ZTM RAG: OpenAI API

segunda-feira, 7 de julho de 2025 18:25

Messages Roles

- system: Comportamento do modelo.
- user: Prompt para o modelo.
- assistant: Resposta anterior do modelo.
- tool: usada em chamadas com ferramentas, functions (essa informação é do chatgpt, pois o curso não fala sobre isso)

No chatgpt, o system prompt não pode ser alterado mas na api é possível.

Prompt Injection (IMPORTANTE)

Se o system role não tiver bem protegido, pode haver prompt injection e vazar informações sigilosas colocadas no rag. É importante proteger o prompt injection. Ver aula 51.

OpenAI Playground

É uma ferramenta para testar parâmetros, modelos e funcionalidades da api da OpenAI. É possível salvar presets e também puxar o código correspondente para colocar no rag. Muito bom. Link: https://platform.openai.com/playground

Documentação da Api da OpenAi (BASTANTE COISA BOA)
https://platform.openai.com/docs/api-reference/introduction /
https://platform.openai.com/docs/overview

Imagens pela API da OpenAI

A Api da OpenAI consegue analisar imagens.
Para analisar imagens com a api da OpenAi, ela precisa estar no formato Base64. Caso não esteja, ela então precisa ser passada por uma url externa.
Base64 fornece segurança e confiabilidade. Para melhores resultados, utilize imagens de alta resolução com baixa compressão. É ideal para lidar com dados sensíveis que não devem ser públicos.

O que é Base64?
Base64 converte dados binários, como imagens, para texto ASCII, tornando-os fáceis para enviar via protocolos baseados em texto. Quando uma imagem é transformada em Base64, seus dados binários viram uma long string de texto. É bastante usado para transmitir imagens em HTML, emails ou API requests pela simplicidade e compatibilidade.