**Como fazer um projeto prático com Langchain, Chatgpt, Python, RAG**

**Transcrito por** [**TurboScribe.ai**](https://turboscribe.ai/pt/?ref=docx_export_upsell)**.** [**Atualize para Ilimitado**](https://turboscribe.ai/pt/subscribed?ref=docx_export_upsell) **para remover esta mensagem.**

Quer ver uma aplicação prática de Learning Chain? Fica comigo nesse vídeo. APLICAÇÃO PRÁTICA DE LEARNING CHAIN Oi, pessoal. Aqui é o Rafa.

Esse é o quarto vídeo da série de Learning Chain. E, dessa vez, eu já trouxe para vocês os conceitos básicos. A gente falou sobre o que é a biblioteca Learning Chain, a gente falou sobre as chains, falou sobre RAG, falou sobre agentes.

Agora, nesse vídeo, o que eu vou fazer com vocês? Vou pegar esses conceitos e torná-los de maneira prática. Eu vou criar um projeto do começo ao fim. Claro que não vai ser nesse vídeo.

Nesse vídeo, eu vou explicar para vocês qual é o projeto. E, ao longo dos próximos vídeos, a gente vai construindo esse projeto passo a passo. Eu não terminei esse projeto ainda.

A minha ideia é que vocês vejam todo o processo criativo. Então, eu vou mostrar para vocês como eu penso quando eu estou prototypando uma ideia, como eu vou testando e fazendo essas hipóteses. E a ideia é trazer aqui para vocês um relato de vida real.

Todos os erros e acertos que teve no projeto, todos os códigos que eu fiz, como eu pensei a arquitetura do sistema, como a gente vai evoluir nesse projeto do início ao fim. E o primeiro ponto é a escolha do projeto. A escolha do projeto é uma coisa que eu sempre falo para vocês.

Eu vou repetir mais uma vez aqui. Vocês têm que começar por coisas que vocês dominem, coisas que façam sentido para vocês, coisas que vocês gostem. Não adianta você estar aqui me assistindo e querer aplicar IA generativa em finanças se você odeia finanças.

Ou se você não gosta de trabalhar com marketing, deixa eu tentar aplicar uma coisa em marketing porque tem uma vaga de marketing, porque dizem que o portfólio de marketing é melhor. Esquece as aplicações de mercado. Pense em coisas que vocês dominem, em coisas que estão no seu dia a dia.

Se você tem um hobby, por exemplo, se você gosta de futebol, faça uma coisa relacionada a futebol. Se você gosta de games, faça alguma coisa relacionada a games. Pegue coisas que sejam úteis do seu dia a dia.

Se você, por exemplo, tem um animal de estimação, faça alguma coisa para reconhecer a carinha do seu animal de estimação. Faça sempre coisas relacionadas ao seu dia a dia. Há uma coisa que vai ajudar você primeiro como indivíduo, como pessoa.

Porque você já vai ter metade do trabalho resolvido, que é a motivação. Projetos grandes precisam de muita motivação. E se você está fazendo uma coisa que você não gosta ou que você não tem uma interação muito legal, isso vai ser um desafio maior ainda, certo? Então, pensando nisso, eu vou propor um projeto que faz sentido para mim, para o Rafael.

Se você quiser replicar esse projeto, fique à vontade. Se fizer ele como um exercício didático, tudo bem. Mas a minha ideia é que vocês peguem o meu jeito de pensar, peguem as técnicas que vocês estão usando e repliquem em algo que faça sentido na vida de vocês.

E não tem nada que seja mais desafiador na vida de uma pessoa que tenha uma carreira profissional em paralelo com a criação de conteúdo que seja a criação de conteúdo. Pode parecer simples, vocês que estão me assistindo aqui, ver um vídeo de meia hora, de dez minutos, um tutorial, mas o trabalho que esse vídeo tem é insano, às vezes. Às vezes são quatro horas de gravação, uma semana de planejamento para sair um vídeo de meia hora, porque a gente tem que pensar em muitos detalhes, pensar na didática, pensar em como vocês vão receber isso.

Eu tenho muita experiência em dar aula em ambientes físicos, até online, mas com alunos no Zoom, real-time. E isso é uma grande diferença em relação ao YouTube, porque eu tenho um feedback real. Eu vivo brincando com meus alunos, tem que ver o cara de ué quando eu explico alguma coisa e eu tenho certeza que eles não entenderam.

Eu tenho esse feedback real-time, eu sei que eles não entenderam, então eu tenho que refinar a explicação, volto, faço perguntas, eu consigo trabalhar mais certos pontos da aula que eu sei que são mais complexos. Mas aqui no YouTube é um grande exercício de esquizofrenia. Eu tenho que imaginar o que vocês não vão entender, eu tenho que imaginar o que vai ser mais difícil e, a partir disso, lapidar melhor os conceitos, lapidar melhor a explicação e ter um feedback talvez uma, duas semanas depois e que já passou.

Por isso que é muito mais complicado fazer criação de conteúdo. E, pensando nisso, eu vou criar uma ferramenta para mim, para me ajudar a criar conteúdo aqui para vocês no YouTube. Eu já tenho uma ideia de como eu quero que essa ferramenta seja, eu já tenho algumas dores que eu sinto quando eu estou pensando em criar conteúdo e agora eu vou mostrar para vocês todo o meu processo de criação.

Lembrando que eu não sou engenheiro de software de formação, sou de engenharia de software por conta, porque eu gosto da área de tecnologia. Acho que é algo necessário para quem quer trabalhar com IA e Data Science. Não é necessário, mas se você tem essa skill, mal não vai te fazer.

Entender onde o componente da IA conecta com o componente da tecnologia é essencial. E eu quero dar esse filhinho aqui para vocês também nesse projeto. Eu vou fazer coisas, por exemplo, hoje eu vou mostrar uma parte de design, que é horrível o design que eu fiz.

Eu não sou design de formação, eu li algumas coisas aleatórias de design, não é uma área que eu gosto muito, mas eu preciso de um design para a minha aplicação. Então eu fiz algo simples, algo funcional para mim e, quem sabe, no futuro, acho um design profissional que me ajude a deixar esse processo mais interessante em termos de UI e UX, mas não é a fonte, não é o foco do vídeo. E sempre que algo não for o foco da tecnologia, eu vou falar para vocês, eu estou falando nessa parte de front, por exemplo, não é o foco, mas a gente precisa entender como essa coisa funciona, entender como vão ser essas conexões para entender o nosso lado, a nossa aplicação.

Sempre tem essa interação entre o time de engenharia, que toca front e back, junto com o time de IA e Data Science, que toca a parte de algoritmos e recomendação. Então é sempre bom a gente ter essa conversa. Numa empresa comum, você tem áreas diferentes com profissionais diferentes.

E saber andar por esses caminhos, conversar com esses profissionais é tão importante quanto você saber o algoritmo de IML. Você só vai crescer na carreira, só vai virar um profissional sênior, um staff, um profissional reconhecido, se você souber de alugar com esses pares, se esses pares estiverem seguros de que o seu trabalho vai se conectar com o trabalho deles de forma suave, tranquilo, sem surpresas. Então eu quero dar essa dimensão para vocês aqui nesse vídeo também.

Então quando eu estiver fazendo uma coisa que não é de IA, mas que impacta no trabalho do projeto, eu vou falar. Quando for uma coisa importante para IA, eu vou chamar a atenção de vocês também. E agora vamos lá conhecer o projeto que a gente vai fazer nesse vídeo.

E o projeto que a gente vai fazer está relacionado, como tudo na parte de negócios, a uma dor, certo? Sempre que você for pensar em projetos no mundo prático, na indústria, você tem que estar resolvendo um problema. Quando você está na academia, você pode se dar o luxo de focar em coisas teóricas, você tem que entrar, às vezes, em meandros que são curiosos, mas você não sabe necessariamente onde você vai aplicar. Mas quando a gente está falando de um projeto, mão na massa, de indústria, você tem que estar resolvendo uma dor.

E a dor, nesse caso, minha, que a gente vai estar resolvendo, é em relação a pensar em novos conteúdos relevantes. Então eu não estou aqui no meu notebook, dessa maneira que eu tento pontuificar e pensar. A minha dor é em relação a pensar em conteúdo.

Isso é uma parte até filosófica, ampla, e a partir dessa dor, eu vou criar uma ferramenta com o IA que vai me ajudar a sanar essa dor aqui. E, basicamente, como é que funciona hoje para fazer conteúdo? Tem uma parte que é dados, certo? Eu olho para o passado, olho para as estatísticas do canal, vejo quais são os vídeos que estão melhores, vejo quais são as séries que estão tendo mais retorno em termos de frequência, de views, em termos de horas assistidas, e a partir dessa análise, eu combino com o que pode estar mais ou menos parecido com isso. As pessoas gostaram de tutorial, então eu vou fazer mais tutorial.

Mas de quê? As pessoas gostam quando eu falo de tecnologias novas. Certo, mas onde eu vou achar essas novas tecnologias? Então, pensando nessa dinâmica de olhar para os dados e tentar fazer esse exercício de futurologia para achar o que dá um match legal com o meu canal ou não, eu vou criar uma aplicação onde eu posso, a partir dessa combinação de dados estatísticos que eu tenho à disposição no YouTube, com dados novos relevantes de outros canais, entender o que as pessoas estão falando em canais do planeta inteiro, eu vou criar uma ferramenta que vai me sugerir pautas para vídeos. Essa é a minha ideia.

Eu analiso as estatísticas do meu canal, a partir disso, eu vou pegar ideias de outros canais, entender o que outras pessoas estão falando de outras fontes, eu vou combinar essas duas fontes de informação e eu quero que minha IA sugira para mim pautas. Olha, Rafa, baseado no sucesso do seu canal, baseado nas métricas que você tem aqui para mim e baseado no que está acontecendo, no que a galera está falando na última semana, por exemplo, eu sugiro que você faça um vídeo DX, certo? Essa é a minha ideia de negócio. Note que aqui eu não tenho nada de IA, tá? Eu sempre gosto que vocês pensem que IA, assim como qualquer outra tecnologia, é uma ferramenta para resolver um problema de negócio.

Quando a gente está na indústria, fazendo mais ou menos esse paralelo, quando você está trabalhando na indústria, quando você está trabalhando com uma empresa, você precisa ter esse viés de negócio. Você tem que ser menos teórico e mais prático. Às vezes, uma regressão linear resolve a sua vida.

Não vai ser o caso aqui. Aliás, eu já pensei nesse projeto justamente para continuar na série de blockchain, LLM, então, provavelmente, as ferramentas que eu vou utilizar aqui para resolver esse problema estão atreladas a Long Chain e LLMs. Já foi um problema pensado para a ferramenta certa.

Mas na vida, geralmente, não é muito assim. A gente tem que lapidar, às vezes, um pouco do problema de negócio, o que as pessoas querem. Por exemplo, uma coisa que eu sempre falo que é essencial quando a gente está construindo um projeto novo é você pensar na métrica de sucesso, porque é muito fácil você se deixar levar pelo negócio quando é uma métrica abstrata.

Ah, o projeto acaba quando a gente tiver sucesso na empresa. O que é sucesso? É aumentar o faturamento em duas vezes, três vezes? É você diminuir o churn? Sempre crie uma métrica mensurável para os seus projetos. Por menor que seja, você tem que ter uma média do que é o final, o que você vai conseguir em sucesso.

Uma vez que eu atingi isso, acabou o projeto. Ou, se eu não conseguir atingir essa métrica, o projeto fracassou, volto para a média de projeto, entende o porquê. Mas sempre tem algo mensurável.

E aqui é mais fácil. O que eu quero aqui? Eu quero que a gente tenha uma possibilidade de eu colocar a URL do meu canal, ele vai fazer essa combinação e eu quero que ele me sugira temas relevantes. O que são temas relevantes? Temas que eu consigo pegar a parte desse tema e fazer um vídeo.

Então, se eu colocar o link do meu canal nessa minha aplicação, ele me deu ver um vídeo de culinária, por exemplo, eu acho que fracassei. Meu foco aqui não é culinária, se eu começar a fazer vídeo de culinária, provavelmente você vai parar de me assistir aqui. Muito embora eu adora cozinhar, não é o foco desse canal, você não está me vendo aqui pelas minhas habilidades como cozinheiro.

Então, se a aplicação conseguir a parte do link do meu canal, me dá ideias de conteúdos relevantes, 10 de 10, meu projeto está acabado com sucesso. E pensando um pouco também em mundo real, a gente nunca acaba um projeto, são etapas. Generalmente, a cabeça da tecnologia de desenvolvimento sempre tem features melhores para você desenvolver.

Ao longo do projeto, você vai descobrindo novas ferramentas, novas possibilidades, e vai criando outras possibilidades de expansão do projeto. E aqui vai uma segunda dica. Pessoal, diminui o escopo, deixa o escopo fechado.

É muito tentador aparecer outra ferramenta ou você descobrir outra possibilidade no meio do projeto, você tentar colocar essa feature também. Existem técnicas de Scrum, técnicas de Kanban, técnicas de gerenciamento de projetos, que eu não vou entrar aqui. Eu estou dando o básico para você sobreviver na vida, que é métricas mensuráveis e escopo fechado, certo? Vou até colocar aqui, a gente tem que pensar no conteúdo.

Vamos falar de métricas e escopo. Essas duas últimas aqui são essenciais para você não se perder no projeto. Eu vou fazer essa feature.

Entra no canal, dá o Enter, me dá a sugestão. Qualquer outra coisa que eu pensar no projeto, o que eu vou fazer? Eu vou anotar e vou colocar no backlog. Um belo dia eu volto e faço essa feature.

Acabei essa primeira parte do projeto, eu volto e faço ela. No dia a dia não é tão bem assim. A gente tem que negociar, às vezes, com o PM, negociar com os executivos da empresa, às vezes outra coisa passa na frente, mas sempre tenha em mente que se você quebrar o escopo e deixar o escopo bem definido, uma métrica mensurável, você vai conseguir avançar no projeto.

E depois que eu já defini essas diretrizes aqui, já pensei no que eu quero fazer, vou fazer algo para me ajudar a pensar em conteúdo, já defini a minha métrica, já defini o que eu quero que eu me escopo, geralmente eu começo a rascunhar o que eu quero de interação. De novo, não sou design, não sou front-end, conheço as técnicas, sei programar em React, sei programar um pouco de JavaScript, mas eu sei o básico para mim, não faço isso profissionalmente, não é a área que me ativa, mas eu sei o básico. Então, pensando um pouco nisso, eu geralmente pego uma folha de papel, como essa, e rascunho mais ou menos o que eu quero.

Então, aqui a gente pode notar que a gente tem aqui um box, um input box, onde a minha ideia é de estar... A princípio, eu vou fazer com o meu canal, mas não sei, de repente o projeto dá certo, no futuro eu posso talvez tentar monetizar esse treco, pensar em modelos de negócio. Por enquanto, estou fazendo algo para casa, estou fazendo aqui fundo de quintal mesmo para o canal acarreiar, mas nada me impede de fazer isso talvez para outros canais. Então, aqui eu quero que o meu input text, eu posso colocar uma URL do meu canal ou qualquer outro canal.

Aqui do lado, eu quero definir qual é a relevância do tempo, então eu quero, por exemplo, que a minha aplicação pegue ou 7, 28 ou 1 ano de conteúdo, não meu. Eu quero lembrar, eu quero comparar as minhas estatísticas com o que está acontecendo no mundo, e a partir disso me gerar insights, sugestões. Então, por exemplo, se eu flagar aqui os 7 dias, ele vai pegar o conteúdo dos últimos 7 dias para fazer essa comparação.

Então, eu digitei o canal aqui, setei a parâmetro de tempo aqui, e ele vai me devolver aqui, por exemplo, a priori, o título do vídeo, quantos views tem tal vídeo e um resumo desse vídeo aqui. Então, essa é a primeira parte que eu vi, são não vídeos do meu canal, tá? Eu quero que ele pegue vídeos correlatos, que não são do meu canal, mas que são relevantes para o meu canal. Olha só que coisa abstrata, eu não falei para ele sobre o que é o meu canal, onde vai entrar AI.

Aqui eu coloquei carreira AI, mas poderia ser, por exemplo, um podcast sobre luta livre, podia ser, por exemplo, um canal sobre culinária. E a partir da URL, só da URL do canal, eu quero que ele infira o que é o canal, o que é relevante para esse canal. Justamente porque eu estou pensando, já tenho algum tempo de carreira na área de tecnologia, e eu poderia, por exemplo, colocar aqui algumas tags e pedir para o usuário digitar o Flegato, meu canal é de culinária, de tech, meu canal é disso, daquilo.

Mas tudo que você pede para o usuário fazer, tudo que gera trabalho para o usuário, a chance do usuário não utilizar é muito grande. Pensa você, quantos softwares você aprendeu a usar no último ano? Deve ser pouquíssimos. Para você aprender a usar um software, para você aprender a fazer alguma coisa, você dedicar tempo, você tem que querer muito fazer essa coisa, essa coisa tem que ser muito relevante para você.

Então, eu quero diminuir o máximo possível de atrito entre o usuário e a pessoa que estiver trabalhando aqui na tecnologia. Então, eu quero que só a partir da URL ele já entenda sobre o que é o canal. E aí, eu já tenho o meu primeiro desafio de IA.

Eu vou utilizar a IA nessa parte aqui, onde entra IA, para descobrir o nicho do canal. A partir do nicho do canal, eu tenho que, para gerar essa outra parte aqui, essa parte dos vídeos relevantes, aqui eu consigo usar a API do YouTube, certo? E eu vou alimentar a API do YouTube com os nichos que eu achar aqui. Isso aqui é fácil.

Isso aqui é programático, essa parte que está circulada em preto aqui é tudo programático. A partir do que eu peço para o YouTube, ele consegue essas informações para mim, like, view e público, uma parte vírgula. Eu consigo, aliás, está impreciso isso aqui, eu acabei de reparar.

Eu consigo esta parte aqui, certo? Esta outra parte aqui, os resumos. Isso aqui eu preciso de uma outra IA. Um LLM para ir resumindo.

Provavelmente vai botar um chat GPT, vai botar alguma outra coisa para ele resumir essa parte do vídeo. Então, eu já tenho duas ferramentas de IA para trabalhar. Uma ferramenta aqui para definir o nicho do canal, e outra ferramenta aqui para resumir isso.

E quando eu estava desenhando no papel, eu falei, cara, eu já tenho uma boa noção do que eu quero com o papel. Chegou a hora de deixar essa coisa um pouquinho mais profissional, quero entender melhor essa mecânica. Então, o segundo passo que eu dei, depois que eu rascunhei, depois que eu pensei onde a IA ia entrar, eu fui aqui para a parte de projetos, propriamente dita, onde a gente foi aqui para o Figma.

E você pode notar aqui que ele está praticamente a mesma coisa que eu queria no papel, só um pouquinho mais bonitinho, mas de maneira que eu quero aqui um checkbox, aqui um push text, aqui que eu botei o retorno, aqui os resumos. Basicamente, essa parte aqui é a mesma coisa que eu peguei lá no papel, só que agora é um pouco mais bonitinho. Designs, me desculpem, eu sei que isso aqui está super básico.

Talvez um dia eu peço para um amigo design profissional dar uma ajudada e deixar esse UX, UI um pouquinho melhor. Mas aí o que eu pensei, o que eu quero como resposta aqui? Eu quero que ele fale, o tema sugerido com base nesses vídeos aqui é Lame Chain. Então, aqui a gente entra mais um outro modelo de IA.

Perceba que eu já tenho três modelos de IA. Um na parte do textbox aqui, essa parte está destacada aqui. Onde eu vou digitar o canal, e ele, a partir do URL do canal, tem que entender qual é o nicho.

Com essa resposta, eu vou mandar para a API do YouTube para ele pegar os vídeos mais relevantes, que são esses três vídeos mais relevantes aqui. Eu quero que ele mostre os views, os likes e quando foi a data de publicação, que eu espero que seja nos últimos sete dias, se tiver plegado isso, 28, 365 dias. Ele vai me dar um resumo sobre o vídeo, resumo textual, então aqui tem o nosso segundo modelo de IA para fazer o resumo desses vídeos.

E a gente tem o terceiro, a aplicação de IA, que é baseada nesses vídeos aqui, que são relevantes para o seu canal. Eu sugiro que você faça um vídeo de tal tópico. E aqui está uma sugestão de roteiro.

Por exemplo, a introdução, 30 segundos, por exemplo, parados, hands-on e conclusão. Essa imagem que vocês estão vendo aqui é o que eu pretendo construir nos próximos vídeos. A gente vai ter um vídeo pensando em cada modelo, como a gente vai usar o Lame Chain, aí eu vou mostrar as decisões de técnicas, se eu vou usar só a Chain, se eu vou criar um agente para fazer isso.

Todas essas decisões, essas minúcias decisões de engenharia, eu quero mostrar aqui para vocês. Eu vou utilizar o front-end mais simples possível que eu achar. Não sei se eu vou usar React, eu acho que talvez seja desnecessário para isso aqui.

Talvez eu pegue um Streamlit ou um Fleet, alguma dessas bibliotecas gráficas de Python para deixar o negócio mais rápido. Não recomendo, porque vocês usam profissionalmente essas bibliotecas. As bibliotecas são boas para prototipar ou, quando muito, para você fazer alguma coisa que você vai usar em home, em casa, no seu time.

Elas têm zilhares de problemas de segurança, não são performáticas, mas elas, para quem entende Python, é a coisa mais rápida do mundo para você fazer um front-end simples. Talvez eu use Fleet ou Streamlit, não tenho certeza. Talvez, num cenário muito raro, eu vou para um React aí.

Não sei se, talvez, me deixem aqui nos comentários, me ajudem nessa parte, se vocês querem ver alguma coisa mais profissional em React, trazer outras tecnologias, porque eu sei que o foco do canal, vocês não me assistem pela programação, vocês me assistem pela IA. Se você quiser, de repente, ver um projeto profissional, do começo ao fim, comenta aqui embaixo, é legal ter React ou não. Não imparte tempo com React, foca na IA, foca nos mais simples para a gente ter mais conteúdo de IA.

Me ajuda nessa parte de construção de conteúdo aqui. Beleza, gente? E a parte de back, vou usar o Flash API, que é um dos backends mais, acredito, de tendência no mundo de Python. Você tem o Django, o Flash e o Flash API.

O Flash API, acredito que é o Xanamas Asynchronous, ele é realmente mais moderno, é um framework bem interessante, vou usar ele. Vou mostrar como integra o modelo de IA no Flash API, acho que essa parte é importante, mas não vou falar sobre backends, não vou falar como você cria um endpoint, não vou falar como você cria um servidor, como você faz o deploy desse track numa cloud, eu não vou focar nessa parte de desenvolvimento de novo. Se você tiver interesse, deixa nos comentários, talvez eu me aprofundo nisso.

Mas, a princípio, eu quero mostrar como se integra o modelo de IA com a Flash API, que eu acho que até aí esse meandro é o limite onde o profissional de IA, de Data Science, de Machine Learning vai. A gente está até entrando mais uma área de engenharia de Machine Learning até, eu diria. Mas acho que é interessante mostrar essa parte, porque essa parte é importante e pouca gente mostra como se integra isso no mundo real.

E aqui nessas integrações, por exemplo, a gente acabou de passar aqui, eu acredito que vão ter três grandes modelos de IA aqui nesse projeto, um para achar o nicho, outro para resumir os textos e um terceiro para me dar a minha resposta. Pretendo desenvolver mais ou menos cada um, um vídeo nas próximas semanas. E eu espero que se você acha relevante essa ideia, se você quer que esse projeto continue, deixe nos comentários, curte o vídeo, compartilhe.

Eu vou ter o maior prazer do mundo em conversar com vocês e ver sugestões. Vocês que assistem esse canal, vocês que fazem esse canal, vocês estavam me pedindo bastante prazer prático, foi daí que eu tirei essa ideia. Então, eu vou ter o maior prazer do mundo em ouvir as ideias de vocês, interajam.

Interação é boa para o canal. E nos vemos na próxima semana, onde a gente vai, de fato, botar a mão na massa, onde eu vou começar a fazer esse projeto hands-on. Até mais, pessoal.

**Transcrito por** [**TurboScribe.ai**](https://turboscribe.ai/pt/?ref=docx_export_upsell)**.** [**Atualize para Ilimitado**](https://turboscribe.ai/pt/subscribed?ref=docx_export_upsell) **para remover esta mensagem.**