

THE TMINUSAI PROMPT SCORECARD

Decision-Grade Evaluation, QA Loops & Adversarial Testing

INSIDE THIS GUIDE:

- The 5-dimension Prompt Scorecard
- Adversarial Prompting Loop
- 10 battle-tested prompt templates
- Diagnostic & Fix Protocol
- Prompt Debugging Decision Tree
- The 10-Shot Consistency Test

4

From Prompt Writing to Prompt Engineering

In professional AI work, the "first-shot" prompt is rarely the "final-shot" solution. The difference between a \$5 output and a \$500 output is not the model — it is the rigor of the iteration loop. This guide teaches you to engineer prompts the way a systems architect engineers software: with specification, testing, and failure analysis.

THE TMINUSAI AXIOM

"AI reliability is a function of the iteration loop, not the initial ask. A mediocre prompt + a rigorous QA loop beats a brilliant prompt with no follow-through."

The 5 Root Causes of Bad AI Output

Root Cause	How It Manifests	Frequency
Vague Persona	Generic expert voice; no domain specificity; off-target tone	Very High
Undefined Objective	Model guesses at the goal; produces plausible but wrong output	Very High
No Failure Mode Plan	Model encounters edge case and hallucinate a solution instead of flagging it	High
No Reasoning Trigger	Model gives shallow first-pass answer without examining the problem space	High
No Output Schema	Output format varies per run; cannot be parsed programmatically	Medium

SECTION 1 — The TminusAI Prompt Scorecard (TPS v2)

Score Every Prompt Before You Send It

Use this 5-dimension rubric to evaluate any prompt. A score of 8 or higher is required before submitting a prompt to a production workflow. Score each dimension 0 (Fail), 1 (Pass), or 2 (Mastery).

Dimension	0 — Fail	1 — Pass	2 — Mastery
Persona Rigor	Generic ("expert")	Specific role ("CRO Analyst")	Domain + bias constraints ("CRO Analyst who avoids vanity metrics")
Objective Clarity	Vague ("help me")	Clear goal ("Draft X")	Goal + success criteria + stated non-goals
Failure Mode Plan	None	"If error, tell me"	"If X occurs, execute Y; else Z. If uncertain, ask."
Reasoning Trigger	None	"Think step-by-step"	Explicit Reasoning Escalation Level (REL L1/L2/L3)
Output Schema	None	"Markdown table"	JSON schema / Pydantic model / Mermaid.js / exact example

HOW TO USE THE SCORECARD

Score your prompt on all 5 dimensions. If total < 8/10: identify the lowest scoring dimension and fix it first. A score of 10/10 is achievable with practice and is required before any prompt runs in an automated, unmonitored pipeline.

SECTION 2 — The Diagnostic & Fix Protocol

When the AI Gives You a Bad Answer

Do not just re-submit the same prompt. Diagnose the specific failure point and apply a targeted fix. Here is the professional operator's protocol.

Case Study 1 — The Vague-Ask Trap

Bad Prompt (Score: 2/10):

```
"Give me some ideas for a new marketing campaign."
```

Diagnosis: Missing Persona (0), Objective (0), Output Schema (0). The model has no domain, no budget, no target audience, and no output format.

Fixed Prompt (Score: 10/10):

```
"You are a Senior Growth Marketer for a B2B SaaS company targeting CTOs. Budget: $3k/month. Channel: LinkedIn only. Goal: Generate 50 qualified demo requests in Q2. Suggest exactly 3 campaign concepts. For each, provide: [Concept Name] | [Hook angle] | [Ad format] | [Expected CPL range]. Do not use generic buzzwords. If any concept requires budget >$3k/month, flag it instead of including it."
```

Case Study 2 — The Missing Schema Trap

Bad Prompt (Score: 4/10):

```
"Analyze this customer feedback and identify the main themes."
```

Diagnosis: Good objective, but no output schema — every run produces a different format that cannot be programmatically processed.

Fixed Prompt (Score: 9/10):

```
"Analyze the attached customer feedback. Output ONLY a JSON array where each object has: {theme: string, frequency: integer, severity: \"low\"|\"medium\"|\"high\", representative_quote: string (max 20 words)}. Sort by frequency descending. Do not include any explanatory text outside the JSON."
```

SECTION 3 — The Adversarial Prompting Loop

Stress-Testing Your AI-Generated Strategies

The most powerful technique in advanced prompt engineering is subjecting your AI-generated outputs to adversarial review before acting on them. This is the professional standard; it is also how top AI teams build reliable pipelines.

The 3-Step Adversarial Loop

Step	Prompt Type	Purpose	Example
1. Generate	Initial prompt	Produce the first-draft solution	"Design a 30-day onboarding sequence for new B2B SaaS users."
2. Adversarial	Role-reversal prompt	Find weaknesses the author's bias would miss	"Act as a skeptical, time-poor user who hates automated emails. Identify 3 points where you would unsubscribe and why."
3. Refine	Synthesis prompt	Incorporate critique into an improved solution	"Based on that critique, rewrite the sequence to be 40% shorter and 100% focused on Time-to-Value in days 1-7."

Advanced: The Red Team Prompt

For high-stakes outputs (legal drafts, financial models, technical architectures), apply a full Red Team pass before using the output:

```
"You are a hostile expert reviewer with 20 years of experience in [domain]. You are deeply skeptical. Review the following output. Identify: 1) Any factual errors or questionable assumptions. 2) Any logical gaps or unstated dependencies. 3) Any sections where the advice could backfire in a realistic worst-case scenario. Score the output's reliability: Unacceptable / Needs Revision / Acceptable / Strong."
```

SECTION 4 — Prompt Debugging Decision Tree

Fix Any Prompt in Under 2 Minutes

When a prompt fails, use this decision tree to find the root cause and apply the correct fix immediately.

Symptom	Root Cause	Fix
Output is factually wrong	Model is hallucinating or lacks source context	Provide source documents. Increase reasoning_effort to medium/high. Ask model to "cite specific evidence for each claim."
Output is too long or verbose	No output constraint defined	Add explicit constraint: "Max 150 words." or "Output must fit in a Slack message (<500 chars)."
Output format is wrong or inconsistent	No output schema or example provided	Provide a "Few-Shot Example" showing the exact desired format. For programmatic use, require JSON schema output.
Model ignores explicit instructions	Critical instruction buried mid-prompt (Primacy Bias)	Move the most important instruction to the END of the prompt (Recency Bias is strong). Repeat it if necessary.
Output is generic and unspecific	Persona and context are too vague	Add company context, role, audience, and a negative example ("Do not produce output like: [bad example]").
Output varies wildly between runs	Prompt has ambiguity the model is resolving differently each time	Run the 10-Shot Consistency Test. Add constraints for every ambiguous dimension (tone, length, scope, perspective).

SECTION 5 — 10 Battle-Tested Prompt Templates

Copy-Paste Templates for Immediate Use

These 10 templates score 8+ on the TminusAI Prompt Scorecard. Fill in the bracketed fields and deploy immediately.

Template 1 — Strategic Analysis

```
"You are a [Senior Consultant / Domain Expert]. Analyze [topic/document] for a [audience]. Identify: 1) The 3 most significant risks. 2) The 2 highest-leverage opportunities. 3) One counterintuitive insight most people miss. Output as a structured brief, max 300 words. Do not use jargon."
```

Template 2 — Structured Research Brief

```
"You are a Research Analyst. Summarize the state of [topic] as of [year]. Cover: 1) Current consensus view. 2) The strongest dissenting argument. 3) The key open questions. 4) The 2-3 sources that matter most. Output as Markdown. Flag any claim you are less than 80% confident in."
```

Template 3 — Code Review

```
"You are a Senior [Python/Go/TypeScript] Engineer. Review the following code. Check for: 1) Security vulnerabilities. 2) Performance bottlenecks. 3) Logic errors. 4) Style violations. For each issue: [Issue] | [Severity: Low/Medium/High/Critical] | [Exact fix]. If the code is clean, say so "
```

Template 4 — Negotiation Prep

```
"You are a professional negotiator with 15 years in [industry]. I am preparing to negotiate [topic] with [counterpart]. My goal: [your goal]. Their likely goal: [their goal]. Prepare: 1) My opening position. 2) My BATNA. 3) 3 likely objections + rebuttals. 4) The single most important concession I should not make."
```

Template 5 — First-Principles Breakdown

```
"Break down [complex topic] to its first principles. Start with the most fundamental truth that cannot be disputed. Build up in no more than 5 logical steps to the practical implication. At each step, flag any assumption that has been historically questioned."
```

THE 10-SHOT CONSISTENCY TEST

Before deploying any prompt in an automated workflow, send it 10 times with identical inputs. If the output structure, length, or key conclusions vary significantly across runs, your prompt is under-specified. Keep adding constraints until you see <20% variance. This is non-negotiable for production pipelines.

GUIDE 4 COMPLETE — SYSTEMS SERIES

You have completed the TminusAI Systems Series. The path forward: combine the GPT-5 model knowledge (Guide 1), the OpenClaw agent stack (Guide 2), multimodal context (Guide 3), and prompt engineering rigor (Guide 4) into a unified AI operating system. Visit tminusai.com for the full Power User System and live workshops.