistakes any of them

for a real facility condition. Five shape the assessment: the `state-doh` evidence store made unreadable (`PGA_003`); the `CMS-482.55` owner role overridden to `ed-medical-director`, a physician with no fetched wage (`EST_004`); `GEDA-BRONZE-EQUIPMENT` flagged as requiring capital (`EST_003`); the restraint-application competency seeded with zero completion rows (`EST_008`); and roster entry 10, `regional-burn-center-verification`, a program key the designation ledger does not know (`PGA_008`). The sixth never reaches the assessment: the malformed import file dry-run in §3, whose three rows all reject with `OWL_INTAKE_016` and which is never committed.

1. **The readiness score is a prioritization aid, not a guarantee of survey outcome,** and its

input is evidence coverage rather than a mock-survey pass rate. No mock survey was run.

1. **Nuthatch is a build disclosure rather than a runtime control,** and it says observed, never

prevented.

1. **Time to value.** Any payback figure assumes immediate full deployment. A real rollout phases

over six to twelve months.

Two defects in the shipped artifacts are open. `dossier.json` reports one suppressed cell where `report.md` reports three, so a machine consumer reading the dossier alone undercounts the suppression (§5). And `state-doh` keeps its below-floor flag false on the unevaluable branch, so its evidence-floor column reads *at or above floor (5)* over a numerator the run never read (§5). Both are filed. Neither changes a figure in this paper; both would change what a reader concludes from the artifact without this paragraph. They are beads `nightingale-19n6z` (the dossier suppression count) and `nightingale-sr0j7` (the unevaluable floor flag).

Open adversarial findings. The adversarial review that produced this revision worked three lenses over the machine and returned twenty-two findings. Twenty-one are repaired in the code this edition runs on. Twenty of those carry a bead: `nightingale-9c5ma`, `-wlhp8`, `-70njd`, `-2rw1h`, `-qag0b`, `-m1uvn`, `-u913i`, `-t0yoy`, `-3eyuz`, `-krqtl`, `-v87as`, `-jeb74`, `-ro9jv`, `-bl8qr`, `-wogi8`, `-j070c`, `-xwlga`, `-h93sq`, `-n59sk`, and the orchestrator half of `-ya1mm`. The twenty-first is an artifact defect with no bead of its own: the engines runner printed one corrected falls figure two different ways, and section G now derives that figure from section B instead of restating it. Between them they cover cross-facility receipt reads, a forged program key that could write rows into the dossier's scoreboard, suppression that was one renderer deep, an evidence floor another facility's rows could clear, and the structural-input escapes that let an invented figure ride out as a counted one. Every one of those beads is still open in the tracker: repaired in code is not the same as closed, and closing them against evidence is a separate step.

Sixteen findings are open and unrepaired, from this review and the reviews before it: `nightingale-fn7uu`, `-fc2il`, `-f6xlt`, `-3zksa`, `-87g3c`, `-kdz8s`, the gap-analysis half of `-ya1mm`, `-qige0`, `-pndzq`, `-4olnl`, `-geegp`, `-1qo19`, `-3a27x`, `-h1i9v`, `-mpyyu` and `-oi7os`. They cover the ambient-documentation claims withdrawn in §1.3, the flow simulator's silent three-unit scope, an HMAC key that falls back to the published development key, a governance composer that never returns inside ten minutes, and related correctness and disclosure defects. They are named here because a limitations section that omits the findings its own audit produced is not a limitations section.

10. Conclusion and next steps

Revision 2 published a projection and called it a pilot. Revision 3 publishes a machine, its output, and the thirty-six statements we withdrew when the machine and the sources disagreed with the paper. The remaining claim is narrower and easier to check: on a synthetic 274-bed facility, the platform's own engines produce a reproducible gap assessment, a costed remediation roadmap with a 28-day critical path, six named refusals where no basis exists, and a never-event category of $112,556.50 where the July model printed $264,870.

The next step is unchanged and now better specified. Nightingale OS will build a custom digital pilot on a partner facility's own census, acuity, staffing and evidence data, using the same two commands, the same run-hash discipline and the same refusal codes — replacing modeled inputs with measured ones one line at a time. The first live deployment is the moment the assumption register in Appendix B starts shrinking.

References

Entries carry a resolution label. **verified-primary** means the article record or the publisher document itself was read; **verified-secondary** means a reliable secondary source was read; **unresolved** means no locator fixes the work as printed, and the entry says so rather than offering a plausible-looking substitute.

1. Cline K. *Nightingale OS: Pre-Registration of Software Artifact.* Nightingale OS Inc.; 8 May
2. **Unresolved (self-published)** — needs a stable locator (repository commit hash, or a

public URL) before a diligence reader can inspect the eight claims audited in §8.1.

1. Walonoski J, Kramer M, Nichols J, et al. Synthea: An approach, method, and software mechanism

for **generating** synthetic patients and the synthetic electronic health care record. *J Am Med Inform Assoc.* 2018;25(3):230-238. doi:10.1093/jamia/ocx079. PMID 29025144. **Verified-primary.**

1. The MITRE Corporation. *Synthea synthetic patient generator.* https://synthea.mitre.org

(accessed 15 June 2026). **Verified-secondary.**

1. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient

mortality, nurse burnout, and job dissatisfaction. *JAMA.* 2002;288(16):1987-1993. doi:10.1001/jama.288.16.1987. PMID 12387650. **Verified-primary.**

1. Johnson A, Bulgarelli L, Pollard T, Celi LA, Mark R, Horng S. *MIMIC-IV-ED* (version 2.2).

PhysioNet; 2023. doi:10.13026/5ntk-km72. **Verified-primary.**

1. Bouldin EL, Andresen EM, Dunton NE, et al. Falls among adult patients hospitalized in the

United States: prevalence and trends. *J Patient Saf.* 2013;9(1):13-17. doi:10.1097/PTS.0b013e3182699b64. **Verified-primary** — 3.56 falls per 1,000 patient-days, 26.1% injurious, so 0.93 injurious falls per 1,000. This supersedes Revision 2's attribution of the 3.5 figure to AHRQ.

1. Dykes PC, Burns Z, Adelman J, et al. Evaluation of a Patient-Centered Fall-Prevention Tool Kit

to Reduce Falls and Injuries. *JAMA Netw Open.* 2020;3(11):e2025889. doi:10.1001/jamanetworkopen.2020.25889. **Verified-primary** — 15% reduction in all falls, 34% in injurious falls.

1. Wong CA, Recktenwald AJ, Jones ML, Waterman BM, Bollini ML, Dunagan WC. The cost of serious

fall-related injuries at three Midwestern hospitals. *Jt Comm J Qual Patient Saf.* 2011;37(2):81-87. doi:10.1016/s1553-7250(11)37010-9. **Verified-primary** — $13,316 attributable cost in 2004–2006 dollars, against which $15,000 is a conservative floor.

1. Saint S, Greene MT, Krein SL, et al. A Program to Prevent Catheter-Associated Urinary Tract

Infection in Acute Care. *N Engl J Med.* 2016;374(22):2111-2119. doi:10.1056/NEJMoa1504906. **Verified-primary** — 14% overall, 32% in the non-ICU arm, ICU rates largely unchanged.

1. Agency for Healthcare Research and Quality. *Estimating the Additional Hospital Inpatient Cost

and Mortality Associated With Selected Hospital-Acquired Conditions,* Exhibit 7. https://www.ahrq.gov/hai/pfp/haccost2017-results.html (retrieved 2026-09-02). **Verified-primary** — CAUTI $13,793 (95% CI $5,019–$22,568); CLABSI $48,108; pressure ulcers $14,506.

1. Berenholtz SM, Lubomski LH, Weeks K, et al. Eliminating central line-associated bloodstream

infections: a national patient safety imperative. *Infect Control Hosp Epidemiol.* 2014;35(1):56-62. doi:10.1086/674384. **Verified-primary** — 43% adjusted decrease, against which the model's 35% is conservative. This replaces Revision 2's Appendix B citation, which resolved to a non-significant result.

1. Zimlichman E, Henderson D, Tamir O, et al. Health care-associated infections: a meta-analysis

of costs and financial impact on the US health care system. *JAMA Intern Med.* 2013;173(22):2039-2046. doi:10.1001/jamainternmed.2013.9763. **Verified-primary** — CLABSI $45,814; CAUTI $896.

1. VanGilder CA, Cox J, Edsberg LE, Koloms K. Pressure Injury Prevalence in Acute Care Hospitals

With Unit-Specific Analysis: Results From the International Pressure Ulcer Prevalence (IPUP) Survey Database. *J Wound Ostomy Continence Nurs.* 2021;48(6):492-503. doi:10.1097/WON.0000000000000817. **Verified-primary** — 2.58% point prevalence across 296,014 patients. This replaces Revision 2's NPIAP attribution.

1. European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel, Pan Pacific

Pressure Injury Alliance. *Prevention and Treatment of Pressure Ulcers/Injuries: Clinical Practice Guideline,* 3rd ed. 2019. Supporting single-site reductions: *Adv Skin Wound Care* 2023;36(12):658-665 (39%) and Hamidi L, 2023;36(8):1-5 (74%). The 25% figure this paper uses is a bounded assumption rather than a measured effect, and is labeled as one.

1. Ma SP, Liang AS, Shah SJ, et al. Ambient artificial intelligence scribes: utilization and

impact on documentation time. *J Am Med Inform Assoc.* 2025;32(2):381-385. doi:10.1093/jamia/ocae304. **Verified-primary** — median daily documentation time fell 6.89 minutes across 45 physicians and 17,428 encounters. No peer-reviewed study has measured ambient documentation savings for inpatient registered nurses.

1. Falvo T, Grove L, Stachura R, Vega D, Stike R, Schlenker M. The opportunity loss of boarding

admitted patients in the emergency department. *Acad Emerg Med.* 2007;14(4):332-337. doi:10.1197/j.aem.2006.11.011. **Verified-primary** — $380.90 per boarding hour in 2004–2005 dollars, against which $250 is conservative.

1. McConnell KJ, Richards CF, Daya M, Weathers CC, Lowe RA. Ambulance diversion and lost hospital

revenues. *Ann Emerg Med.* 2006;48(6):702-710. doi:10.1016/j.annemergmed.2006.05.001. **Verified-primary** — $1,086 per diversion hour (95% CI $611–$1,461).

1. Ashley DW, Mullins RF, Dente CJ, et al. What Are the Costs of Trauma Center Readiness?

Defining and Standardizing Readiness Costs for Trauma Centers Statewide. *Am Surg.* 2017;83(9):979-990. PMID 28958278. **Verified-primary** — Georgia, 2013 survey; Level II average $2,333,113 across nine centers; four cost categories; stratified by American College of Surgeons level rather than by bed count.

1. Evans KJ, Samanta D. Development of a Prediction Model for Trauma Registry Chart Abstraction

Time. *J Trauma Nurs.* 2020;27(6):369-373. doi:10.1097/JTN.0000000000000545. PMID 33156254. **Verified-primary** — 38.95-minute regression intercept over 600 charts at one Level I center.

1. Erlang AK. Solution of some problems in the theory of probabilities of significance in

automatic telephone exchanges. *Elektrotkeknikeren.* 1917;13:5-13. English translation: *Post Office Electrical Engineers' Journal.* 1918;10:189-197. **Verified-secondary.**

1. Cal. Code Regs. tit. 22, §70217 (1:2 critical care; **1:5 medical/surgical**; 1:3 step-down;

1:4 telemetry). Cal. Health & Safety Code §1280.3, as amended by SB 596 (Stats. 2025, ch. 773, approved 2025-10-13): $15,000 first violation, $30,000 second and subsequent, violations on separate days treated separately. Or. Rev. Stat. §441.765 (HB 2697, 2023): 1:2 ICU, 1:5 medical-surgical, 1:4 emergency, oncology, telemetry and pediatrics; amendments operative 1 July 2026. Wash. Rev. Code ch. 70.41 (ESSB 5236, 2023): staffing committees and enforceable staffing plans rather than fixed ratios. **Verified-secondary** against the legislature's own bill text and the California Department of Public Health's all-facilities letter 23-27.

1. Centers for Medicare & Medicaid Services. *Conditions of Participation for Hospitals,*

42 C.F.R. pt. 482, including §482.21 (quality assessment and performance improvement); Form CMS-2567 statement of deficiencies and plan of correction; 42 C.F.R. §424.515 and pt. 489 (enrollment revalidation); the CMS Cell Suppression Policy, applied with HIPAA 45 C.F.R. §164.514(b). **Verified-secondary.**

1. The Joint Commission. *Sentinel Events (SE)*, Comprehensive Accreditation Manual for

Hospitals. **Unresolved (incomplete)** — the current edition and chapter must be fixed before this is citable as printed.

1. US Bureau of Labor Statistics, Occupational Employment and Wage Statistics, May 2025.

Registered Nurses (29-1141) hourly mean $48.76, median $46.90, annual mean $101,420, series `OEUN0000000000000029114103`. Also, as *(median / mean)*: Medical and Health Services Managers (11-9111) $59.55 / $67.77; Compliance Officers (13-1041) $38.81 / $42.50; Health Information Technologists and Medical Registrars (29-9021) $32.70 / $36.04. Wage only, no benefits or overhead multiplier. **Verified-primary** through the BLS public API, 2026-09-02.

1. Centers for Disease Control and Prevention, National Center for Health Statistics. *FastStats

— Emergency Department Visits,* National Hospital Ambulatory Medical Care Survey 2022 National Summary Tables: 155.4 million visits. Retrieved 2026-09-02. **Verified-secondary.**

1. American Hospital Association. *Fast Facts on U.S. Hospitals, 2026* (2024 Annual Survey):

907,216 staffed beds, 6,100 hospitals, 5,121 community hospitals. The 2025 edition figures — 913,136 beds across 6,093 hospitals — are the ones the §6 revenue ceilings use, and the canonical page now serves the 2026 edition. **Verified-secondary.**

1. Centers for Disease Control and Prevention, National Healthcare Safety Network. *National and

State Healthcare-Associated Infections Progress Report* and device-associated module data. **Unresolved (CAUTI)** — no NHSN publication was found stating the 1.5 per 1,000 catheter-days pooled mean Revision 2 printed; 2023 acute-care reporting implies approximately 0.74. **Verified-secondary (CLABSI)** — 1.0 per 1,000 line-days is plausible against the published medical-surgical ICU figure of 0.8.

Appendix A. Machine outputs

Both current artifacts are committed to the repository and both carry a `--verify` mode that byte-compares the committed files and exits 0. The two current commands are in §3; the September 2 artifacts remain historical records and are not expected to verify against current source.

The engines run.

Field	Value
Path	`docs/owl/pilot/control-facility-engines/run-2026-09-19/`
Run hash	`76eedcef9979374366c4b9658c5c7b3b88455a1a516fd7c44ffe0b56cfa8be7b`
As of	2026-09-19T00:00:00Z
Seed	42 · 168-hour horizon · 20 replications
Facility	the `data/demo/` fixture set
Store	none — pure engines, no database
Rates version	BLS OEWS May 2025
Runner	`server/scripts/run-digital-pilot.mjs`
Sidecar	`environment.json` — builder runtime facts; outside the run hash and ignored by `--verify`

The Owl gap assessment.

Field	Value
Path	`docs/owl/pilot/control-facility/run-2026-09-19/`
Run hash	`27bee20243095644903318d7af5f0b9ea3846cbc6e2f6c86d6034c13b76a77cd`
As of	2026-09-19T00:00:00.000Z
Facility	`SYN0274`, Control General Hospital
Store	in-memory fixture store, no database
Rates version	`oews-may2025.api.v1`
Effort model	`owl-effort-sheet.v1`
Code version	`owl-pilot-gap-assessment.v1`
Figures rendered	2,290
Runner	`server/scripts/run-owl-pilot-gap-assessment.mjs`

The engines run's `engines.md` §H prints the SHA-256 digest of all 14 engine modules, 6 fixtures and 2 runner files it touched, and `engines.json` stores them; a change to any one of them moves the run hash.

The Owl run's input receipt is `evidence-inventory.csv`, hash `df912990f9d29f912f9516796786b13f095cf7b13dc20fcffa284a507215e4cf`, 105 rows, attester `emp-049`.

Environment pins on the control run. The build-commit stamp is the literal string `control-run-not-a-build`, because a fixture run claims no commit as its provenance. The hash-chain key is the repository's own published development key (`hmacKeyResolver.js:29`, key id `6980a0ef5062ef17`), which is not and never was a secret. The facility-identifier floor is cleared. The whole run executes under a clock pinned to the as-of instant, because two sibling services stamp the wall clock into a hashed body; pinning changes no figure and makes the artifact byte-reproducible on another machine.

Receipts. 154 rows: 105 `accreditation.recorded`, 29 `owl-designation-ledger`, 18 `owl-regulatory-lifecycle`, 1 `owl-intake`, 1 `owl-pilot-gap-assessment`. 49 rows sit on chained kinds and all 49 verify — `owl-designation-ledger` 29 of 29, `owl-regulatory-lifecycle` 18 of 18, `owl-intake` 1 of 1, `owl-pilot-gap-assessment` 1 of 1 — with no break reported on any chain. The other 105 carry a null content hash because their kind is not high-stakes.

Refusal authorities and unblock paths, complementing the summary in §5:

Code	Authority cited	What would unblock it
`PGA_003`	provenance-liveness honesty	restore read access, then re-run
`EST_008`	authoring half vs per-learner delivery half	supply a head count, zero included
`EST_004`	US BLS OEWS — no row for the occupation	supply a published national wage
`EST_003`	HFMA cost-accounting standards	obtain a facility-specific capital quote
`PGA_008`	designation ledger pass-through	correct the roster key, or drop the row
`OWL_INTAKE_016`	Joint Commission LD.04.01.07	register the requirement, or drop the row

Modeled versus sourced, across the Owl run's 2,290 rendered figures:

Basis	Figures	Kind
`published-time-study`	1	published source
`peer-reviewed-cost-study`	3	published source
`public-procurement`	16	published source
`bls-oews-wage`	555	published source
`modeled-assumption`	1,011	stated assumption
`stated-planning-convention`	274	stated assumption
structural count	430	counted, not modeled

The census falls from the 2,414 an earlier draft of this revision printed, because small-cell suppression now withholds a role's figures in the dossier as well as the report: a suppressed cell is no longer a rendered figure

anywhere.

The one measured anchor is printed as a single value rather than widened into a band, because the band would be the company's and not the measurement's: **0.65 hours per evidence artifact**, from Evans and Samanta 2020 (approximately 39 minutes per case, applied here to evidence-artifact abstraction and assembly).

Sensitivity over the declared grid. Elasticity is measured across the grid shown; the run asserts nothing about which knob dominates.

Knob	Grid	Totals	Elasticity
`scopeMultiplier`	0.75, 1, 1.5	$33,890 / $35,792 / $39,600	15.9%
`rateAxis`	median-only, mean, strict	$35,381 / $35,792 / $35,792	1.1%
`weeklyCapacity`	0.5, 1, 2	$35,792 / $35,792 / $35,792	0.0%

Appendix B. Assumption register

A line appears here when no code in the repository produces it. Each carries the value the July 2026 model published, a basis label, the verdict from the source review, and either a citation or the words *unsourced — assumption*. Nothing in this register is presented as an engine output.

1. **CAUTI events prevented** — July: 1 event / $23,625. Basis: assumption. No CAUTI estimator

exists in the repository; the nearest modules are surveillance and benchmarking. Baseline unresolved, reduction supported only against a non-ICU arm, cost basis contradicted and replaced. Source: Saint 2016; AHRQ Exhibit 7.

1. **CLABSI events prevented** — July: 0.4 events / $19,845. Basis: assumption. No CLABSI

estimator exists, and the count was back-filled: the repository's history shows the prevented count moving from 0 to 0.4 on 2026-07-01 while the dollar figure stayed at $19,845, so the count was fitted to a number that already existed. Source: Berenholtz 2014.

1. **Ambulance diversions avoided** — July: 6 / $300,000. Basis: assumption. No diversion model

exists. The literature prices diversion per hour, so $50,000 per episode implies roughly 46 hours. Restate the lever in diversion-hours or state the assumed hours. Source: McConnell 2006.

1. **Ambient documentation hours returned** — July: 2,457 h / $135,135. Basis: assumption.

Nothing computes documentation time saved. The 35% reduction behind it is contradicted by the measured literature, which reports single digits to the mid-teens and none of it for inpatient nurses. Base case 15%; 35% is an upper bound only. Source: Ma 2025; BLS OEWS May 2025.

1. **Length-of-stay reduction, 0.8 days** — July: 480 bed-days / $1,056,000. Basis: assumption.

The length-of-stay predictor projects a stay and models no reduction in one; the only outcome-anchored calibration in that family scores prediction error, a different quantity. This was the largest single line in the July model at 51% of its total, and it rests on the softest cost basis: $2,200 is a floor against *average* expense per inpatient day, while an avoided bed-day saves marginal cost, or contribution margin if the bed is backfilled. Source: KFF State Health Facts, 2023 AHA Annual Survey — nonprofit $3,288, state/local $2,857, for-profit $2,529.

1. **ED boarding hours cut** — July: 130 h / $32,625. Basis: assumption on the quantity.

Detectors measure open intervals and model no reduction. The unit price is the best-sourced in the model: $250 per boarding hour is 66% of the $380.90 implied by Falvo 2007, in dollars twenty years older. Source: Falvo 2007.

1. **Overtime shifts avoided** — July: 67 shifts / $20,218. Basis: assumption on the count,

derivable on the price. No shift-count model exists, and the same 67 is reused for agency shifts. $20,218 ÷ 67 = $301.76, almost exactly the half-time premium on a 12-hour shift at the May 2025 registered-nurse mean (0.5 × $48.76 × 12 = $292.56). Source: BLS OEWS May 2025.

1. **Agency shifts eliminated** — July: 67 shifts / $134,784. Basis: assumption. The count is

asserted and identical to the overtime count. $134,784 ÷ 67 = $2,011.70 per shift implies a bill rate near $167 an hour, a crisis-peak figure. The deeper problem is the accounting basis: eliminating an agency shift saves the *difference* between it and the staff shift that replaces it, roughly $315 on a 12-hour shift at a $75 bill rate, unless the shift is eliminated outright. Source: unsourced — assumption on the price; BLS OEWS May 2025 on the staff wage.

1. **Ratio violations eliminated, 8 → 0** — July: $120,000. Basis: assumption on the count,

published source on the price. The registry evaluates one unit at one moment and returns zero penalty exposure because its California entry carries no penalty field. Source: Cal. Health & Safety Code §1280.3 as amended by SB 596 (2025).

1. **Census-forecast accuracy, 85%** — Basis: assumption. **Dropped.** Nothing in the repository

measures forecast accuracy, and no observed-outcome series exists to measure one against. Source: unsourced — assumption.

1. **Falls prevented** — July: 12 falls / $176,400. Basis: published-source arithmetic on the

wrong baseline. It reproduces exactly as 3.5/1,000 × 8,400 patient-days × 40% × $15,000; the arithmetic is right and the inputs are wrong. Corrected in §4.2. Source: Bouldin 2013; Dykes 2020; Wong 2011.

1. **Pressure injuries prevented** — July: 4 injuries / $45,000. Basis: published-source

arithmetic, attribution corrected. It reproduces exactly as 2.5% × 600 admissions × 25% × $12,000. The 2.5% was attributed to a body that does not say it; at the defensible 2.58% the line becomes $46,440, and the 25% reduction remains a bounded assumption. Source: VanGilder 2021; EPUAP/NPIAP/PPPIA 2019.

1. **Platform cost, $164,400** — July: 274 beds × $600/bed/yr, as-of 2026-07-01. Basis:

assumption. A price card, not a computed quantity, and every return figure in the July model divided by it. Source: unsourced — assumption.

Appendix C. The eight pre-registered claims, audited

Read-only audit at commit `6afb979ee`, 2026-09-02. *Mechanism present* means the named mechanism exists in code at the cited location; the audit does not execute any claim end to end.

1. **A closed clinical event posts a FHIR R4 bundle without re-charting.**

`clinicalEventFhirSync.js:8` builds a transaction bundle at `:187`; `:150` refuses unless the event is resolved (`EVENT_NOT_RESOLVED`). Mechanism present; the timing half is untested, because no live EHR has ever accepted a bundle.

1. **No clinical event resolves with required elements missing unless each omission carries a

non-empty reason.** `model RequiredElementSkip` at `schema.prisma:12146`; gate at `clinicalEventCloseoutGate.js:13,51`. Mechanism present.

1. **No wellness observation persists with an observer role outside the peer allow-list.**

`pulseScale.js:70` refuses when the role is absent or non-peer. Mechanism present.

1. **No individual wellness row older than 90 days carries an unanonymized observer id.**

`schema.prisma:11599–11601` documents the purge field and the anonymization logic sits in `pulseScale.js`. Mechanism present in the service; the scheduled sweep that enforces it on a clock was not verified in this pass.

1. **Equity placement returns the worst-severity quadrant across cohorts, never the composite.**

`phaseSpaceDiagnostic.js:141,178`. Mechanism present.

1. **A unit firing its slack floor three or more times is ineligible for bonuses.**

`unitCompetition.js:14–15,67,73` — three fires set the multiplier to 0.0 and the ineligible flag to true. Mechanism present.

1. **A sentinel-event recovery cannot close with an incomplete layer or before the six-month

window.** `sentinelEventRecovery.js:63` sets the 180-day constant; `:343–358` and `:417–422` enforce the due date and the required six-month action. Mechanism present.

1. **Every privacy-critical row carries a citation, and the hash-chained audit log is verifiable

end to end. `auditLog.js:48,347` gate the chain on high-stakes kinds, and the `prevHash` and `contentHash` columns are nullable (`schema.prisma:584`).** Narrower in code than in prose** — see §8.1 for the scope clause this paper prints instead.

Appendix D. Commands behind §7

Run from the repository root, at `888975d8c`. Every figure in §7's table is keyed to one of these, and fourteen of the fifteen were executed read-only against that commit while this revision was written. D7 is the exception, and it is the exception on purpose: it reads a git-ignored local artifact that exists only in a checkout where a batched run has been executed, so it was read on the working checkout on 2026-09-03 rather than in the artifact worktree, which carries no such file.

```
# D1  Prisma data models
grep -c '^model ' server/prisma/schema.prisma
# D2  Bundles, surfaced/unsurfaced/unwired, citations resolved
sed -n '3p;6p;7p;8p;11p' docs/feature-catalogue/FEATURE_CATALOGUE.md
# D3  Catalogue families
awk '/^\| Family \| Bundles/{f=1;next} f&&/^\|---/{next} f&&/^\|/{n++} f&&!/^\|/{exit} \
  END{print n}' docs/feature-catalogue/FEATURE_CATALOGUE.md
# D4  Server route patterns (canonical field)
node -e 'console.log(JSON.parse(require("fs").readFileSync(

"server/src/services/_generated/routeRenderMap.json","utf8")).stats.serverRoutePatterns)'
# D5  Server lines of code and file count, excluding generated seeds
git ls-files server/src > /tmp/s.txt
grep '\.js$' /tmp/s.txt | grep -v '_generated/' | tr '\n' '\0' | xargs -0 cat | wc -l
grep '\.js$' /tmp/s.txt | grep -v '_generated/' | wc -l        # => 4042
# D6  Client lines of code — including tests, then excluding them
git ls-files client/src > /tmp/c.txt
grep -E '\.(js|jsx)$' /tmp/c.txt | tr '\n' '\0' | xargs -0 cat | wc -l
grep -E '\.(js|jsx)$' /tmp/c.txt | grep -vE '\.(test|spec)\.jsx?$|/__tests__/' \
  | tr '\n' '\0' | xargs -0 cat | wc -l                        # => 247199
# D7  Test files and results (git-ignored local artifact; overwritten by the next local
run)
node -e 'const r=require("./server/src/services/_generated/testResults.json");
  let p=0,f=0,s=0;
  for(const x of Object.values(r.files)){p+=x.passed;f+=x.failed;s+=x.skipped}
  console.log(r.fileCount,p,f,s)'                              # => 2885 72122 0 17
# D8  Refusal codes: declarations and distinct, over server/src/bundles/*/manifest.js
node -e '
const fs=require("fs"),path=require("path");
let entries=0; const codes=new Set(); let manifests=0;
for(const d of fs.readdirSync("server/src/bundles")){
  const p=path.join(process.cwd(),"server/src/bundles",d,"manifest.js");
  if(!fs.existsSync(p))continue; manifests++;
  let m; try{m=require(p);}catch(e){continue;}
  for(const r of (m.refusalCodes||[])){entries++; if(r&&r.code)codes.add(r.code);}
}
console.log({manifests, declaredEntries:entries, distinctCodes:codes.size});'
# => { manifests: 1347, declaredEntries: 4147, distinctCodes: 3651 }
# D9  Owl programs catalogued
(cd server && node -e 'console.log(Object.keys(
  require("./src/services/owlDesignationLedger.js").knownPrograms()).length)')
# D10 Owl programs modeled end to end
(cd server && node -e 'console.log(Object.keys(
  require("./src/bundles/accreditation-
workflow/seedRequirements.js").SEED_REQUIREMENTS).length)')
# D11 Owl seed requirement elements
(cd server && node -e 'const s=require(
  "./src/bundles/accreditation-workflow/seedRequirements.js");
  console.log(s.listAllRequirements().length)')
```

```
# D12 Owl certifier packs and pack criteria
(cd server && node -e 'const p=require("./src/services/owlProgramPacks.js").listPacks();
  console.log(p.packs.length, p.packs.reduce((a,k)=>a+(k.criterionCount||0),0))')
# D13 Eagle domains data-grounded
sed -n '245,248p' server/src/services/currentStateTwinModeler.js
grep -n 'groundedCount' server/tests/currentStateTwinDataBacked.test.js
# D14 Self-learning reachability
grep -m1 '^| Self-learning' docs/feature-catalogue/FEATURE_CATALOGUE.md
# D15 Commits
git rev-list --count HEAD
git rev-list --count --since=2026-07-01 HEAD
```

A tracked script, `scripts/count-refusal-codes.cjs`, counts a broader population than D8: it also walks service-level `REFUSAL_CITATIONS`, and at the same commit it returns 4,274 declarations across 3,759 distinct codes. The two must not be quoted interchangeably. §7 prints the manifest walk, because that is the population the manifest-declaration claim is about.

Prepared by Nightingale OS — Nightingale OS Inc., a Delaware C-Corp. The certificate of incorporation was executed 12 August 2026 and is held in the company record (register F-002, RESOLVED by the instrument); the Delaware filing itself is founder-attested (F-003, 2026-09-14) and is not yet evidenced by a file-stamped copy. All patients synthetic; all outcomes modeled. UNVALIDATED pending clinical-advisor review and live deployment. September 3, 2026 (Revision 3; supersedes Revision 2 of July 1, 2026).

STATUS: APPROVED by Kristen Cline, 2026-09-15. Corrected after approval, 2026-09-15, on her instruction — three factual corrections, no argument changed. The masthead still carried "founder attestation 2026-09-13; certificate to follow", contradicting this document's own footer, and is now aligned to it — the 2026-09-13 date was unsourced. And the patent was described as filed on a founder statement, when the USPTO Electronic Acknowledgement Receipt for application 64/139,773, received 2026-08-24, is in the company record.

Standing disclosures

These ride every document Nightingale OS renders into its data room. They are load-bearing and are not trimmed for length. Company stage, customers and revenue are stated in full, and in one place, in the market and go-to-market pack (§6, "Not yet — and the sequence that changes it").

1. No hospital's actual feed has ever flowed through the system. The HL7v2 ADT ingestion path is demonstrable on a replayable feed; FHIR and timekeeping integrations are scaffolds.
2. The early-warning composite is labelled UNVALIDATED in its own code (`server/src/services/deteriorationModel.js`, label present at commit f3e28b3c5, checked at build). Retiring or confirming the label requires a named clinical advisor and has not started.
3. No clinical validation. No production SSO. No independent security certification of any kind: SOC 2 Type I and II not attested, no audit engaged, no report; HIPAA is designed-for and is not a certification anyone

can hold; the BAA is at template stage with no production BAA in force; no third-party penetration-test report exists.

4. Every pilot value figure in any company document is a model on synthetic data, not a measured cost or a performance claim.
5. Every downtime path described anywhere has been exercised only by unit tests and seed data; none has run against a real EHR, a real device fleet or a real outage.
6. Nothing in this document states a conclusion about Nightingale's PHI handling or security posture, in either direction. Where the text describes data handling, it describes design intent at the cited commit.

Source: PRODUCT_MANUAL §12.1 and CANONICAL_FACTS F-011, F-033, F-034 (Edition 1 · compiled 2026-09-12). Rendered by `data-room build` on 2026-09-19 from a fact snapshot provided by manual-facts 2.0.0 + data-room-local (fill); the company register, not this document, is the authority on every count.