stellar

A STELLAR PLAYBOOK

UX *of* AI

How to make better decisions about AI throughout your digital ecosystem.

LOS ANGELES
COPENHAGEN
HYDERABAD

2026 EDITION

CONTENTS

What's inside.

Six areas to confidently bring AI into your website and digital platforms.

A little intro about this playbook.

You don't need to be convinced that AI matters. You've sat through the keynotes, read the latest reports, and nodded along in the board meeting where someone said the words "agentic workflows" with great confidence. We know you get it!

What you probably don't have is a clear sense of what to do about it on your website. It's one thing to understand the concept in theory, but it's another entirely to implement it across your actual digital ecosystem, where real customers show up with specific intent, and your brand either helps them or wastes their time. This playbook is built to guide you through how to evaluate a strong AI strategy on your own website.

Bridging Two AI Camps

Most AI advice aimed at marketing leaders falls into two camps. The first explains what AI is, which is well-covered ground by now. The second waves at the future and tells you everything is about to change, which is true but not useful. Neither camp helps you decide what to greenlight next quarter, what to ask your agency, or how to tell whether the chatbot your team built last year is actually earning its keep.

What You Need

You need to know where AI belongs on your digital properties, how it protects your brand voice, what it should do for customers, what it should know about them, and what it should do on its own without being asked. You also need a way to measure whether any of it is working, and the freedom to say "not yet" to things that aren't ready.

A NOTE ON WHAT THIS IS NOT

We're not here to write another AI handbook or add to the technology hype. This document has opinions about where brands are wasting money on AI and where they're under-investing, and we're going to share them. We also know this technology is changing fast, so we're prepared to pivot as the solutions evolve.

Let's get into it.

01 PLAY ONE OF SIX

Where AI Belongs.

Start with what's invisible, not what's visible.

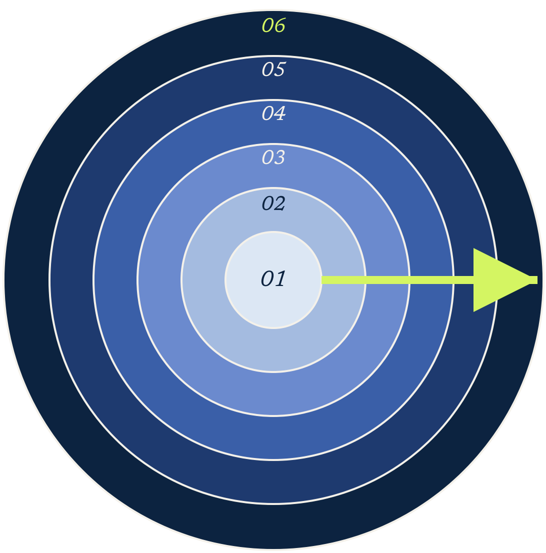
Here's a mistake we often see: a team decides it's time to *"do something with AI,"* and a month later, there's a chatbot on the homepage.

We understand why this happens. A chatbot is visible, and it's easy to point at in a board deck. This gives the organization the satisfying feeling of having made something tangible. If AI progress is being measured by what customers can see, a chatbot looks like progress. **It usually isn't.**

Having a visible AI feature is not going to give your brand real value. The most important work is the work that actually moves the needle for your business. It happens behind the scenes in the unglamorous process of rebuilding your content operations, internal search, recommendation infrastructure, and customer data plumbing. The brands who went chatbot-first often spent 18 months learning that lesson the hard way.

This is the first and most important decision in any AI roadmap. Not "should we use AI," but **"where should we use it first."**

The ecosystem map, from invisible to visible.



01 **Content Operations**
Produce, tag, translate, maintain.

02 **Search & Findability**
On your site and inside external assistants.

03 **Personalization Infrastructure**
Data plumbing first, shiny layer second.

04 **Customer Service & Support**
Only as good as the knowledge base behind it.

05 **Guided Shopping & Journeys**
Service, not just information.

06 **Brand-Voice AI**
Chatbots with personalities. Last, not first.

Think of your digital ecosystem as concentric rings, from internal operations to customer-facing experiences. The innermost work is invisible and undersold. The outermost work is visible and oversold.

Putting this into practice.

If you're starting from zero, start from the invisible ring and work your way out. Each ring addresses a specific part of the ecosystem, and most of the work is less about AI than about getting the underlying material in shape so AI can do something useful with it.

Content Operations. This is the easiest place to start because the work pays back inside a quarter. Translation, image tagging, alt text, metadata, and repeatable copy production at scale are all being handled by AI at every competent agency and every well-resourced in-house team. If your team is still doing this manually, you're wasting resources and falling behind your competitors.

Search and Findability. Two separate jobs, both urgent. The first is fixing what happens when a customer types a natural-language question into your site search. The second is making sure your pages are structured in a way that AI assistants outside your site (ChatGPT, Perplexity, Gemini, Claude) can accurately understand and recommend them. Brands have spent a decade optimizing for Google. The job is much bigger now.

Personalization Infrastructure. Before investing in any AI-driven personalization engine, audit the data feeding it. That means cleaning up duplicate customer records and fixing the points where a single customer becomes two (or more) unconnected records. This happens most often at login-to-guest transitions, cross-device browsing, offline-to-online interactions (call center to website), and anywhere archived customer records fail to merge with active ones. A traditional recommendation engine blindly absorbs noisy identity data and returns mediocre results, but an AI personalization layer amplifies the noise, confidently producing wrong recommendations instead of blank ones. This data cleanup ultimately determines whether every downstream personalization investment succeeds or fails.

Customer Service and Support. An AI support agent is only as accurate as the knowledge base behind it. Before deploying one, review the support documentation honestly. Is it current? Does it contradict itself? Are the edge cases documented? Every flaw in the underlying content will be reflected back by the AI, sometimes louder.

Guided Shopping and Journeys. This ring is worth investing in when the purchase decision is genuinely complex. Insurance, financial products, configured software, paint, renovation materials, enterprise tools. If customers regularly need help deciding, AI can help them decide better. If they just need to click "add to cart," this ring won't deliver much.

Brand-Voice AI. Only worth considering when all five inner rings are solid. Rushing up to this ring can leave brands with a chatbot that customers don't use, or worse, a chatbot that creates incidents the brand then has to manage.

OUR STANCE

Work outward from the center. A year spent improving content operations, site search, and personalization infrastructure will deliver more customer value than a year spent building a flashy chatbot. An infrastructure upgrade also carries zero risk to your brand perception. If you want to be ahead of the game by next year, start the foundational work right now.

A LITTLE WATCH OUT

There's a different risk that doesn't get discussed enough: **doing too little**. The instinct to wait for AI to mature is understandable, but in practice, the gap between AI-mature and AI-immature organizations is widening faster than most leaders realize. Teams that have been iterating in Rings 1 through 3 for 18 months have developed institutional muscle, and those who haven't will struggle to catch up, even if they eventually outspend the earlier movers.

QUESTIONS YOU CAN ASK YOURSELF

Which ring are we actually investing in, and why that one first?

If the answer is "Ring 6, because the executive team wants visible progress," the plan needs rethinking.

What would have to be true in our inner rings for our outer rings to succeed?

This is the sequencing question. Naming the dependencies up front prevents the failure everyone saw coming.

If we only made one AI investment in the next twelve months, which one would pay back the most across the business?

If the answer is obvious, do that one. If it isn't, diagnostic auditing work needs to be done. More on that at the end of this document.

02 PLAY TWO OF SIX

What Your Brand Sounds Like.

The 3 layers of branding your AI voice.

When a brand builds customer-facing AI, the first question usually asked is a technology question. *What model? What platform? What integration?* These matter, but they come second.

The first question should be a brand question: **what does your brand sound like when it's not you?**

For most of branding history, this was answered by hiring good copywriters, writing a voice guide, and trusting that the humans representing the brand had judgment. The guide lived in a PDF nobody opened, but the brand survived because those humans could read a room. They knew when to be warm and when to be brisk. AI has none of that judgment by default. It only has the judgment you give it.

A generic large language model has a voice. However, that voice is the statistical average of the internet, which is to say, it's the voice of everyone and therefore nobody. It apologizes too much, and sounds like a junior employee trying very hard not to get fired. If a brand deploys that voice and puts its logo on it, the brand now sounds like that too. Every customer interaction becomes a chance for the brand to sound exactly like every other brand doing the same thing.

The three layers of brand voice in AI.

ONE

The words it says.

Tone, vocabulary, contractions, sentence length. Does the brand use exclamation points? Does it say "y'all"? Does it refuse to end sentences with prepositions even when it should? This is what most voice and tone guides cover, and it's the easiest to get right.

TWO

The things it talks about.

This layer often gets skipped. A brand AI doesn't just need to know how to speak, it needs to know what it will and won't engage with. Does it answer political questions? Does it discuss competitors? A chatbot that will cheerfully debate anything a customer types is not a feature — it's a lawsuit waiting to happen.

THREE

The things it does.

This is really about brand behavior. It's the deepest layer and the one that most distinguishes a good brand AI from a generic one. Does it proactively help, or wait to be asked? Does it apologize for not knowing something, or direct the customer somewhere useful? Does it end conversations well, or keep trying to upsell?

If you get Layer 1 right and Layers 2 and 3 wrong, your AI will sound like your brand for about eight seconds, until a customer asks it something unexpected. Get all three right, and it'll feel like your brand even in edge cases.

AI CAUTIONARY TALES

In 2024, **Air Canada's** customer service chatbot invented a bereavement discount policy that didn't exist, and a tribunal ruled the airline had to honor it. Around the same time, a **Chevrolet dealership's** chatbot was coaxed into agreeing to sell a Tahoe for one dollar, and the screenshot went viral. **DPD's UK** chatbot, in a now-famous exchange, was talked into writing a poem about how terrible DPD's own customer service was.

Though the failures were different, they shared a common cause: a charming brand voice without the guardrails or judgment needed to keep the system on track.

Putting this into practice.

Training a brand AI is mostly a content and judgment exercise, because the technology itself is a commodity.

Start by auditing what the brand has actually said in the last 24 months across all channels. Customer service transcripts, social responses, email templates, product copy, press responses. This body of work is the raw material for voice training whether or not anyone formalizes it. If it's inconsistent, and it almost always is, that inconsistency will show up in the AI. The AI will reproduce whatever it's trained on, ergo, inconsistent source material will produce inconsistent output.

Then build three kinds of guardrails:

Content guardrails, which govern what the AI knows. Restrict the model to a curated knowledge base rather than the open internet.

Behavioral guardrails, which govern how the AI responds under pressure. Pre-write the hardest scenarios. What happens when a customer is angry? What happens when someone tries to get the AI to say something embarrassing for a screenshot? Issues like these happen every day when dealing with a live AI model.

Escalation guardrails, which govern when the AI hands off to a human. The best brands know their limits and route customers to real people gracefully.

WAYS TO PRESSURE TEST YOUR BRAND VOICE

What does our brand refuse to do? Not what it can't do. What it won't. A brand that can't articulate its refusals in advance will make them under pressure, usually after something has gone wrong in public.

What does our brand sound like when a customer is frustrated? Any voice can be charming when everything is going well. A brand's actual voice reveals itself in difficulty. If the answer is "we haven't thought about that yet," the AI isn't ready to launch.

What would our best employee say here, and why? If a real person representing the brand at its best wouldn't respond a certain way, the AI shouldn't either. The goal is to encode the judgment of your best people, not the average of all of them.

A LITTLE WATCH OUT

Everyone worries about high-stakes failures, but many underestimate a much more common risk: **brand dilution**. Too often a brand's AI integration launches, but it sounds generic, so then customers barely notice it's there. The novelty wears off, and six months later, the brand has spent real money just to make itself less distinctive than it was before. There is no headline-worthy crisis, but the brand is left with a wasted budget and a forgettable customer experience.

QUESTIONS YOU CAN ASK YOURSELF

If a stranger read 50 of our AI's responses without a logo on them, would they know it was us?

If not, the voice work isn't done.

What are the three conversations we'd be mortified to see our AI engaged in, and have we trained the system to refuse them?

Who on our team owns the brand voice of our AI the same way someone owns the brand voice of our advertising?

If the answer is "engineering," something's missing.

03 PLAY THREE OF SIX

What Your Brand Does.

How to be more like the concierge.

Right now, most marketing teams are using AI as an accelerated content machine.

They're cranking out text, images, and translations faster than ever before, and while that's fine, it's a surface-level approach to everything AI has to offer. If you focus entirely on what your AI can say rather than what it can actually do for your customers, your brand is missing the bigger opportunity, which is to build utility.

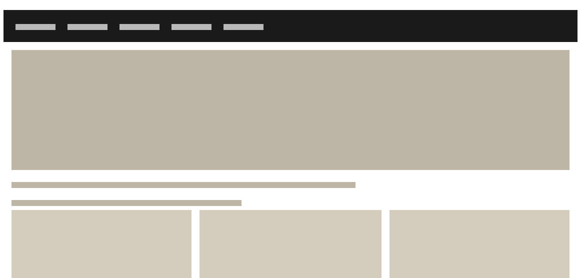
Stop thinking about what your AI can ***say****.*
Start thinking about what your AI can ***do****.*

What this looks like on an actual website.

The shift from content brand to service brand shows up most concretely in how websites are structured.

THEN

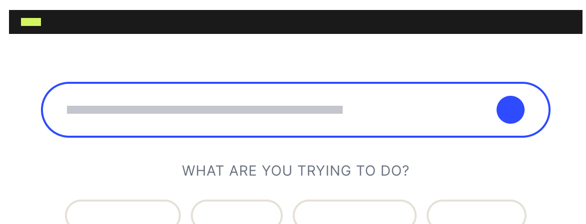
The Catalog.



Organized around what the brand wants to communicate. Navigation strings guide users through static paths.

NOW

The Concierge.



The Service-Oriented Experience. On a service-oriented site, the primary way customers find what they need isn't by clicking through a menu, but by describing what they're trying to do in their own words. The navigation is still there, but just as a fallback.

Most brand websites today are still catalogs with chatbots bolted on. The work of the next three years is turning them into concierges.

What changes does this require?

In practice, this means fewer top-level categories, shorter menus, and a much more prominent input layer.

The product page becomes a decision page. Instead of a detailed specs-and-features layout that assumes the customer already knows what they want, the page asks what they're trying to accomplish and then configures itself around that. A paint brand's product page doesn't lead with thirty color swatches, it leads with "What room are you painting, and what kind of feeling do you want in it?"

The FAQ disappears. Service-oriented sites don't need a wall of static text covering edge cases because the entire interface dynamically answers questions as they arise.

A NOTE ON WHERE THE CONCIERGE LIVES

"Service-oriented website" doesn't mean "website dominated by a big chat window." A chat interface can be part of it, but the more effective patterns are usually embedded. An input bar at the top of the homepage. A "help me figure this out" flow inside a product page. A smart configurator that replaces a spec sheet. A search experience that actually understands what's being asked.

The best service-oriented sites actually feel like they don't have AI on them. For a customer, they just feel dramatically less annoying than they used to.

OUR STANCE

The highest-value AI work on most brand websites right now is replacing **catalog architecture** with **concierge architecture**. Everything else, the chatbots, the content engines, the personalization layers, tends to underperform when grafted onto a site still organized around what the brand wants to say.

Fix the architecture first because the AI features layered on top will work harder. This is also the work most agencies aren't selling, because it's harder than shipping a chatbot and doesn't come with a product demo. It's where the real ROI lives, and the brands doing this work will become untouchable.

A LITTLE WATCH OUT

The risk with service-oriented architecture is that it exposes how weak a brand's underlying knowledge actually is. A catalog site can hide behind marketing copy, but a concierge site cannot. If a customer asks "which of your products is right for me" and the brand doesn't have a real, defensible answer, the concierge breaks immediately.

Teams often discover, partway through the rebuild, that they don't actually have clear answers to the questions customers most want to ask. **That's a product marketing problem that surfaces during an AI build**, and it has to be solved with humans before any AI can help.

QUESTIONS YOU CAN ASK YOURSELF

If a customer could ask our website any question about our category, what's the single question we most want to own the answer to?

This is the question your site architecture should be organized around. Most brands have never explicitly answered it.

Where in the current journey does a customer have to guess, hunt, or compare without help?

These are the places to start. Not the homepage. Not the chatbot. The friction points that already exist.

What do we actually know about our customers' jobs that our competitors don't?

If the answer is "nothing specific," the service opportunity is limited. The concierge can only be as smart as the brand's underlying knowledge.

04 PLAY FOUR OF SIX

What Your Brand Knows.

Personalization is a journey up a trust ladder.

Every customer wants personalization.

The problem is that most personalization initiatives underdeliver, and if this happened at the same rate in any other discipline, it would be a scandal. The brands that get it right see meaningful revenue gains.[12] The brands that get it wrong spend a lot of money making their customers uncomfortable.

The difference between those two outcomes has almost nothing to do with technology. **It comes down to trust.** Personalization works when your customer believes you're using what you know about them in their interest. The moment that belief breaks, personalization starts to feel like surveillance, and it doesn't matter whether you technically complied with privacy regulations. *Legal permission is not the same as customer comfort.*

AI makes this harder, not easier. An AI-powered personalization engine can infer mood from session behavior, predict future purchases, and customize messaging in real time. When it works, it feels like great service. When it misses, it feels invasive in a way that's harder to recover from.

The trust ladder.

Not all personalization carries the same trust cost. Some forms of it customers welcome. Some they tolerate. Some they quietly resent. **Knowing the difference is most of the work.**

Highest Trust	**Explicit** Personalization the customer asked for.	Saved preferences, favorited items, opt-in quizzes. Safest ground, most underused.
High Trust	**Contextual** Based on observable behavior in this visit.	Feels responsive, not predictive. The AI equivalent of a good salesperson paying attention.
Trusted Ground	**Historical** Based on their own prior behavior with us.	Requires honest acknowledgment of the history. Don't pretend not to remember.
Trust Fractures	**Inferred** What we think we know based on patterns.	Often right. When wrong, memorably wrong. When right about something they didn't disclose, they notice.
Lowest Trust	**Surveilled** Data the customer didn't know you had.	The moment they realize you had it, the relationship changes. AI makes that realization more likely.

The brands running the best personalization programs are operating almost entirely in the top three rungs. They're leaving the bottom two alone, not because they can't access the data, but because they've done the math on what it actually costs.

What changes does this require?

Get explicit about what you collect and why. If you haven't already, write down, in plain language, what customer data you actually use and what value the customer gets in return. This should be a legible, public document, not a legal disclosure. Customers are increasingly willing to share data with brands that can clearly articulate the exchange, and by the same token, they're increasingly skeptical of brands that can't.

Move the consent conversation up. The standard industry practice is to treat consent as a checkbox at the bottom of a signup flow, when brands should really treat it as an ongoing, visible part of the experience. This includes a small indicator showing what the brand is currently using to personalize, a "why am I seeing this?" link that actually explains, and a clear way to turn specific kinds of personalization off without losing the account. These may feel like small UX details, but they're actually the most important personalization features a brand can build. Let customers see what you think you know about them.

A NOTE ON THE ZERO-PARTY DATA OPPORTUNITY

The most valuable data a brand can have is data the customer gave them directly and explicitly, for a clear purpose. Zero-party data. This is often dismissed as less sophisticated than inferred behavioral data, because it's volunteered rather than detected. **That framing is backwards.**

Zero-party data is high-trust, high-accuracy, and increasingly scarce. A customer who answers three well-designed questions about what they're trying to accomplish gives a recommendation engine more useful signal than six months of passive tracking, and they do it without any of the trust cost. Learning how to ask well will matter more than extracting insight from exhaust data.

OUR STANCE

Use what you collect more deliberately. The era of data maximalism is ending, and the brands still operating as if more data is always better are going to be caught on the wrong side of a trust reckoning they didn't see coming.

A lot of martech vendors are still selling the dream of the unified customer profile that knows everything. That dream was already fragile before AI. Now that AI can make the implications of all that data visible to customers in real time, the dream is actively dangerous. A customer who realizes a brand knows more about them than they're comfortable with doesn't just disengage. *They tell other people.*

A LITTLE WATCH OUT

The regulatory landscape is moving faster than the personalization tooling. The EU AI Act, evolving state-level privacy laws in the US, and a growing patchwork of sector-specific regulations are all tightening the rules on how customer data can be used for AI-powered decisioning. Personalization programs built on assumptions about what's allowed today may be out of compliance in future months. So it's imperative to keep on top of the latest.

QUESTIONS YOU CAN ASK YOURSELF

If a customer could see exactly what data we use to personalize their experience, would we be comfortable showing it to them?

If the answer is "technically yes, but we'd prefer not to," the trust isn't actually real.

What's the smallest set of data we'd need to deliver 80% of the personalization value?

It's easy to fall into the trap of collecting far more than you need because collection is cheap, but less is really more in this case. Collect only what you actually need and it'll streamline the process in the long run.

05 PLAY FIVE OF SIX

How Your Brand Is Found.

Made for discovery, even if nobody goes to your site.

Content made for discovery, even if nobody ever goes to your site.

For decades, companies have felt the control of building their own website experience, but customer behavior is shifting dramatically. The first (and sometimes only) interaction a customer has with a brand increasingly happens inside general-purpose AI assistants like ChatGPT, Claude, Perplexity, and Gemini. ChatGPT alone draws over 800 million weekly users who use the tool to research purchases and look for recommendations,[3] and if your brand is not prepared for this shift, your customers may never find you.

What "agents" actually means for a brand.

The word "agent" has been doing a lot of work in AI marketing lately, and it means at least three different things depending on who's saying it. Worth separating them, because they have very different strategic implications.

Agents the brand builds and deploys.

A branded AI assistant that lives on the brand's property and acts on the brand's behalf. This is what most of the "agent" conversation in marketing has been about. For most brands, the least important of the three.

Agents the customer runs, that interact with brands.

A customer's personal AI assistant that shops, researches, or transacts on their behalf, potentially across many brands in a single workflow. Already happening in consumer contexts. Will be the dominant interaction model within three years.

Agents operating in the gap between brands and customers.

Third-party AI experiences that recommend, compare, or transact across a category. Travel planning assistants, financial comparison tools, shopping agents. Increasingly the place purchase decisions actually get made.

For most brands, the second and third categories will matter far more than the first. Almost none of the playbooks circulating right now are helping brands prepare for them.

The core strategic shift.

Your brand's job is no longer just to be a *destination*. It's to be a **legible source** that other AI systems can confidently recommend, explain, and transact with on the customer's behalf.

This changes what matters about a brand's digital presence. A decade of web strategy has been about getting customers to your site. The next decade is partly about making sure your brand shows up well when the customer never visits at all.

How to actually influence what AI assistants say about your brand.

This question comes up in nearly every conversation we have with marketing leaders, and the answers are evolving quickly. Here is the current landscape:

Generative Engine Optimization (GEO) shows early evidence that a few techniques work better than others. (You may also hear this work called Answer Engine Optimization, or AEO, particularly when the focus is Google's AI Overviews rather than conversational AI assistants. The tactics overlap substantially. We use GEO as the umbrella term because it covers the broader set of surfaces this playbook addresses.)

What the research shows works. Princeton research tested a range of optimization tactics and found three that consistently increased AI citation rates by roughly 30 to 40 percent:[4]

- **Citing credible sources within your own content.** Pages that reference authoritative third parties are cited more often than pages that don't.
- **Adding specific statistics and data points.** AI systems disproportionately cite content with concrete numbers.
- **Including direct quotations from experts.** Content with attributed expert voices lands in more answers than content without.

The deeper pattern: AI systems are trained to look for content that reads like a verified source rather than marketing copy. The brands that show up most often in AI responses are the ones publishing content that reads like journalism, not like advertising.

The earned media pattern. This is the finding that surprises marketing leaders most. Industry research tracking AI citations has found that AI platforms cite third-party earned media (reviews, press coverage, analyst reports, independent content) 69 to 82 percent of the time, compared to 36 to 45 percent for traditional Google search.[5] Your own website is not your best AI visibility asset. The coverage other credible sources produce about you matters more. This means influencing AI visibility is partly a PR and communications investment, not just a content one.

A note on timelines. AI visibility tends to respond faster than traditional search, though it also decays faster. Realistic expectations, based on current industry data:[6]

0 to 2 months	*3 to 6 months*	*6 to 12 months*
First citation signals after consistent optimization. Early wins usually in long-tail or less-competitive queries.	Meaningful, consistent visibility shifts. Brands start to see themselves cited regularly in category queries.	Durable competitive position. Brands compounding advantages here treat GEO as an ongoing channel, not a one-time project.

AI citation patterns are less stable than search rankings. Roughly half of all content cited in AI answers is less than three months old.[7] Freshness matters more than in traditional SEO, and content requires more frequent maintenance to stay cited.

One caveat. This space is moving faster than any discipline we've worked in. The techniques that work today may work differently in twelve months. Major AI platforms update their models regularly, and each update can shift citation patterns overnight. Which means we need to treat AI visibility as an iterative program with monthly monitoring, rather than a project with a defined end state.

Monitor how your brand shows up, literally. Ask ChatGPT, Perplexity, Claude, and Gemini the questions your customers are asking about your category. See what they say about your brand. See what they say about your competitors. Do this monthly, with a running record. The first time most brands do this, they find surprises.

OUR STANCE

For most brands, investing in how your brand shows up **inside other AI systems** will deliver more strategic value than building your own branded AI assistant. The home-field advantage of your own website and app is smaller than it looks, because increasingly the game isn't being played on home field.

Making your brand legible to third-party AI systems won't produce a demo reel or a good keynote slide. **It's structural work, which is where the compounding advantages live.**

A LITTLE WATCH OUT

The risk in this section is different from the others. It's about being invisible without even realizing it. When a company's visibility depends on search rankings, problems are measurable. You could see where you ranked, watch the traffic, and diagnose what caused a drop. Now that a brand's visibility is shifting to "what AI assistants say about you," problems become much harder to detect. A brand can lose share of voice in their category inside ChatGPT or Claude or Perplexity without anyone in the organization noticing for quarters, because there's no dashboard for it yet. Until things change, a consistent auditing schedule is imperative.

QUESTIONS YOU CAN ASK YOURSELF

What do the major AI assistants currently say about our brand, and how accurate is it?

If no one in the organization has asked this question systematically in the last 90 days, that's the first thing to fix.

What would a well-designed agent, acting on a customer's behalf, need from our digital properties to recommend us confidently?

Structured data, clear pricing, transparent policies, accessible APIs, unambiguous product information. If the answer is "we'd need to build some of that," you've found the work.

Where in our category are purchase decisions starting to happen outside our owned channels?

If decisions are increasingly being made inside AI assistants or comparison agents, your strategy needs to account for those surfaces directly, not just the ones you control.

06 PLAY SIX OF SIX

How You Know It's Working.

Four dimensions for measuring what AI can do.

"We think it's working" shouldn't be an acceptable answer for any business metrics, including your brand's AI initiatives.

If your reports show that engagement is up and the team is moving faster, but you can't actually isolate AI's contribution from the rest of your marketing, it'll be difficult to defend that budget next quarter.

While metrics like page views, time spent on the site, click-through rates, and conversion funnels still matter, they were not built to measure what AI actually changes. If you force your AI initiatives onto those metrics, you'll get incomplete, even misleading, results.

The four dimensions of AI measurement.

DIMENSION 01

Efficiency.

How much time, money, and human effort is the AI saving?

- Content produced per hour
- Support tickets deflected
- Translation costs reduced
- Time to brief

Easiest to measure. Most defaulted to. Table stakes.

DIMENSION 02

Experience.

Are customers actually getting better outcomes?

- Task completion rate
- Right answer vs. wrong answer
- Escalations to a human
- Repeat return for the same question

Harder to measure. Not satisfaction scores. Actual outcomes.

DIMENSION 03

Revenue.

Is AI directly unlocking commercial value?

- Recommendation-driven purchases
- AI-assisted upsells
- Basket size with AI vs. without
- Reactivation from AI-triggered outreach

A/B testing matters. Don't overstate revenue claims.

DIMENSION 04

Risk.

Is the AI increasing the brand's exposure?

- Hallucinated answers per 1,000 interactions
- Brand-voice deviations in outputs
- Privacy edge cases triggered
- Regulatory near-misses

The one nobody wants to measure. The one everybody needs to.

Every serious AI investment should be measurable against at least two of these. Investments that can't be tied to any of them should be scrutinized.

What changes does this require?

Build a measurement plan before the AI investment launches. The number one reason AI measurement fails is that nobody decided what success looked like before the program launched. Retrofitting metrics to a live program is how "we think it's working" becomes the answer. Every AI initiative should have measurable targets in at least two of the four dimensions, defined up front, with a baseline established before launch.

Measure AI against what would have happened without it, not against last quarter. The most common measurement mistake is comparing post-AI results to pre-AI results without accounting for everything else that changed. A 15% improvement in conversion might be because of the AI, or it might be the new homepage, or the market. Isolation matters. For any AI investment big enough to justify serious measurement, running a holdout, a staged rollout, or a proper A/B test is the only way to prove what is actually driving the results.

Make risk a standing agenda item, not a quarterly review topic. The efficiency, experience, and revenue dimensions can survive quarterly review cycles. But the risk dimension cannot. An AI system operating at scale can generate weeks of exposure before anyone notices if risk metrics aren't being tracked continuously. This sounds like over-engineering until the first incident happens, so we recommend getting into practice now.

Measuring AI may take longer than we're used to.

Some AI value shows up immediately (efficiency gains are usually visible within a quarter). The more strategic effects — brand trust, customer lifetime value impact, competitive positioning inside AI assistants — accrue over 12 to 24 months. Programs measured on a quarterly rhythm will systematically undervalue the long-term investments and over-prioritize the short-term ones.

The fix is **dual-timeline measurement**: short-term metrics reviewed frequently to catch problems and validate direction, and long-term metrics reviewed annually to assess strategic position. Most brands only have one of these, but the ones that have both are making better AI decisions.

REFERENCES

1 McKinsey & Company, "The value of getting personalization right—or wrong—is multiplying," *Next in Personalization 2021 Report*, November 2021. mckinsey.com

2 McKinsey & Company, "What is personalization?" May 2023. mckinsey.com

3 Sam Altman, OpenAI Dev Day keynote, October 6, 2025. As reported by TechCrunch: "Sam Altman says ChatGPT has hit 800M weekly active users." techcrunch.com

4 Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., & Deshpande, A., "GEO: Generative Engine Optimization," Princeton University, Georgia Tech, IIT Delhi, and Allen Institute for AI, November 2023. Published in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2024)*. arxiv.org

5 AI citation patterns across ChatGPT, Perplexity, Gemini, and Claude, 2025-2026. Key sources: Muck Rack analysis of 1M+ AI citations finding 82% from earned media (December 2025); Ahrefs study of 75,000 brands finding brand mentions correlate 3x more strongly with AI visibility than backlinks (2025); Chen et al., University of Toronto, arXiv:2509.08919 (September 2025). Aggregated by Machine Relations: machinerelations.ai

6 GEO timeline expectations aggregated from multiple industry reports, 2025-2026. Primary sources: Frase.io, "What is Generative Engine Optimization (GEO): 2026 Guide"; Superlines AI Search Statistics compilation (2026); BrightEdge AI citation volatility data. frase.io

7 Citation recency in AI-generated responses. Perplexity favors content published within 90 days per Frase.io GEO analysis; LLMrefs reports AI citations drop sharply after 3 months; SE Ranking/Superlines (2026) found content updated within 2 months earns 5.0 average citations vs. 3.9 for older pages. frase.io · llmrefs.com

OUR STANCE

Measure across all four dimensions or don't claim to be measuring at all. The brands running AI programs without risk metrics are running AI programs they can't honestly defend. The brands measuring only efficiency are making their case on the cheapest dimension while leaving the strategic ones unclaimed. The brands without revenue isolation are going to lose budget arguments they should win, because they can't tell the CFO what the AI actually earned.

A single-dimension measurement program is a dashboard, not a measurement program.

A LITTLE WATCH OUT

The risk in measurement is falling into the trap of tracking everything just to feel safe. Your metrics must match your business reality. A healthcare company naturally needs stricter guardrails than a clothing retailer, and no framework is a perfect template. Your job is to figure out what actually drives success for your specific business, and focus all your energy right there.

QUESTIONS YOU CAN ASK YOURSELF

What would we measure to know this worked, and did we set that up before it launched?

If you can't answer both upfront, the investment is going to be tough to defend later.

Are we measuring risk as rigorously as we're measuring return?

It's easy to focus on the wins, but if you ignore the risks, you won't have insight into the things that could negatively impact your brand, which makes it difficult to scale safely over the long term.

On which of the four dimensions is our measurement weakest, and what would it take to fix that?

Every brand has a soft spot. Naming it is the first step. Leaving it unnamed is how unmeasured problems become unexpected ones.

WHERE TO GO FROM HERE

The Stellar *6* Point AI Audit.

Reading a playbook is not the same as knowing where your organization actually stands. This playbook covered six decisions, and the audit will tell you where you are on each one.

01 Where AI belongs on your digital properties.

02 What your brand sounds like when it speaks.

03 What your brand does for customers.

04 What your brand knows about them.

05 How your brand is found.

06 How you know any of it is working.

The Stellar 6 Point AI Audit scores your current AI position across all six points, benchmarks you against where peers in your category are operating, identifies the places where the highest-leverage moves are available, and delivers a prioritized roadmap for the next 12 months.

The audit is delivered over 3 to 4 weeks, depending on the complexity of your digital ecosystem. It results in a written strategic assessment, a scored maturity model, and a working session with your leadership team to pressure-test the recommendations.

The Stellar 6 Point AI Audit is for brands ready to take the next step with real diagnostic rigor. Reach out to us directly at uxofai@stellaragency.com *to book.*

Shall we talk?

LOS ANGELES · COPENHAGEN · HYDERABAD

stellaragency.com