ORDINARA

AI engineering opportunity audit

Sample report, watermarked

Prepared for: Norvia Motion Systems, a fictional composite client
Scope: requirements, code review, test, compliance documentation, and engineering overhead across a 140-engineer software organization
Engagement window: 15 business days (illustrative)

What this document is

This is the format, depth and structure of the report an audit client receives. Every number in it is illustrative and every name is fictional: the sample exists so a buyer can judge the deliverable before engaging, not to describe any real client.

SECTION 1

The problem, its scale, and the summary

The situation the audit walked into

Norvia ships software with 140 engineers under production deadlines. Leadership knew AI should be cutting cost somewhere; what it had instead was scattered experimentation, no measured baseline, and no way to defend any investment in front of its own finance team. The question the audit answers is narrow on purpose: where do engineering hours actually go, which of them can AI remove with tooling that exists today, and in what order should the organization act.

The answer, in three lines

Of roughly 61,000 engineering hours per year across the five audited areas, about 11,400 are addressable with production-grade tooling while keeping a human accountable for every released artifact. The three highest-ranked opportunities carry 7,300 of those hours and pay back their implementation cost inside the first quarter of steady-state use. A shadow-mode pilot run during the audit itself, on the client's own documents, measured a 70 percent reduction in drafting effort with no change in review findings.

#	Opportunity	Hours/yr addressable	Payback	Score (see sec. 4)
1	Compliance documentation drafting and traceability	3,120	6 weeks	4.60 / 5
2	First-pass code review on low-risk change classes	2,480	8 weeks	4.25 / 5
3	Requirements consistency and impact analysis	1,700	11 weeks	3.80 / 5

Every score cites the model of section 4, every hour figure traces to the evidence of section 2, and section 7 lists what we refused to recommend. Nothing in this report asks to be believed without its worksheet.

SECTION 2

Method and evidence base

The fifteen days, as executed

Days	Phase	What happened
D1 to D2	Kickoff and live demo	Workflows on the table; the delivery system shown running. Scope frozen in writing.
D3 to D7	Evidence	22 structured interviews across roles; artifact review; hours quantified per area.
D8 to D11	ROI analysis and pilot	Every opportunity scored with the model of section 4; shadow-mode pilot on one document family.
D12 to D15	Report and readout	This document, a 90-minute readout, and the 30-day starting plan of section 6.

Where the numbers come from

Hours per area were reconstructed from three independent sources and cross-checked: time allocation records, ticket and review system history, and the interviews. Where the three disagreed by more than 15 percent we kept the most conservative figure. Every figure in this sample is illustrative, but the method is exactly the one applied in a real engagement.

What we did not do

No repository access, no production data access, no custody of any client artifact. Interview notes were reviewed and approved by each interviewee before entering the evidence base.

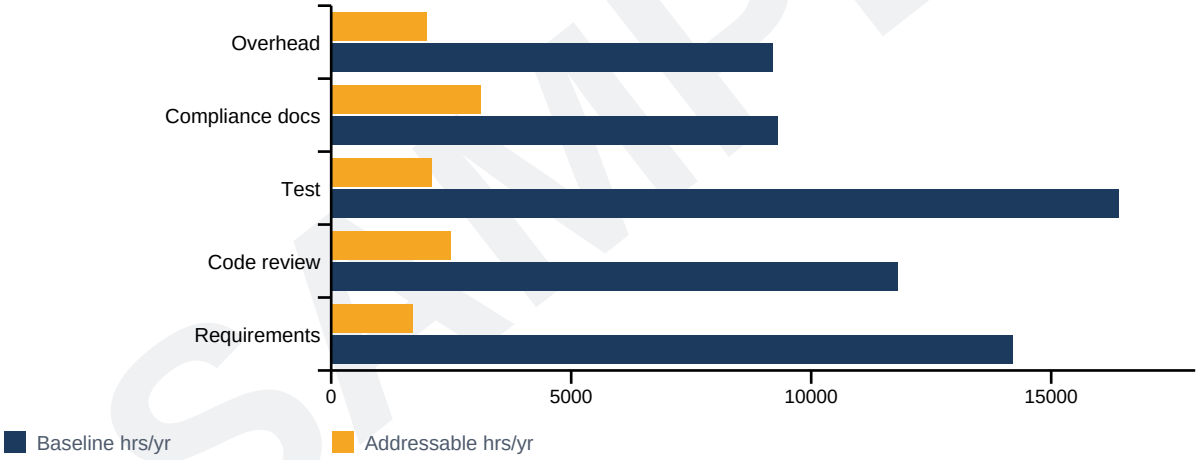
SECTION 3

Diagnosis: where the hours go

The five audited areas, with the baseline effort reconstructed and the share addressable today. Addressable means automatable with production-grade tooling while a human stays accountable for every released artifact.

Area	Baseline hrs/yr	Addressable	Share
Requirements engineering	14,200	1,700	12%
Code review	11,800	2,480	21%
Test design and maintenance	16,400	2,100	13%
Compliance documentation	9,300	3,120	34%
Engineering overhead: status, reporting, handoffs	9,200	2,000	22%
Total	60,900	11,400	19%

Baseline hours per year (navy) vs. addressable hours (amber), illustrative



The pattern matters more than any single bar: the biggest area is not the biggest opportunity. Compliance documentation is fourth in size and first in addressable share, because a third of its effort is drafting documents whose structure is fixed by process and whose inputs already exist in the toolchain. That mismatch between where effort goes and where opportunity sits is exactly what an organization cannot see without measuring, and why scattered experimentation was landing in the wrong places.

SECTION 4

The scoring model, so every rank can be checked

Each opportunity is scored 1 to 5 on four axes. The anchors below define what the endpoints mean, so a reader can dispute a score with evidence instead of taste.

Axis	Weight	1 means	5 means
Impact	35%	Under 500 hrs/yr recovered	Over 2,500 hrs/yr recovered
Feasibility	25%	Research problem, no proven tooling	Production-grade tooling exists today
Risk, inverted	20%	Touches release-critical judgment	No release-relevant decision moves
Adoption	20%	Changes how whole roles work daily	Invisible to current ways of working

The formula, and one score computed in the open

Score = 0.35 x Impact + 0.25 x Feasibility + 0.20 x Risk + 0.20 x Adoption. Opportunity 1 scores Impact 5, Feasibility 5, Risk 4, Adoption 4: 0.35 x 5 + 0.25 x 5 + 0.20 x 4 + 0.20 x 4 = 1.75 + 1.25 + 0.80 + 0.80 = 4.60. Every other score in this report computes the same way, and in a real engagement each axis value cites its worksheet: the interviews, artifacts and measurements behind it.

#	Opportunity	Impact	Feasibility	Risk	Adoption	Score
1	Compliance doc drafting agents	5	5	4	4	4.60
2	First-pass review, low-risk classes	4	5	4	4	4.25
3	Requirements consistency checks	4	4	4	3	3.80
4	Status and reporting automation	3	4	5	3	3.65
5	Test maintenance triage	3	4	4	3	3.45
6	Interface change impact mapping	3	3	3	3	3.00

Risk is scored inverted: 5 means lowest risk. Items are ranked by score, then by hours when scores tie. Items 4 to 6 are documented so the organization can revisit them with its own priorities, not because all of them should be built.

SECTION 5

The solution, and the proof it works here

What opportunity 1 looks like in operation

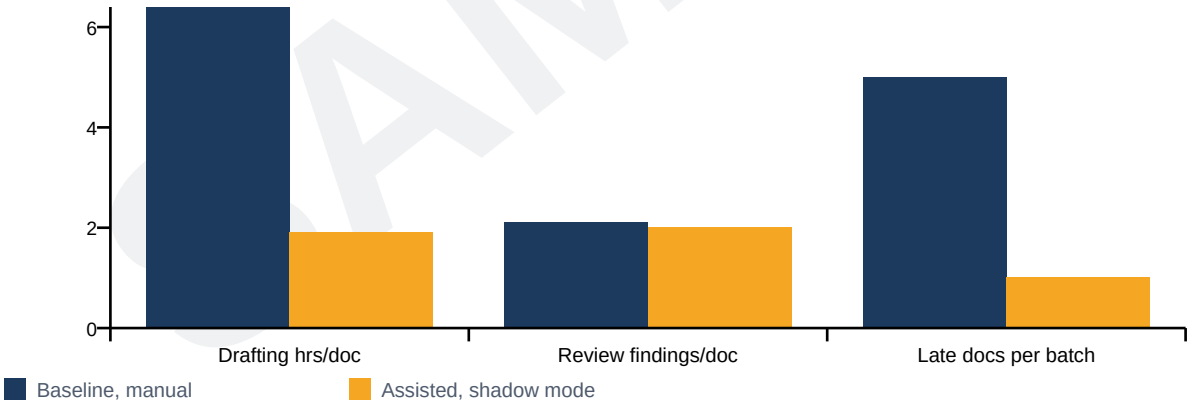
Drafting agents produce the first version of each compliance document from inputs that already live in the toolchain; the accountable engineer reviews, corrects and signs exactly as today. Nothing releases without a human signature, which is why the risk axis scores 4: the judgment stays where the process puts it, only the blank page disappears.

The shadow-mode pilot, run during the audit

Rather than ask the organization to trust a vendor benchmark, the audit ran opportunity 1 in shadow mode on 24 documents of one family, with the organization's own reviewers signing. Illustrative results:

Measure	Baseline, manual	Assisted, pilot	Change
Drafting effort per document	6.4 hours	1.9 hours	minus 70%
Review findings per document	2.1	2.0	no change
Reviewer sign-off rate, first pass	83%	86%	plus 3 pts
Documents delivered late in the batch	5 of 24	1 of 24	minus 4

Pilot, 24 documents: baseline (navy) vs. assisted (amber), illustrative



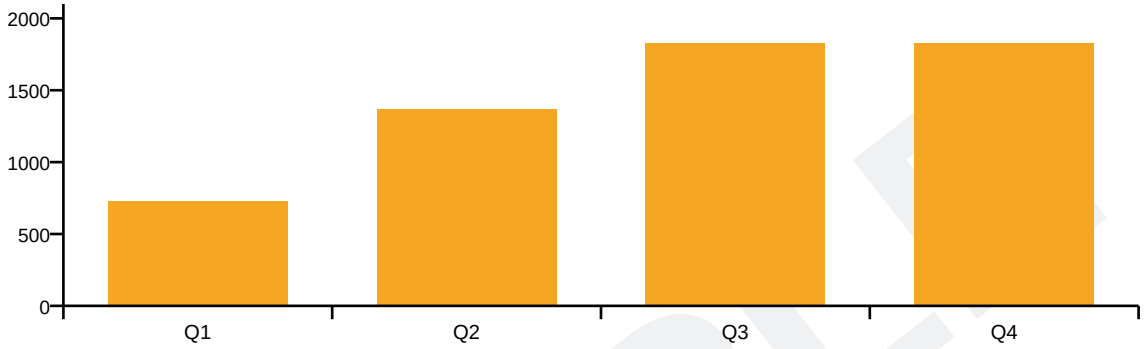
The number that persuades is not the 70 percent: it is the review findings staying flat. Speed with unchanged quality, measured by the client's own reviewers on the client's own documents, is the evidence a finance team and a quality manager can both accept.

SECTION 6

The benefits, quarter by quarter, and the first 30 days

Rolling out opportunities 1 and 2 immediately and 3 after steady state, the recovered hours compound as adoption ramps. The curve below assumes the conservative ramp of the worksheets: 40 percent of steady state in the first quarter, 75 in the second, full from the third.

Engineering hours recovered per quarter, opportunities 1 to 3, illustrative ramp



At steady state that is roughly 7,300 hours per year: the capacity of three and a half engineers, recovered without hiring, applied to the backlog the organization already has. Implementation cost is recovered inside the first quarter for opportunities 1 and 2; every assumption behind that statement lives in the appendix and can be challenged before a single build starts.

The first 30 days

Week	What ships	Who is accountable
1	Opportunity 1 scoped in writing; acceptance criteria signed	Sponsor and Ordinara
2	Drafting agent on one document family, shadow mode	Process owner
3	Human sign-off loop live; first real documents released	Process owner
4	Measured baseline vs. assisted; go or no-go for family two	Sponsor

The plan starts with the highest-confidence, lowest-friction item, in shadow mode, behind a measured go or no-go gate. Momentum in week four is worth more than breadth in week one.

SECTION 7

What we did not recommend, and why

Autonomous test result adjudication. Deciding whether a failed run is a product defect or an infrastructure flake, without human review, moves a release-relevant judgment away from an accountable engineer. The hours saved are real; the failure mode is not acceptable in this organization's release process. Scores 2 on the risk axis, which the model correctly buries.

Unreviewed requirement generation. Generating requirements from stakeholder notes and releasing them into the baseline skips the analysis step where this organization catches its most expensive defects. Assistance yes; unreviewed release, no.

APPENDIX

Assumptions, sources and status of this document

Assumptions behind the illustrative figures

140 engineers in scope; 1,650 productive hours per engineer per year; area baselines reconstructed as described in section 2; pilot measured on 24 documents of one family with the client's own reviewers; ramp of 40, 75 and 100 percent of steady state per quarter. Payback assumes the implementation costs of the fixed-scope builds current at the time of the engagement.

What a real engagement adds beyond this sample

Per-opportunity worksheets with the scoring evidence, the anonymized interview base, the measured baselines per area, and the readout session where every score can be challenged and corrected before the report is final.

Status of this document

This is a watermarked sample. The client, every figure, and every finding are illustrative, constructed to show the structure and depth of the real deliverable. It is not an offer, a contract, or a description of any real organization. The audit it illustrates is published at ordinara.ai with its scope, fee and timeline.