datsugi spark Blueprint

Automating 20k quote requests / year with Claude

How we designed an AI quoting pipeline for a freight forwarder, where a queue-based confidence score decides when a human is brought in the loop.

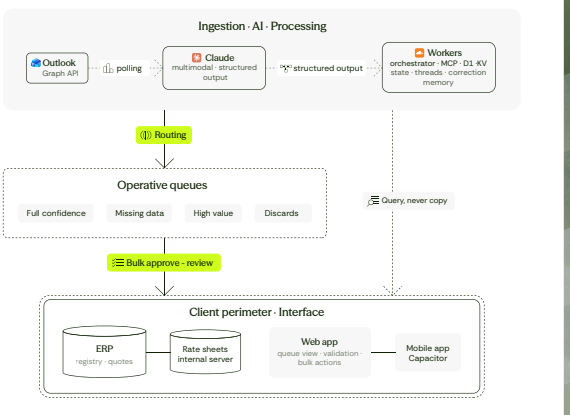
- 20k quote requests / year
- 4 people handling them
- 4 confidence routed queues

A freight forwarder takes twenty thousand quote requests a year by email, handled by four people. Plain text, Excel, scanned PDFs, photographs of paperwork — no two arrive the same way.
Reading the email is the easy part. A quote is a promise, and a system that is right most of the time is worse than none, because nobody can tell which quotes to check.

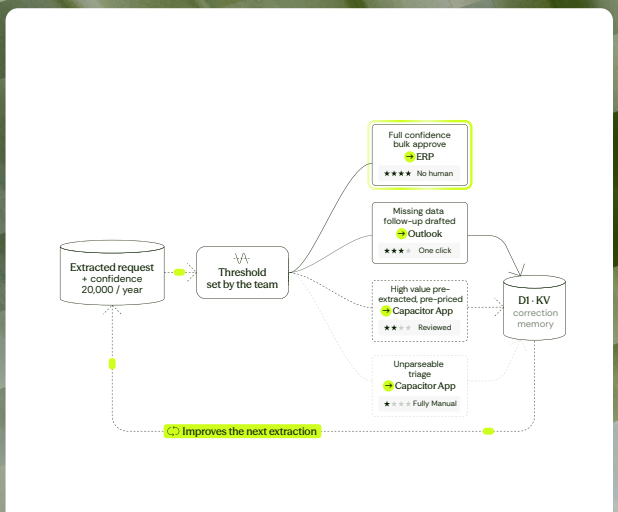
Constraints

- Four input formats, one pipeline
- Rate data that stays on the client's own servers
- Accuracy that has to be provable, not probable

Read. Validate against live data. Route.



The model doesn't decide the quote. It decides how sure it is.



datsugi spark Blueprint

Solutions ✓

1 Confidence As The Router

The model doesn't decide the quote. It decides how sure it is. Full confidence goes straight to the ERP. Missing a field drafts the follow-up. High value waits for a human — but arrives pre-extracted and pre-priced. Unparseable goes to triage. The team sets every threshold.

2 Data That Stays Put

MCP servers run on the client's own infrastructure. The model queries the ERP registry and the rate spreadsheets; it never receives a copy of either, and never holds a credential. Every correction the team makes is stored and fed back into the next extraction.

The Benefits Gained +

- + Volume stops being a headcount question
- + Attention goes only where judgement is required
- + Straightforward quotes clear in a bulk approval
- + Rate data never leaves the client's own servers

