

Enoch Godson

PORTFOLIO

AI Product Thinker ▪ Prompt Engineer ▪ AI Evaluator

I study how AI thinks, where it fails, and how to make it more reliable.

About

Who I Am

I'm an AI product thinker focused on evaluation, prompt engineering, and building AI workflows that are trustworthy and useful. I spend most days testing models — ChatGPT, Claude, Gemini, Grok, Perplexity, Codex — comparing outputs, and turning what I learn into clearer prompts, better product ideas, and more dependable systems.

How I Think

I approach problems as systems: I look for root causes, break issues into parts, test assumptions, and iterate until the solution is simpler and more reliable. This mindset shapes how I evaluate AI models, design prompts, and think about product strategy — always working to make complex systems clearer and more dependable.

Focus Areas



AI Evaluation & Reliability

Testing models for reasoning quality, instruction-following, and hallucination detection through structured benchmark-style comparisons.



Prompt Engineering

Designing, optimizing, and iterating prompts to improve AI output quality and consistency across different models and workflows.



Product Strategy & UX

Shaping product ideas, feature design, and user experience thinking for AI-powered tools and systems.



Human-AI Trust & Safety

Exploring how users interact with AI, where trust breaks down, and how to design systems that are safer and more dependable.

Core Skills — AI

Prompt Engineering

Crafting and refining prompts to guide AI models toward more accurate, consistent, and useful outputs across a range of tasks and use cases.

Model Evaluation & Response Comparison

Systematically testing and comparing outputs across ChatGPT, Claude, Gemini, Grok, Perplexity, and Codex to understand strengths, weaknesses, and failure modes.

Hallucination Detection

Identifying when AI models generate confident but incorrect or fabricated information — a key part of making AI systems more reliable and trustworthy.

Benchmark-Style Testing & Prompt Iteration

Running structured tests against AI models and iterating on prompts to improve performance, consistency, and reliability over time.

Core Skills — Product & Tools

Product Skills

- **Product Planning** — Defining scope, priorities, and direction for AI-powered products and features.
- **Feature Design** — Shaping what gets built and how it works from a user-centered perspective.
- **UX Thinking** — Applying user experience principles to make AI interactions clearer and more intuitive.
- **Workflow Design** — Structuring how humans and AI work together efficiently and reliably.

Tools

- **Notion** — Documentation, project organization, and knowledge management
- **Lovable** — Building and prototyping AI-powered applications
- **Figma** — UI/UX design and prototyping (currently learning)
- **GitHub** — Version control and collaboration (currently learning)

Current Projects



Prompt Engineering Library

A collection of reusable prompts for common workflows — designed to save time and improve consistency when working with AI models. **Status: In Progress**



AI Evaluation Research

Model comparisons across reasoning, creativity, and reliability — testing ChatGPT, Claude, Gemini, Grok, Perplexity, and Codex to understand where each excels and where they fail. **Status: Ongoing**

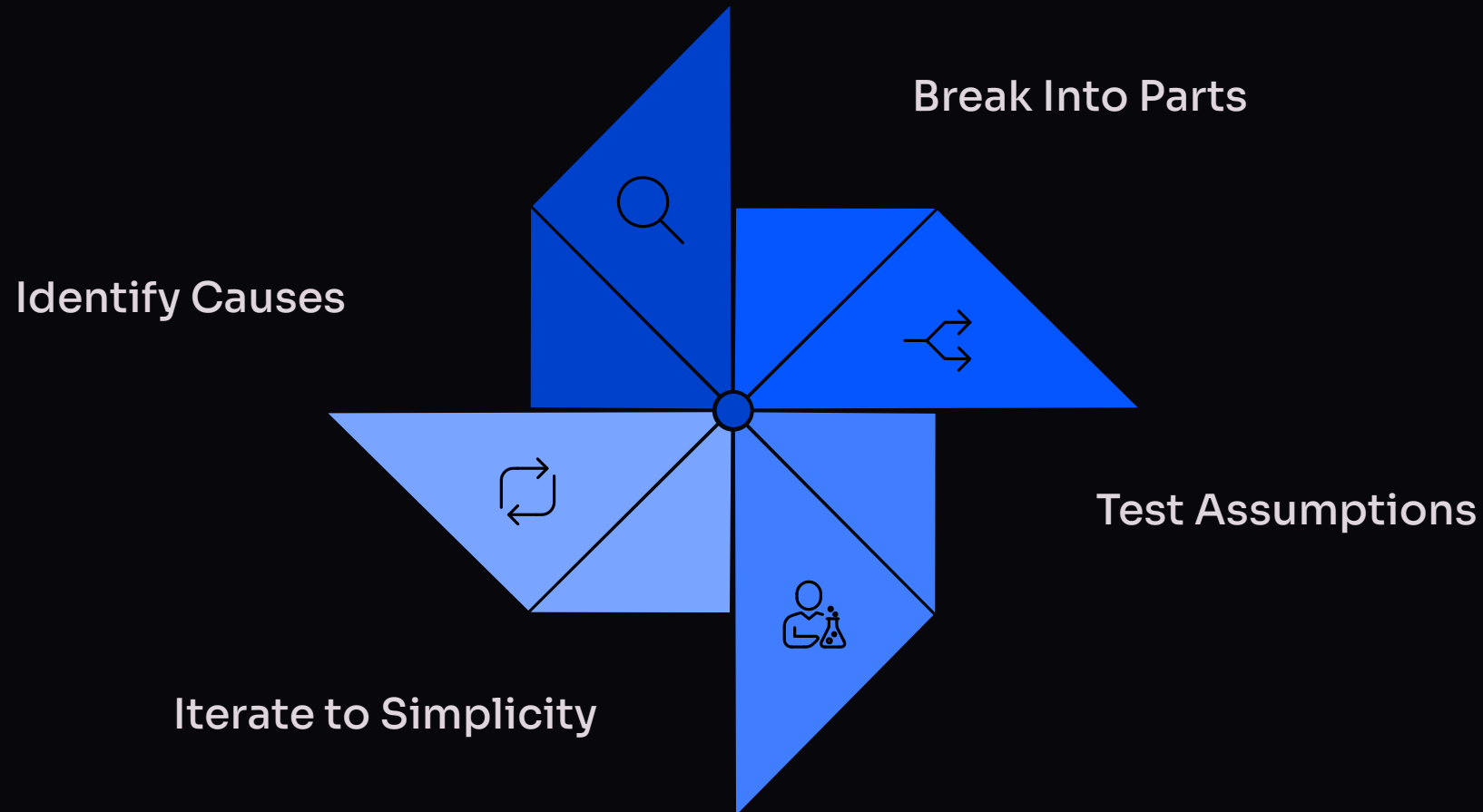


AI Product Case Studies

Product teardowns and improvement opportunities — analyzing existing AI products to identify what works, what doesn't, and how they could be better. **Status: Ongoing**

How I Work

My approach to problems is systematic: I look for root causes, break issues into parts, test assumptions, and iterate until the solution is simpler and more reliable. This applies to how I evaluate AI models, design prompts, and think about product strategy.



This cycle reflects how I approach AI evaluation and product thinking — always working to make complex systems clearer, more dependable, and easier to trust.

Models I Work With

I spend most days testing and comparing outputs across a range of AI models to understand their strengths, weaknesses, and failure modes. This hands-on evaluation informs my prompt engineering, product ideas, and workflow design.



ChatGPT



Claude



Gemini



Grok



Perplexity



Codex

By comparing outputs across these models, I identify patterns in reasoning quality, instruction-following, hallucination rates, and overall reliability — turning those findings into better prompts and more dependable systems.

What Drives Me

The Mission

Technology matters most when it solves real problems — and earns user trust. I'm driven by the belief that AI systems should be reliable, understandable, and genuinely useful to the people who depend on them.

The Approach

I turn what I learn from testing and comparing AI models into clearer prompts, better product ideas, and more dependable systems. Every evaluation, every prompt iteration, and every case study is aimed at making AI more trustworthy and useful in practice.

6+

AI Models Tested

ChatGPT, Claude, Gemini, Grok,
Perplexity, Codex

3

Active Projects

Prompt Library, Evaluation Research,
Case Studies

4

Focus Areas

Evaluation, Prompts, Product, Trust &
Safety

Get In Touch

Technology matters most when it solves real problems — and earns user trust.

Let's Connect

If you're interested in AI evaluation, prompt engineering, or building more trustworthy AI systems, I'd love to hear from you. Whether it's a collaboration, a question, or just a conversation about where AI is headed — reach out.

✉ **Email:** enochgodson75@gmail.com

Enoch Godson

AI Product Thinker

Prompt Engineer

AI Evaluator

I study how AI thinks, where it fails, and how to make it more reliable.