

Building the AI-Ready Foundation:

How IRiS Automates the AI-Ready Lakehouse, Incrementally

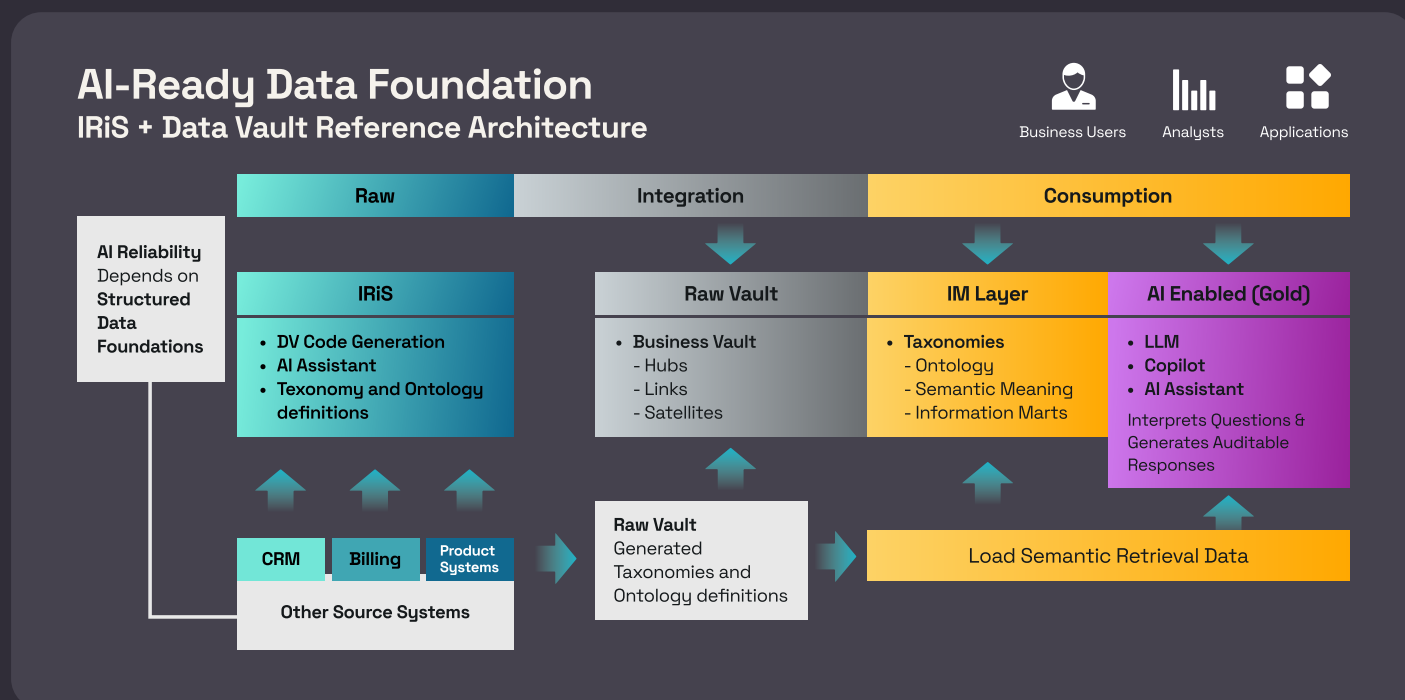
Most enterprise AI projects stall not because the models are inadequate, but because the data underneath them is inconsistent, incomplete, or untraceable. This brief explains how IRiS automates the Silver integration layer of your Lakehouse to solve that problem, delivering the semantic structure AI needs to reason correctly, without the manual effort and specialist dependency that typically makes it prohibitive. It outlines the reference architecture, what a real use case requires, what is and is not in scope, and how to sequence the build so AI capability is unlocked incrementally rather than deferred to a distant go-live.

WHO THIS IS FOR

Data and AI leaders, Lakehouse architects, and programme sponsors who want to understand what their Silver integration layer needs to deliver and how to build it at the pace the business requires.

THE REFERENCE ARCHITECTURE

The architecture below maps a complete enterprise AI stack to Medallion zones. If your organisation already operates a Medallion architecture, IRiS sits within and strengthens the Bronze-to-Silver transition, adding the structural rigour that raw Medallion alone does not provide and that AI workloads specifically require. The critical dependency is simple: every layer is only as reliable as the layer beneath it. Most AI investment targets the top of the stack. Most AI failure originates lower in the stack. Data Vault and IRiS address the integration and semantic layers, the layers that raw Medallion leaves structurally undefined.



Raw Medallion gives you zones. IRiS gives those zones a formal structure, stable identity, complete history, explicit relationships, and versioned definitions. That structure is what AI needs to reason correctly.

A CONCRETE USE CASE: CUSTOMER CHURN PREDICTION

Consider a financial services organisation that wants AI to identify customers at risk of churning within the next 90 days. The AI model requires five foundational data capabilities. Each is delivered structurally by IRiS:



CONSISTENT CUSTOMER IDENTITY

Customer records across CRM, billing, and product systems are resolved through Hub business keys, creating one stable entity per customer.



LONG-TERM BEHAVIOURAL HISTORY

Twenty-four months of behaviour is preserved in full through Satellites, timestamped and never overwritten.



CROSS-ENTITY RELATIONSHIP CONTEXT

Relationships such as products held, complaints raised, and campaigns responded to are explicitly encoded in Links, rather than inferred through joins.



CONSISTENT BUSINESS DEFINITIONS

A single definition of “churned” is implemented in the Business Vault as a versioned rule, ensuring that training labels remain consistent and reproducible.



AUDITABILITY

Every prediction input is traceable to its source system and the Business Vault rule applied. In regulated industries, this is a compliance requirement, not a nice-to-have.

Without these properties: Models train on inconsistent labels, historical context is incomplete, and outputs cannot be fully explained. The result may look plausible during testing, but it will not survive operational or regulatory scrutiny.

WHAT THIS IS AND WHAT IT ISN'T

IRiS automates the integration and semantic foundation of your Lakehouse. It is the foundation of the AI stack, not the entire stack itself. The table below clarifies what is included in the foundation and what belongs to later layers of the AI architecture:

CAPABILITY	DV + IRIS?	WHAT DELIVERS IT?
Stable entity identity across sources	✓ Included	Hub Business Keys
Complete temporal history	✓ Included	Satellites
Explicit relationship structure	✓ Included	Links
Versioned business definitions	✓ Included	Business Vault rules
End-to-end data lineage	✓ Included	DV structural audit trail
Natural language query interface	X Not included	LLM / Copilot layer (built on top of an Infomation Mart asset with a Semantic Model)
Retrieval / RAG pipeline	X Not included	RAG engine (scoped separately)
Business consumption dashboards	X Not included	Information Marts & reporting layer — Power BI, Tableau, etc.
ML model training and inference	X Not included	ML platform - Azure ML, Databricks MLflow, etc.

HONEST EXPECTATION

After Phase 1, you will have built the foundation the AI needs, not the complete AI product. Capabilities such as RAG pipelines and LLM interfaces sit in the next layer of the architecture. These can be scoped either alongside the foundation or immediately afterwards with your IRiS implementation partner.

WHERE IRIS FITS

IRiS is a Lakehouse automation platform that generates the Silver integration layer of your Medallion architecture, consistently, incrementally, and without requiring deep Data Vault specialisation in every team member. It implements Data Vault 2.0 methodology under the hood, which is what gives the generated structures their semantic rigour and AI readiness.



4.5× FASTER DELIVERY WITH 65% LESS ENGINEERING EFFORT

Automation removes repetitive engineering tasks so teams can focus on business logic.



SINGLE-SOURCE INCREMENTAL DELIVERY

IRiS generates one source table at a time, allowing AI capability to begin unlocking from the first sprint rather than after a multi-year build.



CONSISTENT SEMANTIC INTEGRITY

Every generated structure follows the same structural conventions. Quality does not depend on individual engineer discipline.



IRIS ASSISTANT

The assistant guides teams from conceptual model to Data Vault design in a single conversation, capturing business definitions along the way.

THE IRIS ASSISTANT IS AVAILABLE NOW

PHASED ROADMAP

The most effective approach is to build the foundation for the highest-priority subject area first, unlock AI capability for that domain, and then expand incrementally. A typical Phase 1 timeline with IRiS is four to eight weeks for one subject area across two to three source systems.

#	FOCUS	DELIVERABLE	AI CAPABILITY UNLOCKED
1	Foundation - first subject area	Raw Vault + Business Vault, core DQ and definition rules	AI can answer questions on this subject area with traceable, consistent results
2	Expand - additional sources	Further Hubs, Links, Satellites as business priorities drive scope	Cross-entity AI reasoning, customers and products, orders and suppliers
3	AI Stack - retrieval and interface	RAG pipeline from Business Vault to LLM; consumption layer	Business users query the platform in natural language with auditable responses
4	Scale - full platform + advanced AI	Complete semantic model; ML model integration	Enterprise-wide AI with full history, consistent semantics, end-to-end auditability

PHASE 1 SCOPE

2–3 source systems. One core subject area (e.g. Customer, or Customer + Product). Raw Vault generated via IRiS. Business Vault with core definitions and DQ rules for quality visibility. Timeline: 4–8 weeks. *Note – DQ management will be addressed in detail in a future article of this series.*

THE INVESTMENT CASE

Building the AI stack before the Lakehouse Silver layer is properly structured often means building twice: first to get something running, then again to make it trustworthy. Organisations that are succeeding with enterprise AI got the integration layer right first. They now reuse that investment across every new subject area and use case. Data Vault provides the foundation. IRiS makes building it achievable in weeks. An experienced IRiS implementation partner ensures that the methodology, automation platform, and AI stack operate as a coherent programme, rather than separate initiatives that must later be reconciled.

You do not have to choose between doing this properly and doing it fast. With IRiS, you can do both.



IRiS is a Lakehouse automation platform available on the Microsoft Marketplace, supporting Microsoft Fabric, Snowflake, and Databricks.

The IRiS Assistant is available now.

IRiS is implemented by a global network of authorised partners with deep Data Vault and AI stack expertise.

**READY TO SCOPE YOUR FOUNDATION?
TALK TO AN IRIS PARTNER**

ignition-data.com/iris