4·Linux Open Software Specialists™

CERTIFICADO

Certificate

Cleber Reis Do Bomfim

Concluiu seus estudos no curso **250 - LLMs Open Source: Do Zero ao Seu Primeiro Modelo Corporativo** ministrado pela empresa 4Linux e cumprindo a carga horária de 24 horas.

21 de agosto de 2026



RODOLFO GOBBI

DIRETOR GERAL

Para validar a autenticidade deste certificado acesse aia.4linux.com.br/admin/tool/certificate/ e digite o código: 7354717463CR

Ementa de Curso

Cleber Reis Do Bomfim
Concluiu seus estudos no curso **250 - LLMs Open Source: Do Zero ao Seu Primeiro Modelo Corporativo** ministrado pela empresa 4Linux e cumprindo a carga horária de 24 horas.
21 de agosto de 2026

A autenticidade deste documento pode ser verificado em aia.4linux.com.br/admin/tool/certificate/ digitando o código 7354717463CR

MÓDULO 1 — O Que É um LLM de Verdade (Sem Mistério)

Carga horária: 2h

Objetivo de aprendizagem: Explicar como um LLM funciona internamente usando linguagem visual e analogias, eliminando o misticismo que paralisa decisões técnicas.

Pré-requisito do módulo: Nenhum. Ponto de partida absoluto.

Tópicos teóricos:

- O que um LLM faz, de verdade: previsão de tokens, não "inteligência"
- Como texto vira número: a ideia de token e embedding explicada visualmente
- O que é contexto e por que ele tem limite (janela de contexto)
- Diferença entre inferência (usar o modelo) e treinamento (criar o modelo)
- Por que LLMs erram: alucinação, viés, data de corte
- A diferença entre um modelo open source e uma API proprietária — em uma frase

Laboratório prático obrigatório: Atividade: O aluno executa um prompt técnico idêntico em três interfaces diferentes — ChatGPT (API proprietária via browser), Hugging Face Inference API (modelo open source hospedado) e Ollama local (modelo rodando na própria máquina). Compara as respostas, a latência e o que cada abordagem revela sobre controle de dados. Ferramentas: Browser, Ollama (instalação guiada passo a passo), Hugging Face Inference API com token gratuito Entregável: Tabela comparativa preenchida pelo aluno com respostas, tempo de resposta e observações sobre privacidade de dados em cada cenário

Critério de avaliação: O aluno consegue explicar em uma frase por que a mesma pergunta pode gerar respostas diferentes nos três ambientes, e identifica qual dos três cenários seria inviável para uso com dados confidenciais da empresa.

MÓDULO 2 — Open Source vs. Proprietário: Uma Decisão de Negócio, Não de Preferência

Carga horária: 2h

Objetivo de aprendizagem: Avaliar cenários reais para decidir quando um LLM open source faz mais sentido do que uma API proprietária e/ou pública, usando critérios técnicos e econômicos objetivos.

Pré-requisito do módulo: Módulo 1 concluído.

Tópicos teóricos:

- O que "open source" realmente significa em LLMs: modelo, pesos e licença
- Visão geral das licenças mais comuns: LLaMA (Meta), Mistral, Qwen, Falcon — o que cada uma permite em uso comercial
- Lock-in de fornecedor: o risco que ninguém calcula até ser tarde
- Quando APIs proprietárias fazem sentido — e quando não fazem
- Os 5 critérios para decidir: privacidade dos dados, custo de escala, controle de versão, latência e compliance regulatório
- Mitos populares sobre LLMs open source (são piores? são mais baratos sempre? são mais difíceis?)

Laboratório prático obrigatório: Atividade: O aluno recebe 3 cenários corporativos fictícios (empresa de saúde, e-commerce de médio porte, escritório de advocacia) com restrições de dados, orçamento e regulatório descritas. Preenche uma matriz de decisão com os 5 critérios e justifica, por escrito, a recomendação para cada cenário. Ferramentas: Planilha de matriz de decisão fornecida, documentação resumida de licenças (fornecida pelo curso) Entregável: Matriz de decisão preenchida com justificativa para os 3 cenários

Critério de avaliação: O aluno justifica corretamente ao menos 2 dos 3 cenários com base nos critérios apresentados, sem contradições internas.

MÓDULO 3 — O Ecossistema Open Source: Quem São os Jogadores

Carga horária: 2h

Objetivo de aprendizagem: Navegar no ecossistema de LLMs open source e identificar modelos adequados a casos de uso corporativos com base em critérios técnicos objetivos.

Pré-requisito do módulo: Módulo 2 concluído.

Tópicos teóricos:

- O Hugging Face Hub: o que é, como funciona, o que tem lá
- Visão geral dos modelos mais usados em contexto corporativo: LLaMA 3, Mistral, Qwen, Phi,

Gemma
- Como ler um model card: parâmetros, benchmark, licença, requisitos de hardware
- Contrapartidas práticas: qualidade × custo × latência × hardware necessário
- Modelos quantizados: o que é GGUF e por que ele importa para quem não tem GPU dedicada
- Como a comunidade open source mantém e evolui modelos — e o que isso significa para estabilidade corporativa

Laboratório prático obrigatório: Atividade: O aluno acessa o Hugging Face Hub, navega até a aba de modelos de linguagem e seleciona dois modelos diferentes usando filtros de licença e tamanho. Lê o model card de cada um e preenche uma ficha comparativa com: tamanho em parâmetros, licença comercial (sim/não), hardware mínimo recomendado, caso de uso documentado e última atualização. Ferramentas: Hugging Face Hub (browser), ficha comparativa fornecida pelo curso Entregável: Ficha comparativa de 2 modelos com parecer do aluno sobre qual usaria para um assistente interno de RH, justificando com base nos critérios da ficha

Critério de avaliação: O aluno identifica corretamente a licença e os requisitos de hardware de cada modelo e justifica a escolha com ao menos dois critérios técnicos válidos.

MÓDULO 4 — Rodando um LLM: Da Teoria ao Terminal

Carga horária: 4h

Objetivo de aprendizagem: Executar localmente um LLM open source e interagir com ele via API, compreendendo o que acontece por baixo de cada requisição.

Pré-requisito do módulo: Módulos 1 a 3 concluídos. Ter terminal disponível (Linux, Mac ou WSL no Windows).

Tópicos teóricos:

- O que acontece quando você faz uma requisição a um LLM: fluxo completo de inferência
- Ferramentas de inferência local: Ollama, LM Studio e llama.cpp — quando usar cada uma
- O que é uma API de inferência e por que padronizar no formato OpenAI-compatible muda tudo
- Parâmetros de geração que todo profissional técnico precisa conhecer: temperature, top_p, max_tokens, stop sequences
- Hardware e realidade: o que roda em CPU, o que precisa de GPU, o que é inviável sem investimento

Laboratório prático obrigatório: Atividade em 3 etapas. Etapa 1: instalação guiada do Ollama e download do modelo Mistral 7B (ou Phi-3 Mini para máquinas com menos recursos). Etapa 2: interação via terminal com o modelo usando diferentes valores de temperature (0.1, 0.7, 1.2) para a mesma pergunta técnica — observando e documentando a variação nas respostas. Etapa 3: requisição ao modelo via curl para a API local compatível com OpenAI, visualizando o JSON de resposta e identificando os campos relevantes. Ferramentas: Ollama, terminal (bash/zsh), curl ou Postman/Insomnia Entregável: Print das 3 respostas com temperature diferentes + JSON da resposta via API, com comentário do aluno sobre o impacto do parâmetro temperature no resultado

Critério de avaliação: O aluno executa com sucesso as 3 etapas, demonstra via entregável que o modelo respondeu localmente sem conexão externa, e articula a diferença de comportamento entre temperature baixa e alta.

MÓDULO 5 — Arquitetura Corporativa: Como LLMs Entram em Produçã

Carga horária: 3h

Objetivo de aprendizagem: Desenhar a arquitetura básica de um LLM como serviço interno corporativo, identificando os componentes necessários e as decisões de design que impactam custo e segurança.

Pré-requisito do módulo: Módulo 4 concluído. Familiaridade básica com o conceito de container é suficiente.

Tópicos teóricos:

- LLM como componente de plataforma: a diferença entre um experimento e um serviço
- Componentes de uma arquitetura corporativa de LLM: modelo, servidor de inferência, gateway de API, autenticação, logging
- Inferência local vs. cloud vs. híbrida: quando cada modelo faz sentido
- Padrões arquiteturais: single-tenant vs. multi-tenant, sidecar de LLM, serviço compartilhado
- O papel do Docker: porque conteinerizar a inferência resolve problemas reais de portabilidade e versionamento
- Kubernetes: visão arquitetural — o que ele faz e por que aparece nas decisões de escala (sem código, só conceito)

Laboratório prático obrigatório: Atividade: O aluno recebe um diagrama de arquitetura incompleto de um serviço interno de LLM e deve completá-lo, posicionando corretamente: o container de inferência, a camada de autenticação, o endpoint de API interno e o componente de logging. Em seguida, identifica, no diagrama, qual componente seria responsável por impedir que dados de um departamento vazem para outro. Ferramentas: Diagrama editável (Excalidraw ou draw.io, fornecido pelo curso), gabarito comentado entregue após submissão Entregável: Diagrama completado com identificação e justificativa do componente de isolamento de dados

Critério de avaliação: O aluno posiciona corretamente ao menos 3 dos 4 componentes e identifica corretamente o ponto de isolamento de dados, com justificativa coerente.

MÓDULO 6 — Custos, Riscos e a Decisão de Avançar ou Não

Carga horária: 3h

Objetivo de aprendizagem: Estimar os componentes de custo de um projeto de LLM corporativo e aplicar critérios objetivos para decidir se um projeto deve avançar, pivotar ou ser abandonado.

Pré-requisito do módulo: Módulos 4 e 5 concluídos.

Tópicos teóricos:

- Componentes de custo em projetos de LLM: infraestrutura, operação, manutenção e custo oculto de engenharia
- TCO simplificado: como calcular o ponto de equilíbrio entre open source on-prem e API proprietária e/ou pública.

- Riscos operacionais: o que pode dar errado em produção e qual é o custo de cada falha
- Segurança e privacidade: os três vetores de risco em LLMs corporativos (dado de entrada, dado de saída, modelo em si)
- LGPD e LLMs: o que a lei exige e o que a maioria das empresas está ignorando
- Anti-padrões clássicos: os erros mais comuns que encarecem e atrapalham projetos de LLM
- Como defender (ou enterrar) tecnicamente um projeto de LLM perante gestores e diretoria

Laboratório prático obrigatório: Atividade: O aluno recebe o briefing de um projeto fictício de LLM corporativo — uma empresa de 500 funcionários que quer um assistente interno para consulta de documentos de RH. O aluno deve: (1) estimar o custo mensal de dois cenários (API OpenAI vs. Ollama local com servidor dedicado), usando uma planilha de TCO simplificada fornecida; (2) identificar 3 riscos de LGPD no projeto; (3) dar um parecer técnico de "avança / pivotar / não faz" com justificativa. Ferramentas: Planilha de TCO fornecida (Google Sheets ou LibreOffice Calc), briefing do projeto em PDF Entregável: Planilha preenchida + parecer técnico escrito (máximo 1 página)

Critério de avaliação: O aluno produz estimativas de custo com premissas explícitas (não precisa ser perfeito, precisa ser consistente), identifica ao menos 2 riscos de LGPD válidos e emite um parecer que não contradiz os dados da própria planilha.

MÓDULO 7 — Casos Reais e Projeto Integrador

Carga horária: 4h

Objetivo de aprendizagem: Integrar os conhecimentos do curso em uma análise completa de viabilidade técnica e econômica de um caso de uso corporativo real de LLM, defendendo a recomendação com argumentos técnicos estruturados.

Pré-requisito do módulo: Todos os módulos anteriores concluídos.

Tópicos teóricos:

- Casos de uso corporativos realistas para LLMs open source: assistente interno, sumarização de documentos, triagem de chamados, geração de rascunhos
- Casos onde LLMs são a solução errada — e o que usar no lugar
- Como apresentar uma decisão de arquitetura para stakeholders técnicos e não técnicos
- O que vem depois: MLOps, RAG governado e agentes de IA como evolução natural desta base (preview dos Cursos 2 e 3 da trilha)

Laboratório prático obrigatório — Projeto Integrador: Atividade: O aluno escolhe um dos 3 cenários corporativos disponibilizados (saúde, jurídico ou serviços financeiros) e produz um documento técnico de 2 a 3 páginas contendo: (1) modelo open source recomendado com justificativa de licença e hardware; (2) esboço da arquitetura de serviço (baseado no Módulo 5); (3) estimativa de TCO simplificada (baseada no Módulo 6); (4) principais riscos e como mitigá-los; (5) recomendação final fundamentada. Ferramentas: Todos os artefatos produzidos nos módulos anteriores, template de documento fornecido Entregável: Documento técnico de análise de viabilidade — este documento pode ser usado pelo aluno como portfólio profissional real

Critério de avaliação: O documento apresenta consistência interna entre recomendação de modelo, arquitetura, custo e riscos. A recomendação final é sustentada por pelo menos 3 critérios técnicos apresentados ao longo do curso.

MÓDULO 8 — Módulo Bônus: Seu Próximo Passo na Trilha

Carga horária: 2h (assíncrono, autoguiado)

Objetivo: Apresentar ao aluno o que os Cursos 2 e 3 da trilha oferecem, com demonstrações práticas dos laboratórios mais impactantes de cada um — funcionando como vitrine técnica da sequência.

Formato: Videoaulas demonstrativas + acesso de leitura às ementas dos cursos seguintes. Sem avaliação formal.