

【今さら聞けないRAG入門】現場で役立つRAGの基礎と設計 ～社内情報資産をビジネス価値に変えるAI活用の考え方～

Q

ベクトルストアの構築後に、一部ファイルの情報が誤っていることが判明した場合にベクトルストアから該当ファイルの情報だけを削除や更新したりすることは可能なのでしょうか。それともベクトルストア自体の再計算が必要でしょうか。

A

削除や更新はできるが、基本的にはチャンクの再計算が必要になる。利用しているベクトルストアやフレームワークやライブラリにもよってくるが、ベクトルストアはベクトルの値だけではなくメタ情報もチャンクとして保存している。チャンキングで分割した際にそれぞれのチャンクに、分割する前のファイル名やファイルIDがメタ情報として入っている。例えば、ファイル名やファイルIDの情報を削除したい場合はその削除したいファイル名やファイルIDを渡すと、そこから分割されたチャンクが削除可能です。更新したいファイルを読み込んで一連の流れを通ってもう一度再登録する必要がある。

Q

RAGで登録したデータの容量は、コンテキストウィンドウのサイズ利用にも影響するのでしょうか？

A

RAGを使うことによって、関連データだけをAIに渡すので、全ての情報を渡すわけではないので、その分のコンテキストウィンドウの有効活用ができる。