

eu li estudei e testei as técnicas de mais de 35 artigos científicos de nomes e pessoas como Google dem Mind meta Microsoft alibaba e também de pesquisadores das Universidades de Washington Singapura Universidade da California E tantas outras e também debole toda a documentação de alguns dos maiores modelos de linguagem para garantir que eu Não perdia nada na hora de criar um bom prompt e extrair o melhor resultado além de já ter criado mais de 20 agentes customizados que eu uso aqui na minha empresa para automatizar as tarefas e ganhar tempo e que eu passo para todos os meus alunos usarem também o que já gerou milhares de usos em alguns dos meus agentes que eu disponibilizo nas plataformas como por exemplo no po onde eu tenho o terceiro bote mais usado da língua portuguesa ali isso fez com que eu me tornasse o primeiro Conselheiro de inteligência artificial para empresas no Brasil e combinando isso com um estudo gigante de todos os cursos que eu a a o dem aqueles mais obscuros que tu só acha depois de horas no Twitter para finalmente testar todos eles os prompts e a saída em literalmente milhões de palavras para agora te trazer tudo mastigadinho e pronto para usar e por isso aqui a minha promessa para você nesta masterclass você vai aprender a como fazer a engenharia de prompt ou seja como extrair o máximo possível desses modelos de linguagem essa Ferrari que a gente tem nas mãos mas que poucos sabem usar tudo de uma maneira simples direta e prática para pessoas normais como você e eu que só querem aproveitar o máximo dessas ferramentas para nos ajudar nossas vidas economizar trabalho e produzir mais toda essa aula foi desenhada para ajudar desde o iniciante que talvez nunca tenha criado um prompt na vida até os mais avançados aqueles que de repente tem até alguns SAS ou desenvolvendo algum tipo de chatbot tudo para que você não precise nunca mais ir a atrás de prompts e lista de 5000 prompts que eu vejo as pessoas vendendo por aí Nunca entre aspas porque é um período bem longo se falando em Inteligência Artificial mas com certeza o suficiente para isso está resolvido e você ter confiança Total nas suas habilidades para conseguir e gerar os resultados que você quer e como a gente vai fazer isso são sete módulos que a gente tem aqui que eu recomendo que você faça na ordem Mas se você quiser pular algum ou quiser visitar alguma coisa você pode usar os capítulos aqui abaixo para pular então Quais

são esses modos o primeiro deles fundamentos onde a gente vai ver a parte básica e já ver como trabalhar da maneira certa com documentos longos o segundo é uma estrutura básica de prompt para já te colocar no caminho certo o terceiro é o processo recomendado o que que faz sentido pensar no prompt e o que faz sentido pensar fora dele como por exemplo a escolha do modelo certo pros resultados que você quer depois a gente vem pra parte de técnicas básicas que já vão ajudar você a ter respostas muito melhores dos seus prompts até a gente chegar nas técnicas mais avançadas que até por exemplo criar menus interativos para fazer prompts muito mais complexos que requerem um raciocínio maior depois eu vou mostrar alguns dos exemplos práticos tanto meu quanto de outros para ver como é que é tudo isso na prática e por último as ferramentas recomendadas para te ajudar a acelerar tudo isso e para caso você não me conheça aqui quem fala com você é Bruno Picinini eu já vendi mais de oito dígitos online enquanto eu visitava mais de 70 países ao redor do mundo e aqui nesse canal eu quero ajudar outros empreendedores e experts a fazer o mesmo a ter um negócio do tipo sed que é um negócio simples enxuto divertido escalável com qual objetivo ter não só resultados financeiros mas também mais tempo e qualidade de vida bora começar Então por aquilo que mais faz diferença e é o que a gente vai mais usar agora ao longo das aulas que são os fundamentos começando pela definição Afinal o que é engenharia de Prom a lição mais simples que nós vamos adotar agora pro resto é engenharia de prompt é a ciência empírica ou seja Com base no resultado no mundo real e as respostas que a gente tem de lá e não só teoria de planejar Criar e testar prompts para gerar melhores respostas em grandes modelos de linguagens ou large language models que é na versão como chama em inglês e por que que ela é tão importante porque ela é a meta habilidade que destrava Todas as outras aprendendo esta depois tu pode usar ela para fazer o que quiser para os seus próprios prompts para alunos para estudantes para sua empresa para software sas o que for e ela é tão importante que empresas como por exemplo a Anthropic que faz o cloud eles tinham antes até no passado uma proposta de emprego para engenheiro de prompt e agora mais atual que eles têm é esta aqui que é um um pesquisador de engenharia de prompt com um pequeno salário de dar

Salma molhada 300 a 10.000 por ano isso se a gente botar em em Reais dá mais ou menos R 2 mil 700.000 no Brasil óbvio que esse é um valor lá nos Estados Unidos Mas serve para mostrar a importância de tudo isso que você vai aprender aqui agora e para aprender o que a gente vai fazer é usar o que eu vi uma vez num livro de um cara chamado Josh witkin que ele fez um livro com o TIM feries inclusive sobre como aprender e lá ele contou a história de quando ele aprendeu xadrez a maneira que ele começou era um pouco diferente das outras que era mais algo Tipo isso ou seja Ele começava pelo final por quê Porque nessa situação de um rei e um peão contra só um rei ele foca nos fundamentos sem tantas coisas complicadas sem tantas regras para entender a essência do jogo além de aprender essa parte que poucos treinam porque às vezes dedicam tempos demais para aberturas e o jogo O midterm Ali do jogo e não o final então aqui que nessa aula em si a gente vai usar um processo parecido nós vamos começar com o mais básico do básico tanto que nós já estamos aqui nos fundamentos e vamos construindo em cima dessa habilidade para que tanto os iniciantes possam aproveitar mas aqueles que quiserem avançar e usar para cenários mais complicados ou com casos mais complexos também tenham as ferramentas e justo para isso eu gosto de usar este exemplo aqui que é um ovo Esse é uma uma referência que eu uso quando eu vou palestrar em eventos eu gosto até para interagir com o pessoal que eu cito para colocar a mentalidade certa na hora de pensar os prompts que de novo fundamentos se eu falar para você como é que faz um ovo pode vir em várias respostas a nessa hora eu peço pro público falar Alguns falam fritar cozinhar Mas você está querendo dizer comprar ou a questão do ovo uma galinha questão filosófica e eu falo Exatamente esse é o ponto se você não deixa claro o que você quer não tem como saber isso vale tanto para humanos Que Nós já vamos ver em seguida como para máquina Então a primeira regra de ouro que a gente vai levar em consideração seguinte que instruções você daria para alguém fazer porque se tu mostrar aquelas instruções que seria o teu prompt para uma pessoa ela conseguiria executar aquilo se ela não conseguiria dependendo o que é Talvez os grandes modelos de linguagem como chat GPT cloue gem e tantos outros também não então lembra sempre disso que é uma regra que se a gente levar em consideração que é o como

eu faria para uma pessoa do meu lado para instruir ela tem grandes chances disso ao traduzir para um prompt funcione muito bem e você vai ver como essa simples regra depois vai ter aplicações diretas desde os prompts mais simples Até técnicas bem mais avançadas que você vai aprender aqui mas de início o que que é mais importante comece com o simples para depois complicar por isso esta ordem que a gente vai seguir Não complique além do necessário faça o básico prompt simples vão funcionar para resolver a maioria dos problemas como dois exemplos Olha aqui eu tava no chat GPT perguntando justamente porque eu não tinha certeza se esse era o autor que dava este exemplo do xadrez que eu lembrava que eu tinha lido mas não tinha certeza então não precisei complicar muito fazer uma técnica nem nada disso simplesmente perguntei com a informação necessária para que o chat GPT encontrasse a resposta que eu precisasse e tem algumas coisas aqui do por que ele já dá uma resposta maior com outras lições que depois a gente vai aprender porquê Ou um outro aqui um outro prompt que eu tava na dúvida como na hora do inglês dizer don't invent Stuff que eu tava com que parecia um pouco do português do seu cisson tentando falar não inventa coisa e a maneira que isso Aria mais americanizada natural que eles falariam é don't make things up então prompt simples que eu não preciso complicar Além disso outro detalhe importante que vai te ajudar a fazer melhores prompts é na medida do possível começar com modelos mais caros que tendem a performar melhor por exemplo aqui do cloude Se tu entrar na página deles eles têm estes três modelos atuais o Opus que é o mais poderoso deles ele é bem mais caro do cerca de cinco vezes mais caro Dependendo do que você faz então isso vai fazer diferença por isso eu prefiro começar com os melhores modelos mesmo que sejam mais caros para depois reduzir isso eu tô falando de prompts mais complexos tá aquele outro como eu mostrei simples não precisa mas quando a gente quiser levar ao extremo do do que pode ser feito e de uma resposta muito precisa que a gente necessita aí vai fazer diferença então começa com os mais caros para eliminar isso não tentar otimizar com os mais baratos porque esses geralmente são os mais caros e depois a gente vai pros mais baratos para refinar para escalar agora antes de a gente partir para uma estrutura de prompt básica super simples um aviso tá vai ficar cada vez

mais fácil porque nós estamos agora em certos modelos que eles estão sendo treinados já os próximos e os próximos talvez quando você esteja assistindo esse vídeo Nós já estamos no GPT 5 6 ou 7 imagina quando for o 10 então cada vez mais será menos necessário complicar tanto e ser mais direto e simples mas algumas dessas técnicas que você vai aprender ainda assim vão perdurar porque elas ajudam a organizar o texto tanto visualmente para nós como pra máquina interpretar só que a grande moral que eu quero deixar claro aqui é para não exagerar porque não vale a pena porque vai ser cada vez mais fácil e como o naval costuma falar aá significa que os computadores estão aprendendo a nossa língua ao invés de nós termos que aprender os deles então por exemplo quando eu comecei no marketing digital lá em 2010 tinha que fazer HTML na mão escrever todo o código integrar o meu site hoje clica arrasta integra automatiza e tudo se resolve porque já se tem ferramentas no code e low code que resolvem isso e em termos de prompt é cada vez mais isso é saber mais se expressar e comunicar boas ideias do que uma programação complicada portanto sim aprenda como criar bons prompts como eu vou te mostrar aqui mas Além disso em ter bons processos por trás ter uma boa comunicação a regra de ouro se esse é um processo que seria fácil passar para alguém fazer eventualmente se o modelo não pode fazer agora ele vai ficar melhor que inclusive eu acabei de participar de uma reunião de uma empresa que queria me ajuda justamente para automatizar os seus processos com agentes e o que que eu recomendei para eles não pensa só o que que os modelos podem fazer agora porque se eu fizer isso quando tiver pronto tu já tá atrasado Tenta olhar e já chuta porque as leis de escala e de exponencialidade como a gente tá indo que eu uso esta fórmula aqui no pulso para me lembrar disso é isso que vai acontecer então se programa para o que vai estar disponível Talvez daqui 1 do anos porque aí o teu tempo de construir a tecnologia já vai te acompanhar para entregar exatamente aquele o que antes era sonho que agora vai ser realidade e com isso você fica com essa vantagem competitivo porque enquanto os caras estão olhando para trás você está fazendo o que o Peter Drucker falava qual que era melhor melhor maneira de prever o futuro que é inventá-lo agora avançando um pouco nós vamos ver uma estrutura básica ainda de como gerar um bom

prompt que na dúvida usa ela que ela vai servir pra maioria dos casos porque aí depois nós vamos entrar com técnicas mais avançadas para expandir em cima dela então essa técnica eu coloquei com acrônimo de prompt que serve para englobar tudo isso como é que ela funciona cada uma dessas letras representa uma das partes de uma estrutura de um prompt quando este precisa ser um pouco mais complexo lembrando para não complicar além do necessário o que que é a primeira letra do P é de Persona você é um especialista em marketing redes sociais e psicologia humana como um exemplo Então como o treinamento desses grandes modelos de linguagem tem um conteúdo muito grande se nós não direcionarmos ele não tem bem certeza do que a gente tá falando então nós estamos dando contexto para ele para depois ver pra segunda letra que é o roteiro eu quero ajuda para criar um roteiro para um reels no Instagram Esse é o exemplo do que eu quero fazer pro terceiro objetivo o objetivo é chamar atenção com uma história cativante e depois fazer uma chamada para se inscrever para depois passar pro M modelo como é que eu quero este resultado Quero o resultado dividido em 8it até 10 slides como aqui só um exemplo depois a gente vai ver nas técnicas mais avançadas o quanto que dá para trabalhar em cima disso especialmente se você vai usar para outros fins Panorama Meu cliente é um Expert vende curso online pá P pá P pá ou Além disso depois em algumas das técnicas que a gente vai olhar aqui dá para colocar mais informações como exemplo e tantos outros que vai ajudar e muito nos resultados e por último a questão do T transformar que isso para entender quando a gente coloca lá o prompt Não espere que logo de primeira ele já dê a resposta perfeita às vezes vai ser uma interação base inicial vai mandar vai dar feedback vai fazer e Inclusive eu vou te mostrar algumas das técnicas de como a gente pode fazer com que essa melhoria essa otimização seja feita tanto com o seu feedback como a própria máquina se treinando isso a gente vai ver agora com calma então nessa parte transformar a gente vai dar um feedback Troca isso faz assim faz assado implemente isso implemente assim esse então é a estrutura básica de um bom prompt que só com isso aqui tu tem digamos o 2080 que já serve pra maioria dos casos Ah e como aviso não tenha medo de mudar a ordem tá essa aqui é uma da da ordem que eu coloquei para facilitar inclusive muitas vezes quando

eu vou criar um prompt mais robustos eu começo literalmente digitando essas seis letras e dali eu começo a digitar meu prompt porque da eu sei que ele vai seguir uma estrutura tem algumas coisas que a gente vai ver aqui por exemplo se é um documento muito longo isso tem que vir antes ou depois das instruções isso é importante e faz bastante diferença no resultado que agora vou explicar com calma visto isso qual que é agora o processo que a gente faz que se um prompt simples não resolveu se essa estrutura que eu acabei de passado também não resolveu Qual que é o processo que a gente vai para desenvolver prompts mais robustos que requerem um processo de raciocínio lógica maior para gerar o resultado que a gente quer então o processo recomendado por um estudo que eu fiz de diversas empresas tutoriais e cursos de como fazer é o seguinte ele envolve alguns passos o primeiro de todos o mais básico e Fundamental defina a tarefa e os critérios de sucesso isso é muito importante porque muitos começam a desenvolver o Prom até o negócio inteiro e nem sequer sabem onde tá o Gol para marcar não sab onde querem chegar e não tem como dar certo então por que que isso é importante tem três critérios aqui pra gente olhar primeiro lugar desempenho e precisão o quão bom precisa ser esse modelo Porque dependendo que você precisa pode ser só uma resposta genérica agora por exemplo eu tenho no com os meus mentorandos onde a gente ajuda a rentabilizar um conhecimento escala com agentes de a o tipo de resposta que tem que dar com julgamento desde cop marketing tantos outros requer um grau de um grau de precisão bem grande a ponto de eu nem usar o chat GPT e hoje ter usado mais o cloud para algumas das coisas que a gente precisa inclusive no modelo mais claro então isso influencia a minha decisão de como pensar o processo para que é meu prompt segundo latência qual o tempo de resposta aceitável porque modelos mais caros não sempre mas por exemplo no caso do clode é bem dá para ver isso bem claro tendem a ser mais lerdos porque eles têm um processo de raciocínio uma base maior não sei exatamente como é que funciona maior então ele demora mais para dar resposta Mas é uma resposta com mais precisão agora se tu tá num chatbot na internet Talvez tu precisa de uma resposta mais prática sacrifica um pouco da qualidade mas para ter uma resposta mais instantânea se por exemplo está conectado com o WhatsApp então isso também

influencia e por último preço Qual é o seu orçamento para isso porque Dependendo do que você tá fazendo vai fazer muita diferença como eu mostrei aqui na tabela dos custos só que no Cloud aqui do haq pro Opus dá uma diferença brutal em termos de custo muito grande então tudo isso vai mudar e muito da maneira que você pensa o prompt definido a tarefa e os critérios do Sucesso a gente vem pro segundo passo que é desenvolver casos de teste ok se eu quero chegar lá como é que eu sei que eu estou perto disso como é que eu sei que esse prompt está gerando os resultados que eu espero então eu ten que desenvolver esses casos de testes de principalmente os padrões aqueles que correspondem a 80% mas também aqueles que a gente chama de borda os Ed cases ou seja aquela exceção da exceção Será que o nosso prompt também conseguiria vale a pena incluir aqui porque aí a gente pega e escreve um prompt Inicial pode usar aquela estrutura que eu acabei de mostrar do do acrônimo prompt para pelo menos ter uma base e ver se Estamos no caminho certo porque lembrando A ideia é iteração não é numa primeira vez que a gente acerta é um processo constante porque a gente vai testar contra os exemplos que a gente pode julgar nós mesmos ou botar outras llms outros grandes modelos de linguagem para julgar e ver se tá certo inclusive sugerindo melhorias refinar este prompt conforme os resultados tem alguns dos prompts aqui que eu estou constantemente refinando eu tenho por exemplo um editor de e-mails que me ajuda a todos os e-mails que sou eu 100% que escrevo eu uso como editor e quase sempre que eu vou lá eu adiciono mais alguma modificação que me ajuda a identificar e melhorar cada vez mais então é como se eu treinasse alguém mas ao invés de treinar uma pessoa eu estou treinando essa ia para que este Prom agora pronto fique cada vez melhor melhor e melhor e ainda mais e agora consigo compartilhar com os meus alunos O que é bem legal então refino prompt e por último ponho em produção depois que tá pronto Agora eu tenho muitos dos que eu tenho aqui que eles já estão em produção que eu uso constantemente paraas tarefas que eu precisar visto essa introdução básica nós vamos agora pra parte de técnicas que acredito que muita gente que ao ouvir as palavras engenharia de prompt tenderia a pensar que é que a gente vai ver agora mas o importante é como eu falei antes ter estes fundamentos que a gente acabou de aprender porque a

gente vai aprender como o xadrez que a gente vai construindo de trás para a frente para agora a gente adicionar em cima Então é isso que a gente vai fazer agora a primeira delas que a gente vai aprender não é uma técnica específica de engenharia de prompt mas que ajuda muito mas é que é a questão de Mark Down para caso você não conheça markd é esse tipo de texto aqui ó que por exemplo tem esse o sustenido aqui e outras que é um tipo tipo de texto que ele pode ser convertido ele é usado em texto plano que a gente chama mas ele pode ser interpretado para títulos e subtítulos então por exemplo este aqui se eu tirar esse sustenido ele vira um texto normal se eu colocar um ele vira dois ele vira um heading um cabeçalho um se eu botar dois ele vai diminuindo porque é um cabeçalho dois TR e assim por diante que é o que a gente usa em páginas HTML tem também por exemplo que tem no WhatsApp de certa maneira se eu colocar por exemplo ao redor dessa palavra aqui um asterisco e um asterisco do outro lado ele faz o Itálico Se eu colocar tirar meu mouse da frente dois asterisco ele vira agora Se eu colocar um por exemplo um sublinhado ele vira Itálico também e se eu colocar este símbolo aqui de do tio dois tios de cada lado neste programa aqui ele faz esse Strike ele faz o taxado E tem também outros comandos como por exemplo aqui no obsidia que eu estou usando se eu colocar dois símbolos de iguais ele vira um highlight então tem diferentes tipos de markdown e depois a gente pode até procurar um sheet sheet que é um um guia rápido e prático de markdown procurando rapidamente no Google já encontrei aqui dois que eu posso deixar o link depois aqui abaixo sen não só digital markd sheet sheet no Google que você deve achar então tem heading Itálico código que você vai ver que a gente vai usar bastante listas links e tantos outros o principal são esses aqui em cima tem aqui uma outra imagem que você pode usar mas o que que é o importante de tudo isso aqui é entender que markdown é importante primeiro porque visualmente eu fazendo aqui no obsídio é mais fácil de a gente se localizar e por esses grandes modelos de linguagem aprenderam Com base no texto eles replicam isso então um texto bem organizado para nós tem grandes chances que para a máquina também serve lembra da regra de ouro que eu falei lá no início aqui serve para você ver como é uma das aplicações um texto bem organizado ajuda e muito para você e para os modelos

responderem melhor então por exemplo neste prompt aqui que é de sumário e tarefas para mentorandos que eu mostrei num outro vídeo aqui no canal eu mostro e pode ver como ele tá organizado eu tenho headings para separar algumas das sessões Eu tenho esse tipo de tag que a gente já vai falar que são tags XML xmls tem tem Itálico eu dou indicativos que eu quero que a resposta venha em markdown porque depois é melhor o tipo de resposta que ele dá dou exemplos e tantos outros e assim eu vou formatando então visualmente me ajuda e eu sei que isso tira uma resposta melhor do chat GPT cloude ou que for inclusive isso é algo que eu descobri agora nas minhas pesquisas é que um prompt bem formatado dá uma resposta diferente na documentação do cloude por exemplo eles falam isso e dá para entender porque se esses modelos de linguagem eles são criados para simular como o nosso cérebro funciona de redes neurais é só você pensar como você adapta a maneira que você como você fala dependendo com quem você está falando se você vai falar com a sua avó ou com grande amigo ou amiga o seu jeito e as palavras que você escolhe são diferente então no modelo de linguagem é a mesma coisa um prompt bem formatado é como se o chat de GPT o cloude ou gmin entendesse aqui está uma pessoa educada vou responder da mesma maneira então isso aqui é um tipo de resposta que ele deu baseado neste prompt bem organizado ele deu este tipo de resposta pode ver que ele é organizado que é o que eu mostrei do do outro vídeo usando markdown usando lista Itálico e tudo que precisa de uma maneira que visualmente é melhor para mim e pros meus alunos agora antes de fechar aqui algo importante duas coisas que eu quero que você preste atenção este aqui E este aqui O que que é isso esse símbolo assim é uma linguagem de marcação extensiva XML extensive markup language isso aqui é importante porque isso Tá bastante na documentação do Cloud não tanto chat GPT mas faz sentido ele se localiza muito bem porque algumas vezes e principalmente a gente já vai ver com documentos longos Ele se perde no meio do caminho entender o que o que que é referência o que que é instrução o que que são exemplos e se você usar isso para separar ele ajuda Então pode ver que aqui eu coloquei aqui Aqui está a transcrição coloquei ela com essa chave e embaixo pode ver que eu coloquei a barra para fechar esse esse comando e quando eu me referencio a ela eu diria

Por exemplo essa maneira assim leve em consideração essa aqui transcrição fecha a transcrição isso aqui tá literalmente na documentação Então faz diferença Então isso é uma das coisas dois isso aqui está entre essas duas Chaves porque aqui a parte que eu sei que eu ou os meus alunos quando eu for passar alguns dos prompts que eu passo para eles eles precisam vir e modificar isso aqui é muito bom porque olha só se eu copio e colo aquele prompt jogo aqui no console da antropic que é para testar o modelo deles pode ver que ele já formatou aqui bonitinho essa parte da transcrição O que que significa isso isso é agora uma variável que se eu vier aqui no site e vir nesta parte aqui de variáveis e cá ele já me dá o espaço para eu só colar minha transcrição aqui então faz bastante sentido deixar estes espaços com esse comando porque visualmente é mais fácil de achar e algumas das Ferramentas também já interpretam da maneira certa e pra gente fechar aqui mais dois exemplos eu coloquei isto aqui porque quando eu quero que ele crie o insites eu tô dando alguns exemplos e você vai ver depois nas técnicas mais avançadas o quanto a gente vai usar esse tipo de marcação para ajudar e aqui o do exemplos exemplo do insite original exemplo do Insight melhorado porque às vezes ele não tava produzindo o que eu queria Então isso ajuda bastante e por último já dando uma dica extra caso você trabalhe com documentos muitos longos Porque neste caso aqui tá quando eu uso este prompt o que vai aqui em transcrição é uma transcrição de uma mentoria nossa do meu grupo de mentoria de 3 horas que dá mais ou menos 40.000 palavras é texto para caramba e até no outro vídeo eu ainda não tinha feito isso e aqui eu já corriji quando você tiver isso as são as próprias instruções dos modelos falam isso quando você tiver documentos muito longos inclua ele antes das instruções que é por exemplo o que antropic recomenda aqui estruturação de documentos longos quando estiver trabalhando com documentos muito longos use tags XML que isso ajuda bastante ou XML e em termos de posicionamento posicione eles antes para depois colocar as instruções abaixo que é exatamente o que eu fiz aqui vou te passar a transcrição preste atenção porque depois eu vou te questionar aí que teria a transcrição e depois as instruções desse prompt do que eu quero que ele faça portanto documentos longos coloca antes e as instruções depois só que tem alguns casos quando o

documento não só é longo como ele tá dividido e quebrado em seções aí tem dá alguns probleminhas que a gente tem técnicas de como fazer para resolver isso que eu vou mostrar

Principalmente nos exemplos que eu vou mostrar aqui no final da aula quando a gente vai juntar todas ou quase todas as técnicas num único lugar como segunda técnica digamos assim a gente vai ver o que é prompt do sistema isso aqui ajuda bastante a refinar o resultado que tu consegue para que que serve isso se tu olhar por exemplo aqui no console da antropic tem aqui no sistema prompt do sistema Esse é quando tu coloca uma mensagem de como tu quer que o sistema atue então alguns dos exemplos tem por exemplo instruções para uma tarefa responder em determinada língua responder de determinada maneira responder de maneira socrática ou em perguntas ou Em lógica e tantas outras coisas então isso aqui ajuda bastante na interpretação do teu prompt para que ele Gere melhores resultados como eu falei responder em língua alguma coisa assim então aqui no meu se eu vier dentro do meu chat GPT vier aqui na parte de customizar chat GPT aqui tem algumas das coisas que parte desse prompt aqui eu peguei com o meu amigo Alan que a gente tinha criado no passado junto então adaptei algumas das coisas para mim aqui ele pergunta o que que ele deve saber sobre mim para dar uma resposta melhor e aqui como ele deve responder no caso tá desativado agora poderia ativar e ele vai fazer algumas dessas coisas e a aqui nessa parte dos prompt de sistema que é uma das técnicas que a gente vai usar para diminuir bastante a Alucinação e ajudar no nível de qualidade das respostas que a gente já vai ver em seguida aqui em algumas das técnicas esses prompts do sistema em si não é sempre que se usa mas vale a pena considerar porque se é algo muito repetitivo de algum estilo de language que você quer que o modelo atue vale a pena incluir como terceira técnica ou estratégia de prompt a gente tem o que a gente chama de zero shot e você vai ver esse termo se repetindo muitas vezes o que que significa é que o modelo é treinado em bilhões de parâmetros quando a gente faz uma pergunta para ele sem nenhuma referência ou exemplo a gente chama isso de zero shots ou seja zero exemplos então interprete shots como exemplos para o modelo seguir pode ver aqui que nesse prompt por exemplo que eu tava querendo uma ajuda para otimizar o título que eu tava

criando para como você organizaria melhor esta frase para que ela fosse mais clara e simplesmente coleí meu título eu não dei mais exemplos ou shots de como é um título esperado de como eu acredito que deveria ser porque isso com certeza melhoraria a resposta mas nesse caso eu só quis ir direto ao ponto me diz como melhorar e aqui ele deu algumas sugestões de como ele faria só que obviamente e a gente já vai ver em seguida com exemplos a resposta é absurdamente melhor que é uma das técnicas que a gente usa então aqui ele fez essa deu mais sugestões aqui num outro prompt a mesma coisa eu mandei um um outro título que eu tava trabalhando e me diz como melhorar não dei muitas alternativas inclusive pedi para ele me gerar alternativas Então esse é o zero shot que eu não dou nenhuma referência só faço minha pergunta bem direta e agora a gente começa a entrar em algumas das técnicas que não são super avançadas mas eu quero te mostrar a origem delas para explicar melhor e também para entender que não é simplesmente de alguns videozinhos no YouTube gerando lista de prompts tem estudos científicos de pessoas sérias na área e eu quero te mostrar alguns deles porque Foi a pesquisa que eu fiz e já interpretei para ti para que você não tivesse que ler tudo isso então um deles aqui neste papel aqui de de um pessoal da Universidade de Califórnia na Santa Bárbara e da Microsoft eles demonstram e provam mostram aqui essa técnica chamada de estímulo de prompt direcional algo n parece complicado 27 páginas de um estudo que eu já entendi já traduzir para você que é esta ideia aqui e eu já vou te mostrar um exemplo o que que é ideia num prompt normal eu falaria por exemplo PR faça um resumo do artigo em duas ou três frases e dá uma certa frase e aqui um score que atingiu se você simplesmente fizer uma modificação super simples que é baseada nas dicas que eu tô traduzindo direto aqui do inglês e dar estas dicas a resposta dele já vai ser muito melhor porque agora você direcionou através desse estímulo o que você quer deixa eu te mostrar um exemplo na prática eu peguei um outro paper que a gente já vai ver em seguida aqui nas técnicas que é uma chamada skeleton of thought que é chamada de esqueleto de pensamento e para me ajudar a interpretar tudo isso eu coloquei aqui através do po que é uma das Ferramentas de que eu uso no modelo Opus que é o melhor do Cloud de 200k com uma janela de contexto maior e aqui no final primeiro dica que eu dei

antes coloquei o documento antes e aqui pro final eu pedi para ele fazer um resumo de tudo isso só que ao invés de só dar as instruções de fazer o resumo eu vim aqui e falei Organize a marcação de maneira organizada usando as keywords abaixo que é justamente que veio daí então coloquei como keywords coloquei aqui um comando para diferenciar um pouco promos exemplos aplicação prática resultado e comparações porque assim eu sei que com isto com este estímulo direcional ele tem uma chance maior de procurar as referências certas e organizar da melhor maneira que me interessa essa é uma técnica Super Simples mas que Ajuda muito principalmente na hora de trabalhar com grandes documentos que a gente vai usar e em seguida esse próprio aqui que eu mostrei que eu usei como referência aqui desse esqueleto de pensamento você vai ver como ele atua para fazer algo ainda mais robusto a próxima técnica que a gente tem é uma relativamente simples mas que bem usada com prompt bem calibrado já vai gerar a maior parte dos resultados que tu precisa ela se chama F shot que significa que que a gente falou antes zero shot que eu não dou nenhuma referência portanto no f shot eu vou vir aqui e fazendo isso olha a diferença que pode dar isso aqui é os resultados que eles mostraram de quanto de precisão com zero shot One shot que é um só exemplo ou fio que é alguns exemplos e a melhora é significativa é realmente muito grande então como exemplo um prompt que eu uso que é um gerador de títulos infinitos que eu passo pros meus alunos e mentorados ao invés de eu simplesmente dar as instruções que eu já tenho aqui eu também aqui dentro eu dou exemplos do que são bons títulos algumas das referências que eu peguei com grandes nomes como John Benson Stephen George Craig clemens e outros caras aí do copywriting então eu coloco aqui esses títulos como exemplo a explicação de cada um deles porque o que que é isso lembra da regra de ouro se eu fosse pedir para uma pessoa fazer se eu mostro para ela crie um título aqui exemplo de 20 títulos vencedores que funcionaram para nós qual que é a chance que essa pessoa produziria um resultado melhor então todo esse estudo é basicamente para dizer isso ponha exemplos e a gente até vai ver que tipos de exemplos funcionam melhor mas simplesmente dar exemplos positivos já ajuda e muito e voltando na nossa estrutura básica que eu mostrei antes esses shots ouos exemplos era o que eu colocaria

aqui na parte de Panorama onde eu dou não só informações do meu cliente ou de que mais eu quero mas também exemplos para guiar melhor o modelo e produzir um um resultado uma resposta melhor a próxima técnica que a gente vai aprender é uma das mais famosas e se chama cadeia de pensamento essa aqui é muito boa para se usar quando quer uma linha de raciocínio e lógica por exemplo para aritmética ou problemas complexos maior do que simplesmente uma resposta digamos bruta como é que ela funciona eu já vou te mostrar um exemplo Então nesse papel aqui do pessoal da do de pesquisa do Google como pode ver aqui eles fizeram esses testes então aqui um prompt normal onde um problema de lógica e a resposta errada nesse cadeia de pensamento eles incluem no input Ou seja no prompt que eles colocaram um exemplo da pergunta e como o modelo deve se comportar que nesse caso é fazer essa cadeia de pensamento que é como a gente faria isso seria um pouco como aquelas perguntas de entrevistas de raciocínio e lógica que às vezes fazem por aí ou dizem que fazem do tipo Quantos balões cabem nessa sala você não daria uma de maluco e falaria 26 não né provavelmente você faria um passo a passo pensando de que Ok qual o tamanho do balão Qual o tamanho dessa sala Quantos metros faria uma multiplicação e aí tentaria estimar com base nisso e não daria um chute qualquer Então essas técnicas tentam simular esse processo de raciocínio porque é como a gente faz e com isso a gente consegue um resultado mais preciso para problemas mais complexos repara de novo Quantas vezes a gente volta na regra de ouro tente pensar como seria o processo mais adequado para nós resolvermos o problema e é isso que isso aqui faz ele mostra como é o processo de como esse modelo deve fazer vou te mostrar um exemplo que não é exatamente isso mas de certa maneira é onde eu estava tentando resolver um problema no meu site que eu tinha um código de um vídeo de vendas do vurb que é uma ferramenta que a gente usa que tava dando problema e eu em vez de eu esperar pela resposta eu vim aqui no chat GPT joguei o código aqui e expliquei meu problema isso aqui tá dando erro por favor me ajude porque eu não entendo suficiente ele modificou o código e depois porque isso está nas instruções por trás que a gente vai ver depois na hora de evitar a Alucinação ele me ajuda a explicar o que foi feito porque isso não só me ajuda a entender o que ele

fez para implementar a solução como para ele mesmo pensar e me dar melhores respostas e a gente continuou ainda não tá funcionando a ele fez uma nova alteração e me explicou e me deu sugestões de depuração se isso não está funcionando vamos tentar isso para tentar ajustar Então isso que é a cadeia de pensamento quando tiver um problema de raciocínio lógica maior deixar os modelos pensarem explicitamente funciona melhor e isso aqui no pessoal do cloude eles dão mais algumas sugestões de como levar isso adiante e produzir ainda melhores resultados que por exemplo para uma pergunta que eles colocaram aqui que requer um certo tipo de raciocínio eles falam para dividir em etapas passo por passo para conseguir essa resposta só que o que você pode fazer E aí dependendo onde é que você vai ter a aplicação você pode usar as tags XML que eu mostrei antes para separar o que é pensamento e o que é resposta e você vai ver como a gente vai usar isso depois nas alucinações então aqui ele instrui para pense usando um espaço para Thinking de pensar e depois coloca a resposta dentro de respostas Isso é muito bom se por exemplo você quer desenvolver um SAS ou alguma plataforma no WhatsApp para dar resposta você pode ter um espaço e aí fazer algum filtro onde aqui o modelo pensa para por exemplo se precisar fazer alguma conta e a resposta é só dentro destas tags específicas para que o cliente não veja aquilo que ele não deve ver e nesse caso eu tava falando aqui com o professor sandec aqui do canal Dele quem quiser acompanhar recomenda um cara bem fera entende do que faz realmente entende do assunto e não só como muitos viu alguns vos vídeos no YouTube e agora se acha especialista ensina e a professor no assunto e eu conversava com ele sobre isso Será que o modelo consegue fazer este tipo de raciocínio internamente Porque aqui no Cloud eles falam explicitamente é importante notar que esse pensamento não pode acontecer sem o output os resultados o cloud deve produzir seu seu pensamento para realmente pensar que que é desvantagem disso custo porque o output é o mais caro \$5 por 1 milhão de tokens no caso do cloud no modelo mais avançado então sim se necessário beleza mas tem um custo e aí a gente tava tentando ver se isso faria diferença ou não em outros modelos a resposta é pros outros modelos pro chat GPT parece que tem algum tipo de processo interno mas não tenho 100% de

certeza pro Cloud eles explicitamente falam você tem que mandar o cloud produzir este pensamento porque senão ele não vai usar como referência e em cima dessa técnica nós vamos para a próxima que é a cadeia de pensamento contrastiva digamos assim que é um pessoal de Singapura e do alibaba que é super simples pensando que a gente acabou de aprender em como botar os passos para fazer o que ajuda bastante aí para alguns modelos mais outros menos é não só dar o modelo da explicação certa de assim é a resposta que eu quero como colocar a explicação errada assim é como eu não quero porque assim o modelo pode aprender tanto com exemplo certo como com errado legal Bruno entendi isso da cadeia de pensamento mas como é que eu faço essas perguntas respostas indico algumas vezes já é óbvia você já tem exemplos que você quer colocar mas tem um atalho de um outro estudo de como atalhar um pouco disso que a gente pode usar inclusive nas instruções do prompt de sistema que são zero shot reasoner que eles chamaram de cadeia de pensamento com zero shot o que que é isso isso aqui é bem simples Tá tudo que a gente falou antes na questão de cadeia de pensamento Onde colocar dando aqui as instruções de como o modelo deve pensar se eu não tenho isso simplesmente instruindo o modelo a falar vamos pensar isso passo a passo respire com calma e pense logicamente passo a passo às vezes em modelos mais avançados já é suficiente para naturalmente o modelo fazer todo esse processo sem tu ter que explicitamente dar os exemplos não é sempre que funciona depende o que você quer mas muitas vezes ajuda e agora com isso Você entende quando por exemplo quando lançar o Gemini 1.5 pro a interpretar tabelas dessa O que que significa essas tabelas isso aqui são os nomes de testes qualquer coisa você pode colocar no Google ou perplexity para pesquisar o que que é tipos de testes diferentes de do que que eles comparam os modelos e é isso que significa esses 10 shots five shots 11 shots que são significa que nos inputs dos modelos eles colocaram 11 exemplos do que eles queriam e inclusive Se você olhar aqui embaixo por exemplo neste aqui formatados como C problems ou seja problemas de cadeias de pensamento é isso que significa essa sigla inclusive alguns com zero shot pra comparação de quando você não dá nenhum exemplo como é que eles performam O legal é que ao entender tudo isso agora você consegue

interpretar melhor quando eles postam gráficos Tipo isso Ou aquelas tabelas comparativas de modelos foi por isso que quando saiu inclusive o Gemini Ultra eu não fiquei tão impressionado porque tem um detalhe aqui que não sei se você reparou mas que eu vou vou te mostrar agora primeiro aqui eles estão reportando contra o GPT 4 certo sota significa State of the art Model ou seja o modelo melhor do mercado no momento que era o GPT 4 só que repara só lembra que a gente aprendeu o que que é shots nesse aqui eles estão comparando um com cadeia de pensamento CT de 32 shots comparado um de cinco shots no GPT 4 isso faz muita diferença então eles postam isso aqui como se fosse algo ível Mas será que é mesmo não dá para ter 100% de certeza até porque o Google tá meio desesperado e é engraçado Além disso um p comentário Extra esse gráfico assim explodindo um foguete rumo à Lua mas se você colocar em perspectiva ele tá mais para algo assim não é uma mudança tão grande comparativamente sim É eu sei que na escala logarítmica das coisas vai ficando cada vez melhor mas porque a gente precisa está muito bem obrigado e se você enxergar de outras maneiras já não é tão incrível mas de qualquer maneira Vale frisar de novo a escolha do modelo certo vai fazer muita diferença no resultado final o mesmo prompt que você roda no GPT no gemini ou no Cloud vai dar uma resposta completamente diferente no geral eu tenho usado o Claudião da massa na maioria das vezes mas tem certos comandos que eu já usei no gemini ou no GPT e produzi o resultado melhor depende dos seus objetivos não só depende como muda a cada atualização Então tem que ir testando e acompanhando conforme as coisas mudam agora neste momento um aviso Lembra que eu falei que a gente ia construindo em cima das habilidades O que você viu até aqui na boa é mais do que o 2080 pra maioria dos casos das pessoas aqui neste canal já vai ser suficiente o que a gente vai ver agora é algumas técnicas que vão aprofundando para prompts mais complexos para produzir respostas melhores de qualquer maneira mesmo dito isso eu recomendo que você continue assistindo por quê porque primeiro é super interessante dois vai te dar ideias de como expandir a mente para depois para um prompt mais simples tu conseguir produzir respostas melhores e terceiro fica como referência para caso você queira criar algo maior ou impressionar os amigos com

o seu conhecimento você já vai saber como é que funciona e também porque muitos dessas próximas técnicas vão ser mostradas e incorporadas nos exemplos que eu vou mostrar aqui no final começando então com as mais avançadas a primeira delas é uma que se chama consistência própria ou self consistency que é esse nesse paper aqui que eles divulgaram pelo Google onde o que que eles falam aqui logo de cara eles já comentam que a técnica de cadeia de pensamento tem produzido resultados ótimos mas o risco dela é que aqui como ela faz ela é um pouco inocente às vezes ela gera uma linha de raciocínio mas que não é a certa e isso dá erro como é que eles uma das maneiras de resolver isso e a gente tá falando aqui Mais especificamente de prompts mas talvez se você for mais avançado podia até executar em paralelo é que eles sugerem que que ao invés de gerar só uma cadeia de pensamento você Gere três e dessas o próprio modelo compare as três para ver qual que tem a maior chance de produzir o resultado final então isso não é algo que eu uso o tempo inteiro mas achei interessante mostrar por motivo que eu acabei de falar para ter como ideia porque se você tiver dificuldade e Verê que a resposta que você quer não vem custa a vir de repente incorporar alguma dessas técnicas mais avançadas gerando múltiplas respostas talvez te ajude Mesmo que não seja para o modelo se julgar por conta própria mas para que você visualmente julgue e fale eu gostei da 2 3 e 7 que eu costumo usar bastante prompts assim eu mando gerar 10 sugestões como alguns que a gente vai ver aqui escolho três e trabalho em cima desses três porque assim eu tenho uma chance maior de acerto como próxima nós temos agora da cadeia de pensamento para a árvore de pensamento que aqui tem dois papéis aqui um do pessoal da deep Mind do Google e o outro de um laboratório em na Califórnia em São José que eles falam dessa técnica de árvore de pensamentos que é de novo expandindo na cadeia de pensamentos para ter resultados ainda melhores como é que funciona visualmente é fácil de entender mas a aplicação prática nem tanto aqui eles mostram um exemplo inicial de o prompt básico input output aqui eles mostram como a cadeia de pensamento com abreviação CT como é que funciona e o que que era o que a gente acabou de ver de consistência própria onde a gente gera múltiplas respostas e dessa escolha aquelas que tem a maior chance de estarem certas o que eles recomendam aqui é

que isso para problemas mais complexo e por isso que muitas dessas técnicas mais avançadas às vezes elas soam muito legal parece incrível mas não tem tanto uso prático Ou pelo menos não pra maioria das pessoas então eu vou mostrar algumas delas aqui porque dá sim Para incorporar algumas das coisas como eu vou mostrar mas não tanto quanto que a gente viu antes mas a ideia é que se desenvolve uma uma árvore aqui de conhecimento onde essas alternativas estão sendo estudadas para que dessas possibilidades porque é um processo que se assemelha como Nós pensamos vamos resolver tal problema hum a gente pode fazer um dois ou três mas essa tem problema tal essa sabe como a gente vai desenvolvendo essa a ideia tentar replicar este processo de raciocínio dentro desses modelos então nisso esse tipo de estudo aqui esse aqui é um outro paper que eles fizeram parecido mas tem Digamos um roteador que eles usam de maneira interna para gerar os resultados como é que isso aqui seria na prática em termos de prompt tem um cara que eu achei aqui no github que eu achei que ele fez uma sugestão que eu adaptei e aqui ele comentou que leu sobre esse papel e ele sugeriu este prompt aqui que incorpora essa ideia que ele fala imagine três experts diferentes e eles conversam entre si e a gente deixa com que o modelo faça isso e olha só como eu fiz na prática aqui porque eu fui inclusive além eu usei essa ferramenta chamada chathub que eu consigo conversar com duas três quatro ou seis modelos ao mesmo tempo e nesse caso aqui eu coloquei para conversar com o chat de GPT e com o cloud para comparar a resposta então o que eu fiz que era uma das técnicas que eu tava estudando para testar a aplicação dessa eu criei três personas uma que é o copywriter que escreve os títulos o prospecto como se fosse os meus possíveis clientes julgando e um gerente que dá as instruções pro copywriter aí eu crio um loop aqui por comando usando de novo repara que eu tô usando as tags XML como eu falei e aqui eu vou fazer esse processo eu quero que você Execute E aí conform o copywriter escrever o prospecto julga o gerente pega esse feedback e passa adiante aí eles começam aqui e fazem esse processo entendi as informações me passa que eu vou começar aí eu beleza passei aqui alguns dos títulos que eu tinha mostrado no outro e ele começa o copyright começa a fazer algumas sugestões aí vem sugestões aí vem feedback do prospecto sugestões do

gerente e assim adiante e aqui ele faz a próxima aí eu só falei continue e ele fica repetindo esse processo então eu podia fazer isso diversas vezes até que eu ficasse satisfeito para ver outras ideias sugestões só que ao invés de ter que ser eu ou alguém da minha equipe ou alguma outra pessoa dando feedback ou se eu quisesse eu podia vir aqui onde eu digitei continue eu dar o meu feedback próprio para este gerente que simular de certa maneira como seria a estrutura da minha empresa para que eles incorporassem isso que eu falei e colocassem aqui para jogar Então essa é uma das aplicações que aqui eu tô mostrando mais como exemplo até quero fazer uns testes mais avançados em cima disso porque eu vejo o potencial de como tu pode pegar essa essa técnica que a gente acabou de aprender aqui de árvore do pensamento e ter uma aplicação real aí pros seus prompts o próximo que a gente tem é um que Eu mencionei antes que se chama esqueleto de pensamento esse aqui é um papel do do pessoal da China Estados Unidos Bélgica e outros lugares e o que que é a ideia disso aqui que eu já aplicava de certa maneira mas foi interessante eu aprender como fazer isso de uma maneira mais estruturada porque na verdade esse papel aqui esse paper aqui esse estudo ele é mais para um treinamento interno dos modelos do que para o prompt em si só que o que que eu fiz eu peguei tudo isso aqui que são 51 páginas dei uma bela olhada e lida vi os prompts que eles deram como exemplo e sugestão E aí onde que eu apliquei primeiro eu fui lá lembra que eu tinha mostrado que usei o po aqui para interpretar me ajudar a interpretar Esse estudo e me dar sugestões e aí depois algo que eu já fazia que eu quero te mostrar agora um exemplo de como eu uso isso já misturando algumas técnicas então aqui eu tô no guia então aqui eu tô no guia de implementação da da minha mentoria que é do Expert exponencial e aqui na parte de marketing eu tenho o que eu chamo da causa surpreendente principal que é um que a gente cria um mecanismo único do problema se você não sabe que é isso não tem problema agora mas é para diferenciar da concorrência basicamente e nesse aqui para me ajudar a criar ao invés de dar uma resposta direta O que que eu faço e eu vou mostrar aqui como exemplo então aqui eu tenho um agente específico que faz isso aqui ele dá as instruções eu primeiro mandei os dados do meu cliente que eu gerei com outra agente que eu mostrei antes e

aqui agora ao invés de pedir logo de cara já a resposta definitiva eu peço para lhe gerar um esqueleto do pensamento que é basicamente gerar uma lista das possibilidades que a gente pode fazer então até tem esse nome esqueleto do pensamento com estudo científico mas é um processo de raciocínio de como a gente tomaria boas decisões um Brainstorm de ideias pra gente colocar agora julgar elas com uma explicação rápida mas que isso também podia ser adaptado por exemplo por um um processo ento de algum raciocínio para julgar e ir para os próximos passos que o que eu faço agora Inclusive inclusive nesse modelo aqui eu crio uma espécie de menuzinho para que eu coloco para criar por exemplo dar feedback explorar mais opções escolher mais e eu beleza botei letra A Vamos explorar um pouco mais mandei mais detalhes lembra o que eu falei antes não só feedback da máquina mas feedback humano junto o meu feedback eu estou guiando como se fosse uma pessoa do do meu lado pro que eu quero Entendi fez mais sugestões achei bacana me deu mais sugestões agora combine essas Lembra que eu falei que isso a gente ia fazer e agora ele começa a fazer isso eu vou desenvolvendo porque assim é mais rápido e eu economizo também tokens se eu tiver usando api do que eu mandar ele desenvolver o que a gente vai fazer agora todas as vezes escolhi aqui alguma específica E aí eu peço agora eu quero detalhar isso que está embutido nesse agente a explicação do que eu quero que é criar uma definição técnica criar uma definição simples e um pit do elevador mais metáforas e sugestões de marcas tudo isso eu já Já pré-programados e já tenham essas sugestões para depois a gente explorar em cima junto nós virmos lá no nosso trabalho que a gente faz semanalmente e quinzenalmente pra gente aí explorar junto quais são as melhores ideias inclusive já achando algumas sugestões de argumentos de quais aquelas informações e provas que eu preciso usar para que este argumento que eu acabei de criar sejam mais não é que eles sejam mais verídicos para eu conseguir provar o que eu tô querendo dizer inclusive achando argumentos para me ajudar a provar isso que a gente acabou de criar então ele acha tudo isso faz aqui o todo o processo me dá de novo as opções do que eu quero fazer e Assim Segue Então essa é a ideia ao invés de começar com algo já super complexo eu desenvolvo pelo menos uma uma estrutura como se fosse o índice do que eu quero fazer e depois

eu vou passo a passo desenvolvendo Isso é muito bom e eu uso bastante por exemplo na hora de criar uma copy mais longa que eu sei que tem ali os sete passos que eu tenho que bater as coisas que eu tenho que fazer em vez de fazer de uma vez só que é muito difícil e com grande risco de errar a gente subdivide e com essa referência nós vamos trabalhando ponto a ponto o próximo que a gente tem aqui é esse aqui que é a geração de conhecimento através de prompt e você vai ver que conforme a gente avança cada vez menos é a praticidade de algumas dessas técnicas são realmente só para casos extremos mas nesse aqui eu vou mostrar como é que eu usei Qual a ideia desse é gerar um conhecimento que a própria ia o próprio modelo gera para isso usar como base para depois gerar as próximas respostas postas e apesar de ser algo assim Digamos um pouco mais avançado isso em larga escala é o que empresas como Open ai e meta consideram fazer para treinar os seus modelos ou seja gerar de certa maneira um conhecimento um conteúdo sintético para esse colocar e treinar os modelos então não é uma ideia toda ruim como é que eu usei isso lá no outro prompt que eu tinha mostrado aqui quando eu fui criar um perfil do meu cliente ideal que seja para fazer um título de 10 palavras se o modelo Tem mais informações do meu cliente a ch de ele acertar essas 10 palavras é maior então é isso que eu faço aqui para gerar um perfil do cliente ideal passei algumas informações e ele foi gerando todo este perfil com diversas e diversas informações que depois eu uso ele para gerar os próximos passos às vezes não é gerar grandes textos assim mas que seja um título já é mais informação e contexto que ajuda a melhorar a resposta o próximo é o prompt mático é um que eu não usei tanto ele tem sim a aplicação para de novo certos casos que você esteja tenha dificuldade a ideia dele é usar uma uma técnica de ensino que a gente já usa na escola de novo lembrando da regra de ouro mais parecida como a gente aprende que é o justifique a sua resposta então a gente pede pro modelo dar a resposta e justificar porque daí talvez ele próprio entenda quais erros ele tá cometendo E com isso também melhora a resposta o próximo que a gente tem aqui é a geração argumentada de recuperação que eu já traduzi aqui direto esse aqui é um artigo da Meta onde eles falam sobre qual que é uma das limitações de todos esses modelos a gente falou das técnicas de fio shot zero shot cadeia de

pensamento mas tem um limite o quanto eu consigo colocar no prompt porque eu tenho por exemplo toda a minha base de conhecimento das minhas aulas de transcrições de arquivos vídeos e podcasts então isso aqui eles fazem a sugestão que esse é o próximo passo isso aqui já é bem mais avançado que eu sei que a maioria das pessoas aqui não vai aplicar mas para quem quiser realmente avançar e ter respostas ainda melhores não se coloca só no prompt se combinam os modelos de linguagem para que estes atuem com uma base externa de conhecimento para aumentar a resposta e com isso de novo produzir um resultado muito melhor vou te mostrar dois exemplos agora um um pouco assim não exatamente o que eles falam Nesse artigo e o outro um pouquinho mais avançado mas também nada maluco para tu ver como é que funciona isso na prática então eu tô aqui no obsidian que é a ferramenta que eu uso para anotações e eu vou te mostrar o que que é a ideia porque como eu falei só depender do prompt em si tem um certo limite Sim a gente vai colocar essas informações no prompt mas o ideal é ter uma maneira de consultar outras informações outras bases de conhecimento de uma maneira dinâmica vou mostrar uma aqui Super Simples então aqui no obsidium tem um plugin chamado CoPilot que é esse aqui que eu tô usando que ele integra diferentes modelos que eu conecto a minha api nesse caso tô usando o clod porque eu gosto bastante e aqui eu posso vir no chat e conectar algumas dessas informações que eu já tenho salvas que eu uso para fazer o que eu preciso então eu diria algo como use as informações que você encontre Quando aperto o colchete colchete ele me traz esse menu que eu posso literalmente linkar este artigo que está aqui e ele vai botar essa informação no meu prompt Então usa informações que você encontra em premissa persuasiva e depois consulte o perfil do meu cliente que tá nesse aqui usando o que você aprender cliente 7D E aí agora mando enter e isso por trás ele vai fazer e vai mandar essas informações para agora criar então a a vantagem que com o promet que em teoria se você for olhar ele é relativamente pequeno é só isso aqui só que o que que acontece por trás eu estou mandando todas as informações que estão aqui aqui no que tem alguns detalhes do que eu passo na minha mentoria de sete critérios que eu quero que a pessoa olhe para desenhar o perfil de um cliente ideal e aqui é uma visão Geral do meu cliente que

eu já desenhei desde mecanismo oferta quem é o meu cliente com informação digamos mais específica E aí com base nisso ele consegue me dar uma resposta muito melhor como você pode ver aqui então Essa é uma das maneiras de usar agora tem outras que por exemplo se eu vier aqui nessa ferramenta que é o chatbase que talvez você conheça que eles usam api da openai por trás para criar um chatbot aqui eu mostro como que é isso usado mas de uma maneira mais dinâmica onde eu tenho um chat que podia estar aqui incorporado na minha página no WhatsApp no Instagram ou de outras maneiras e eu fiz uma conversa simples só testando o que é a regra 1 1 1 um que é um dos conceitos que eu ensino lá na minha mentoria e ele dá essa resposta pode ver que isso aqui não é uma resposta que ele puxou da base de treinamento dele ele foi lá nos documentos que eu subi aqui que a gente treinou a Flávia me ajudou a fazer isso e dessa E daí que ele deu a resposta e deu aqui então um exemplo do que eu queria estratégias de como usar e depois o que é o perfil 7D de um cliente também que é esse conceito que eu acabei de mostrar que tá lá no obsídio tá nas minhas aulas ele pega isso aqui e devolve como uma resposta isso aqui tudo aqui tem algumas informações do modelo a temperatura que é usada que eu comentei antes isso aqui a gente vem aqui nesse painel dessa ferramenta que é o chatbase e o que que a gente fez a gente treina ele para que ele possa responder com essas informações então agora a gente não dá só resposta Com base no prompt ou no que aquele modelo tem de informação mas sim a gente puxa um conhecimento externo e sobe através de arquivos pode ver que a gente subiu a transcrição na verdade não só a transcrição tá tem uma mania certa de tratar essas informações para subir aqui e já dando a dica a melhor maneira é fazendo um formato de perguntas e respostas Então tem prompts que a gente usa que transformam a transcrição que eu tinha das minhas aulas em perguntas e respostas porque se assemelha mais o que vai acontecer quando a pessoa vir aqui no chat ela provavelmente vai fazer perguntas se essa pergunta for o mais parecido possível com o que já está na base de treinamento dela a chance de ela responder com uma maior confiança é maior que é o que acontece aqui então eu tenho aqui não só não só os arquivos da transcrição mas os arquivos da transcrição tratados que a gente subiu aqui podia subir mais também o meu site que

aí vai ficar mais ou menos um perguntas e respostas e até do no então aqui no perguntas e respostas a gente vem para adicionar algumas daquelas perguntas que de repente não cobriu ou não ficaram boas suficientes que a gente consegue ver aqui em atividades pelo menos nessa ferramenta aqui aqui eu venho e pode ver que tá aqui o meu no chat e aqui ele me dá um índice de confiança de como ele tem feito ah essa resposta minha confiança é baixa aqui é 0.888 aqui é uma 9.50 Então dependendo do quanto que ele acredita que tem dessas respostas na base de treinamento dele ele vai dar uma resposta melhor então essa digamos essa esse artigo que eles descrevem aqui é uma maneira que tu não resolve só no prompt é o prompt com uma base de conhecimento extra para adicionar essas informações e com isso gerar a resposta mais perto possível do que baseada no que você tem como conhecimento seja como aula seja como transcrição seja com PDF seja artigos que você já criou que você alimente alguma ferramenta aqui eu só mostrei uma delas mas tem tantas outras para ir gerar a resposta perfeita agora esse aqui eu achei bem interessante que é como usar técnicas de linguagem de programação só que nos modelos de linguagem então o que que é a ideia aqui eles demonstram aqui nesse exemplo como eles usaram por exemplo a criação de variáveis para isso usar na cadeia de raciocínio do modelo como é que eu usei quando eu faço por exemplo anúncios que eu passo para alunos e mentorandos ao invés de ter que colocar no meio do texto eu usei essa ideia de variáveis e eu já Crio e coloco no início usando aqui as minhas minhas tags de XML eu coloco aqui variáveis que eu criei por exemplo do público promessa principal o culpado enganoso e depois aqui embaixo eu tenho um template de anúncio Onde eu deixo esse espaço para essas variáveis Assim fica mais fácil porque eu peço pro modelo aqui que ele pode sim ajustar a linguagem mas sem modificar tanto então se eu tenho um template de uma copy específica que eu quero que os meus alunos se inscrevam sem mudar tanto essa é uma das melhores maneiras de fazer Então essa é uma das maneiras de usar essa técnica que vem da linguagem de programação mas para prompts para criação de texto copy e outros e agora o último para esse momento é este aqui que é excelente é a maneira de se ar para agentes mais sofisticados que não fazem uma única tarefa mas múltiplas

como alguns exemplos que eu vou mostrar que é o react que ele usa a a parte de raciocínio com ações ou seja se eu quisesse executar uma tarefa maior eu não faria simplesmente numa única passada eu teria o primeiro passo um segundo um terceiro e um quarto e cada um desses Passos talvez tenha tarefas E subtarefas então esse aqui essa aqui é a ideia desse papel aqui desenvolvendo pelo pessoal de prinston e do Google onde eles mostram com uns exemplos aqui e eu vou mostrar mais ou menos um exemplo na prática onde Para uma determinada tarefa eles dividem começando do pensamento para a primeira ação e observação se você já usou algum tipo de mini agi baby agi alguma coisa assim é esse tipo que faz onde ele gera esse primeiro pensamento e daí ele vai desenvolvendo uma cadeia tá ok mas isso gera um segundo pensamento que eu preciso fazer Tal Coisa e Tal Coisa e Tal Coisa que é como a gente racionar Então essa é a a moral dessa dessa técnica aqui de prompt onde a gente coloca razão e ação não é simplesmente executar de uma vez só mas é questionar e ver o que que cada uma das nossas ações implica para decidir os próximos passos não é exatamente isso Tá Mas por exemplo nesse aqui que eu tenho um prompt e-mail de que ele me ajuda a editar meus e-mails o que eu faço é dividir que é uma das técnicas que a gente viu antes em pensamento e sugestões porque assim eu consigo ver algumas das ideias do que ele tá analisando para depois ver as sugestões Gerais e como a gente viu anteriormente pelo fato de eu dar um espaço pro modelo pensar a chance é que os comentários dele depois sejam melhores e como um exemplo de como isso tá mais ou menos embutido eu tô aqui no site de de um desses Agentes do Baby Adi onde ele executa que de um plano simples Eu só dei essa instrução aqui de comando ele começa a desenvolver subtarefas então ele faz esse processo de raciocinar ação raciocinar ação e ele percebe que tem as determinadas tarefas que ele precisa fazer fazer e a dependência de cada uma delas e depois ele começa a executar uma por vez pode ver que algumas aqui delas deu erro então não é um sistema ainda perfeito mas essa é a ideia dessa técnica que pra GES mais sofisticados é a maneira de resolver e agora para fechar essa parte uma das grandes dificuldades que muitos têm que é como evitar alucinações porque se você usa para criar copy emails ou até como um chatbot às vezes esse bote pode inventar coisas

que não é nada legal tem até algumas histórias engraçadas disso como é que a gente faz tem algumas técnicas eu tô por exemplo aqui no chat GPT se eu vier aqui no meu balãozinho e clicar para customizar chat GPT ele abre esta tela aqui que a gente conversou anteriormente e aqui que tá o segredo Como você quer que o chat GPT responda primeiro pode ver que eu coloquei da cadeia de pensamento a estratégia de ao raciocinar use o passo a passo e pode ver que Este comando aqui me ajuda com isso para evitar alucinações Gere conteúdo preciso e factual depois eu tenho outras quando quando citar Fontes indique o link ou referência e essa aqui principalmente se você especular ou prever algo informe-me fique à vontade e essa aqui é a parte principal tá recomendação do próprio da documentação deles Fique à vontade para dizer que você não sabe a resposta porque se tu não disser isso os modelos vão inventar porque ele quer tentar te agradar Então deixe ele à vontade para que se ele não souber a resposta te informar só isso aqui já ajuda bastante a evitar muitos das alucinações para que ele não invente co coisas que não é verdade outra técnica que ajuda bastante é parecido com o que a gente viu anteriormente de antes de ele processar resposta peça que ele procure citações ou insights no texto que você está mandando porque daí com isso ele usa isso aqui e você pode instruir Tente se basear o máximo possível dentro das situações ou senão de novo dê permissão para que ele fale eu não sei fazendo isso e colocando aqui no output no resultado explícito diminui as chances de L inventar alguma coisa e por último uma outra coisa que poucos sabem quando eu faço por exemplo o Sumário e tarefas da minha mentoria que eu ensinei no último vídeo eu não quero que ele invente muita coisa eu quero que ele fique o mais perto possível do que eu falei que é neste prompt aqui no caso então uma configuração que tem que ser usada é vir aqui nas configurações de modelo no caso eu tô por exemplo aqui do cloud da un tropic Mas seria mais uma coisa de api porque no chat GPT em si não tem isso especificamente então é mais para api quem usa por exemplo chatbot e é diminuir a temperatura Porque quanto maior ess a temperatura maior a criatividade então se eu quero ideias títulos e outros eu vou aumentar porque eu quero que ele invente se eu não quero quero que ele seja mais factual e mais parecido com o texto que eu enviei eu vou diminuir tipo até 01 02 máximo

03 para que ele não desie muito do texto Recebido e a última para garantir é essa técnica que a gente viu antes de consistência própria em que você pode pedir para ele gerar mais de uma resposta e e mandar que ele próprio desenvolva um consenso do que que é a resposta mais esperada Beleza agora vamos ver tudo isso na prática num prompt assim relativamente complexo mas bem grande pra criação de Landing pages que é esse aqui que eu desenvolvi em parceria com o pessoal do Great pages que é um Construtor aí de páginas do mercado que talvez você conheça Então vou pegar aqui esse prompt que eu já tenho aqui elele salvo no bloco de nota para te mostrar e te most mostrar o que que ele tá fazendo e como é que eu programei ele para fazer isso antes de a gente avaliar o prompt em si essa aqui é a tela de configuração de Agentes se você não sabe como fazer depois vai ter um link aqui em cima ou logo em seguida ou na descrição para ver uma aula de como criar o seu primeiro agente pegar tudo isso que a gente vem fazendo agora e colocar aqui dentro mas nesse momento eu só quero estudar mais o prompt em si Porque é importante tá E algumas das outras configurações que a gente tem aqui principalmente porque tem este arquivo aqui neste formato que é o que a gente já vai olhar o Primeira coisa eu quero te mostrar antes o tipo de resultado que ele gera tá Porque daí a partir daqui eu quero que tu já tente pensar pô mas como é que ele fez isso E aí eu vou te mostrar como é que é na prática no prompt para você entender não só como gerar algo mais complexo do que uma Simples Resposta mas inclusive Passos multipass caso você seja necessário por exemplo pedir informação do cliente pedir mais detalhes e trabalhar em cima disso que é algo que muitos querem E aí algumas das técnicas que a gente estudou até aqui por exemplo aquela do react podem fazer bastante sentido então aqui de Inicial Eu eu só cliquei em iniciar que são um dos botões que são os iniciadores de conversa que eu defini aqui com algumas das funções básicas aí coloquei esse aqui iniciar e ele começou este aqui é um texto padrão que eu botei lá para ele falar todas as vezes inicialmente num próximo exemplo que eu vou mostrar de um dos gpts mais usados do mundo eles usam algo parecido vou te mostrar como é que eu fiz aqui eu deixo dois links número um esse aqui pra pessoa clicar e aí já tem uma maneira de faz fazer um opin e eu consegui uma Lead caso outras pessoas usem porque

eu tenho agentes que são públicos e outros que não e também deixo aqui que nesse caso eu desenvolvi um modelo de Landing page que tá em PDF como é que eu anexo aqui na verdade eu só Criei um link e eu vou te mostrar como é que eu fiz isso lá dentro do prompt usando uma das técnicas que a gente aprendeu e aqui é só um PDF um PDF simples que eu coloquei com essa estrutura colocando com numerais cada uma das sessões isso aqui é mais uma referência visual para quem estiver usando o agente se localizar mais fácil então eu criei isso pros meus alunos e pro pessoal do Great pages para ficar mais fácil para eles e achei interessante colocar aqui logo de cara para que a pessoa possa clicar nesse link e já acessar esse modelo aí conforme ela preenche ela vem aqui e ela sabe de qual sessão estou falando como eu falei mais referência visual aí começa primeiro você já tem o Avatar do cliente eu programei isso para perguntar mandei aqui uma informação básica do que eu tenho e ele começa show entendi vamos começar as tarefas aqui estão as sessões que temos como ele puxa essas sessões e cada e o detalhe de cada uma delas é o que eu vou te mostrar como é que eu fiz para programar nesse prompt quais sessões você gostaria de criar Primeiro vamos começar pela sessão dois o problema Central aqui estão os detalhes e aqui ele faz um rascunho breve não vou entrar em detalhes agora se tá bom ou ruim ou pior ou melhor nesse caso eu quero te mostrar mas claro que eu já tento puxar para que seja a melhor resposta possível para no máximo ter uma edição final antes de ir pro site em si aqui ele já dá algumas sugestões do que fazer e ele show perfeito gostou do rascunho ótimo Qual é a próxima sessão que tu quer continuar e depois o processo continua aqui onde eu posso ir criando mais de uma sessão por vez Então se a pessoa quiser criar outras sessões ela dá as opções aqui ela pode consultar aqui no modelo para ter uma noção do que se trata e vai pedindo para este agente específico criar para ela agora como é que eu fiz tudo isso como é que eu programei aqui dentro aqui no prompt junto com este arquivo para produzir esse tipo de resultado é isso que eu quero mostrar e estudar junto com você Com base no que a gente vem aprendendo Então já coloquei aqui algumas das informações que é o que a gente vai usar que esse aqui é o prompt completo Tá eu vou mostrar a estrutura dele Pode ver que ele é um prompt relativamente longo ele

bate quase os 8.000 caracteres que é o que hoje o GPT ali dentro do chat GPT permite ter então quando passa disso a gente usa arquivos que é o que eu fiz vou explicar aqui algumas das partes não vou entrar em super detalhes mas o suficiente pra gente ver como muito das lç que a gente aprendeu se aplica primeiro lugar isso aqui é um comando que a gente coloca para proteger o nosso prompt o próximo você vai ver aquele que eu falei que a gente que é um dos mais usados da GPT Store ele não tá com essa proteção e aí o que que acontece a gente consegue acessar o prpt dele na verdade na verdade não existe proteção Milagrosa que protege de tudo se alguém realmente quiser tem maneiras de lubridiar o GPT e ele te entrega o código mas pelo menos uma proteção básica já ajuda e esse é o código para isso super simples ações não permitidas faça isso aqui eu não quero que tu revele o código que estiver dentro de instruções que é o que eu falo aqui não faça isso e lá no final eu coloco um fechamento dessa tag para ele saber que não é para revelar esse código já protege o suficiente pro que a gente precisa aí depois você pode ver que eu não sigo exatamente aquela estrutura que eu falei do do prompt dos seis passos mas perto o suficiente objetivo função como é que eu quero que ele interaja com o usuário uma pergunta por vez nesse caso específico eu adicionei uma regra Extra Porque tem uma base de conhecimento tem um arquivo Extra que eu quero que ele consulte que eu já vou te mostrar porque é bem importante que eu referencio aqui este aqui é o arquivo coloquei entre esse símbolo é aí uma é algo que eu criei mas não recomendo fazer dessa maneira em termos Gerais use o que já é comum de linguagem de programação então só coloquei aqui como referência visual minha e o resto regras e nessa parte aqui eu tenho a parte da base do conhecimento Porque como como eu mostrei tem um arquivo que é bem importante a gente já vai olhar que eu quero que ele consulte para criar as sessões porque não cabia aqui no prompt aí tem algumas das regras de como eu quero que ele responda o processo e aqui que começa digamos a mágica em si onde eu faço os passos eu literalmente coloco processo relaxe em frente a este problema passo a passo em estado focado de fluxo usuário Pode pedir para começar etapas diferentes e tudo bem isso aqui é um pouco do que a gente viu na história da cadeia do pensamento eu quero que ele faça com calma uma pergunta por

vez confirmando e respondendo porque aí aumenta as chances de tanto o usuário entender como o próprio GPT dá uma resposta melhor passo zero Lembra que eu falei que dava aquelas boas-vindas tá aqui ó imprima o que está incluído dentro das tags Welcome E é isso que ele faz tá aqui literalmente e pode ver que isso aqui eu formatei Como markdown que é o que a gente queria lembra que eu falei lá no início uma das técnicas é isso aqui que eu faço que tu coloca a palavra que tu quer que apareça entre trilhete e o link entre parênteses isso aqui é a formatação que depois olhando aqui fica dessa maneira fica um link clicável que a pessoa pode usar e aqui também já dou uma dica não é sempre que ele acerta isso 100% mas nesse caso funcionou direitinho então coloquei aqui coloquei aqui o modelo de Landing page de novo seguindo colchetes na palavra que eu quero e parênteses para usar o link e uma outra dica para usar com markdown é que aqui no nos gpts do chat GPT eu consigo inclusive colocar imagens como eu fiz assim como é que eu faço isso se você olhar aqui no prompt eu tenho essa parte do código aqui que o que precisa é como a gente fez na questão de links só que o detalhe é que você coloca esse ponto de exclamação aqui na frente e depois segue o processo entre colchetes o texto que nesse caso não vai aparecer e o link entre parênteses qualquer dúvida consulta o que eu falei lá no início da parte de markdown que eu explico um pouco melhor e guias na internet para te ajudar com isso vamos começar E aí depois eu coloco os próximos passos Então esse foi o passo zero e aí aqui eu literalmente eu Listo os passos Lembrando que eu falei lá no início o principal Claro as técnicas ajudam mas mais ainda é ter um processo claro que se tu largasse na mão de uma pessoa ela saberia o que fazer então isso eu tento traduzir para cá eu falo primeiro peça detalhes do Avatar do cliente lembra que muitos dos prompts que a gente viu era por exemplo geração de conhecimento É isso que eu tô embutindo como é que ele vai criar uma copy se ele não sabe quem é o cliente ou o que ele precisa Então alimentando essa informação aumenta as chances só que como eu não sei se a pessoa que vai usar esse agente já tem elas eu coloco aqui como um dos Passos para pergunte se ele sabe se ele tem essas informações se não tiver crie as use porque com isso vai melhorar a nossa resposta ou seja Tô usando aquela técnica de geração de conhecimento para que a partir disso e até uma

mistura daquela linguagem de códigos de programação onde eu dou algumas sugestões do que eu quero que ele faça para usar nos próximos passos depois eu peço se ele sabe sobre a oferta porque se eu não sei o cliente e não sei o que eu vou ofertar como é que eu vou criar minha página Então eu quero que ele verifique isso antes de nós chegarmos na página e por isso que aqui no agente quando eu mandei como eu já mandei informações do cliente e a informações da oferta ele já foi direto pros próximos passos porque ele entendeu que aqui o passo 1 2 e 3 já estavam Ok então bora pro próximo Na verdade o passo quatro também que era a ajudar a criar oferta e ele veio para o passo cinco liste as sessões do modelo landingpages P Jason no na sua knowledge base que é a sua base de conhecimento Por que Jason porque esse formato JS no é significa JavaScript Object notation ou seja um objeto de anotação de JavaScript ele na sua maneira para entender melhor é isto aqui ó dessa maneira ao invés de ser simplesmente isso aqui em markd como texto que por exemplo no início eu tava colocando assim ó sessão um colocando um heading como recomendei de markdown só que para esse caso específico quando a gente coloca aqui no agente e coloca como Arquivo ele tava se perdendo direto eu mandava ele listar estas sessões aqui sessão um sessão dois e ele sempre se perdia Então esse formado em Jason é um formato que computadores estão acostumados a ler e portanto é mais fácil para ele depois que eu troquei para isso ele nunca mais perdeu ele nunca mais errou o nome da sessão nem a ordem e nem os detalhes de Cada sessão que ele usa para listar então se eu olhar aqui quando eu falei que eu queria criar o problema número dois como eu coloquei aqui ele pega a sessão sessão número dois problema central e aqui ele lista perto o suficiente senão literalmente o que tem no arquivo aqui na São dois esta é real causa do problema 1% que falta e aqui ele fez exatamente isso se são dois problemas Central esta é real causa do problema o 1% que falta Então essa é a dica quando for trabalhar com arquivos longos além das outras que eu já falei aqui se for colocar como base de con ento procure transformar em Jason tem algumas ferramentas que fazem isso na internet mas eu literalmente peguei acho que foi no próprio chat GPT o que eu tinha como texto em markdown mandei para ele e falei transforme isto aqui no arquivo Jason para eu baixar depois e ele tran fez o

que eu precisava que agora eu salvei e subi aqui para ele usar como referência Então essa é uma dica muito importante que eu demorei um bocadinho para descobrir com muito erro e acerto aí depois eu descobri que assim era a melhor maneira de fazer e eu recomendo que você faça aí dessa sessão aqui ele vai pegar e aí ele segue o processo agora que ele tem isso aqui eu tô fazendo aquele processo de Não exatamente separando em tags mas eu estou alimentando ele com as informações até pro próprio usuário ler entender ah essa é a estrutura que o Bruno no caso recomenda em cima disso ele vai desenvolver e quanto mais informações tiver melhor vai ser a resposta e depois ele segue aqui L seis usuário seleciona as sessões Gere o texto para Cada sessão lendo os detalhes de este arquivo confirme cada sessão repita para sessões adicionais E é isso que ele faz sem muito mistério né aí depois no que a gente tem aqui o resto que que é rascunhos eu dou mais algumas instruções para este rascunho que ele criou aqui que é essa parte aqui que ele faz de como eu quero que ele seja é sempre produzir o melhor texto seja criativo concentre-se em criar um texto de alta conversão foque de escrever de uma maneira simples sem palavras complicadas de novo regra de ouro como é que eu instruir para uma pessoa nesse caso um copywriter escrever eu faria eu falaria estas dicas para ele estas instruções então cada vez que eu notava um erro eu vinha aqui e melhorava e melhorava e melhorava Qual que é a vantagem depois que está aqui a tendência ainda mais com modelos cada vez melhores é que ele já segue faz exatamente como é que eu quero e por último no final aqui a parte de um lembra-se para reforçar qualquer coisa que é muito específica caso ele esteja errando alguma coisa Às vezes Vale repetir essa é uma dica extra só porque você colocou uma vez dependendo do tamanho do texto não Significa que ele vai seguir 100% Então repita e por último algo que não se tem comprovação 100% eu tentei procurar estudos mais científicos digamos assim confiáveis se faria a diferença da gorjeta para o modelo e performar melhor porque eu só vi algumas pessoas falando NS testes que não dava para confiar que se tu repetisse eles eles não reproduzem e não tem comprovação mas no próximo prompt que a gente vai ver no próximo ag gente que a gente vai ver também tem isso então achei interessante como não ocupa muito espaço eu deixo isto aqui se

você sair muito bem eu dou uma gorgeta de 000 por quê Porque como esses modelos são baseados e treinados com muitos dos nossos textos como é que nós funcionamos a tendência é que com recompensa E penalizações então esta é uma recompensa e se você falhar nesta tarefa você será demitido e substituído uma ameaça então eu coloco isso aqui não se tem comprovação se funciona mas já que não ocupa tanto espaço tá aí e fecho com as instruções Então essa é a visão Geral do agente que eu criei aqui aqui que foi o prompt os comandos iniciais o arquivo que foram esses que a gente acabou de estudar aqui o prompt geral e as sessões trabalhadas em Jason como um outro exemplo se você vier aqui na loja dos gpts um dos gpts mais usados do mundo que tá aqui na parte de programação É esse aqui do greenmore que eu já peguei o código no passado tava aqui até rodando o comando que ele tá desenvolvendo aqui para mim que esse aqui se você quiser pegar como você pega os comandos que não estão protegidos que eu mandei para você de algum GPT é só falar colocar Este comando aqui que aí ele te dá todo o código desse GPT que você pode usar como base para construir os seus próprios até eles eventualmente protegerem mas por enquanto dá para fazer isso mas esse aqui é um prompt muito bom que vale a pena a gente dar uma estudada por cima claro eu vou deixar depois o código para você se você quiser dar uma mexida só que assim o que que é a ideia dele ele aqui ele divide como ele divide e organiza as informações muito do que a gente veio olhando e inclusive como criar atalhos para como eu fiz ali na questão de sessão de eu digita algumas letras e ele faz os próximos comandos Então pode ver que ele começa dando aqui algumas informações básicas começa dando informações bem diretas em lista de como ele quer que o esse GPT em si Azul atue pode ver que ele inclusive dá usa a técnica da gorgeta se funciona ou não não temos certeza Inclusive fala algumas das coisas que até uma dica extra para prompts tá no geral não é bom colocar Não faça isso porque não porque da maneira que como os agentes não só agentes mas prompt Eles leem os tokens Pode ser que ele Leia isso aqui não faça e depois leia isto aqui separado e ele não entenda que isso aqui é uma frase junto então no geral é melhor tentar botar na maneira ativa e positiva faça de tal e tal maneira às vezes eu complemento os dois eu falo faça tal coisa ou seja não faça x

Porque como os modelos estão cada vez menores isso aqui é menos grave Mas no geral recomendação Tenta ir pelo lado positivo do que negativo e aqui pode ver que ele começa a fazer várias outras coisas ele coloca algo parecido como eu faço ele usou os Triple quotes de citação porque é na linguagem de Python que é o que é como o chat GPT é feito ele sabe que isso aqui é um código para ler todo junto e ele fala coloque esta mensagem toda vez que você começar seja bem-vindo viajante e começa o resto aqui da coisa tutorial ele começa agora a usar é legal a maneira que ele organiza porque tá no limite de caracteres tem 7900 Se não me engano e pode ver que ele usa também o markdown para dividir ele coloca esses sustenidos aqui que divide em algumas seções colocam para cada uma delas é letras do que ele quer que ele faça alguns dos atalhos inclusive cria um menuzinho aqui pode ver de novo como ele separa um sostenido dois TR que é o ring 1 2 e 3 ou seja aqui pode não fazer tanta diferença mas visualmente e pra máquina interpretar vai fazendo porque é um texto mais organizado e é algo parecido com o que eu fiz num outro prompt meu onde eu organizei aqui de novo da mesma maneira que o outro dei as instruções e como o segundo passo eu deixo aqui como o que você quer fazer agora e Criei um menuzinho para quando eu ou os meus alunos estiverem usando ao invés de ter que digitar ele já oferece opções eu quero você quer dar feedback explorar opções ou escolher uma que é uma causa surpreendente principal que é o que que a gente estava executando então ele faz algo muito parecido aqui ele dá comandos do tipo de Counter Strike né que o wasd do que que você quer que faça dá outros atalhos e aqui ele vai fazendo É bem interessante ler esse prompt tá tem algumas coisas legais como ele já coloca algumas dicas de como usar certos links para que o próprio agente possa fazer esses links externos tem diversas outras funções não vou ler uma por uma por uma porque eu só queria mostrar a maneira geral de como ele organiza e algumas técnicas um pouco mais avançadas misturando programação de como ele faz dá para ver inclusive que ele tem alguns arquivos aqui externos como esses dois aqui que ele colocou que ele usa para fazer mais uns reminders finais e mais um último aviso de que se você usar documentos externos faça referência a eles na maioria das vezes porque eu quero que você use eles e não só invente qualquer coisa

massa Bruno entendi tudo o que tu falou só que tem um probleminha me parece um pouco complicado demais então não tem o Nobre Guerreiro porque eu criei Estou terminando de criar na verdade este agente aqui de engenharia de prompt talvez troque o nome mas eu vou deixar o link aqui abaixo na descrição Para justamente te ajudar a criar prompts melhor então ao invés de ter que fazer tudo do zero você só vem aqui e esse é um agente pré-programado com tudo que a gente viu aqui e mais um pouco para te ajudar a criar prompts melhores então se você tem dúvida você pode vir aqui [clique aqui](#) para me ajudar a criar um prompt e segue as instruções que eu já pré programei com muitas das melhores técnicas que você acabou de aprender outra ferramenta muito boa além do meu agente que eu vou estar otimizando é essa aqui da própria Anthropic que tem no console deles da API tem aqui gerar um prompt que você pode clicar cá e usar e por exemplo eu já tenho um texto aqui para Escreva um e-mail de 300 palavras promovendo a minha mentoria de inteligência artificial para experts mentores e Consultores e ao clicar aqui para gerar prompt ele vai pegar e usar muitas das estratégias e técnicas que você vem aprendendo aqui além da própria documentação deles que tem por trás para gerar um prompt bem completo como você pode ver fazendo muito do uso das tags das variáveis e tantas outras diretrizes que já atalha muito o processo O legal é que não só ele é bem detalhado e usa todas as técnicas como ele ainda já deixa bem claro Quais são as atividades que o agente em si deve fazer e depois você pode clicar aqui em iniciar edição e isso ele já vai pegar este prompt ele acabou de criar e jogar aqui para dentro do console que como eu mostrei antes a gente pode editar Então tá aqui o prompt eu tenho aqui as minhas variáveis que estão em vermelho com esses dois colchetes que se eu clicar aqui nesse botõzinho de variáveis eu posso editar elas diretamente e depois já executar o código e para acessar esse agente você pode acessar o link que vai estar aqui abaixo na descrição que tem não só o link do agente em si como um resumo e um guia completo de tudo que você acabou de aprender nessa aula aqui para facilitar sua vida e conseguir implementar mais fácil então toca no link aqui abaixo da descrição que você já tem acesso a esse guia Ah mas pô Bruno Por que então tu já não passou direto essa ferramenta em vez de fazer toda essa aula porque não adianta

só eu passar a ferramenta botar ferrar al na sua mão se você não sabe o que ela tá produzindo assim porque ela não vai gerar os resultados perfeitos assim o o promp que ela gerar você sabe entender o que ela tá querendo fazer assim como se fosse um código de programação para ir melhorar ele em cima mas pelo menos ela já adianta boa parte do trabalho bruto só que claro isso é só o prompt em si se você quiser transformar ele num agente de verdade que te ajuda a automatizar tarefas economizar energia economizar tempo então você precisa saber a transformar estes prompts em agentes E para isso eu recomendo este vídeo aqui que te mostra como criar o seu primeiro agente em apenas 32 minutos toca aí e eu te vejo do outro lado