

AI Marketing Operating Model

Ownership, evidence and human decisions



HOW TO USE THIS RESOURCE

AI Marketing Operating Model

CMOs, marketing operations and product marketing leaders

Your practical outputs

- Workflow portfolio and decision map
- Named ownership and handover contracts
- Evidence and review operating cadence
- Release gates and economics worksheet

Read, choose, apply

Start with the overview and decision rules. Choose a concrete constraint in your business. Work through the relevant templates using your own evidence, record open assumptions and agree a review with the accountable people. The complete online edition provides linked sources and contextual navigation.

Editorial context

Explicitly fictional cases and example figures are working aids, not client results or universal benchmarks. Frameworks structure decisions; they are not scientifically validated scores. Platform capabilities may change. Sources and review context are disclosed.

Read online: [AI Marketing Operating Model](#)

CONTENTS

1. An operating model is a set of decisions	4
2. Build a portfolio of workflows, not a list of tools	4
3. Create a service catalogue with acceptance contracts	5
4. Define the boundary before choosing technology	6
5. Make ownership visible at every handover	7
6. Measure total work and conservative economics	7
7. Design a cadence that produces decisions	8
8. Separate shared standards from domain judgment	9
9. Turn evaluation into a decision system	10
10. Protect the evidence chain and action boundary	10
11. Plan adoption around the actual bottleneck	11
12. Install the model with a working transition plan	12
13. Diagnose failures before changing prompts	12
14. Release in stages and preserve the manual exit	13
Turn this resource into a team brief	15
Sources & further documentation	16

1. An operating model is a set of decisions

The core question is who can decide what, with which evidence, at which point in the work. A tool catalogue cannot answer it. Start by tracing a marketing deliverable from request to accepted use. Mark every moment where someone interprets a business objective, chooses a source, makes a claim, approves an audience, or authorizes distribution. Those moments are the real design surface.

The original framework in this guide is OED: Ownership, Evidence, Decision. Ownership names the person accountable for the next transition. Evidence defines what that person must see. Decision states the permissible dispositions. Apply OED to a campaign brief before considering organization charts. A researcher may assemble evidence, but the product marketer decides whether a claim is valid. An AI-generated score may inform prioritization, but a budget owner approves allocation. Record these boundaries in plain language that a new colleague can follow.

Build the first decision map

List the current steps using actual recent work. Next to each step, name the decision that makes the next step legitimate. Ask which decisions can be made from approved evidence and which require judgment about strategy, relationships, or reputational consequences. Automate preparation where useful; retain explicit judgment where context is not represented reliably.

This framework is an original operating aid, not a certified standard. The [NIST AI RMF](#) offers a broader voluntary risk-management reference. A small marketing team should adapt its operating depth to its actual exposure. The purpose is reliable accountability, not generating ceremonial documentation that nobody uses.

2. Build a portfolio of workflows, not a list of tools

Inventory repeatable work before collecting subscriptions. Include content adaptation, campaign research, sales enablement, customer insight synthesis, lifecycle copy, and reporting commentary. Divide broad activities into units that have stable inputs and an acceptance decision. “Content” is not a workflow; “adapt an approved launch narrative into an internal webinar brief” can be.

Score candidates on input readiness, business importance, review burden, data sensitivity, action exposure, and reversibility. Keep each dimension visible rather than collapsing everything into one impressive score. A high-value workflow may be a poor first pilot because evidence is scattered or the cost of a wrong answer is substantial. A modest, repetitive workflow may provide a better learning environment.

Candidate	Evidence readiness	Human decision	Initial disposition
Internal brief adaptation	Approved narrative available	Product claim acceptance	Consider pilot
Public competitor claim	Evidence changes often	Comparative substantiation	Require source work
Automatic prospect outreach	Personal/contextual data	Recipient and send approval	Exclude initially
Monthly commentary	Definitions need agreement	Business interpretation	Repair metrics first

Portfolio worksheet

For each candidate record the request type, accepted output, current monthly volume, source steward, business owner, sensitive fields, manual route, and expected bottleneck. Mark every volume or benefit estimate as measured, inferred, or unknown. Unknown is useful information. It tells the team what to investigate instead of inviting an invented number.

Review the portfolio whenever a workflow changes permission or source class. Remove retired candidates and merge duplicates. A manageable portfolio gives operations a way to see aggregate review demand, while individual teams retain ownership of the decisions that matter in their domain.

3. Create a service catalogue with acceptance contracts

A service catalogue makes a workflow requestable and predictable. Each entry states what the team will receive, what it must provide, what will be refused, who reviews the output, and what completion means. Avoid promising response times before observing arrival patterns and reviewer capacity. A model can generate drafts quickly while the human queue grows without limit.

For NorthstarDesk, the FICTIONAL B2B SaaS company used throughout this guide, the initial service is “approved narrative to internal campaign brief.” All volumes and examples in this guide are illustrative assumptions, never client results. The requester supplies an approved narrative version, product scope, audience, and campaign purpose. The service returns a structured brief with claim-source pairs and open questions. It does not return a ready-to-publish campaign or authorize any spend.

Acceptance contract template

- Entry requirement: approved source ID, intended use, required language.
- Delivered object: draft brief, source list, unresolved questions, version stamp.

- Acceptance criteria: no unsupported claim, correct scope, usable structure.
- Return reasons: missing source, contradictory evidence, unclear audience, capacity limit.
- Service owner and substitute reviewer: named people, with availability recorded.

Pilot the request form with an actual colleague. If filling it out requires more work than the old process, simplify the interface without deleting essential evidence. Allow a requester to say “unknown” and route that item to clarification rather than forcing guesses. Measure returns because repeated incomplete requests indicate a service-design problem. At handover, the requester should know whether the work is accepted, waiting, or returned. Clear status reduces coordination effort and prevents a draft from acquiring authority merely because it looks finished.

4. Define the boundary before choosing technology

This resource uses a deliberately bounded definition of an AI workflow: an input arrives, a permitted transformation takes place, an accountable person accepts or rejects the result, and the work reaches a defined destination. An assistive workflow drafts or analyses within this boundary. An autonomous agent can choose intermediate actions or tools. These are different operating arrangements, even when both use the same language model. The [Anthropic architecture discussion](#) is a useful starting point for that distinction; its older tooling examples should not be treated as current product instructions.

Write a workflow contract before testing prompts. Specify the initiating event, eligible inputs, forbidden information, permitted outputs, tools, human decisions, and expiry date. “Help marketing” is too broad. “Prepare an internal comparison of approved product claims for the product marketer to review” has a visible boundary. Do not silently extend an internal drafting permission into authority to publish, email prospects, update a CRM, or spend budget.

Copyable boundary card

- Work item and intended reader: ____
- Input owner and permitted source classes: ____
- Output destination and acceptance owner: ____
- Actions explicitly outside scope: ____
- Manual alternative, escalation channel, review date: ____

Walk one ordinary case and one awkward case through the card. If the team disagrees about whether the awkward case belongs, tighten the definition before building. Record the disagreement because it often reveals an unresolved business policy rather than a model problem. The boundary should also say what happens when the person responsible is unavailable. Waiting, returning the work to a manual queue, or declining the request are valid outcomes. Inventing a substitute approval is not.

5. Make ownership visible at every handover

A workflow needs an owner with authority to stop it, not merely an enthusiastic operator. Separate the business outcome owner, source steward, technical maintainer, reviewer, and incident lead. One person can hold several roles in a small team, but the responsibilities must remain distinguishable. Assign a named deputy and a route for questions when a case crosses boundaries.

Decision	Accountable role	Evidence delivered	Receiver confirms
Admit a source	Source steward	Version, scope, permission	Workflow owner
Accept an output	Domain reviewer	Claim checks, corrections	Publishing owner
Change configuration	Technical maintainer	Test comparison, rollback	Workflow owner
Resume after incident	Incident lead	Containment, retest	Business owner

The sender completes the evidence packet before transfer. The receiver explicitly accepts the work; a task moving into a new column is not proof of acceptance. Use “accepted,” “returned with reason,” and “waiting for decision” rather than a vague “done.” Save the decision timestamp and the configuration version, while limiting sensitive content in operational logs.

Run a handover rehearsal

Choose a plausible difficult output, such as an unsupported performance claim hidden in otherwise good copy. Have the reviewer return it with a reason code. Ask the maintainer to distinguish a source defect from a prompt defect. Then ask the business owner whether the workflow continues while that defect is investigated. Rehearse the deputy taking over without an oral briefing. If the deputy cannot locate the approved source, current boundary, and stop procedure, the handover is incomplete. This exercise tests the organization, not the intelligence of the model. Budget for it as operational work rather than expecting it to happen after the launch.

6. Measure total work and conservative economics

Use the same unit for the manual baseline and the assisted process: an accepted work item, not a generated draft. Measure active production, source preparation, checking, correction, coordination, failed attempts, and incident work. Record waiting time separately because reducing elapsed time does not automatically release labour capacity. Keep exceptional cases in the record instead of quietly dropping them.

The following is a FICTIONAL B2B SaaS planning example for NorthstarDesk. Every number is an assumption, not a client result, benchmark, forecast, or current vendor price. Assume 60 accepted items per month, a manual baseline of 70 minutes per item, and a loaded labour rate of €60 per hour. Assisted production takes 20 minutes, review 20, and correction 10: 50 minutes total. The nominal capacity difference is 20 hours, valued at €1,200. Assume €250 tooling and €300 monthly maintenance. The recurring net capacity value would be €650. A fictional €2,600 setup effort would need four comparable months to recover at that rate.

Scenario assumption	Assisted minutes/item	Recurring net value	Interpretation
Difficult review	65	–€250/month	Stop expansion
Planning case	50	€650/month	Continue measuring
Easier review	40	€1,250/month	Check quality parity

These values exclude revenue attribution and cash savings from staffing changes. Freed time only creates value when someone uses it productively. Track where that time goes. If quality differs, compare costs per accepted item at the required quality floor rather than declaring a winner on speed alone. Keep setup costs separate, disclose missing cost categories, and recalculate with observed ranges before committing to expansion.

7. Design a cadence that produces decisions

The operating cadence should answer different questions at different intervals. At the work-item level, the reviewer decides whether a result is acceptable. In a weekly operating review, the owner examines defects, queue age, reviewer effort, and blocked decisions. In a periodic portfolio review, leadership changes permissions, funding, or priorities. Do not send every copy edit to the steering meeting or hide a serious incident in a monthly slide.

Use a short decision log rather than a long narrative status report. Each record names the issue, evidence, options, decision, owner, due date, and recheck trigger. “The team discussed AI quality” is not a decision. “Keep external publication disabled until the product-claim regression set passes and the publishing owner accepts the review process” is actionable.

A useful operating agenda

Start with open STOP conditions and overdue containment work. Then review changes in source quality and permissions. Examine quality and cost by workflow segment, including rejected work. Resolve ownership conflicts next. Finish by choosing the smallest useful experiment and explicitly stating what remains unauthorized.

Do not interpret low incident counts as proof of a healthy process. People may be bypassing reporting, reviewers may be rubber-stamping, or the workflow may barely be used. Pair incident

counts with usage, sampling, and qualitative interviews. Ask a practitioner to demonstrate the process rather than merely reporting adherence.

The cadence fails when it becomes a presentation ritual without authority. Give the meeting owner the right to pause a workflow and a route to resolve cross-functional disputes. Keep disagreements visible until decided. A consistent decision process makes cautious expansion possible without pretending that every uncertainty can be removed in advance.

8. Separate shared standards from domain judgment

A central operations function can provide source metadata conventions, permission patterns, evaluation templates, incident routing, and a register of approved configurations. It should not become the sole judge of product truth, brand interpretation, privacy questions, and revenue strategy. Those decisions require domain owners. Conversely, a local team should not bypass shared access restrictions because its campaign is urgent.

Design a two-layer contract. The shared layer specifies minimum evidence, traceability, testing, and containment. The domain layer defines acceptable claims, audience needs, channel constraints, and review expertise. Changes to the shared layer require cross-workflow assessment. Changes to a domain rule require the relevant subject expert and a targeted retest.

Test a conflict explicitly

Imagine that a product marketer wants to use a roadmap statement in a campaign brief, while the source steward marks that document “internal planning only.” The model should not resolve the disagreement through persuasive rewriting. The item returns to the product and source owners. They can approve a different public statement, narrow the task, or decline the request. Record the outcome in the source register so the same disagreement does not recur in every campaign.

Small teams can combine roles while preserving this distinction in their records. Large teams may need delegated reviewers by product or region. In both cases, keep a discoverable authority map. A colleague should be able to identify the owner of a source, a permission, and a release decision without knowing the organization's informal relationships. This reduces reliance on heroic individuals and makes transitions between agencies, internal teams, and new hires less fragile.

9. Turn evaluation into a decision system

A test set is a deliberately selected sample, not the entire future population of work. Its pass rate describes those cases under the recorded conditions. It does not establish that a workflow will perform identically on new products, languages, customers, sources, or model versions. Treat a small pilot as a way to discover failure modes and estimate operating effort, not as statistical proof of safety.

Build a case register from ordinary tasks, boundary tasks, missing-evidence tasks, conflicting sources, and intentionally hostile instructions inside retrieved content. Remove or mask information that the evaluation environment should not receive. Give each case an expected disposition: answer, qualify, ask a question, decline, or escalate. For each important claim, record an acceptable source and the limits of that source. A perfectly fluent refusal may be correct; a confident answer may fail.

Practical scoring sequence

- Check hard failures first: unauthorized disclosure, forbidden action, invented evidence.
- Score factual support, task completeness, audience fit, and useful uncertainty separately.
- Record reviewer time and material edits alongside quality.
- Have a second reviewer inspect disagreements and a subset of accepted cases.
- Freeze a holdout set before selecting the preferred configuration.

[OpenAI's evaluation documentation](#) illustrates test data and grading criteria. The operating rubric here is an original management design, not a vendor benchmark. If an automated judge is used, calibrate it against human decisions and keep its disagreements visible. Do not let an average hide a catastrophic case or a weak language segment. Release decisions should combine separate risk gates, quality evidence, economic evidence, and the ability to reverse the change. Save the case IDs behind every aggregate so a later investigation can reconstruct the result.

10. Protect the evidence chain and action boundary

Treat source content as evidence to inspect, not as authority to alter system behaviour. A document can contain valid product facts and hostile instructions in the same file. [OWASP's prompt-injection guidance](#) describes this class of risk; retrieval alone does not remove it. Access control, permitted tools, and approval checks belong outside the model's persuasive text output.

For each important statement, preserve source identity, version or date, relevant passage, intended geography or product scope, and the reviewer decision. A citation is only a pointer. The reviewer must verify that the passage actually supports the claim and has not expired or been

superseded. Retrieval relevance is not truth, and a high similarity score is not permission to disclose a document.

Control	Before processing	During review	On failure
Access	Check user/source permission	Inspect unexpected material	Quarantine and investigate
Evidence	Admit approved versions	Match claim to passage	Qualify or remove claim
Actions	Restrict tool permissions	Confirm exact destination	Block the action
Logging	Minimize sensitive fields	Preserve decision metadata	Restrict incident access

Apply proportional safeguards

Start with read-only access and a draft destination. Keep credentials out of prompts. Use the narrowest source collection that can serve the task. Test whether a restricted user can retrieve material that only a privileged user should see. Ask the technical owner to explain enforcement in the actual system rather than accepting a policy statement as implementation evidence. This is operating guidance, not legal advice or a claim of regulatory compliance. Route legal, privacy, contractual, or sector-specific questions to qualified owners before using affected data. A checklist helps find gaps; it cannot certify that none remain.

11. Plan adoption around the actual bottleneck

Adoption is not the number of accounts created. Useful adoption means eligible work enters the agreed route, reviewers can make informed decisions, and accepted outputs reach their intended use. Start with practitioners who perform the existing workflow and ask them which parts cause delay or uncertainty. A workflow that saves drafting time but doubles review burden may simply move the bottleneck to the scarcest person.

Observe a full work cycle. Compare instructions, context switching, source retrieval, editing, review, and final acceptance. Ask users to record why they chose the manual route. Bypass may reveal a missing feature, unclear permission, poor source coverage, or an unreasonable queue. Do not assume resistance is a training problem.

Adoption worksheet

Capture eligible work, attempted assisted work, accepted work, manual fallbacks, and reasons for abandonment. Pair those counts with reviewer workload and quality. Keep the denominator visible: a high acceptance rate among a tiny subset says little about the wider workflow. If users select only easy cases, describe that limitation before leadership generalizes the result.

Give new users a real demonstration of an accepted answer, a corrected answer, an appropriate refusal, and a stopped workflow. Teach them where to find evidence and how to challenge it. The

[NIST Generative AI Profile](#) is a broad reference for risk conversations; it does not replace local operating decisions. The operating model succeeds when ordinary colleagues can use, question, and stop the service without depending on its original champion.

12. Install the model with a working transition plan

Convert the design into a transition backlog with a deliverable and acceptance owner for every item. Begin with the decision map and a single workflow contract. Add the source register, role assignments, review rubric, baseline measurements, and stop rehearsal. Only then ask for wider use. Keep the backlog tied to operational gaps rather than tool features.

For NorthstarDesk, an illustrative transition would keep campaign briefs in a supervised route while operations collects correction reasons and source owners resolve ambiguous claims. The product marketer accepts content quality; operations accepts traceability and queue handling; the business owner accepts the remaining economic uncertainty. None of these acceptances gives the system authority to publish. That would require a separate decision and a new assessment of the action boundary.

Final operating-model checklist

- Every live workflow has a named owner, deputy, and expiry or review date.
- Every source class has an admission and retirement rule.
- Every accepted output has a visible decision and destination.
- Every permission change has evidence and a reversible deployment path.
- Every important metric has a denominator, owner, and known limitation.
- Every incident route has been rehearsed by someone other than the designer.

At the first portfolio review, ask what should be removed. Retiring an unhelpful workflow improves the operating model. Keep the shared standards small enough to maintain and the domain rules specific enough to guide real decisions. The model is a living arrangement of authority and evidence. It should evolve when work changes, while preserving a clear account of why permissions changed and who accepted the consequences.

13. Diagnose failures before changing prompts

When an output fails, preserve a minimally sufficient reproduction: case ID, configuration version, permitted input, retrieved source identifiers, observed behaviour, expected behaviour, and reviewer rationale. Redact or restrict sensitive material. Separate the first visible symptom from the causal stage. An unsupported answer might come from an absent document, a retrieval miss, an ambiguous source, a generation error, or an unreasonable task definition. Rewriting the prompt for every symptom hides this distinction.

Symptom	First diagnostic check	Likely repair path	Retest scope
Missing fact	Is approved evidence present?	Source admission or abstention	Missing-evidence cases
Wrong version	Was supersession metadata applied?	Index and version control	Time-sensitive cases
Unsupported claim	Does the passage entail it?	Claim constraint and review	Claim-heavy cases
Unauthorized action	Which boundary allowed it?	Permission and approval repair	All action tests

Contain, investigate, decide

Stop the affected route immediately for suspected unauthorized disclosure or action. Preserve the evidence, identify who received the output, and notify the designated incident owner through the approved channel. Do not erase the record by simply regenerating a better answer. The incident owner determines containment and any obligations with qualified specialists.

For less severe defects, record frequency, severity, affected segment, correction effort, and whether the manual process has the same weakness. Change one causal layer at a time when practical, rerun the failed cases, and then run a wider regression set. A fix for one product language can break another. Close a defect only when the owner has accepted the repair and the operational documentation matches it. Unknown cause is a legitimate status. It should lead to a narrower permission or a paused rollout, not a fictional explanation.

14. Release in stages and preserve the manual exit

Use a release ladder with explicit permissions: offline experiments, shadow operation alongside the manual process, a supervised pilot, bounded team use, and only then a considered expansion of scope. These are permissions, not calendar promises. Passing a date does not pass a gate. A team may stay permanently at assistive drafting if autonomous actions add cost or risk without enough useful benefit.

The owner signs a release packet containing the boundary card, approved sources, configuration version, test results by segment, unresolved defects, measured economics, reviewer capacity, and rollback instructions. Define STOP rules before reviewing the results. Recommended hard stops for this operating design are an unresolved unauthorized disclosure, a forbidden external action, absence of an accountable owner, and inability to restore the manual route. Quality or cost gates can trigger redesign rather than permanent abandonment, but the decision must be recorded.

Copyable release decision

- Decision: stop / redesign / continue bounded / expand one dimension.
- Evidence reviewed and important limitations: ____
- Unresolved defects accepted by a named owner until a date: ____
- Permission change, affected users, rollback trigger: ____
- Next observation window and person who can stop the release: ____

Expand one dimension at a time: volume, user group, product family, language, data sensitivity, or action permission. Training is not a substitute for technical restrictions, and a good pilot in English does not release a German workflow automatically. The receiving owner should run a real case, a missing-evidence case, and a stop rehearsal before accepting the handover. At the next review, examine actual use, bypasses, reviewer workload, and stale sources. Retire a workflow when maintenance exceeds value or ownership disappears. A graceful shutdown is a successful operating decision.

Primary references and limits

Sources checked 24 October 2026. Technical documentation changes; verify implementation details before release. Frameworks, worksheets and fictional assumptions are original operating proposals, not vendor guarantees.

- [NIST AI Risk Management Framework](#)
- [NIST Generative AI Profile](#)
- [OpenAI: Working with evals](#)
- [OpenAI: Agents](#)
- [Anthropic: Building effective agents](#)
- [OWASP: Prompt injection](#)
- [Lewis et al.: Retrieval-Augmented Generation](#)
- [OpenAI: File search](#)

YOUR TRANSFER CANVAS

Turn this resource into a team brief

Use this page for a joint review. Agree one practical output, observable evidence and an accountable owner. Add the specific methods and templates from the relevant chapters.

Working field	Your agreement
Priority output	Workflow portfolio and decision map
Your starting point	Finding / source / date: _____
Smallest useful test	_____ _____
Evidence & counter-evidence	_____ _____
Ownership & review	Owner / reviewer / date: _____
Stop or change when ...	_____ _____
Decision & next action	_____ _____

Keep unresolved assumptions explicitly open. A completed worksheet is not a substitute for evidence or accountable judgement.

Sources & further documentation

Primary sources checked 24 October 2026. Sources support the linked technical or factual statements; the working models and fictional cases are original editorial decision aids. Recheck changing platform details before implementation.

1. [NIST AI Risk Management Framework](#)
2. [NIST Generative AI Profile](#)
3. [OpenAI: Working with evals](#)
4. [OpenAI: Agents](#)
5. [Anthropic: Building effective agents](#)
6. [OWASP: Prompt injection](#)
7. [Lewis et al.: Retrieval-Augmented Generation](#)
8. [OpenAI: File search](#)

Next step

If you want help adapting the working plan to your product, evidence and team capacity, request a non-binding strategy conversation.

patrickwings.com