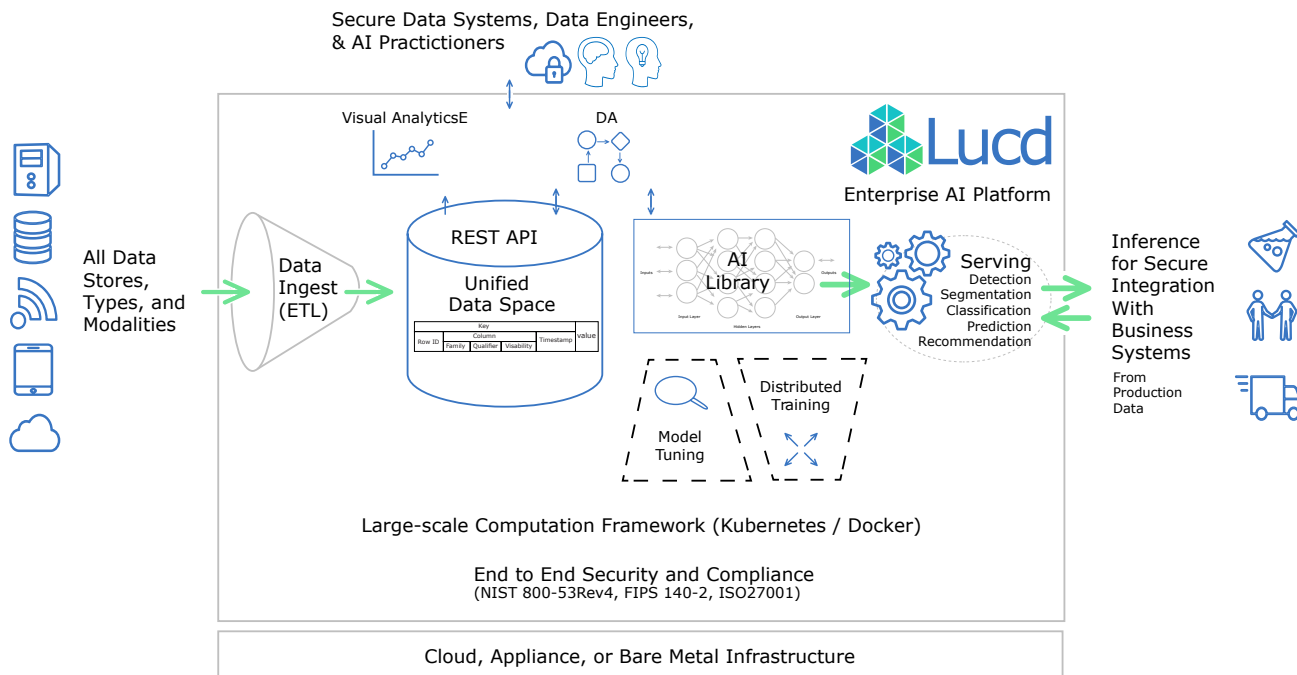# Lucd

The Platform for AI Innovation
Driving Enterprise AI Success

DATA INGEST ENGINE

The Lucd Data Ingest Module greatly eases the collection, feature extraction and use of many and multiple data sources to then be leveraged as part of creating AI/Machine Learning Business Benefit. The Lucd Data Ingest Module can ingest all data types and uses ML to extract features from unstructured data. The Lucd Data Ingest Module handles ALL data modalities, including binary and heavy data (i.e., large-scale datasets such as imagery databases).

PRODUCT ARCHITECTURE

BUSINESSES NEED A SECURE AI PLATFORM

In AI, more data usually means better results. ETL (Extract Transform and Load process) is the most difficult, tedious, and time-consuming part of the AI/Machine Learning Pipeline. *"As this New York Times article "For Big- Data Scientists, 'Janitor Work' Is Key Hurdle to Insights" describes, loading a big data platform with quality data with enough structure to deliver value is a lot of work. Data scientist spend a comparatively large amount of time in the data preparation phase of a project. Whether you call it data wrangling, data munging, or data janitor work, the Times article estimates 50%-80% of a data scientists' time is spent on data preparation."*
Link

The Platform for AI Innovation
Driving Enterprise AI Success

BUSINESSES NEED ETL AS PART OF THEIR AI PLATFORM

While an AI Platform gives AI Innovators in business the means to capture, secure, and harness data, the platform must also provide the means to deliver to solutions based on the results of their work. The AI Platform needs to provide the applications to extract, transform, and load data, perform exploratory data analysis and model building. But, the AI Platform also needs to be able to provide the means to build applications and deploy models as well as host third party applications that can then be transferred to the businesses environment. The AI Platform needs to present AI Solutions.

THE LUCD DATA INGEST ENGINE PROVIDES THE FULL ETL CAPABILITIES BUSINESSES NEED.

As with all parts of Lucd, the Data Ingest Module is also modular and extensible. There is interfaces to all parts of the ingestion framework so that support personnel can extend and customize all parts of the ingestion process, including ETL. All data types can be enqueued by the Data Ingest Module. This module provides the over-the-wire connectivity and server-side buffering required to ensure data is properly persisted and managed once it has entered into the Lucd platform. One of the core functions of the Data Ingest Module is to extract and maintain all data pedigree information. This includes all of the source information regarding the data feed, any historical information, data provenance details, as well as any data ingest engineer comments and additional information.