

Prova prática - Analista de IA Sênior (Atlantyx)

Kit de recrutamento (pronto para envio ao candidato).

Contexto

A Atlantyx vai implantar um assistente corporativo de IA para uma grande empresa. O objetivo é responder perguntas a partir de documentos internos com rastreabilidade (citações), controle de acesso por área e padrões de produção (segurança, observabilidade e custo).

Tempo sugerido

6 a 10 horas (take-home).

Materiais fornecidos

- Prova_Atlantyx_IA_Senior_10_Perguntas_Respostas.csv (10 perguntas + respostas esperadas + fontes)
- Doc1_Politica_IA_Grandes_Empresas_v1_2.docx
- Doc2_Playbook_Implantacao_IA_Enterprise_v0_9.docx
- PDF1_ArquiteturaReferencia_RAG_Enterprise.pdf
- PDF2_Matriz_Riscos_Controles_IA.pdf
- HTML1_FAQ_Glossario_IA_Grandes_Empresas.html
- HTML2_Caso_Uso_Roadmap_IA_Empresa_X.html

Observação: os documentos contêm pontos deliberadamente ambíguos e/ou conflitantes para testar governança, rastreabilidade e tomada de decisão do candidato.

Objetivo da prova

- Construir um MVP funcional de RAG (Retrieval-Augmented Generation).
- Demonstrar padrões de implantação enterprise (segurança, observabilidade, custo e escala).
- Entregar avaliação mínima com métricas e recomendações.

Requisitos obrigatórios

- Pipeline de ingestão e indexação (extração, chunking, embeddings, index).
- API com endpoint /ask que retorne resposta, citações (fonte + trecho), confiança (heurística) e métricas (tokens/latência).
- Filtro por permissão (ex.: user_role) no momento do retrieval.
- Estratégia mínima anti-alucinação (se não houver evidência, responder 'não encontrei nos documentos fornecidos').
- Avaliação automática usando o CSV (10 perguntas) com métricas e relatório.
- Logs estruturados (JSON) e correlação por conversation_id.
- Mínimo de segurança: mitigação de prompt injection e política de logs (evitar vazamento de conteúdo sensível).

Entregáveis

- **Código-fonte** (repositório): ingestão + API + avaliação (scripts/CLI).
- **README**: como executar localmente (Docker opcional), decisões técnicas, trade-offs e limitações.
- **Relatório** (Markdown ou PDF): métricas da avaliação, análise de erros e próximos passos.
- **Exemplo de execução**: saída real de 3 perguntas (com citações).

Critérios de avaliação (pontuação)

Dimensão	Peso	O que será observado
Qualidade do RAG e rastreabilidade	30	Citações consistentes, recuperação relevante, respostas 'grounded'
Arquitetura e padrões de produção	20	Separação de camadas, configuração, versionamento, escalabilidade
Avaliação e observabilidade	20	Métricas, logs, diagnose, reprodutibilidade dos testes.
Segurança e privacidade (LGPD)	15	RBAC/ABAC, anti-injection, política de logs, secrets.
Código, testes e documentação	15	Qualidade do código, testes, clareza do README, automação.

Entrevista técnica (ao vivo - 60 a 90 min)

Após a entrega, o candidato apresenta o design para escalar a solução (5.000 usuários/dia), incluindo estratégia de custo, governança, segurança, observabilidade e operação (LLMOps).

Regras

- Não usar dados reais de clientes. Trabalhar somente com o material fornecido.
- É permitido usar bibliotecas e serviços gerenciados, desde que as decisões estejam explicadas.
- A solução deve rodar localmente ou com instruções claras. Se usar cloud, fornecer um modo local equivalente (mock).
- A resposta deve ser em português e com postura corporativa.

Como enviar

Enviar link do repositório (ou zip) + relatório + instruções de execução. Opcional: vídeo curto (5-8 min) mostrando execução e resultados.

Atlantyx - Recrutamento (documento interno).