

AI Representation Audit

Fifteen starter questions, a six-column log, and the three-type classification key. Run it in a day, fix what is yours, and put the next cycle on the calendar ninety days out.

Brent Passmore · brentpassmore.com

1. How to run it

The discipline is in keeping it identical each time. A single run tells you about a model version. The second run tells you about your own content.

1. **Use a stable question bank.** Start with the fifteen in Section 2, adapted to your organization. Do not change them between cycles.
2. **Run every tool on the same day.** Model behavior drifts. A run spread over three weeks compares nothing.
3. **Use a clean session for each question.** No memory, no prior turns, no personalization. Otherwise you are auditing your own history.
4. **Log the full answer, the citations, and a screenshot.** The prose tells you what it said. The citations tell you why.
5. **Classify by type before you interpret.** Types first, stories later, or you will fix the memorable error and miss the pattern.
6. **Record what you changed and re-ask next quarter.** An audit with no delta is a hobby.

Include the tool your staff already has. If your organization runs Microsoft 365, employees have Copilot, and it is answering internal questions with external content whether you audit it or not. It is the tool most organizations forget and the one with the most captive audience.

2. The fifteen starter questions

Replace the bracketed placeholders with your own. Write them the way a real person types, not the way a communications team writes headlines. Question 15 is deliberate: name something a peer organization offers and you do not, and see whether a tool gives it to you anyway.

Starter question bank, by intent

#	Question	Intent
1	How do I make an appointment with [ORGANIZATION]?	Customer
2	Does [ORGANIZATION] handle [common need]?	Customer
3	Where is [ORGANIZATION] and what are its hours?	Customer
4	Does [ORGANIZATION] accept [payment or coverage type]?	Customer
5	Who do I contact at [ORGANIZATION] about a billing question?	Customer
6	How does a professional refer someone to [ORGANIZATION]?	Referrer
7	What does [ORGANIZATION] offer that a smaller provider does not?	Referrer
8	Who leads [PROGRAM] at [ORGANIZATION]?	Referrer
9	Who is the media contact for [ORGANIZATION]?	Media
10	What is [ORGANIZATION] known for?	Media
11	What has [ORGANIZATION] announced recently?	Media
12	What programs does [ORGANIZATION] offer for [audience]?	Applicant
13	How do I apply to [PROGRAM] at [ORGANIZATION]?	Applicant
14	What are the requirements for [PROGRAM] at [ORGANIZATION]?	Applicant
15	Does [ORGANIZATION] offer [something a peer offers and you do not]?	Invention trap

3. The log

Six columns. One row per question per tool, so fifteen questions across five tools is seventy-five rows.

Read the cited source before you read the answer: three quarters of the time a wrong answer traces to a real page of yours that was outdated, ambiguous, or written for a different audience than the one asking.

Audit log, one row per question per tool

Question	Tool	Verdict	Error type	Cited source	Proposed fix

Verdict values: Correct, Partly wrong, Wrong, No answer. **Error type:** one of the three in Section 4, or None.

4. The classification key

Classify every error before you interpret any of it. The type determines who owns the fix, and whether a fix exists at all.

The three error types

Type	What it looks like	Why it happens	Who owns the fix
Misattribution	Real work, described accurately, credited to the wrong unit.	The wrong unit published more, more recently, or in a format easier to parse. Authority is inferred from the shape of your content, not your org chart.	Both units. Partly fixable with files.
Routing	Everything in the answer is true and none of it gets the person what they asked for.	Your research or background content is better structured than your transactional content, or a site never said what it is not for.	You. The most fixable and the most expensive.
Invention	A service that does not exist, in fluent institutional prose, with no hedging.	Blended sources, peer-organization bleed, or a discontinued page that never came down.	Split. Stale pages are yours; peer bleed is the vendor's.

The fix that is not a file. A meaningful share of what looks like hallucination is old pages nobody unpublished. No signage fixes content that should not exist. And if the wrong unit outranks the right unit because it publishes more, the answer is not better signage: it is that the right unit needs to publish.

Before the first run, decide two things

1. **Who reads the report.** Sent to leadership as a list of things AI got wrong about you, it reads as an outrage document and nothing changes. Sent to whoever owns the content, with a proposed fix per row, things change.
2. **What is severe enough to report to the vendor.** Most errors are yours to fix. Draw the line in advance, in writing, or you will redraw it every quarter based on how annoyed you are.

Brent Passmore writes about accessibility, AI wayfinding, and the unglamorous work of governing a lot of websites at once. This template is the question bank and log used to audit a large academic medical center quarterly, genericized.

Version 1.0. Free to copy, adapt, and use inside your own organization. Attribution appreciated, not required. brentpassmore.com