



DataXpertIA

Dados & Inteligência Artificial

APRESENTAÇÃO ESTRATÉGICA

DataXpertIA

Plataforma de Dados Enterprise

Sua plataforma de dados, sem dependência de fornecedor. Unificamos seus dados e aceleramos analytics e IA — com custo previsível e arquitetura 100% sob seu controle.

100% OPEN SOURCE

SEM VENDOR LOCK-IN

CUSTO PREVISÍVEL



O PROBLEMA

Dados Espalhados, Projetos Lentos, Custos Crescentes



A Dor

- **Dados espalhados** — cada área guarda seus números em um sistema diferente, e nenhum conversa com o outro
- **Projetos lentos** — iniciativas de analytics e IA levam meses porque a base de dados nunca está pronta
- **Custos crescentes** — licenças e infraestrutura sobem todo ano, sem retorno proporcional para o negócio



O Risco

- Decisões atrasadas, risco operacional e equipes sobrecarregadas em operação reativa, em vez de inovar
- LGPD: multas de até 2% do faturamento, limitadas a R\$ 50 milhões (Lei 13.709/2018, art. 52)
- 80% do dado corporativo não gera valor por falta de integração e governança (Forrester Research)

80%

dados ociosos (Forrester)

40%

crescimento anual de dados (IDC)

2%

do faturamento (LGPD)

R\$ 50M

limite de multa (Lei 13.709)



A SOLUÇÃO

DataXpertIA: A Plataforma de Dados sob o Seu Controle

01



100% OPEN SOURCE

Transparência e independência, sem licenças proprietárias pesadas. Código auditável, padrões abertos.

02



CUSTO PREVISÍVEL

Sem licenças por usuário ou volume: o custo cresce com a sua operação — não com o sucesso dela.

03



GOVERNANÇA POR DESIGN

Catálogo, lineage e políticas de acesso desde o primeiro dia, alinhados à LGPD.

04



OPERAÇÃO CONTÍNUA

Implantamos, operamos e evoluímos a plataforma junto com o seu time. Não é consultoria pontual.

05



PRONTIDÃO PARA IA

Base estruturada para ML e IA generativa — feature store, vetores e modelos open source.

CAPACIDADES

Uma Plataforma Completa, em Operação Contínua

CAPACIDADE	DESCRIÇÃO	DIFERENCIAL
Ingestão Unificada	Pipelines monitorados conectando todas as suas fontes	Batch e streaming
Armazenamento Aberto	Dados em formatos abertos, portáteis e sob seu controle	Zero lock-in
Processamento Escalável	Capacidade ajustada à demanda, sem desperdício	Escala horizontal
Governança Integrada	Inventário ativo, rastreabilidade (lineage) e controle de acesso nativos	LGPD by design
Analytics e BI	Dashboards e análises self-service para o negócio	Self-service
Prontidão para IA	Vetores, feature store e modelos prontos para ML e IA generativa	IA em produção



A ARQUITETURA

Arquitetura Lakehouse em Camadas Abertas

- **Fontes** — ERP, CRM, bases operacionais, APIs, eventos em streaming e sistemas legados (SAP, Salesforce, PostgreSQL, MongoDB)
- **Ingestão** — pipelines orquestrados em batch e streaming, com versionamento e testes (Apache Airflow, Spark, Kafka)
- **Armazenamento** — object storage S3 durável e econômico, com formatos abertos (MinIO, Apache Iceberg)
- **Lakehouse** — flexibilidade de data lake com governança de data warehouse (Apache Iceberg + Apache Polaris)
- **Processamento** — ETL, análises interativas e preparação de dados (Apache Spark, Spark Connect, Apache Kyuubi)
- **Metadados e Governança** — catálogo ativo, lineage, qualidade e políticas LGPD (DataHub, Schema Registry)
- **Analytics e IA** — BI self-service, notebooks SQL e base para IA generativa (DataXpertIA UI, PostgreSQL + pgvector)

SOBERANIA DIGITAL

Cada camada evolui de forma independente, sem prender seus dados. Código aberto, padrões abertos e controle total do stack.



A TECNOLOGIA

Stack Open Source de Classe Enterprise

- **Apache Iceberg** — tabelas abertas com ACID, evolução de schema e time travel; seus dados abertos e portáteis
- **Apache Spark + Apache Kyuubi** — processamento distribuído para ETL e análises interativas, batch e streaming
- **Apache Airflow** — orquestração de pipelines com agendamento, monitoramento e reprodutibilidade
- **MinIO** — object storage compatível com S3, em infraestrutura própria ou cloud
- **Apache Polaris** — catálogo REST do Iceberg, com registro central de tabelas e controle de acesso
- **Apache Kafka** — streaming de eventos em tempo real para pipelines de baixa latência
- **DataHub** — catálogo de dados e governança: inventário, lineage e metadados em um só lugar
- **PostgreSQL + pgvector** — metadados operacionais e os vetores que sustentam as aplicações de IA



COMPARATIVO

Plataforma, Não Projeto Nem Licença Proprietária



Modelo Proprietário

- Licenças por usuário ou volume que sobem todo ano, sem retorno proporcional
- Vendor lock-in: dados e metadados presos a formatos e APIs fechados
- Implantação em meses, com caixa-preta e dependência total do fornecedor



Modelo DataXpertIA

- Custo previsível, sem licenças — investimento em engenharia e operação, não em royalties de software
- Zero lock-in: padrões abertos, código disponível e controle total do stack
- Implantação em semanas, com operação e evolução contínuas junto ao seu time

SERVIÇOS

Serviços para Cada Etapa da Sua Plataforma de Dados

SERVIÇO	O QUE ENTREGA	FOCO
Consultoria em Plataforma de Dados	Diagnóstico de maturidade, roadmap priorizado por impacto e caso de negócio	Decisão
Arquitetura e Design	Desenho do lakehouse, modelagem, fluxos de integração e governança	Fundação
Implementação e Deploy	Instalação completa, infraestrutura automatizada (IaC) e pipelines produtivos	Go-live
Migração de Legado	Plano gradual e seguro, sem parar a operação, com descomissionamento	Modernização
Suporte e Operação Contínua	Monitoramento proativo, suporte técnico com SLA e evolução contínua	Operação
Treinamento e Capacitação	Treinamento prático por perfil e documentação para autonomia do time	Autonomia



METODOLOGIA

Do Diagnóstico à Operação, em Semanas

- **1. Diagnóstico** — mapeamento do cenário atual, dores e prioridades
- **2. Arquitetura** — design do lakehouse, stack e roadmap sob medida
- **3. Deploy** — infraestrutura automatizada e primeira carga produtiva
- **4. DataOps** — pipelines versionados, testes automatizados e observabilidade
- **5. Governança** — catálogo, lineage e políticas de acesso alinhadas à LGPD
- **6. Operação Contínua** — suporte N2/N3, evolução contínua e capacitação do seu time

VALOR EM SEMANAS, NÃO ANOS

Cada fase tem entregáveis mensuráveis — você sabe o que recebe antes de avançar. Sprints curtas, comitê de decisões técnicas e de negócio, e transferência progressiva para o time interno.



MODELOS DE CONTRATAÇÃO

Três Formas de Contratar, um Compromisso com Resultado

Escolha o modelo que se encaixa no seu momento — todos sem licença de software, 100% open source



Projeto

Escopo Fechado

O QUE É

Escopo fechado e entrega por fases, com prazo e resultado definidos desde o início.

IDEAL PARA

Ideal para implantações, migrações e iniciativas com escopo claro.



Assinatura de Operação

Operação Contínua

O QUE É

Operamos, monitoramos e evoluímos sua plataforma de forma contínua.

IDEAL PARA

Ideal para empresas sem time interno dedicado à plataforma.



Consultoria sob Demanda

Horas de Especialista

O QUE É

Horas de especialista para resolver desafios pontuais de arquitetura ou engenharia de dados.

IDEAL PARA

Ideal para times que precisam de reforço técnico específico.



OBJEÇÕES

“E se...?” — Antecipando as Principais Dúvidas



Objeção: Suporte a Open Source

“Quem garante o suporte se algo der errado?”

- Operação contínua com SLA, monitoramento proativo e suporte N2/N3 — especialistas ao seu lado
- Comunidades globais ativas (Apache Software Foundation, Linux Foundation) e código aberto auditável — você nunca fica refém de um único fornecedor



Objeção: Dependência de Consultoria

“Vamos ficar dependentes de quem implantou?”

- Transferência progressiva de conhecimento: treinamento prático por perfil, documentação viva e acompanhamento pós-capacitação
- Seu time opera com total autonomia — a plataforma é sua, o código é aberto, os dados são seus



QUEM SOMOS

DataXpertIA — Dados, eXpertise e IA



Nossa Proposta de Valor

- **Plataforma, não projeto** — um ativo operacional que evolui com o seu negócio
- **Open source com suporte** — independência de vendor, com especialistas ao seu lado
- **Metodologia validada** — um caminho testado, do diagnóstico à operação contínua
- **Foco em resultados** — começamos pelos casos de maior valor para o negócio
- **Time sênior** — engenheiros e arquitetos certificados, com governança e conformidade LGPD

NOSSA VISÃO

Toda organização — pública ou privada — pode ter uma plataforma de dados moderna, governada e pronta para IA, sem abrir mão do controle sobre arquitetura e custos.

NOSSO COMPROMISSO

Entregar transformação inteligente de dados, com soberania tecnológica, custo previsível e decisões aceleradas por IA.

PRÓXIMO PASSO

Agende uma avaliação técnica gratuita — mapeamos o cenário atual e propomos uma arquitetura inicial em poucos dias.
contato@dataxpertia.com · +55 61 98657-1054 · dataxpertia.com