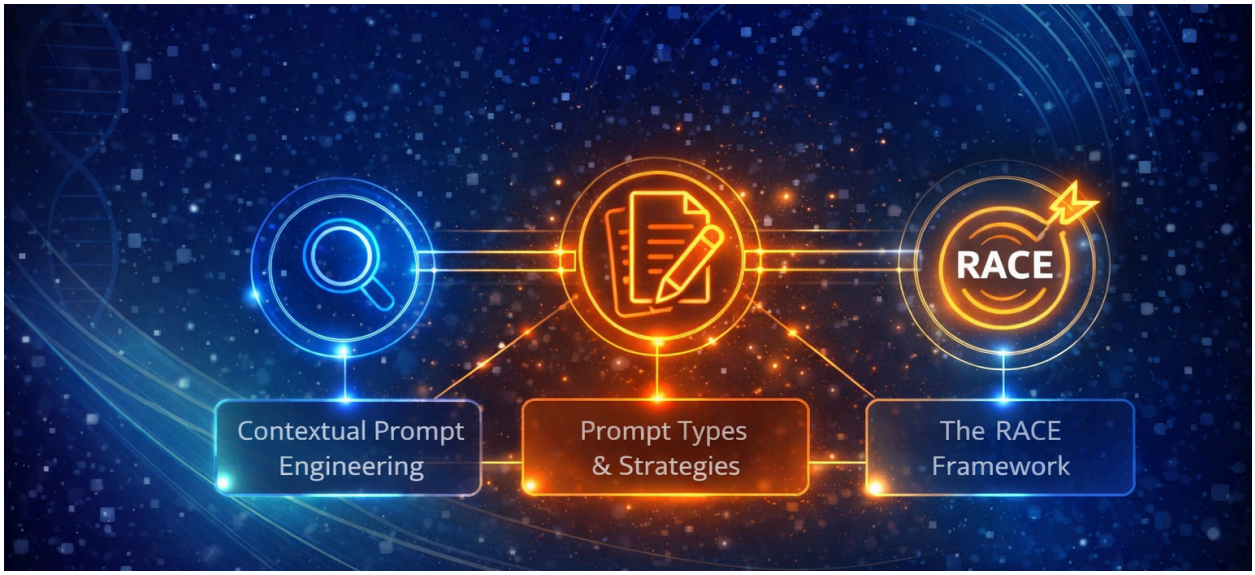
Module 5: Interacting with AI - Prompts and Process



Purpose: To teach individuals how to structure effective context for high-quality, tailored responses.

Topics/Learning Objectives
Upon completion of this module, individuals will have learned:

- Introduction to Contextual Prompt Engineering
- Prompt Types: Single-Shot, Few-Shot, Chain-of-Thought
- The RACE Framework for Effective Prompting
- Prompt Iteration for Refining Output

Note: The text and graphics in these modules were co-developed with the assistance of generative AI tools such as OpenAI's ChatGPT, Google's Gemini and NotebookLM and Microsoft's CoPilot, drawing on the indicated reference materials. Materials were then edited for relevancy and accuracy.

Topic 1: Introduction to Contextual Prompting

Introduction to Contextual Prompting (Context Engineering)

Large Language Models (LLMs) like Gemini, ChatGPT and others are powerful tools, but their utility in medicine and medical education hinges on a critical skill: **contextual prompting**. Think of it not just as asking a question, but as conducting a high-quality, information-rich consultation with a specialist.

Prompt Engineering vs. Context Engineering

You may already be familiar with *prompt engineering*: crafting specific queries or instructions for the AI to generate a desired output.

Example Prompt: "What are the causes of chest pain?"

The AI will respond, but likely with a generic list that is not relevant to your case or learning goal.

This is where *context engineering* becomes essential. Context engineering is the art and science of embedding the prompt within a well-developed clinical or academic scenario. It includes not only the question but also the background, intent, key constraints and expected output. It's like providing a detailed case file to a consultant—you are guiding the AI toward focused, meaningful thinking.

But… Critical Thinking First!

Before using AI, take a moment to pause and reason through the problem yourself.

Ask: What do I already know? What's my initial hypothesis? What would I do next if I couldn't ask for help?

The goal is to strengthen your clinical judgment, not bypass it. AI can help test your reasoning, challenge your assumptions, or fill gaps—but it should not replace your own analysis.

Treat each case as a learning opportunity to practice decision-making. Let AI help you develop or refine your thinking, not define it.

Example: A Practical Clinical Prompting Framework

To practice context engineering effectively, you can use a structured approach to prompting, similar to how you would prepare for case-based discussions or structured oral exams.

Here's a four-part framework to help you create a contextual prompt:

---

1. *Clinical Scenario and Goal*

- What is the specific situation you're analyzing?
- What is your role (e.g., learner, explainer, researcher)?
- What is your objective—differential diagnosis, mechanisms, exam questions?

Ethics Reminder: Responsible Use of AI with Case Information

When using AI tools to support your case-based learning, always state clearly in your prompt that the case is simulated and for educational purposes only. This models professional communication and helps prevent the AI from interpreting the scenario as real clinical care.

Example statement to include in your prompt:

"This is a fictional case created for educational purposes. No real patient data is included."

Never enter, upload or share any real patient information—whether in text, images or documents—in any AI tool. This includes de-identified data unless explicitly permitted by institutional policy. Doing so may violate HIPAA regulations and school guidelines on data privacy and academic integrity.

Using AI responsibly means treating it as a tool for learning—not for clinical documentation or patient care.

Example:
“I’m a medical school faculty member preparing for a case-based learning session. I need help generating and organizing a differential diagnosis for a patient with shortness of breath. This is a fictional case created for educational purposes. No real patient data is included."

---

2. Relevant Case Information

- What are the key features of the patient case?
- Include pertinent positives and negatives.
- Stick to de-identified, realistic data if referencing real or simulated cases.

Example:
“The case involves a 68-year-old man with known COPD and a 40-pack-year smoking history. He presents with two days of worsening dyspnea and a productive cough. Vitals: HR 110, RR 24, BP 145/90, SpO2 89% on room air. He has wheezing, no chest pain or leg swelling.”

---

3. Intended Output and Audience

- What do you want the AI to do?
- Who is the output for? (Yourself? A classmate? A simulated teaching context?)

Example:
“Generate a ranked list of possible diagnoses with rationales and suggest the first few steps of diagnostic work-up, appropriate for a student-led discussion group.”

---

4. Tone, Role, and Constraints

- Ask the AI to take on a role (e.g., expert educator, medical tutor).
- Set boundaries (e.g., “Don’t recommend treatment,” “Limit to 5 suggestions,” “Use plain language”).

Example:
“Respond as a clinical reasoning coach for second-year students. Focus on the diagnostic process and don’t include treatment recommendations.”

Putting It All Together: A Sample Contextual Prompt

> "I’m a medical school faculty member preparing for a case-based learning session. I need help generating and organizing a differential diagnosis for a patient with shortness of breath. This is a fictional case created for educational purposes. No real patient data is included. The case involves a 68-year-old man with known COPD and a 40-pack-year smoking history. He presents with two days of worsening dyspnea and a productive cough. Vitals: HR 110, RR 24, BP 145/90, SpO2 89% on room air. He has wheezing, no chest pain, or leg swelling. Generate a ranked list of possible diagnoses with rationales and suggest the first few steps of diagnostic work-up, appropriate for a student-led discussion group. Respond as a clinical reasoning coach for second-year students. Focus on the diagnostic process and don’t include treatment recommendations."

This prompt is more likely to generate a useful and targeted response than a generic question—and will model the kind of clinical thinking you’ll need to build a case.

Why Contextual Prompting Matters in Medical Education

Patient Safety
The more context you give, the less likely AI is to make reasoning errors.

Efficient Learning
AI can serve as a quick-access curriculum development helper, tutor or study partner—but only when it knows what you're actually working on.

Critical Thinking
Crafting a well-contextualized prompt is itself a thinking exercise. You must prioritize, select and clarify—just like in critical thinking and clinical reasoning.

Topic 2: Single Shot, Few Shot, Chain of Thought

Single-Shot or Zero-Shot Prompting

Single-shot or zero-shot prompting is a style where you provide the AI with **one clear instruction or question**, **without example inputs or outputs**. You rely on the model's prior training and generalization to generate a response.

- **Why use it:** Efficient for retrieving factual knowledge, definitions or straightforward reasoning—especially useful during quick review or fact-checking.

Example Prompt: "Define the functional zones of the cerebral cortex."

This style works well when the task is well-scoped and doesn't require the AI to infer an expected format.

Few-Shot or Multishot Prompting

Few-shot prompting means showing the AI a few examples first, so it can learn the **pattern** you want it to follow. Each example includes both the input and the correct output. This helps guide the AI's response when tasks involve classification, reasoning, or formatting. You can use text samples, rubrics, templates, images, etc. For example, if you would like it to render a practice test question in the NBOME format, then provide an example question.

- **Why use it:** Improves accuracy when you're asking the model to perform a **specific task with subtle distinctions**—like categorizing, translating formats or mimicking a teaching style.

Example Prompt:

> **"Classify each sentence as a Diagnosis, Prognosis or Treatment.**
> Example 1: The patient has pneumonia. → Diagnosis
> Example 2: The patient is expected to recover fully. → Prognosis
> Example 3: The patient was started on azithromycin. → Treatment
>
> **Now classify: The patient reports chronic joint pain."**

➤ **Likely AI output:** Diagnosis

Chain-of-Thought Prompting: Think Step-by-Step

What it is:
With **Chain-of-Thought (CoT) prompting**, you ask the AI to **reason through a problem step-by-step**, mimicking how clinicians think through differential diagnoses or complex cases.

Best for:
Multi-step clinical reasoning tasks, exam-type questions, diagnostic analysis and ethically nuanced scenarios where you want the AI to lay out its thought process.

Example Prompt (Medical Scenario):

> **"Think step-by-step: A patient presents with hemisection at T_{10}. What is the most likely condition, and how would you explain the resulting motor and sensory deficits?"**

Expected AI Output Structure:

1. Explain spinal cord hemisection at T_{10}
2. Describe ipsilateral motor weakness below the level
3. Describe contralateral pain/temperature loss
4. Name the condition: *Brown-Séquard Syndrome*

Why it's Effective:
CoT prompting lets the model simulate clinical reasoning by breaking complex problems into smaller analytic steps, improving transparency and accuracy in multi-step tasks.

In medicine, empirical work shows that CoT-style or diagnostic reasoning prompts allow models to more faithfully mimic clinician thinking than simpler prompts—without sacrificing diagnostic accuracy.

Topic 3: The RACE framework for effective prompting

RACE Prompt Framework: Setting the Stage for High-Quality AI Output

Let's discuss a simple prompting framework called RACE.

RACE stands for:

- **Role**: Define the professional persona or expertise you want the AI to embody.
- **Action**: Specify the task you want the AI to perform—be explicit and precise.

- **Context**: Provide only the relevant background information or scenario.
- **Expectations (Execute)**: Clarify the desired output format, quality criteria and style.

This structured framework teaches the model to produce high-quality, domain-specific responses by anchoring it to a clearly defined role and task within a specific context and format.

RACE is an especially helpful framework to remember because it starts by assigning the AI assistant a specific role. Assigning a role is important because it tells the model *who* it should act like: a board-certified specialist, an anatomy professor, a USMLE question-writer or a peer tutor. When you specify the role, you're anchoring the response's tone, depth and perspective.

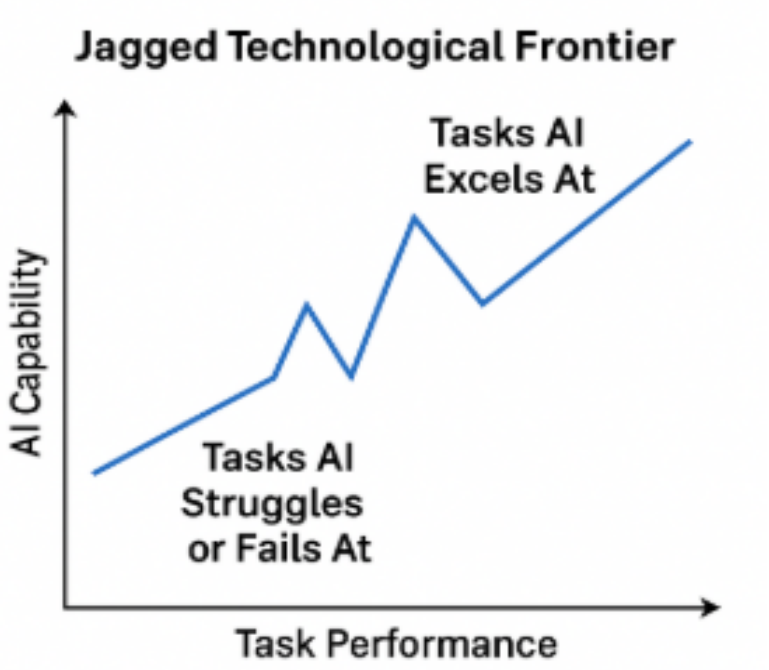
Think of it this way: asking, "Explain Bell's palsy" will get a very different result than:

> *"You are a neurologist teaching second-year medical students. Explain Bell's palsy using clinical examples and anatomical reasoning."*

Sample Prompt Using RACE in Medical Education:

> **"As a pediatric neurologist (Role), review this EEG report and provide a differential diagnosis (Action). The patient is a 5-year-old with recurrent seizures and focal slowing in the left temporal region (Context). Output a prioritized list of possible diagnoses with brief supporting rationale, formatted as bullet points with severity estimates (Expectations).**

Topic 4: Prompt Iteration for Refining Output



Prompt Iteration: It's Normal, and It's Necessary

Even experienced educators don't get perfect AI responses on the first try—and that's okay. **Prompt iteration** is the natural, expected process of refining your instructions to get more accurate, useful or relevant results.

Think of it like writing a test question, giving clinical feedback, or revising a lecture: the first draft is a starting point—not the final product.

Sometimes you'll need to:

- Clarify the **level** of complexity (e.g., pre-clinical vs. board review)
- Specify the **output format** or role
- Add or remove **details** in the context
- Use a framework like **RACE** to restructure the prompt

Prompting is not a one-and-done task, it's a **dialogue** between you and the model. Iteration is where your discernment as an educator comes in.

Flip the Script: Ask AI to Help You Create the Prompt

Prompting isn't always about knowing what to say—it's about **knowing how to shape the interaction**. If you're unsure how to get started, you can ask the AI itself:

> *“Help me write a clear and specific prompt to generate a case-based MCQ about cardiac tamponade.”*

This meta-level use of AI helps you:

- Learn **prompt design strategies**
- See examples of effective **structure and scope**
- Reduce trial-and-error early in the process

In other words: You can prompt the AI to help you prompt better. This is a collaborative, co-intelligent process. Sometimes it's more efficient because AI does not think like people when it comes to prompt architecture for generating output.

Example

Prompt:

> **“Can you help me write a well-structured prompt that uses the RACE framework to generate an NBOME-style question on Brown-Séquard Syndrome for second-year students?”**

AI Output:

“Sure. Here's a draft:
***Role*: You are a medical educator preparing board-style questions.**
***Action*: Write a multiple-choice question on Brown-Séquard Syndrome.**
***Context*: The patient has a T10 spinal cord hemisection. Include clinical findings.**
***Expectations/Output format*: Five plausible options, with rationales for the correct and incorrect choices.”**

This approach supports self-regulated learning and shows students that *prompting is a creative and iterative process*—one where AI can also be a design partner, not just a responder.

AI Can Only Do Well What It Was Designed to Do

When AI seems to "fail," it’s not always because of a bad prompt or user error. Sometimes it’s because the task you’re asking it to do falls **outside its core strengths**.

This is known as the **Jagged Technological Frontier** (Dell’Acqua et al., 2023):

AI excels at some tasks, struggles with others and completely fails at a few—all within the same domain. It's not a smooth curve of competence. It's jagged.

For example:

- AI can write high-quality multiple-choice questions—**but** may repeat, forget formatting rules or miss cognitive alignment without oversight.
- It can generate long reports or logs—**but** it might lose consistency, skip sections or time out when processing large outputs.
- It may promise to "bundle," "export" or "format" content—**but** the actual performance depends on the limitations of the interface and backend (e.g., Copilot, ChatGPT, Claude).

What to Do Instead:

When this happens, don't fight the AI—adapt your approach:

- Break big tasks into smaller steps that the model can handle reliably.
- Save drafts manually at critical junctures.
- Ask the AI to write the prompt you can use to re-import, reformat or revise your content later.
- Use external tools (like Word or OneNote) to preserve structure and formatting when output errors persist.

Key Takeaway:

AI is a set of trained systems with strengths and limits. Your job is to understand where the edges are jagged and steer thoughtfully around them.

Supplementary Materials & Resources

- **Supplementary Materials**
  - PowerPoint
  - Flashcards

- Google NotebookLM or GEM

- **Resources**
  - Articles
    - Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality (Harvard Business School Working Paper No. 24-013). Harvard Business School Technology & Operations Management Unit. Available at SSRN: https://ssrn.com/abstract=4573321
    - Savage, T., Nayak, A., Gallo, R., Rangan, E., & Chen, J. H. (2024). Diagnostic reasoning prompts reveal the potential for large language model interpretability in medicine. NPJ Digital Medicine, 7(1), 20. https://doi.org/10.1038/s41746-024-01010-1
    - Sivarajkumar, S., Kelley, M., Samolyk-Mazzanti, A., Visweswaran, S., & Wang, Y. (2024). An empirical evaluation of prompting strategies for large language models in zero-shot clinical natural language processing: Algorithm development and validation study. JMIR Medical Informatics, 12, e55318. https://doi.org/10.2196/55318 medinform.jmir.org+11valueinhealthjournal.com+11arthroscopyjournal.org+11.
    - Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E. H., Le, Q. V. & Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. arXiv preprint arXiv:2201.11903. https://doi.org/10.48550/arXiv.2201.11903 arxiv.org+4arxiv.org+4home.cse.ust.hk+4Videos
  - Videos
    - Google Career Certificates
      - A Simple Framework for Writing Effective Prompts
    - James Briggs
      - Prompt Engineering: Zero, One, and Few-Shot Prompting
  - Websites
    - Anthropic. (2024).
      - Prompt engineering overview: Multi-shot prompting.
    - Acronymat. (2024).
      - RACE Prompt Framework: Create Better Prompts with Role, Action, Context, and Expectations.
    - Learn Prompting
      - Prompt Engineering Guide
    - Medium (2026)

- Mastering LLM Prompting: The RACE Framework for Actually Getting What You Want