

JUNCTURE RESEARCH · WHITEPAPER

# The Content Intelligence Playbook

How a pharma team turns an approved-content estate into measurable AI visibility, and closes the loop on what the machine repeats back.

---

JUNE 2026

JUNCTURE.HEALTH

# The two programs are already one

Most commercial pharma organizations run two content programs that the org chart treats as separate.

The first is internal and operational. Brand and medical teams build a library of pre-approved modules, core efficacy claims with their graphics and references, safety statements, indication blurbs, mechanism descriptions, and reassemble them into emails, detail aids, web pages, and congress materials. Each module carries its own identifier and business rules, gets approved once through MLR, and is reused across markets and channels. Veeva describes this modular content approach as a way to skip full reviews when most of an asset is already pre-approved, attacking a content delivery process that often takes around three weeks ([Veeva](#)).

The second program is external and, until recently, invisible. Health care professionals increasingly ask an answer engine before they read your page. An American Medical Association survey of nearly 1,200 physicians found that 66 percent reported using health care AI in 2024, a 78 percent increase from the 38 percent who used it in 2023 ([AMA](#)). The traffic pattern is shifting underneath the question. Adobe Analytics reported that, in February 2025, traffic from generative AI sources to US retail websites had increased 1,200 percent compared with July 2024, with those visitors browsing 12 percent more pages per visit and showing a 23 percent lower bounce rate than visitors from other sources ([Adobe](#)). The retail numbers are the leading indicator. The behavior, ask the machine first, then maybe click, is general.

The thesis of this playbook is that these are not two programs. They are one. The approved core you reuse inside is the same core the machine learns to repeat outside. Once you accept that, modular content stops being a back-office efficiency play and becomes the asset that determines your AI visibility. Content Intelligence is the discipline of managing that asset on both sides of the line.

# Why approved content is the machine's preferred raw material

Answer engines do not invent a brand's clinical profile from nothing. The credible ones ground their answers in retrieved sources, and the research on how to influence them points in one direction: structure, evidence, and attribution win.

The foundational generative engine optimization study (Aggarwal et al., accepted to KDD 2024) tested which on-page tactics raise a source's visibility inside AI-generated answers. Its top-performing methods, Cite Sources, Quotation Addition, and Statistics Addition, lifted visibility by up to roughly 40 percent on the paper's primary metric, and the authors show the effect varies by domain, so the optimization has to be tailored rather than generic ([GEO paper, arXiv](#)). Empirical work on cross-engine citation behavior, the GEO16 framework analysis, extends this: how often a source gets cited across multiple engines depends on measurable on-page features ([GEO16, arXiv](#)).

Underneath the optimization sits the retrieval mechanism. Retrieval-augmented generation grounds model output in documents the system fetches at answer time, and the attribution literature frames the goal as substantiating each claim with a traceable source, so that a factual statement can be tied back to evidence a reader can verify ([LLM attribution survey](#)).

Read those three findings together and the implication for pharma is precise. A well-formed approved module is already the thing these systems want: a discrete, evidence-backed, source-cited statement that is consistent with the label. The same properties that earn a module its MLR approval, a substantiated claim, a clear reference, fair balance of benefit and risk, are the properties that make it retrievable and citable. The FDA's Office of Prescription Drug Promotion describes its mission as helping to ensure that prescription drug promotion is truthful, balanced, and not misleading ([FDA OPDP](#)). Compliant content and machine-preferred content are converging on the same shape. Your claims library is not exhaust. It is the corpus.

# The estate: from modules to a measured asset

To act on this, the approved estate has to be more than a folder of approved PDFs. Content Intelligence treats it as a structured system of record with three layers.

The first layer is the module and claims library: discrete, identified, MLR-approved units of meaning, each tied to its supporting references and its usage rules. This is the PromoMats-shaped spine that internal reuse already depends on.

The second layer is reuse scoring. Not every approved module is equal. Some are load-bearing, reused across dozens of assets and markets, others approved once and never touched again. Reuse scoring ranks modules by how central they are to the brand's actual communication, which tells you which claims to protect, refresh, and prioritize for external exposure. A claim reused in forty assets is a claim the organization has effectively decided is its brand story.

The third layer is exposure measurement, and this is where Content Intelligence connects to the outside. For each high-value module and claim, you want to know: does this approved statement appear in what the answer engines say? That measure is AI Pickup, the share of your approved content the models echo. AI Pickup converts a soft worry, are the models getting us right, into a number tied to specific modules.

The discipline here is to resist over-building. The point is not a giant content warehouse. It is a small, ruthlessly maintained spine of the claims that matter, scored by reuse and instrumented for exposure, so you can see which approved statements are doing the work inside and outside.

## A worked example: Varigel

Varigel is a fictional brand, used here to make the mechanics concrete.

The Varigel team has a modular estate of about 180 approved modules in their PromoMats-shaped library. Reuse scoring surfaces a short list of load-bearing claims: a primary efficacy claim with its pivotal-trial reference, a tolerability statement, a dosing-convenience claim, and the required safety language. These four

modules appear in most Varigel assets across email, web, and field materials. By the internal definition, they are the brand story.

Content Intelligence then measures AI Pickup for those modules. The team runs the questions an HCP would actually ask, how does Varigel work, what is the dosing, how does Varigel compare on tolerability, across ChatGPT, Gemini, Perplexity, Google AI Overviews, and Claude. The result is a join: for each approved claim, do the engines repeat it, in approved language, with a source they attribute?

The findings split three ways. The efficacy claim has high AI Pickup: the models repeat it close to the approved wording and several cite the brand's own clinical page. Good. Claim uptake here is strong, so the work is to defend it. The dosing-convenience claim has low pickup: the engines either omit it or describe an older regimen, because the most retrievable public source predates a label update. That is a content-exposure gap, an approved claim the brand owns internally but has not made retrievable externally. The tolerability picture is the dangerous one: one engine blends Varigel's profile with class-level statements in a way that drifts from the approved claim and brushes against an unapproved comparison. Because promotion is expected to be truthful, balanced, and not misleading ([FDA OPDP](#)), this is a risk signal, not just a marketing miss.

The compliant-push response is constrained and specific. The brand cannot edit the model. It can make the approved version more retrievable and more clearly attributable: publish the updated dosing claim as a structured, source-cited, label-consistent statement using the already-approved module, exactly the citations-and-statistics shape the GEO research shows the engines prefer ([GEO paper](#), [arXiv](#)). No new claim is invented. The approved core is simply made into the most citable source on the question, and the next measurement cycle checks whether AI Pickup and claim uptake moved. The tolerability drift becomes a flagged item for medical and regulatory review, with the exact prompts and engines documented.

## The phased adoption playbook

Standing this up is a sequence, not a launch.

Phase 1, structure the estate. Get the claims library into a real spine: identified modules, attached references, usage rules, and reuse scoring. Most teams already have the raw material in their content management system. The work is to make reuse and centrality visible so you know what to protect. This phase pays for itself internally through faster assembly and review before any external measurement happens.

Phase 2, instrument exposure. Define the HCP questions per indication and measure AI Pickup for the high-reuse modules across all five engines. Establish a baseline for claim uptake (are the approved claims echoed) and the headline Answer Monitor metrics, Share of Answer, Precision of Answer, and Risk of Answer. The output is a map of where the machine agrees with your approved core, where it is silent, and where it drifts.

Phase 3, close the loop. Route the gaps. Silence and outdated statements go to a compliant-push workflow that makes the already-approved module the most retrievable, citable source on the question. Drift and off-label-adjacent signals go to medical and regulatory review with full evidence. Crucially, both directions reuse the same claims spine, so the team is never inventing new messaging, only changing the retrievability and attribution of approved content.

Phase 4, operate it. Make measurement recurring, not a one-off audit. Label updates, new data, and competitor moves all change the answer landscape, so AI Pickup and claim uptake become standing metrics reviewed alongside traditional channel performance. Resist the urge to build heavy infrastructure for scale you do not have. A focused, well-governed loop on the claims that matter beats a sprawling system nobody maintains.

## Where this leaves you

The strategic point is simple and uncomfortable. For decades, pharma controlled its message by controlling its assets. That control is now mediated by systems you cannot edit, that answer HCPs before your page loads, and that prefer exactly the structured, evidence-backed, label-consistent content your MLR process already

produces. The approved core you built for internal reuse is the lever you have on external visibility. Nothing else you publish is as trusted, as citable, or as defensible.

Juncture's Content Intelligence product is built for this estate: the module and claims library as a system of record, reuse scoring to find the load-bearing claims, and AI Pickup to measure how much of that approved content the models echo. The defensible part is the join. Because Content Intelligence and Answer Monitor share one claims spine, the same approved claim can be scored for reuse inside, measured for uptake outside, checked for drift and off-label risk, and routed to a compliant-push workflow, all without inventing new messaging. Pre-Check keeps the spine clean by delivering pre-MLR verdicts (Cleared, Review, Blocked) against the label, and Answer Monitor headlines the six Core KPIs, Share of Answer, Ecosystem Share of Answer, Precision of Answer, Risk of Answer, Claim Uptake, and Top References, so the loop is measured, not assumed. The platform is built to support 21 CFR Part 11 controls for that governed estate. Start with the estate you already own, instrument what the machine repeats, and operate the loop. The brands that treat their approved content as the corpus, not the exhaust, will be the ones the machine quotes correctly.

## Sources

1. Veeva, Getting Started with Modular Content.  
<https://www.veeva.com/blog/getting-started-with-modular-content/>
2. American Medical Association, 2 in 3 physicians are using health AI, up 78% from 2023. <https://www.ama-assn.org/practice-management/digital-health/2-3-physicians-are-using-health-ai-78-2023>
3. Adobe, Adobe Analytics: Traffic to U.S. Retail Websites from Generative AI Sources Jumps 1,200 Percent.  
<https://blog.adobe.com/en/publish/2025/03/17/adobe-analytics-traffic-to-us-retail-websites-from-generative-ai-sources-jumps-1200-percent>
4. Aggarwal et al., GEO: Generative Engine Optimization (KDD 2024), arXiv.  
<https://arxiv.org/abs/2311.09735>

5. AI Answer Engine Citation Behavior: An Empirical Analysis of the GEO16 Framework, arXiv. <https://arxiv.org/pdf/2509.10762>
6. HITsz-TMG, A Survey of Large Language Models Attribution (awesome-llm-attributions). <https://github.com/HITsz-TMG/awesome-llm-attributions>
7. U.S. FDA, The Office of Prescription Drug Promotion (OPDP). <https://www.fda.gov/about-fda/cder-offices-and-divisions/office-prescription-drug-promotion-opdp>