

I put AI agents into production. Here is what they do not do.

Today, an experienced architect can steer a large software portfolio alone with AI agents, provided the reflex of “I ask, it codes” is replaced by a complete engineering system: memory, rules, tests, reviews and traceability.

2,048

commits

121

active days

226requests triaged in a
single pass**60+**architecture
decisions**~40**

.NET libraries

20+

data sources

40%of the code is test
code

Measured on four of the five repositories, over 121 days of activity.

WHO DECIDES, WHO EXECUTES?

Mine

- Business decisions and trade-offs
- Accountability and architecture
- Final approval, above all before anything is written to production

The agent's

- Explore the code and the data, implement, test
- Review, document, deploy
- Repetitive monitoring, every night

The agent can propose, implement, review and document. Accountability stays human and cannot be delegated.

FROM ISSUE TO PRODUCTION

Human decision, GitHub issue, agent, tests, review, pull request, human approval, production, observation, new decision. Every request first becomes an issue, scored on the financial risk and the number of records affected:

$$\text{Score} = (\text{Value} \times \text{Impact} \times \text{Risk} \times \text{Environment}) / \text{Effort}$$

THE SYSTEM

Five repositories and a shared memory: a data project for an SME, a cross-reference search engine, a framework of about 40 .NET libraries, the web showcases and a governance repository (decisions, ADRs, journal, audit reports).

- **Guardrails:** specification before the code, self-review backed by mutation testing, full test suite, independent review before merge.
- **No silent errors, no guessing.** My process aims to write to production only after approval on Test.
- **Automations:** a morning routine that reads the execution logs of the test and production environments, writes the overnight report and proposes the next request (its only write action: a pull request), a read-only agent, workflows and hooks. An automation observes and reports before it earns the right to act.

THE LIMITS

- Roughly one commit in five is a fix: the agent is fast and sometimes wrong.
- In one triage, it proposed 21 closures: 12 held up under verification, 9 were refuted.
- The system has an upkeep cost and takes experience.

AI has not reduced accountability. It has increased the capacity to steer.

A problem in an SME is rarely only technical or only human: it is both, seen from two angles. Over 30 years in the field, in data architecture and IT for SMEs.