

# Prompt Engineering

Designing the system prompts, context and evaluations that make AI outputs reliable enough to ship

Course Duration ~50 Hours · 8 Weeks · Live Online · Workshop-Style · Portfolio Reviewed 1:1

System Prompts

Context Engineering

Structured Outputs

Evals

Prompt Security

**In short:** This is a live, workshop-style program in prompt *and* context engineering — designing the whole information environment a model sees, then proving quality with evaluations. It is not a course of clever phrasing tricks. Across 8 modules you ship versioned system prompts with evals attached, extraction prompts that turn messy documents into validated JSON, a golden eval set with a calibrated LLM-as-judge, a prompt-injection red-team exercise, and a reusable prompt library for your own field — reviewed 1:1 as a portfolio capstone. Built for fresh IT graduates who want an AI-ready skill before their first role and for working professionals in any domain. It is the entry point to our AI track: Generative AI & LLM Engineering is the broad builder foundation, and AI Agent Development is the specialist deep-dive into autonomous systems.

|                     |                                                                                                      |                   |                                                                                 |
|---------------------|------------------------------------------------------------------------------------------------------|-------------------|---------------------------------------------------------------------------------|
| <b>Duration</b>     | 8 weeks · ~50 hours · live + hands-on                                                                | <b>Format</b>     | Live online · workshop-style · portfolio reviewed 1:1                           |
| <b>Prerequisite</b> | None beyond comfort with everyday software. Python helpful for the optional API track, not required. | <b>Curriculum</b> | 8 modules, from how models read prompts to safety, security and applied domains |
| <b>Portfolio</b>    | 6 shipped artifacts plus a 1:1-reviewed capstone                                                     | <b>Ideal for</b>  | Fresh IT graduates and working professionals in any domain                      |

## What This Course Is Not

Techniques we deliberately skip, because they no longer earn their token cost on modern models:

- Role-play preambles, incantations and fake incentives
- Megaprompts that dump everything "just in case"
- Hard-coded few-shot examples for every call
- Jailbreak-style hacks

## Course Highlights

- Every module leads with what you build and ship, not what you "learn about"
- Prompt *and* context engineering: design everything the model sees, not just the wording
- Evaluation-first: golden test sets and LLM-as-judge before any optimization
- Structured outputs downstream systems can depend on — schemas, validation, repair loops
- Prompts treated as production software: versioned, reviewed, regression-tested
- Prompt injection defense and a red-team exercise against your own system prompt
- Applied patterns for coding, analysis, writing, support, research and agent instructions

## Hands-on Projects & Portfolio

- **Versioned System Prompt** — a production system prompt with a changelog and an eval attached to every revision
- **Structured Extraction Prompt** — reliably turns messy documents into validated JSON
- **Golden Eval Set + Judge Harness** — scores prompt variants automatically with a calibrated LLM-as-judge
- **Optimization Case Study** — a before/after on a real workflow, with measured improvement
- **Injection Red-Team Exercise** — attack your own system prompt, then ship the fixes
- **Domain Prompt Library** — a reusable, domain-specific library for your own field or job function

## Tools & Stack

ChatGPT · Anthropic Claude · Google Gemini · Claude Code · OpenAI Playground · Anthropic Console · LangSmith · Langfuse · Promptfoo · DSPy · Pydantic · JSON Schema · Git · Python (optional track)

## Career Outcomes

An honest framing: prompt engineering today is mostly a force multiplier inside an existing role rather than a standalone job title. The strongest career position is prompting plus a durable specialty — and this course is built to combine with the domain you already work in. Where the skill pays:

- AI Engineer · Applied AI Engineer · AI Product Manager · AI Automation / Workflow Specialist
- AI Solutions Consultant · Content & Knowledge Systems Specialist
- Domain expert applying AI in healthcare, legal, finance, support or education

## Who This Course Is For

- Final-year students and fresh IT graduates who want an AI-ready skill before their first role
- Software engineers and QA professionals getting more out of AI coding tools
- Data analysts, marketers, HR, legal, finance and support professionals automating real workflows
- Product managers and team leads writing AI feature specs and evaluating AI vendors
- Anyone who has been let down by inconsistent AI output and wants a repeatable method

## Module 1: How Models Actually Read Your Prompt

6 Hours

### YOU SHIP

- A failure autopsy on a real broken prompt — traced to tokens, position, sampling or context rot — plus the fixed version, tested across GPT, Claude and Gemini

### TOPICS

- Tokens, context windows, attention and why position matters; temperature and sampling
- Context rot: why more input often means worse output
- Model-specific behaviour across GPT, Claude and Gemini; how reasoning models change what you should and shouldn't spell out

**Assessment:** Quiz + Workshop

## Module 2: The Anatomy of a Reliable Prompt

6 Hours

### YOU SHIP

- A production-grade system prompt for a real task — role, constraints, success criteria, escape hatches and edge-case handling, structured with XML tags and numbered sections

### TOPICS

- System prompt vs user message vs injected context; role and constraint design
- Task decomposition; explicit success criteria; escape hatches for when the model doesn't know
- Refusal and edge-case handling; XML tags and numbered sections for structure

**Assessment:** Quiz + Workshop

## Module 3: Core Techniques That Still Earn Their Keep

6 Hours

### YOU SHIP

- A technique decision card for your own workflow — which of zero-shot, few-shot, chain-of-thought, self-consistency or decomposition actually wins on your task, with cache-friendly example placement

### TOPICS

- Zero-shot vs few-shot and when each wins; examples chosen for format rather than knowledge

- Chain-of-thought and its limits on reasoning models; self-consistency; reflection and self-critique; decomposition chains
- Prompt caching and where examples must sit for caching to work

**Assessment:** Quiz + Workshop

## Module 4: Structured Outputs & Reliable Formats

6 Hours

### YOU SHIP

- An extraction prompt that reliably turns messy real-world documents into validated JSON, with a repair loop for the failures

### TOPICS

- JSON schemas, enums and constrained generation
- Validation and repair loops; extraction prompts that survive messy real-world input
- Designing output contracts a downstream system can depend on

**Assessment:** Quiz + Workshop

## Module 5: Context Engineering

7 Hours

### YOU SHIP

- A context plan for a multi-step workflow — what lands in the window at each step, what deliberately stays out, with citation-forcing and a token budget

### TOPICS

- Deciding what lands in the context window at each step — and what you deliberately keep out
- Retrieval-injected context and just-in-time examples; memory and conversation compaction
- Document and table handling; citation-forcing; token budgeting as a design constraint

**Assessment:** Quiz + Workshop

## Module 6: Evaluation — Proving a Prompt Is Better

7 Hours

### YOU SHIP

- A golden eval set and LLM-as-judge harness that scores prompt variants automatically — plus a before/after optimization case study with measured improvement

### TOPICS

- Building a golden test set before you optimize; pass/fail criteria
- LLM-as-judge and calibrating the judge; A/B comparison; regression testing so a "small tweak" can't silently break yesterday's behaviour
- Automated prompt optimization with DSPy-style tooling and where it beats a human

**Assessment:** Quiz + Workshop

## Module 7: Prompts as Production Software

6 Hours

### YOU SHIP

- A versioned prompt repository — templates, variables, changelog, an eval result attached to every change, and a prompt-review checklist your team can adopt

### TOPICS

- Version control and changelogs for prompts; an eval result attached to every change
- Prompt templates and variables; environment and model migration
- Cost and latency trade-offs; model routing; reviewing a prompt the way you'd review code

**Assessment:** Quiz + Workshop

**Module 8: Safety, Security & Applied Domains**

6 Hours

**YOU SHIP**

- A prompt-injection red-team exercise against your own system prompt plus the fixes — and your portfolio capstone: a reusable prompt library for your own field, defended in a 1:1 review

**TOPICS**

- Prompt injection and indirect injection; jailbreak resistance; PII handling; bias and tone control; guardrails and output validation
- Applied prompt patterns for coding, data analysis, business writing, customer support and research
- Instructions for agents and AI coding tools

**Assessment:** Portfolio Capstone + 1:1 Review

www.progmatrix.com