Class1

Five-page executive manual

AI cost controls before merge.

Canadian English. Public-facing. Large type.

Free to estimate. Paid to enforce.

The core question

Can we afford this pull request at P90 after scale?

Class1 turns a code change into a cost decision before the change is merged. It shows likely cost, budget-case risk, escalation, model fit, environmental footprint, and the control needed before approval.

Plain-language definition

Pull request	A proposed code change. It is the right moment to review cost because the change has not shipped yet.
P90	A budget-case estimate: about nine out of ten simulated outcomes are below this number.
After scale	What happens after users, requests, documents, tools, and agent steps grow.

PRODUCT

What Class1 does in the pull request

Class1 reads the code difference and asks whether the change creates new recurring AI spend. It does not try to replace engineering review. It adds a cost-control review beside it.

P50 LIKELY

+$420/mo

The middle case. Useful, but not enough for approval.

P90 BUDGET

+$1,850/mo

The number finance should treat as the review line.

12-MONTH P90

+$9,800/mo

What the risk can become after adoption and structure growth.

What it reads	What it means in plain English	Why it matters
model	Which AI model the code calls.	Different models have different price and capability.
max_tokens	How long the answer is allowed to be.	Longer answers can widen the cost tail.
retries	How many times the system tries again after failure.	Failures can multiply cost.
fallback model	Backup model used when the first path fails.	Fallbacks can silently move work to a higher-cost path.
tool schemas	Descriptions of tools sent to the model.	Agents may pay recurring input tokens even when tools are not used.

What the PR comment should answer

- How much AI workload did this PR create?
- Is the selected model fit for the task, or merely cheap per token?
- Does the P90 exceed the policy gate?
- What controls would reduce the risk before merge?

METHOD

The simple formula behind the report

The product is easier to understand if the estimate is split into pieces. The model price is only one piece. The pull request can change the amount of AI work, the uncertainty, and the future growth path.

AI Cost = Q x R + C + E

Quantity x Rate + Contingency + Escalation

Part	Plain-language meaning	Example
Q: Quantity	How much AI work the PR creates.	Calls, tokens, retries, tools, agent steps.
R: Rate	The price basis for the model and pricing structure.	Input/output, cache, batch, tiers, date.
C: Contingency	The visible risk allowance.	P90 minus P50.
E: Escalation	How the cost can grow over time.	More users, deeper agents, bigger context.
Estimate class	How mature the estimate is.	Class 5 early; Class 1 calibrated.

The code islands that matter

Model fit Cheap per token is not always cheap per completed task.

Structured pricing Rates may include cache, batch, tiers, and effective dates.

MCP/tool overhead Tool definitions can become recurring input tokens.

Actuals loop Estimate -> actual -> variance -> calibration.

Audit basis Every number should say whether it is measured, inferred, assumed, excluded, or unpriced.

COMMERCIAL MODEL

Free to estimate. Paid to enforce.

The free layer proves the engine. The Business Pilot sells control: private repositories, blocking gates, actuals ingestion, a private Blue Book, and a monthly report that a CTO, CFO, and CEO can all read.

Open core

$0

- Browser-local demo
- Local estimate from diff
- Advisory PR comment
- Open data and price index
- Self-hosted Action

Business Pilot

$799/mo

- Private repos
- Blocking budget gates
- Actuals and variance
- Private Blue Book ledger
- Monthly CTO/CFO/CEO report

Deliverable	What the buyer receives
PR cost comment	P50, P90, P95, contingency, escalation, class, footprint, and gate.
Policy config	Budget thresholds, retry caps, quantity requirements, review owners.
Actuals pipeline	CSV/provider/OTel/FOCUS inputs normalized against estimates.
Blue Book ledger	Private history of estimates, actuals, variance, rates, and evidence.
Monthly report	Board-readable summary for engineering, finance, and leadership.

Business rule

The buyer should not be sold a generic AI dashboard. The buyer should be sold a cost-control loop: estimate, approve, measure, compare, calibrate, and enforce.

PRODUCT HONESTY

Say what is live, what is pilot, and what is coming

A strong product can be ambitious and honest at the same time. The safest public language is precise: browser-local demo, self-hosted Action, Business Pilot for private actuals, and marketplace install coming.

Claim	Status	Safer wording
Browser demo	Live	Real estimation logic, browser-local.
Self-hosted GitHub Action	Live	Runs inside the customer's CI environment.
Marketplace one-line install	Coming	Package and Action release required.
Private actuals calibration	Pilot	Available through onboarding.
Self-serve billing	Coming	Do not promise until ready.

Five messages to keep

- The invoice starts in the pull request.
- No quantity basis, no credible estimate.
- Contingency is not padding; it is risk you can see.
- Cheap per token is not cheap per completed task.
- The estimate does not end at merge. It starts there.

The executive summary

Class1 is project controls for AI systems. The PR gate is the wedge: it shows the cost and risk of a code change before merge. The Business Pilot adds the part companies will pay for: enforcement, private data, actuals, calibration, and a report that connects engineering decisions to financial accountability.