

The Slop Audit

Paste this file into an agent chat and point it at your work.

It checks writing and screens for the things that make work look machine-made. Claims with no proof. Numbers with no source. Crowded layouts. Screens with no error or empty state. You read the rules, find what fails, and fix it.

The rules come from four places. Wikipedia's style guide, which has flagged puffery and weasel words since 2010, and the cleanup project that logs how generated articles read. Corpus studies comparing human and machine text, on sentence variability and on vocabulary shifts across 15 million abstracts. WCAG 2.2 for contrast, zoom, and target size. My own work, which is marked as such.

Inside: how to run a pass and rate what you find, 22 rules for copy, 19 for screens, 16 industry overlays, wording for the six screens most builds skip, search lists, a block to paste when you build, a one page card, and the sources.

CONTENTS

The rule set

Protocol	How to run the audit, how much each signal is worth, how to rate what you find, and what a fixer is not allowed to change.	3
The four gates	Four fast checks every sentence and every screen clears before anything ships.	8
Baseline	What the floor covers everywhere, and what your organisation sets for itself.	9
C-01 to C-22	Copy rules. What the words claim, who is claiming it, and whether the proof is there.	10
V-01 to V-19	Screen rules. Spacing, alignment, type, density, color, components, states, and access.	16
D-01 to D-16	Domain overlays. What tends to go wrong in one field, added on top of the copy and interface rules. Not universal.	23
P-01 to P-06	Patterns. Six screens most builds skip, written out so you can copy them.	32
A-01 to A-10	Prompting. What to settle before an agent generates anything: context, evidence, what may change, what must not.	39
Appendix	Search lists, the text to paste when invoking an audit, build constraints, and a one page card.	49
Sources	What the rules rest on, what changed in version 3, and the references.	55

Identifiers are stable, the contents are not

Use the numbers in review comments, commits, and tickets. C-09 will mean the same thing next quarter.

The word lists inside the rules will go stale. Wikipedia editors who track AI writing sort their notes by model, because the words change when the models change. Put a date on yours.

PROTOCOL

Running the audit

Follow these steps in order. Answer in the format on the next page.

STEP 1 ESTABLISH WHAT YOU ARE LOOKING AT

Settle these five first. If you cannot tell from what you were given, ask. Many rules give a different answer depending on them.

Artifact	Screen, landing page, document, deck, email, README, or case study.
Audience	The role and the task. An operator who lives in the tool is not a visitor.
Domain	The closest overlay in D-01 to D-16, or none.
Stage	Proposed, prototype, tested, built, shipped, or stopped. This limits every claim.
Evidence	What can be verified: analytics, tickets, research, logs, or nothing.

STEP 2 RUN THE PASSES IN THIS ORDER

- ❑ The four gates. A gate failure is a finding regardless of which rule it maps to.
- ❑ Copy rules C-01 to C-22 over every string, heading, caption, and paragraph.
- ❑ Interface rules V-01 to V-19 over layout, components, states, and styles.
- ❑ The domain overlay if one applies. Overlay rules win where they conflict.
- ❑ Patterns P-01 to P-06 to check the states exist, not only that they read well.
- ❑ The appendix search lists, run literally, to catch what reading missed.

STEP 3 CLASSIFY EVERY FINDING

BLOCKER	MAJOR	MINOR
A claim beyond the evidence, an invented fact, a missing required state, an accessibility failure, or an unguarded destructive action. Ships nothing.	A break a user will notice: abstraction instead of substance, no actor, unreadable density, broken alignment, an undefined metric.	Local wording, one stray spacing value, a weak caption. Fix in a batch.

PROTOCOL

Answering

One finding per problem, in this shape. No opening about how careful the review was, no encouragement at the end.

Rule	The identifier: C-09, V-14, D-03, P-02.
Location	File, line, screen, component, or section heading. Enough to find it in one step.
Evidence	The exact string, or the element and its property values.
Why	One sentence naming the failure, not restating the rule.
Fix	The new wording, or the change to make. If you need a fact you were not given, name it.
Severity	Blocker, Major, or Minor.

WORKED FINDING

FINDING	C-10, Major. Marketing page, hero section. "Teams save 40% of their time." No baseline, no population, no measurement window, and no source in the repository.
FIX	Either cite the measurement, or write what is actually known: "In the pilot, the weekly reconciliation went from two sittings to one." If neither is available, cut the claim. Do not soften it to "save significant time," which is the same claim with the evidence hidden.

SUMMARY LINE

Start with one line: what you looked at, how many blockers, majors, and minors, and the one change worth making first. Then the findings, blockers at the top. A blocker buried under forty small notes will be missed.

PROTOCOL

How strong is the signal

Some rules rest on published standards. Some rest on habit. Sort what you find into these three groups before you write it up.

STRONG MEASURED OR PUBLISHED

- ☐ Contrast ratios, target sizes, zoom behavior, and keyboard access. These come from published specifications (W3C, 2023; TetraLogical, 2022), so they are pass or fail, not taste.
- ☐ Claims that outrun their evidence, and metrics with no baseline. The oldest rule here, predating the technology: Wikipedia's puffery and weasel-word guidance has said the same since 2010.
- ☐ Flat variance across a document. Several corpus studies converge on spread as the difference that survives model upgrades (Zanotto & Aroyehun, 2024).
- ☐ Missing states. Either the empty, error, and permission screens exist in the code or they do not.

WEAK COUNTS ONLY ALONGSIDE OTHERS

- ☐ The em dash. A 2026 preprint measured GPT-4.1 at about 10.6 em dashes per thousand words and two Llama models at zero, against Mark Twain at 10.1 in Huckleberry Finn (Freeburg, as cited in HumanizeAI, 2026). Twain is not a chatbot. Never open a finding with punctuation.
- ☐ Any single vocabulary hit. Wikipedia's own guidance says a sign like the em dash appears in human writing too, and only becomes telling when it co-occurs with several others (Wikipedia contributors, 2026a).
- ☐ Rule-of-three lists, boldface, and headers in the form of X and Y. Common in generated text, also common in readmes and sales decks written by people.

DEAD DO NOT REPORT

- ☐ Authorship. This document audits work, not writers. Never conclude or state that something was written by AI. Report which rules failed and let the author decide what to do about it.
- ☐ Detector scores as evidence. Detection is contested and gets less reliable as models improve (Kujur, 2025), two studies of the same platform in the same month can differ by a factor of two (Originality.AI, 2026b; Pangram Labs, as cited in Vynce Digital, 2026), and false positives fall hardest on people writing in a second language.
- ☐ A person's writing habits. Someone who has used semicolons for thirty years is not producing slop.

PROTOCOL

What a fixer may not do

These matter more than the rules. A rewrite can pass every check and stop being true.

- ☐ Never invent a fact to satisfy a rule. No numbers, names, quotes, dates, sources, or user counts that were not supplied.
- ☐ Never replace a vague claim with a specific one you cannot verify. Say which fact is missing and stop there.
- ☐ Never change domain terminology to something friendlier. The ugly internal word is usually the correct one.
- ☐ Never rewrite legal, regulatory, clinical, or safety strings. Flag them and route them to the person who owns them.
- ☐ Never delete a caveat, a limitation, or an uncertainty because it reads as weak.
- ☐ Never swap one template for another. Cutting a pattern and installing your own across every section leaves the same problem.
- ☐ Never restructure files, rename components, or change behavior when the request was an audit.
- ☐ Never mark a rule as passed when you could not check it. Say you could not check it.

When to stop and ask

- ☐ The claim depends on evidence only the author has.
- ☐ The density target is unknown and the layout verdict depends on it.
- ☐ The voice appears to break a rule deliberately and consistently, which is a style decision rather than a defect.
- ☐ The domain is regulated and the copy carries obligations.

PROTOCOL

Three ways to run it

The rules stay the same. What changes is when you run them and what you may touch.

AUDIT

Read only. Find problems, change nothing. Use it on work you did not write and on anything about to ship. This is the default.

FIX

Fix what you can without new facts. Turn the rest into questions. Show before and after for every line you change, so the author can reject a rewrite that drifted.

BUILD

Use the rules while you build. Write the prompt with the A rules, take the constraints block in the appendix as part of the instructions, then check your own output before showing it. This is the cheapest of the three.

Scope control

Audit one screen at a time. A report on forty screens gets ignored. Start with the screen that carries the most traffic or the most risk, fix it, and use it as the model for the rest.

GATES

The four gates

Four questions that catch most of what the rules catch, faster. Failing one is a finding on its own.

GATE 1 SAY IT ALOUD

Could someone say this sentence to a colleague while pointing at the screen?

If it would sound odd said out loud, it was written for an audience that is not there. Keynote voice and launch post voice both fail here.

GATE 2 LAND ANYWHERE

If a reader lands on this exact section with no context, do they know where they are?

Who it is for, what it does, and whether it is proposed, built, or shipped. Nobody should scroll up to work out what a pronoun points at.

GATE 3 NAME THE REASON

Can someone name the reason for every visual decision on this screen?

Each spacing value, weight, color, and radius came from a system or came from nowhere. If the answer is that it looked fine, it was a default, not a choice.

GATE 4 SIZE THE CLAIM

Does the confidence of this claim match what can actually be proven?

A prototype is not adoption. Approval is not validation. A percentage with no baseline is decoration. Claims that match the proof read as more credible, not less.

Fail one and you rebuild the line or the screen. Do not smooth it over.

BASELINE

The floor, and what is yours to set

Everything in this document is a minimum. It says what any product owes the person using it, in any industry. It does not say how your organisation should implement that.

THE FLOOR APPLIES EVERYWHERE, NO EXCEPTIONS BY INDUSTRY

- ☐ Claims sized to the evidence behind them, including claims about AI.
- ☐ Accessibility: contrast, target size, keyboard paths, focus, and zoom.
- ☐ Privacy and security behavior stated in plain words where a person can act on it.
- ☐ Safety, clinical, legal, and regulatory language left as the owner wrote it.
- ☐ Interface copy that names the action, the rule, and what happens next.
- ☐ The states a real system reaches: empty, loading, partial, error, denied, stale.
- ☐ Numbers with a baseline, a window, and a source, or no numbers.

YOURS TO SET VARIES BY COMPANY, EVEN INSIDE ONE INDUSTRY

- ☐ Policy and how your legal team reads a regulation.
- ☐ Approval structure, review gates, and who signs off.
- ☐ Terminology, naming, and the words your customers already use.
- ☐ Workflow design, roles, and how work is queued.
- ☐ Design system values: scale, radius, palette, density targets, motion.
- ☐ Technical architecture and the limits that come with it.

When they disagree

Your organisation decides implementation. The floor still holds. A company rule can say which words to use for a state; it cannot remove the state. Two companies in the same industry will not work the same way, and this document does not assume they do.

Overlays are overlays

D-01 to D-16 describe what tends to go wrong in a field. They add to the copy and interface rules and never replace them, and they are not a description of how your team operates. Read the one that matches the work in front of you and ignore the other fifteen.

RULES C

Copy

Twenty-two rules for words. They cover interface labels, errors, marketing pages, docs, decks, and reports, which all go wrong in the same ways.

Each rule says what to look for and what to do. The last one says when to stop cutting.

A CROSS-INDUSTRY CHECK FOR POLISHED WORK WITH NO EVIDENCE, CONTEXT, OR STATES BEHIND IT

Gates / **C copy** / V interface / D overlays / P patterns / A prompting / Appendix

C-01 Manufactured contrast

A sentence built on a reversal that carries no information. It sounds like a point because it is shaped like one.

FLAGS not X but Y / this wasn't X, it was Y / more than just / rather than / instead of / X is not Y. It is Z.

FIX State the decision without the foil. If the sentence survives with the contrast deleted, delete it. Wikipedia's Manual of Style flagged but, despite, however and although as false connectives in 2010, long before any of this (Wikipedia contributors, 2010).

C-02 Rhetorical symmetry

Balanced clauses, mirrored structures, and groups of three. Models write this way. People rarely do.

FLAGS Three verbs, nouns, adjectives, or benefits in a row. Sentence fragments in sequence. Every paragraph the same length.

FIX Break the pattern. Let one item stand alone, let one sentence run long, and let one paragraph be two lines.

C-03 Flat variance

The clearest difference between machine and human writing is spread. People write unevenly. Generated text stays level.

FLAGS Sentence lengths clustered near the mean. Paragraphs of near-equal size. Constant enthusiasm with no dips. Repeated openings.

FIX Measure it rather than sensing it. Zanutto and Aroyehun (2024) profiled 250 linguistic features across four domains and found human writing considerably more variable, with more emotional content, and the gap widest in loosely constrained genres such as Wikipedia and Reddit. Worth knowing that they also found humans with a lower type-token ratio than the models, which cuts against other studies, so treat vocabulary richness as contested and variance as the solid part. Fix by rewriting passages, not by lengthening one sentence.

C-04 Fake drama

Reveals, pivots, and stakes added to work that did not have them. Usually a sign that there was little to report.

FLAGS The challenge? / The result? / the real problem / the deeper issue / That's when / Enter X / every click mattered

FIX Say what happened in order. A project that stopped can end with why it stopped, with no lesson attached.

C-05 Abstraction instead of change

A quality standing in for the thing that actually changed. This is what gets written when nobody checked.

FLAGS improved clarity / reduced friction / built trust / enhanced efficiency / empowered users / better experience

FIX Name what became easier to see, compare, or act on, or which step, delay, or repeated action is gone.

C-06 Prestige vocabulary

Words whose job is to sound senior. Real domain terms are fine.

FLAGS leverage, utilize, streamline, optimize, reimagine, elevate, empower, seamless, holistic, robust, intuitive, frictionless, ecosystem, journey, touchpoint, pain point, north star, unlock

FIX Use the plain verb. Keep technical vocabulary only where it changes what the reader understands.

C-07 Word inflation

Every observation called an insight, every change a decision, every result impact. Small work starts lying one word at a time.

FLAGS insight, impact, strategic, constraint, alignment, feasibility, rapid iteration, transformation

FIX Use the smallest accurate word. Noticed, changed, asked, removed, tried, and stopped are usually the true ones.

C-08 Missing actor

Passive construction hiding who did the work, including the work done by other people.

FLAGS was created / was designed / the product evolved / the system became / it was decided

FIX Name who did it. Use I for owned work and we only for genuinely shared decisions. Name the partner whose action changed a decision.

C-09 Process theater

Naming a method, a function, or a ceremony as though naming it were evidence that something happened.

FLAGS collaborated closely with / aligned stakeholders / partnered with engineering / cross functional / conducted user research

FIX Say what was exchanged: a constraint raised, a requirement changed, a state that was expensive, a ticket that clarified behavior.

C-10 Claim beyond evidence

The worst failure in this document. A word swapped for a bigger one the evidence does not support.

FLAGS Outcomes attributed to unshipped work. Approval called validation. Testing called success. Handoff called launch. Figma called production ready.

FIX State the stage exactly: proposed, tested, built, shipped, adopted, or stopped. If the stage is unknown, ask before writing.

C-11 Metrics without provenance

A number with nothing behind it: no baseline, no window, no source.

FLAGS Round percentages. 40%, 3x, 10,000 users, 99.9% uptime. Any figure with no citation in the repository or research.

FIX Attach the baseline, the population, and the window, or remove the number. Directional truth beats false precision. A number worth keeping looks like Kobak et al. (2025): a stated corpus, a pre-period comparison, and a published word list, arriving at at least 13.5% of 2024 abstracts processed with an LLM.

C-12 Context debt

A section that only works for someone who read the one above it. Every section is someone's first.

FLAGS Unresolved it, this, they, the system. Headings like The solution, What happened next, Making it work.

FIX Name the subject locally. Headings name the task, workflow, user, or artifact. Label the artifact's stage.

C-13 Filler transitions

Correct sentences that tire the reader by making the same move again and again.

FLAGS This informed / This enabled / This revealed / By doing X / Recognizing this / With this in mind / As the project evolved

FIX Start sentences with the subject. Cut the transition and check whether anything was lost. Usually nothing was.

C-14 Domain gravity

An opening that borrows weight from the industry instead of starting with the work.

FLAGS Healthcare decisions can be overwhelming / In banking, trust is everything / AI is transforming how we work / In an increasingly connected world

FIX Open on the specific task, screen, or decision. The stakes are already in the domain and do not need announcing.

C-15 Self praise

The work praising itself instead of showing the decision. Captions and headings do this most.

FLAGS thoughtfully designed / carefully considered / intuitive / best in class / This demonstrates my ability to / end to end ownership / wore many hats

FIX Delete the assessment and leave the decision. Let the artifact carry the competency.

C-16 Interface strings

Buttons, labels, and help text written as filler instead of as answers to what the person is doing.

FLAGS Submit / OK / Click here / Learn more / Get started / labels that repeat the database column name

FIX Buttons name the action and its object. Labels use the user's word. Help text explains a rule or a consequence or is deleted.

C-17 System messages

Errors, confirmations, and empty states written to protect the product instead of helping the person.

FLAGS Oops / Something went wrong / Are you sure? / We apologize for any inconvenience / No data to display

FIX Say what happened, whether the work was lost, and what to do next. See patterns P-01 to P-03 for the full shape.

C-18 Voice drift

One product using different words for the same thing in the app, the docs, and the email.

FLAGS Two names for one object. Marketing adjectives inside the product. Jokes in system messages. Formal register in one surface and casual in another.

FIX Keep one glossary of product nouns. Register can shift by surface, vocabulary cannot. Have one person read all strings before release.

C-19 Plain language

Sentences that only work for a rested reader on a big screen.

FLAGS Long sentences with stacked clauses. Idioms and sports metaphors. Unexpanded acronyms. Strings assembled from fragments.

FIX One idea per sentence, common words, acronyms expanded on first use in every section, and room for translated text to grow.

C-20 Units, dates, and identifiers

Formats that change between screens, so the reader has to check what they are looking at.

FLAGS Mixed date formats. Currency without the code in a multi currency product. Timestamps without a zone. Inconsistent truncation of account or order numbers.

FIX Pick one format per type and apply it everywhere, including in emails, exports, and error messages.

C-21 Final sentence scan

Phrases that survive every other pass because each one looks fine on its own.

FLAGS at its core, ultimately, notably, importantly, what emerged, the key insight, the missing piece, at the intersection of, bridging the gap, turning X into Y

FIX Run the search list in Appendix B literally. Delete on sight unless the phrase carries a fact.

C-22 Do not scrub the work to nothing

The counter-rule, and the one most likely to be ignored. Wikipedia removes editorializing because it has a neutral point of view policy. A landing page, a pitch, or an essay does not have one, and having a position is the reason anyone reads it.

FLAGS Prose that passes every rule in this document and says nothing. All claims hedged into vagueness. Every judgement removed. A voice indistinguishable from any other cleaned draft.

FIX Cut what is unsupported, keep what is argued. A position stated plainly with the reasoning attached is not slop. If a full pass leaves text that could belong to any company, the pass went too far and the draft needs an argument, not another edit.

RULES V

Interface

Nineteen rules for how a screen looks and behaves. Generated layouts fail the same way every time: crowded, centered, evenly weighted, full of values nobody picked.

Check the screen and the stylesheet separately. The screen hides loose values, the code hides crowding.

A CROSS-INDUSTRY CHECK FOR POLISHED WORK WITH NO EVIDENCE, CONTEXT, OR STATES BEHIND IT

[Gates](#) / [C copy](#) / [V interface](#) / [D overlays](#) / [P patterns](#) / [A prompting](#) / [Appendix](#)

V-01 Spacing scale

Every gap comes from one scale. Without that, every screen needs hand tuning forever.

FLAGS Values outside the scale in the stylesheet. 13px, 27px, 5px. Inline style overrides on spacing.

FIX One base unit and its multiples. 4, 8, 12, 16, 24, 32, 48, 64. The 8dp convention comes from Material Design, and any consistent scale beats an inconsistent one. Fix at the source, not on the element.

V-02 Grouping and proximity

Leave more space between groups than inside them. When the gaps match, the grouping disappears and the reader takes the screen one item at a time.

FLAGS Uniform gaps everywhere. A label as far from its field as from the field above. Cards separated by the same gap as the content inside them.

FIX Set internal spacing one or two steps below group spacing. Draw the groups before touching numbers.

V-03 Alignment and edges

Nothing makes a build look careless faster, and nothing is cheaper to fix.

FLAGS More than four vertical start positions on a screen. Text aligned to a neighboring component's padding. Centered body copy. Centered numeric columns.

FIX Reduce shared edges. Align text to text, numbers right or on the decimal, and nudge icons optically.

V-04 Grid and responsive

A layout that only works at the width you built it on is not a layout yet.

FLAGS No max width on reading content. Breakpoints named after devices. Stacking order that reverses meaning. Sticky elements that stack and eat the viewport.

FIX Keep content in a grid with a reading max width. Set breakpoints where the content breaks. Test narrowest first.

V-05 Type scale and hierarchy

Three to five sizes cover almost any product. More than that and the hierarchy stops meaning anything.

FLAGS One off font sizes inside components. Adjacent levels within a point or two. Body measure past 75 characters. Justified text. All caps paragraphs.

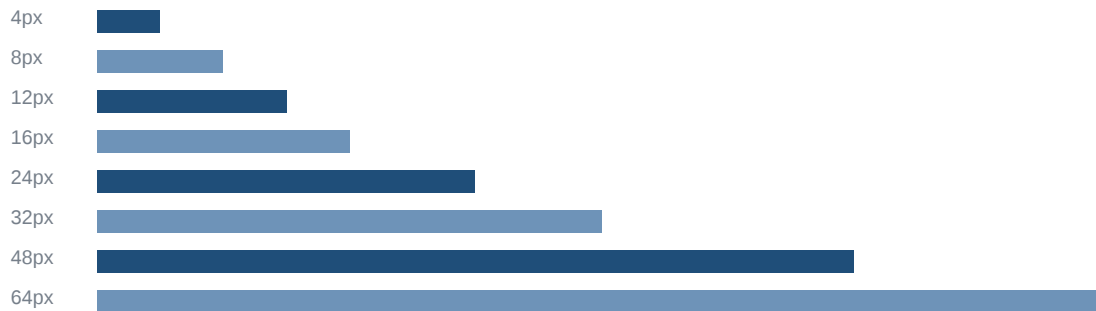
FIX Define the steps and use only those. Let weight carry hierarchy in dense screens. Line height falls as size rises. The 45 to 75 character measure is typographic convention rather than a tested threshold, so treat it as a starting range (Bringinghurst, 2004).

V-06 Color roles

Color either means something or it is noise. Generated palettes use one hero color everywhere and give it no job.

FLAGS Status colors reused decoratively. Meaning carried by hue alone. Multiple accents competing. Gradient backgrounds. Inverted dark mode with pure black and pure white.

FIX Define surface, border, primary text, secondary text, accent, and the status set. One accent per screen. Pair color with a label or icon.



A working scale for V-01. Eight values, no midpoints invented during the build.

V-07 Density decision

Consumer screens and work screens need opposite answers. Pick one before you set row height.

FLAGS A comfortable consumer layout in a tool used for six hours a day. Fewer than twenty rows on an operator screen. Padding added to fill an empty page.

FIX Name the target: compact for daily operator work, comfortable for occasional use. Compact reduces padding, not the spacing scale.

V-08 Component discipline

Repeating yourself is the point. A new variant should be a decision, not something a prompt added.

FLAGS Mixed border radii. Shadows on every surface. Two primary buttons in one view. A fourth button style. Three components doing the same job.

FIX One control per job, one radius scale, two or three elevation levels, one primary action per screen.

V-09 Emphasis budget

If three things stand out, nothing does. You only get so much emphasis per screen.

FLAGS Bold, color, size, and a badge on the same element. Every card with a call to action. Alerts used for non urgent information.

FIX Pick the one element the screen exists for and reduce the weight of everything competing with it.

V-10 Forms and input

Forms take up most of the day in enterprise products, and they are the thinnest part of a generated build.

FLAGS Uniform field widths. Validation on keystroke. Placeholder used as label. Data lost on failure. Autofill blocked. Required marking inconsistent.

FIX Width signals expected length, validation runs on blur, labels stay visible, and entered data survives every failure.

V-11 Navigation and wayfinding

People should be able to say where they are, how they got there, and how to get back.

FLAGS No current location marker. Navigation labels differing from page titles. State not in the URL. Primary navigation hidden behind an icon on desktop.

FIX Mark the current location, match labels to destinations, keep the back button working, and make search real in any product with depth.

V-12 Tables and data grids

In enterprise work the table is the product. Treat it as a surface, not a component that was dropped in.

FLAGS Horizontal scroll that loses the identifying column. Mixed date formats. Sort state hidden in a panel. Row actions that move between row types. Counts that ignore filters.

FIX Freeze the identifier, align numbers right, show sort and filter state in the header, keep actions in one place, and state matched and shown counts.

V-13 State coverage

The happy path is the easy part. Missing states are where a build goes quiet and people stop trusting it.

FLAGS No empty, loading, partial, error, permission denied, or stale state in the code. A spinner in place of a skeleton. Incomplete data presented as complete.

FIX Design all six. See P-01 to P-06. A missing required state is a blocker, not a polish item.

V-14 Charts and numbers

This is where a good looking build turns into a wrong one.

FLAGS Truncated bar chart axes. Pie charts with more than three slices. No units, window, or source. Precision beyond what the data supports. Undefined metrics.

FIX Pick the chart from the question, start bars at zero, label units and window near the chart, and make every metric definition reachable in one interaction.

V-15 Motion

Animation should show where something came from or went. The rest costs time on every use.

FLAGS Any transition you notice twice. Staggered lists on every render. Animated layout size on content being read. No reduced motion handling. Motion covering a slow response.

FIX Animate position and opacity, respect the reduced motion setting, and show real progress instead of decoration.

V-16 Icons and imagery

Decoration is the first thing a build adds and the first thing a user ignores.

FLAGS Mixed icon sets or weights. Emoji as interface icons. Faceless people illustrations. Stock team photography. Icons that repeat their own label.

FIX One set, one weight, labels next to ambiguous icons, and no image that carries no information.

V-17 Access and input targets

The most common defect in generated interfaces, and the easiest to check.

FLAGS Text under 4.5 to 1 contrast. Gray placeholder used as content. Pointer targets under 24 by 24 CSS pixels with no spacing exception, which is the Level AA floor in WCAG 2.2 (W3C, 2023, SC 2.5.8); the widely quoted 44 by 44 is the AAA criterion and matches Apple's platform guidance, with Material asking for 48 dp (TetraLogical, 2022). Focus outlines removed. Keyboard traps. Clipping at 200 percent zoom (SC 1.4.4).

FIX Verify contrast against real backgrounds including hover and dark mode, tab the whole flow, and zoom before shipping. These are published requirements, not preferences, which is why they sit at blocker severity.

V-18 Destructive and irreversible actions

The one place where friction is right. A build that treats delete like save will eventually delete something that mattered.

FLAGS Destructive actions styled like routine ones. Confirmation dialogs on reversible actions. No undo. Bulk actions with no count.

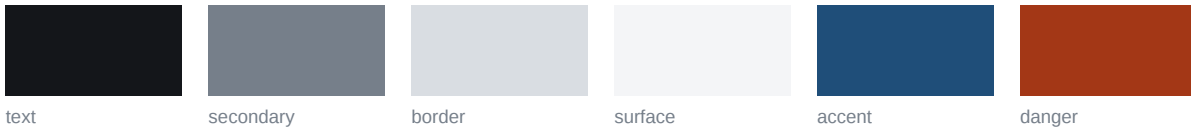
FIX Confirm with the object named and the consequence stated, require typing for irreversible actions, and offer undo where the system can honor it.

V-19 **Generated default signature**

None of these is wrong alone, and the same caution Wikipedia applies to writing tells applies here: a single sign means little, several together mean a lot (Wikipedia contributors, 2026b).

FLAGS Violet to indigo hero gradient. Three equal icon cards. Uniform large radius with shadows everywhere. Centered everything. Pill badges with no state. Unsourced statistics. Glass panels over blur. Heading, subheading, three items, button, repeated.

FIX Rebuild with the layout protocol: one job, real groups, the scale, one primary, aligned edges, and delete the filler.



Roles for V-06. Six values, each with a job, before any of them is called a brand color.

Domains

Sixteen overlays. Everything in C and V still applies. These add what goes wrong in a particular field, and how dense the screen has to be before people can work in it.

If an overlay disagrees with a general rule, the overlay wins. Cite both numbers.

A CROSS-INDUSTRY CHECK FOR POLISHED WORK WITH NO EVIDENCE, CONTEXT, OR STATES BEHIND IT

[Gates](#) / [C copy](#) / [V interface](#) / [D overlays](#) / [P patterns](#) / [A prompting](#) / [Appendix](#)

D-01

Healthcare and clinical

An ordering physician scanning a worklist between appointments, a medical assistant reconciling a requisition, a patient opening a result before anyone has called.

DEAD ON ARRIVAL “Healthcare decisions can be overwhelming.” / “Patients deserve better.” / “The care journey.”

COPY

- ☐ Name the role. Physician, assistant, and patient are not one user.
- ☐ Use chart vocabulary: specimen, requisition, panel, order, result status.
- ☐ Never imply a diagnosis the product is not cleared to make.
- ☐ State when a result becomes visible and to whom, and who to call.

INTERFACE

- ☐ Status and collection date sit next to the value, not in a drawer.
- ☐ Reference range shown with the result. Abnormal flags are never color alone.
- ☐ Never truncate a medication name, dose, or unit.
- ☐ The printed and faxed view is a real surface.

DENSITY Compact worklists, fitting as many rows as stay legible. Comfortable patient views, one action per screen, and no critical value below the fold.

D-02

Banking, payments, fintech

A customer moving rent money on a Sunday night, an analyst working a fraud queue, an operator reversing a batch that posted wrong.

DEAD ON ARRIVAL “In banking, trust is everything.” / “Money, simplified.” / “Banking built for you.”

COPY

- ☐ Two dates: when money leaves and when it can be used.
- ☐ Pending, posted, and available mean what the ledger says they mean.
- ☐ Fees and rates appear before confirm, in the charged amount.
- ☐ Error copy states whether money moved. That is the only question.

INTERFACE

- ☐ Amounts right aligned with tabular figures.
- ☐ One format for account identity everywhere.
- ☐ Confirmation shows source, destination, amount, fee, and date together.
- ☐ Credit and debit never signaled by color alone.

DENSITY Compact for operations and fraud queues with keyboard paths. Comfortable retail, one money movement per screen.

D-03

Insurance, claims, billing

A reviewer working a document queue against a clock, a specialist appealing a denial, a patient decoding an explanation of benefits.

DEAD ON ARRIVAL "Claims processing is broken." / "Billing doesn't have to be painful." / "Chaos into clarity."

COPY

- ☐ Name the document: remittance, denial letter, prior authorization.
- ☐ Quote denial reasons with the payer code and wording.
- ☐ Give the deadline and the next action with dates.
- ☐ Never state coverage the system has not confirmed.

INTERFACE

- ☐ Source document and extracted fields side by side.
- ☐ Every extracted value links to its region on the page.
- ☐ Low confidence fields start empty and focused.
- ☐ Queue counts and ages are real, not decorative badges.

DENSITY Compact throughout. A reviewer works a document in under a minute or two and judges the tool by keystrokes, so ask yours for the real number before you set row height.

D-04

Enterprise data and internal tools

An analyst who lives in the tool eight hours a day, an admin fixing someone else's record, an agent reading the screen while a customer waits.

DEAD ON ARRIVAL "One source of truth." / "Data is the new oil." / "Unlock your data."

COPY

- ☐ Use internal vocabulary, including the ugly words.
- ☐ Headings name the job, not the feature.
- ☐ Errors carry the identifier that can be pasted into a ticket.
- ☐ Permission and audit language must be exact.

INTERFACE

- ☐ The table is the product: sort, filter, column choice, saved views.
- ☐ Bulk selection and bulk action anywhere rows exceed twenty.
- ☐ Filter state lives in the URL so screens can be sent.
- ☐ No dashboard tiles unless a decision was made from one.

DENSITY Compact by default with an optional comfortable mode. Fit as many rows as stay readable on the laptop these people actually use, and give keyboard paths to the handful of actions that fill their day.

D-05

AI product surfaces

A person reviewing model output inside work they are accountable for, deciding whether to accept, edit, or discard it.

DEAD ON ARRIVAL “AI is transforming how we work.” / “Your intelligent assistant.” / “Let AI handle the busywork.”

COPY

- ☐ Say what the model did and did not do.
- ☐ Do not write certainty the system cannot hold.
- ☐ Label generated content in place, not in a footer.
- ☐ State what data leaves the workspace and what is retained.

INTERFACE

- ☐ Generated content is visually distinct and reversible.
- ☐ Citations sit next to the claim, linked to the source row.
- ☐ A stop control exists and partial output is kept.
- ☐ Nothing writes to a record of consequence without a human action.

DENSITY Comfortable for conversational surfaces, compact where output lands in a queue. The human action stays visually dominant.

D-06

Security, identity, cryptography

An admin configuring single sign on, an engineer rotating a key under pressure, a person locked out at the worst possible moment.

DEAD ON ARRIVAL “In an increasingly connected world.” / “Security without compromise.” / “Military grade.”

COPY

- ☐ Name algorithms only where they change what the user does.
- ☐ Say exactly what an action revokes, for whom, and how fast.
- ☐ Write recovery copy for someone already locked out.
- ☐ State the property: at rest, in transit, end to end.

INTERFACE

- ☐ Secrets show once, with a clear copied state.
- ☐ Key lists show state and dates: created, last used, rotating, revoked.
- ☐ Irreversible actions require typing the object name.
- ☐ Session lists show client, approximate location, and last activity.

DENSITY Comfortable setup flows, one decision per screen, done rarely and under stress. Compact audit logs. Dangerous actions separated by space.

D-07

Developer tools and infrastructure

An engineer three tabs deep in documentation at eleven at night, finding out why yesterday's request returns a 403 today.

DEAD ON ARRIVAL "Built by developers, for developers." / "Ship faster." / "The modern way to build."

COPY

- ☐ Every sample runs as written, with its tested version.
- ☐ Error messages name the cause and the next command.
- ☐ Reference states behavior, defaults, and limits. No marketing verbs.
- ☐ Changelogs say what breaks and what to do, with dates.

INTERFACE

- ☐ Monospace anything copied, with a confirmed copy state.
- ☐ Logs need compact rows, wrapping off, and real search.
- ☐ Syntax highlighting correct in both themes.
- ☐ Status and latency shown with the time window.

DENSITY Compact everywhere. Developers read dense text professionally. They need stable navigation and working search, not padding.

D-08

Commerce and retail

A shopper comparing two options on a phone, and the same shopper two weeks later trying to return one of them.

DEAD ON ARRIVAL "Shopping made simple." / "Curated for you." / "Discover your new favorite."

COPY

- ☐ Total cost before the final step, including shipping and tax.
- ☐ Give an arrival date, not only a shipping speed name.
- ☐ Stock language is literal or absent.
- ☐ Return window, return shipping, and condition rules on the product page.

INTERFACE

- ☐ Consistent image crop and background across a category.
- ☐ Selected variant unmistakable, unavailable combinations shown as unavailable.
- ☐ Line by line cart total as soon as it can be calculated.
- ☐ One overlay at a time. Price visible while options change.

DENSITY Comfortable and image led for browsing, compact and factual for checkout and order history. Checkout is one column.

D-09**Consumer subscription and media**

Someone opening the app for the fourth day in a row, and someone else opening it to cancel before Friday's charge.

DEAD ON ARRIVAL "Your journey starts here." / "Level up your life." / "Become your best self."

COPY

- ☐ Onboarding asks only what changes the product today.
- ☐ Notifications say why they arrived and what is being asked.
- ☐ Paywalls name the price, the renewal date, and the trial end behavior.
- ☐ The cancel path reads as plainly as the signup path.

INTERFACE

- ☐ Cancel is as visible as upgrade. No gray on gray retention traps.
- ☐ First run shows real content, not a tour over an empty shell.
- ☐ One modal at launch. Permission prompts arrive with the feature.
- ☐ Celebration is rare and short.

DENSITY Comfortable with large targets and one action per screen. Density belongs in lists, history, and settings.

D-10**Public services, civic, education**

Someone applying for a benefit on a phone with a weak signal, a student on a shared laptop, a caseworker on the thirtieth form today.

DEAD ON ARRIVAL "Government made easy." / "Empowering learners." / "Bridging the digital divide."

COPY

- ☐ Short sentences and common words for a tired reader in a second language.
- ☐ List the documents and numbers needed before the form starts.
- ☐ Say what happens next, how long it takes, and how they will be told.
- ☐ Name the office, the number, and the hours at every dead end.

INTERFACE

- ☐ One question per screen for long forms, with progress and autosave.
- ☐ Upload states are explicit: received, unreadable, wrong type, too large.
- ☐ Test on an old device on a slow connection.
- ☐ Provide a print view. People keep paper.

DENSITY Comfortable and generous for applicants, compact for caseworker queues. Both sides use the same words.

D-11

Logistics, field service, supply chain

A dispatcher rebalancing routes at six in the morning, a driver checking the next stop on a mounted phone, a warehouse lead hunting one pallet.

DEAD ON ARRIVAL "Supply chains are more complex than ever." / "End to end visibility." / "Logistics, reimagined."

COPY

- ☐ Use operational nouns: stop, load, manifest, exception, dwell, proof of delivery.
- ☐ Say when the data was last updated.
- ☐ Exceptions name the cause, the next action, and the owner.
- ☐ Estimates are labeled as estimates, with a window.

INTERFACE

- ☐ Field screens: large targets, high contrast, one handed, readable in sun.
- ☐ Assume dropped connections. Show queued, syncing, and failed.
- ☐ Status on the row. The job is spotting the exception.
- ☐ Map and list are one dataset with shared selection.

DENSITY Compact for dispatch, glanceable for the cab and the floor. Both survive a cracked screen and a bad signal.

D-12

HR, payroll, workforce

An employee looking for one number on a pay statement, a manager approving time off between meetings, an administrator running a cycle that cannot be wrong.

DEAD ON ARRIVAL "People are our greatest asset." / "Reimagining the employee experience." / "Work, simplified."

COPY

- ☐ Say what an action means for pay and for the calendar, in dates and amounts.
- ☐ Use the plan and legal terms exactly as the documents use them.
- ☐ Approval copy says what happens if nobody acts, and by when.
- ☐ State deductions, denials, and limits plainly and link the policy.

INTERFACE

- ☐ Pay statements are tables. People compare periods line by line.
- ☐ Effective dates next to anything that changes money or time.
- ☐ Separate personal actions from actions taken for someone else.
- ☐ Privacy by default in shared and screen shared spaces.

DENSITY Comfortable and calm employee side, one task per screen. Compact administration and approvals with audit trails.

D-13

Legal, compliance, risk

A contract reviewer against a close date, an analyst evidencing a control, an auditor who will only accept what the system can prove.

DEAD ON ARRIVAL "Compliance doesn't have to be a burden." / "Risk into opportunity." / "Legal, modernized."

COPY

- ☐ Quote the clause. Do not summarize an obligation into friendly language.
- ☐ Name the regulation, control, and version with effective dates.
- ☐ Record who approved, when, and on what basis.
- ☐ Generated summaries never replace source text in a decision record.

INTERFACE

- ☐ Source document beside the extracted term, with links to location.
- ☐ Versions need a real diff, not a rewritten paragraph.
- ☐ Status, owner, and due date on every obligation row.
- ☐ Nothing is deleted. Superseded items are marked and kept.

DENSITY Compact and document first, built for long reading. Wide reading column, high contrast, no motion. Print and export are primary.

D-14

Sales, marketing, CRM

A rep updating a record between calls, a manager reading a forecast they do not trust, an operations lead cleaning the fields both ignored.

DEAD ON ARRIVAL "Supercharge your pipeline." / "Data driven growth." / "Close more deals, faster."

COPY

- ☐ Stage definitions live in the product, where the rep is standing.
- ☐ Say what each field is used for. Fields with no consumer collect noise.
- ☐ Forecast copy states the basis and the snapshot date.
- ☐ Automated actions are never logged as a person's action.

INTERFACE

- ☐ Data entry sits inside the workflow the rep already has.
- ☐ List rows show last meaningful activity, not created date.
- ☐ Few required fields, enforced. Bulk edit and undo are required.
- ☐ Charts state their filter and window on the chart.

DENSITY Compact lists and pipeline views, comfortable record pages. Optimize for the short gap between calls, which is the only time this screen gets.

D-15

Travel, hospitality, booking

Someone comparing three itineraries with a deadline, and the same person at an airport at eleven at night trying to change one.

DEAD ON ARRIVAL "Your journey begins here." / "Travel made effortless." / "Discover the world your way."

COPY

- ☐ Total price with taxes and fees as early as it can be calculated.
- ☐ Times carry the zone and the date, including next day arrivals.
- ☐ Change and cancellation rules appear before payment.
- ☐ Say what is included: bags, seats, resort fees, deposits.

INTERFACE

- ☐ The itinerary stays visible while options change.
- ☐ Duration, stops, and layover length on the result row.
- ☐ Confirmation is printable and legible offline, locator in large type.
- ☐ Critical screens load on airport wifi without a video header.

DENSITY Comfortable while browsing, compact and factual once a booking exists. Standing up with luggage means large type, no scrolling.

D-16

Industrial, energy, monitoring

A control room operator watching several hundred signals, and a technician at the equipment deciding whether to shut it down.

DEAD ON ARRIVAL "Industry 4.0 is here." / "Smart operations." / "Data into actionable intelligence."

COPY

- ☐ Use plant tag names and units. Do not invent friendly labels.
- ☐ Alarm text names the signal, threshold, value, and required response.
- ☐ Say whether a reading is live, delayed, or last known, with its age.
- ☐ Never soften a safety message. It is an instruction.

INTERFACE

- ☐ Reserve color for state. A colorful dashboard has no alarm state left.
- ☐ Alarm priority readable by position and shape, not hue alone.
- ☐ Trend charts default to the operator's window with limits drawn.
- ☐ Commanded and measured values are visibly different things.

DENSITY Dense and quiet. Normal should look boring so abnormal is the only thing that draws the eye. No animation.

Patterns

Six screens every system has and most builds skip. Each page gives the situation, what the screen should say, and the rules behind it.

A missing state is a blocker under V-13. Check against these pages and copy the wording.

A CROSS-INDUSTRY CHECK FOR POLISHED WORK WITH NO EVIDENCE, CONTEXT, OR STATES BEHIND IT

[Gates](#) / [C copy](#) / [V interface](#) / [D overlays](#) / **[P patterns](#)** / [A prompting](#) / [Appendix](#)

P-01

Errors

An error answers three questions, in this order. What happened. Is my work safe. What do I do now.

Validation	Name the rule, not the failure. "Account numbers are 10 digits" beats "Invalid input."
Server or network	Say the request did not arrive, say the work is saved, and offer a retry.
Permission	Say who can grant access and how to reach them.
Conflict	Say who changed it and when, and offer both versions.
Unknown	Say what failed, give a reference code, and say what support will need.

RULES UNDERNEATH

- ☐ Never blame the person. Something went wrong is the product's problem, not theirs.
- ☐ Put the message where the failure happened. A banner for a field error makes people hunt.
- ☐ Keep the entered data. An error that clears a form is a second, larger error.
- ☐ One apology at most, and never in place of an explanation.

SLOP

Oops! Something went wrong. Please try again later.

REBUILT

We could not reach the payment service, so the transfer was not sent. Your details are saved. Try again, or call the number below with reference 4471 if it keeps failing.

P-02

Empty states

An empty screen is the first screen most people see. Say what goes here and how to add the first one.

Nothing yet	Say what belongs here and give the single action that creates the first one.
--------------------	--

Nothing found	Say what was searched, show the active filters, and offer to clear them.
----------------------	--

Nothing assigned	Say who assigns work here and what happens when it arrives.
-------------------------	---

Cleared	Say the queue is clear and when it refreshes. This one can be short.
----------------	--

Not available	Say why it is unavailable to this account and who can change that.
----------------------	--

RULES UNDERNEATH

- ☐ Do not use an illustration and a slogan in place of an explanation.
- ☐ The first action in an empty state is the real onboarding.
- ☐ Never show an empty state while data is still loading. Those are different screens.
- ☐ A filtered empty result is not an empty product.

SLOP	No data to display. Get started by adding your first item!
-------------	--

REBUILT	No invoices yet. Invoices appear here once a job is marked complete, and you can also add one by hand.
----------------	--

P-03

Confirmation and destruction

Confirm what cannot be undone. Stop confirming what can. A dialog on everything teaches people to click through the one that mattered.

Reversible	Do not confirm. Act, then offer undo for a few seconds.
Slow to reverse	Confirm with the object named and the consequence stated.
Irreversible	Require the object name to be typed, and list what else is removed.
Affects others	Name who is affected and how many, before the button.
Bulk	State the exact count and the filter it came from.

RULES UNDERNEATH

- ☐ The dialog button says the action. Delete 3 users beats Confirm.
- ☐ Keep destructive actions out of the position where people expect Save.
- ☐ Say what will not be deleted when that is unclear.
- ☐ Never make undo the only safeguard for something the system cannot undo.

SLOP

Are you sure you want to continue? This action cannot be undone.

REBUILT

Delete the Northside account? This removes 214 records and revokes access for 6 people. Exports already downloaded are not affected. Type NORTHSIDE to confirm.

P-04

Permission and audit

Two people read this. One is locked out and in a hurry. The other is deciding what to hand over.

Denied	Say what is restricted, who administers it, and how to request access.
Granting	Say what the role can see and change, in examples, not role names.
Elevated	Say the session is elevated, for how long, and what is logged.
Acting for another	Mark it on every screen. Never let an agent forget whose account they are in.
Revoked	Say when access ends and what happens to work in progress.

RULES UNDERNEATH

- ☐ Do not hide a feature without saying it exists and who can unlock it.
- ☐ Role names mean nothing alone. Show two or three concrete permissions.
- ☐ Log actor, time, and object, and let people read the log without a ticket.
- ☐ Impersonation requires consent, a visible banner, and an audit entry.

SLOP You do not have permission to access this resource.

REBUILT Only workspace owners can change billing. Dana Ruiz and Priya Shah own this workspace. Request access and they will get a message with this page attached.

P-05

Notifications and email

A notification spends someone else's attention. Make it worth the interruption and say why it came.

Action needed	Name the action, the deadline, and what happens if nobody acts.
Status change	Name the object, the old state, the new state, and who changed it.
Digest	Group by object, not by event, and cap the length.
Alert	State the threshold crossed and the current value, and link to the source.
Marketing	Keep it out of the transactional channel and honor unsubscribe immediately.

RULES UNDERNEATH

- ☐ The subject line carries the whole message on a lock screen.
- ☐ Say why this person received it, especially when someone else added them.
- ☐ One notification per event, not one per channel per event.
- ☐ Put the controls in the message. Turning it off should not require a hunt.

SLOP

You have a new update! Check it out in the app.

REBUILT

Invoice 8841 was rejected by Accounts Payable. It needs a corrected PO number before Friday, or it returns to the requester. You get this because you submitted it.

P-06

Onboarding and money

The first session, and the moment money is involved. People read both for honesty in about four seconds.

First run	Show the product with real content in it, not a tour over an empty shell.
Setup questions	Ask only what changes the product today. The rest can wait.
Trial start	State the end date, the price after, and whether a card is required.
Trial ending	Say the date, the amount, and what happens to the data if they do nothing.
Cancel	Say when access ends, what is kept, what is deleted, then let them finish.

RULES UNDERNEATH

- ☐ Write the cancel path as plainly as the signup path. That comparison is the whole test.
- ☐ Never disguise the decline option as a quiet link beside a full button.
- ☐ State the renewal date and amount on the payment screen, not in a later email.
- ☐ A locked feature says what unlocks it once, then stops following the person around.

SLOP

Start your free trial today. Cancel anytime!

REBUILT

Free for 14 days. We will charge 29 dollars on 12 October unless you cancel before then, and we will email you three days ahead. Cancelling keeps your data readable for 30 days.

RULES A

Prompting

Most slop is decided before anything is generated. A prompt that leaves out the domain, the user, the problem, the evidence, or the limits gives an agent room to guess, and it will guess with defaults.

This is scope control, not prompt engineering. No structure here guarantees a good result.

A CROSS-INDUSTRY CHECK FOR POLISHED WORK WITH NO EVIDENCE, CONTEXT, OR STATES BEHIND IT

Gates / C copy / V interface / D overlays / P patterns / **A prompting** / Appendix

A / CONTEXT

The smallest useful context

Four inputs. They do not make a production prompt, but without them an agent has to invent the situation before it can start the work.

Industry	The domain and the constraints that come with it: terminology, regulation, evidence rules, what may not be invented.
User	The specific person doing the work. Not users. An oncology nurse, a claims reviewer, a dispatcher.
Verb	What that person is trying to get done. Review, reconcile, approve, compare, submit.
Adjective	The quality the experience needs. Quickly, safely, without losing their place, without a second screen.

WORKED EXAMPLE

Healthcare + oncology nurse + review + quickly.

FOUR INPUTS, WRITTEN OUT

Design a healthcare workflow for an oncology nurse to review incoming specimen information quickly. Prioritize scanability, clear status, error recovery, and the information required to make the next decision. Use the supplied workflow and evidence as the source of truth. Do not invent clinical requirements, workflow states, metrics, regulatory constraints, or product behavior that have not been provided.

The last sentence is doing as much work as the first. Naming the source of truth and the things that may not be invented is what keeps the output inside the evidence you actually have.

A / ARCHITECTURE

Nine parts of a working prompt

The context layer says who and what. This says what is wrong, what may change, and how the result gets checked. Skip a part and the agent fills it in for you.

Problem	What is failing in the current output or experience, stated as a behavior rather than a feeling.
Evidence	The research, screenshot, workflow, requirement, ticket, artifact, or test result that shows the problem.
Diagnosis	What is specifically wrong. Not “make it cleaner” or “improve the UX,” which leave the agent to decide what was broken.
Proposed solution	What should change.
Why it is better	How the change helps the user, the workflow, comprehension, accessibility, or the operation. Different is not better.
Scope	What the agent is allowed to modify.
Exclusions	What must not be added, removed, redesigned, inferred, or reinterpreted.
Constraints	Fixed dimensions, terminology, design-system rules, product behavior, domain requirements, technical limits, evidence boundaries.
Validation	How to check the result against the original problem.

A / SURFACES

Change surface and preservation surface

Every prompt should name both. The change surface is what the agent may touch. The preservation surface is what has to come back unchanged.

This matters most on work that already exists. Coding and design agents tend to read a request to improve one area as permission to regenerate what surrounds it: neighbouring components, layout, copy, routes, dimensions, states, interactions. Naming the preservation surface is usually cheaper than reviewing the diff afterward.

WHAT EACH KIND OF REQUEST MAY TOUCH

New build	Structure, layout, hierarchy, and presentation are open. Research, metrics, domain rules, product requirements, regulations, workflow states, and behavior are not, unless supplied.
Modification	Preservation is the default. Only the named surface changes. Functionality, routes, components, content, dimensions, interactions, visual rules, responsive behavior, and implementation patterns stay as they are.
Visual refinement	Hierarchy, spacing, typography, alignment, composition, emphasis. Not information architecture, workflow logic, data meaning, interaction behavior, or content intent.
Bug fix	The observed failure. Nothing else gets redesigned or refactored on the way past.
Copy update	The wording. Layout and functionality hold unless the new content genuinely needs room.

A / SURFACES

What each kind of request may touch, continued

Responsive work	Recompose for the target viewport. Keep hierarchy, functionality, information boundaries, important proportions, and core interactions. Shrinking desktop, hiding essential actions, and reducing tap targets are not adaptations.
Iteration	Say what is already working and must survive, not only what is wrong.
Design system	State whether the change is local or system-wide. A one-off fix should not propagate unless that is the intent.
Research-led	Keep observation separate from interpretation. Qualitative findings do not become numbers, and speculation does not become a requirement.

A / EXCLUSIONS

Telling it what not to do

Negative instructions are part of the structure, not a footnote at the end. One constraint prevents more bad output than another adjective.

- ❑ Do not invent metrics.
- ❑ Do not introduce new workflow states.
- ❑ Do not change existing interaction logic.
- ❑ Do not rewrite clinical, legal, regulatory, financial, or safety terminology.
- ❑ Do not add decorative treatments unrelated to the identified problem.
- ❑ Do not simplify the interface by removing required states.
- ❑ Do not infer missing requirements. Surface the missing information instead.
- ❑ Do not use a copy change as permission to redesign the component.
- ❑ Do not change established dimensions or proportions.
- ❑ Do not refactor unrelated systems while fixing a local bug.
- ❑ Do not create claims, baselines, outcomes, or user findings that the source material does not support.
- ❑ Do not add components because they are common defaults in generated interfaces.

Exclusions change with the job

In healthcare the thing to block is invented clinical logic. In developer tools it is a changed API. During a visual pass it is logic edits. During responsive work it is functionality quietly disappearing at a smaller width. Write the two or three that apply to what you are asking for, rather than pasting the same list every time.

A / FRAMEWORK

A prompt you can reuse

Fill the brackets. It is one paragraph on purpose: Claude's text field handles heavy formatting badly, and a prompt should survive having its bullets and line breaks stripped out.

PASTE AND FILL

This is a [industry] product used by [user] to [verb] [adjective]. The current problem is [observable problem]. The evidence for that problem is [research, screenshot, workflow, requirement, behavior, artifact, or source]. What is specifically wrong is [diagnosis]. The proposed solution is [specific change]. This is a better solution because [reason tied to the user, workflow, accessibility, comprehension, product, technical, or operational need]. You may change [change surface]. Preserve [preservation surface]. Do not [prohibited actions, unsupported assumptions, invented states, unwanted visual defaults, or unrelated changes]. Exclude [elements that do not belong in this solution]. Respect these constraints: [dimensions, terminology, product rules, design-system rules, domain requirements, evidence boundaries, technical limitations]. Use [source material] as the source of truth. Validate the result by checking [conditions that prove the change addresses the problem]. If information required to complete the change is missing, identify the gap instead of inventing an answer.

Prefer instructional clunks over formatted prompts

If a prompt stops making sense once the indentation and line breaks are gone, it was depending on the formatting to carry meaning. Write it as plain sentences that read in order. It will be uglier and it will hold up better.

A / EXAMPLE

Weak against architected

Same intent, same screen. The difference is how much is left for the agent to decide.

WEAK

Make this dashboard cleaner and easier to use.

ARCHITECTED

This oncology review dashboard gives status, urgency, supporting metadata, and next actions nearly equal visual weight, which forces nurses to scan several fields before identifying what requires action. Increase the hierarchy of status and next action and reduce visual emphasis on secondary metadata so the next operational decision becomes easier to identify. Preserve the existing workflow, terminology, navigation, states, actions, data, and product logic. Do not introduce new clinical requirements, new workflow states, new metrics, gradients, decorative components, or unrelated layout changes. Use the supplied workflow and interface as the source of truth, and validate the result by confirming that the primary action and current status are visually identifiable before secondary metadata, without removing required information.

The first version asks for an opinion about cleanliness. The second names the failure, the change, what stays, what is off limits, and the check that decides whether it worked. Neither one guarantees a good screen. The second makes a bad one easier to catch.

A / RULES

Prompting rules

Cite these the way you cite the others. A-04 in a review comment means the prompt had no preservation surface.

A-01 Describe the failure before prescribing the interface

A prompt that opens with the solution skips the part only you know: what is going wrong now.

FLAGS make it cleaner / improve the UX / modernize this / make it more intuitive

FIX State the observable problem and the evidence for it, then say what should change.

A-02 Name the evidence and the source of truth

Without a named source, the agent treats its own defaults as the source.

FLAGS No workflow, screenshot, requirement, ticket, or research attached. No sentence saying which one governs.

FIX Point at the artifact and say to use it as the source of truth. Say what may not be invented around it.

A-03 Say why the change is better, not how it differs

Different is not an outcome. The reason is what lets an agent choose between two ways of doing it.

FLAGS Rationales that describe the new design rather than the effect on the user, workflow, or operation.

FIX Tie the change to the person doing the work: what gets faster, clearer, safer, or possible to recover from.

A-04 Define a change surface and a preservation surface

Agents read a request to improve one area as permission to regenerate what surrounds it.

FLAGS A prompt with no "preserve" clause. Diffs touching routes, components, or copy nobody asked about.

FIX Name what may change and what must come back untouched, in the same sentence if possible.

A-05 Write the exclusions that apply to this job

Negative instructions are part of the architecture. A constraint prevents more bad output than another adjective.

FLAGS No “do not” anywhere in the prompt. The same pasted exclusion list regardless of the task.

FIX Pick the two or three failures likely here: invented clinical logic, an API change, logic edits during a visual pass, features vanishing at a smaller width.

A-06 Set the constraints that are already fixed

Dimensions, terminology, design-system rules, and technical limits exist whether or not you mention them.

FLAGS Output that breaks the grid, renames domain objects, or ignores a settled platform limit.

FIX List the fixed values and the words that may not be changed. Keep it to what actually binds.

A-07 Expose gaps instead of filling them

The most damaging thing an agent can do with a missing requirement is supply a plausible one.

FLAGS Confident output covering something you never provided: a state, a metric, a rule, a regulation.

FIX Instruct it to identify what is missing and stop. Then supply the fact, or decide the work is not ready.

A-08 Match the guardrails to the kind of request

A new build, a modification, a bug fix, and a responsive pass need different defaults.

FLAGS One prompt style for every task. Refactors arriving with bug fixes.

FIX Use the change surface table in the A section. Preservation is the default everywhere except a new build.

A-09 Keep prompts readable as plain text

Heavy formatting and frequent line breaks do not survive the trip into most text fields.

FLAGS A prompt that only parses with its indentation intact.

FIX Write instructional clunks: plain sentences in order, understandable with every line break removed.

A-10 A prompt reduces guessing, it does not replace review

Better input narrows the range of bad output. It does not verify anything.

FLAGS Work shipped because the prompt was thorough, with no audit between generation and release.

FIX Run the result through the gates and the C, V, D, and P rules. Prompting is upstream governance, not a replacement for the audit.

Reference

Search lists to run literally, the text to paste when invoking an audit, the constraints to paste when generating, and a card small enough to keep open while working.

A CROSS-INDUSTRY CHECK FOR POLISHED WORK WITH NO EVIDENCE, CONTEXT, OR STATES BEHIND IT

Gates / C copy / V interface / D overlays / P patterns / A prompting / **Appendix**

APPENDIX A

Invoking an audit

Attach this file and paste one of these. The first is for review, the second for repair, the third for generation.

REVIEW

AUDIT. Use the attached rule set. Artifact: [what it is]. Audience: [role and task].

Domain: [D-xx or none]. Stage: [proposed, built, shipped]. Evidence available: [what].

Run the gates, then C-01 to C-20, then V-01 to V-19, then the domain overlay, then P-01 to P-06.

Report findings in the protocol format. Change nothing. Blockers first.

Where a rule is not checkable from what I gave you, say so rather than passing it.

REPAIR

FIX. Same rule set and same passes. Apply only the fixes that need no facts I have not given you.

Show before and after for every string you change. List the rest as questions, naming the missing fact.

Do not invent numbers, names, quotes, or sources. Do not touch legal, clinical, or safety strings.

GENERATE

BUILD. Follow appendix C as hard constraints while generating.

Before showing me anything, audit your own output against V-01 to V-19 and P-01 to P-06,

fix what fails, and tell me which states you implemented.

APPENDIX B

The search lists

Run these as literal searches at the end, not as a reading pass. Each entry is a phrase that survives review because it is individually defensible.

COPY CURRENT AS OF THE 2025 TO 2026 MODEL GENERATION

Date this list and expect to rewrite it. Wikipedia's editors sort their vocabulary notes by model era because the words turn over that fast, and Kobak et al. (2025) publish an annotated set of 900 excess words that is a better starting point than anything hand-written.

not just, more than just, rather than, instead of, wasn't X it was Y

The challenge, The result, The opportunity, the real problem, the deeper issue

That's when, That's where, Enter, Meet, Introducing

seamless, intuitive, robust, holistic, frictionless, thoughtful, powerful, effortless, magic

leverage, utilize, streamline, optimize, reimagine, elevate, empower, foster, facilitate

ecosystem, journey, touchpoint, pain point, north star, unlock, solution, surface, solve for

insight, impact, transformative, data driven, research backed, best practice, production ready

at its core, ultimately, notably, importantly, critically, interestingly

what emerged, the key insight, the missing piece, the bigger picture, one thing became clear

at the intersection of, where X meets Y, bridging the gap, closing the loop, from X to Y

collaborated closely, aligned stakeholders, cross functional, end to end, wore many hats

Oops, Something went wrong, Are you sure, No data to display, Get started, Learn more, Click here

INTERFACE RUN AGAINST THE STYLESHEET AND THE RENDERED SCREEN

Spacing values outside the scale. Inline spacing overrides. Equal gaps inside and between groups.

More than four vertical start positions. Centered paragraphs. Centered numeric columns.

More than five type sizes. Line length over 75 characters. Justified text.

Gradient backgrounds. More than one accent. Status colors used decoratively. Meaning by hue alone.

Mixed radii. Shadows on every surface. Two primary buttons. A fourth button style.

Missing empty, loading, partial, error, denied, or stale state. Spinner instead of skeleton.

Truncated bar axes. Unlabeled units or windows. Metrics with no definition.

Contrast under 4.5 to 1. Targets under 44 points. Removed focus outlines. Placeholder used as label.

Un sourced statistics, placeholder logos, fake testimonials, emoji icons, faceless illustrations.

APPENDIX C

Build constraints

Paste above any generation request and fill the bracketed lines. Most failures in V happen because a screen was requested with no system to build it from.

PASTE THIS

Build this for [role] doing [task]. Density: [compact or comfortable]. Domain: [D-xx or none].

Spacing: 4px base. Use only 4, 8, 12, 16, 24, 32, 48, 64. Group gaps exceed internal gaps.

Type: four sizes, two weights, body measure 45 to 75 characters, left aligned.

Color: neutral surface and border, one accent, reserved status colors. No gradients.

Components: one button hierarchy with a single primary per screen. One radius scale. Two elevations.

Layout: content inside a max width, aligned to no more than four vertical edges.

Implement these states: empty, loading skeleton, partial data, error, permission denied, stale, longest realistic content, and 40 rows of data.

Do not add: hero gradients, icon feature cards, centered body copy, pill badges, invented statistics, testimonials, placeholder logos, emoji icons, or stock illustration.

Copy: buttons name action and object. Errors say what happened, whether work was lost, and what to do.

No marketing adjectives anywhere in the interface.

APPENDIX D

The card

The whole rule set, short enough to keep on a second screen.

WORDS

- ☐ No manufactured contrast, no reveals, no three item rhythm, no inspirational endings.
- ☐ Name the actor, the action, and the observable change.
- ☐ Smallest accurate word. State the stage: proposed, tested, built, shipped, adopted, stopped.
- ☐ Every number carries a baseline, a window, and a source, or it goes.
- ☐ Report failed rules, never authorship. No finding says a person used AI.
- ☐ Stop when the draft is true, not when it is neutral. A position is not slop.
- ☐ Every section readable by someone who landed there first.
- ☐ Buttons name the action. Errors say what happened and what to do next.

SCREENS

- ☐ One job per screen, one primary action, one accent.
- ☐ Every value from the scale. Group gaps larger than internal gaps.
- ☐ Four vertical edges at most. Text aligns to text, numbers align right.
- ☐ Density chosen for the person who uses it daily.
- ☐ Empty, loading, partial, error, denied, stale, and longest content all built.
- ☐ Contrast, focus, target size, and keyboard paths verified rather than assumed.

If the prose is cleaner than the truth, or the screen is prettier than the thinking behind it, rebuild it from evidence.

SOURCES

Where this came from

What the rules rest on, what I cut from the earlier versions, and the references. A document about evidence should show its own.

PROVENANCE

What changed in version 3

Versions 1 and 2 came from my own practice, with no research behind them. So the document broke its own rule about unsourced numbers, C-10, several times over.

REMOVED

- ❑ Eight density figures that read as measured and were not: clinical worklists at 32 to 40 rows, reviewers at 20 to 90 seconds per document, internal tools at 25 or more rows on a laptop, and the rest. They now describe direction and tell you to go get the real number from the people who do the work.
- ❑ The 120 to 240 millisecond motion window. I have no source for it. The rule now says what the constraint is for and leaves the number to whoever is testing the screen.
- ❑ The implication that a target must be 44 by 44 points to pass. That is the AAA criterion and the Apple platform guidance. The Level AA floor in WCAG 2.2 is 24 by 24 CSS pixels with a spacing exception, which is a different claim and the one the rule now makes.

ADDED

- ❑ C-03 on flat variance, which the corpus research supports better than any vocabulary rule in here.
- ❑ C-22, the counter-rule. A full scrub can leave prose that passes every check and argues nothing, and nothing in versions 1 or 2 stopped that from happening.
- ❑ A signal-strength page in the protocol, splitting what is measured from what is folklore.
- ❑ Citations in the rules themselves, so a reader can check a claim without taking my word for it.

CORRECTED

- ❑ The em dash. I had it as a tell. A 2026 measurement of twelve models puts GPT-4.1 near Mark Twain's rate in Huckleberry Finn and two Llama models at zero, so it is now filed as a weak signal that only counts alongside others.
- ❑ Vocabulary lists are now marked as perishable and dated by model generation, which is how the Wikipedia editors who track this actually maintain theirs.
- ❑ The document no longer reports authorship. It reports failed rules. Detector output is not evidence and the false positives land hardest on people writing in a second language.

PROVENANCE

What is still thin

Matching claims to evidence works both ways. Here is what I would not defend if you pushed on it.

- ❑ The sixteen domain overlays come from pattern knowledge and from work I have done, not from field research. D-01 and D-03 are the ones I would trust most because I have worked in those tools. The industrial and legal pages are the ones I would check with a practitioner before quoting.
- ❑ Every scenario in the slop-and-rebuilt pairs is invented. They are illustrations of a shape, not case studies, and no client, ticket, or incident sits behind any of them.
- ❑ The platform prevalence numbers are detector output. Two studies of LinkedIn in the same month reported 81 percent and 41 percent using different thresholds and sampling. The direction is solid, the precision is not, and one of the vendors sells detection.
- ❑ The generational reading in the copy rules rests mostly on trade research. The peer-reviewed part is narrow: a period at the end of a short text message reads as less sincere than no punctuation, which held in text but not in handwriting.
- ❑ The audience research on whether readers punish AI-sounding work is genuinely split. Some studies find a brand authenticity penalty, at least one finds none when the text is disclosed. The behavioral evidence, meaning engagement and platform distribution, is stronger than the survey evidence.

REFERENCES

Works cited

APA 7 author-date, which converts to Chicago author-date with almost no change. Where I found something through someone else's summary, the entry says so.

STANDARDS AND PLATFORM GUIDANCE

TetraLogical. (2022, December 20). *Foundations: Target sizes*.

<https://tetralogical.com/blog/2022/12/20/foundations-target-size/> [Secondary source for Apple's 44 by 44 point and Material Design's 48 dp guidance.]

W3C. (2023). *Web Content Accessibility Guidelines (WCAG) 2.2*. World Wide Web Consortium.

<https://www.w3.org/TR/WCAG22/> [SC 1.4.3 contrast, 1.4.4 resize text, 1.4.11 non-text contrast, 2.5.5 target size enhanced, 2.5.8 target size minimum.]

Wikipedia contributors. (2010). *Wikipedia: Manual of style/Words to watch*. Wikipedia.

https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style/Words_to_watch

LANGUAGE AND CORPUS RESEARCH

Gunraj, D. N., Drumm-Hewitt, A. M., Dashow, E. M., Upadhyay, S. S. N., & Klin, C. M. (2016). Texting insincerely: The role of the period in text messaging. *Computers in Human Behavior*, 55, 1067-1075.

<https://doi.org/10.1016/j.chb.2015.11.003>

Kobak, D., González-Márquez, R., Horvát, E.-Á., & Lause, J. (2025). Delving into LLM-assisted writing in biomedical publications through excess vocabulary. *Science Advances*, 11(27), eadt3813.

<https://doi.org/10.1126/sciadv.adt3813>

Kujur, A. (2025). *A comparative analysis of AI-generated and human-written text: Linguistic patterns, detection accuracy, and implications for modern communication* [Preprint]. SSRN.

Zanotto, S. E., & Aroyehun, S. (2024). *Human variability vs. machine consistency: A linguistic analysis of texts generated by humans and large language models* (arXiv:2412.03025). arXiv. <https://arxiv.org/abs/2412.03025>

REFERENCES

Works cited, continued

AI CONTENT IN THE WILD

Brooks, C., Eggert, S., & Peskoff, D. (2024). The rise of AI-generated content in Wikipedia. In *Proceedings of the First Workshop on Advancing Natural Language Processing for Wikipedia* (pp. 67-79). Association for Computational Linguistics. <https://aclanthology.org/2024.wikinlp-1.12/>

HumanizeAI. (2026, July 14). *Wikipedia's signs of AI writing: The full list*. [Secondary source for E. M. Freeburg's 2026 preprint measuring em dash rates across twelve models.]

Originality.AI. (2024, November 27). *Over half of long posts on LinkedIn are likely AI-generated since ChatGPT launched*. <https://originality.ai/blog/ai-content-published-linkedin>

Originality.AI. (2026a, January 22). *50%+ of LinkedIn posts were likely AI in 2025 plus engagement insights*. <https://originality.ai/blog/linkedin-ai-study-engagement>

Originality.AI. (2026b, July 30). *LinkedIn AI content study: 81% of long-form posts are likely AI*. <https://originality.ai/blog/ai-content-published-linkedin>

Vynce Digital. (2026). *LinkedIn AI content: What marketers need to know in 2026*. [Secondary source for Pangram Labs' July 2026 feed study of 1,002,627 posts.]

Wikipedia contributors. (2026a). *Wikipedia: Signs of AI writing*. Wikipedia. https://en.wikipedia.org/wiki/Wikipedia:Signs_of_AI_writing

Wikipedia contributors. (2026b). *Wikipedia: WikiProject AI Cleanup/Guide*. Wikipedia. https://en.wikipedia.org/wiki/Wikipedia:WikiProject_AI_Cleanup/Guide

TYPOGRAPHY

Brighurst, R. (2004). *The elements of typographic style* (3rd ed.). Hartley & Marks. [Conventional source for the 45 to 75 character measure, which is craft consensus rather than a tested threshold.]

If you find a claim in here with no citation and no hedge, that is a defect. Tell me which rule and I will either source it or cut it.

RELEASE GATE

Say it aloud, size it to the evidence, and put a reason behind every value.

Nothing ships with an open blocker. Nothing ships while a claim outruns its evidence, a required state is missing, or a value on the screen came from nowhere.

The goal was never to sound less polished. It is to sound and look like people who were there, making specific decisions for reasons they can still explain a year later.