REONE AI: A Data-Grounded, Bilingual Real Estate Intelligence and Sales Technology System for Egypt

Architecture, capabilities, and potential market impact in 2027

The founder of Real Estate One (REONE) · Real Estate One (REONE)

Technical report · Working paper · Version 1.0 · 2026-11-03 · Working draft · not peer reviewed

https://www.reone-ai.com/research/reone-ai-technical-report

> **Disclosure:** this report is written by the founder of REONE, who built the system alone and has a commercial interest in it. It is descriptive: it documents what is implemented and proposes an evaluation framework; it reports no experiments that have not been run. It is not peer reviewed, and deposit in a repository such as Zenodo does not make it so.

Architecture, Capabilities, and Potential Market Impact in 2027

Technical Report / Working Paper — version 1.0 (draft), 3 November 2026

Author: The founder of Real Estate One (REONE). Publication name, affiliation and ORCID are added when the founder approves the final version. **Status:** Not peer reviewed. No DOI has been reserved at the time of writing. **System snapshot:** All descriptions of REONE refer to the source repository as implemented at commit `745652d` (3 November 2026) unless another date is given.

> **Conflict of interest (short form).** The author is the founder of REONE (Real Estate One) and has a direct commercial interest in the system described. This report describes a system and proposes how to evaluate it. It reports no evaluation results, because no evaluation has been run. See the full Disclosure statement at the end.

Abstract

Buyers of residential property in Egypt face a fragmented market. Prices are spread across developer brochures, portals, brokers and messaging apps, and figures are often undated, unsourced, or quoted for different units under the same name. This report describes REONE AI (Real Estate One, www.reone-ai.com), a bilingual (Arabic/English) real estate intelligence and sales technology system built by a single founder between September and November 2026. REONE combines (1) an evidence pipeline that stores every observed asking price as an append-only, dated, source-attributed record, normalizes it, matches it to a canonical

catalogue of developers, projects and units with explicit confidence thresholds and human review, and computes a weighted "verified price" with median-absolute-deviation outlier exclusion and an explicit status; (2) a Market Index of median asking price per square metre with sample-size gates; (3) deterministic payment-plan, mortgage and investment calculators; (4) two conversational advisors, "Nour" and "Zaki", available by text chat and live voice, that obtain every figure through database-backed tool calls rather than generating it, run on provider-agnostic, server-minted ephemeral sessions, and request microphone access only when the user presses "start"; and (5) a lead-capture and human-advisor handoff layer. We describe the architecture from code, separate existing functions from proposed ones, and set out mechanisms by which such a system could reduce search costs and information asymmetry, stated as eight testable hypotheses. We then propose an evaluation framework with metrics, denominators, baselines, sampling plans and observation windows, and present three illustrative 2027 scenarios whose assumptions are explicit. **No accuracy, latency, conversion, satisfaction or usage results are reported: none have been measured under a defined protocol.** The author is the founder and has a commercial interest in the system.

Keywords: REONE AI; PropTech; SalesTech; Egyptian Real Estate; Artificial Intelligence; Real Estate Intelligence; Conversational AI; Arabic AI; Property Search; Real Estate Data; Sales Automation; Real Estate Marketing; Digital Transformation; Business Intelligence.

---

1. Introduction

1.1 The problem

A household buying a home in Egypt's new-city and compound market usually assembles price information from many places: developer sales offices and brochures, national portals, brokerage teams, social media and messaging apps. The founder of one of Egypt's largest property technology companies described the market in 2025 as fragmented and opaque, with buyers approaching developers one at a time and "everyone … incentivized to push you one way or another" (Kene-Okafor, 2025). The Egyptian state has also framed transparency as a policy goal: in 2025 the Housing Ministry promoted an official Egyptian Real Estate Platform as a national hub for verified property data (Daily News Egypt, 2025), and in August 2026 the Cabinet announced a mobile application for property transactions and real estate tax services (Ahram Online, 2026). These are stated intentions, not measured outcomes; we found no independent study measuring how such tools have changed price discovery in Egypt.

Three features of the setting make the problem harder than "put the listings online":

1. **Asking prices are not transaction prices.** Most observable prices are developer or seller asking prices. Completed-sale prices are not systematically public.
2. **The same name refers to different things.** A compound name may cover several phases, unit types, finishing levels and payment plans. A price quoted "for project X" is rarely comparable to another quote for "project X".

3. **Prices are time-sensitive in an inflationary economy.** Egypt has been through a period of high inflation and monetary tightening; the IMF reported inflation of 11.9 percent in January 2026 after earlier highs (International Monetary Fund, 2026). An undated price quickly becomes misleading.

Economic theory predicts that such conditions raise search costs (Stigler, 1961), can degrade quality signals when one side knows more than the other (Akerlof, 1970), and let better-informed intermediaries capture value (Levitt & Syverson, 2008). Large language models add a new risk: a fluent assistant can state a plausible but false price, a well-documented failure mode called hallucination (Ji et al., 2023).

1.2 What this report is

This report is a **descriptive technical report with a proposed evaluation framework**. It documents what REONE is and how it works, as implemented in its source repository, and proposes how its claims could be tested. It does not claim the system has improved any market outcome. Where documentation and code disagree, the code is preferred and the text says "as implemented at commit 745652d".

REONE was built by one founder, using AI coding assistants extensively (most commit messages in the repository carry an AI-assistant co-author line). The repository's first commit is dated 4 September 2026; commit `745652d` is the 751st, dated 3 November 2026. This short, fast history matters for interpretation: the system is young, its data coverage is changing week to week, and several components are recorded by the project itself as "not proven" (Section 10).

1.3 Evidence conventions

Following REONE's editorial standards, claims carry one of five labels where relevant: **Observed** (as published by a named source on a stated date), **Verified** (checked against REONE's published methodology), **Calculated** (reproducible from stated inputs), **Estimated** (our estimate, assumptions stated) and **Forecast** (a scenario, never a fact). Every REONE price discussed is an **asking price**. This report types no live REONE price; the worked example in Appendix C uses invented, clearly hypothetical inputs.

1.4 Structure

Section 2 states the objectives and research questions. Section 3 describes the architecture. Section 4 covers the conversational advisors. Section 5 covers business model and sales technology, separating existing from proposed capabilities. Section 6 develops mechanisms and hypotheses. Section 7 compares the platform landscape. Section 8 proposes an evaluation framework. Section 9 sets out illustrative 2027 scenarios. Section 10 covers challenges, ethics and limitations, followed by discussion, conclusion, references, appendices, disclosure and version history.

---

2. Research objectives and questions

Objective. To document, at a level another engineer or researcher can inspect and criticise, how a data-grounded bilingual real estate system can be built for Egypt, and to define how its effects could be measured honestly.

The report addresses three research questions:

- **RQ1 — How does it work?** What architecture turns scattered, inconsistent asking-price observations into dated, source-attributed, confidence-labelled figures, and how are conversational advisors constrained to those figures?
- **RQ2 — How could it improve real estate business and sales workflows?** Which existing functions plausibly reduce work for buyers, advisors and sales teams (search, comparison, payment calculation, qualification, handoff), and which are only proposed?
- **RQ3 — Under what conditions could it materially change the Egyptian market?** What would have to be true (coverage, data quality, adoption, regulation, trust) for a system of this kind to change search costs, price dispersion or bargaining outcomes at market scale, rather than only for its own users?

RQ1 is answered descriptively from code (Sections 3–4, Appendices A, C, D). RQ2 is answered by a capability inventory (Section 5) and testable hypotheses (Section 6). RQ3 is answered only conditionally, through hypotheses, a proposed evaluation (Section 8) and illustrative scenarios (Section 9). None of the three is answered with measured outcomes.

---

3. System architecture

3.1 Overview

REONE has two surfaces over one governed database:

- a **public marketplace** (English and Arabic, including right-to-left layout) with search, maps, project/unit/developer/area pages, a Market Index, Market Watch (news), calculators, and the Nour and Zaki advisors; and
- a **staff control centre**, separate from the public site, for data sources, crawling, field mapping, manual review, price verification, catalogue editing, providers, voice profiles, leads, offers and automation.

The technology choices are deliberately small:

Layer	Implementation (as implemented at commit 745652d)
Frontend	Framework-free static HTML, CSS and JavaScript; numbered modules under one namespace; most SEO pages generated at build time by Python scripts
Database	PostgreSQL (managed by Supabase); 240 migration files in the repository
Access control	Row-level security on tables; role-based staff permissions; public reads through constrained views and functions
Server logic	Postgres functions plus 21 Deno Edge Function directories (Section 3.12)
Secrets	Provider API keys held server-side in an encrypted secret store and used only inside server functions; never sent to the browser
Hosting	Static hosting with build-time page generation and scheduled jobs

Appendix A gives the architecture diagram.

3.2 Data ingestion and source identification

Every source of price or catalogue data is a row in a `data_sources` table. A source has a type (developer website, portal, marketplace, broker website, listing website, official project, internal, external API, CSV, Excel, manual, custom), a **trust score** from 0 to 100 (default 30), a per-minute rate limit, an adapter key, an enabled flag and free-text notes used to record compliance evidence. Test sources are flagged and excluded from all public figures.

Ingestion paths in the repository include:

- a single-source crawl function that fetches only a source's own configured base URL, checks `robots.txt`, and records a job, logs and raw records;
- per-source adapters for unit research and new-launch synchronisation, which read structured data the source's own pages use, enforce a database-backed per-source request clock shared across function instances, and send anything without an exact canonical identity to manual review;
- an owner-authorised partner feed and offline import scripts; and
- RSS feeds of news publishers (headline, short excerpt and link only; article bodies are not fetched).

The architecture document requires, for every crawler, an explicit domain allowlist drawn from `data_sources`, rejection of private, loopback and cloud-metadata addresses, `robots.txt` compliance, and low concurrency with real delays.

Source identity is kept in the backend. Since 29 September 2026 public pages show each price's date and verification status but do not name the third-party source; source columns are readable by staff only. This report follows the same practice and describes sources by category. Section 10 discusses why the compliance status of some sources is an open matter.

3.3 Normalization

Raw payloads are stored in `raw_records` (source, URL, payload, crawl time, crawler version) and parsed into `normalized_records`. Three database functions do the core normalization (as implemented at commit 745652d):

- `normalize_text_key(text)`: lower-case, strip non-alphanumeric characters, collapse whitespace. Used as the matching key.
- `normalize_area_sqm(text)`: strip everything but digits and the decimal point.
- `normalize_currency_amount(text)`: recognises "N million"/"N m" forms and otherwise strips formatting, so that "EGP 8,500,000", "8.5 Million" and "8,500,000 LE" all yield 8,500,000. The currency itself is tracked in a separate column, not inferred.

A `field_mappings` table maps each source's external field names to canonical fields (developer, project, unit, payment plan, offer) with transform rules. Later additions normalise domain vocabulary; for example, triggers map finishing labels from several sources and both languages to one canonical bilingual set (fully finished, semi-finished, core and shell, fully finished and furnished, super lux).

A limitation is visible in the code: `normalize_text_key` keeps only `a–z` and `0–9`, so Arabic-script names produce an empty key and cannot match by exact key; Arabic names rely on aliases and on English names. This is a documented gap, not a hidden one.

3.4 Entity matching and duplicate handling

Matching links an observation to a canonical developer or project, a modern instance of the record-linkage problem (Fellegi & Sunter, 1969). `match_developer_candidates` and `match_project_candidates` return ranked candidates using the cheapest-first strategy set out in the architecture document (as implemented at commit 745652d):

Method	Confidence assigned
Exact normalized-name match	99
Alias table match	the alias row's stored confidence
Fuzzy trigram similarity (above 0.35 for developers, 0.30 for projects)	similarity × 100
Project co-signals	+8 if the candidate's developer matches, +5 if its destination matches, capped at 100

Methods are ranked exact > alias > fuzzy before confidence, because a 2026-09-04 bug fix found that a row's perfect trigram self-similarity was outranking its own exact match and mislabelling the method.

Routing uses thresholds stored in `system_settings` rather than code: matches at or above an auto-approval level (default 98) can be linked; matches in a review band (from 70) go to a

`manual_reviews` queue that staff resolve in a grouped review screen; anything lower is recorded as unmatched. Nothing below the threshold is merged automatically. Matched aliases are stored in `entity_aliases` so the next observation of the same name resolves deterministically.

Duplicates are handled at several levels: a database rule treats the same project, normalized unit name/type and price as one unit for an owner-authorised feed; in verification, the same source repeating the same price counts as one observation (Section 3.6); a findings-only backend agent lists likely duplicate projects for staff; and the project documentation records that observation counts are not distinct-unit counts.

3.5 Provenance and observation timestamps

Price evidence lives in `price_records`, documented in the schema as "immutable event data — one row per observed price per source per capture". Each row references exactly one unit or one project and records the source, source URL, crawl job, price, currency, area, a generated price per square metre, unit facts at the time (type, bedrooms, bathrooms, floor), payment terms (down payment, instalment years and amount, maintenance, discount, offer text), delivery date, availability, match confidence, **the source's own update date** and **the crawl date**.

Append-only behaviour is enforced by a trigger, not only by convention (as implemented at commit 745652d): the only permitted update is a retraction that sets a retraction time, the retracting user and a non-empty reason; every other column must remain identical, and a retraction cannot be edited or undone. Retracted rows and test-source rows are excluded from public reads and from verification. Staff actions write an `audit_logs` row.

This design follows from a recorded incident. On 16 September 2026 the project found that 142 observations had been approved through manual review into the wrong projects, and that the then verification function counted repeated records from one listing as many "sources", producing a confident "verified" figure far below the catalogue's real starting price. The fix introduced retraction with reasons, test-source exclusion, and counting of independent sources (Section 3.6). The incident is a useful reminder that provenance alone does not guarantee correctness; it makes errors traceable and reversible.

3.6 Verification and freshness rules

`recompute_verified_price` produces one current row per unit or project in `verified_prices`. It is never a plain average. The algorithm, reproduced in full in Appendix C, is (as implemented at commit 745652d):

1. **Usable evidence**: not retracted, not from a test source, match confidence at or above the review floor (default 70).
2. **De-duplication**: one observation per (source, price) pair, keeping the most recent.

3. **Outliers**: compute the median price and the median absolute deviation (MAD); an observation is an outlier if its absolute deviation exceeds 3 × MAD. Outliers are excluded from the estimate but kept and listed. MAD-based rules are a standard robust alternative to standard-deviation cut-offs (Leys et al., 2013).
4. **Recency weight** by age of the observation (source update date if known, else crawl date): 1.00 (same day), 0.95 (≤7 days), 0.80 (≤30), 0.50 (≤90), 0.20 (older).
5. **Weight** = (source trust / 100) × recency weight × (match confidence / 100). Verified price = weighted mean of non-outlier prices.
6. **Confidence (0–100)** = 40 × min(1, independent sources / 4) + 25 × max(0, 1 − spread% / 15) + 20 × (mean trust / 100) + 15 × freshest recency weight, where spread% = (max − min) / median × 100 over non-outliers.
7. **Status ladder**, first match wins: fewer independent sources than the minimum (default 2) → `insufficient_sources` and **no published price**; freshest observation older than 180 days → `outdated`; any outlier → `major_price_difference`; otherwise `verified` (≥95), `high_confidence` (≥85), `medium_confidence` (≥70) or `needs_review`.

The status thresholds, match floor and minimum source count are configurable settings; the recency bands, the 3 × MAD rule, the 180-day cut-off and the confidence weights are fixed in code. Two statuses defined in the schema, `conflicting` and `potential_error`, are documented as reserved and not produced by the current function.

3.7 Search and filtering

Public search runs as database functions callable anonymously and constrained by row-level security: unit search with filters for destination, type, price, bedrooms, delivery, down payment, instalment years, unit size, finishing and listing type (developer sale or resale), a map search, and a recommendation function. The recommendation ranking, as documented, rewards use of the stated budget, prefers requested bedroom counts, shows at most one unit per project in the top three, and excludes from recommendations (but not from search) units priced below 40 percent of the destination-and-type median price per square metre, a guard against mis-scraped prices. A trigger-maintained keyword index supports bilingual text search. A feature flag lets staff hide resale inventory from the interactive surfaces; generated SEO pages deliberately keep the same content for all visitors to avoid serving search engines something different from people.

3.8 Market Index calculation

The Market Index reports **median asking price per square metre** by destination and property type, separately for developer sales and resale (as implemented at commit 745652d, function `get_market_dashboard`):

1. Take the latest EGP price evidence per published, available unit in a published project and destination.
2. Compute each project's median price per square metre.

3. Mark a unit **eligible** only if its price per square metre lies between 0.5× and 2× its own project's median. Ineligible units stay listed but are excluded from every aggregate.
4. Publish a median only when at least **five** eligible units back it; otherwise return no figure.

The public page flags destinations with fewer than 30 priced units as small samples, states that supply counts describe REONE's coverage rather than Egypt's housing stock, and says that medians are asking prices, not sales. A "verified asking-price ranges" panel reads from independently corroborated evidence and also requires five distinct units per row. A rental-yield/NPV/ROI/IRR calculator on the same page uses editable modelling assumptions and states that rental data is not available in the catalogue.

3.9 Financial and instalment modelling

Three deterministic browser engines, each with automated tests in the repository, do all arithmetic. No language model performs calculations.

- **Payment plan** (`payment-model.js`): Egyptian developer plans are modelled as interest-free; price minus down payment (and minus an optional payment on delivery) is split into equal instalments at a chosen frequency; maintenance is a separate deposit on top of the price. The calculator loads terms on record for a unit, project or area and labels each by origin ("verified by REONE", "from the listing", "developer offer", "seen on units in this project", "most common in the area").
- **Mortgage** (`mortgage-model.js`): reducing-balance amortisation over published lender products whose limits and source links are coded with a check date (2026-10-10 in the code); commercial rates require a user-entered quote.
- **Investment** (`investment-model.js`, formula version `reone-invest-1.0.0`): dated cash flows, NPV and XIRR (actual/365), comparing property with gold, an equity benchmark and bank certificates. Any input the caller does not supply is reported as missing, never filled with a hidden default.

3.10 AI tool invocation and answer grounding

The advisors follow a tool-use pattern (Schick et al., 2023; Yao et al., 2023) that resembles retrieval-augmented generation (Lewis et al., 2020; Gao et al., 2023) but retrieves **structured rows** rather than text passages. A language model decides which tool to call; the tool queries the database through the same anonymous, row-level-secured functions the website uses; the model narrates the returned values.

As implemented at commit 745652d, static inspection finds 13 function declarations in the text-chat server function and 18 in the shared voice prompt module, plus two "screen" tools shared by both channels. They include `search_catalog`, `recommend_units`, `get_payment_terms`, `calculate_property`, `compare_units`, `get_project_features`, `get_market_index`, `get_market_trends`, `get_system_knowledge`, `get_price_analysis`, `get_investment_analysis`, `compare_investment`, `read_page`, `recall_conversation`,

`capture_lead`, `prepare_contact`, `end_session`, `extend_session`, `show_on_map` and `open_page`.

Grounding is enforced by several layers, none of which alone is sufficient:

- **Instruction layer.** Shared prompts forbid stating any price, availability, payment term or delivery date not returned by a tool, and forbid any figure in a "let me check" bridge before the tool result arrives.
- **Data layer.** Tools return only published rows; market figures come from the Market Index and verified views; finishing is "not recorded" when null.
- **Rendering layer.** Every money field in a tool result is given a deterministic spoken-word sibling in English and Egyptian Arabic, so the voice model reads out words computed by code rather than composing numbers.
- **Navigation layer.** `open_page` builds links only from references a tool returned, checks that the page exists, and stays on the same site.
- **Untrusted page text.** `read_page` passes page content to the model fenced as untrusted data.

These measures reduce, but cannot eliminate, the chance of an ungrounded statement; the project's own notes record a case where an advisor described the cheapest of three returned villas as "the cheapest villa available" in an area, an over-generalisation the tools did not support. Measuring the residual rate is the first proposed evaluation (Section 8, E1).

3.11 Bilingual conversational interaction, human handoff and analytics

These are covered in detail in Sections 4 and 5. In brief: one conversation spans text and voice; advisors speak Egyptian Arabic or English; staff receive leads with a deterministic summary; and analytics are limited to consented tracking tags, property-intent signals that reject contact fields, a lead-capture funnel view, and lead-capture outcome counts. No analytics results are reported here.

3.12 Server functions and background agents

The repository's Edge Functions (as implemented at commit 745652d) fall into five groups:

Group	Functions (purpose)
Secrets and providers	provider secret handling (masking, connection tests)
Data acquisition	single-source crawl; new-launch sync; unit research adapters; news RSS ingestion; bank-rate research (proposals only, stored for human approval); daily gold collector
Conversation	text-chat brain; Gemini Live and OpenAI Realtime ephemeral-session minting; two legacy gateway adapters kept for versioning
Leads and privacy	lead capture; minimal contact capture; phone validation (computation only); buyer-intent signals; data-subject request intake
Audit and publishing	SEO audit of the deployed site; UX/accessibility audit of REONE's own pages; research worker; social publishing

Background "agents" are scheduled database or server jobs registered in an agent registry. They are fenced by design: findings-only agents write to a findings table and never change units, projects or verified prices (examples: price anomalies against destination medians, stale listings, catalogue completeness); financial observations pass a state machine (collected → normalized → verified or review_required or rejected → published → stale) in which published rows are immutable and a jump of more than 10 percent is routed to human review. The public agent page states that twelve agent roles exist and labels which are working; this report did not verify the live registry.

---

4. Nour and Zaki: conversational capabilities

4.1 What they are

"Nour" and "Zaki" are two named advisor personas stored as rows of a `voice_profiles` table (name, provider, model, voice, language, personality instructions, turn detection). Renaming a profile renames the advisor without a code change. The public assistant lists only profiles that are marked realtime and whose provider is enabled and holds a key; visitors can choose either advisor and the choice persists in the browser. Both use the same tools and data.

4.2 Channels and providers

- **Text chat.** The browser sends the conversation to a server function that adds the persona and tool declarations, calls a language model with a server-held key, and returns text and/or tool calls; the browser runs the tools and loops (up to six steps). If the AI call fails, an older rule-based router answers instead. As implemented at commit 745652d the text path is coded for Google Gemini models with an ordered fallback list and multiple keys with automatic failover (models configured at the time of writing; staff can change the preferred model in settings).

- **Live voice.** The browser never receives a provider key. A server function mints a short-lived, constrained client token (Gemini Live) or client secret (OpenAI Realtime) locked to the profile's model, voice, instructions and tools; the browser then connects directly to the provider. Which provider is used is a per-profile setting. Project documentation records that on 25 October 2026 the default public profile was switched to OpenAI Realtime, with Gemini one setting away; this report did not inspect the live database, so the provider in use on any given day should be treated as configuration, not architecture.
- **Voice behaviour.** Calls have a configurable time limit (15 minutes in documentation), a visible countdown, reconnection on provider-initiated session handover, a jitter buffer for playback, and server-side voice-activity detection with barge-in. Languages are constrained to Arabic (Egypt) and English (US) locales. Voice access can be configured to require a Google sign-in.
- **On-screen actions.** Advisors can open the map over the conversation with a project pinned, or place a button to a unit, project, developer or area page or a filtered catalogue; during a live call the call keeps running.

4.3 Bilingual interaction

Arabic is written and spoken for Egyptian users, not translated: the prompts ask for Cairene Arabic; money is spoken through code-generated Egyptian-Arabic words; catalogue names, unit types and finishing labels have Arabic equivalents in the database; and the site serves an indexable `/ar` version with right-to-left layout. Arabic natural-language processing remains comparatively under-resourced, particularly for dialects (Darwish et al., 2021), which is one reason REONE keeps arithmetic and data retrieval outside the model.

4.4 Missing-data behaviour

The advisors are instructed, and the public methodology states, that when the data does not contain a figure they say so rather than estimate it. Concretely: a unit with no recorded finishing returns null and is described as not recorded; a verified price with fewer than two independent sources is not published (`insufficient_sources`); a Market Index cell with fewer than five eligible units returns no median; the investment engine reports missing inputs by name; calculator input errors return a structured "input needed" result. Rental data, transaction prices and completed-sale histories are not in the catalogue, and the advisors are told not to present asking prices as sale prices.

4.5 Microphone-permission invariant

A product invariant states that microphone permission is never requested on page load or when the voice dialog opens, only at the moment the user presses "start", with a visible indicator while recording is live. **Verified by code inspection at commit 745652d:** the only call to `getUserMedia` in the public JavaScript is inside the voice controller's `start()` function, after an access check (and sign-in, if required) and after the dialog has announced "requesting permission". The site-wide assistant frame is loaded only on first click and grants

the microphone feature to the frame, but does not request it. Whether the invariant holds on every mobile browser has not been tested with real devices (Section 10).

4.6 Memory and the shared thread

Text and voice share one conversation. As implemented at commit 745652d, when a visitor first writes, the browser creates an anonymous authenticated identity and stores chat and voice turns in `conversation_threads` and `conversation_turns`. Row-level security limits each visitor to their own rows; turns are append-only for their owner, size-limited, and a scheduled job deletes threads and profiles idle for 90 days. A "forget me" function deletes the caller's memory on request. Project documentation also records an open decision on durable server-side persistence of voice transcripts in other contexts; the report treats retention and consent as an ethical issue in Section 10.

4.7 Human oversight

Human oversight appears at four points: (1) all below-threshold entity matches go to staff review; (2) financial and rate figures proposed by research agents are stored as "review required" and published only after staff approval; (3) leads are handed to human advisors, who own every sales decision; (4) every staff change is audit-logged. The advisors are instructed to say they are REONE's AI advisor if sincerely asked whether they are human, and public pages tell users to confirm price, availability and terms with the developer before any reservation.

4.8 Limitations recorded by the project

The project's own documentation lists as **not proven**: a real spoken call crossing a provider session handover, lead capture during a real spoken session on every provider, mobile voice on iPhone Safari and Android Chrome, acceptance of language hints on the production model, barge-in quality for slower Arabic speakers, and how natural the voice sounds. Server-side truncation of unplayed audio after barge-in is recorded as not implemented on one provider path. Provider quotas have caused at least one recorded outage when a model without quota was selected. No user-facing latency figures are reported in this paper; internal engineering timings exist in project notes but were not collected under a protocol and are not results.

---

5. Business model and sales technology

5.1 Position

REONE presents itself as a real estate intelligence platform with AI advisors and human consultants. It does not lend money, does not publish completed-sale prices, and is not affiliated with the developers or portals it compares. Its public comparison describes its intended use as "checking whether a price is fair before you talk to a seller".

The repository contains no payment, commission, subscription or invoicing system, and no contract with any developer, broker, bank or portal is evidenced in the code. **Revenue mechanisms are therefore not described as existing.** The last row of the table below lists proposals only.

5.2 Existing functions versus proposed capabilities

Area	Existing (in code at commit 745652d)	Proposed (not built; for evaluation or future work)
Discovery	Bilingual search and filters; map; area, project, developer and unit pages; new launches; offers landing pages	Saved-search alerts on price changes (price-alert table designed in architecture, not built)
Price transparency	Dated asking-price evidence; verified-price status; Market Index with sample gates; public methodology	Published source-level transparency view for users (architecture Phase 5; currently staff-only by design)
Conversational advice	Nour and Zaki by chat and live voice; data tools; map and page actions; shared memory with 90-day retention	Measured grounding audits; human review of sampled transcripts (Section 8)
Calculation	Payment-plan, mortgage and investment calculators; NPV/XIRR/IRR; CSV export	Lender pre-qualification or application (not offered; would need regulatory review)
Lead capture	Conversational capture after three meaningful questions; server-side phone normalisation and deduplication; deterministic lead summary; contact and callback forms; overseas-buyer brief	Consent re-approach policy (open owner decision)
Sales handoff	Staff Leads screen with readiness indicator (distinct from AI score), claim, owner, next action, notes, due filters; user-opened WhatsApp link with receipt	CRM integration; appointment booking; response-time commitments
Learning	Thompson-sampling selection among honest wordings of the contact request, with outcome logging and no contact details in the learning table	Controlled experiments on advisor behaviour (Section 8)
Marketing	Offers wizard with bilingual landing pages and popup; SEO page generation; social publishing connector (recorded in development mode)	Paid acquisition measurement; attribution model
Data operations	Source registry, adapters, manual review, retraction, audit log, findings-only agents, automation watchdog	CSV/Excel bulk import of price evidence (documented as not built); external worker fleet at scale
Monetisation	None in code	Broker or developer referral fees; data or index licensing; advisor subscriptions — all hypothetical, none agreed

5.3 How the system could support sales workflows (RQ2)

Three workflow effects are plausible, and each is testable:

1. **Pre-qualified conversations.** A lead arrives with an area, budget, type, payment preference, shortlist and conversation summary already assembled. If accurate, this shortens the first human call.
2. **Fewer price disputes.** If the buyer and advisor look at the same dated, sourced figure and its verification status, disagreements shift from "what is the price" to "is this unit right for me".
3. **Consistent bilingual first response at any hour.** The advisors answer in Arabic or English when no human is available, then hand over.

Each could also go wrong: a confident advisor could anchor buyers on a stale asking price; automated contact requests could feel pushy; and staff could over-trust an AI summary. Section 8 proposes measurements for both directions.

---

6. Buyer behaviour and market efficiency

6.1 Mechanisms

We consider five mechanisms through which a system like REONE could affect buyer behaviour and, if adopted widely enough, market efficiency.

- **M1 — Search-cost reduction.** Consolidating projects, units, payment terms and benchmarks in one place, with conversational access, lowers the cost of each additional comparison (Stigler, 1961). Lower search costs predict more comparisons per buyer and, at scale, lower price dispersion for comparable units.
- **M2 — Information-asymmetry reduction.** A dated benchmark and a visible verification status reduce the information advantage of the better-informed party (Akerlof, 1970; Levitt & Syverson, 2008). This predicts better-calibrated buyer expectations and fewer abandoned negotiations caused by mismatched expectations.
- **M3 — Comparability through canonicalisation.** Matching observations to canonical units and projects, and labelling finishing and payment terms, makes like-for-like comparison possible. Without it, a lower quote may simply be a different product.
- **M4 — Platform competition.** Evidence from other housing markets shows that different listing platforms can produce different outcomes for sellers (Hendel et al., 2009), and that information technology changes the roles of agents rather than removing them (Sawyer et al., 2005; Baum, 2017). This predicts complementarity with human advisors, not substitution.
- **M5 — Trust through uncertainty disclosure.** Showing "insufficient sources" instead of a number could raise long-run trust even if it lowers short-run engagement.

6.2 Testable hypotheses

ID	Hypothesis	Primary measure (see Section 8)
H1	Advisor answers that contain a price, availability or payment figure match the tool output for that turn in at least a pre-registered share of sampled turns, in both Arabic and English.	Grounded-figure rate (E1)
H2	The Arabic grounded-figure rate is not lower than the English rate by more than a pre-registered margin.	Language gap in E1
H3	Verified prices with status `verified` or `high_confidence` are closer to independently obtained current asking quotes than prices with lower statuses.	Absolute percentage error by status (E2)
H4	Users who view a Market Index benchmark before contacting an advisor state budget expectations closer to the benchmark than users who do not.	Expectation gap (E5)
H5	Leads that arrive with an AI-assembled brief require less advisor time to reach a qualified state than form-only leads, without a lower subsequent qualification rate.	Time-to-qualification (E4)
H6	Showing an explicit "insufficient sources" state does not reduce the share of sessions that proceed to a comparison or calculator.	Progression rate (E3)
H7	Contact-request wordings chosen by the learning policy do not increase complaint, opt-out or "decline" rates relative to a fixed honest wording.	Harm metrics in E6
H8	At market scale, destinations with higher REONE coverage show lower dispersion of asking price per square metre for comparable unit types over time. (Observational; causal identification not possible without external variation.)	Dispersion trend (E8)

H1–H7 concern REONE's own users and can be tested by the operator. H8 concerns the market and can only be tested observationally, with the caveat that coverage and dispersion may both be driven by a third factor such as developer marketing cycles.

---

7. Competitive landscape

The table below compares features that each company shows on its own public website. It reproduces the comparison REONE recorded on **27 October 2026** in its build scripts, in which a dash means "not found on the company's own website at the check date" and not "does not have". For this report, the Nawy and Aqarmap homepages were re-checked on **3 November 2026**; the observations were consistent with the table (Aqarmap shows a prices guide by area, a real estate index and an AI property estimate; Nawy shows its financing product, a fractional-ownership product and an "800+ compounds" claim, and no visible AI assistant or price-per-metre index). The other cells were not re-checked for this report. Company counts

are each company's own published figures. No ranking or market-share claim is made; names and trademarks belong to their owners and REONE is not affiliated with any of them.

Feature (as shown on each company's own site)	REONE	Nawy	Aqarmap	Property Finder (Egypt)	Bayut (Egypt)	Coldwell Banker Egypt
What it is	Real estate intelligence platform: listings with dated price evidence and AI advisors	Proptech marketplace with its own brokerage team	Property classifieds portal for owners, brokers and developers	Portal for agent and developer listings	Portal, part of Dubizzle Group, in Egypt since 2023	Brokerage and consultancy with branch offices
Inventory focus	Compounds, units for sale, resale, new launches	Primary and resale; "800+ compounds" (company figure)	Sale and rent; compounds; "1,200+ projects" (company figure, 27 Oct)	Sale, rent, new projects, commercial	Sale, rent, new projects	Primary, resale, rent, commercial; "400+ compounds" (company figure)
Price per m² by area	Median asking price per m² by area and type, with date and outlier exclusion	—	Prices Guide by area	—	—	—
Shows how a price was checked	Evidence with source trust weighting; public methodology	—	—	—	—	—
Market data and news	Market Index; Market Watch headlines	Blog; offers page	Real estate demand index; research	—	Blog	Market insights; blog; media room
AI help	Chat and live-voice advisors (Arabic/English) answering from catalogue data	—	AI property value estimate	—	—	—
Financing	Calculators and mortgage guide; does not lend	Home financing product (company describes it	—	—	—	—

Feature (as shown on each company's own site)	REONE	Nawy	Aqarmap	Property Finder (Egypt)	Bayut (Egypt)	Coldwell Banker Egypt
		as FRA-licensed)				
Who the buyer talks to	REONE consultants	Company sales team	Listing owner or broker	Listing agent	Listing agent	Branch consultants

Sources: each company's public website as listed in the References (Nawy, Aqarmap, Property Finder Egypt, Bayut Egypt, Coldwell Banker Egypt), checked 27 October 2026 by REONE and, for Nawy and Aqarmap, 3 November 2026 by the author. Public-sector initiatives (the official Egyptian Real Estate Platform and the Cabinet's transaction app) are relevant context but are not compared feature by feature because their public feature sets were not checked for this report (Daily News Egypt, 2025; Ahram Online, 2026).

Two observations follow without implying superiority. First, REONE's distinctive combination, in this snapshot, is a published price-verification methodology plus advisors constrained to database tools; Aqarmap also publishes area price guides and an AI estimate, and Nawy operates at a far larger funded scale (Kene-Okafor, 2025). Second, several competitors operate brokerage teams and financing products that REONE does not have.

---

8. Evaluation framework (proposed)

No results yet. Nothing in this section has been run. It is a protocol for future work, written before data collection so that it can be held to.

8.1 Principles

- Pre-register metrics, thresholds and analysis plans before looking at outcome data.
- Report denominators and exclusions with every rate.
- Separate **controlled** evaluations (fixed prompts, fixed data snapshot, scripted test users, no real leads written) from **real-world** observation (live traffic, consented analytics).
- Test both languages separately.
- Never write synthetic leads into production tables; where a test needs the lead path, mark and exclude the rows, as the project already does for learning-table test rows.
- Use online controlled-experiment practice for any A/B test (Kohavi et al., 2020), and frame AI risks with an established structure such as the NIST AI Risk Management Framework (National Institute of Standards and Technology, 2023).

8.2 Evaluations

ID	Question	Metric and denominator	Baseline	Sampling	Window	Setting
E1	Are advisor figures grounded?	Grounded-figure rate = turns whose every price/area/term/date figure matches that turn's tool output ÷ turns containing at least one such figure	Same prompts answered by the same model **without** tools (ablation)	Stratified prompt set: ≥200 prompts per language across search, payment, market, comparison, missing-data and adversarial cases	One fixed data snapshot	Controlled
E2	Are verified prices close to reality?	Median absolute percentage error between `verified_price` and an independently obtained current asking quote, by status	Unweighted mean of the same evidence; latest single observation	Random sample of units stratified by status and destination	Quotes collected within 14 days of the snapshot	Controlled (field collection)
E3	Do users progress?	Share of sessions reaching a comparison, calculator or advisor answer containing a grounded figure ÷ sessions with a search	Pre-launch period of the same pages, or a holdout	All consented sessions	8 weeks	Real-world
E4	Does the brief help advisors?	Median minutes from lead creation to qualified status; qualification rate ÷ leads contacted	Form-only leads in the same period	All leads, excluding test and spam	12 weeks	Real-world
E5	Are expectations better calibrated?	Absolute gap between a user's stated budget per m² and the relevant Market Index median	Users who did not view the index before contact	Leads with both values present	12 weeks	Real-world (observational)

ID	Question	Metric and denominator	Baseline	Sampling	Window	Setting
E6	Is contact-request learning harmless?	Decline, opt-out and complaint rates ÷ asks; data-request submissions ÷ leads	Fixed single honest wording (randomised holdout share)	All asks	Until a pre-registered sample size	Real-world, randomised
E7	Is voice usable and safe?	Session completion ÷ started sessions; microphone requested before "start" (must be zero); device/browser breakdown	Text chat for the same task set	Scripted tasks on named devices; consenting testers only	2 weeks	Controlled
E8	Does coverage relate to dispersion?	Interquartile range of asking price/m² for comparable unit types per destination	Earlier snapshots	Monthly destination-level aggregates	12 months	Real-world (observational)

8.3 Data-quality audits supporting all evaluations

- **Match audit:** sample auto-linked matches each month and report precision with a confidence interval; sample manual-review decisions for inter-reviewer agreement.
- **Freshness audit:** distribution of observation age for published figures.
- **Coverage audit:** share of published units with area, finishing, payment terms and delivery date recorded.

Protocols are detailed in Appendix B.

---

9. 2027 business-impact scenarios (illustrative)

These scenarios are **Forecasts** in the sense of Section 1.3: possible futures, not predictions, and not commitments. They are written as conditions rather than numbers, because no measured baseline exists.

Dimension	Conservative	Base	Optimistic
Data coverage	Coverage stays concentrated in a few sources; many verified prices remain `insufficient_sources`	Coverage broadens through authorised feeds and staff entry; most priced units in major destinations have two or more independent sources	Developers or the official platform provide authorised feeds; coverage becomes market-representative in major destinations
Compliance	One or more sources must be withdrawn after terms review, shrinking coverage	Source agreements or permissions are documented for the main sources	Formal data partnerships exist
Advisor quality	E1 shows grounded-figure rates below threshold in Arabic; voice limited to chat-first use	E1 meets thresholds in both languages; voice used for a minority of sessions	E1 and E7 meet thresholds on mobile; voice becomes a common entry point
Sales effect	Leads are few; human follow-up capacity is the bottleneck	Briefs measurably shorten qualification (H5 supported)	Briefs and benchmarks measurably improve qualification and reduce disputes
Market effect	None detectable	Effects limited to REONE's own users	H8 shows a consistent association between coverage and lower dispersion, prompting further causal study
Key assumptions	Funding and staffing remain at founder scale; provider costs constrain voice	Founder plus a small advisory team; voice costs controlled by quotas	Partnerships, staff and provider costs all favourable

Illustrative arithmetic (not a projection). If a period has S sessions with a search, a progression rate p (E3), a lead rate l per progressing session, and a qualification rate q per lead, then qualified leads = $S \times p \times l \times q$. Any business case for 2027 must replace each symbol with a measured value from Section 8; until then, no revenue or lead forecast is offered.

Conditions for material market change (RQ3). The scenarios suggest that REONE alone changes the Egyptian market only if (a) its coverage becomes representative, which depends on authorised data rather than engineering; (b) its grounding is measured and holds in Arabic; (c) buyers and advisors actually use the benchmarks in negotiations; and (d) the regulatory environment for data and AI, including Egypt's personal data protection regime (CMS, 2026), permits the necessary processing. Without (a), the most plausible effect is a better-informed minority of buyers.

---

10. Challenges, ethics and limitations

10.1 Data rights and source compliance

At least one third-party source's published terms restrict automated access without written permission, and another signals that its content should not be used as AI input. The project records owner-level permissions and risk-acceptance decisions and keeps the matter as an open decision; a database-enforced source compliance gate exists as a migration file but is recorded as not applied. Some third-party listing images are served through a same-origin caching path after an owner decision; image rights are not independently cleared. These are material legal and ethical risks that this report does not resolve.

10.2 Asking prices are not market values

All REONE prices are asking prices. Verified prices measure agreement among asking-price observations, not transaction values. Treating them as valuations would mislead.

10.3 Trust scores are judgement

Source trust scores are set by staff with justification notes; they directly weight verified prices. They have not been calibrated against outcomes (E2 proposes how).

10.4 Persuasion and lead capture

The contact-request learning policy optimises for obtaining a mobile number. The implementation constrains wording to be honest (sharing is optional; no promised callback time, price or reservation) and never repeats an approach to the same visitor, but optimising persuasion is a known route to manipulative design (Mathur et al., 2019). Bandit methods such as Thompson sampling (Russo et al., 2018) are efficient precisely because they shift exposure toward what works; E6 and H7 are therefore necessary safeguards, not optional extras. Whether a visitor who declines may be re-approached later is recorded as an open owner decision on consent.

10.5 Privacy and retention

Conversation turns are stored under anonymous or Google-linked identities for up to 90 days, with self-service deletion; leads contain personal data handled by server functions; analytics tags are gated by a consent manager; a data-subject request form exists. Egypt's Personal Data Protection Law (Law No. 151 of 2020) has executive regulations issued by Ministerial Decision No. 816 of 2025, with a compliance period that commentators place at or near 1 November 2026 (CMS, 2026). A formal legal review of REONE against those regulations is not evidenced in the repository and should be treated as outstanding.

10.6 Model and provider dependence

Advisors depend on third-party model providers whose models, quotas and prices change; the project has recorded an outage caused by selecting a model without quota. Provider-agnostic configuration mitigates but does not remove this risk. Model behaviour can drift between versions without code changes.

10.7 Engineering and organisational limits

The system was built in two months by one founder with AI coding assistance. The project records concurrent-session conflicts, known-broken test harnesses, and several features verified only in local or synthetic conditions. Real-device voice testing, real spoken calls and signed-in staff flows are partly unverified. There is no independent security audit referenced in the repository; security is described here only at the level of principles (row-level security, server-held secrets, origin and rate limits, audit logging).

10.8 Limitations of this report

The author is the founder (see Disclosure). The description is based on static inspection of code and project documentation; no live database was queried for this report, so live counts, current provider settings and the current state of open decisions may differ. Competitor features were checked on stated dates and may have changed. External literature on Egyptian real estate information problems is thin; claims about the Egyptian market rely partly on journalism and official statements rather than peer-reviewed studies.

---

11. Discussion

The central design choice in REONE is to keep **numbers out of the language model**. Prices, areas, instalments, medians, yields and spoken numerals are computed by database functions or deterministic code, and the model's job is to choose tools and narrate. This is a narrower and more testable promise than "an AI that knows the market". It also moves the hard problems to places where they can be inspected: entity matching thresholds, source trust scores, recency bands, outlier rules, sample-size gates.

The September 2026 data-integrity incident illustrates both the value and the limits of that move. Append-only provenance made the wrong figures traceable and reversible, but it did not prevent them: a human review interface that hid which project an approval linked to, combined with a verification function that counted repeated records as independent sources, produced a confident and wrong "verified" price. Grounding an AI in a database transfers trust to the database; the database then needs its own audits (Section 8.3).

For sales teams, the most immediate value is likely to be mundane: a lead that arrives with a structured brief, a shortlist and a summary composed deterministically from stated facts.

Whether that shortens qualification is an empirical question (H5) that a small operator can answer within a quarter.

For the market, the honest answer to RQ3 is conditional. Information tools reduce asymmetry only for those who use them, and benchmarks discipline prices only when they are representative and trusted. Official initiatives such as the Egyptian Real Estate Platform point to a possible future in which authorised, structured data is more widely available (Daily News Egypt, 2025); a system like REONE would benefit from such data far more than from additional crawling.

---

12. Conclusion

REONE AI is a working, bilingual real estate intelligence and sales technology system for Egypt whose architecture is designed around dated, source-attributed, append-only asking-price evidence; thresholded entity matching with human review; a transparent weighted verification formula; sample-gated market medians; deterministic financial engines; and conversational advisors that obtain every figure from database tools and request the microphone only on explicit user action. The report has described these components from code, separated existing from proposed capabilities, offered eight testable hypotheses, and proposed an evaluation framework and illustrative scenarios.

It has not shown that the system is accurate, fast, persuasive, profitable or market-changing. Those are open questions. The next version of this report should contain the results of at least E1 (grounding) and E2 (verified-price error), run under the protocols in Appendix B, ideally with an independent reviewer.

References

Ahram Online. (2026, August 19). *Egypt launches real estate mobile app to streamline transactions, expand digital tax services*. https://english.ahram.org.eg/News/574894.aspx (accessed 2026-11-03)

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics, 84*(3), 488–500. https://doi.org/10.2307/1879431

Aqarmap. (n.d.). *Aqarmap* [Website]. https://aqarmap.com.eg/en/ (accessed 2026-11-03)

Baum, A. (2017). *PropTech 3.0: The future of real estate*. Saïd Business School, University of Oxford. https://www.sbs.ox.ac.uk/news/property-tech-30-ground-breaking-report-looks-future-real-estate (accessed 2026-11-03)

Bayut Egypt. (n.d.). *Bayut* [Website]. https://www.bayut.eg (feature check recorded 2026-10-27)

CMS. (2026, January 13). *Egypt's PDPL: Executive regulations issued – one year compliance countdown begins* (M. Ooijevaar & B. Gibson). https://cms.law/en/are/legal-updates/egypt-s-pdpl-executive-regulations-issued-one-year-compliance-countdown-begins (accessed 2026-11-03)

Coldwell Banker Egypt. (n.d.). *Coldwell Banker Egypt* [Website]. https://coldwellbanker-eg.com (feature check recorded 2026-10-27)

Daily News Egypt. (2025, November 17). *Official Egyptian Real Estate Platform set to oversee 400,000 new units as digital transformation accelerates*. https://www.dailynewsegypt.com/2025/11/17/official-egyptian-real-estate-platform-set-to-oversee-400000-new-units-as-digital-transformation-accelerates/ (accessed 2026-11-03)

Darwish, K., Habash, N., Abbas, M., Al-Khalifa, H., Al-Natsheh, H. T., Bouamor, H., Bouzoubaa, K., Cavalli-Sforza, V., El-Beltagy, S. R., El-Hajj, W., Jarrar, M., & Mubarak, H. (2021). A panoramic survey of natural language processing in the Arab world. *Communications of the ACM, 64*(4), 72–81. https://doi.org/10.1145/3447735

Fellegi, I. P., & Sunter, A. B. (1969). A theory for record linkage. *Journal of the American Statistical Association, 64*(328), 1183–1210. https://doi.org/10.1080/01621459.1969.10501049

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2023). *Retrieval-augmented generation for large language models: A survey* (arXiv:2312.10997). arXiv. https://doi.org/10.48550/arXiv.2312.10997

Hendel, I., Nevo, A., & Ortalo-Magné, F. (2009). The relative performance of real estate marketing platforms: MLS versus FSBOMadison.com. *American Economic Review, 99*(5), 1878–1898. https://doi.org/10.1257/aer.99.5.1878

International Monetary Fund. (2026, February 26). *IMF Executive Board completes the fifth and sixth reviews under the extended arrangement under the Extended Fund Facility and first review under the Resilience and Sustainability Arrangement for Egypt* (Press Release No. 26/064). https://www.imf.org/en/news/articles/2026/02/26/pr-26064-egypt-imf-completes-5th-and-6th-revs-under-ext-arrange-under-eff-and-1st-rev-under-rsa (accessed 2026-11-03)

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), Article 248. https://doi.org/10.1145/3571730

Kene-Okafor, T. (2025, May 11). Egypt's Nawy, the largest proptech in Africa, raises $52M to take on MENA. *TechCrunch*. https://techcrunch.com/2025/05/11/egypts-nawy-lands-a-52m-series-a-to-take-on-mena (accessed 2026-11-03)

Kohavi, R., Tang, D., & Xu, Y. (2020). *Trustworthy online controlled experiments: A practical guide to A/B testing*. Cambridge University Press. https://doi.org/10.1017/9781108653985

Levitt, S. D., & Syverson, C. (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *The Review of Economics and Statistics, 90*(4), 599–611. https://doi.org/10.1162/rest.90.4.599

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems 33* (pp. 9459–9474). https://doi.org/10.48550/arXiv.2005.11401

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), Article 81. https://doi.org/10.1145/3359183

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

Nawy. (n.d.). *Nawy* [Website]. https://www.nawy.com (accessed 2026-11-03)

Property Finder Egypt. (n.d.). *Property Finder* [Website]. https://www.propertyfinder.eg (feature check recorded 2026-10-27)

Russo, D. J., Van Roy, B., Kazerouni, A., Osband, I., & Wen, Z. (2018). A tutorial on Thompson sampling. *Foundations and Trends in Machine Learning, 11*(1), 1–96. https://doi.org/10.1561/2200000070

Sawyer, S., Wigand, R. T., & Crowston, K. (2005). Redefining access: Uses and roles of information and communication technologies in the US residential real estate industry from 1995 to 2005. *Journal of Information Technology, 20*(4), 213–223. https://doi.org/10.1057/palgrave.jit.2000049

Schick, T., Dwivedi-Yu, J., Dessì, R., Raileanu, R., Lomeli, M., Zettlemoyer, L., Cancedda, N., & Scialom, T. (2023). *Toolformer: Language models can teach themselves to use tools* (arXiv:2302.04761). arXiv. https://doi.org/10.48550/arXiv.2302.04761

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy, 69*(3), 213–225. https://doi.org/10.1086/258464

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations (ICLR 2023)*. https://doi.org/10.48550/arXiv.2210.03629

Appendix A. Architecture diagram

```
flowchart LR
  subgraph Sources["Sources (categories)"]
    S1[Developer websites]
    S2[Portals and listing sites]
    S3[Owner-authorised partner feed]
    S4[Staff entry]
    S5[News RSS feeds]
  end

  subgraph Pipeline["Evidence pipeline (Postgres + server functions)"]
    G[Source guard: allowlist, robots.txt, rate clock]
    R[raw_records]
    N["normalized_records<br/>normalize_text_key / area / currency<br/>field_mappings"]
    M{"Entity matching<br/>exact / alias / fuzzy + co-signals"}
    MR["manual_reviews<br/>human decision"]
    P[("price_records<br/>append-only, dated, retractable")]
    V["recompute_verified_price<br/>trust x recency x match, MAD outliers"]
    VP[("verified_prices<br/>status + confidence")]
    MI["get_market_dashboard<br/>median EGP/m2, 0.5x-2x gate, n>=5"]
  end

  subgraph Catalog["Canonical catalogue"]
    C[("destinations / developers / projects / units<br/>payment_plans / offers / launches")]
  end

  subgraph Public["Public surface (EN/AR)"]
    W[Search, map, pages]
    K[Calculators: payment, mortgage, investment]
    A["Nour and Zaki<br/>text chat + live voice"]
  end

  subgraph Server["Server functions"]
    T["Text-chat brain<br/>server-held key"]
    E["Ephemeral voice session minting<br/>Gemini Live / OpenAI Realtime"]
    L["Lead capture<br/>phone normalisation"]
  end

  subgraph Staff["Staff control centre"]
    H[Leads and handoff]
    Q[Review, retraction, audit log]
    AG[Findings-only agents]
  end

  S1 --> G
  S2 --> G
  S3 --> G
  G --> R --> N --> M
  S4 --> P
```

```
S5 --> W
M -- "at or above auto threshold" --> C
M -- "review band" --> MR --> C
M --> P
C --> P
P --> V --> VP
P --> MI
C --> W
VP --> W
MI --> W
W --> A
A -- "tool calls (read-only, RLS)" --> C
A -- "tool calls" --> VP
A -- "tool calls" --> MI
A --> K
A <--> T
A <--> E
A --> L --> H
Q --> P
AG --> Q
```

Text description. Third-party and partner sources pass through a source guard into raw records, are normalized, and are matched to the canonical catalogue. High-confidence matches link automatically; review-band matches go to human review; prices are stored as append-only, dated price records. A verification function combines independent observations into a verified price with a status; a Market Index function computes sample-gated medians. The public site and the two advisors read the catalogue, verified prices and Market Index through read-only, row-level-secured functions. Server functions hold provider keys, mint short-lived voice sessions and normalise lead data; staff see leads, review queues, the audit log and findings from fenced background agents.

Appendix B. Evaluation protocols (proposed)

B.1 E1 — Grounded-figure audit

1. Freeze a database snapshot and record its date.
2. Build a prompt set of at least 200 prompts per language, stratified: catalogue search (30%), payment terms and calculation (20%), Market Index (15%), comparison (15%), missing-data cases where the correct answer is "not recorded" (10%), adversarial prompts asking for invented prices or guarantees (10%). Arabic prompts are written by Egyptian Arabic speakers, not translated.
3. Run each prompt through the text channel and, for a 25% subsample, the voice channel with synthetic audio (no real microphone, no real leads).
4. Log every tool call and result alongside the response.

5. Two annotators independently label each figure as *grounded* (matches tool output, allowing stated rounding), *ungrounded* (no supporting tool output), or *contradicted*; and each missing-data case as *correctly declined* or *fabricated*. Report Cohen's kappa.
6. Primary metric: grounded-figure rate with 95% Wilson interval, per language. Secondary: fabrication rate on missing-data and adversarial cases. Ablation: same prompts with tools disabled.
7. Pre-register pass thresholds before running.

B.2 E2 — Verified-price error

1. Draw a random sample of published units stratified by verification status and destination.
2. Obtain a current asking quote for each from an independent channel (developer sales office or published price list), recording date and channel; do not use any source already feeding the unit's evidence.
3. Compute absolute percentage error of `verified_price` (and of the unweighted mean and the latest single observation as baselines).
4. Report median error and interquartile range by status; test whether higher statuses have lower error (H3).

B.3 E6 — Contact-request harm check

1. Randomly assign a pre-registered share of sessions to a fixed honest wording (holdout); the rest use the learning policy.
2. Measure decline, opt-out, complaint and data-request rates per ask; and capture rate.
3. Stop the learning policy if harm metrics exceed the holdout by a pre-registered margin.

B.4 E7 — Voice permission and usability

1. On named devices and browsers (at least iPhone Safari, Android Chrome, desktop Chrome and Safari), load each page with the assistant, open the dialog, and record whether any permission prompt appears before "start" (expected: never).
2. Consenting testers complete scripted tasks in Arabic and English; record completion, errors and handover behaviour.

B.5 Reporting

Every evaluation reports: snapshot date, code commit, provider and model configuration in use, sample sizes, exclusions, metric definitions, and raw counts. Results are published regardless of direction.

Appendix C. Calculation example: verified-price confidence

C.1 Formula (reproduced from code, as implemented at commit 745652d)

For one unit or project, with usable evidence E (not retracted, not test, match confidence ≥ review floor, default 70), de-duplicated to one row per (source, price):

- median price m; absolute deviation $d_i = |p_i - m|$; MAD = median(d_i)
- outlier if MAD > 0 and $d_i > 3 \times$ MAD
- recency weight r_i: 1.00 if age ≤ 0 days; 0.95 if ≤ 7; 0.80 if ≤ 30; 0.50 if ≤ 90; else 0.20
- weight w_i = (trust_i / 100) × r_i × (match_i / 100) over non-outliers
- verified price = $\Sigma\, w_i\, p_i \,/\, \Sigma\, w_i$
- spread% = (max − min) / median × 100 over non-outliers
- n = number of distinct sources among non-outliers
- confidence = 40·min(1, n/4) + 25·max(0, 1 − spread%/15) + 20·(mean trust/100) + 15·max(r_i), rounded to one decimal
- status: n < minimum sources (default 2) → `insufficient_sources` (price withheld); freshest age > 180 days → `outdated`; any outlier → `major_price_difference`; confidence ≥ 95 → `verified`; ≥ 85 → `high_confidence`; ≥ 70 → `medium_confidence`; else `needs_review`.

C.2 Worked example (hypothetical inputs, Calculated)

Inputs are invented for illustration and do not describe any real unit.

Source	Trust	Price (EGP)	Age (days)	Recency r	Match	Weight w
A	85	10,000,000	3	0.95	99	0.85 × 0.95 × 0.99 = 0.799425
B	65	10,400,000	20	0.80	100	0.65 × 0.80 × 1.00 = 0.520000
B (repeat, same price, older)	65	10,400,000	25	—	100	removed by de-duplication
C	55	9,800,000	45	0.50	99	0.55 × 0.50 × 0.99 = 0.272250

- Median m = 10,000,000; deviations 0; 400,000; 200,000; MAD = 200,000; 3 × MAD = 600,000 → no outliers.
- Σw = 1.591675; Σwp = 16,070,300 → **verified price ≈ EGP 10,096,471**.
- spread% = (10,400,000 − 9,800,000) / 10,000,000 × 100 = 6.0.
- confidence = 40 × min(1, 3/4) + 25 × (1 − 6/15) + 20 × (68.33/100) + 15 × 0.95 = 30.0 + 15.0 + 13.67 + 14.25 = **72.9**.

- Status: 3 sources ≥ 2; freshest 3 days ≤ 180; no outliers; 72.9 ≥ 70 → **`medium_confidence`**.

Variant with an outlier. Add source D (trust 30) at EGP 15,000,000. The median of the four prices becomes 10,200,000; deviations are 200,000; 200,000; 400,000; 4,800,000; MAD = 300,000; 3 × MAD = 900,000. D is an outlier and is excluded from the estimate (the verified price is unchanged at about EGP 10,096,471), but the status becomes **`major_price_difference`**, so the figure carries a warning rather than a verified label.

This example shows two properties of the design: a third agreeing source raises confidence less than one might expect because the formula saturates at four sources and also rewards freshness and trust; and a single disagreeing source changes the status even when it does not change the number.

Appendix D. Data dictionary of core public entities (non-sensitive)

Only entities and fields relevant to public figures are listed. Personal-data tables, secret references and staff-only columns are omitted.

Entity	Purpose	Key fields (selection)
`destinations`	Areas/cities	names (EN/AR), slug, status, map position
`developers`	Property developers	names (EN/AR), slug, logo reference, status
`projects`	Compounds/projects	developer, destination, names (EN/AR), slug, status, starting price, hero image reference, map position
`units`	Individual units	project, unit type (EN/AR), bedrooms, bathrooms, area, price, currency, finishing (canonical label), delivery date, availability, listing type (developer sale / resale), publication status
`payment_plans`	Plan terms	scope (unit/project), down payment %, years, frequency, delivery payment, maintenance %, validity
`offers`	Developer offers	developer scope, projects, validity window, eligibility rules, bilingual copy
`launches`	New launches	name, developer, area, link to canonical project
`price_records`	Price evidence (public reads exclude retracted and test rows; source identity is staff-only)	unit or project, price, currency, area, price per m² (generated), terms snapshot, source update date, crawl date, match confidence
`verified_prices`	Current verified figure per unit/project	verified price, confidence, status, min, max, median, spread %, source count, outlier count, computed at
Market Index (function output)	Aggregates	destination, unit type, units, priced units, excluded outliers, median price/m², median price
`voice_profiles` (public subset)	Advisor personas	id, name, presentation, default flag
`feature_flags` (public subset)	Site switches	key, enabled

Appendix E. Reproducibility notes

- **Code snapshot.** All system descriptions refer to commit `745652d` (3 November 2026). The repository is private; it is not part of this package.
- **What can be reproduced without the code.** The verified-price formula (Appendix C), Market Index rules (Section 3.8), matching confidences (Section 3.4) and calculator conventions (Section 3.9) are fully specified here and can be re-implemented.

- **What cannot be reproduced.** Live catalogue contents, source data, provider configurations and model behaviour. Provider models change over time; any evaluation must record the configuration in use.
- **Tests in the repository.** The project maintains unit tests for the payment, mortgage, investment and market models, lead-capture logic, voice transport and resampling, provider key handling, and SEO generation; this report did not run them and reports no test results as evidence.
- **Not done for this report.** No database query, no live traffic analysis, no user study, no interviews and no external review.

Disclosure and conflict of interest

The author is the founder of REONE (Real Estate One) and has a direct financial and commercial interest in the system described and in its commercial success. The report was drafted with the assistance of AI writing tools and must be reviewed and approved by the founder before publication; the system itself was built with extensive use of AI coding assistants. No external funding was received for this report. The report has not been peer reviewed. No participants were recruited, no interviews were conducted, no datasets were collected for this report, and no evaluation has been run; all evaluations are proposed. Competitor information is limited to what each company shows on its own website on the dates stated; REONE is not affiliated with any company named. Nothing in this report is financial, investment or legal advice.

Version history

Version	Date	Changes
1.0 (draft)	2026-11-03	First draft. System described as implemented at commit `745652d`. No evaluation results. Author, affiliation, ORCID, licence and DOI pending.