ProjectX

INTELLIGENT DATA OS · PILLAR OVERVIEW

An open lakehouse you can walk away from.

Apache Iceberg and S3 semantics, with pipelines, SQL, notebooks, machine learning and governed generative AI on top. Portable by construction: the storage format is open and the catalog is standard, so leaving is a copy rather than a migration project.

What it does

Iceberg storage with time travel

Snapshots, schema evolution and point-in-time reads. A query against last Tuesday returns what was true last Tuesday, not what is true now.

SQL, pipelines and notebooks

Trino and Spark over the same tables, with pipelines that declare their lineage rather than having it inferred afterwards.

Governed machine learning

Experiments, a model registry, a feature store with point-in-time correctness, and promotion gated on a model card and an approval. Every served model traces back to the snapshot it was trained on.

Generative AI with guardrails

Retrieval over your own governed data, with output schemas, budgets, a sovereign mode that keeps inference in-region, and the same entitlement filtering as every other read.

Catalog, lineage and sharing

Every table registered, every movement recorded as an edge, and sharing that is a grant rather than an export.

How it works

01 **Land your data once** · Console
Tables are stored in the open Apache Iceberg format on S3-compatible storage — readable by any engine, and yours to leave with.

02 **Work where the data lives** · Data workspace
SQL, pipelines, notebooks and machine learning run over the same governed tables, with lineage recorded as the work happens.

03 **Share by grant, not by export** · Catalog
The catalog registers every table and records every movement; sharing is an auditable grant, never a copied file.

Products

Lakehouse
Open Iceberg tables you can always take with you.

SQL & Analytics
SQL warehouses, dashboards and a query history you can trust.

ML & GenAI
Models you can trace, and AI that reads only what it may.

Questions buyers ask

Can we leave with our data?
Yes — that is the design. Tables are stored in the open Apache Iceberg format on S3-compatible storage, readable by any engine, so leaving is a copy, not a migration project.

Which tools can we use?
SQL, pipelines, notebooks and machine learning all run over the same governed tables, with lineage recorded as the work happens.

How is machine learning governed?
Experiments, a model registry and a feature store with point-in-time correctness; promotion is gated on a model card and an approval, and every served model traces to the snapshot it trained on.

What does it cost?
Query and storage usage at published rates on the pricing page, drawing from the same credit balance as the rest of the platform.