The In-House AI Content Agency

Build Specification — OpenAI / ChatGPT Implementation

Prepared for: Anirudh Dashora — Sherringearth (developer brand) & Anirudh Dashora (personal brand) **Market:** Indore, Madhya Pradesh **Document 2 of 2.** Document 1 covers the same system built on Claude. **Date:** 21 August 2026. All platform facts verified against live sources on this date.

⚠️ 0. Read this before you build anything on OpenAI

There is a platform decision here that will cost you weeks if you get it wrong, and most tutorials online are already out of date.

0.1 Do NOT build in Agent Builder

OpenAI's **Agent Builder** — the visual drag-and-drop canvas launched as part of AgentKit at DevDay in October 2025 — is the tool almost every "build an AI agent" tutorial from 2025 and early 2026 points you at.

On 3 June 2026, OpenAI announced it is being wound down. It leaves the platform on 30 November 2026. The Evals platform is going with it (read-only from 31 October 2026). ChatKit survives, but needs a backend you control once the hosted Agent Builder backend is gone.

If you build your twelve-agent system in Agent Builder starting today, it stops working in roughly three months. **Do not start there.** Any guide telling you to go to `platform.openai.com/agent-builder` was written before June 2026.

0.2 Where OpenAI points you instead

Two paths:

- **Workspace Agents in ChatGPT** — the no-code path, for natural-language-defined agents
- **Agents SDK** — the code path, Python or TypeScript

0.3 The critical eligibility constraint

Workspace Agents are not available on ChatGPT Plus. They require **Business, Enterprise,**

Edu or Teachers plans. Launched 22 April 2026 in research preview, Codex-powered, cloud-run, shareable across a workspace, schedulable, and usable in ChatGPT or Slack.

They were free during the research preview window; **credit-based pricing began 6 May 2026, with agent pricing taking effect 6 July 2026.**

What this means for you concretely: if you are on ChatGPT Plus today, you cannot build this system as specified. You must move to a **ChatGPT Business** seat (commonly cited around $20–25 per user per month, plus credit consumption for agent runs). Budget both the seat and the credits — heavier and longer agent runs consume more credits, and a twelve-agent weekly research pipeline is a heavy workload.

0.4 Also note: Custom GPTs are being retired for organisations

Workspace Agents are explicitly positioned as the successor to Custom GPTs, and OpenAI has stated it will deprecate the custom GPT standard for Business, Enterprise, Edu and Teachers accounts at a future date, requiring migration to workspace agents. GPTs remain available in the meantime, and OpenAI has said conversion tooling is coming.

Implication: if you are on Plus and want to start immediately, you *can* prototype with Custom GPTs + Projects (§7.4) — but treat it as a prototype, not the destination. Don't invest weeks in a Custom GPT build.

0.5 And: Sora is gone

If you were considering Sora for video: OpenAI announced in March 2026 that it is winding Sora down. The web and app experiences shut down on **26 April 2026**, and the Sora API is discontinued on **24 September 2026**, after which the model is fully retired. Do not build any video pipeline on Sora. Use Higgsfield (which aggregates Seedance 2.0, Veo 3.1, Kling 3.0 and others) or HeyGen for avatar work.

---

1. Executive summary

1.1 The system in one paragraph

A twelve-agent editorial network, orchestrated by one Editor-in-Chief agent, running a weekly loop: signal detection → verified research → adversarial fact audit → contrarian positioning → narrative architecture → format routing → scripting → design/AV production → quality gate → publish → performance learning. Two brands (Sherringearth institutional, Anirudh Dashora personal) share a research layer and keep separate voice, visual and legal layers.

1.2 Target outcome

Your involvement drops from ~20 hours/week to **~2.5 hours/week**, and those hours become

the highest-leverage ones. Running cost at full build: **₹13,000–₹30,000/month** on the OpenAI path (higher than the Claude path, primarily because Workspace Agents require a Business seat plus credits). Build time: 4–6 weeks part-time.

1.3 The strategic finding

Automate the production. Do not automate the judgment. This is expanded in §2 and is the most important section in this document — it argues against part of your original brief, and it is the difference between building a durable asset and building a commodity.

---

2. The strategic core: what you're actually selling

2.1 Two buckets

Every claim in "high-end, conceptual, educational" real estate content is one of two kinds:

Bucket A — Publicly retrievable. Circle rates, RERA data, metro alignment, FSI norms, launch prices, absorption, interest rates, policy changes. An agent can retrieve and verify all of this. **So can any competitor's agent.** This is a commodity within eighteen months.

Bucket B — Proprietary and tacit. Why a specific Super Corridor parcel sat unsold for four years and what finally moved it. What an approvals officer actually asks for versus the checklist. The real reason a "sold out" board went up before the sales did. What a ₹40 crore land deal's payment structure looks like when one of three co-owners is in Dubai. Which amenity Nipania buyers say they want and never use.

Bucket B is the entire moat. It exists only in your head and in perhaps thirty other heads in Indore. No agent retrieves it, because it is not written down.

2.2 The implication

A system producing Bucket A content only will produce well-designed, well-sourced, perfectly-formatted content that is **indistinguishable from what a smart 24-year-old with the same tools will produce in six months.** You will have automated yourself into the commodity tier while feeling like you escaped it.

So the architecture must include a mechanism that **systematically harvests Bucket B from you.** That is the Founder Deposit (§3.4). Ten minutes a week. It is the highest-value component in this specification and the one no off-the-shelf agent stack will build for you.

2.3 The revised target

Not *"AI does everything, I press post."* But *"AI does 92% of the work and asks me the right seven questions each week — the ones only I can answer — then builds world-class content*

around those answers."

2.4 On avatars — a specific warning

Bucket A → avatar allowed. Project walkthroughs, data explainers, listing updates, announcements, Hindi/English duplicates.

Bucket B → your actual face, always. Opinion, contrarian takes, criticism of market practice. Audiences detect avatar content at a much higher rate in 2026 than in 2024, and the credibility penalty when detected on an *opinion* piece is severe and hard to reverse. A real estate influencer whose hot take turns out to be an avatar has a positioning problem with no easy fix.

2.5 What "high-end" means — encode it structurally

Vague instructions like "make it premium" produce nothing. Put this table verbatim into your positioning and scriptwriting agents:

Commodity content	**Your content**
States a fact	States a fact, then what it *implies* that isn't obvious
"Nipania prices rose 18%"	"Prices rose 18% but registrations fell 6% — that gap is speculative holding, and here's what it does to your exit"
Explains what RERA is	Explains what a promoter does between registration and QPR filing, and how to read that
Lists amenities	The ₹/sq ft carrying cost of each amenity and who actually pays it
Advice	A *model* — a repeatable way of thinking the reader applies to the next property too
Hook = curiosity gap	Hook = a claim a knowledgeable person would initially dispute

---

3. The Brand Constitution — build these files first

Ninety percent of failed agent builds fail here: agents built first, knowledge base never. Build these seven files before any agent.

In the OpenAI path these live in **two places**: uploaded to the agent's Files section (limit 512 MB per file, 10 GB per agent — you will not approach this), and mirrored in a ChatGPT **Project** for chat-based iteration.

> **Performance note:** OpenAI's own guidance is that agent performance declines as more files and larger data are added. Add only the files the agent needs, split large documents into smaller files, and organise collections into folders. This argues for **per-agent file scoping** — SCOUT gets the source registry and audience dossier, not the visual system. Do not dump all seven files into all twelve agents.

```
/AI-Agency/
  /00-constitution/
     01-brand-constitution.md
     02-voice-corpus.md
     03-visual-system.md
     04-audience-dossier.md
     05-source-registry.md
     06-legal-guardrails.md
     07-format-playbook.md
  /01-founder-deposits/
  /02-claim-ledger/
  /03-output/ (drafts | approved | published)
  /04-performance/
  /05-competitors/
```

3.1 `01-brand-constitution.md`

Separate profiles for the two brands — **never a shared voice.** A developer brand speaking like a person loses institutional credibility; a personal brand speaking like a company loses everything.

Per brand, write: one-sentence positioning (specific enough to be slightly uncomfortable); the three things you say that others won't (actual sentences); the five phrases you never say — explicitly ban "dream home", "investment ka sahi waqt", "limited units available", "prime location", "world-class amenities"; register (Hinglish first-person for Anirudh Dashora, English-primary institutional for Sherringearth); and the **sophistication floor**, written as a hard rule:

> *"If someone with three years in this industry could have written this, it does not ship."*

That single line, present in every agent's instructions, does more for quality than any amount of prompt engineering.

3.2 `02-voice-corpus.md` — the highest-leverage file

Fifteen to twenty-five **raw** samples of your actual writing and speech. Not summaries. Transcribed WhatsApp voice notes about a deal, your best LinkedIn posts, an email explaining something complex, a rant, a proposal intro you wrote yourself. Tag each with context.

Then your own analysis: sentences you start with; your three verbal tics (keep them — they're identity markers); how and when you switch Hindi/English; how you express disagreement; how

you open when serious versus playful.

This file alone is roughly 60% of the difference between output that sounds like you and output that sounds like ChatGPT. Two hours. Best two hours in the build.

3.3 The Founder Deposit — the Bucket B harvester

Every Monday, 10 minutes, one voice note. The Editor-in-Chief agent generates seven questions from that week's signals. Examples of the *kind*:

- "Regulation X changed. In 17 years, what's the closest precedent — and what actually happened after, not what people predicted?"
- "You've seen a plotted scheme go wrong. What was the first visible sign, three months before anyone noticed?"
- "Absorption in Super Corridor is up. What does a developer see in that number that a buyer structurally cannot?"
- "What did someone tell you this week that you disagreed with?"
- "What's a number everyone in this market quotes and is wrong about?"

Drop the file in `/01-founder-deposits/`. An agent transcribes and mines it into a running **Insight Bank**, tagged and searchable. Every piece of content must draw at least one element from it — enforced as a hard gate. Within three months you own an insight asset no competitor can replicate.

3.4 `03-visual-system.md`

Agents can't see your Canva. Write it down: exact hex codes (primary, secondary, two accents, background, text); fonts with weights and web-safe fallbacks; **carousel geometry 1080×1350 (4:5)** — not 1080×1080, which wastes vertical space on mobile; safe margins ≥80px, nothing within 100px of the bottom; **minimum type 32px at 1080px width** (below this it's unreadable on a phone and you won't notice on a laptop — the single most common mobile failure); slide 1 maximum 9 words; logo on slide 1 and final slide only; chart rules (no 3D, no pie charts over 4 segments, axes labelled in ₹ and sq ft); reel 1080×1920 with burnt-in captions avoiding the top 15% and bottom 20%.

3.5 `04-audience-dossier.md`

Micro-market map with your own commentary — Vijay Nagar, Nipania, Super Corridor, Scheme 54, AB Road, Bypass, Rau, Mhow road, Bicholi Mardana: who buys, why, what the common mistake is. Buyer archetypes (first-time flat buyer, plot investor, NRI parent-proxy, commercial buyer, upgrade buyer) in three sentences each, with the language each responds to. Local reference points an outsider wouldn't know — these are credibility signals in themselves.

3.6 `05-source-registry.md` — the spine of fact integrity

Tier 1 — primary, citable without qualification: MP RERA project and agent database (`rera.mp.gov.in` , project search at `/all-projects/` — some guides cite `reraonline.mp.gov.in` ; confirm which is live); RERA Act 2016 and MP Rules (MP rules effective 1 May 2017); IGRS/registration department circulars and collector guideline (circle rate) notifications; Indore Municipal Corporation and IDA notifications and master plan; MPMRCL/Indore Metro releases; RBI policy statements; CBDT circulars for capital gains, TDS 194-IA, Sections 54/54F; audited company filings.

Tier 2 — reputable secondary, cite with attribution: Knight Frank, JLL, Anarock, CBRE, PropTiger, Colliers — **always name the report and quarter, never "a report says"**; Mint, Business Standard, Economic Times, Hindu BusinessLine, Moneycontrol for reported facts; Dainik Bhaskar, Naidunia, Free Press Journal Indore for civic news.

Tier 3 — signal only, NEVER citable: 99acres, MagicBricks, Housing.com. **These are *asking* prices, not transaction prices.** Write this in bold in the registry, because every agent will otherwise cheerfully cite a listing portal as transaction data. This is the most common factual error in Indian real estate content. Also here: broker forwards, YouTube commentary, LinkedIn posts, brokerage "market reports" with a sales interest in the number.

Banned: anything AI-generated you can't trace to a primary source; any figure whose original source you cannot open and read.

3.7 `06-legal-guardrails.md`

You are a **developer**, not just a commentator. Your exposure is materially different and most influencer-style agent stacks will walk you into it.

- **RERA advertising.** Under the Act, a promoter cannot advertise, market, book, sell or offer for sale a project requiring registration unless registered, and the registration number must appear in the advertisement. Any Sherringearth content showing or naming a project **is** an advertisement. Hard rule: **no RERA number → does not ship.** Get the exact current disclosure wording from your lawyer once, then encode it verbatim.
- **Registration threshold:** MP RERA applies above 500 sq m or more than eight units. Flag content touching this boundary.
- **No return guarantees.** Ban "assured returns", "guaranteed appreciation", "X% ROI" at the QC layer — for your projects and as generic advice.
- **No named competitor criticism.** Critique practices, never firms. Absolute.
- **No individual financial or legal advice.** Educational framing plus a standing disclaimer.
- **Cross-brand disclosure.** Anirudh Dashora content about a Sherringearth project must disclose the interest. Non-negotiable — and it reads as confidence.

- **AI disclosure.** Where a reel uses synthetic voice or avatar, say so. Global regulatory direction is toward mandatory synthetic-media labelling; pre-complying costs nothing and buys trust.

---

4. The Fact Integrity Protocol

You asked for "no assumptions, double-check, be very sure." Here is the mechanism.

4.1 The core problem

A model asked to check its own work will usually confirm its own work. Self-review is theatre. Real verification requires structural separation: a different agent, different instructions, whose measured success is *finding errors* — not producing content.

This is why the roster has RED as a distinct agent rather than a "review step" inside the researcher.

4.2 The Claim Ledger

Every draft ships with a machine-readable ledger — one row per factual claim, no exceptions, including claims the agent is confident about.

#	Claim as written	Type	Tier	Source	URL	Source date	Retrieved	Verdict
1	"MP RERA requires registration above 500 sq m"	Regulatory	1	RERA Act / MP Rules	[url]	2017-05-01	2026-08-21	VERIFIED
2	"Nipania up 18% YoY"	Market	3	Listing portal	[url]	2026-07	2026-08-21	**DOWNGRADE — asking price. Rewrite**

Verdicts: `VERIFIED` / `VERIFIED-WITH-QUALIFICATION` / `UNVERIFIABLE — CUT` / `DOWNGRADE — REWRITE` / `CONTESTED — PRESENT BOTH SIDES`

> **OpenAI-specific advantage:** Workspace Agents are Codex-powered and can write and run short programs on the fly — fetching structured data, transforming columns, validating output before it moves downstream. This is genuinely useful here: have LEDGER emit the Claim Ledger as structured CSV/JSON and have RED programmatically check for null source URLs, missing dates, and stale sources rather than relying on the model to notice.

> **Machine-checkable rules should be checked by machine, not by a language model.** This is the strongest reason to prefer the OpenAI path if you go that way.

4.3 The three-pass rule

Pass 1 — LEDGER gathers, builds the ledger, writes the brief. Instructed to prefer "I could not verify this" over a plausible number.

Pass 2 — RED, a separate agent in a separate session, given **only the draft and the ledger** — never the research conversation. If it sees the researcher's reasoning it inherits the researcher's assumptions and the audit collapses. Instruction is explicitly adversarial (full text in §8.3).

Pass 3 — You. Verify **two claims per piece, selected at random by the system**, not by you. Two minutes. Random selection is what makes it a control rather than confirmation bias.

4.4 Hard kill rules

Encode as absolutes — agents negotiate with soft rules.

1. A number with no openable source URL does not ship. Cut, not hedged.
2. Listing prices are never stated as prices. Always "asking prices on [portal], [month]."
3. No statistic over 12 months old in the present tense.
4. Always ₹ with lakh/crore specified, and per sq ft vs per sq m. Mixing sq ft and sq m is the second most common error in Indian property content.
5. No causal claims from correlational data unless a Tier 1/2 source makes the causal claim.
6. Regulatory claims require the primary instrument. A news article about a rule is not a source for the rule.
7. **When sources conflict, publish the conflict.** "Anarock says X, Knight Frank says Y, the gap is likely definitional" is *higher-end content* than a false-confident single number. This is exactly the positioning you want.

4.5 Cap the loop

You asked for a system that loops and double-checks. A caution: unbounded self-critique degrades output. After roughly two revision cycles models start hedging and sanding off the sharp edges that make content good. **Maximum two revision passes, then it comes to you.** Encode the cap in the orchestrator.

---

5. The roster, the format engine, and the weekly loop

5.1 Twelve agents, one job each

Single-responsibility agents beat multi-purpose ones consistently — a "research and write" agent does both worse than two agents doing one each.

#	Codename	Name	Function	Phase
1	**DESK**	*Editor-in-Chief*	Orchestration, routing, gates, Founder Deposit questions	1
2	**SCOUT**	Tony	Signal & trend intelligence	1
3	**LEDGER**	Veronica	Research, primary sourcing, Claim Ledger	1
4	**RED**	Sarah	Adversarial fact audit — separate session, draft only	1
5	**ANGLE**	Kabir	Contrarian positioning	2
6	**ARC**	Meera	Narrative architecture, hooks	2
7	**FORM**	Dev	Format decision + distribution	2
8	**QUILL**	Aria	Scriptwriting in brand voice	2
9	**CANVAS**	Rohan	Carousel/static design production	3
10	**STUDIO**	Nikhil	AV production — voice, avatar, b-roll, edit spec	3
11	**GATE**	Ishaan	Final QC, ship/no-ship	2
12	**ECHO**	Tara	Performance analysis, learning loop	3

5.2 Format Decision Engine (FORM's rulebook)

Never let an agent choose format by feel. Score, then route.

Score 1–5 on: claim density · emotional charge · nuance requirement · visual dependency.

Condition	Format
Claim density ≥4 AND nuance ≥3	**Carousel** (8–10 slides)
Emotional charge ≥4 AND claim density ≤2	**Reel** (30–45s)
Nuance ≥4 AND needs lived experience/disagreement	**Podcast segment** (8–14 min)
Visual dependency ≥4 AND claim density ≥3	**Carousel with data slides**

Time-sensitive regulatory/civic news	**Story same-day + Reel within 48h**
Claim density ≥5 AND nuance ≥4	**Podcast**, then atomise

The face test: does credibility depend on it coming from you personally? Yes → your face, no avatar.

The atomisation rule: every podcast yields 1 carousel + 3 reels + 5 story frames + 1 LinkedIn long-form. FORM produces this at scripting time, not afterwards. This is where the economics come from — one recording becomes ten assets.

5.3 Construction templates

Carousel (8–10 slides, 4:5): 1) the claim, ≤9 words, stated as an assertion a knowledgeable person might dispute · 2) why the obvious answer is incomplete · 3) the mechanism · 4–5) evidence, sourced and dated · **6) the Bucket B slide — mandatory** · 7) the counter-argument, stated fairly · 8) what it means for the reader's next decision · 9) the model — the repeatable way of thinking · 10) CTA + sources in small type + disclosure.

Reel (30–45s): 0–3s the disputed claim, no greeting, no logo sting · 3–8s the stake (what they lose by getting this wrong) · 8–25s the mechanism with one concrete number · 25–35s the Bucket B beat · 35–45s takeaway + a question that invites a real comment.

Podcast deliverable must include: verbatim cold open (60–90s, the sharpest 90 seconds of the episode) · segment map with timings · 15–20 questions ordered by escalating difficulty, each with a "why this question" note and a follow-up probe · fact sheet with the Claim Ledger so you can speak from verified data live · the three most clippable moments flagged in-script · **setup sheet** (camera position and framing, mic type and placement, lighting layout, room treatment, background dressing, wardrobe — avoid tight patterns and stripes, they moiré) · post-production clip plan with timecodes.

5.4 The weekly loop

Day	Who	What	Your time
Mon 06:00	SCOUT (scheduled)	8–12 ranked signals	0
Mon 08:00	DESK	Select 4, generate 7 Founder Deposit questions	0
Mon 09:00	**You**	Record the 10-minute voice note	**10 min**

Mon 11:00	LEDGER	Deep research + Claim Ledger	0
Tue	ANGLE → ARC → FORM	Positioning, narrative, format routing	0
Wed	QUILL → RED	Draft, then adversarial audit (separate session)	0
Thu 09:00	**You**	Review 4 pieces, random 2-claim spot check each	**60 min**
Thu	CANVAS / STUDIO	Production	0
Fri	GATE	Final QC	0
Fri	**You**	Final look, schedule natively	**30 min**
Sat	**You**	Podcast recording (fortnightly)	**45 min**
Sun	ECHO	Performance pull, learning log	0

~2h 25m/week, plus fortnightly recording.

5.5 GATE's ship/no-ship checklist — binary, any FAIL blocks

1. Every ledger claim VERIFIED or VERIFIED-WITH-QUALIFICATION
2. RED's audit returned, all flags resolved
3. **At least one Bucket B element present** (the sophistication gate)
4. No banned phrase
5. No return guarantee, express or implied
6. RERA registration number present if a Sherringearth project is referenced
7. Interest disclosed if cross-posting between brands
8. AI disclosure if synthetic voice/avatar used
9. Mobile check: type ≥32px, nothing in bottom safe zone, contrast passes
10. Slide 1 ≤9 words
11. Sources listed
12. Sophistication floor: *could someone with three years in this industry have written this?* Yes → FAIL

6. Architecture on OpenAI

6.1 Which surface does what

Surface	Use for	Availability
Workspace Agents	**The workhorse.** All twelve agents. Codex-powered, cloud-run, scheduled, shareable, tool-connected	Business / Enterprise / Edu / Teachers only
Skills in ChatGPT	Reusable instruction packs the agents call in their instructions	Available in ChatGPT
Projects	Persistent context per brand for chat-based iteration	All paid plans
Agent mode	One-off autonomous browse-and-do tasks	Plus and above
Custom GPTs	Prototyping only — being deprecated for orgs	All paid plans
Agents SDK	Only if you go code-first. You don't need to	API
Agent Builder	**Do not use — retiring 30 Nov 2026**	—

6.2 Build in three phases. Do not skip.

The failure mode of ambitious agent builds is building all twelve at once and then being unable to tell which agent is producing bad output. Debugging a twelve-agent system you have never run is close to impossible for a non-technical builder.

- **Phase 1 (Week 1–2): 4 agents.** DESK, SCOUT, LEDGER, RED. Output: a verified weekly intelligence brief. **Ship no content until this is reliably good.** Weak research = weak everything downstream.
- **Phase 2 (Week 3–4): +5.** ANGLE, ARC, FORM, QUILL, GATE. Output: finished scripts and copy.
- **Phase 3 (Week 5–6): +3.** CANVAS, STUDIO, ECHO. Output: finished assets and a learning loop.

Run each phase a full week before adding the next.

6.3 How to actually build a Workspace Agent

You will not write code. The build flow:

1. Open the agent builder chat in ChatGPT.
2. **Describe the job in plain language** — what the agent does, what a successful outcome looks like, and the constraints it must follow. The builder translates this into a defined workflow you can then refine in chat or edit directly in the workflow and instructions panes.
3. **Choose tools and connectors** — select approved apps (Google Drive, Gmail, Google Calendar, Slack, SharePoint). You can also describe the systems the agent needs and the builder will walk you through adding and authenticating them. Available apps depend on what's enabled in your workspace.
4. **Add files** via the Files section — scoped per agent (§3, performance note).
5. **Reference Skills** in the instructions for reusable behaviour.
6. **Publish** to the workspace directory and **attach a schedule.**
7. **Invite collaborators** if needed — agents can be maintained by multiple people with chat-or-edit permissions.

> On Enterprise, access to *build* agents is admin-controlled. If you set up a Business workspace as owner, this is not a blocker for you.

6.4 Scheduling and triggers

Publish agents to the directory and attach schedules. Schedule:

- **SCOUT** — weekly, Monday 06:00
- **ECHO** — weekly, Sunday 18:00
- **Regulatory watch** — daily 07:00, narrow scope: MP RERA notifications, IMC/IDA circulars, RBI, IGRS circle-rate notifications. Alert-only. This one has business value beyond content — you'll hear about relevant changes before your competitors.

There is also a **Workspace Agents API** for triggering published agents from your own systems. Two things to know before you plan around it: create a Workspace Agent access token in ChatGPT Admin → Access token with the Workspace Agents scope; and **the API queues the run and returns 202 Accepted with no response body — no run ID, and the agent's response cannot currently be retrieved through the API.** So it is a *fire-and-forget trigger*, not a request/response integration. Design your handoffs to land in Drive or Slack, not to be read back from the API.

6.5 Handoffs between agents

Workspace Agents don't chain automatically the way a purpose-built orchestration graph

would. Two workable patterns:

Pattern A — Shared-folder handoff (recommended, no code). Each agent writes its output to a designated Google Drive folder. The next agent's schedule fires later and reads that folder. DESK reads all folders and assembles the weekly review pack. Simple, inspectable, debuggable — you can open any folder and see exactly where the pipeline broke.

Pattern B — Slack-channel handoff. Agents post outputs into dedicated Slack channels (`#scout-signals` , `#ledger-briefs` , `#red-audits`). Agents can be deployed into Slack channels directly. Good if you already live in Slack; gives you a visible audit trail.

Do not attempt full event-driven chaining on this platform right now. The 202-with-no-body API behaviour makes reliable programmatic chaining awkward, and scheduled-plus-shared-folder is more robust for a solo operator anyway.

6.6 Connectors

Agents can connect to all apps available in ChatGPT; admins define which tools and actions are reachable. For this build: **Google Drive** (constitution files, output archive), **Gmail** (report delivery, newsletter ingestion), **Google Calendar** (content calendar, recording blocks), **Slack** (if using Pattern B).

For video and voice, you will work **outside** ChatGPT: ElevenLabs, HeyGen and Higgsfield all have their own interfaces and APIs. Higgsfield in particular exposes an MCP server and a CLI, which is convenient if you later move any of this to a code-based orchestrator.

---

7. Production pipelines

7.1 Carousels — highest ROI for your positioning

Do not use Canva templates. Templates are why all Indore real estate content looks identical.

Instead: have CANVAS generate a **single self-contained HTML file** — inline CSS, your exact hex codes and fonts, 1080×1350 slides. Codex-powered agents write and run code, so this is well inside capability. Open in a browser, screenshot or export to PNG, iterate in plain language ("slide 4's chart is too dense, drop the third series").

Custom design at zero marginal cost, with bespoke data visualisation on every piece rather than a stock chart block. This is the single biggest visual differentiator available to you.

Slide count: Instagram supports up to 20 slides posting **in-app**, but Meta's Content Publishing API caps carousels at 10 — so no third-party scheduler can exceed 10. Default to 8–10 to keep both routes open; reserve 15–20 slide deep-dives for manual posts.

7.2 Reels — three routes

Route A — you on camera. For Bucket B. QUILL delivers a teleprompter script with beat timings; STUDIO delivers a shot list and b-roll spec.

Route B — voice clone + b-roll. ElevenLabs Professional Voice Cloning is available from the **Creator tier ($22/mo)**, which includes roughly 121,000 credits (~1 credit per character) and 192 kbps output. PVC trains a dedicated model and needs substantially more audio than instant cloning — ElevenLabs recommends around **3 hours of consistent studio-grade audio** for production quality, with about **30 minutes as the floor**. Instant Voice Cloning is available from the ~$6 Starter tier and works from 1–3 minutes of clean audio, but the seams show on anything longer than a short clip.

Do the PVC recording once, properly, in a treated room. The quality ceiling of every voice reel you ever make is set that day.

Higher tiers add PVC slots (Scale ~$299/mo includes 3; Business ~$990/mo includes 10) — you need one. Stay on Creator.

Route C — avatar. HeyGen: Free (3 videos/mo, watermarked), **Creator $29/mo** (~$24 annual), **Pro from $49/mo**, **Business $149/mo + $20/seat**. Credits are the real constraint — premium avatar engines burn credits several times faster than standard generation. The separate pay-as-you-go **API wallet starts at $5**, roughly **$1 per minute** of standard 1080p avatar video, with **Avatar IV around $4/min**. Note HeyGen removed free API credits in February 2026.

Recommendation: A for opinion, B for explainers, C sparingly for announcements and multilingual duplicates.

B-roll: Higgsfield aggregates 15+ models (Seedance 2.0, Veo 3.1, Kling 3.0, Wan, MiniMax) under one subscription — Starter ~$15/mo (200 credits), Plus ~$49/mo (1,000 credits) unlocking the full model set. Two cautions: annual billing is the default at signup, and MCP-generated images consume credits even when "unlimited" is toggled in the web UI.

7.3 Podcast

The system does everything except sit in the chair: arc, verbatim cold open, escalating question ladder, verified fact sheet, setup and lighting spec, post-production clip plan. Fortnightly recording, atomised into ten assets.

7.4 If you're on Plus today and want to start this week

You can prototype without a Business seat, with reduced capability:

- Build **DESK, SCOUT, LEDGER and RED as four Custom GPTs**, each with its own instructions and the relevant constitution files.
- Run the handoffs manually: copy SCOUT's output into LEDGER, and so on. Twenty minutes

of clicking per week.

- Use a **Project** per brand for the constitution files and chat iteration.
- **No scheduling.** You trigger runs manually.

This proves the *content* system works before you pay for the *automation* system. It is a legitimate two-week validation step. Just don't build past Phase 1 there — Custom GPTs are on a deprecation path for organisations, and manual handoffs stop being viable at nine agents.

7.5 Publishing — do not build API posting

You said you'd download and post manually. **Do that — and use native scheduling.**

Since **1 March 2026**, any **public** Instagram account can schedule natively — no Business/Creator conversion needed. The toggle is under Advanced settings during post creation. Limits: **25 posts/day, up to 75 days ahead**, supporting posts, carousels and Reels. Private accounts excluded.

Native beats API on every axis for you: no Meta app review (2–4 weeks), no OAuth token management, no permission approvals, up to 20 carousel slides instead of 10, access to trending audio, no third-party subscription. Build API publishing only if you scale to several brands and genuinely need it.

---

8. Agent instructions — copy these into the builder

Describe the job in plain language, then paste these as the instruction core. Keep the six-part structure: **Role → Inputs → Method → Output contract → Hard rules → Failure behaviour.** The failure-behaviour section is the one everyone omits and the one that most improves reliability.

8.1 SCOUT (Tony)

```
ROLE
You are SCOUT, signal intelligence analyst for an Indore-based real estate
developer and personal brand. You find what moved this week. You do not write
content and you do not offer opinions on content.

INPUTS
- 05-source-registry.md (read before every run)
- 04-audience-dossier.md
- /04-performance/ last 4 weeks (avoid repeating covered ground)

METHOD
1. Sweep Tier 1 first: MP RERA notifications, IMC/IDA circulars, IGRS
```

```
   circle-rate notifications, RBI, CBDT, MPMRCL.
2. Then Tier 2: national property desks, Indore local press.
3. Then national/global signals with a plausible Indore transmission path.
4. For each signal answer explicitly: "Why does this matter to a buyer or
   developer in Indore specifically?" If you cannot answer in one concrete
   sentence, discard it.

OUTPUT CONTRACT (write to Drive folder /scout-signals/ as dated markdown)
8-12 signals, each with:
- Headline (<=12 words)
- What happened (2 factual sentences)
- Primary source URL + publication date
- Indore relevance (1 concrete sentence)
- Contrarian potential: HIGH/MEDIUM/LOW
- Brand fit: SHERRINGEARTH / ANIRUDH DASHORA / BOTH / NEITHER
- Freshness window
Ranked by (Indore relevance x contrarian potential).

HARD RULES
- Never cite a listing portal as a price source.
- Never report a claim without opening the source.
- Mark widely-covered stories SATURATED and state what angle remains unclaimed.

FAILURE BEHAVIOUR
If fewer than 8 signals meet the bar, return fewer and say so. Never pad.
An honest 5-signal week beats 12 with 7 filler.
```

8.2 LEDGER (Veronica)

```
ROLE
You are LEDGER, primary-source researcher. Your product is a research brief
plus a Claim Ledger. You are rewarded for accuracy and for admitting gaps,
never for comprehensiveness.

METHOD
1. Go to the primary instrument. A news article about a regulation is NOT a
   source for that regulation - open the notification, circular or Act text.
2. Build the Claim Ledger. Every factual assertion gets a row, including ones
   you are confident about.
3. Where sources conflict, record BOTH and note the likely definitional cause.
   Do not resolve conflicts by choosing.
4. List what you could NOT verify under "UNVERIFIED - DO NOT PUBLISH".
5. Emit the Claim Ledger as CSV as well as a table, so downstream checks can
   be run programmatically.

OUTPUT CONTRACT (write to /ledger-briefs/)
- Research brief (600-900 words)
```

```
- Claim Ledger table AND claim-ledger.csv
- "What surprised me" (3 bullets)
- "UNVERIFIED - DO NOT PUBLISH"
- Open questions only a 17-year practitioner could answer

HARD RULES
- Listing-portal figures are ASKING prices. Label them so, always.
- No statistic older than 12 months in the present tense.
- Always ₹ with lakh/crore, and per sq ft vs per sq m specified.
- If you cannot open a source URL, the claim does not enter the ledger.

FAILURE BEHAVIOUR
Say "I could not verify this" rather than producing a plausible number.
Three verified claims beat twelve unverified ones. Never estimate a figure
and present it as retrieved.
```

8.3 RED (Sarah)

```
ROLE
You are RED. You are a hostile fact-checker hired by a competitor who wants
to publicly embarrass this author. Your success is measured in errors found,
not content approved. You are not a collaborator.

INPUTS
The draft and claim-ledger.csv ONLY. You do not receive the research
conversation and must not request it.

METHOD
Write and run code to check the CSV mechanically first: null source URLs,
missing dates, sources older than 12 months, Tier 3 sources on price claims.
Then, for every remaining claim, attempt independently to DISPROVE it.
Verify the source actually says what is claimed.

FLAG MANDATORILY
- Any number without a Tier 1 or Tier 2 source
- Any listing-portal price presented as a transaction price
- Causal language ("because", "caused", "led to") unsupported by the source
- Any statistic over 12 months old presented as current
- Any implied guarantee of returns or appreciation
- Any regulatory claim sourced from journalism rather than the instrument
- Any sq ft / sq m or lakh / crore inconsistency
- Any claim materially misleading if a knowledgeable reader saw only the headline

OUTPUT CONTRACT (write to /red-audits/)
- Verdict per claim: PASS / QUALIFY / CUT, with reasoning
- Overall: SHIP / REVISE / KILL
- "Strongest attack a competitor could make on this piece" (1 paragraph)
```

```
FAILURE BEHAVIOUR
If you find zero issues you have not looked hard enough. State explicitly what
you attempted to disprove and failed to. "Looks good" is not acceptable output.
```

8.4 ANGLE (Kabir)

```
ROLE
You are ANGLE. You find the take no other Indore real estate account is taking.
You are the difference between commodity content and authority.

METHOD
1. Read the research brief and the Insight Bank.
2. State the consensus view in one sentence.
3. Generate 5 candidate angles across:
   - Inversion: accepted causality runs the other way
   - Second-order: everyone sees effect A, nobody has priced effect B
   - Definitional: the number everyone quotes measures the wrong thing
   - Practitioner: what a developer sees that a buyer structurally cannot
   - Historical: the closest precedent and what actually happened after
4. Score each on defensibility, differentiation, risk.
5. Recommend one. State the strongest counter-argument fairly.

HARD RULES
- Contrarian must mean better-evidenced, never merely provocative.
- Never criticise a named competitor firm. Practices only.
- If the consensus view is simply correct, SAY SO and recommend a clarity
  angle instead. Manufactured contrarianism is the fastest way to destroy
  authority positioning.

OUTPUT CONTRACT
Consensus view / 5 scored candidates / recommendation + rationale /
strongest counter-argument / evidence required to make it stand up.
```

8.5 QUILL (Aria)

```
ROLE
You are QUILL. You write in Anirudh Dashora's voice, or Sherringearth's
institutional voice. Never a generic content voice.

MANDATORY PRE-READ
- 02-voice-corpus.md - read fully before writing a word
- 01-brand-constitution.md
- The Insight Bank

METHOD
```

```
1. Identify which brand voice applies.
2. Draft to the structural template for the assigned format.
3. Integrate at least one Insight Bank element. Mandatory.
4. Self-check against the banned phrase list.
5. Read your draft against three voice-corpus samples. If it reads more like
   a content marketer than the corpus, rewrite.

HARD RULES
- Banned: "dream home", "investment ka sahi waqt", "limited units",
  "prime location", "world-class amenities", "don't miss out",
  "game-changer", "in today's fast-paced world"
- No opening greeting on reels. First line is the claim.
- Every claim traces to a Claim Ledger row. Insert [CL-3] markers.
- Sophistication floor: if someone with 3 years in the industry could have
  written this, rewrite it.
- Hinglish must be natural code-switching as it appears in the voice corpus,
  never Hindi words sprinkled decoratively into English sentences.

OUTPUT CONTRACT
Full script/copy, caption, [CL-n] markers, alt text, and one line on which
voice-corpus sample you calibrated against.
```

8.6 DESK — Editor-in-Chief

```
ROLE
You are DESK, editor-in-chief. You do not write. You route, sequence, enforce
gates and protect quality. You are the only agent that can kill a piece.

METHOD
1. Read /scout-signals/. Select 4, balancing brand mix, format mix and topic
   diversity against the last 4 weeks.
2. Generate exactly 7 Founder Deposit questions. Each must be answerable ONLY
   by someone with 17 years of on-ground development experience. Reject any
   question answerable by research.
3. Route work. Enforce the two-revision cap.
4. Assemble the weekly review pack. For each piece include TWO RANDOMLY
   SELECTED claims for human spot-check. Select them yourself, at random.
   Do not let the human choose.

HARD RULES
- Maximum 2 revision passes, then it goes to the human regardless.
- No piece past GATE with an open RED flag.
- No piece proceeds with zero Bucket B content.
- If the week's signals are weak, publish less. Volume is never a reason to
  lower the bar. Say so explicitly in the review pack.

OUTPUT CONTRACT
```

```
Weekly review pack: 4 pieces with draft, Claim Ledger, RED verdict, 2 random
claims for spot-check, format rationale, recommended publish slot.
Deliver to Gmail Thursday 09:00.
```

(ARC, FORM, CANVAS, STUDIO, GATE and ECHO follow the same six-part structure. Build them in Phase 2/3 using these five as the pattern.)

---

9. Two brands, one system

Layer	Shared?
SCOUT signals	**Shared** — one sweep serves both
LEDGER research	**Shared**
RED audit	**Shared**
Insight Bank	**Shared**
Brand constitution	**Separate**
Voice corpus	**Separate**
Visual system	**Separate**
ANGLE / ARC / QUILL	**Separate agent instances**, same instructions, different files
Legal guardrails	**Separate** — Sherringearth carries RERA advertising obligations the personal brand does not

On OpenAI this is cleanest as **two agent sets in one workspace**, distinguished by the files attached to each agent, with the shared research agents publishing to folders both sets read.

Cross-posting rule: Anirudh Dashora content referencing a Sherringearth project must disclose the interest. Sherringearth content featuring you stays institutional in register. GATE enforces both.

---

10. Cost model

Item	Monthly (USD)	Monthly (₹ approx)	Notes
ChatGPT Business seat	~$25	~₹2,200	**Required** for Workspace Agents
Workspace Agent credits	Variable	₹2,000–10,000+	Credit-based since 6 May 2026; agent pricing effective 6 July 2026. **Heavy weekly research runs consume meaningfully more.** Measure in month one
ElevenLabs Creator	$22	₹1,900	PVC voice clone
HeyGen Creator	$29	₹2,500	Optional; skip if not using avatars
Higgsfield Starter/Plus	$15–49	₹1,300–4,300	B-roll
Native IG scheduling	$0	₹0	§7.5
Scheduler for LinkedIn/FB	$0–15	₹0–1,300	Optional
Phase 1 (Plus prototype, §7.4)	**~$20**	**~₹1,750**	Manual handoffs
Phase 3 realistic	**~$150–350**	**~₹13,000–30,000**	Credit consumption is the swing factor

Comparison: a competent Indore social media agency runs ₹80,000–₹2,50,000/month and would not produce Bucket B content, build a Claim Ledger, or know MP RERA advertising rules.

Platform cost note: this path runs higher than the Claude path in Document 1, principally because Workspace Agents require a Business seat plus consumption-based credits, whereas Claude Cowork is included on individual paid plans. That is a real difference and worth weighing. **Run one full week and measure actual credit burn before committing to annual anything.**

Prices verified 21 August 2026, subject to change. FX ~₹87/USD. Indian GST may apply.

---

11. Measurement and the learning loop (ECHO)

Do not optimise for likes. Optimise for authority.

Primary:

- **Share rate** (shares ÷ reach) — the strongest proxy for "worth passing on." When shares outpace likes, content has moved from passive consumption to active distribution.
- **Save rate** — reference value for educational carousels.
- **Comment depth** — count comments over 15 words. Ten thoughtful comments beat 200 emoji replies for your positioning.
- **Profile-visit rate** and **DM-initiated conversations** — where the business actually is.

Secondary: reel retention curve (where do they drop?), follower quality by archetype, inbound of the right type.

ECHO's job: correlate performance against *structural* variables — angle type, format, hook style, Bucket B presence, day/time, brand. Write to `/04-performance/learning-log.md`. Every agent reads it.

Critical discipline: require **six or more data points** before ECHO changes any rule. Single-post variance is enormous; an agent that rewrites strategy after one good reel sends you chasing noise. Encode the threshold explicitly.

---

12. Honest limits and what will break

12.1 What this system cannot do

- **Generate original insight.** It retrieves, structures, verifies, positions, produces. Bucket B comes from you.
- **Predict virality.** See §12.2.
- **Replace on-camera presence** for opinion content without a credibility cost.
- **Access competitor Instagram data.** See §12.4.
- **Exercise legal judgment.** Have a lawyer review the RERA guardrails once. Encode their words. Do not have an agent interpret the statute.

12.2 On virality scoring

Some platforms genuinely derive patterns at scale — Higgsfield, for example, defines virality

internally by engagement-to-reach ratio with a focus on share velocity, analysing short-form video at scale to extract repeatable structures.

But those patterns come from *their* corpus, not Indore real estate, and they measure correlation with past performance. **Treat any virality score as a hypothesis generator, never a gate.** Your differentiation strategy is explicitly to do what the pattern-matched majority is not doing; a scorer trained on the majority will systematically penalise exactly that. Never let an agent kill a piece because a scorer disliked it.

12.3 Platform risk — the most important non-content risk

This document opens with a deprecation warning for a reason. In roughly eight months OpenAI shipped Agent Builder, deprecated Agent Builder, deprecated Evals, launched Workspace Agents, announced the deprecation of Custom GPTs for organisations, and retired Sora.

Mitigations:

1. **Keep the intelligence in files, not in the platform.** Your constitution files, Insight Bank, Claim Ledgers and learning log live in your own Drive as plain markdown. If the platform changes, you re-point agents at the same files. This is the single most important architectural decision in this document. A system whose value lives inside a vendor's UI is a system you rent.
2. **Prefer MCP-speaking tools.** Tool integrations that speak MCP port to any runtime.
3. **Keep agent instructions as text files you own**, versioned, not only inside the builder UI.
4. **Do not sign annual contracts** on any component in the first quarter.

12.4 Instagram competitor analysis is structurally limited

Official Instagram APIs in 2026 only permit access to accounts that have authorised you. Public discovery endpoints for arbitrary users and broad hashtag search are not available. You **cannot** legitimately build an agent that scrapes competitor Instagram accounts.

Workaround: a manual weekly ritual. Ten minutes, screenshot the five most notable competitor posts, drop them into `/05-competitors/` with one line of context. An agent analyses positioning, format and gaps from your observations. Lower volume, fully legitimate, and honestly more useful — you'll notice things a scraper wouldn't.

12.5 Where sources conflict — verify before relying

Per §4.4 rule 7, flagging rather than picking:

1. **Instagram API publishing cap.** Meta's own documentation states 100 API-published posts per 24-hour moving period, carousels counting as one. Third-party sources report effective caps of 50 or 25. Irrelevant if you follow §7.5 and schedule natively (25/day).

2. **HeyGen credit allocations.** Sources report Creator at both 200 and 600 monthly credits, with materially different burn rates for premium engines. Check `heygen.com/pricing` live and measure one real video before annual commitment.
3. **MP RERA portal URL.** Both `rera.mp.gov.in` and `reraonline.mp.gov.in` appear in current sources. Confirm the live one and hard-code it into the source registry.
4. **Workspace Agent credit pricing.** Public reporting on the exact rate card is inconsistent. Confirm against your own OpenAI billing page before budgeting.

12.6 Failure modes and fixes

Failure mode	Symptom	Fix
Voice drift	Output becomes generic content-marketing English	Re-anchor QUILL to the voice corpus monthly; add 3 samples per quarter
Verification theatre	RED approves everything	Track RED's flag rate. Below ~15% means it's being agreeable — sharpen the adversarial framing
Loop degradation	Content gets more hedged each revision	Enforce the 2-pass cap
Bucket A drift	Content is accurate, well-designed, forgettable	Check the Bucket B gate is actually blocking. It's the gate people quietly disable
Handoff breakage	An agent runs on stale input	Timestamp every folder write; DESK checks freshness before assembling the pack
Credit surprise	Bill much higher than expected	Measure burn in month one before annual commitments
Volume creep	Publishing more, engaging less	DESK is authorised to publish less in a weak week. Hold that line
Platform change	A tool is deprecated	§12.3 mitigations

12.7 Verification log

Everything factual here was checked against live sources on 21 August 2026: OpenAI Agent Builder and Evals deprecation notices (OpenAI deprecations page, AgentKit announcement, developer community); Workspace Agents availability, plans, pricing dates, API behaviour and

admin controls (OpenAI announcement, help centre, developer docs, chatgpt.com/features); Custom GPT deprecation direction (OpenAI statements and reporting); Sora shutdown dates (OpenAI help documentation and reporting); ElevenLabs pricing and PVC requirements (elevenlabs.io/pricing and docs); HeyGen pricing and API rates (heygen.com/pricing, developers.heygen.com, help centre); Higgsfield models, MCP and plans (higgsfield.ai); Instagram publishing rules and native scheduling (Meta developer documentation and current platform reporting); MP RERA portal and thresholds (rera.mp.gov.in and current guides). Four unresolved conflicts in §12.5. Re-verify anything cost-relevant before subscribing.

---

13. Your first three actions

1. **Write `02-voice-corpus.md` today.** Two hours. Nothing else works properly without it, and it's the only file no tool can generate for you.
2. **Decide the plan question this week.** If you're on Plus, either upgrade to Business (needed for Workspace Agents) or run the §7.4 Custom GPT prototype for two weeks to validate the content system before paying for the automation system. Do not build past Phase 1 on Custom GPTs.
3. **Record your first Founder Deposit next Monday.** Ten minutes. The entire strategic argument of this document rests on that habit.

Appendix — Claude vs OpenAI, honestly

Neither platform is wrong. The differences that actually matter for *this* build:

Factor	Claude path (Doc 1)	OpenAI path (Doc 2)
Entry cost	Included on individual paid plans (~$20)	**Requires Business seat** + credits
Non-technical build	Cowork + Skills; folder-native	Workspace Agents; plain-language builder
Scheduling	`/schedule` , cloud-run	Schedules on published agents, cloud-run
Code execution for verification	Available	**Codex-powered — stronger here**
Agent chaining	Sub-agents, parallel execution	Shared-folder or Slack handoffs

Design production (HTML→PNG)	Artifacts render inline for fast iteration	Generate file, open in browser
Platform stability risk	Lower recently	**Higher — see §12.3**
Long-form writing quality	Generally strong	Generally strong

If forced to choose one: the Claude path is cheaper to start, has lower recent platform churn, and its inline artifact rendering makes carousel iteration materially faster. The OpenAI path has the stronger code-execution story, which is genuinely valuable for mechanical Claim Ledger validation.

A pragmatic hybrid worth considering: build the editorial system on the Claude path (Document 1) for cost and stability, and keep a ChatGPT Business seat for the one or two agents where Codex-style structured-data validation genuinely earns its keep. Do not run twelve agents twice — that doubles cost and halves your ability to debug either system.

Whichever you pick, §12.3 mitigation 1 is the rule that protects you: keep the intelligence in your own files. Then the platform is a replaceable component rather than a dependency.

End of Document 2.