

WHITE PAPER

Precision Communication

Machine-scale. Human-signed.

*How commercial pharma organizations will write next —
and how to get there first, responsibly*

Greg Johnsen, lead author · Bradford Lee, contributing editor

Executive summary

Specialty pharma spent two decades building precision medicine — for the molecule — then wrapped its therapies in form letters. This paper describes the capability that closes the gap: **Precision Communication** — adaptive, personalized language for every Patient, prescriber, and payer, at scale, at the moment it matters. The technology has arrived. What's missing is the *permission structure*: personalization generates thousands of communications no reviewer ever sees, and until MLR, privacy, and Legal can say yes to that, the capability stays unapprovable.

We built that permission structure by an unusual route: we assembled the strongest case *against* AI writing from its sharpest critics, tested it against a sample of the discourse it was not built from, then converted the eleven harms they identify into eleven operating disciplines — **Good Writing Practice**, the GxP for the written word (an internal practice standard, not a regulatory designation) — instrumented by a provenance record behind every artifact, a published AI Writing Policy, and human-judgment gates wherever the stakes demand a person. The payoff lands in numbers you already track: activation, abandonment, cycle time.

Machine-scale is coming to pharma communication either way. The only decision is whether it arrives human-signed.

1 Two letters

One day, in early January, a woman we'll call Ana received a letter from the patient support program attached to her specialty therapy. It was accurate. It was approved. It had passed every review it was required to pass. It explained, in the program's standard language, that her benefits had been reverified for the new plan year and that her prior authorization status was being confirmed with her insurer.

Ana read it twice and understood one thing: something about her insurance had changed, and it sounded like a problem. The letter did not say what she owed, whether she could keep taking her medication, or what would happen next — or rather, it said all of those things in the language of the people who wrote it, not the language of the person who needed it. So she did what a reasonable person does with an official letter that alarms without informing. She set it aside. She missed her February fill. In the program's quarterly review, she appeared in a row labeled *abandonment* — *cause unknown*, one unit of variance in a metric everyone agreed was disappointing.

Now let us imagine the letter Ana *could* have received — one that knew, because the program knew, that she had already met her deductible; that her PA renewal was submitted and pending, with a decision expected within ten days; that nothing was required of her; and that said so, in three short paragraphs, at a reading level matched to how people actually read, in the language she speaks at home. A letter written, in a meaningful sense, *for her* — one Patient, one moment, one letter.

Of course, not every Ana is saved by a better letter. A lot can go wrong, and often does. Some abandonment is money — a deductible reset that no paragraph can soften. Some is process — a lapsed PA no letter was

ever going to fix. Comprehension failure is the *avoidable* slice of abandonment, and the hard truth is that nobody knows how big that slice is, because no program has ever been instrumented to isolate it reliably and systematically. This paper includes the instrument that finds out. Expert operators may disagree on the size of that slice, but whatever the size, it's a slice worth chasing, for one simple reason: a Patient who understands bad news can appeal, apply, call, plan. A Patient who is confused does nothing. Clarity doesn't fix affordability — it converts helplessness into action, and it is the portion of the problem entirely within your control.

Every executive in specialty pharma believes in this idea already. It is the founding conviction of the industry's proudest achievement: that one-size-fits-all is primitive, and one-patient-at-a-time is the future. We call it precision medicine, and we spent two decades and billions of dollars building it — for the molecule.

Then we wrapped our precision therapies in form letters.

We spent two decades and billions of dollars building precision medicine — for the molecule. Then we wrapped our precision therapies in form letters.

This paper is about closing that gap. We call the capability **Precision Communication**: adaptive, precise, personalized language — one Patient, one prescriber, one payer at a time, at scale, at the right moment in every journey. The technology to do it has arrived. The method for doing it responsibly, inside the most regulated commercial environment in the world, is what most organizations are missing — and it is buildable. That method is the second half of this paper. The first half is about why you should want it badly enough to build it.

2 The capability, seen whole

Let's start with what a fully instrumented, AI-enabled communications capability actually looks like inside a commercial pharma organization — not the demo, the operation.

Every written artifact the organization ships — benefits letters, PA status updates, welcome kits, field communications, payer dossiers, portal notifications, executive memos — is produced through a governed writing environment. Human professionals direct the work: they set intent, make the governing choices, verify the claims, and sign the result. The machine does what machines now do astonishingly well: it researches, drafts, adapts, structures, translates, edits, and fact-checks at a volume no human team can approach. Every asset carries a named accountable author. Every claim traces to a source. Every communication to a Patient reflects what the program actually knows — her plan design, her deductible status, her PA status, her language, the date that matters — and says what it needs to say in words she can understand.

One engineering truth we should acknowledge here, rather than in the fine print: the letter that “knows” things is only as good as the timeliness, or “freshness,” of its knowledge in the moment. Benefits data goes stale; accumulators lag; PA statuses flip daily. A personalized letter that is confidently wrong about a Patient's own coverage is worse than the boilerplate it replaced — boilerplate never claimed to know her. In this capability, data freshness is a verification obligation, engineered and audited like any other claim, and when a fact can't pass its freshness test the system says the honest general thing instead of the precise wrong thing.

A personalized letter that is confidently wrong about a Patient's own coverage is worse than the boilerplate it replaced — boilerplate never claimed to know her.

The volumes change by an order of magnitude. The trust doesn't have to move at all — and for the first time, the organization can prove how its words are made.

What does that buy, concretely?

Activation. In specialty pharma, the prescription is not the economic event — activation is. Between clinical intent and first fill stands a gauntlet of benefits investigations, prior authorizations, appeals, copay enrollment, and scheduling, and at nearly every step the instrument of success or failure is a piece of writing. The benefits letter a Patient actually understands is the difference between a pending PA and a quiet abandonment booked later as variance. Comprehension is an activation event. Writing quality moves the numbers a CFO already tracks: time-to-therapy, first-fill abandonment, PA cycle time, appeal uptake, persistence at the moments where persistence is actually decided — the first ninety days, the denial moment, the January reverification cliff.

Velocity with your reviewers, not around them. Content operations in pharma are gated by review capacity, and review capacity is gated by trust. A writing operation that arrives at MLR with every claim pre-verified and source-linked, every asset carrying its provenance record, changes the first-pass approval rate and the cycle time — the two metrics every content organization complains about and few know how to move. The stewards stay stewards. They just stop re-doing work the system already proved.

Distinctiveness, at the moment it becomes scarce. Here is the strategic irony of the AI writing era: as baseline fluency becomes free, fluency stops being worth anything. When every company's letters are grammatical, every company's emails polished, the scarce asset is a voice that sounds like *someone* — specific, intentional, recognizably yours. Organizations that simply turn on the tools will converge on the same competent, hollow register: the model's default voice, arriving in everyone's inbox at once. What we want to be is an organization that governs voice as an asset, so we stand out more, not less, as the flood rises.

The other readers. The same architecture serves everyone the organization writes to. The prescriber — drowning in undifferentiated rep email — receives the one message written to her actual panel and practice reality, and reads it. The payer receives a dossier written for the medical director who has to defend the decision internally, instead of a document written for nobody. This paper stays with the Patient, where the outcome data is cleanest and the bar is highest; but the capability is one architecture serving many reader contracts, and precision communication of this caliber is an opportunity for every function across the commercial team.

And the equity dividend, which is also a performance dividend. Today, in most programs, the English version of a Patient letter gets four reviews and the Spanish version gets a vendor invoice. A governed writing operation produces comprehension parity across languages and reading levels as a *designed, budgeted* outcome — parity means staffing the same review rigor for the Spanish letter as the English one, and the practice costs it in rather than wishing it in. The return is that comprehension parity shows up in abandonment data, by language, as a number. This is the rare initiative where the ethical improvement and the

commercial improvement are the same line item.

And beneath all of it, the capability no program has ever had: **attribution**. Today, a Patient Services operation cannot tell a confused Patient from a broke one from a gone one — it books all three as *variance, cause unknown*. An instrumented communication operation closes that loop for the first time: every letter is an experiment, every outcome attributable, every “cause unknown” a question the system can finally answer. Ana’s program didn’t just fail to inform her; it failed to *learn* from her. Precision Communication is the end of cause unknown.

That is the what-if, seen whole. Not better documents for less money — though you will get that too, and it is the least interesting thing you will get. A different operating capability: **machine-scale, human-signed**.

3 Why you don't have it today: the permission structure

If the technology arrived and the prize is that large, why doesn't every commercial organization in pharma already run this way?

Because personalization at scale is, today, **unapprovable** — and every experienced pharma executive knows it in their bones, even if they haven't said it out loud. Your review model approves *artifacts*: a letter, a screen, a script, each one reviewed, each one signed. Personalization produces *instances* — ten thousand letters that each exist for exactly one reader, generated at the moment of need, reviewed by no one. Every “segment of one” pitch of the past decade — and you have heard that pitch, in some form, twice a year for ten years — has died at that gate. The vendors sold the generator. Other industries have started on the governance: newsrooms have published AI-use policies since 2023, and the Authors Guild now registers human-authorship warranties title by title. But no one has built a permission structure that a regulated commercial health organization could actually approve — one that records, for every asset and every personalized instance, who is accountable, what roles AI played, what was verified and by whom, and which version of the policy it was produced under. Nobody built the thing your MLR committee, your privacy officer, and your general counsel could say yes to.

The vendors sold the generator. Nobody built the thing a regulated health organization's general counsel could say yes to.

Meanwhile, the pressure builds from below. Your teams are already writing with AI — more than you know, and mostly outside the purview of Legal and MLR. Disclosure mandates keep producing the same shape: mandates everywhere, compliance near zero — in the largest study to date, seven in ten academic journals had adopted AI policies, and roughly one paper in a thousand disclosed AI use. This isn't defiance; it's what happens when capability arrives before governance. Your most capable people are using the most powerful writing technology ever built, quietly, with no rules that fit, and every unreviewed artifact they ship is spending trust the organization doesn't know it's spending — with prescribers, with payers, with partners, with each other.

And your readers have changed. This is the piece of timing most analyses miss: **suspicion arrived before acceptance**. Readers now assume AI involvement by default — in your letters, your emails, your LinkedIn posts — but they have not stopped thinking worse of it, because nothing tells them what kind of human stands behind the text. The research on news readers is precise about this: more than eight in ten say they want AI use disclosed; shown an article carrying the label, they trust it measurably less, while rating it no less accurate; and the penalty largely disappears when the label arrives with the sources the text was built from. Readers do not punish the machine. They punish the absence of a person. We are living in the gap between those two moments: the assumption is universal, the tolerance hasn't landed. That gap is where trust is being won and lost right now. It is also the window:

organizations that install visible, provable accountability *before* acceptance arrives will define what acceptance looks like — and inherit the trust advantage — the way early leaders on data privacy converted a compliance topic into a brand asset.

So the honest diagnosis is this: the barrier to Precision Communication is not technological, and it is not (as commonly assumed) that “Legal will never allow it.” The barrier is that no one has built, for this industry, the **permission structure** — the operating discipline, the records, the gates, and the proof — that lets the people whose job is to say no say yes. Your reviewers’ caution is not obstruction; it is rational stewardship, and it is currently the only control system you have. The work is to give the stewards something better than caution: instruments.

The rest of this paper describes that permission structure — and the route we took to build it.

4 Taking the critics seriously

This method began in an unusual place: with the critics of AI writing. Before we wrote a single guideline, we went looking for the strongest published case against the thing we intended to build — and made a discovery that shaped everything after it. The case against AI writing had been made many times, but piecemeal: as a craft objection, a trust objection, a labor objection, a regulatory objection, each in its own sector, each answered, if at all, with a label or a ban. Essays and criticism from some of the sharpest writers working, each carrying real insight, many mixing distinct objections together — and none assembled into a single argument that a commercial pharma organization would have to answer, with an answer a compliance function could sign. So we did that work ourselves. We built our opposition's best case for them — and then set out to answer it.

Over the past two years, some of the sharpest writers and thinkers in American letters have mounted a sustained critique of AI-assisted writing — Ted Chiang, in *The New Yorker*, arguing that art is the accumulation of thousands of choices and that prompting delegates nearly all of them; Dan Chiasson, in *The New York Review of Books*, arguing that AI offers shortcuts through the very gaps and impasses where thinking happens; the journalist Charlotte Alter arguing flatly that “any portion done by AI is less thinking done by you.” And in 2026 the discourse got its cautionary tale, chronicled by Katy Waldman in *The New Yorker*: the Commonwealth Short Story Prize, where AI-generated stories won a distinguished literary award through concealment, detection platforms flagged them publicly within days, and an institution's ability to certify its own field collapsed in real time — complete with accused writers who denied AI use and had no way to prove it.

The case we assembled from this literature took the form of a formal framework (Figure 1.0, on the following page) — a taxonomy of eleven distinct harms, disentangled from arguments that often ran several together: from inauthentic authorship to eroded reader trust to the degradation of the shared information commons.

FIGURE 1.0 FOLLOWS

Figure 1.0 — From the critics' harms to the pharma disciplines

The eleven harms as culled and classified from the critical literature, how each translates to commercial pharma writing, the discipline each becomes (detailed in Section 5), and how the harm's weight shifts crossing from the critics' world into this one. The critics named in each row are those whose writing carries that harm in our source corpus; the translation and the weight-shift columns are our own judgment, not the critics', and are argued rather than evidenced. Where a harm's evidence is thin, the row says so.

#	The harm, as the critics frame it	How it translates to commercial pharma writing	The discipline it becomes	Weight shift (literary — pharma)
1	Authorship — the work isn't authentically human-made; prompting delegates the thousands of choices that constitute writing (Chiang; Oremus on the confusion of the AI-writing scandals)	Institutional writing was never solo-authored — ghostwriters settled this generations ago. What remains is ownership: a named human who directed and approved	Accountable Authorship [A]	Collapses — the romantic objection doesn't transfer; accountability replaces it
2	Trust & Transparency — readers deceived about how the text was made (Oremus; Waldman on the Commonwealth Prize attestations)	Every modality carries a reader contract; disclosure is decided deliberately per mode, backed by the standing AI Writing Policy	Fit-for-Context Disclosure [A]	Transforms — from per-work confession to institutional instrument
3	Quality — a flood of competent, hollow, synthetic prose (Horowitch; Oremus on what comes after slop; Kirsch on the didactic default)	When fluency is free, fluency stops being the bar; quality judged blind to provenance, against a raised standard	Blind Quality Standards [A]	Holds — but the bar moves from "well-written" to "true, specific, useful"
4	Development — AI short-circuits the impasses where thinking happens; writers never form (Chiasson; Herman on the end of high-school English; Horowitch)	A talent strategy question: which writing is <i>product</i> (automate) and which is <i>practice</i> (protect)? Junior team members still need to learn to think in writing	Craft Development [O]	Holds — relocated from the writer's soul to the training budget
5	Commercial — the value chain restructured without consent: training data appropriated (Chiasson; Chiang, endorsing Willison's "money laundering for copyrighted data"); platform markets distorted by AI volume, bogus reviews, click farms (Oremus); whole institutions displaced (Horowitch)	The agency and med-comms value chain restructured openly around orchestration and verification, not hollowed by attrition	Value-Chain Stewardship [O]	Holds — different victims (agencies, vendors), same obligation of honesty
6	Meaning & Sincerity — text that means nothing because no one meant it (Chiang; Kirsch). <i>Confirmed in our own reading only; see Appendix A</i>	The central harm: the field email, the Patient letter no one stands behind spends trust and returns nothing. The intent test, per artifact, per variable	Intent Ownership [A]	Intensifies — becomes the keystone, enforced down to the personalization variable
7	Evaluative & Gatekeeping Integrity — prizes and journals can no longer certify what's real (Waldman on the Commonwealth Prize; Herman; Oremus)	MLR and review capacity governed as a ratio against generation volume; the gatekeeper failure mode is scaled unreviewed publishing	Review-Capacity Governance [O]	Intensifies — the gatekeeper is regulated, and the failure is a warning letter, not an embarrassed jury
8	Cultural & Aesthetic — homogenization; everything converges on the model's default voice (Horowitch; Kirsch; Chiasson)	Brand voice managed as an asset precisely when distinctiveness becomes the scarce good	Voice Stewardship [A]	Holds — recast from cultural loss to competitive moat
9	Epistemic Integrity — fabrication, invented references, confident falsehood (Oremus, on a published book built on fabricated quotations). <i>One documented case, reported in two outlets — the pharma translation supplies the force</i>	Claims substantiation extended to AI-drafted content; the model is never the source; freshness as a verification obligation	Verification Discipline [A]	Intensifies sharply — fabrication here is a regulatory matter and a patient-safety event
10	Fairness & Wrongful Accusation — detection regimes accuse on the strength of a flag that is evidence, not proof, and the accused has nothing to exhibit in defense (Waldman; Kirsch). <i>Confirmed in our own reading only; see Appendix A</i>	No AI detectors run on employees' or partners' work, ever; judge the work and its accountability, never its provenance	Judgment Without Detection [O]	Holds — and becomes a hard product and policy constraint
11	Civic & Systemic Trust — the shared information commons degraded by synthetic text at scale (Horowitch, building on Ong, Postman, and Meyrowitz; Chiang). <i>The thinnest harm in the framework: argued by the critics, documented in our sample only once</i>	Publishing health information at scale carries a duty of care beyond compliance: unverified, unmeant, unowned content doesn't ship	Public Trust Stewardship [O]	Intensifies — the commons here includes Patients making health decisions

[A] = artifact-level (assessable per piece of writing) · [O] = organization-level (policy and process). Sources are cited in full in the Notes to Section 4.

SECTION 4 · CONTINUED

We then did what a framework built from hand-picked reading owes its reader: we tested it against reading we had not picked. The eleven categories were written down in advance, along with the result that would falsify them, and then applied blind to a probability sample of *New York Times* coverage drawn from a census of every piece the paper had published on the subject. Every one of the eleven harms appeared in that sample; nine of the eleven were confirmed there to the standard of a documented, evidenced instance, and no twelfth category was needed. Two — the harm of text that means nothing because no one meant it, and the harm of the false accusation — were not confirmed independently, and we say so wherever they appear, because the reason is itself a finding: the first is an interior harm that self-report can only testify to, and the second by construction leaves the harmed person nothing to exhibit. Any governance regime that acts only where harm is well evidenced will under-weight exactly these two. Appendix A gives the coverage matrix and the method; the full instruments and per-source readouts are part of the Good Writing Practice resource. Two findings from the framework changed how we approach everything else.

First: the critics' objections sort into four categories, and three of them can be conceded without damage. The machine may lack meaning in the human sense; it may lack intention behind the sentence; it may tempt writers to bypass thought. All three are claims about the *instrument*, and all three may be true — and none of them settles whether a human being can authentically author *through* the instrument. The only objection that survives contact is the fourth: AI makes it harder for a reader to know what kind of human judgment, labor, and presence stands behind a text. That harm is real. And it is an argument for a framework, not a prohibition — **the risk of contract erosion is precisely**

why a new contract is needed.

Second: the standard that matters is not authenticity. It is accountability. “Did you personally write this?” was never the standard in institutional writing — no one believes the CEO drafts the shareholder letter, and this industry settled the ghostwriting question generations ago. In an AI-mediated world, authorship shifts from the act of writing to the act of judgment: the writer becomes the person who directs the work, recognizes what is right and wrong in it, verifies it, and takes responsibility for it by name. Authenticity is a property of the writer; you cannot govern it. Accountability is a property of the system; you absolutely can. The question every artifact must answer is not “did a machine touch this?” but: **did a named human own the intent, direct the work, verify the claims, and stake their judgment on the output — and were the reader’s expectations, in this mode of writing, honored?**

In an AI-mediated world, authorship shifts from the act of writing to the act of judgment.

That question, operationalized, is the whole method.

5 Good Writing Practice: the eleven disciplines

Pharma already knows what it looks like to do powerful, risky things safely at scale. It writes the discipline down, assigns accountability, keeps records, trains people, and audits the practice — and it names these systems Good Manufacturing Practice, Good Clinical Practice, Good Pharmacovigilance Practice. What we propose is the same move for the written word: **Good Writing Practice** — not a regulatory requirement, not a GxP in the legal sense that binds its namesakes, but an internal practice standard built to the same grammar, because the grammar is the point.

Each of the eleven harms the critics identified becomes a discipline the organization installs. Six operate at the level of the individual artifact; five at the level of the organization.

The artifact disciplines:

- **Accountable Authorship.** Every asset has one named accountable author — the person who directed the work, made the governing choices, and approved the result. “The AI wrote it” is never an explanation.
- **Fit-for-Context Disclosure.** Every mode of writing carries an implicit contract with its reader about what kind of human presence stands behind the text. Disclosure obligations flow from that contract — deliberately decided, consistently applied, honestly defensible. (The formats and the standing public policy are described below.)
- **Blind Quality Standards.** Quality judged against standards that *rise* because baseline fluency is now free. The bar is no longer “is it well

written” but “does it say something true, specific, and useful that a competitor’s model-default boilerplate does not.”

- **Intent Ownership.** The heart of the practice: can the accountable author stand behind this, sentence by sentence, as something they meant to say? A polished message nobody meant is worse than no message — it spends the reader’s trust and returns nothing.
- **Voice Stewardship.** The brand and the team sound like themselves — documented, defended in review, tested by whether a reader could identify the writing as yours with the logo removed.
- **Verification Discipline.** Every claim traces to a source before it enters review, and the model is never the source. In pharma terms: claims substantiation, extended to AI-drafted content, with no exceptions.

The organizational disciplines:

- **Craft Development.** A deliberate answer to the atrophy question: which writing is *product* (automate freely) and which is *practice* (protected, because it is how your people learn to think)? This is a talent strategy decision, not nostalgia for a bygone era.
- **Value-Chain Stewardship.** The agency and writer ecosystem gets restructured openly — renegotiated around orchestration, judgment, and verification — rather than hollowed out by silent attrition.
- **Review-Capacity Governance.** Generation capacity and review capacity are governed as a ratio. An organization that 10x’s content volume against flat review has not scaled writing; it has scaled unreviewed publishing — and the Commonwealth Prize showed how that ends.

An organization that 10x's content volume against flat review has not scaled writing — it has scaled unreviewed publishing.

- **Judgment Without Detection.** The organization judges work and accountability — never provenance. No AI-detection tools run against employees' or partners' writing, ever. A detector's output is the fraction of a text it finds machine-like, not the probability that a person did anything; it sees the artifact and never the process. A flag is evidence, not proof, at any plausible prevalence — the Commonwealth Prize's accused writers learned that the accused has nothing to exhibit in defense — and suspicion regimes destroy precisely the trust this practice exists to protect.
- **Public Trust Stewardship.** What the organization publishes at scale into the health information commons carries a duty of care beyond compliance: if it isn't verified, meant, and owned, it doesn't ship — no matter how cheap it was to generate.

One rule inside these disciplines deserves special notice, because it turns out to be vitally important everywhere. Personalize to what the reader needs to *understand* — their benefit design, their language, their moment — and you are serving them. Personalize to their psychology to move their behavior for revenue, and you have crossed a line that is simultaneously an ethical boundary, a HIPAA marketing boundary, and — should it ever surface in litigation — the single worst sentence your organization could have written down. That's one rule with three jurisdictions, and it compresses to a form worth memorizing: **comprehension-personalization is support;**

persuasion-personalization is demand creation, plain and simple. We treat it as the keystone of the entire practice.

And certain content is gated entirely: anything shaping what a Patient is told about coverage interpretation, affordability determinations, or denials, and anything adjacent to adverse-event reporting, requires a human to *decide* — not review, decide — every time. AI may assist structure and clarity in these modes. It never determines what is said.

6 The instruments

Disciplines without instruments are aspirations. Four instruments make Good Writing Practice operational — and together they are the permission structure that makes Precision Communication approvable.

The Declaration of Authorship and Provenance (DAP). A structured record attached to every asset — the batch record of writing. It captures the named accountable author; how the text was made (a two-part provenance record: which roles AI played, and a composition class — we label them C0 through C4 — from *human-composed* through *AI-generated per instance under governed template*); the author's one-line intent statement; the verification map linking every claim to its source; the review chain; and the policy and consent versions in force when it shipped. The governing design rule: **the record is a byproduct of the work, never an additional task.** The Commonwealth Prize contestants attested twice that no AI was used, and the attestations were worthless — a bare after-the-fact attestation, unstructured and unverifiable, always is. In an instrumented writing environment the DAP is computed rather than declared: it accrues the way git history accrues for software, and the author's single new keystroke is the intent line. Where the environment is not yet instrumented — and today that is most places, our own included — the DAP is attestation with better structure: named, dated, specific about which roles AI played, bound to a policy version, and signed by someone who read the whole thing. Structured attestation beats none; instrumentation beats attestation; the path runs from one to the other, and the record's fields don't change along the way. At personalization scale, the DAP graduates from per-asset to per-instance — you cannot pre-review ten thousand letters that each exist for one reader, so you govern the generator and *audit the*

instances, statistically, the way QA samples calls. Review doesn't disappear. It changes form — and gets stronger.

The AI Writing Policy. The standing public document — treated the way a privacy policy is treated — that tells the world how your organization writes: what AI assistance is allowed at each stage, what is gated for human decision, how disclosure works, how personalization data is used, and what records stand behind it all. Versioned, dated, owned by a named executive, written in plain language over a detailed internal SOP. This is what transforms per-asset disclosure from an awkward per-message confession into institutional transparency: readers who want to know how your writing is made have a real, inspectable answer. The research on disclosure says this is the form readers actually reward: a bare “AI-generated” label costs trust, and a disclosure that arrives with its sources and its method gets that trust back. Newsrooms have been publishing policies of this kind since 2023; what is new here is binding the policy, by version number, to every asset produced under it. And a published policy is an enforceable promise — the privacy-policy enforcement era proved that — which is precisely why it works: **publish only what your records can prove, and instrument everything your policy promises.** Policy without instrumentation is manufactured liability. Policy with instrumentation is armor.

*Policy without instrumentation is manufactured liability.
Policy with instrumentation is armor.*

The Variable Ledger. The data contract behind personalization, one row per variable: what the system may vary on, drawn from which data source, held under which consent, justified by which comprehension purpose. The governing gradient runs from *declared* data (the Patient

told you) through *observed* (seen in authorized operations) to *inferred* (the model guessed) — and consent thins exactly where inference gets clever, so the practice gets less clever: universal plain-language design first, which needs no consent and captures most of the comprehension benefit for every reader; personalization on declared and observed facts second; inference treated as a flag for human follow-up, never an automated switch. Psychological profiles, sentiment targeting, and persuasion susceptibility are presumptively out — see the keystone rule. And the consent artifact itself gets rewritten under this same practice, in plain language, scored for comprehension: a consent nobody understood is a broken reader contract at the very first touchpoint. The Ledger also carries each variable's **freshness rule** — how current the data must be for the system to assert it, and what the letter says when it can't. Stale precision is worse than honest generality; the generator downgrades to the safe general statement whenever a variable fails its freshness test.

The review architecture. Human-judgment gates on the content that requires them; statistical sampling of generated instances; generation-to-review ratios governed as policy. Worth naming plainly: this architecture is also where regulation is heading, on two continents. California now requires health providers' generative-AI communications about a patient's clinical information to carry a prominent AI disclaimer *unless a licensed or certified human provider read and reviewed the communication first*. The European Union's AI Act, in force for these provisions since August 2026, requires AI-generated public-interest text to be labeled *unless it has undergone human review and a person holds editorial responsibility for it*. Neither statute reaches a pharma program's benefits letter today — the California law covers providers and clinical content, not manufacturers and coverage — but both draw the same line: accountable human review is the statutory alternative to a warning

label. The organizations that build this review architecture now aren't just managing risk. They are building the compliance path that lets a personalized letter arrive as a letter.

A note on what these instruments deliberately are not: none of them is an AI detector, and none of them requires one. The entire practice judges writing by its accountability and quality, never by its provenance — because a detector flag is evidence and never proof, because the cost of its errors falls on people who cannot defend themselves, and because in the world that is coming, “did a machine touch this?” will be the least interesting question you can ask about a piece of writing. The interesting question — the only one your readers will ultimately care about — is *who stands behind it*.

7 The path: five stages

Organizations do not leap to Precision Communication. They climb to it, and the climb has a structure we can assess as we step up through its stages:

1. **Ad hoc.** Individuals use AI covertly; no policy; quality and risk invisible. (More of your organization is here than you think.)
2. **Permitted.** Tools sanctioned, training delivered, a policy memo circulated — but no writing-specific governance. This is where most of commercial pharma sits today, and it is the most dangerous stage: sanctioned volume, ungoverned trust.

This is where most of commercial pharma sits today, and it is the most dangerous stage: sanctioned volume, ungoverned trust.

3. **Governed.** Accountable authorship, disclosure policy, and verification installed; the DAP operating as standard practice; the AI Writing Policy running internally as SOP. Risk is controlled — and the foundation is laid.
4. **Orchestrated.** The quality bar raised, intent enforced, review capacity governed, voice defended. Segmented communication ships; governed-generator pilots run on the highest-yield Patient assets; the AI Writing Policy goes public because the records can now back the promise.

5. **Compounding.** Personalization at scale — the reader contract of one, operating as everyday practice — plus the durable disciplines: talent, value chain, fairness, public trust. This is the stage competitors cannot reach by buying tools, because everything that matters here was installed in the three stages before.

Two practical notes on the climb. First, it starts smaller than you expect: an executive's personal writing discipline, then a one-page team writing charter, then the reconciliation of charters into an enterprise policy — we have watched it begin with the question every executive eventually asks in an orientation one-on-one: *"What do I tell my team?"* Second, the pilot that proves the economics is compact: instrument the seven assets where Patients actually fall out — the first ninety days, the denial moment, the January reverification — and let abandonment be the scoreboard. If it moves, the strategic argument is over. If it doesn't, writing wasn't your binding constraint, and you'll have learned that honestly — which is also worth knowing.

8 The window is open

In July 2025, one of us wrote in a dated note that AI authorship would soon be assumed everywhere, and that nobody would think worse of the author, because it was the author who directed, approved, and published the work. A year on, the first half has arrived and the second has not: readers do assume the machine, and they still penalize the disclosure, while “human authored” has become a certification worth paying for. The evidence for both halves appeared earlier in this paper. The gap between them is what this paper aims to close.

Here is the situation, in a nutshell. The technology for one-Patient-at-a-time communication arrived. The permission structure didn't. Your teams are already writing with AI, ungoverned. Your readers already assume it, mostly. Your reviewers are guarding the gate with the blunt instruments available to them — delay, prohibition, restrictive rules — while whole categories of writing, the field email among them, pass no gate at all. And the content volume climbs toward your launch, your indication expansion, your next fight for attention in an industry that runs on trust and words.

Every element of the response is now specifiable: the disciplines, the records, the policy, the gates, the pilot, the path. What it requires is what any new operating capability requires — focus, sequence, and expertise at an intersection (AI-native craft, pharma commercial operations, trust and governance design) that no internal function owns and no generic consultancy occupies.

Machine-scale is coming to pharma communication either way. The only decision any commercial leader gets to make is whether it arrives *human-signed* — under rules you authored, with records you can stand

behind, in a voice that is recognizably yours — or under rules nobody wrote, in the voice of everyone at once.

Machine-scale is coming to pharma communication either way. The only decision any commercial leader gets to make is whether it arrives human-signed.

The first organizations through that gate won't just avoid the harms. They'll own the category of trusted communication in their therapeutic areas — the way the first organizations that took privacy seriously stopped talking about compliance and started talking about competitive advantage.

The window is open now. As Horace, the Roman poet, wrote two thousand years ago: *carpe diem* — seize the day, and put little trust in tomorrow.

— *Odes* 1.11

Genflare advises commercial pharma organizations on AI transformation — including the assessment, program design, and instrumentation of Precision Communication capabilities built on Good Writing Practice. The capabilities assessment maps your organization against the eleven disciplines and five stages; the pilot proves the economics on your own assets in a single quarter.

Notes & Sources

Keyed by section; no superscripts appear in the text. This section is the paper's own verification map, rendered.

Section 3. On disclosure mandates versus disclosure practice: Yongyuan He and Yi Bu, "Academic journals' AI policies fail to curb the surge in AI-assisted academic writing," *Proceedings of the National Academy of Sciences* (2026), doi:10.1073/pnas.2526734123; preprint arXiv:2512.06705 — 5,114 journals and more than 5.2 million papers; about 70% of journals had adopted AI policies; in full-text analysis, 76 of roughly 75,000 papers published since 2023 (~0.1%) explicitly disclosed AI use. — On prior governance work in other sectors: Hannes Cools and Nicholas Diakopoulos, "Towards Guidelines for Guidelines on the Use of Generative AI in Newsrooms," *Generative AI in the Newsroom*, July 10, 2023 (21 published newsroom policies coded); Emily Goligoski, "How to Develop AI Guidelines," *Columbia Journalism Review*, July 16, 2026; The Authors Guild, "Human Authored" certification program, expanded to all authors March 2, 2026, and "AI-Related Model Publishing Contract Clauses," revised May 8, 2026. — On readers and disclosure: Benjamin Toff and Felix M. Simon, "'Or They Could Just Not Use It?': The Dilemma of AI Disclosure for Audience Trust in News," *International Journal of Press/Politics* 30(4), 2025 (working paper 2023; N = 1,483 US adults): 81–84% said news organizations should alert readers that AI was used and, of those, 78% wanted an explanatory note describing how; articles carrying an "AI-generated" label were rated less trustworthy (5.9 to 5.5 on an 11-point scale, $p < 0.001$) but not less accurate or fair; the penalty was largely counteracted when the article disclosed the sources used. Reported by Sarah Scire, *Nieman Lab*, December 5, 2023. The finding concerns US news audiences and is applied here by analogy.

Section 4. The eight sources of the framework's derivation corpus, all cited in Figure 1.0: Ted Chiang, "Why A.I. Isn't Going to Make Art," *The New Yorker*, August 31, 2024 (the "money laundering for copyrighted data" phrase is Simon Willison's, which Chiang endorses). — Dan Chiasson, "Think for Yourself," *The New York Review of Books*, June 25, 2026. — Daniel Herman, "The End of High-School English," *The Atlantic*, December 9, 2022. — Rose Horowitz, "The End of Reading Is Here," *The Atlantic*, July 8, 2026 (August 2026 issue). — Adam Kirsch, "Why AI Is Incurably Didactic," *The Atlantic*, June 15, 2026. — Will Oremus, "AI-Writing Scandals Are Getting Very Confusing," *The Atlantic*, May 23, 2026, and "Is This What Comes After AI Slop?," *The Atlantic*, July 27, 2026. — Katy Waldman, "Did a Chatbot Write a Prize-Winning Story? Does It Matter?," *The New Yorker*, June 10, 2026 (the Commonwealth Short Story Prize account, including the Pangram

detection findings and the contested accusations). — Charlotte Alter, quoted in Katie Parrott, “Writing With AI Is Harder Than You Think,” *Every*, April 6, 2026. — The validation corpus (twelve *New York Times* pieces, 2022–2026) and the method are described in Appendix A.

Section 5. On detector output as a segment fraction rather than a probability of authorship, and on the consequence for the accused: Waldman (above); The Authors Guild, “Human Authored FAQ” (“we are not aware of any truly reliable way to test whether a work includes AI-generated material”); Genflare, *LLM Technical Baseline* v1.1, §8 (Good Writing Practice resource).

Section 6. California Assembly Bill 3030 (Stats. 2024, ch. 848; approved and filed September 28, 2024), codified at Cal. Health & Safety Code §1339.75, effective January 1, 2025: a health facility, clinic, physician’s office, or office of a group practice that uses generative AI to generate written or verbal patient communications pertaining to *patient clinical information* must include a disclaimer that the communication was AI-generated and clear instructions for reaching a human health care provider, employee of the covered entity, or other appropriate person; the requirements do not apply where the communication is read and reviewed by a human licensed or certified health care provider; “patient clinical information” excludes administrative matters such as scheduling and billing. Verified against the codified text, September 22, 2026. — Regulation (EU) 2024/1689 (the AI Act), art. 50(4), second subparagraph, as amended (consolidated version of July 27, 2026; the amending Regulation (EU) 2026/1744 replaced Article 50(7) and did not alter Article 50(4)); Article 50(4) has applied since August 2, 2026: deployers of AI systems that generate or manipulate text published with the purpose of informing the public on matters of public interest must disclose that the text was artificially generated or manipulated, except where the content has undergone a process of human review or editorial control and a natural or legal person holds editorial responsibility for its publication; a separate exception applies to certain law-enforcement uses authorized by law. Verified against the consolidated text, September 22, 2026.

Section 8. The prediction quoted is from a dated personal note by the lead author — “Gen AI Predictions,” Apple Notes, July 3, 2025, 7:07 AM — shared within Genflare that day and not edited since; the note and a verbatim transcription are part of the Good Writing Practice resource. The evidence for the two halves: on the assumption and the disclosure penalty, Toff & Simon (Notes to Section 3); on the standing of “human authored,” the Authors Guild’s Human Authored certification (Notes to Section 3). — Horace, *Odes* 1.11 (c. 23 BC), cited inline.

Appendix A — The evidence base

What was tested. The eleven harm categories in Figure 1.0 were formed as a working hypothesis over months of reading, before any corpus was assembled. They were then tested in two rounds. The first — the *derivation* round — worked through eight long-form pieces of criticism from *The Atlantic*, *The New Yorker*, and *The New York Review of Books*, selected by the lead author, extracting 97 distinct claims and classifying each by category, by how strongly it was evidenced, and by whether it described a harm occurring or a harm forecast. This is where the classification instruments were built, broken, and rebuilt. The second — the *validation* round — tested the finished framework against reading it had not been built from: a probability sample of twelve *New York Times* pieces drawn under a pre-registration that fixed the hypothesis, the falsifier, and four predictions before a single article was captured. The *Times* component is a census: 172,542 articles across 43 months were examined, 401 were eligible, 62 survived a two-stage relevance screen, 47 cleared a 1,000-word floor, and 12 were drawn from those 47 by allocation locked in advance. Extraction and scoring were blind to the hypothesis. The two rounds together cover 20 sources and 341 claims.

What was found. All eleven categories carry at least one anchor-grade claim — evidenced, documented as occurring, technically sound — across the two rounds. Nine of the eleven were confirmed independently in the validation sample. Two were not: **Meaning & Sincerity** (Category 6) and **Fairness & Wrongful Accusation** (Category 10) drew nine claims each in the *Times* sample but no anchor, and rest on the derivation reading alone. We report that rather than absorb it, because the pattern is structural. The first is an interior harm — eight of its nine validation claims are first-person testimony, and testimony cannot reach the top

evidentiary grade by rule. The second is the false positive — the claim that anchors it is correct on the technical record and nonetheless asserted without a study, because the wrongly accused have nothing to exhibit. A governance regime that acts only where harm is well evidenced will systematically under-weight these two categories, and Good Writing Practice makes them a keystone (Intent Ownership) and a hard constraint (Judgment Without Detection) for that reason. A third category, **Civic & Systemic Trust** (11), is the thinnest in the framework: argued in three derivation sources, anchored once in twenty, and that once by an incident from a different generation of systems. Figure 1.0 says so in its row.

What did not go as predicted. Three of the four pre-registered predictions failed. We expected four to seven categories to be empty in a small newspaper sample; none were. We expected at most one candidate for a twelfth category; four surfaced, two of them inside the framework's declared scope, though no two named the same object of harm, so the framework's exhaustiveness test was passed under more pressure than we predicted. We expected news reporting to be better evidenced than long-form criticism; it was not — the share of top-grade evidence was flat across the two corpora. We expected the thin categories to be Development and Epistemic Integrity; they were 6 and 10. We record the misses because a framework that reports only its hits is the kind of writing this paper argues against.

Figure A.1 — Coverage of the eleven categories across both corpora

Anchor-grade claims per category, by corpus. A count is coverage, not weight: a category with many anchors is one the discourse happens to document often — usually because it produces litigation, which produces records — not one that matters more.

#	Category	Derivation (8 sources)	Validation (12 sources)	Provenance
1	Authorship	1	2	both corpora
2	Trust & Transparency	1	5	both corpora
3	Quality of Product	2	3	both corpora
4	Development	1	2	both corpora
5	Commercial	2	16	both corpora
6	Meaning & Sincerity	2	0	derivation only
7	Gatekeeping Integrity	3	2	both corpora
8	Cultural & Aesthetic	1	4	both corpora
9	Epistemic Integrity	1	9	both corpora — one underlying case reported in both
10	Fairness & Wrongful Accusation	2	0	derivation only
11	Civic & Systemic Trust	0	1	validation only

Derivation counts are on each claim's primary category. Category 5's sixteen validation anchors trace almost entirely to litigation records. Category 9's nine validation anchors are seven reports of a single case — a published book built on fabricated quotations — which is also the case behind its one derivation anchor; the pre-registration anticipated this: "a second outlet reporting the same underlying case is not a second case."

What is not in this appendix. The scoring instruments (severity, legitimacy, technical accuracy, modality), the per-source readouts, the full claim inventories, the reliability statistics, and the analysis of how severity and evidence relate across the corpus are part of the Good Writing Practice resource, where they can travel with the caveats they

require. Two of those caveats belong here in brief. Every “independent” scoring pass in this work was a blind re-run of the same AI system, so agreement figures measure replication, not inter-rater reliability in the human sense. And the frame of the whole exercise is US elite anglophone print criticism and reporting, 2022–2026: sampling within that frame was tested exhaustively; the frame itself was not.

How this appendix was made. The reading was the lead author’s; the selection of the derivation sources was his; the extraction, classification, and scoring were performed by an AI system under written instruments and blind protocols he directed; the verification of every figure here against the underlying claim tables was performed on September 5–6, 2026. The verbatim source archives, not the classification tables, are the authority for any quotation.

How this paper was made

Consistent with the practice it describes, this paper ships with its own declaration.

LEAD AND ACCOUNTABLE AUTHOR	Greg Johnsen, Genflare. I use AI for research, drafting, and revision. Whatever the degree of assistance, everything I send is directed, reviewed, and approved by me — it says what I mean to say, and I am its author. I directed this work, made its governing choices, verified its claims, and stand behind every sentence. Approved September 24, 2026.
CONTRIBUTING EDITOR	Bradford Lee, Genflare. This paper does not publish without his review and approval. Approved September 24, 2026.
COMPOSITION	AI-drafted, human-directed and human-transformed (classes C2–C3 under the taxonomy in Section 6), across an extended working dialog. AI sat in all six seats of our practice vocabulary: research, sparring, outlining, drafting, red-teaming, and fact-checking. The framework and positions are ours and Genflare’s; the July 2025 prediction cited in Section 8 is the lead author’s alone; the drafting was collaborative; the judgment and approval are human and named. The evidence base in Appendix A was extracted and scored by an AI system under written instruments the lead author directed, and its figures were verified against the underlying tables before publication.
VERIFICATION	Every checkable claim in this paper traces to a published source or a named Genflare record; the map is the Notes & Sources section above. The model is never the source. Where a claim rests on a study of one field and is applied to another by analogy, the note says so.
REVIEW	Bradford Lee, contributing editor (September 24, 2026), and Greg Johnsen, lead and accountable author (September 24, 2026). No additional reviewers.

RECORDS	This paper was not produced in an instrumented writing environment. This declaration is therefore attestation with structure, as Section 6 describes — named, dated, specific about roles, and bound to a policy version — not a computed record.
POLICY	Written under Genflare’s AI Writing Policy v1.0, effective September 22, 2026. This declaration is the reader-tier rendering of the paper’s provenance record.

© 2026 Genflare. Precision Communication and Good Writing Practice are Genflare program names describing internal practice standards; they are not regulatory designations and imply no regulatory status.



genflare.ai/precision-communication