



AI ENGINEERING ACADEMY

Prompt Engineering na Prática

Como Conversar Melhor com LLMs e
transformar interações em resultados de alta
performance

AULA 01 • 60 MINUTOS • INICIANTE/INTERMEDIÁRIO

Instructional
Prompts

Prompt
Engineering
Techniques

Control
Codes

Contextual
Prompts

O QUE VAMOS APRENDER HOJE?

OBJETIVO: Capacitar você a criar prompts claros, eficientes e reutilizáveis para dominar qualquer LLM.

- CHIP Fundamentos e limitações dos LLMs
- </> Estrutura de um prompt de alta performance
- ⚡ Técnicas avançadas (Zero-shot, Few-shot, CoT)
- ⟳ Ciclo de iteração e avaliação
- ⚠️ Introdução ao RAG

Prompt Engineering Techniques



NÃO É MÁGICA, É COMUNICAÇÃO

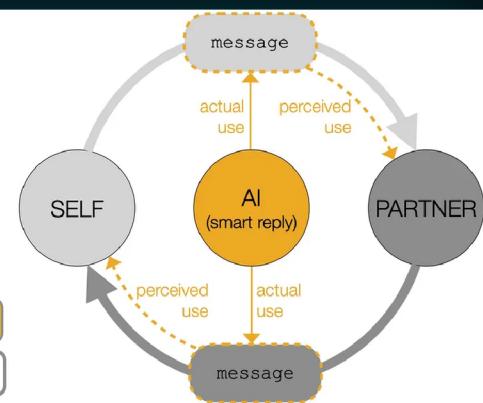
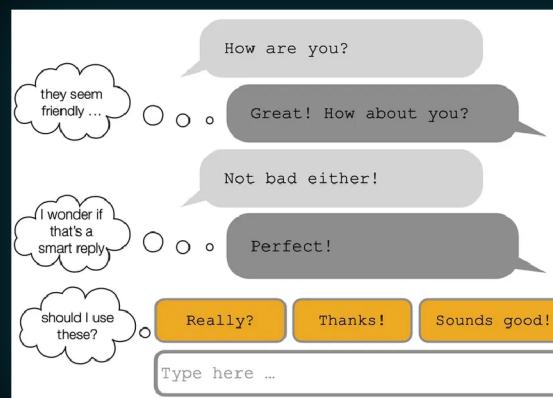
TRADUÇÃO DE INTENÇÕES

Traduzir objetivos humanos em instruções claras para a máquina.

INTERFACE DE CONTROLE

O prompt é o volante: controla comportamento e qualidade da IA.

"Se você controla o prompt, você controla o resultado."



CONHEÇA SEU "ESTAGIÁRIO": LIMITAÇÕES DOS LLMS

▀ Memória e Contexto

Modelos possuem uma "Janela de Contexto" limitada. Se o prompt for longo demais, a IA começa a "esquecer" as instruções iniciais.

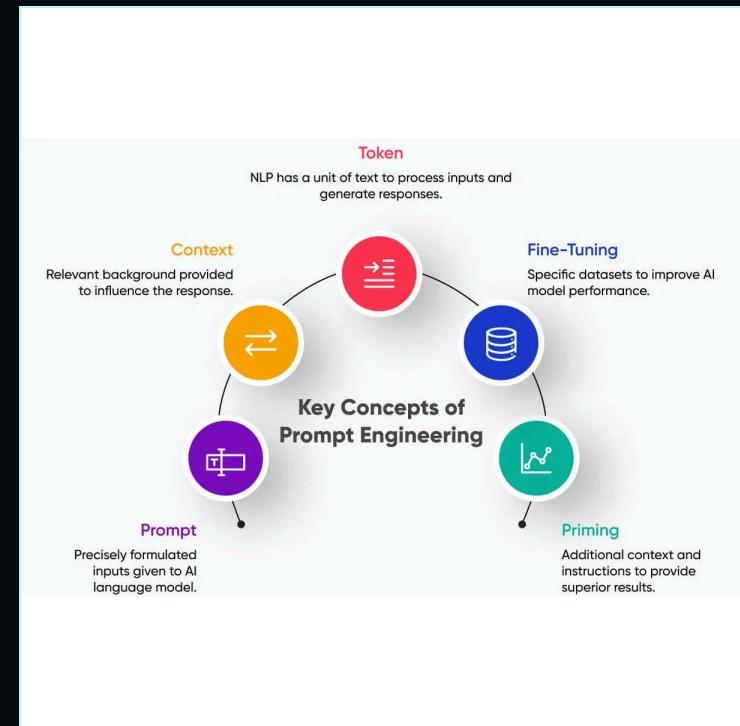
▀ Natureza Estatística

LLMs não "pensam" como humanos. Eles são motores de probabilidade. Instruções claras reduzem a variabilidade indesejada.

👻 Alucinação

A IA não sabe o que é verdade; ela prevê o próximo token. Se o prompt for ambíguo, ela pode inventar fatos com total confiança.

INSIGHT: INSTRUÇÕES CLARAS REDUZEM DRASTICAMENTE O RISCO DE ERROS.



A ANATOMIA DE UM PROMPT PERFEITO

Persona + Contexto + Tarefa + Formato + Restrições

1. PERSONA (ROLE)

Quem o modelo deve ser? Define o tom e a expertise.

Ex: "Você é um Engenheiro de Software."

2. CONTEXTO

Qual a situação e o público-alvo?

Ex: "Explicando para desenvolvedores iniciantes."

3. TAREFA (TASK)

O que exatamente deve ser feito?

Ex: "Resuma este artigo técnico."

4. FORMATO (OUTPUT)

Como a resposta deve ser estruturada?

Ex: "Tabela", "JSON", "Lista de tópicos".

5. RESTRIÇÕES

O que evitar ou quais limites seguir?

Ex: "Máximo de 200 palavras", "Não use jargões", "Responda em Português".

DO RUIM AO EXCELENTE

✗ PROMPT VAGO

"Explique RAG."

PROBLEMAS:

- Falta de contexto
- Sem definição de público
- Resultado genérico e imprevisível

✓ PROMPT ESTRUTURADO

"Você é um engenheiro de IA.
Explique o conceito de RAG para
desenvolvedores iniciantes, usando
uma analogia simples, em até 200
palavras."

DIFERENCIAIS:

- Persona definida (Engenheiro)
- Público claro (Iniciantes)
- Formato e restrição (Analogia, 200 palavras)

MENOS É MAIS: DECOMPOSIÇÃO

✖ O PROBLEMA: PROMPT "FRANKENSTEIN"

"Analise esse texto, melhore a gramática, gere um resumo de 3 parágrafos, crie 5 títulos criativos e explique o racional por trás de cada alteração feita."

Prompts sobrecarregados aumentam a carga cognitiva do modelo e diluem a atenção, resultando em saídas genéricas ou incompletas.

✓ A SOLUÇÃO: DIVIDIR PARA CONQUISTAR

1 Prompt 1: Revisão

Focar exclusivamente em clareza e correção gramatical.

2 Prompt 2: Resumo

Gerar a síntese baseada no texto já revisado.

3 Prompt 3: Títulos

Criar opções criativas a partir do resumo final.

REGRA DE OURO: UMA TAREFA POR PROMPT.

ZERO-SHOT VS. FEW-SHOT PROMPTING

ZERO-SHOT

Apenas a instrução direta, sem exemplos prévios. Ideal para tarefas simples e modelos potentes.

PROMPT:

"Classifique o sentimento do seguinte texto como Positivo ou Negativo:

'O atendimento foi excelente e rápido!'"

FEW-SHOT

Fornece 1 a 3 exemplos do padrão desejado. Ensina o modelo a seguir um formato ou estilo específico.

PROMPT:

"Texto: 'Horrível' -> Negativo
Texto: 'Adorei' -> Positivo

Texto: 'Poderia ser melhor' -> Negativo
Texto: 'Incrível' -> "

Insight: Exemplos são a forma mais rápida de "ensinar" um padrão de resposta complexo ao modelo sem precisar de fine-tuning.

CHAIN-OF-THOUGHT (COT): ESTIMULANDO O RACIOCÍNIO

O QUE É?

Técnica que solicita ao modelo que explice seu processo de raciocínio antes de fornecer a resposta final.

COMANDO CHAVE:

"Pense passo a passo antes de responder."

Nota: Modelos de raciocínio modernos (como o o1) já executam esse processo nativamente.



Redução de Erros Lógicos

Ao decompor o problema, o modelo evita saltos intuitivos que levam a conclusões erradas.



Transparência

Permite ao usuário auditar o "pensamento" da IA e identificar onde ocorreu uma falha.



Complexidade

Ideal para tarefas de matemática, lógica, programação e análise de dados profunda.

O CICLO PROFISSIONAL: ITERAÇÃO E AVALIAÇÃO

01

CRIAR

Escreva a primeira versão do prompt usando o framework estruturado.

02

AVALIAR

A saída atende aos requisitos de clareza, tom e precisão factual?

03

AJUSTAR

Refine as instruções, adicione restrições ou melhore o contexto.

04

REAVALIAR

Teste novamente e repita o ciclo até atingir a performance ideal.

DICA PRO

Use um **segundo LLM** para avaliar a qualidade da resposta do primeiro. Peça para ele analisar clareza, precisão e aderência às restrições impostas.

PARÂMETROS TÉCNICOS E JANELA DE CONTEXTO

🌡 TEMPERATURE

Controla a aleatoriedade.

0.0 - 0.3: Fatos, código, precisão.

0.7 - 1.0: Criatividade, marketing, ficção.

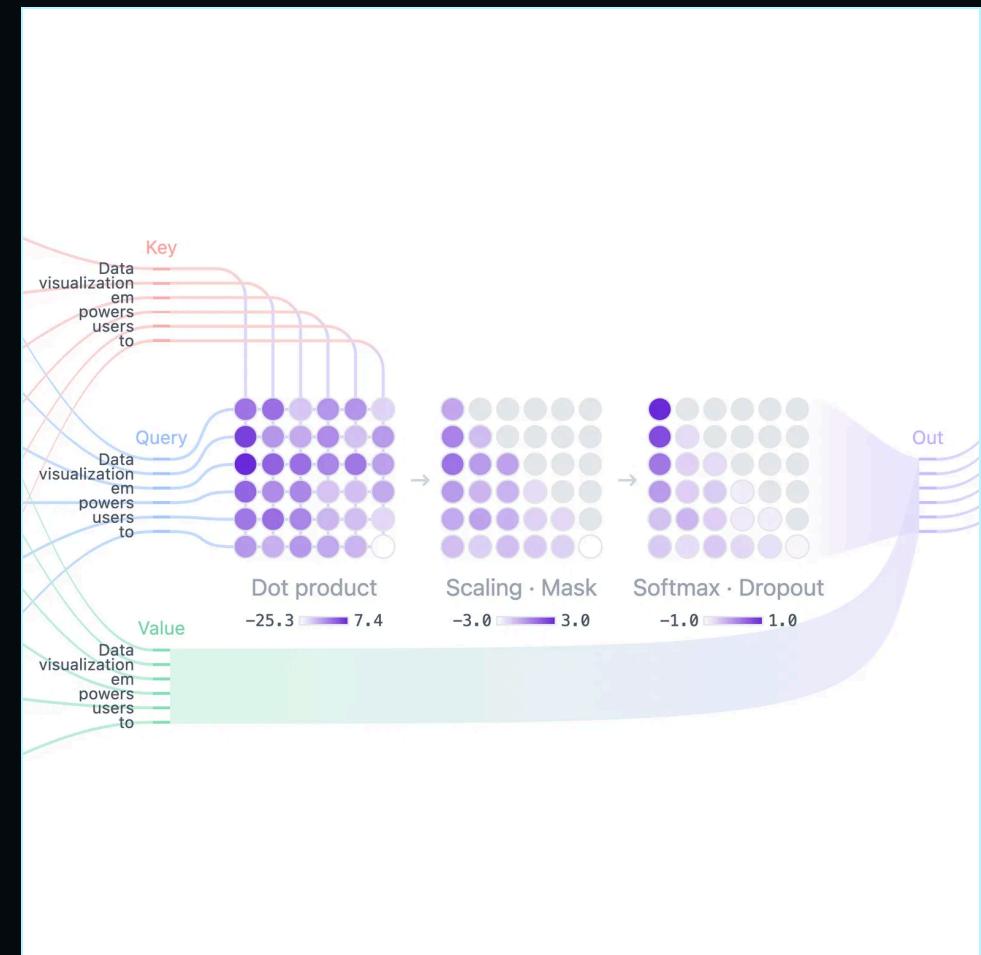
▼ TOP-P / TOP-K

Filtram o vocabulário do modelo. O **Top-p** limita a escolha aos tokens que somam uma probabilidade acumulada específica, garantindo diversidade controlada.

▣ JANELA DE CONTEXTO

O limite de "tokens" (pedaços de palavras) que o modelo processa por vez. Se excedido, o modelo "esquece" o início da conversa.

Atenção: Prompts excessivamente longos diluem a atenção do modelo (Lost in the Middle) e aumentam o custo de processamento.



O PRÓXIMO NÍVEL: INTRODUÇÃO AO RAG

CONHECIMENTO DINÂMICO

A IA busca informações em documentos externos em tempo real, superando a data de corte do treinamento.

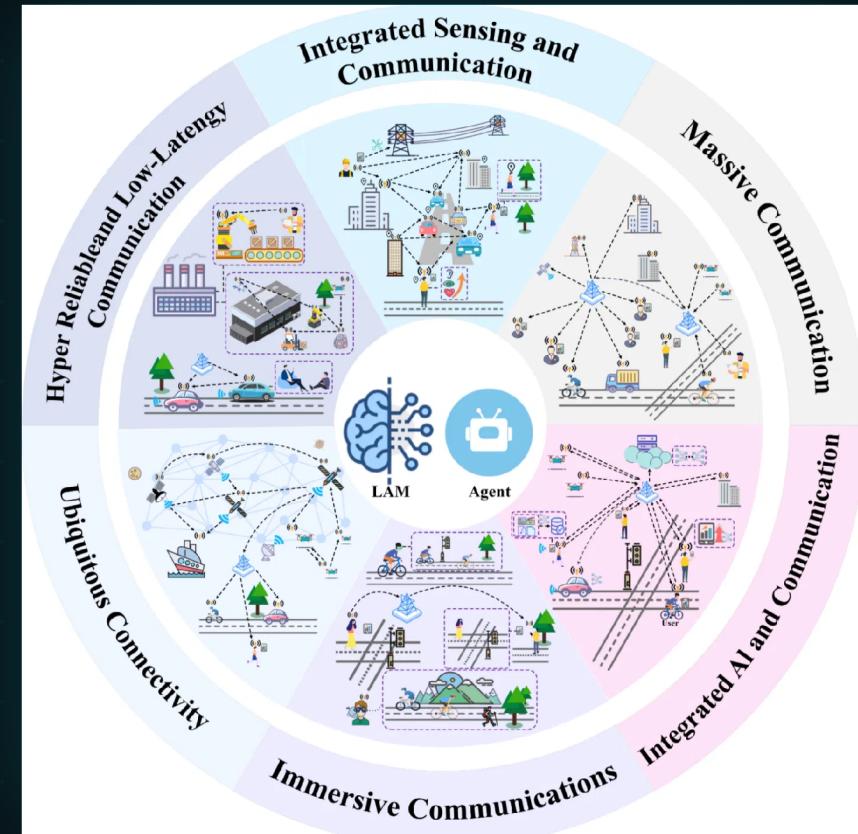
FACTUALIDADE E FONTES

Reduz drasticamente alucinações ao basear as respostas em evidências recuperadas e citar fontes.

ESCALABILIDADE

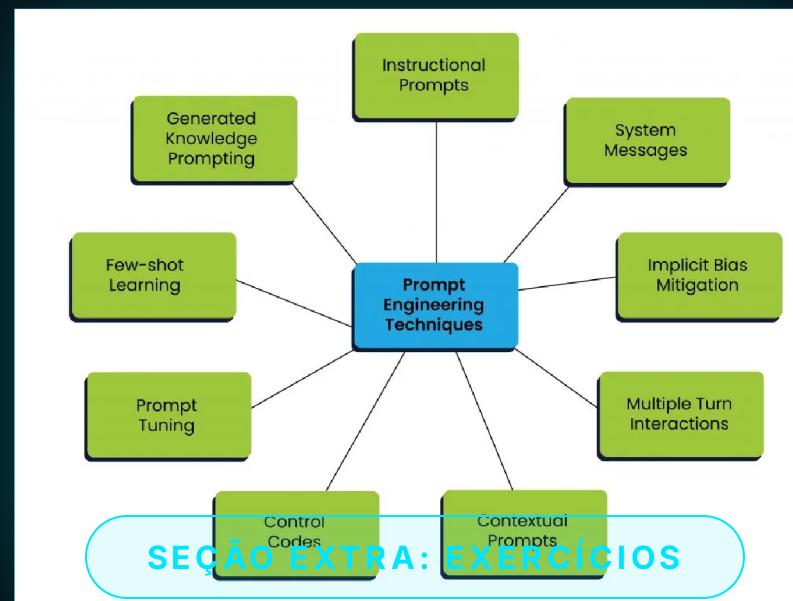
Permite processar volumes massivos de dados que não caberiam em uma janela de contexto padrão.

Prompt Engineering escala até certo ponto.
RAG leva você além.



ZONA DE PRÁTICA

Do Conceito à Execução: Consolidando Frameworks e Técnicas de Prompt Engineering.



DESAFIO 1: O ARQUITETO DE PERSONAS

CENÁRIO

Você deve criar um assistente para uma empresa de advocacia que resume contratos complexos para clientes leigos.

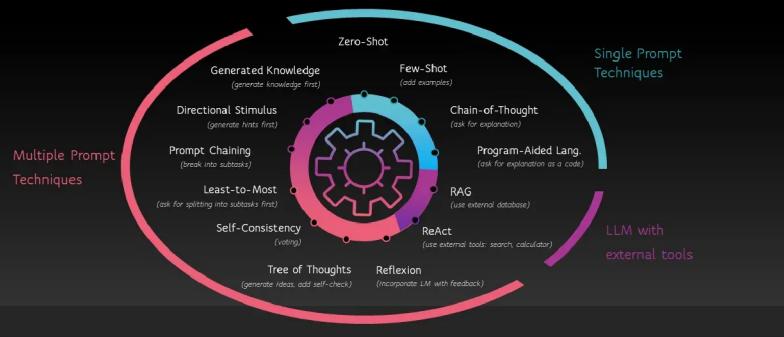
TAREFA

Escreva um prompt estruturado aplicando o framework completo:

Persona + Contexto + Tarefa + Formato + Restrições.

- ✓ Sem juridiquês
- ✓ Foco em riscos financeiros
- ✓ Formato em tópicos
- ✓ Máximo 200 palavras

Prompt Engineering Techniques

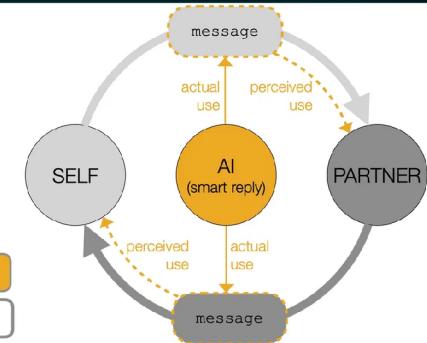
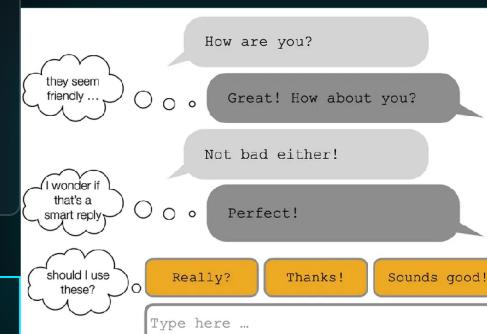


DESAFIO 2: FEW-SHOT PROMPTING

```
# Classifique o sentimento dos feedbacks:  
Texto: 'Amei a nova meditação!' -> Positivo  
Texto: 'O app travou no meio.' -> Negativo  
Texto: 'Poderia ter mais opções.' -> Neutro  
Texto: 'Não sei se gostei da voz.' -> [?]
```

SUA MISSÃO

Crie um prompt de **Few-shot** que ensine o modelo a classificar feedbacks de um app de meditação. Forneça 3 exemplos claros de padrões de resposta antes de pedir a classificação final.



DESAFIO 3: O LABORATÓRIO DE PARÂMETROS

CENÁRIO

Você está criando um gerador de slogans para uma nova marca de roupas sustentáveis. O objetivo é ser inovador e "fora da caixa".

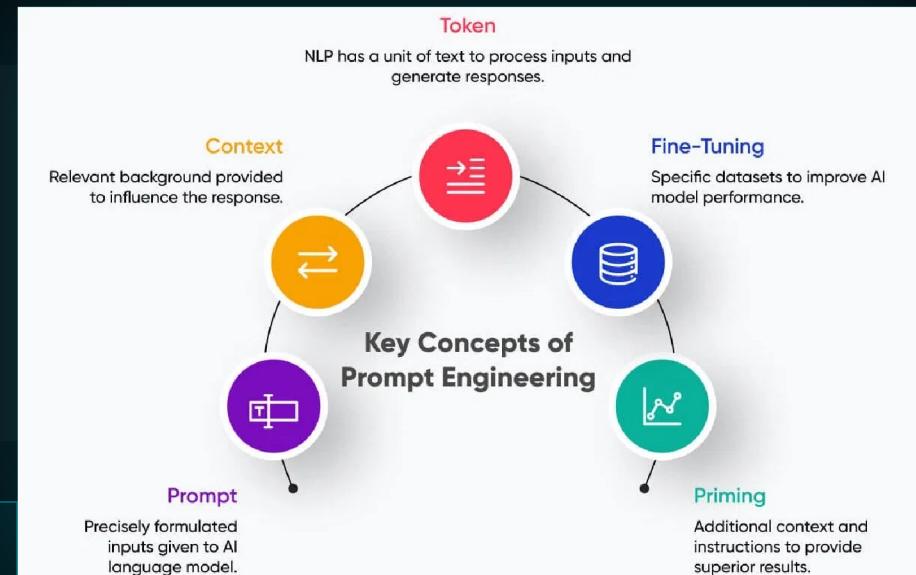
0.2

TEMPERATURE BAIXA

0.9

TEMPERATURE ALTA

- ?
- Qual configuração trará resultados mais criativos? Justifique com base na teoria de entropia e probabilidade de tokens.



DESAFIO 4: DECOMPOSIÇÃO DE PROBLEMAS

O PROBLEMA: PROMPT "FRANKENSTEIN"

"Crie um plano de marketing completo para um novo SaaS de CRM, incluindo análise de concorrência, 5 posts para LinkedIn e uma estratégia de SEO."



PASSO 1: ANÁLISE

Focar exclusivamente na pesquisa de mercado e diferenciais do CRM.



PASSO 2: CONTEÚDO

Gerar os posts baseados nos diferenciais validados no passo anterior.



PASSO 3: SEO

Estruturar as palavras-chave para sustentar o plano de conteúdo.

Sua Missão: Reescreva o prompt acima em 3 instruções sequenciais e lógicas.