**Agentes de IA com Python: As 3 ferramentas indispensáveis**

**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.**

Essas são as três principais bibliotecas para a criação de agentes de IA com Python. Vamos, dentro desse vídeo, entender como funciona a criação de um agente de IA, quais são as três principais bibliotecas, que você já teve um spoiler aqui, para que cada uma delas serve e algumas bibliotecas extras que você pode usar na hora de criar os seus próprios agentes de IA e por que você deveria estar prestando atenção nisso, afinal de contas, isso aqui está mudando a forma como as empresas usam e criam soluções para elas e para os seus clientes. Vamos lá.

Bom, galera, vídeo bem direto ao ponto para trazer para vocês algo que você deveria se preocupar em aprender agora. Só para trazer primeiro um exemplo para vocês. Como que funciona um agente de IA? Se você entrar, por exemplo, no primeiro link aqui da descrição, vai ser o link onde você vai ver uma página de uma lista de espera para o nosso curso completo de Python.

Entrando nessa lista de espera, você vai cair nessa página aqui para você poder se inscrever, ver os detalhes e tudo mais. Se você reparar, tem um chatzinho aqui na ponta. E esse chatzinho aqui na ponta, ele pode ser respondido por pessoas, por seres humanos, mas ele também pode ser respondido por uma inteligência artificial.

E como que isso daqui acontece? Você chega aqui e envia uma mensagem. Por exemplo, quero a ementa de Python. E o que acontece? Quando você envia essa mensagem, essa mensagem vai para o nosso sistema.

O nosso sistema vai analisar essa mensagem e vai escolher se vai ser uma inteligência artificial que vai te responder ou se vai ser um ser humano, dependendo do momento e tudo mais. A gente criou essa lógica. E aí, digamos que seja uma inteligência artificial que vai te responder no momento que você mandou essa mensagem daqui.

O que vai acontecer? Ele vai esperar alguns segundos para ver se você vai ter alguma outra mensagem, alguma outra coisa para te enviar, e ele vai pegar essa sua mensagem e vai enviar para o nosso agente de inteligência artificial. Esse agente está recebendo essa sua mensagem e instruções padrões que a gente passou para ele. Com base nisso, ele vai usar algum sistema de LLM, ou seja, algum sistema de inteligência artificial, seja o da OpenAI, seja o do Cloud, seja o do Gemini, seja o do Google, seja o da Meta, enfim, vai usar algum deles para escolher como ele vai responder, para entender o que você mandou e para escolher a resposta que ele vai te mandar.

Como que ele vai escolher a resposta que ele vai te mandar? Com essa inteligência e com base nas informações que a gente passou para ele. Porque, repara, nesse nosso caso aqui, olha lá, ele me respondeu, foi uma inteligência artificial que me respondeu aqui. Olha lá o que ele vai falar para mim.

Ele vai me dar aqui uma ementa completa com base nas informações que eu passei para ele. Como que ele sabe que eu mandei e quero a ementa completa? Como que ele sabe a informação da ementa? Ele não está pesquisando essa informação na internet. Ele está pegando isso daqui dentro de memória e dentro de tools que eu passei para ele, uma base de conhecimento que eu passei para ele.

Como que você constrói isso tudo daqui? Para você construir isso tudo, você vai precisar conectar algumas soluções. Então, são três bibliotecas as principais que você precisa dominar. A primeira delas é a biblioteca da OpenAI.

Por quê? Porque hoje em dia, quando a gente cria as nossas soluções usando código, você pode ver que a OpenAI já vem com códigos em Python, já previamente aqui na documentação, dentro dela. E repara, não é OpenAI ou ChatGPT você usar ele por lá. Não, é você entrar na OpenAI Platform, API Platform.

Ou seja, é a parte de API da OpenAI. Você pode consumir ela por meio de requisições API ou por meio de um pacote mesmo, uma biblioteca Python chamada OpenAI, que você vai ter que instalar e usar. Dentro dela, você vai conseguir usar as soluções da OpenAI para a criação dos seus agentes de IA.

Então, estou falando de soluções de texto. Você vai poder mandar um texto e ele vai responder aquele texto para você, que foi o que vocês viram algo parecido aqui. Soluções de imagem, soluções de áudio.

Então, para gerar legenda, para gerar transcrição de vídeo, transcrição de áudio, gerar um áudio a partir de um texto, tudo isso você vai ter acesso por meio das soluções aqui que você tem da OpenAI. E a OpenAI, a gente acaba sendo uma das principais usos, não só porque ela é mais famosa, mas porque ela é muito boa, realmente. O resultado final costuma ser muito bom.

E porque dentro da API da OpenAI, você tem diferentes soluções. Você não tem só a solução de mandar um texto e ter um texto de resposta. Você tem soluções, por exemplo, para o processo de Embedding, que é o processo que a gente chama de pegar a sua base de conhecimento e extrair, transformar essa base de conhecimento em um formato de informações que as inteligências artificiais conseguem trabalhar de forma eficiente.

A OpenAI tem todos os tipos de soluções que você precisa para criar seus agentes de IA que você tem dentro da OpenAI. Então, acaba sendo uma das principais que a gente usa quando a gente vai criar os nossos agentes de IA. Então, a primeira delas é a OpenAI.

Beleza? Então, você vai usar a biblioteca da OpenAI e dentro dela você vai ter uma série e uma série de soluções. A segunda delas é a LangChain. Para que serve o LangChain? O LangChain é o cara que vai conectar todas essas pontas dos seus agentes aqui.

Você vai usar a OpenAI para conseguir puxar sua LLM. Você vai usar a OpenAI para conseguir fazer o Embedding da sua base de dados. Você vai usar a OpenAI para algumas coisas.

Mas o que conecta todas essas pontas aqui? Conecta essa mensagem com um formato específico de mensagem que vai enviar essa informação de uma maneira padronizada para a inteligência. Essa inteligência vai enviar uma resposta de uma maneira padronizada, vai criar uma memória para sua inteligência artificial para ela saber o que você falou antes e vai te responder depois. Quem que gerencia essa arquitetura toda, ou seja, a conexão de todas essas pontas? Esse cara é o LangChain.

Se você entrar aqui no LangChain, o LangChain hoje é a maior solução para criar esse tipo de estrutura, para conectar todas as pontas de um agente. E se você descer aqui na página deles, eles vão explicar de uma forma resumida. O LangChain serve para você orquestrar, integrar, avaliar e observar e fazer o deploy das suas soluções.

Ou seja, você vai pensar a sua solução toda, construir a lógica dela toda e gerenciar os seus diferentes agentes. Você vai conseguir integrar todas as pontas para garantir que o seu agente vai funcionar de uma forma lisa para o seu usuário final. Igual vocês viram aqui, eu mandei uma mensagem, chega uma resposta aqui, mas que por trás dele está acontecendo uma série de outras coisas.

Você pode avaliar e observar. Por exemplo, nesse prompt que eu mandei aqui, ele poderia ter dado algum erro. Por que ele deu esse erro? Eu vou poder chegar aqui e vou avaliar.

Por que ele deu esse erro? Por que ele demorou tão tempo para responder? Vocês viram que ele não respondeu tão instantâneo? Foi proposital isso ou foi alguma falha no sistema que ele deveria estar ajustando? Então, avaliar e observar e deploy, que é simplesmente colocar isso aqui no ar. A partir do deploy, e cada uma dessas partes aqui, existem várias soluções que permitem você fazer isso, mas o linkchain é a ferramenta mais utilizada, pelo menos, para usar isso tudo e construir isso tudo com uma única ferramenta. Essa é a segunda ferramenta que vocês vão ter que dominar se quiserem aprender a criar agentes de AI.

E a terceira é o Hugging Face. Por que o Hugging Face? Porque muitas das vezes a solução que você está criando, você não vai querer consumir uma API de uma inteligência artificial, como por exemplo a API do OpenAI. A API do OpenAI é paga.

Para cada requisição, você tem que pagar uma solução para ela e tudo mais. Às vezes, você vai querer hospedar a sua própria solução, ou não vai querer usar a solução do OpenAI, vai querer usar a solução de um outro modelo, de um outro sistema, ou você vai querer usar bancos de dados para poder testar e treinar os seus próprios modelos. O Hugging Face existe para isso.

Então, o Hugging Face, ele essencialmente tem duas grandes soluções aqui. Tem uma terceira, que é o Spaces, mas duas grandes soluções, pelo menos, para o objetivo desse vídeo, que é você vai ter aqui diferentes modelos de inteligência artificial para cada tipo de resposta que você quer dar. Então, eu quero fazer uma tarefa de gerar texto, eu quero fazer uma tarefa de gerar imagem, eu quero fazer uma tarefa, por exemplo, de visão computacional, eu quero fazer uma tarefa de processamento de linguagem natural, uma tarefa de áudio, uma tarefa de vídeo.

Você tem uma série de diferentes modelos com diferentes soluções que são open source daqui. Então, você vai encontrar modelos do Google, você vai encontrar modelos Mistral, você vai encontrar modelos do Entropiq, você vai encontrar vários modelos diferentes aqui dentro. E, além disso, ele vai ter também bancos de dados, bases de dados, datasets, para você poder treinar os seus modelos ou testar as suas soluções que você está criando.

Então, o Hugging Face acaba sendo uma alternativa se você não quiser ir pela solução do OpenAI, mas ele permite também você pegar esses modelos e rodar eles localmente na sua máquina ou no servidor que você mesmo hospedou. Então, você não fica dependente de uma API externa, você hospeda isso tudo dentro do seu próprio sistema para você ter controle total, digamos assim, sobre tudo o que está acontecendo ali dentro. Você não precisa se preocupar de enviar essa informação para uma API externa nem nada do tipo.

Aí que entra o uso do Hugging Face. Então, esses três são as três principais ferramentas na hora que a gente tem para construir agentes de AI. São justamente as três ferramentas que lá no nosso curso de Python a gente ensina lá na sessão de construção de agentes de AI e a gente foca principalmente nessas três ferramentas.

Existem outras? Sim, existem e eu vou comentar de alguma delas para vocês agora. Uma delas é a Crue AI, que está crescendo bastante, onde basicamente é uma ferramenta com o objetivo de você integrar múltiplos agentes. Então, por exemplo, digamos que você trabalha em um portal de notícia e você quer criar um agente de AI para criar as notícias para você automaticamente.

Na verdade, para criar os posts das notícias especificamente. Nesse caso, você não vai ter necessariamente um único agente, porque normalmente quando você está criando agentes de AI, você tende a preferir criar agentes específicos. Quanto mais específico for a inteligência artificial, melhor tende a ser o resultado dela.

Então, na prática, para você automatizar um processo inteiro, às vezes você não tem um único agente de AI. Você tem vários agentes de AI conectados um depois do outro para você automatizar um processo inteiro. Como o exemplo que eu estava falando para vocês.

Digamos que eu precise conectar diferentes agentes para eu poder fazer um processo de criação de notícias, posts dentro de um blog. O que você vai precisar fazer? Por exemplo, você poderia ter três agentes de AI diferentes. Um agente que vai ser um agente pesquisador, que o objetivo dele vai ser pesquisar uma série de informações dentro da internet, uma série de lugares para verificar a veracidade daquilo, ele vai ser otimizado para conseguir fazer pesquisas, ele vai ser otimizado para conseguir verificar a veracidade das informações, ele vai ser otimizado para trazer essas informações e tudo mais.

Ele vai pegar isso tudo e vai passar isso para um agente que vai ser um analista. Esse agente analista, ele vai analisar todas aquelas informações, e a partir daí ele vai escolher quais vão ser as informações que vão ter que ter dentro do post final, do post de blog ou do post de notícia final. E você vai ter um agente que é um agente escritor, que ele é muito bom em conseguir escrever títulos eficientes, títulos chamativos, textos engajantes e tudo mais.

Ele vai pegar aquelas informações que o analista destilou para ele e vai conseguir trazer isso para um texto super didático, super chamativo, super fácil de ler e tudo mais. Então repara, são três objetivos completamente diferentes, mas para você automatizar um único processo. E o Crew AI serve exatamente para isso, como o próprio nome diz aqui dele, né? Crew AI, né? Então ele serve para você criar a sua própria crew, ou seja, a sua própria, o seu próprio grupo de pessoas de inteligência artificial para você conseguir automatizar um processo.

Então a sua própria equipe de IAs para elas conseguirem executar a tarefa que você quer que elas executem. Aí que entra o Crew AI. E existe uma outra solução também muito usada, inclusive tem vídeo aqui no canal sobre ela também, que é o Lama Index.

O Lama Index é como se fosse uma substituição para o LinkedIn. Então ele vai fazer esse processo todo de orquestrar, de conseguir ler arquivos em PDF, de conseguir tratar bases de dados em formato diferente, de conectar com inteligência artificial de uma maneira diferente para você poder criar as suas próprias soluções de inteligência artificial. De novo, o LinkedIn acaba sendo mais usado que o Lama Index, pelo menos no momento em que eu gravo esse vídeo, mas o Lama Index também é uma excelente solução.

E além dessas, tem as outras LLMs, porque repara que eu falei para vocês da OpenAI, mas você poderia conectar direto com a API do Gemini, você poderia conectar direto com a API do Cloud, da Ontropic, você poderia conectar direto com APIs de outras LLMs diferentes. Então sim, existem todas essas outras gamas de APIs, mas esses três que eu passei para vocês, OpenAI, LinkedIn e Hugging Face, são os mais usados na prática para a gente criar esse tipo de solução e o Crew AI e o Lama Index acabam sendo boas alternativas e os outros modelos de inteligência artificial são as outras ferramentas que você pode precisar aprender. O que é o core? O que é o little? Precisa aprender para criar agentes de IA.

Essas três. OpenAI, LinkedIn e Hugging Face. Aprendendo essas três, você vai conseguir implementar e criar suas próprias soluções, seus próprios agentes de IA dentro da sua empresa e eu espero que vocês consigam.

Se dediquem, aprendam sobre isso, porque isso está realmente mudando muito o mercado de trabalho. A gente aqui na hashtag já está usando muito internamente e ensinando para os nossos alunos. Então, de novo, se você quiser aprender sobre isso, dá uma olhada no primeiro link da descrição, que o nosso curso vai estar lá.

O mais importante é vocês saírem daqui sabendo o que vocês precisam aprender e sabendo que vocês precisam aprender esse conteúdo para se destacar no mercado de trabalho. Beleza? Espero que vocês tenham gostado desse vídeo. Um grande abraço e até a próxima aula.

**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.**