



Daqui, chegamos na aula mais importante deste módulo. Daqui pra frente é só pra trás, não tô brincando. Essa aula é muito importante porque, infelizmente, ela vai filtrar muitas pessoas. É uma aula que, se você não. Souber lidar e entender, você não vai. Conseguir seguir no N8n. E a gente vai combinar assim, antes de te aterrorizar. Eu quero que essa aula você só preste atenção, não vai fazer nada junto. Comigo, e a próxima aula a gente vai fazer junto. Essa é uma aula que a gente. Vai fazer requisições para todas as LLMs. Que a gente utiliza aqui na academia. A gente tem aqui requisição para o Chat EPT, requisição para o Cloud, requisição para o Gemini, requisição para o Gurok, requisição para o Perplex, requisição para o DeepSeek. Então, você vai aprender como utilizar todas. Essas requisições sem ser no Node padrão aqui. Se você digitar chat.ept, OpenAI, tem os Nodes da OpenAI. Se você digitar aqui, Anthropic, tem o Node da Anthropic. Se você digitar, tem Perplexity aqui? Perplexity, não tem. Então, assim, Grok, por exemplo, não tem. Então, você vai aprender da forma mais profissional, a fazer as requisições sem precisar desses modelos prontos, tá bom? Vamos fazer o seguinte, eu quero que. Você aprenda, entenda o conceito, o fundamento. E não a ferramenta. A ferramenta pode mudar, mas o conceito vai permanecer o mesmo, o fundamento vai permanecer o mesmo. O que vai acontecer aqui? Eu vou fazer seis requisições de texto, porque aqui eu estou utilizando texto, e dois tipos de aquisição de imagem a. Gente vai gerar imagem com o chat. E pt e com o grok então eu vou falar já é uma imagem um gato vai chegar uma imagem um gato para o nosso whatsapp tanto pelo chat e pt quanto pelo grok certo então vou primeiro mostrar para você o funcionamento do primeiro tá vou ligar aqui. Ó esse aqui nesse aqui o que. Vai acontecer eu vou mandar uma pergunta. Chamada qual o sentido da vida. Então, vou botar aqui para testar. Lembrando, você não precisa fazer nada agora. Perguntar aqui, qual o sentido da vida. Para o meu WhatsApp, tá? Vou mandar, vai começar a rodar. Recebeu. Vamos dar uma olhada aqui no nosso WhatsApp. O que chegou? Bom, já foi enviado, então aqui eu mandei qual o sentido da vida, ele já mandou o sentido da vida, uma questão complexa, profundamente pessoal, variando de acordo. Com as crenças, valores e experiências individuais. O que eu fiz? A mesma requisição HTTP request que a. Gente aprendeu na aula anterior. Então, como está feita esta requisição? Presta bastante atenção, quero que você aprenda o formato de requisição. Você não precisa decorar nome, código, só o formato, só a anatomia da requisição. Percebe? Nós temos uma URL, que essa URL é o endereço padrão para fazer requisição. Para esse tipo de resposta em texto no chat APT. Nós temos o header, o header é um authorization, e tem barrier e a API da OpenAI. Depois, nós temos um body, com o. Nome do modelo, que é o gpt4-omni, e o input, que é a entrada. O que ele vai processar? Nesse caso, ele está processando a minha mensagem. Isso aqui é uma variável. Está vendo? Eu perguntei qual é o sentido da vida, e aqui aparece qual é o sentido da vida. Vou fazer a mesma coisa, só que. Agora para Anthropic, para o Cloud. Vou mandar aqui, vou conectar no Cloud. Novamente, vou test your flow. E vou perguntar qual o sentido da vida pra você,

Clode? Mandei. Ele já tá processando. Tá vendo? Tá rodando aí. Vamos ver essa mensagem chegando aqui, ó. Pronto, acabou de chegar. Como uma IA não tem experiências pessoais. Ou consciência para desenvolver um sentido próprio da vida, experimento alegrias, tristezas e buscas existenciais como humanos. Beleza. Vamos fechar aqui. Vamos dar uma olhada na requisição do Cloud. Vou abrir aqui. Então, no Cloud, a gente tem o URL, que é o endereço quando a gente quer fazer requisição de texto. Percebe? Tem messages aqui. E aqui no chat EPT tem responses. Então, muda de um para o outro. No formato, na anatomia, a gente tem o URL. A gente tem o headers, que aqui. Diferentemente do chat EPT, Aqui está xAPIKEY. Lá era authorization, lembra? Então, vamos dar uma olhada aqui. Lá era authorization e API com esse barrier na frente. E aqui é xAPIKEY, só com API, sem o barrier na frente. Ainda tem no header o nome, que. É Anthropic Version, o value, que é 20230601. Isso é padrão, é estático, não vai mudar, tá bom? Não se preocupa que a próxima aula a gente vai fazer passo a passo cada uma dessas requisições. E aqui o JSON, que é a parte do body, percebe que é um pouquinho diferente. Então, eu tenho aqui um modelo, eu tenho a quantidade máxima de tokens e a mensagem. A mensagem é de um usuário. E o prompt do usuário é a mensagem que eu mandei, tá? Então, aqui ó, user significa é um. Prompt de usuário, ou seja, um prompt de entrada. E aqui o content é o conteúdo enviado por esse usuário. A única coisa que a gente muda. É aqui em content. O que eu poderia mudar também? O modelo, se eu quisesse mudar por um modelo 3.5, e aí eu teria que escrever exatamente da forma que a empresa escreve o modelo. Eu não posso simplesmente colocar aqui Cloud 3.7. Ele não vai funcionar. Existe a forma escrita de modelo na documentação. E, mais uma vez, na próxima aula. A gente vai ver isso. Máximo de tokens seria a quantidade máxima de tokens que ele pode gerar. Eu posso mudar aqui. Se eu botar 10 tokens, vamos ver. O que vai acontecer? Vou mandar novamente. Vou dar aqui um test workflow. E vou mandar aqui. Qual o sentido da vida? ele processou como é 10 tokens ele. Processou bem rapidinho só que tem um. Problema eu limitei 10 tokens o sentido. Da vida é uma questão pro 10. Tokens ele chegou no limite então eu. Limitei ele 10 tokens o que é muito pouco então assim percebe que qualquer ajuste que a gente faça modifica totalmente a resposta vamos colocar aqui novamente. E 24 tokens, só para continuar os testes. Entendeu que até aqui cada LLM vai ter uma requisição? Vamos fazer com o Gemini agora? Vou conectar no Gemini, vou colocar teste workflow, vou perguntar qual o sentido da vida. Qual o sentido da vida, Gemini? Ele vai processar, e aí já vou. Abrir aqui para a gente dar uma olhada. Beleza, perguntei aqui. Nossa, a resposta é gigante. Qual o sentido da vida... Ah, está aqui. Qual o sentido da vida de Gemini? Está vendo? E aqui a resposta de Gemini. Essa é uma pergunta de um milhão de dólares, não é mesmo? Uma pergunta que filósofos, cientistas, indivíduos como. Você e eu têm feito há séculos. Mais uma pergunta de um lendário aqui. Então, resposta do Gemini. Vamos dar uma olhada na requisição do Gemini? Requisição do Gemini, URL do Gemini, endereço padrão. Percebe que, diferentemente dos outros, ele tem. Um query parameters, que é onde coloco a key e a API do Gemini. Então, aqui a API fica no header, aqui a API fica no header, mas aqui no Gemini a API fica no query parameters. E depois, headers não tem, tem o. Body, que é bem parecido com o Cloud, tá? Então, tem contents, paths e texto. Essa é a forma de requisição do Gemini. Não se preocupe, essas requisições você não vai digitar no dedo da mão, não vai precisar codificar, tá bom? Vamos fazer agora para o grok. E agora eu

vou só fazer e. De forma muito rápida. Qual o sentido da vida? Qual o débito? Sentido da vida? Grok. Mandei aqui para o Grok. Ele já processou. Vamos dar uma olhada na resposta dele. Aqui, qual o sentido da vida, Grok? O sentido da vida é 42, claro, mas falando 42. Mas falando sério, essa é uma pergunta profunda e filósofas, pensadores 42 sentido da vida é 42. Eu acho que alguma piada que eu, Zé, não peguei. Deve ter alguma piada por aqui. O sentido da vida é 42. Curiosidade, você vai matar junto comigo. Resposta 42. Questão fundamental de vida do universo e de mais é uma referência do livro O Grito do Michelle. Ah tá, ok.

Vamos dar uma olhada aqui no prompt dele. Aqui no prompt, role system, you are a test assistant. O Grok foi treinado como parte do livro, como é o nome do livro? O Guia do Mochileiro das Galáxias. Então, por isso que ele traz essas respostas. Ele tem uma forma de interação mais divertida. Vamos dar uma olhada na requisição do Grok, URL. Não tem Query Parameters, tem headers. E aí, é parecido com o chat EPT. Authorization, Berry, a chave de API, e. No JSON é mais parecido com o Cloud. Então, eu tenho aqui messages, whole system. Aqui no content é o prompt do sistema. É o prompt onde eu coloco, você é um professor de matemática, filósofo... E. Aqui no user eu coloco o prompt do usuário. Então, qual vai ser a entrada do usuário? E aí, eu tenho o modelo aqui. String false, temperatura zero. Não vamos mexer nisso. Então, percebe que, para a gente finalizar. A requisição para o perplexity vai ser igual, para o deepseq vai ser igual, o que muda é só o formato. Tá vendo aqui? O formato de requisição, porque cada um tem um formato. O que você tem que saber? A requisição HTTP request tem um padrão, que é esse padrão aqui. Ela tem uma URL, ela tem um query parameters, ela tem um header e. Ela tem um body. Essa é a anatomia do HTTP request, que é a ferramenta que a gente mais utiliza no mundo de inteligência artificial. Não especificamente de automação, sempre diferenciando. Para IA, a gente usa muito ferramenta de HTTP request. Vamos imaginar que a gente tem uma carta. Essa carta tem um endereço chamado Rua dos Bobos número 0, que era uma casa muito engraçada, Ela tem entregar somente para o João, ou seja, só vai ser entregue essa carta se for para o João. Ela tem o remetente, que é de José Carlos Amorim. E ela tem o conteúdo dela, que é... Fala João, como está a Maria? Espero que esteja tudo bem por aí. Soube que um lobo anda aterrorizando a vizinhança. Se cuidem. O formato dessa carta é... O URL é sempre o endereço de onde a gente está tentando acessar. Query parameters é sempre observações para essa entrega. Então, aqui nesse caso, entregar somente para o João, entregar somente em dias úteis. É o query parameters. Dificilmente os query parameters são utilizados no HTTP request nas LLMs. header quase sempre tem, com exceção do. Gemini, como a gente viu, que é o remetente. Então, eu preciso no remetente sempre colocar ali um padrão e a minha chave de API, que é para me certificar. De que eu sou o remetente. O remetente é a chave de API... Aquela sopa de letrinha lá. E o body é o conteúdo. É o conteúdo que vai fazer parte dessa requisição. Então, percebe que aqui... Vamos dar uma olhada em cada parte. Nós temos aqui o RL, que é o endereço. Não temos query parameters, nós temos o. Header, que é o remetente. Então, authorization, remetente é esse aqui, que é a minha chave de API. Então, é a chave de API do. Zé, o remetente aqui é o Zé. E o body é o conteúdo. Então, eu tenho aqui de conteúdo o modelo, que é o GPT-4 Omni, e um input. Cada LLM vai ter um estilo de body content type. Isso daqui você não precisa decorar, porque vai ser leitura de documentação, que na próxima aula a gente vai fazer. Então, vamos ver aqui no Cloud. Qual é o URL do

Cloud? Essa daqui. Qual é o headers do Cloud? É esse daqui. Headers é remetente. Então, XAPI aqui, esse aqui é o zero. Eu sou o remetente. E aqui também tem do remetente, Anthropic Version 2030601. especificamente para o Cloud tem essa diferenciação. Então, não é decoreba. Não tem como você decorar. E é por isso que aqui a gente filtra muita gente que é entusiasta e que vai ter um limite cognitivo. Isso é natural. A vida, ela filtra quase sempre. E não tô dizendo que você não vai conseguir se você não tiver capacidade cognitiva. O que eu tô dizendo é que, para algumas pessoas, esse conceito vai ser mais difícil. Para outras pessoas, não. Vamos fazer uma requisição de imagem? Olha só, vamos primeiro fazer uma requisição. De imagem do chatpt. Vou conectar aqui. Vou tirar esse pin. Tá vendo que tem um azulzinho aqui? Vou tirar o pin. Então, vamos lá. Vou dar... Aqui é geração de imagem. Então, vou pedir. Gerem uma imagem de um gato sapeca. Fez a requisição, está processando. Processamento de imagem geralmente demora um pouquinho mais. Então, aqui especificamente está rodando o modelo Dolly 3. O Chat EPT, neste momento, lançou uma versão nova de imagem que está absurdamente. Boa e ainda não tem na API. Vamos dar uma olhada aqui. Como está? Ele está enviando para o Evolution. Pronto, chegou aqui. Imagem de um gato sapeca. Essa aqui é a imagem, tá bom? Vamos agora pedir do grok essa requisição. Vou só esperar finalizar, percebe que já chegou, mas às vezes dá um bugzinho no envio. Pronto, finalizou aqui. Agora vou fazer o seguinte, vou excluir aqui, vou vir pra cá, e vou. Conectar aqui no grok. Mesma coisa, vou pedir pra gerar a. Imagem de um gato sapeca. Gere uma imagem de um gato sapeca. Pronto, mandei. Vamos abrir aqui para a gente ver a requisição sendo processada em tempo real. Percebe que geração de imagem, geração de áudio, tudo aquilo que exige um pouco. Mais de processamento demora um pouquinho. Nossa, que gato estranho. Veio com um bigodinho aqui. Sei lá, é uma minhoca? Enfim, veio assim. Veio assim a imagem, tá? O que eu quero mostrar para você é o seguinte. Vou tirar aqui. Como é a requisição de imagem do Chat EPT? Tenho aqui, ó. O endereço é diferente. Aqui eu tenho V1 imagens de Narrations. Aqui eu tenho V1 Responses. Então, percebe que o endereço modifica totalmente. O que você está solicitando. Headers é o mesmo. Authorization e o Barrier, ou seja, é o mesmo remetente. Esse aqui é o Zé. Eu quero gerar imagem. O body, o body muda um pouquinho. Então, eu tenho o JSON, modelo dolly-e3. Vamos dar uma olhada lá na requisição de texto. Na requisição de texto, modelo gpt4-omni. gpt4-omni-llm. O dolly-e3 é modelo de imagem. Outra parte da requisição é o prompt. Então, o prompt gera uma imagem digital. Lá no texto, vamos dar uma olhada. Aqui é o prompt também, que é input, o nome. E lá, como é que é o nome? É prompt. Então, na mensagem é input e na imagem é prompt. Vamos dar uma olhada na requisição do grok. Do grok, images generation. Vamos dar uma olhada na requisição de texto. Aqui do grok, chat completions. E aqui, images generation. Vamos dar uma olhada no headers. O headers aqui chama accept application.json. Tem um outro header chamado authorization e o berry API. Esse é o formato padrão do grok. Vamos ver o body modelo grok2-imagem. Vamos dar uma olhada na requisição do modelo de texto. Aqui não tem. Já tem direto. o json, aqui o modelo, tá bem. Aqui, o grok2-latest e aqui já modifica Aqui já é grok2-imagem. Lá era grok2-latest, aqui é grok2-imagem. E o prompt, que é a imagem. Que vai ser gerada. Então, muito importante que nesta aula você entenda que a requisição tem um formato. E esse formato de requisição vai mudar. De LLM para LLM. Para a gente finalizar, aqui no Evolution eu

coloquei também o que modifica quando você envia um texto e quando. Você envia uma imagem? Quando você envia um texto, a URL é sendText e o body é parâmetro, o body parâmetro é number e text. Vamos dar uma olhada dentro aqui. O que muda? sendText, tá vendo aqui, message, sendText. E eu preciso do número e preciso do texto. Beleza. Todos esses daqui que geram texto é assim. Todos são padrão. Mas aqui é diferente. Por quê? Porque eu vou enviar uma imagem, uma mídia. Pode ser áudio, pode ser imagem, pode ser vídeo. Então, percebe que aqui a URL acende mídia. Aqui a URL acende text. Vamos dar uma olhada? Aqui, ó. Message, SendMedia. E lá é Message, SendTest. Vamos dar uma olhada aqui no que é enviado. Número. MediaType, diferentemente do outro. MediaType, Image, que é padrão, estático. E a mídia, que é o resultado do que é gerado pelo ChatEPT. Então, percebe que é diferente. Então, aqui eu tenho Number, MediaType, Image e Mídia. Lá eu tenho apenas number e text. É muito importante que você entenda essas diferenças de requisição, porque percebe que todo mundo, toda a geração, toda a conexão é feita por meio desses códigos. Quando você acerta o código, você consegue. Conectar uma ferramenta com a outra. Na próxima aula, a gente vai gerar isso passo a passo para você entender como ler a documentação e como fazer. Isso de uma forma mais intuitiva.