

warning "Atualização de backend (outubro de 2026)" A inferência do GitHub Models foi descontinuada em 30 de julho de 2026. As referências abaixo são mantidas apenas como contexto histórico enquanto este lab é migrado. Use o Lab 013 ? Backends de LLM gratuitos para configurar o Foundry Local ou o Ollama antes de continuar.

Lab 009: Início Rápido com Azure OpenAI Service

Nível: L100 Caminho: ? Microsoft Foundry Tempo: ~30 min ? Custo: Azure Free ? Conta gratuita do Azure, uso mínimo

O Que Você Vai Aprender

O que é o Azure OpenAI Service e como ele difere do OpenAI e do GitHub Models Como criar um recurso Azure OpenAI e implantar um modelo Como chamar o Azure OpenAI via SDK Python e API REST Autenticação por chave de API vs. Microsoft Entra ID (identidade gerenciada) Quando usar Azure OpenAI vs. GitHub Models vs. Ollama

Introdução

O Azure OpenAI Service fornece acesso aos modelos da OpenAI (GPT-4o, GPT-4o-mini, o1 e embeddings) por meio da infraestrutura Azure. Isso significa:

Conformidade empresarial: Os dados permanecem no seu tenant Azure, governados pelas suas

políticas Rede privada: Implante atrás de uma VNet, sem exposição à internet pública

Integrações Azure: Funciona nativamente com Azure AI Search, Azure AI Foundry, Key Vault e muito mais SLAs: Garantias de disponibilidade de nível empresarial

info "Azure OpenAI vs. GitHub Models vs. OpenAI" | | GitHub Models | OpenAI (direto) |

Azure OpenAI | |--|-----|-----|-----| | Custo | Nível

gratuito | Pago por token | Pago por token | | Autenticação | GitHub PAT | Chave de API |

Chave de API ou identidade gerenciada | | Residência de dados | GitHub/Azure | OpenAI EUA |

Sua região Azure | | Recursos empresariais | ? | Limitado | ? Completo | | Limites de taxa

| Baixo (apenas dev) | Padrão | Configurável | | Melhor para | Aprendizado, dev | Apps para

consumidor | Produção empresarial |

Para todos os labs L200+ neste hub, você pode usar GitHub Models (gratuito). Azure OpenAI é o caminho recomendado ao migrar para produção.

Configuração de Pré-requisitos

Criar uma Conta Azure

? Conta Gratuita do Azure ? crédito de \$200, 12 meses de serviços gratuitos

warning "Limitações do nível gratuito" O Azure OpenAI no nível gratuito tem cota muito limitada. Para estes exercícios, algumas chamadas de API são tudo o que você precisa.

Solicitar Acesso ao Azure OpenAI

O Azure OpenAI requer uma solicitação: 1. Acesse [aka.ms/aoi/access](https://aka.ms/aoai/access) 2. Preencha o formulário com seu caso de uso 3. Aguarde a aprovação (geralmente algumas horas a 1 dia)

tip "Enquanto aguarda a aprovação" Complete as seções conceituais deste lab e use GitHub Models para os exercícios práticos ? o código é idêntico.

Passo 1: Criar um Recurso Azure OpenAI

1. No Portal do Azure, pesquise por Azure OpenAI 2. Clique em + Criar 3. Preencha: -

Assinatura: sua assinatura - Grupo de recursos: criar novo ? rg-learning-hub - Região: East

US 2 (melhor disponibilidade de modelos) - Nome: aoai-learning-hub-[seunome] (deve ser

globalmente único) - Tipo de preço: Standard S0 4. Clique em Revisar + Criar ? Criar

Passo 2: Implantar um Modelo

1. Abra seu recurso Azure OpenAI 2. Clique em "Ir para Azure OpenAI Studio" ? oai.azure.com

3. Clique em Implantações ? + Criar nova implantação 4. Configure: - Modelo: gpt-4o-mini -

Nome da implantação: gpt-4o-mini (mesmo nome do modelo, mais fácil de lembrar) - Tokens por

minuto: 10K (suficiente para os labs) 5. Clique em Implantar

Implante também um modelo de embedding: Modelo: text-embedding-3-small Nome da implantação: text-embedding-3-small

Passo 3: Obter Seu Endpoint e Chaves

No seu recurso Azure OpenAI: 1. Clique em Chaves e Endpoint no menu à esquerda 2. Copie a CHAVE 1 e a URL do Endpoint

Armazene-os como variáveis de ambiente:

```
=== "Windows (PowerShell)" [diagram or code sample]
```

```
=== "macOS / Linux" [diagram or code sample]
```

danger "Nunca faça commit de chaves de API" Use variáveis de ambiente, Azure Key Vault ou arquivos .env (adicione .env ao .gitignore). Nunca codifique chaves de API diretamente no código-fonte.

Passo 4: Sua Primeira Conclusão de Chat

[diagram or code sample]

[diagram or code sample]

tip "Mesma API, cliente diferente" O código é idêntico ao OpenAI ? basta usar AzureOpenAI em vez de OpenAI, adicionar azure_endpoint e api_version. Só isso.

Passo 5: Embeddings com Azure OpenAI

[diagram or code sample]

Passo 6: Alternar Entre Backends com Uma Variável

Este padrão permite que você escreva o código uma vez e execute-o com GitHub Models (gratuito) em desenvolvimento e Azure OpenAI em produção:

[diagram or code sample]

Uso: [diagram or code sample]

Passo 7: Identidade Gerenciada (Melhor Prática para Produção)

Em produção, nunca use chaves de API. Use Identidade Gerenciada em vez disso:

[diagram or code sample]

[diagram or code sample]

Atribua ao seu app a função Cognitive Services OpenAI User no Azure RBAC.

Azure OpenAI no Azure AI Foundry

O Azure AI Foundry fornece um portal unificado para gerenciar o Azure OpenAI e outros serviços de IA:

1. Acesse ai.azure.com 2. Crie um Hub (uma vez por organização) e um Projeto (uma vez por aplicação) 3. Implante modelos por meio do Catálogo de Modelos 4. Use o Playground para testar interativamente 5. Acesse ferramentas de Avaliação e Monitoramento

? Veja o Lab 030 ? Foundry Agent Service + MCP para um passo a passo completo do Foundry.

? Verificação de Conhecimento

question "1. Qual é a principal diferença de código entre usar OpenAI diretamente e Azure OpenAI?" Use AzureOpenAI em vez de OpenAI, e forneça azure_endpoint e api_version. O parâmetro model recebe o nome da implantação, não o nome do modelo. Todo o resto (messages, temperature, etc.) é idêntico.

question "2. Por que você deve usar Identidade Gerenciada em vez de chaves de API em produção?" Chaves de API são segredos estáticos que podem ser vazados, roubados ou acidentalmente enviados ao controle de versão. A Identidade Gerenciada usa o sistema de identidade integrado do Azure ? nenhum segredo é armazenado em código ou arquivos de configuração. O acesso é controlado por meio do Azure RBAC, é rotacionado automaticamente e deixa trilhas de auditoria.

question "3. Quando você escolheria GitHub Models em vez de Azure OpenAI?" O GitHub Models é ideal para desenvolvimento, aprendizado e prototipagem ? é gratuito com apenas uma conta

GitHub, sem necessidade de assinatura Azure, sem implantação para configurar. Escolha Azure OpenAI quando precisar de: requisitos de residência de dados, rede privada, SLAs empresariais, integração com Azure AI Search ou Foundry, ou confiabilidade de nível de produção.

Resumo

Conceito | Principal conclusão | Cliente AzureOpenAI | Mesma API do OpenAI, apenas parâmetros de construtor diferentes | Nome da implantação | Você implanta um modelo com um nome ? use esse nome, não o nome do modelo | Versão da API | Parâmetro obrigatório ? verifique a mais recente em aka.ms/oai/docs | Identidade gerenciada | Use em produção ? sem segredos estáticos | Foundry | O portal que envolve o Azure OpenAI com avaliação, monitoramento e ferramentas de agente |

Próximos Passos

Construa um agente completo com Foundry: ? Lab 030 ? Foundry Agent Service + MCP Adicione RAG ao seu app Azure OpenAI: ? Lab 031 ? Busca Semântica com pgvector no Azure Alternativa gratuita para os labs: ? Lab 013 ? GitHub Models

Modernização da Fase 3

warning "Verified 2026-10-27" Use o nome Microsoft Foundry e o Agent Service GA baseado na Responses API. Avaliações estão GA; hosted agents continuam preview e devem ser identificados assim. Não crie dependência nova de Workflows do portal, descontinuados em 01/12/2026.

[diagram or code sample]

Foundry Agent Service GA announcement Foundry Agent Service Foundry evaluation Foundry Workflows retirement

Kit de apoio

O build do site gera downloads estáveis em PDF e ZIP. O ZIP contém o guia do lab e todos os arquivos de apoio declarados para este lab.

Baixar PDF gerado Baixar ZIP gerado

Supporting file: CHEATSHEET.md

Arquivos de apoio: Lab 009: Início Rápido com Azure OpenAI Service

Comece pelo caminho local ou gratuito documentado. Registre premissas e evidências de validação. Mantenha serviços hospedados opcionais, salvo quando o lab os exigir.

Supporting file: README.md

Lab 009: Início Rápido com Azure OpenAI Service

Use estes arquivos com o guia do lab e o kit gerado.