03 · COMMERCIAL BUILD · LIVE IN INTERNAL USE

Lead & Competitor Intelligence Platform

Automated lead and competitor research for a Gulf B2B AI-software company: public web data plus narrow LLM extraction, quote-verified, delivered to a shared sheet for the sales team.

ROLE AI Engineer — designed and built both workflows (Oct–Nov 2026)

STACK n8n · Claude models via OpenRouter · JavaScript · Python · Google Apps Script · Google Sheets · Railway · maps, web-search and email-verification APIs

95	166	0	519	58	~$0.001
workflow nodes (61 + 34)	live businesses processed	lost IDs or duplicates	automated checks	baseline facts re-checked	model cost per lead (est.)

THE PROBLEM

The sales team needed to know, before the first call, which local businesses fit, what technology each one runs, who decides, and sourced facts about competitors. It was manual, scattered and not auditable. Plausible-sounding LLM facts were the main risk.

LEAD ENGINE PIPELINE

Form: industries · areas · tech → Discover (maps + web) → Read site · fingerprint tech → Score & pick top N → Find decision maker → LLM note → QA → retry → fallback → Sheet upsert (Leads · Sources · Runs)

WHAT I BUILT

- **Lead engine (61 nodes).** Discovers, reads, fingerprints and scores local businesses, finds the decision maker, and writes one sourced research note per lead; SHA-256 lead IDs make reruns update instead of duplicate.
- **Competitor research (34 nodes).** Renders pricing, feature and docs pages; the LLM extracts facts that must each carry a verbatim quote; code verifies the quote (and the numbers in any price) on the page, assigns one of 7 evidence levels and compares against the client baseline in 8 verdicts.
- **Guardrails.** The LLM is treated as an untrusted extractor: no quote on the page, no fact. Page text is treated as data (injection defence), rule-based QA with one reasoned retry and a deterministic fallback, people only counted when named verbatim on a page, and monthly cost caps per API.
- **Safe by construction.** Build-time checks fail the build if anything could message a prospect; all outreach stays human.

EVIDENCE

Live integrity check	166 leads: 0 lost IDs, 0 duplicates, 0 protected columns changed; 120 refreshed, 8 flagged unreadable
Technology detection	Shopify confirmed on 19 sites, Odoo on 4; tuned on 79 real clinic websites
Competitor extraction	One real competitor run produced 88 verified facts; 58 baseline facts quote-checked every release
Test harness	519 checks incl. end-to-end simulations of both workflows in an n8n-like sandbox

Client name, data and screenshots withheld under confidentiality.