Engineer Software

Custom AI agents built for your trade

Custom AI agents for an agreed workflow, connected to your systems and evaluated on your cases, with human review and scoped access.

Build an AI agent around one useful workflow: the systems it connects to, the cases it is evaluated against, the actions a person approves and the access it receives. We agree that scope with your domain experts and choose a model using those requirements. The integration and evaluation record remain part of your deliverable when the model changes.

Turn the service into a result

Make “Custom AI agents built for your trade” serve a business outcome. We can connect the technical work to the users, operations and decisions it needs to support. An agent wired to the systems you use, tested on cases your own experts write, overseen by your people and never tied to one model.

Explore our services Get in touch →

Examine the relevant work

Use the related work on this page to examine our approach. Ask about the tradeoffs and the conditions behind each result. Built a 268-check guard engine for the repo — 268 commit checks, and 146 tests of the checks themselves

Built a 268-check guard engine for the repo Get in touch →

Connected: to the systems the work lives in

An agent that cannot reach your systems can only talk. We connect it to the ones the work happens in — the ERP, the CRM, the booking system, the mailbox, the document store — each through a described tool with a narrow job: look up a vessel, draft a quote, file a case. Tools are offered over MCP or an API, every input is typed and validated, and an agent that needs your company’s own records reads them from data made ready for it. Memory of past cases lives in your database, not in a vendor’s chat history.

Measured: against cases your experts write

Without tests, “it seems better” is the only measurement anyone has. We sit with the people who do the work and write down real cases from your trade, each with the answer a senior person would give. Every change — a prompt, a tool, a model — runs against all of them before it ships, and the score is kept per case over time, so review can reveal regressions in the agreed cases. It is the quality gate we run on our own code, applied to an agent’s answers.

Supervised: a person decides where it matters

Not every action should be the agent’s alone. We write down, with you, what it may do on its own, what it drafts for a person to approve, and what it never touches, and the rules are enforced in the tools rather than asked for in the prompt. A review screen shows a person what the agent proposes and why; a correction made there becomes a new test case. Every step is traced with OpenTelemetry, so when an answer goes wrong you can see which tool, which document and which model produced it.

The model is **the part you rent** — the cases, the tools and the rules are the part you own.

Keep model-specific code behind a clear interface, so a model change can be evaluated and scoped before deployment.

Evaluate model changes before rollout

Model choice is configuration where the integration permits it. We compare a candidate against your agreed cases for answer quality, failure behavior, cost and latency. The evaluation informs a deployment decision; a newer model does not automatically improve your workflow. Changes go through the agreed review and rollout process, with the previous configuration available for rollback.

The work behind it

The habits that make an agent trustworthy are the ones we already keep in our own engineering. We built a guard engine of 268 commit checks and a commit gate of 22 linters with no warning tier, then wrote tests for the checkers themselves so a check that reads nothing cannot report clean. We built first-party error monitoring and tracing, made six collectors state what they skipped, and ran a whole development programme as 51 written initiatives, each with its measurements and the decisions it declined. This site's own readiness for agents was built the same way.

How an engagement runs

It starts with one task and its cases. We sit with the people who do it, write down thirty to fifty real examples with the right answers, and agree what the agent may do alone. Then we build the smallest agent that passes them, behind a review screen, in your infrastructure or your cloud account, with its access locked to what it needs. Then the next task. You own the code, the cases and the data, so the agent stays yours when the model under it changes. It is bought like the rest of our work, on the terms in what it costs.

If there is work in your trade that an agent should take off somebody's desk, tell us what it is.

Take one task **off somebody's desk** — and measure it before you trust it.

It starts with thirty to fifty real cases from your trade, written with the people who do the work.

Ask about one task What it costs →

Scope the work before commissioning it

Describe the workflow, the existing interfaces and the users who need the change. We can define a testable delivery and a handover that lets your team maintain it.

See what it costs Get in touch →

Talk through your requirements

Tell us what you need from “Custom AI agents built for your trade”, including the deadline and the team that will maintain it. We can discuss a practical way forward.

Explore our services

welcome@engineer.company +45 50 10 61 84 Copenhagen, Denmark

Take this page with you

Contact card

Carefully Crafted With Love by D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Copyright © 2025 – present · Engineer ApS · All rights reserved.

Printed from https://software.engineer.company/ai-agents/