

AI Opportunity Audit

Retail & E-commerce

A practical assessment of where AI can reduce friction, where it should not be trusted yet, and which pilot deserves investment first.

The recommendation is not “add AI everywhere.” It is to remove one measurable bottleneck without weakening customer trust.

Prepared by

Keith Soyka

AI Implementation & Systems Advisory

Scenario

Fictional mid-market home and lifestyle retailer

Online store + two physical locations

01	Decision brief	2
02	Current-state diagnosis	3
03	Opportunity portfolio	4
04	Pilot design & controls	5–6

Start with the service desk, not a storewide transformation

This sample audit models a retailer with growing online order volume, a lean service team, fragmented policy information, and frequent “where is my order?” contacts. The cleanest first investment is a retrieval-grounded service copilot that helps staff answer routine questions faster while keeping refunds, exceptions, and customer-facing sends under human control.

PRIORITY

01

Service copilot

Best balance of impact, data readiness, and reversibility.

NEXT

02

Catalog enrichment

Useful after product data quality is measured and ownership is clear.

LATER

03

Demand forecasting

Potentially valuable, but dependent on clean history and planning discipline.

Go/no-go recommendation: approve a constrained discovery sprint. Do not approve full implementation until ticket samples, policy sources, integration access, and baseline measures are verified.

WHAT APPEARS READY

- Repeated, classifiable questions create a stable pilot boundary.
- Existing policy and order information can ground responses.
- Staff review provides a natural quality-control layer.
- Value can be measured against a pre-pilot baseline.

WHAT REMAINS UNPROVEN

- Actual contact volume by intent and channel.
- Quality and freshness of policy documentation.
- Application programming interface access to order data.
- Team adoption constraints and escalation rules.

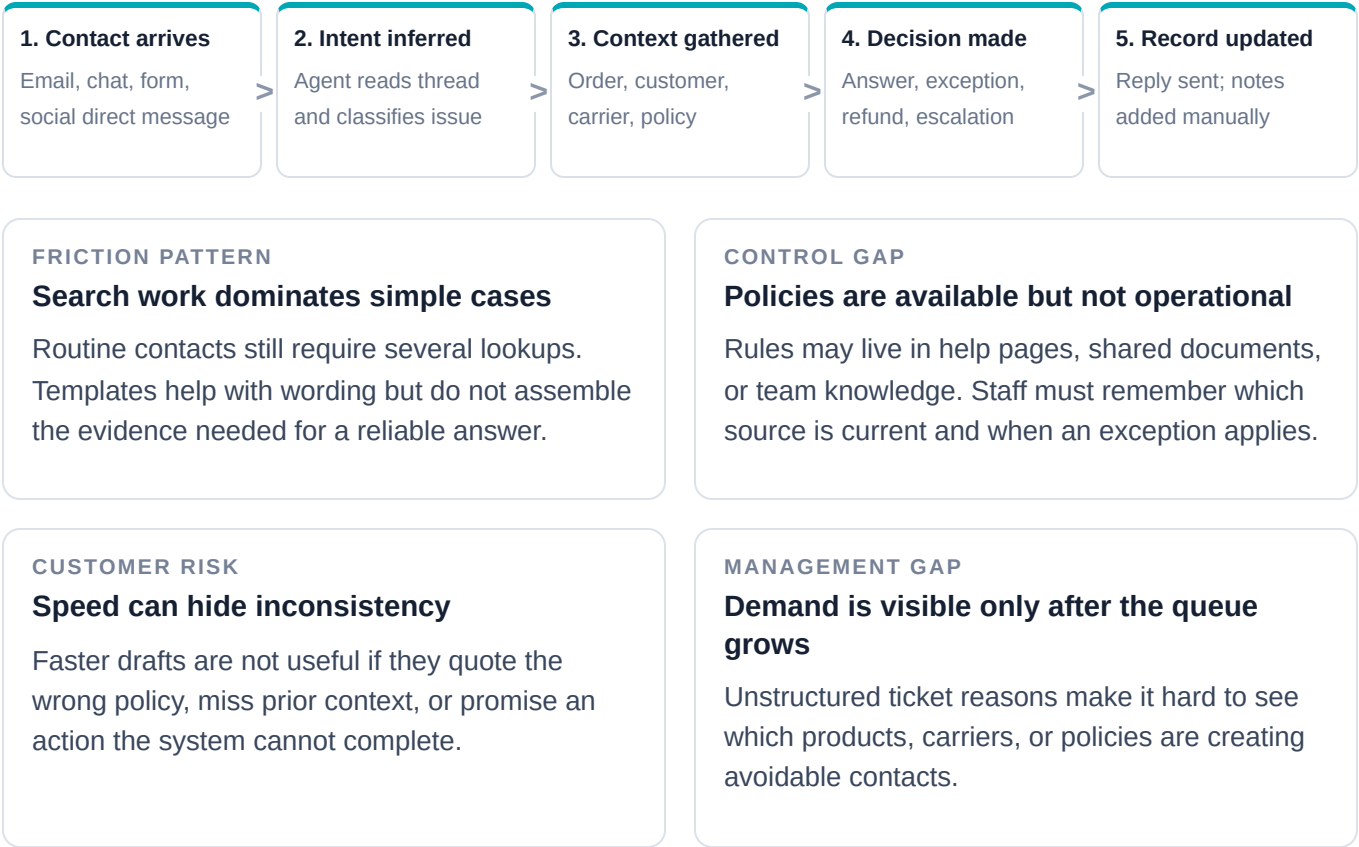
Audit standard

A real engagement would replace every modeled assumption in this sample with client evidence: ticket exports, workflow interviews, system access, policy documents, and baseline performance data. Recommendations would change if those inputs contradict the scenario.

CURRENT-STATE DIAGNOSIS

Customers wait while staff reconstruct context across systems

The central bottleneck is not the act of writing a reply. It is the search: locating an order, checking delivery status, finding the right policy, interpreting exceptions, and documenting the outcome across disconnected tools.



OPPORTUNITY PORTFOLIO

Five ideas are viable; two are ready to test

Opportunities are ranked using a weighted sample rubric: business impact (35%), feasibility (25%), data readiness (20%), risk (10%), and time to evidence (10%). Scores illustrate the method; they are not client results.

OPPORTUNITY	IMPACT	FEASIBILITY	DATA	RISK	SCORE	DISPOSITION
Service agent copilot Retrieve context + draft response	5	4	4	4	4.35	Pilot
Product catalog enrichment Attributes, taxonomy, draft copy	4	4	3	4	3.85	Validate
Voice-of-customer analysis Theme and defect signals	4	4	3	3	3.75	Validate
Marketing content adaptation Channel-specific draft variants	3	5	3	3	3.65	Queue
Demand forecasting assist Planning scenarios + anomaly flags	5	2	2	2	3.10	Defer

WEIGHTED SCORE



PORTFOLIO DECISION

High value / Ready
Service copilot
Catalog enrichment

High value / Prepare
Demand forecasting assist

Useful / Ready
Voice-of-customer analysis
Content adaptation

Low priority
No initiative enters delivery without a measurable owner.

Why the top idea wins: it is bounded, reversible, and observable. It can create value without autonomous customer communication or financial authority.

RECOMMENDED PILOT

A governed copilot assembles evidence; staff own the decision

The pilot should support three narrow intents: order status, return eligibility, and product-policy questions. It should retrieve approved information, show its sources, draft a response, and route exceptions to a human. It should not send, refund, cancel, or edit customer records autonomously.

Discover
Map intents, sample tickets, owners, policies, and access limits.

Prototype
Retrieve approved sources and draft answers in a safe test set.

Pilot
Deploy to a small agent group with review required.

Decide
Compare results to baseline and approve, revise, or stop.

IN SCOPE

- Intent classification for the three approved categories.
- Read-only lookup of order and shipment context.
- Retrieval from versioned policy content.
- Draft answer with source links and confidence cues.
- Human approval before any customer-facing action.
- Logging for evaluation and incident review.

EXPLICITLY OUT OF SCOPE

- Autonomous refunds, cancellations, or credits.
- Fraud decisions or account restrictions.
- Free-form legal, medical, or safety advice.
- Training on private customer data by default.
- Unreviewed policy changes or inferred exceptions.
- Replacing the help desk or commerce platform.

DECISION LOGIC

CONDITION	COPILOT BEHAVIOR	HUMAN RESPONSIBILITY
Source is current and case is routine	Retrieve facts and propose answer	Review accuracy and send
Sources conflict or confidence is low	Do not improvise; surface conflict	Resolve source of truth
Financial or policy exception requested	Summarize context and escalate	Approve or decline exception
Customer disputes prior action	Preserve full thread and flag	Investigate and own resolution

Design principle: the copilot should make the evidence easier to see, not make accountability harder to locate.

Success means better resolutions, not more generated text

Baseline measures should be captured before the pilot. Targets should then be set from observed performance rather than borrowed benchmarks. The measures below define what to compare and how to interpret it.

Median handling time
Minutes from agent open to completed response, segmented by intent.
Watch for: apparent speed gains caused by premature closure or rework.

First-contact resolution
Share of cases resolved without repeat contact inside the agreed window.
Watch for: channel switching that hides repeat demand.

Grounded-answer rate
Share of evaluated claims traceable to an approved source or live record.
Watch for: fluent answers with unsupported policy claims.

Escalation precision
How often the system escalates cases that truly need judgment.
Watch for: over-escalation that simply moves the queue.

Agent adoption
Eligible cases where staff used, edited, or rejected assistance.
Watch for: passive non-use caused by workflow friction.

Customer trust signals
Reopens, complaints, satisfaction comments, and policy disputes.
Watch for: aggregate scores masking high-severity failures.

MINIMUM CONTROL SET

CONTROL	PURPOSE	EVIDENCE RETAINED
Approved-source registry	Prevents retrieval from stale or unofficial content	Owner, version, effective date, review date
Access minimization	Limits customer data and system permissions	Role matrix and access log
Human approval gate	Keeps consequential actions accountable	Draft, edits, approver, final response
Evaluation set	Tests common, edge, and adversarial cases	Case, expected behavior, result, reviewer
Kill switch and rollback	Stops unsafe behavior without platform disruption	Trigger, owner, action, recovery record

Privacy, retention, consent, and vendor terms require review against the retailer's actual jurisdictions, contracts, and policies. This sample is not legal advice.

DECISION & NEXT STEPS

A short discovery sprint can replace assumptions with evidence

The audit has identified a defensible pilot boundary, but the recommendation remains conditional until the retailer's real workflow and data are inspected. Discovery should end with a written go/no-go decision, not an automatic commitment to build.

DELIVERABLE 1

Validated current-state map

Confirmed channels, systems, handoffs, exception paths, owners, and access constraints.

DELIVERABLE 2

Evidence-backed baseline

Intent mix, handling time, repeat contact, quality review, and failure categories.

DELIVERABLE 3

Pilot specification

Inputs, outputs, integrations, permissions, controls, evaluation set, and acceptance criteria.

DELIVERABLE 4

Implementation decision

Proceed, revise, or stop - with costs, dependencies, risks, and an accountable owner.

DISCOVERY AGENDA

WORKSTREAM	ACTIVITIES	EXIT CONDITION
Workflow	Interview operators; observe representative cases	Current-state map approved by frontline owner
Data	Profile tickets, policies, and integration fields	Required sources are usable or gaps are named
Risk	Define prohibited actions and escalation rules	Control owner accepts pilot boundary
Economics	Capture baseline effort and operating costs	Value hypothesis can be measured
Delivery	Estimate build, test, training, and support	Decision package is specific enough to approve

Recommended decision rule: proceed only if approved sources cover the selected intents, access can remain read-only, the evaluation set passes agreed thresholds, and frontline staff confirm the workflow helps rather than adds another screen.

ABOUT THIS DOCUMENT

This is an illustrative portfolio sample for a fictional retailer. It demonstrates audit structure, reasoning, prioritization, and governance. It does not represent a real client's systems, data, or results.

PREPARED BY

Keith Soyka
AI Implementation & Systems Advisory
Retail operations + practical AI systems