**Introdução à Fundamentos de IA generativa com a Universia**

# Como Trabalhar com Inteligência Artificial

## Aplicações e Impacto da IA no mundo Atual

### IA no seu dia a dia

Machine Learning é a ideia da máquina aprender com você ou com uma massa de dados gerados por outras pessoas;

Ex: teclados antigos dos celulares(T9), depois evoluiu para teclados grandes e agora só teclado em tela(não tem o teclado físico)...troca de mensagens...teclados inteligentes na troca de mensagens onde você vai escrevendo e aparece sugestões das próximas palavras. Isso é um exemplo de treinamento de modelo de IA

* Teclados;
* Filtros de spam;
* (Algoritmos) modelos de sugestão;

### Os modelos de algoritmos de IA

Machine Learning é uma área da IA que busca ensinar máquinas e aprender a melhorar suas performances a partir de dados.

Tipos de algoritmos:

- Aprendizado supervisionado: tem um conjunto de dados rotulados, para um conjunto de dados eu já tenho uma saída determinada e com base nessas entradas e saídas já rotuladas é ensinado os padrões e orienta o modelo durante o seu processo de aprendizado. Relação matemática entre as entradas e saídas.

- Aprendizado não supervisionado: aqui os dados não são rotulados, tem se uma massa de dados e é necessário que algoritmo aprenda os padrões desses dados e os agrupe de acordo com suas proximidades/semelhanças. Ex: segmentar clientes.

- Aprendizado por reforço: um agente que interage com o ambiente e aprende de acordo com os feedbacks que eles recebem. Ele recebe recompensas quando faz ações corretas e punições quando erra. O aprendizado se da através de erros e acertos.

### Visão Computacional

Essa é uma área da IA, elas permitem que os computadores enxerguem e interprete conteúdos visuais através de imagens e vídeos (olhos e cérebro para as máquinas).

É usado Redes Neurais Convulocionais (são um tipo especializado de rede neural artificial projetada para analisar dados com uma estrutura de grade).

Um dos grandes desafios da Visão Computacional é a segmentação de objetos, isto é, separar diferentes objetos em uma imagem e identificar o contexto de uma imagem.

* Técnica amplamente usada em carros autonômos;
* Detecção facial;

O algoritmo aprende a detectar padrões em imagens e vídeos

### Linguagem Natural

Processamento de Linguagem Natural (PLN) é uma subárea da IA que que permite que computadores compreendam, interpretem e interajam com a linguagem humana em suas variadas formas. É ensinar o computador e se comunicar como se fosse um humano. São algoritmos baseados em modelos de redes neurais, são modelos que entendem o significado das palavras, frases e contexto do que está sendo falado/escrito/escutado.

Como as máquinas nos entendem: modelos são treinados através de textos e esses treinamentos ensinam as máquinas a aprender as palavras e os contextos em que elas são inseridas ou mais utilizadas.

- tokens

* Análise semântica: busca capturar o significado de uma palavra ou frase através da sua relação com outras palavras que já estão naquele texto, dessa forma, é possível inferir o significado mais provável dessa palavra com base no contexto ao qual ela aparece.
* Desambiguação léxica(semântica): é uma técnica que busca resolver a ambiguidade das palavras que possuem mais de um sentido. Nesse cenário a máquina usa o contexto na qual a palavra está sendo empregada para determinar qual o significado mais adequado para aquela situação em questão.

Ex de uso: geração de texto (geração de Linguagem Natural). O processo de geração de texto envolve Redes Neurais de Linguagem.

### Robótica

Uma área que é um casamento entre a IA e a Engenharia. Maior conhecimento agregado, envolve várias áreas como IA, Mecatrônica, Visão Computacional e busca desenvolver robôs que buscam interagir com o ambiente de forma autônomas.

Como os robôs conseguem ver e entender o mundo a sua volta: através das visão computacional e de sensores. Os robôs são equipados com câmeras e sensores de proximidade para compreender melhor o ambiente a sua volta. São equipados com algoritmos de aprendizado de máquina(ML) como **aprendizado por reforço** e deep learning

- Interação Humano-robô

## Aplicações Práticas da Inteligência Artificial

### Revolução das Ias Generativas

Ias Generativas são Ias capazes de criar, adpatar e aprimorar conteúdos de maneira autônomas, sempre aprendendo e se aperfeiçoando.

### IA no Cotidiano

IA na Educação

* Acessibilidade;
* Personalização;
* Facilidade de Uso.

**Cases DIO e algumas IAs interessantes**

* Case 1: o próprio curso trouxe a criação de avatares, transcrição de voz e texto em mais de um idioma; melhoria e auxílio na criação do curso;
* Curso Green Software;
* Uso de Ias generativas para a criação de Ias em múltiplos idiomas;
* Legendas, transcrições, libras, avatar;
* **Chats:** ChatGPT, Bing Chat(Microsoft), Bard(Google);
* **Geração de imagens:** Bing Image Creator, Leonard.Ai, Midjourney; Gama.app(Gera slides)
* **IAs de fala e Avatares Virtuais:** ElevenLabs, D-ID, HeyGen;

### Conhecendo a API do OpenAI

\*Documentação oficial da OpenAI: platform.openai.com/docs/overview

- dar fork no postman da collection da Openai das requisições da sua api(minuto 8)

### Geração de Imagens(Bing, Leonardo AI e MidJourney)

//Técnica que o Leonardo.Ai usa para gerar as imagens

### Voz e Avatares Virtuais(ElevenLabs, D-ID e HeyGen)

//Apresentação das ferramentas

### IA na Indústria

**Robótica e Automação Inteligente:** A IA está revolucionando a robótica e a automação inteligente, como na produção de veículos autônomos e robôs para a indústria.

**Saúde e Finanças:** A IA está transformando áreas como a saúde, tornando mais preciso o diagnóstico de doenças, e finanças, tornando mais precisa a análises e previsões de investimentos.

**Marketing:** A IA ajuda a indentificar os usuários ideais para campanhas de marketing e personalizar a experiência do usuário, analisando o seu comportamento e histórico nas mídias sociais.

## Introdução a Inteligência Artificial

### Como a Inteligência Artificial Nasceu

1950 – Alan Turing começa a se questionar se um computador poderia se passar por uma pessoa -> Teste da imitação

1956 – John McCarthy – em uma conferência ele disse que para um computador se passar (dentro da discussão levantada) por uma pessoa real era necessário que houvesse uma **Inteligência Artificial**, dai surgiu o termo da IA

1966 – 1º chatBot(Eliza) – Joseph Weizenbaum – Substituição de palavras por padrão – começo de um aprendizado de ML e entendimento que precisariam de um processamento de linguagem natural(usuário e computador conversar na mesma linguagem) – Levantamento de questões sobre Ética e Psicologia.

1972 – Parry (Kenneth Colby) – um psiquiátra, um chatBot voltada para a área específica

1984 ficou bem conhecido, um gerador de texto (Hector?)

Anos 2000 início de discussões maiores

2010 – IBM lançou IBM Watson, uma IA(ChatBot) voltada para perguntas mais complexas, um processamento de linguagem natural (PLN)

\*Tudo é um grande processamento matemático – probabilidades de palavras

Com essa evolução foi se criando os termos que conhecemos atualmente(PLN, ML etc.)

2018 – acesso via navegador/app -> começo do ChatGPT

### Como uma inteligência Artificial é treinada e o que são LLMs

Zerado -> contexto: o que ela vai precisar faze/saber -> guarda tokens(parâmetros) para que a IA aprenda os padrões daquilo que ela vai aprender e vai formando uma base de conhecimento(knowledge base) com esses tokens, quando a IA recebe um parâmetro novo e não sabe identificar, saber o que é aquilo pode receber reforços positivos para aumentar seu conhecimento

Definição da palavra “Artificial” = produzido pela mão do homem, não pela natureza; postiço / Que envolve artifício.

Quanto mais tokens, maior sua base de conhecimento e assim vai se tornando um Grande Modelo de Liguagem ou, em inglês, Large Language Model (**LLM**)

**SLM** = Small Laguage Model, são modelos menores, geralmente modelos mais especializados

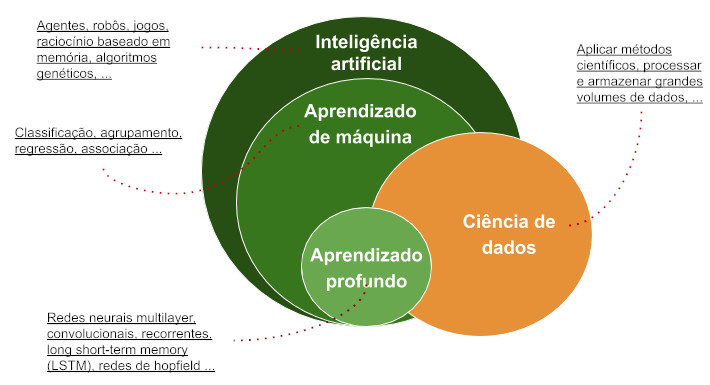
### Entendendo Deep Learning

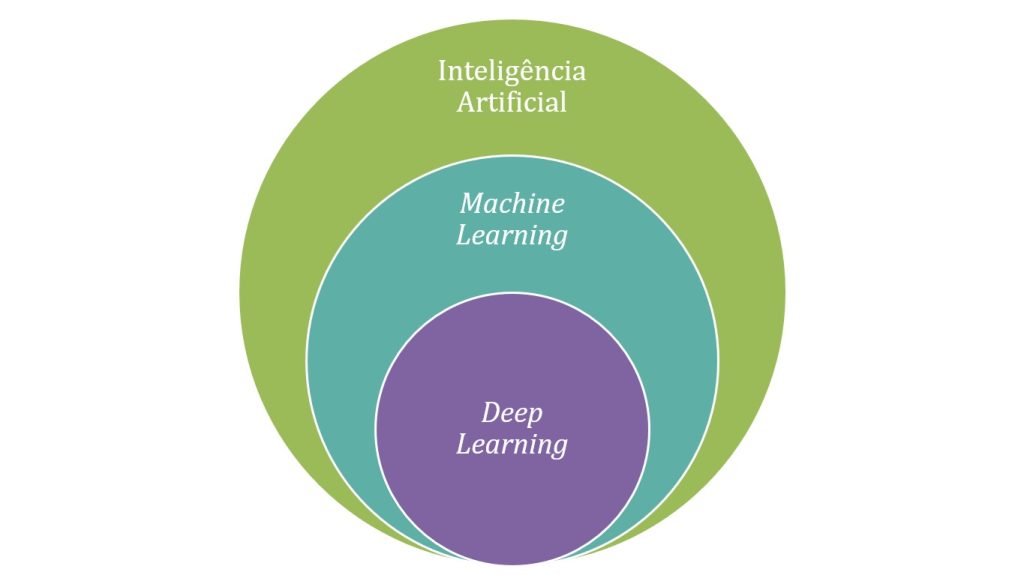
**Inferência:** é um processo de raciocínio no qual se chega a uma conclusão, ou seja, a um reconhecimento desconhecido, a partir de premissas ou observações que são consideradas verdadeiras.

**Neurônios artificiais –** vai conseguir inferir coisas

// Temos a Inteligência artificial, que é um conceito, ai tem o ML que é onde a máquina aprende e tem o Deep Learning, onde tem os neurõnios artificiais para que a IA aprenda mais, aprenda a inferir outras coisas por meio de um princípio.

//Atualmente as IA já são **Multimodais**, elas não trabalham somente com texto, mas também imagens, audios e mais.





### A Era das IAs Generativas

“A Era das IAs Generativas marca um ponto de inflexão na história da tecnologia, redefinindo a produção de conteúdo e remodelando indústrias inteiras. Diferentemente da IA tradicional, que analisa e reproduz dados existentes, a IA generativa é capaz de criar conteúdo novo e original, como textos, imagens, vídeos, músicas e código, a partir dos padrões aprendidos em vastos conjuntos de dados. ”

Estamos vivendo uma revolução histórica dentro da área de tecnologia que está afetando todas as outras áreas.

# Introdução à Engenharia de Prompts

## Introdução a Engenharia de Prompts

//Resposta gerada por IA: “A engenharia de prompt é a **arte e ciência de criar e otimizar instruções (prompts)** para sistemas de inteligência artificial (IA) generativa, como grandes modelos de linguagem (LLMs), para obter as respostas mais precisas, úteis e relevantes”

Pré-requisitos par o entendimento:

* Noção básica de Processamento de Linguagem Natural que trata da interação entre máquinas e a linguagem humana;
* Modelos de Linguagem de Grande Escala: que são modelos treinados com grande volume de dados para gerar e entender textos de forma eficiente;
* Ferramentas: ChatGpt e Microsoft Copilot

### Visão Geral da Engenharia de Prompts

Os LLM calculam probabilidades com base nos padrões aprendidos durante o treinamento.

//Arquitetura transformers -> permite que os LLM entendam o contexto de cada palavra em relação a todas as outras do texto e não de forma linear.

Texto escrito por um humano(prompts) -> tokenização(unidades básicas) pelo modelo de linguagem(IA) -> sequência de **tokens** processados pela camada de atenção e por outras transformações até gerar a resposta.

Os tokens são tranformados em embedings(representação vetorial que capturam o significado do token) -> depois esses embeddings passam por camadas de redes neurais com trasnformadores e então o modelo aplica essas transformações para poder entender o contexto, então o modela gera uma distribuição de probabilidades para tentar prever o próximo token e esse processo de tentar prever o token seguido é repetido até gerar a sequência.

**Como os modelos de linguagem “entendem” um prompt?**

De maneira geral os modelos de linguagem não possuem memória persistente entre interações, ou seja, eles não lembram o que foi dito em conversas passadas depois que a interação termina. No entanto dentro de uma mesma interação o modelo consegue manter o contexto do que foi dito por meio da **janela de contexto**, que é o limite de tokens que o modelo consegue processar simultâneamente(cada modelo tem seu limite, quando esse limite é atingido o modelo limpa os tonkens passados para dar espaço para novos).

**Elementos Essenciais de um Bom Prompt**

* **Instrução clara:** tarefa específica que quer que o modelo faça;
* **Contexto adequado:** fornecer as informações que vai ajudar o modelo a entender melhor essa tarefa que ele precisa realizar e produzir uma resposta mais precisa e relevante;
* **Exemplos:** para orientar o modelo sobre o formato ou estilo que queremos que a resposta tenha, pode se usar exemplos;
* **Dados de entrada:** informações ou o problema específico que você quer que o modelo crie, análise ou resolva;
* **Formato da saída:** formato que você quer que tenha a resposta(longa, curta, em tópicos, parágrafos etc.)

### Aplicações Práticas da Engenharia de Prompts

### Engenharia de Prompts no Dia a dia

Prompts para todas as áreas: saúde, educação, financeira, vendas etc.

**EX1 – Planejamento de Metas Pessoais:**

Crie um plano detalhado de desenvolvimento pessoal para os próximos [período de tempo] considerando as seguintes características:

- Foco: [área de interesse ou hobby];

- Objetivos: [objetivos específicos];

- Limitações de tempo: [disponibilidade semanal ou diária];

- Recursos financeiros: [orçamento disponível];

Estruture o plano com:

1. Metas específicas e mensuráveis;
2. Recursos de aprendizado dentro do orçamento disponível;
3. Cronograma semanal adaptado à disponibilidade;
4. Métricas de acompanhamento para medir o progresso;
5. Identificação de possíveis obstáculos e estratégias de superação.

**EX2 – Preparação para Entrevistas de Emprego:**

Simule uma entrevista de emprego para o cargo de [cargo desejado], considerando:

- Perfil: [perfil profissional, experiência e qualificações];

- Descrição da vaga: [detalhes do cargo, como responsabilidades, habilidades e qualificações necessárias];

- Empresa: [tipo de empresa, setor ou foco principal];

- Nível de senioridade: [nível de experiência esperado];

Passo 1: Pergunta Inicial -> pergunte a primeira pergunta mais provável de ser feita na entrevista;

Passo 2: Resposta do Usuário 2. -> O usuário responde à pergunta.

Passo 3: Feedback 3. Após cada resposta do usuário, forneça feedback detalhado sobre a resposta:

* O que foi bom;
* O que pode ser melhorado;
* Sugestões para aprimorar a resposta com exemplos contextualizados;
* Dicas de como melhorar ainda mais a performance na entrevista.

Passo 4: Continuar com as perguntas 4.

Continue o processo, fazendo uma pergunta por vez e oferecendo feedback após cada resposta.

### Cuidados na Aplicação de Prompts

* **Prompts enviesados:** perguntas que possam induzir ou sugerir a resposta;
* **Alucinações:** quando o modelo retorna algo falso ou inventado para tentar preencher lacunas de conhecimentos;
* **Considerações Éticas:**  ;
* **Privacidade e Segurança:** não compartilhar dados pessoais.

Preocupação em criar modelos de IA responsáveis que forneçam respostas, façam tarefas tendo como princípios:

* Imparcialidade;
* Confiabilidade;
* Segurança;
* Privacidade;
* Inclusão;
* Transparência;
* Responsabilidade.

### Leve seus Prompts ao Próximo Nível

## Técnicas de Engenharia de Prompt

Titulo maior - 3

Titulo dentro – negrito sublinhado

### Conhecendo o Prompt

**O que é um prompt:** instrução fornecida para que o modelo de IA gere uma resposta ou realize uma tarefa/ ação específica.

**A importância de um Prompt**

Os prompts vão guiar o modelo para nos dar uma resposta ou realizar uma tarefa que desejamos.

**Como os Prompts transformam Interações**

Um prompt bem formulado(claro) possiblita um retorno mais assertivo e detalhado.

**Principais Aplicações de Prompts**

* Educação: criação de material didático, explicações mais claras, analogias, tutorias especializadas, criação de cenários, ajuda em pesquisas etc;
* Desenvolvimento de software: possibilita geração de trechos de código, explicação de código de forma mais clara, possibilita documentação do software, geração de casos de teste, teste automatizado etc.;
* Marketing: geração de conteúdo(posts), ajuda em pesquisas de mercado, redação persuasiva, campanhas de publicidade etc.
* Ciência e pesquisa;
* Serviços jurídicos;
* Entre outras áreas.

### Componentes de um Prompt

**Instruções**

Orientam o modelo sobre a tarefa a ser realizada, especificando o que fazer, como “Resuma o texto abaixo em três frases objetivas.”

**Exemplos (Few-short Learning)**

Apresentamos pares de entrada e saída esperadas, mostrando ao modelo o comportamento ou formato ideal para a tarfea definida.

//Quando se fala de 0 short se refere a capacidade do modelo de realizar uma tarefa sem ter sido esplicitamente treinado com exemplos específicos para aquela tarefa. Já o few short é quando conforme você vai interagindo com o modelo e passando alguns exemplos ele vai aprendendo com a conversa que está tendo atualmente

**Contexto ou Configuração**

Define o papel do modelo e o cenário, como “Você é um assistente de IA que simplifica explicações técnicas para iniciantes em tecnologia.”

**Restrições ou Limitações**

Limitam a resposta em termos de extensão ou escopo, como “Responda em até 50 palavras, sem incluir opiniões ou dados fictícios.”

**Conteúdo Principal**

Texto central para processamento, como um parágrafo, tabela ou documento, que o modelo deve analisar, traduzir ou resumir.

**Indicações**

Guiam o modelo para gerar saídas específicas, usando estímulos “Liste pontos principais:” ou “Responda como um parágrafo único.”

//Se refere a sugestões, direções ou instruções que apontem para algo bem específico como recomendações ou orientações.

**Formato de Saída**

Especifíca a forma da resposta como “Apresente no formato JSON:” ou “Liste os resulatdos como uma sequência de tópicos claros.”

//Estilo ou estrutura na qual os dados são apresentados no final do processo.

**Conteúdo de Suporte**

Dados extras que ajudam a tarefa, como datas ou preferências, por exemplo: “Use o contexto atual de dezembro de 2025 nas respostas.”

### Técnicas de Engenharia de Prompt

**Introdução as Técnicas de Engenharia de Prompt**

Quando se fala de Engenharia de Prompt se refere a todo processo estratégico para criar, ajustar e otimizar os prompts com o objetivo de obter respostas mais úteis, precisas e relevantes dos modelos de linguagem. Todo esse processo de linguagem envolve refinar as intruções, modificar o contexto, adicionar exemplos claros e bem objetivos, interar e testar até que alcancemos o resultado desejado.

//Refinar não é sinal de falha, é uma parte fundamental.

//Componenetes de prompts são elementos básicos que compõe o prompt. Já nas práticas aboradará estratégias usadas para estruturar, ajustar o prompt para garantir a obtenção dos melhores resultados conforme o desejo do usuário.

**Entendendo o objetivo na prática**

A ideia de exemplo será usar o modelo de linguagem para gerar o conceito de um jogo.

**Aplicando instruções e Repetição**

* Instruções Claras: Técnica que organiza o prompt para que as orientações sejam objetivas, detalhadas e colocadas no início para maior clareza.

Exemplo da intrução:

Ajude a criar uma proposta de jogo narrativo chamado "Adventure DIO Quest". O jogo será aventura onde os jogadores tomam decisões que afetam o desenrolar da história. O jogo deve incluir : - Uma introdução à história; - As principais mecânicas de jogo; - Uma descrição do público-alvo

Juntando com outras técnicas, poderia utilizar a técnica de

* Repetir Instruções no Final: Reforça as orientações no final do prompt para garantir que o modelo compreenda e siga a tarefa proposta sem desvios.

Ajude a Ajude a criar uma proposta de jogo narrativo chamado "Adventure DIO Quest". O jogo será aventura onde os jogadores tomam decisões que afetam o desenrolar da história. O jogo deve incluir:

- Uma introdução à história;

- As principais mecânicas de jogo;

- Uma descrição do público-alvo.

---

Lembre-se, após a geração revise a proposta de "Adventure DIO Quest" para que ela inclua todas as seções: história, mecânica e o público alvo.

**Entendendo o Guardrails**

* Guardrails: Técnicas para limitar respostas, evitando conteúdos prejudiciais, irrelevantes ou incorretos, alinhando o modelo a padrões éticos.

**Preparando nossa saída**

* Preparar a Saída: Define palavras ou frases no final do prompt que ajudam a moldar o formato da resposta, tonrnando-a mais estruturada e clara.

Organize a proposta no seguinte formato:

- Nome do Jogo;

- Introdução à história;

- Mecânicas Principais;

- Público-Alvo;

- Estilo Gráfico ou Visual;

- Conclusão.

**Solicitação de cadeia de pensamento**

* Solicitação de Cadeia de Pensamento (Chain of Thought Prompting): Instrução para que o modelo responda passo a passo, explicando o raciocínio até chegar à conclusão final.

Explique por que as mecânicas escolhidas são interessantes para o público-alvo e como elas se alinham ao tema da aventura.

**Especifícando a estrutura de Saída**

* Especificar Estrutura de Saída: Indica o formato exato da resposta, como JSON ou listas, ajudando a obter resultados organizados e úteis.

Crie a proposta do jogo em formato de tópicos

**Dividindo a tarefa**

* Dividir a Tarefa: Quebra uma tarefa complexa em várias etapas menores, facilitando para o modelo responder de forma lógica e organizada.

Descreva a introdução da história de "Adventure DIO Quest". Comece com o jogador acordando em lugar misterioso.

//Nem sempre, para todos as atividades, será necessário usar todas essas técnicas, vai depender da complexidade e da necessidade de cada tarefa.

**Adicionando Sintaxe Clara**

* Adicionar Sintaxe Clara: organiza o prompt com formatações como listas, títulos ou tabelas, tornando as intruções mais visuais e compreensíveis.

Formate a proposta do jogo em uma tabela visual para uma apresentação rápida.

**Aplicando prompts para gerar imagens com o Microsoft Copilot**

Crie uma imagem de uma ilha misteriosa com árvores sombrias, estilo pixel art, cenário de jogo.

## Criando um Podcast com IAs Generativas

**Por que criar um podcast?**

Produza conteúdos sobre o que você gosta e para construir sua autoridade sobre algo, pode ser uma ferramenta de trabalho, um meio de você destravar sobre algo.

**Definição de um grupo**

Fala sobre algo que você goste: Java, dados, linguagens de programação etc. Difina um grupo, a quem você quer atingir e o nicho do assunto.

**O que é prompt Engineering**

<https://helpful-jump-17b.notion.site/PAS-Podcast-AI-Studio-210489e15d7a4a73b743bb159e45d06f>

Um Engenheiro de Prompt é um especialista em criar, refinar e otimizar as instruções (prompts) que são dadas a modelos de inteligência artificial (IA), como os grandes modelos de linguagem (LLMs). A função principal é servir de ponte entre a intenção humana e a capacidade da máquina para garantir que a IA produza resultados precisos, relevantes e de alta qualidade.

**Como escrever prompts melhores**

* Seja mais específico;
* Forneça um contexto adequado;
* Estimular a aplicação prática do conhecimento;
* Estimular a comunicação clara e concisa.

**Conceitos avançados de prompt**

* Prompts negativos: são partes do seu prompt onde você diz o que não deseja na sua resposta.
* Passagem variáveis(prompt pattern): Para deixar os resultados menos aleatórios e mas direcionados, sempre envie através de linguagem natural blocos de variáveis. Exemplo:
  + “Olá [NOME]! Como posso te ajudar hoje?” -> durante a interação, peça para substituir o bloco “[NOME]” pelo nome real da pessoa.

**Criando um título poderoso**

<https://helpful-jump-17b.notion.site/Roterista-ee83819ca5274e7684498ea7e8a7865a>

**Imagem de capa e dicas de Midjourney**

//Acessar pelo discord

**Criando um roteiro com variáveis**

//Usar os roteiros do Notion

**Gerando audio com Eleven Labs**

//Logar no Eleven Labs assim você consegue créditos no primeiro uso.

**Editando seu podcast de maneira fácil**

Exemplo de ferramenta: Capcut – tem como usar no celular e no computador.

**Como entregar meu projeto e dicas de onde publicar**

* Fazer o podcast e gerar o audio;
* Fazer o clone do git\_hub e colocar na pasta output o audio
* <https://github.com/felipeAguiarCode/prompts-for-podcast-generate-by-ia>

**Entendendo o Desafio**

## Criando um Ebook com ChatGpt & Midjourney