

Então, galera, agora para fechar com chave de ouro as nossas ferramentas interessantes de inteligência artificial, nós vamos trazer algumas inteligências artificiais generativas relacionadas a voz e avatares virtuais. Então a gente vai começar falando do Eleven Labs, que é uma ferramenta bem interessante para a sintetização de voz, então para fazer aquele processo de text-to-speech, usando inteligência artificial, ele tem também funcionalidades de clonagem de voz bem interessantes e bem efetivas em termos de qualidade, a gente vai dar uma olhada aqui. Além dele, a gente vai também explorar um pouquinho do DID, para vocês entenderem um pouco de como ele funciona, ele se propõe também a fazer sintetização de voz, se propõe também a fazer a virtualização de avatares, assim como o Heijan, que é um concorrente direto do DID nessa parte de avatares virtuais, mas que tem uma qualidade significativamente melhor do que o DID. Então a gente vai explorar um pouco, mostrar um pouco das características, lembrando que todas elas são ferramentas pagas, que tem uma free tier, vamos dizer assim, então aqui no caso do Eleven Labs, a gente tem 10 mil caracteres por mês gratuitos, o que dá mais ou menos uns 5 minutos, então se a gente for colocar aqui, então por exemplo, 10 minutos, eu quero criar um conteúdo com 10 minutos, vai dar 10 mil caracteres, então até 10 minutos de conteúdo, não 5 como eu tinha dito, mas até 10 minutos de conteúdo, a gente conseguiria fazer de maneira gratuita por mês, então essa é a camada gratuita do Eleven Labs. Olhando para o DID, ele também tem a camada gratuita para a parte de criação dele aqui, são 5 minutos de vídeo, então a gente conseguiria 20 créditos por mês com esses 5 minutos de vídeo, e o Hey Jan também tem a precificação dele com um crédito gratuito de no máximo 1 minuto de duração, então cada um tem as suas possibilidades, lembrando que alguns deles não renovam mensalmente, então é algo que eles estão ajustando os modelos de negócio deles, mudam isso o tempo todo, então tanto o Eleven Labs, quanto o DID, quanto o Hey Jan, eles estão evoluindo muito significativamente em termos de funcionalidades que eles estão oferecendo, assim como todas as IAs generativas nesse momento que a gente está vivendo, então não é nada incomum haverem mudanças no modelo de negócios, no modelo de precificação, então é sempre legal ficar atento e entender um pouco de como cada um funciona, beleza? Então eu vou logar aqui no Eleven Labs, para a gente começar a brincar aqui um pouquinho, então o Eleven Labs tem uma função que é super interessante galera, que é o seguinte, aqui ele tem essa parte de Voice Labs, que é onde você consegue clonar a sua própria voz, então olha que interessante, aqui por exemplo, é o que eu usei para clonar a minha voz, então olha que bacana, aqui eu peguei alguns trechos de conteúdos que eu já havia criado aqui na DIO, que estavam com um áudio bem legal, e coloquei aqui por exemplo, vou dar um play em um deles para vocês verem aqui. Como que nós já estamos com o projeto devidamente importado e configurado aqui no IntelliJ, uma coisa que a gente pode fazer, que é uma boa prática de desenvolvimento, é versionar o nosso código fonte, e para isso a gente vai usar, então percebam que com uma quantidade de exemplos aqui, 14 exemplos de trechinhos de código, de código não né, trechos de fala que eu tenho aqui, a gente já conseguiu clonar a minha voz, então percebam que são trechos bem pequenininhos né, trechos de, no caso aqui de 5 minutos, então basicamente com um conjunto de trechos de áudio aqui, a gente já consegue fazer a sintetização de voz, então quando eu venho aqui em use, basicamente o que eu posso fazer aqui é fazer algumas configurações de toning de voz aqui, configurações para melhorar por exemplo a estabilidade da voz, eu quero que tenha uma similaridade maior com um pouco dos meus trejeitos, do meu sotaque, então enfim, tem como fazer uma série de ajustes aqui, e obviamente fazer aqui o uso da conversão de texto em fala galera, então aqui ó, por exemplo eu poderia colocar aqui ó, testando o Eleven Labs como ferramenta de conversão de texto em fala, ou seja, Speech to Text, então basicamente eu pude criar aqui um textinho e eu gostaria de gerar com a minha própria voz né, então como eu já clonei a minha voz, eu poderia vir aqui e dar um generate, testando o Eleven Labs como ferramenta de conversão de texto em fala, ou seja, Speech to Text, então olha que legal né galera, a gente já conseguiu fazer esse teste usando o Eleven Labs, então ele gerou com uma voz num timbre muito próximo da minha né, então isso é

bem legal, e aqui se você não gostou dessa versão, quiser gerar de novo, você pode gerar, e aí todo o histórico fica aqui ó, legal, então fica todo aqui o histórico, inclusive eles têm estruturas de projetos né, onde a gente consegue fazer a organização em projetos, tem funcionalidades novas de dublagem né, onde você consegue pegar um vídeo no YouTube, upar aqui, falar por exemplo que esse vídeo está em português, e que você quer traduzir para espanhol por exemplo, chinês, inglês, e ele faz a dublagem usando a sua própria voz né, já clonando a voz da pessoa que está no vídeo, então assim, algumas funcionalidades bem interessantes né, mas o que é mais interessante mesmo assim do Eleven Labs é essa qualidade em termos de clonagem de voz né, que é onde a gente consegue clonar várias vozes aqui, então dá para vocês verem aqui que tem uma galera da D.I.O. que já teve a voz clonada, e todas as vozes têm essa mesma característica né, tem pedacinhos de áudio para treinar aqui o modelo de inteligência artificial para poder sintetizar a voz, então além né, dessa parte de clonagem de voz aqui no próprio Eleven Labs existem vozes que são vozes profissionais, vamos dizer assim, gravadas em estúdio por pessoas que você pode utilizar também né, então aqui por exemplo a gente poderia usar essa voz do Valentino por exemplo, aqui ó, como ele... Então, uma voz bem marcante né, bem legal aqui, a voz aqui também tem uma outra voz feminina aqui que a gente pode ver como é que tá... Então bem legal né galera, assim, as vozes têm uma qualidade muito boa, então não necessariamente você precisa clonar a sua voz para poder usar o Eleven Labs, você pode vir aqui e usar uma dessas vozes profissionais né, olhar aqui qual tem a maior quantidade de usuários, qual tem a qualidade ou se adequa melhor a necessidade que você tem, então bem bacana né, ver aqui essas possibilidades e aqui ó, a gente vai inclusive já fazer o download desse exemplo aqui que a gente fez no Eleven Labs, desse teste para poder mostrar como o GID e o Radeon funcionam. Então vamos agora logar aqui no GID, eu já deixei uma conta zerada aqui né, então vou até fazer o login nela, então aqui galera, a criação de conta é basicamente usando aí a sua conta do Gmail, então é bem tranquilo, então aqui basicamente eu tenho meus 20 créditos gratuitos no GID, e aí no GID aqui ó, quando a gente quer criar um vídeo, eu já tinha até upado aqui uma imagem, mas eu vou até refazer o processo para vocês verem que é bem simples, então aqui basicamente ele tem uma série de avatares já pré-definidos vamos dizer assim né, inclusive o nosso amigo aqui ó, o Venuton do Dioverso tá aqui, é o avatar que a gente utilizou aí nesse curso em específico, então a qualidade desses avatares aqui é bem legal, então depois a gente vai fazer até um teste com o avatar da foto aqui né, que a gente pode criar um avatar a partir de uma foto, e esses avatares HQ aqui que são avatares de maior qualidade, então aqui ó, eu vou adicionar uma foto minha, vou pegar exatamente aquela mesma foto e vou adicionar aqui, e aí o que a gente pode fazer? A gente pode tanto descrever um textinho aqui né, e escolher a língua né, falar assim putz, é um texto em português por exemplo né, então eu posso vir aqui falar com o texto em português e posso escolher algumas vozes aqui ó, então tem a voz do Antônio, tem a voz do Valério, Daniel, vozes masculinas e femininas, então tem vozes bem legais aqui para a gente poder colocar, e aí basicamente seria pegar esse texto, vamos dizer assim né ó, e daria para a gente até ouvir aqui ó, vamos ver como é que tá. Testando o Eleven Labs como ferramenta de conversão de texto. Então aqui ele limitou a 10 palavras né, mas deu para ver que a voz está bem legal né, e a gente poderia vir aqui também e mudar, por exemplo para o Júlio. Testando o Eleven Labs como ferramenta de conversão de texto. Da hora né galera, então poderia ser feito dessa forma né, usando o conceito de text to speech né, com as vozes que já existem aqui dentro, ou a gente poderia pegar aquele áudio que a gente gerou no próprio Eleven Labs e colocar aqui né, então eu posso vir aqui ó, e testando o Eleven Labs como ferramenta de conversão de texto em fala, ou seja, speech to text. Então aqui a gente pegou basicamente o áudio que foi gerado no Eleven Labs, que é esse áudiozinho aqui, o pô, e com isso a gente já consegue gerar um vídeo com um avatar baseado em imagem, vocês vão ver que a qualidade final não é lá aquelas coisas, mas vocês vão ver que já é um aspecto interessante para a geração de avatares virtuais, então tem uma série de recomendações relacionadas a foto aqui mesmo né, do avatar, então o

ideal é que a foto não tenha sombra, que você fique com a boca fechada na hora da foto para não ter uma expressão facial que distorça muito o seu rosto, tem aspectos de luz né, também, testando o Eleven Labs, então ó, vamos ver aqui ó, aqui ele botou a marca d'água do DID né ó, porque eu estou usando a versão gratuita, então ele colocou a marca d'água em todos os lugares, mas vamos ver como é que ficou a qualidade final do meu avatarzinho ali falando né ó, testando o Eleven Labs como ferramenta de conversão de texto em fala, ou seja, speech to text. Então dá para ver que às vezes quando a cabeça se movimenta, ele deforma um pouco, então não fica um negócio tão, tão, tão interessante né, tão natural, mas como ele está convertendo uma imagem em um vídeo né, no final das contas, é algo até que se você colocar condições boas na foto, uma qualidade legal, tirar sombra, trabalhar na iluminação, muito provavelmente você vai ter um resultado melhor do que esse aqui que eu peguei de uma foto aleatória minha aqui né, então vamos criar um segundo vídeo agora usando o nosso conhecido aqui, o nosso amigo já, que a gente está cocriando conteúdos nesse curso aqui ó, então eu vou colocar a minha voz nele ó, aquele mesmo áudio, e vou pedir para gerar, aqui eu poderia definir um background verde por exemplo, para simular um chroma key, para a gente poder trabalhar melhor na edição, e poderia gerar esse vídeo né, então ele vai consumir mais um créditozinho meu aqui, então percebam que dá para brincar bastante com esses 20 créditos que o GID disponibiliza para nós né, e aqui vocês vão ver que a qualidade final desses avatares de alta qualidade do GID já é algo bem melhor. Então vejam que as mãos já se movimentam né, o movimento da boca já é mais natural, da cabeça, então assim, a qualidade é bem superior. Então esse é basicamente o que o Eleven Labs né, oferece de, o Eleven Labs, não desculpem, o GID oferece como possibilidades né, então além disso que eu já apresentei, ainda tem essa parte de gerar um apresentador usando IA né, então você poderia vir aqui e colocar um prompt né, falando sobre algo que vocês gostariam de ter, então a gente poderia gerar um avatar aqui usando AI, então a gente poderia colocar aqui ó, a portrait of, aqui alguns exemplos ó, então uma imagem de, e aí a gente poderia colocar né, de, of a, a gente poderia, a programmer, in anime style, vamos ver aqui se ele entende né, se a gente colocar aqui que é uma foto de um animador no estilo de anime, então vamos ver o que ele gera para a gente no final das contas né, então essa daqui é uma outra possibilidade que é, tenta trazer aspectos de geração de imagem né, junto com uma questão de avatares virtuais, então olha que legal galera, então aqui ele já trouxe algumas opções né ó, vou pedir para ele adicionar essa na minha galeria, e eu vou tentar usar ela, então vou pegar o mesmo áudio aqui ó, e vamos gerar para ver como é que fica aqui né, usando um avatarzinho que nós mesmos geramos aqui, usando um prompt de geração de imagens aqui né, então percebam que essas ferramentas de inteligência artificial tentam integrar diversos conceitos né, por exemplo, essa galeria que a gente está vendo, ele integrando desde conceitos de prompting né, para a gente poder gerar um avatar, para gerar um avatar virtual, até conceitos de text to speech né, e também questões relacionadas à geração desses avatares virtuais, então essas ferramentas muitas vezes se propõe a fazer várias coisas, e vocês vão ver isso até de maneira mais evidente quando a gente chegar no Heijin, que ele já se propõe a fazer muito mais coisas do que todas as ferramentas que a gente buscou aqui, obviamente isso tem um custo né, vocês vão ver que o custo do Heijin é mais elevado, a gente até viu ali na página de pricing, e é tudo uma questão de análise para cada contexto em que vocês estão querendo explorar a inteligência artificial, vamos ver como é que ficou, testando o Eleven Labs como ferramenta de conversão de texto em fala, ou seja, speech to text, então percebam que ficou bem estranho né, a gente teria que colocar no prompt para a imagem ser gerada de frente né, para ele não ficar de lado assim, mas deu para entender a ideia né galera, e as possibilidades aqui que existem no DID, e aí indo para a última aqui né, e mais não menos importante é o Heijin, a gente pode vir e logar nele também, eu tenho uma conta também limpa aqui ó, que eu deixei prontinha, e aqui o Heijin é um pouco mais complicado em termos de créditos grátis né, porque a gente conseguiria fazer dois videozinhos só né, então enquanto aqui no DID a gente tem 20 créditos, o que daria para a gente criar até 20 vídeos bem curtinhos, aqui a gente conseguiria gerar só dois

né, porque ele sempre vai arredondar os videos para 30 segundos né, e a gente conseguiria gerar dois videos de 30 segundos cada, totalizando esse um minuto de crédito que a gente tem aqui, mas o que é interessante do Heijin, é que ele tem tanto a parte de avatares, onde você pode criar o seu, e essa criação de avatares, ela é um pouco diferente, então vocês vão ver que ela, vou até pegar aqui o Text Instructions, mas basicamente o que a gente tem que fazer, gravar um video de dois minutos em alta resolução, seguindo todas essas recomendações né, e fazendo gestos mais genéricos, vamos dizer assim, e com isso o modelo de inteligência artificial consegue entender e fazer algo muito mais natural do que aquilo que a gente viu no DID por exemplo, então aqui a forma de treinar o modelo de inteligência artificial é diferente né, é baseado em um video de dois minutos, e aí com esse video de dois minutos ele já consegue entender suas expressões faciais, clonar sua voz, entender a cadência de gestos que você está fazendo de maneira genérica com as mãos, e conseguir colocar aquilo de uma maneira adequada dentro dos videos né, então isso tudo faz parte de como o Heijin cria as experiências usando os avatares aqui, mas além disso ele se propõe também a ser um estúdio de edição, então aqui ele tem questões relacionadas a templates, para edição de video, além dos avatares que a gente já viu, tem a parte de clonagem de voz também, que você pode clonar a voz ou usar vozes profissionais aqui dentro, percebam que ele se propõe a ser o Eleven Labs e também o DID e mais uma série de outras coisas né, e aqui na parte de videos você também vai ter todos os videos que você criou, além de ter uma área aqui relacionada a colocar os assets da sua empresa para poder trabalhar na edição aqui de uma maneira mais tranquila, buscando esses assets com o branding kit da sua empresa, então percebam que aqui o Heijin ele tenta meio que abraçar o mundo vamos dizer assim, então tem até uma opção em beta aqui ó, que ele traduz um vídeo do YouTube e faz o lip sync né, a sincronia labial desse vídeo né, então vamos ver aqui essa demo ó. Então percebam que o Elon Musk obviamente não tá falando francês aqui, ele tá falando inglês, mas aqui foi feita a sincronia labial né, para ele poder falar em francês né, então percebam as possibilidades que existem aqui né galera, então esse vídeo translate é bem interessante, tá em beta ainda, mas abre uma série de possibilidades bem legais para a gente poder trabalhar nessa questão de acessibilidade de conteúdos né, então um conteúdo que tá em espanhol, em inglês, em russo, em chinês, em árabe, nós poderíamos trazer para o português de uma maneira muito mais simples né, e realmente democratizar o acesso a esse conteúdo. E olha que legal ó, a voz é provida pelo próprio Eleven Labs, então o Eleven Labs é uma referência nessa parte de sintetização de voz, tanto que o próprio Heijin possibilita a integração e utiliza esse serviço em alguns aspectos dentro da própria plataforma deles. Então aqui para a gente testar, eu não vou gerar um avatar aqui meu nem nada, mas eu vou usar um aqui para vocês verem a qualidade do Heijin, então eu vou gerar aqui em landscape, esse daqui ó, e aqui da mesma forma galera, eu posso usar o text to speech aqui né, usando uma voz por exemplo, dentro da galeria de vozes aqui né, então aqui se eu der o play ele vai falar um pouco. Bem-vindo à nova era da criação de vídeos com o Heijin, simplesmente escreva o seu script para começar. Então a gente vai usar meio crédito do um crédito que a gente tem, e percebam que o Heijin, ele é um pouquinho mais demorado para processar o vídeo né, para gerar o vídeo final, mas como é um vídeo curtinho de 6 segundos, vocês vão ver que ele não vai demorar tanto assim, mas para um vídeo maior, vocês veriam que esse tempo demoraria, aumentaria né, exponencialmente. Então, uma vez que você pediu para gerar aqui, você pode continuar navegando né, e ele vai continuar nessa parte de vídeo aqui, disponível até que ele esteja pronto, aí você vai receber uma notificaçõzinha aqui, mas basicamente era esse o contexto que eu tinha para passar para vocês relacionados a voz e avatares virtuais, enquanto a gente espera aqui, queria já aproveitar para fazer alguns disclaimers interessantes né, sobre as IAs generativas no nosso dia a dia, um exemplo prático galera, para quem usa o Windows, o Copilot, ele está disponível aqui ó, numa versão prévia, então você só clica aqui por exemplo, ele já aparece e você já consegue trocar uma ideia, usando o linguagem natural aqui por exemplo né, então, sei lá, resumir uma página no Edge né, que ele mostra



várias possibilidades ó, de prompt né, então sei lá, código em Python de Bubble Sort, sei lá, então você pode já trocar uma ideia aqui né, usando linguagem natural, que esse Copilot basicamente vai já trazer uma integração mais fluida com o sistema operacional aqui né, então olha que bacana, ele inclusive gerou um código aqui, mostrando esse algoritmo de ordenação que a gente pediu para ele gerar né, o Bubble Sort, com a opção de copiar e tal, então percebam né, as possibilidades e o quanto as IAs tendem a estar mais integradas no nosso dia a dia né, então o próprio sistema operacional hoje já vai trazer algumas possibilidades, como é o caso do Copilot da Microsoft, essa mesma funcionalidade de Copilot né, que a gente costuma dizer, está presente em outras soluções da Microsoft, como o próprio Visual Studio Code, onde você consegue usar o Copilot como um copiloto de programação mesmo né, então ele traz ali a inteligência artificial para dentro da parte de codificação no seu dia a dia, assim como outras IDAs também né, o IntelliJ tem extensões para isso, o Android Studio também tem, o Xcode, enfim, uma infinidade de outras ferramentas estão trazendo as inteligências artificiais para dentro do contexto de uso das ferramentas, para realmente facilitar, acelerar e dar produtividade para nós aí no dia a dia, legal? Então vamos ver aí como é que está a geração do nosso vídeo, então ele está em 70% já, vamos ver aqui, já já ele deve terminar, geralmente nesse finalzinho ele costuma já ir um pouco mais rápido, mas na essência é isso galera, cada vez mais as inteligências artificiais vão estar diante de nós no nosso dia a dia, e a intenção é que elas sejam facilitadores e aceleradores da nossa produtividade, e isso que é o mais interessante né, de usar esse tipo de ferramenta, que ela seja realmente algo que nos ajude a nos tornar profissionais melhores, então é nessa perspectiva que a gente trouxe esse conteúdo, desse curso em específico, para realmente mostrar para vocês essas possibilidades né, então aqui o Radiant está demorando bastante né, para finalizar, vamos ver, já já ele deve estar fechando, e a gente já volta para mostrar o resultado final. 2.000 anos depois. Bom galera, enfim, depois de quase 15 minutos aí, o vídeo gerou aqui no Radiant, acredito que tenha sido por conta de alguma demanda um pouco maior, e como a gente está na conta gratuita né, muito provavelmente eles não priorizam o processamento dos nossos vídeos obviamente, então vamos ver como é que ficou a qualidade final, só para vocês entenderem ó. 3. Testando o Eleven Labs como ferramenta de conversão de texto em fala, ou seja, speech to text. Então, percebam que o avatar é bem mais natural né galera, olhem os detalhes assim, da sincronia labial, movimentação de cabeça, respiração, movimentação das mãos. 4. Testando o Eleven Labs como ferramenta de conversão de texto em fala, ou seja, speech to text. Legal né, então basicamente para fechar aí né, esse conteúdo, era isso galera que eu tinha para passar para vocês, o nosso objetivo com esse conteúdo e as aplicações práticas dessas entrevistas artificiais é trazer realmente um conjunto de ferramentas incríveis para vocês poderem criar coisas cada vez mais interessantes e disruptivas no dia a dia de vocês. Beleza, então espero que vocês tenham curtido e a gente continua aí com o conteúdo do nosso curso para que vocês possam explorar e entender a aplicação das entrevistas artificiais em múltiplos contextos. Valeu demais, estamos juntos e até a próxima.