

## **Case Técnico – Arquiteto de IA**

### **Empresa fictícia: NexoBank**

#### **Contexto**

O NexoBank é uma fintech digital com foco em crédito personalizado e investimentos automatizados. A empresa já utiliza machine learning tradicional para recomendação de produtos, mas quer expandir o uso de IA generativa.

Stack atual:

- Cloud principal: AWS (EKS, Lambda, S3, SageMaker, Redshift, API Gateway)
- Cloud secundária (legado): Oracle Cloud (OCI), onde estão os bancos Oracle e documentos regulatórios (PDFs/XML)
- Orquestração: Apache Airflow
- CI/CD: GitHub Actions + Terraform
- Autenticação: Auth0 by Okta (OAuth2 + JWT)
- Aplicação mobile: React Native com backend Node.js

#### **Desafio**

Criar uma proposta arquitetural para um novo módulo de insights financeiros com IA generativa, que será exibido dentro do app mobile dos clientes Private e Select.

Esse módulo deve:

- Analisar os dados financeiros do cliente (transações, perfil, contratos)
- Cruzar com regras e documentos internos
- Gerar textos explicativos e personalizados (ex: “Você pode economizar com...”, “Sua previsão de gastos em setembro é...”)
- Exibir esses insights com segurança, apenas para usuários autenticados
- Operar com base em dados armazenados tanto na AWS quanto na OCI
- Ter reprodutibilidade garantida — ou seja, qualquer insight gerado deve poder ser reproduzido com os mesmos dados, modelo, prompt e contexto.

## **Entregável**

O candidato deverá produzir um documento de Desenho de Solução (Solution Design Document), contendo:

1. 1. Diagrama de Arquitetura

- Nível de Contexto (visão macro)
- Nível de Contêineres (principais serviços e suas interações)
- Padrão AWS + indicações de onde o legado (OCI) entra

2. 2. Estratégia de IA Generativa

- Abordagem sugerida (ex: RAG, fine-tune, prompt engineering)
- Onde o modelo roda e como os dados são preparados
- Como a solução garante governança e reprodutibilidade

3. 3. Fluxo de Segurança e Autenticação

- Visão de como a autenticação via Auth0 protege os dados
- Como garantir que cada insight é acessado apenas pelo cliente correto

4. 4. Exemplo prático simplificado de insight

- Texto gerado
- Dados utilizados como base
- Onde isso aparece no app

Importante: Não será necessário construir código, nem realizar POC. Avaliaremos apenas o raciocínio, estrutura da solução e clareza técnica.

## **Prazo**

Entrega em até 3 dias corridos.

## **Critérios de Avaliação**

- Clareza e estruturação da proposta
- Arquitetura escalável e realista, considerando uso multicloud (AWS + OCI)
- Governança e segurança de dados (com autenticação e rastreabilidade)
- Reprodutibilidade: capacidade de explicar como os insights podem ser reproduzidos
- Criatividade na aplicação da IA generativa ao contexto bancário