Making Intent Legible Before Generation

Created by Pat Glatz · 2025

1. Project Overview

Image generation systems reliably produce fluent output even when user intent is underspecified, silently resolving ambiguity in ways users cannot inspect, audit, or correct. In practice, this shifts responsibility from the user to the system while obscuring where intent was inferred, distorted, or invented.

This project was developed in response to that failure mode. Rather than optimizing prompt quality or creative output, it introduces a pre-generation boundary that forces intent to become explicit before delegation occurs.

The system was initially deployed in middle-school graphic design classes, where the consequences of silent resolution were immediately visible. Students would receive plausible but incorrect images—misaligned compositions, inverted relationships, unintended text, or semantic errors—without any clear indication of which assumptions the model had filled in on their behalf. The issue was not lack of creativity, but loss of diagnosability.

To address this, I built a guided, one-question-at-a-time prompt-building system (December 2024) that externalizes decision-making prior to image generation. The goal was not to teach prompting, but to prevent generative systems from resolving ambiguity invisibly.

My Role

I designed and built the full tiered intent-governance system, including workflow logic, question architecture, schema design, guardrails, validation checks, UX scaffolding, and all supporting prototypes and artifacts.

---

2. The Problem

Image-generation tools such as Adobe and DALL·E accept natural-language prompts but do not require intent to be fully specified. When prompts are ambiguous, these systems proceed anyway, filling gaps autonomously and producing outputs that appear confident but may not reflect the user's actual goal.

In classroom use, this failure surfaced repeatedly. Students would generate images that were technically coherent but semantically wrong—animals wearing sweaters instead of animals on sweaters, inverted foreground–background relationships, merged objects, or unintended foreign-language text. The systems did not signal uncertainty or request clarification; they resolved ambiguity silently.

The core problem was not that users "prompted poorly," but that the systems provided no mechanism for making intent inspectable before generation. Once an image was produced, it was often impossible to determine whether the error originated from unclear intent, model assumption, or misinterpreted constraints. Correction became iterative guesswork rather than accountable revision.

This pattern mirrors a broader system risk: when delegation occurs before intent is legible, failures appear as user error or output quality issues rather than as governance gaps.

---

3. Users & Needs

The initial users were middle-school students learning introductory graphic design, working under time constraints and limited teacher availability. They needed a way to express visual intent without specialized vocabulary, but more critically, they needed a system that would not proceed until intent was sufficiently specified.

As development progressed, the same failure pattern appeared across adults, small-business owners, and hobbyists. These users were capable of forming clear goals but still encountered silent assumption-making by image generators. Across contexts, the underlying need was consistent: a structured mechanism that surfaces gaps, contradictions, or uncertainty in intent *before* generation, rather than allowing the model to resolve them implicitly.

---

4. Research & Analysis

Informal observation during remote classes revealed recurring categories of failure: unintended actions, inverted compositions, merged elements, and incorrect text. These errors were consistent with models extrapolating from incomplete instructions rather than misunderstanding explicit ones.

To understand what information was missing at the point of delegation, I consulted two graphic designers to identify the components typically required in a professional design brief. Their input surfaced stable variables—purpose, audience, style, composition, constraints—that are rarely all present in a casual prompt but are necessary for accountable interpretation.

The resulting analysis suggested that the primary intervention point was *before generation*, not during revision. The system therefore needed to externalize intent formation itself, rather than optimize post-hoc correction.

---

5. Design Criteria

- Interrupt generation until intent is sufficiently specified
- Make decision points explicit and reviewable
- Prevent silent resolution of ambiguity
- Scale language and depth without changing structural rigor
- Align outputs with Adobe and DALL·E interpretation requirements
- Produce structured, transferable representations of intent (JSON-ready)

Accessibility and cognitive-load reduction were design constraints, not primary objectives.

---

6. The Solution

The solution is a guided prompt-building system that functions as an intent-governance layer rather than a prompt tutor. It uses sequential, one-question-at-a-time interaction to externalize decisions that image-generation systems would otherwise infer silently.

This structure is intentional. By isolating each decision, the system makes gaps, contradictions, or uncertainty in user intent visible *before* delegation occurs, rather than allowing the model to resolve them autonomously.

The tool operates across three tiers:

- **Basic**: simplified language for novice users
- **Plus**: expanded depth for general users and small projects
- **Pro**: design-forward terminology and constraint handling for professional workflows

All tiers follow the same underlying pathway: clarify intent, surface constraints, confirm assumptions, and only then assemble a final prompt. Each session produces a structured, JSON-formatted representation of intent designed for direct transfer into Adobe and DALL·E workflows.

Guardrails prevent topic drift and enforce alignment with the user's stated goal, ensuring that generation occurs only after intent is legible.

---

7. Implementation

The prototype was developed within a GPT environment, with emphasis on enforcing role discipline between questioning, validation, and assembly. I began with the Plus tier to establish structural completeness, then constrained it downward for students by reducing field count and simplifying language without altering decision structure.

Input from two graphic designers informed expansion into the Pro tier, adding contradiction handling and stricter validation aligned with professional design briefs.

The system was piloted during a remote middle-school graphic design unit. Students required no direct instruction to use it, allowing observation of how the interaction model performed when human intervention was minimal.

---

8. Outcomes & Observations

In baseline assignments completed without the system, students frequently produced images that were plausible but misaligned with intent, leading to frustration and repeated trial-and-error.
When the system was available, a subset of students chose to use it. These users produced images that aligned more closely with their stated goals, completed work with fewer corrective iterations, and finished assignments more consistently.
While these outcomes resemble instructional improvements, they are better understood as *secondary effects* of restored diagnosability. By making intent explicit before generation, users could see what decisions were being made and why, reducing both confusion and corrective churn.
Adult users applied the system to professional and personal projects (e.g., book covers, logos, artwork), further validating that the failure mode extended beyond the classroom.

---

9. Reflection

This project clarified that many frustrations attributed to "prompting difficulty" are in fact governance failures. When systems accept delegation under ambiguity without surfacing assumptions, responsibility becomes diffuse and correction becomes performative.
The most significant technical insight emerged during tiering. Reducing language complexity without enforcing tier-specific constraints allowed higher-tier logic to leak downward, reintroducing hidden complexity. Stability required not just simplified language, but strict field limits and tier-locked question sets to preserve accountability at each level.
Ultimately, the system's value lies not in helping users say the "right thing," but in preventing systems from deciding what the user meant without making that process visible.

10. Artefact Gallery (Appendix)

These artifacts collectively demonstrate the system's instructional scaffolding, tiered logic, schema depth, and how user intent is shaped into precise, AI-ready outputs.

Purpose: This appendix provides a coherent set of artefacts that demonstrate the system architecture, user experience flow, and instructional design decisions behind the tiered AI prompt-builder.

A. Tier Comparison Matrix

A structured comparison of the three tiered experiences.

Feature	Basic	Plus	Pro
Question Depth	Low	Medium	High (34 fields)
Language Level	Student-friendly	General consumer	Professional/technical
Schema Control	Light JSON	Mid-weight	Expanded schema with contradiction handling
Validation	Minimal	Soft validator	Strict validation rules
Contradiction Handling	None	Limited	Active contradiction resolution
UX Scaffolding	Single-step	Guided + suggestions	Structured intake flow
Output	Simple prompt + JSON	Summary + variants	Structured prompt variants

B. Workflow Diagrams

Clear conceptual diagrams describing how each tier processes user intent.

Basic Flow

1. User idea → 2. One-question intake → 3. Summary → 4. Prompt → 5. JSON

Plus Flow

1. Context setup → 2. Subject details → 3. Setting + mood → 4. Stylistic cues → 5. Framing + format → 6. Constraints → 7. Variants → 8. Summary → 9. Final prompts

Pro Flow

1. Extended Intake → 2. validation → 3. Contradiction Check 4. file consolidation → 5. Prompt assembly