

Chevron: Harmonizing facility and equipment data on the MarkLogic platform

Lance Stafford, Enterprise Architect, Chevron
Damon Feldman, Solutions Director, MarkLogic



human energy®

Harmonizing facility and equipment data on the MarkLogic platform

Lance Stafford, Enterprise Architect
Damon Feldman, Solutions Director

Marklogic World 2019
Washington DC,

This document is intended only for use by Chevron for presentation at the Marklogic World Conference 2019. No portion of this document may be copied, displayed, distributed, reproduced, published, sold, licensed, downloaded, or used to create a derivative work, unless the use has been specifically authorized by Chevron in writing.

A leading international energy company

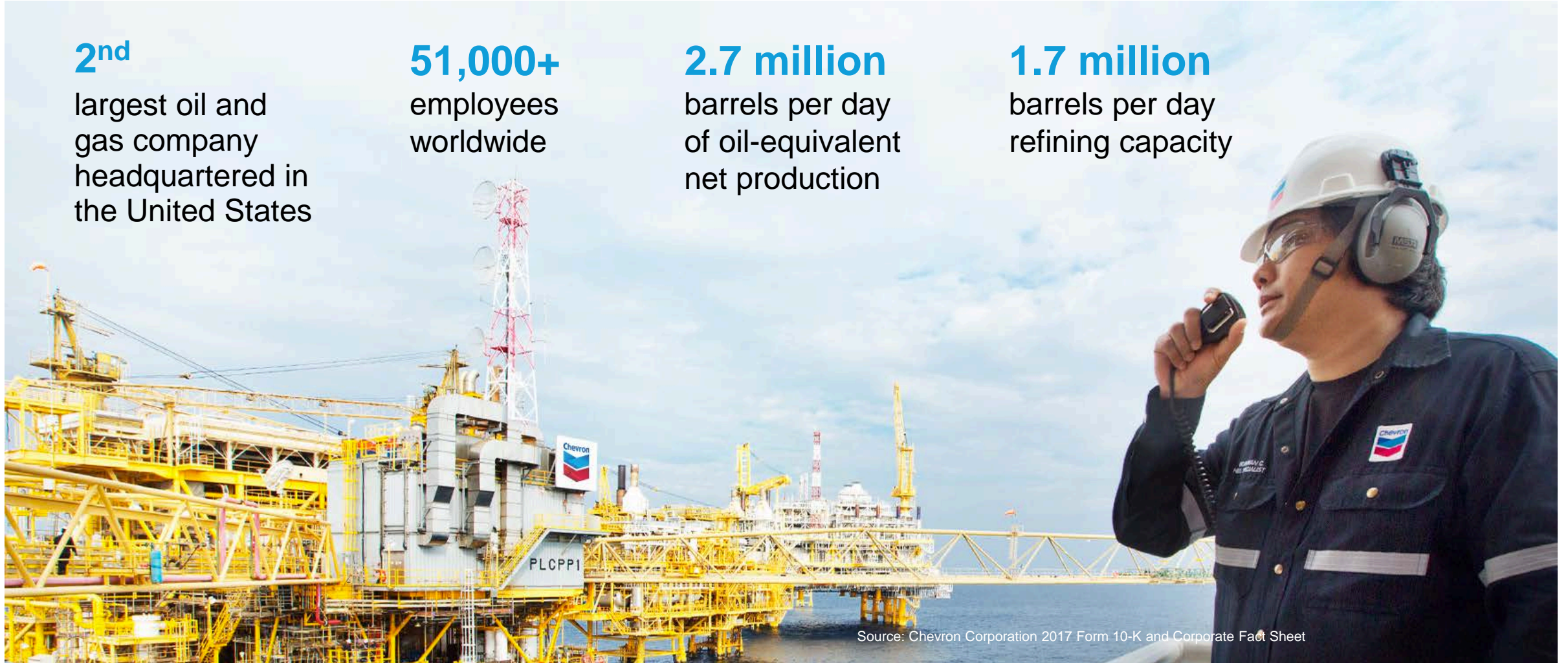
2nd

largest oil and
gas company
headquartered in
the United States

51,000+
employees
worldwide

2.7 million
barrels per day
of oil-equivalent
net production

1.7 million
barrels per day
refining capacity

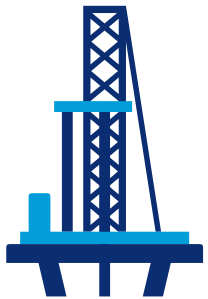


Source: Chevron Corporation 2017 Form 10-K and Corporate Fact Sheet



Involved in every facet of the energy industry

Upstream



Explore



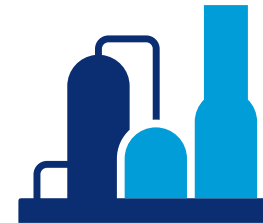
Produce

Midstream

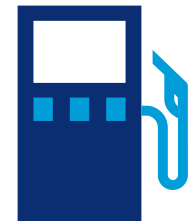


Transport

Downstream



Refine



Market



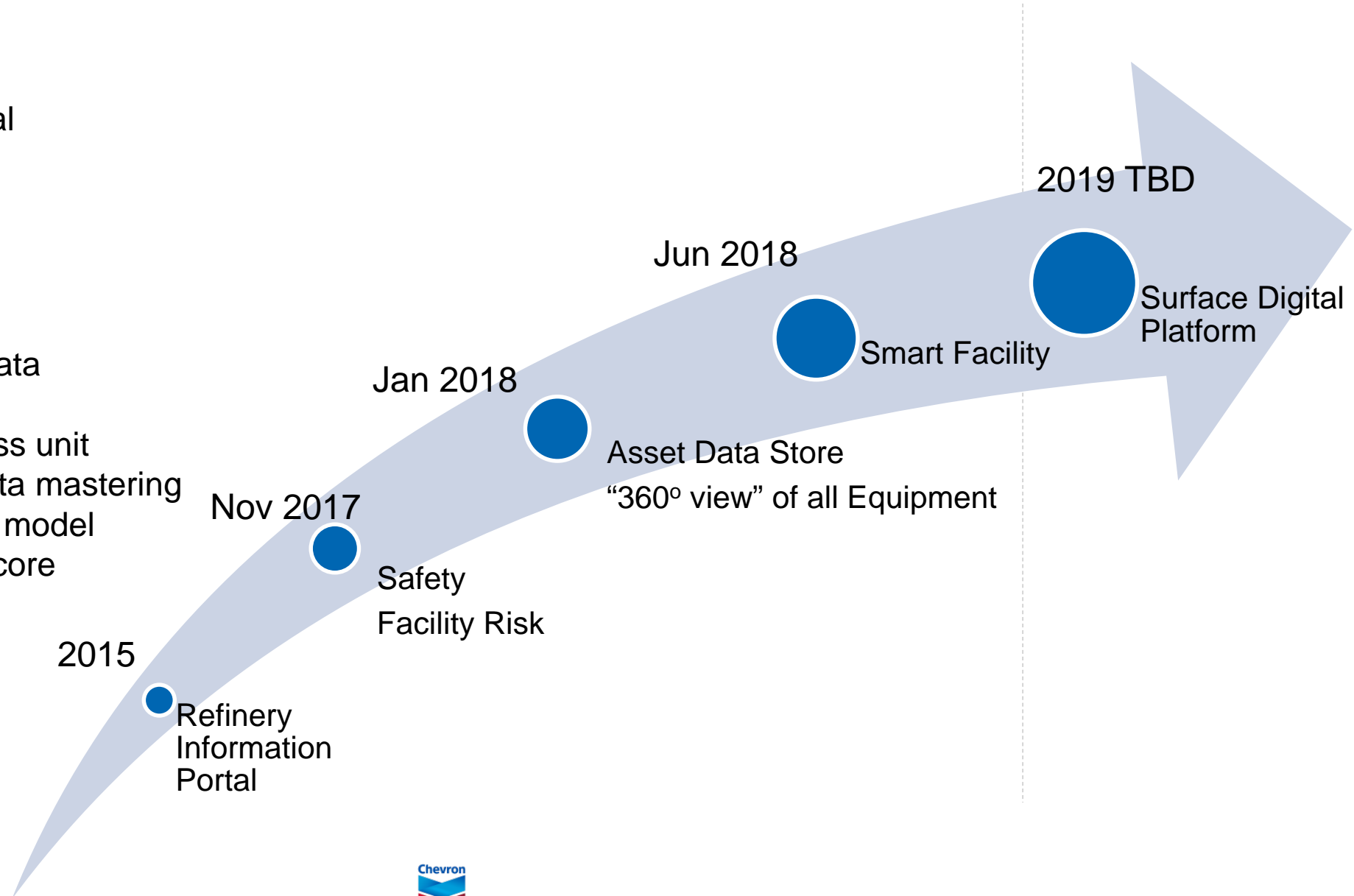
Marklogic deployment and projects

Project

- Refinery Information Portal
- Facility Risk
- Asset Data Store
- Smart Facility

Capabilities

- Structured/unstructured data discovery
- Security based on business unit
- Facility and equipment data mastering
- Enterprise canonical data model
- Data quality confidence score
- Data mapping GUI
- Linking sensor data to equipment



Equipment mastering implementation - Asset Data Store (ADS)

Value

- Silo-busting
- Provides one trusted place for harmonized information
- A standard format

Breadth

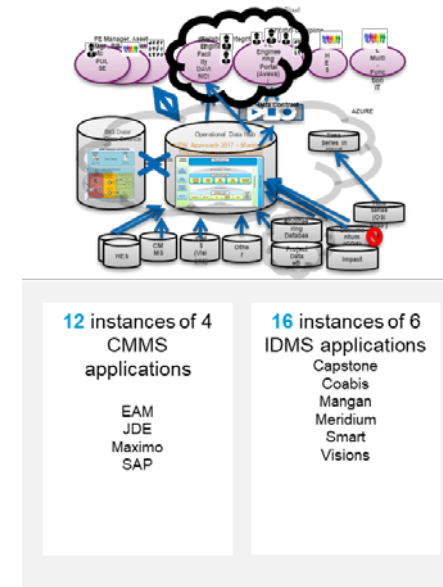
- Combines: safety, operations, maintenance, construction for Facilities

Trust

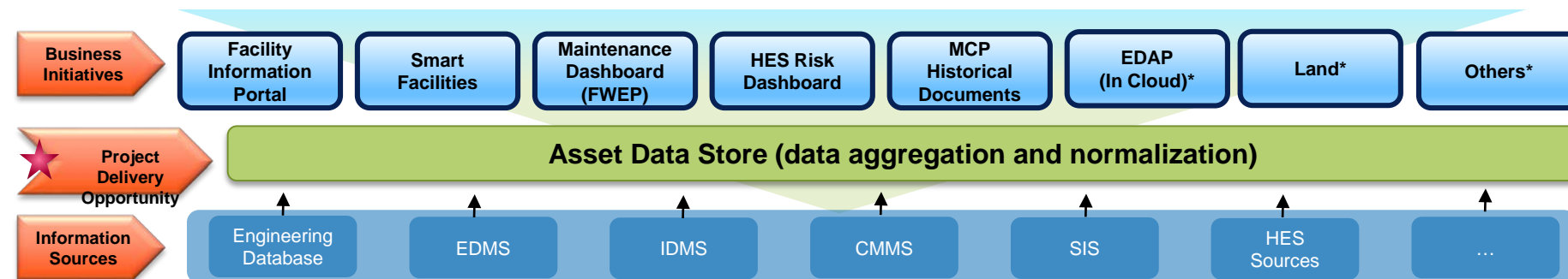
- Resolve data inconsistencies

Technology

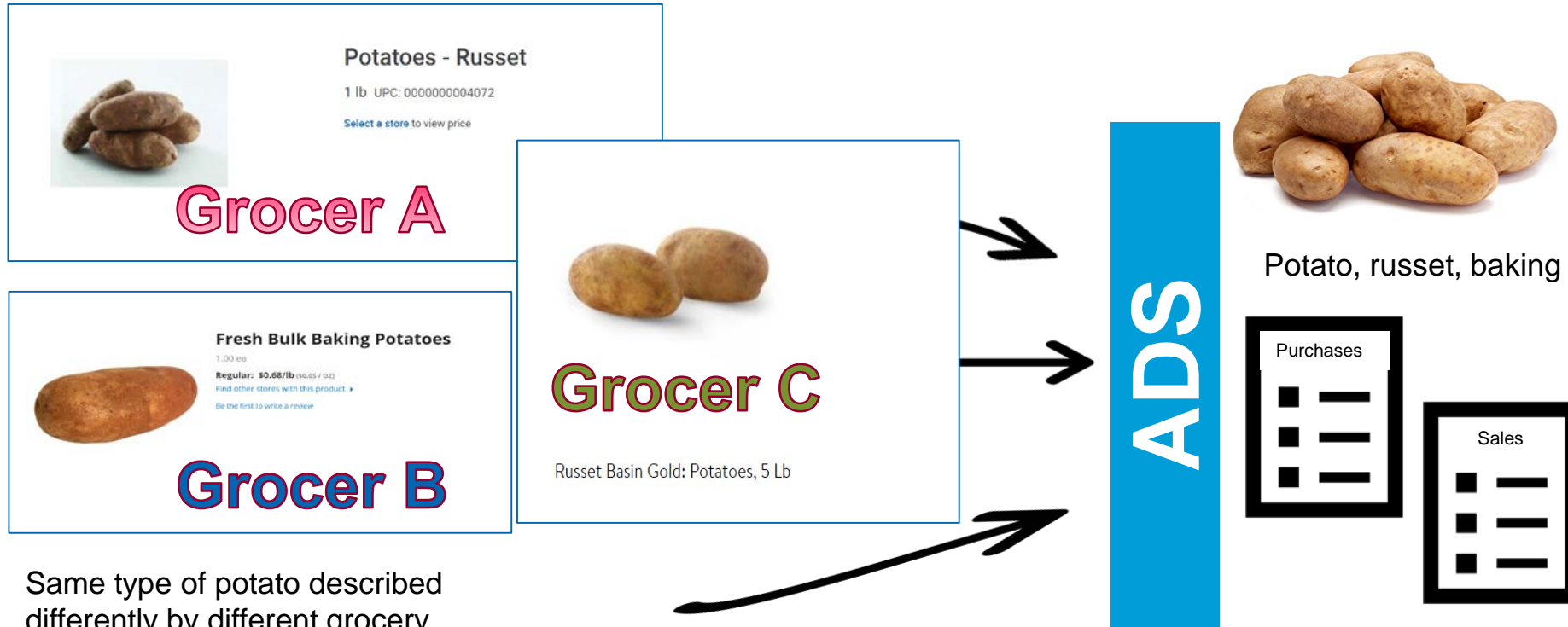
- One-stop shop for cross-functional, consolidated, standardized information



*28 different instances
Across 10 system types*



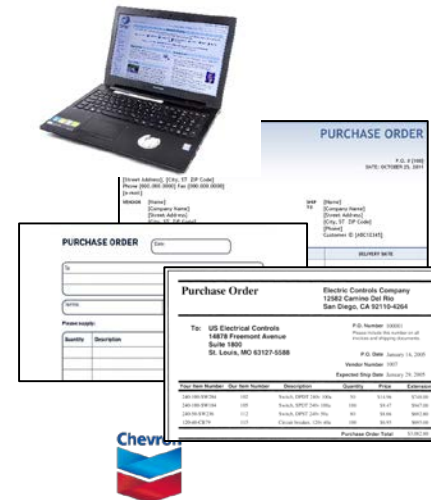
Mastering concept - you say potayto, I say “pototoe”



Same type of potato described differently by different grocery stores

ADS concept combines and harmonizes the description of the potato across grocery stores. Point of Sale and Procurement transactions are aggregated and harmonized against the standard potato description.

Point of Sale transactions



Why mastering matters

Chevron's IT Guiding Principles highlight the importance of proper master data management.

Operational Excellence (OE) - Operational Excellence requires getting the right information to the right person at the right time. Master data allows a single, comprehensive, high quality record for each piece of equipment, facility or activity.

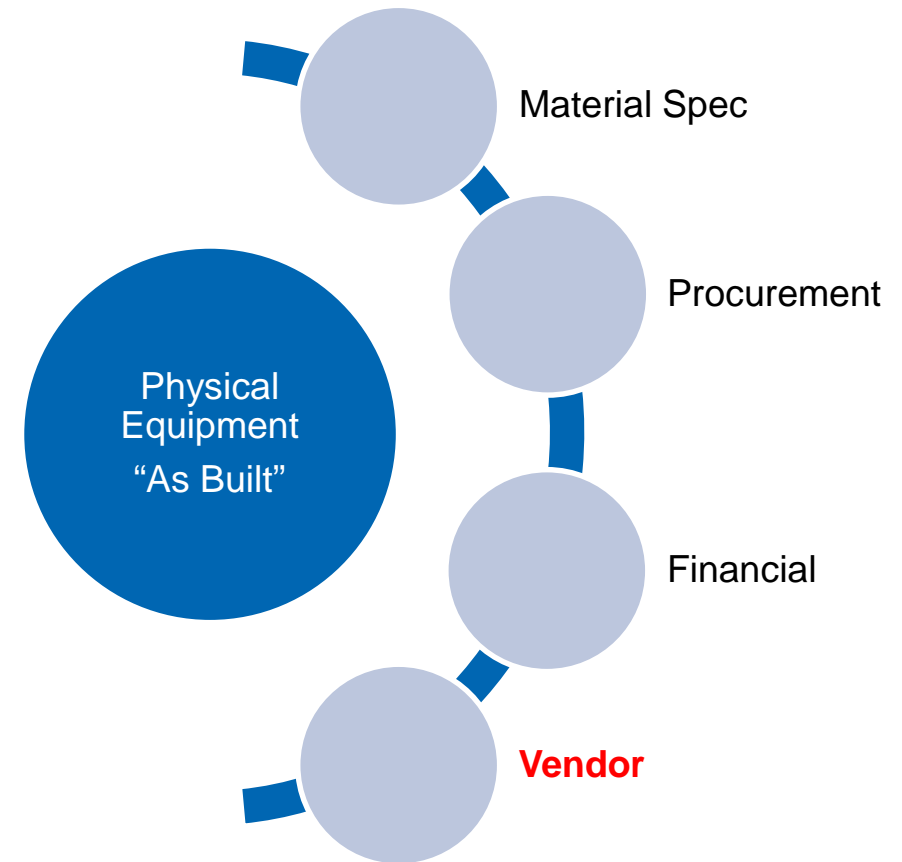
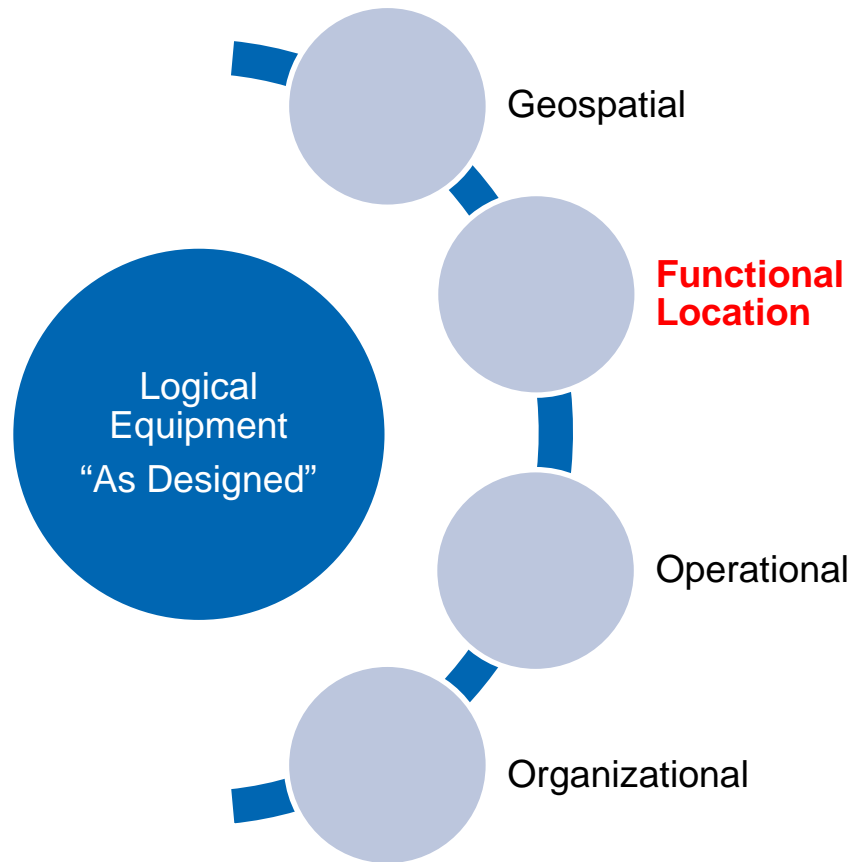
Standardization - Standardization is the keystone for master data and makes information integration technically feasible.

Integration - Master data provides consistent semantics (meaning) between data repositories. Without master data, there is little possibility of information integration.

Business process optimization - Business process optimization can occur when high quality master data enables improved decision making. There is a direct link between data quality and the ability to understand and optimize the business.



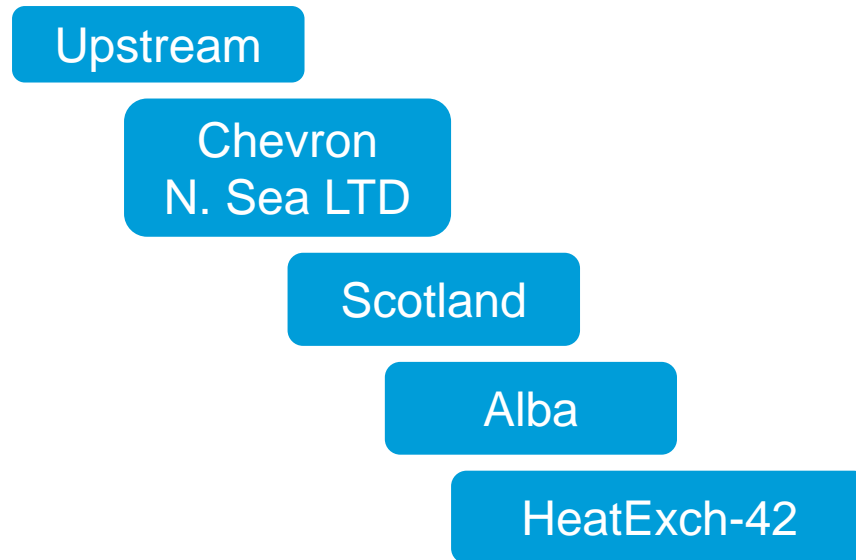
Mastering addresses multiple hierarchies and one universal key



Example: align disparate data by using the natural key hierarchy

Inspections system

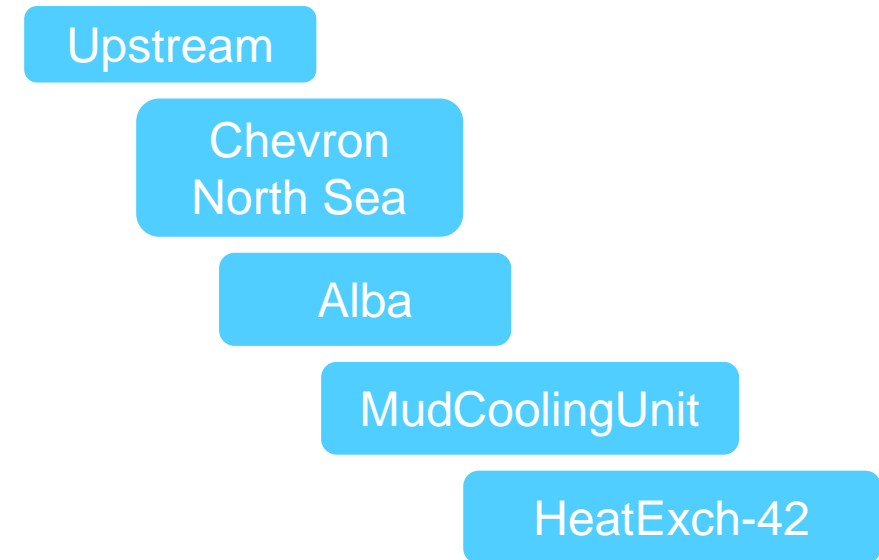
“How many heat exchangers fouled in the North Sea area last year? Which models?”



Country/region is needed for the inspection work processes and fleet analytics

Maintenance system

“What are some work orders Jamie the underwater welder can do when she’s at Platform XYZ?”

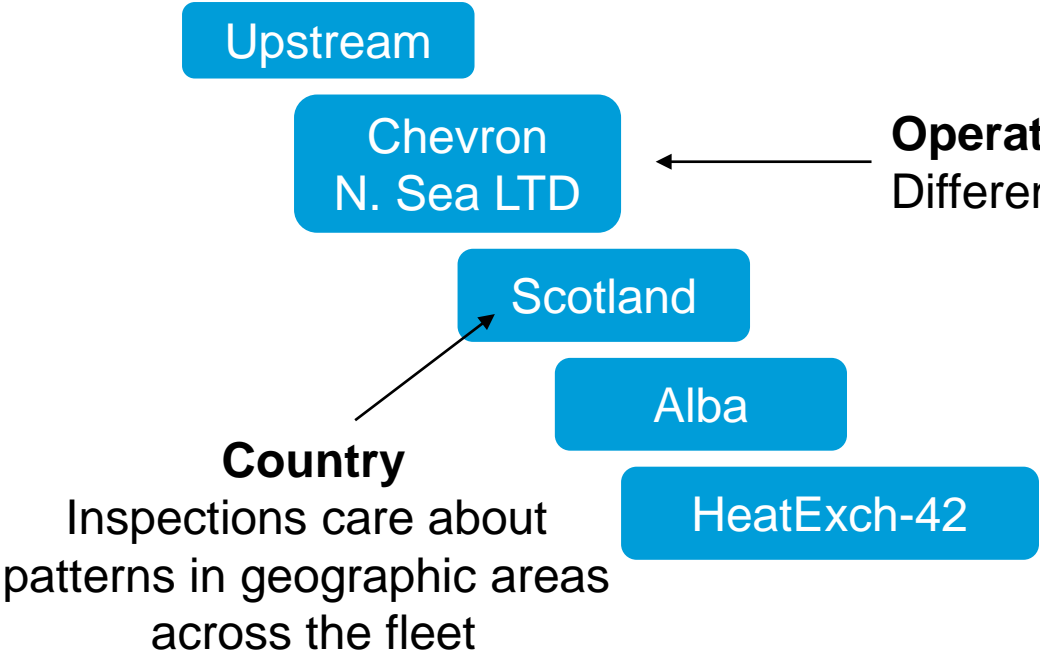


No country, but the maintenance systems need units (areas on the platform) to localize equipment locations for work order activity

Example: align disparate data by using the natural key hierarchy

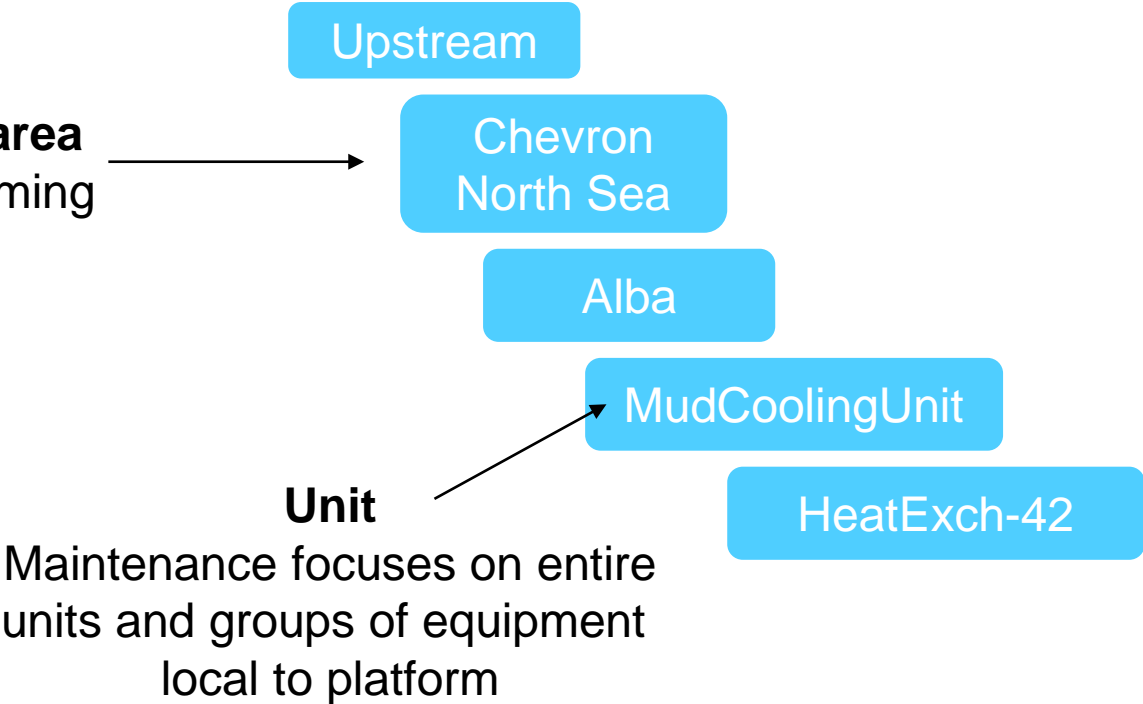
Inspections system

“How many heat exchangers fouled in the North Sea area last year? Which models?”



Maintenance system

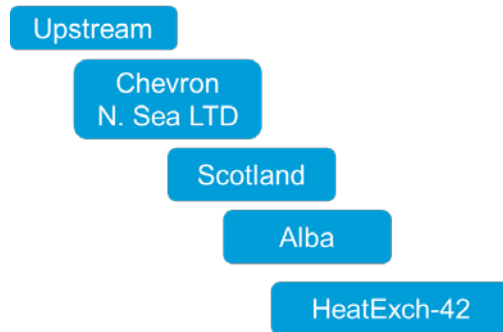
“What are some work orders Jamie the underwater welder can do when she’s at Platform XYZ?”



Example: align disparate data by using the natural key hierarchy

- Inspectors and Maintenance groups can retain use of their own hierarchies internal to their systems
- The Chevron ADS Data Hub aligns disparate systems to a single, universal human readable natural key

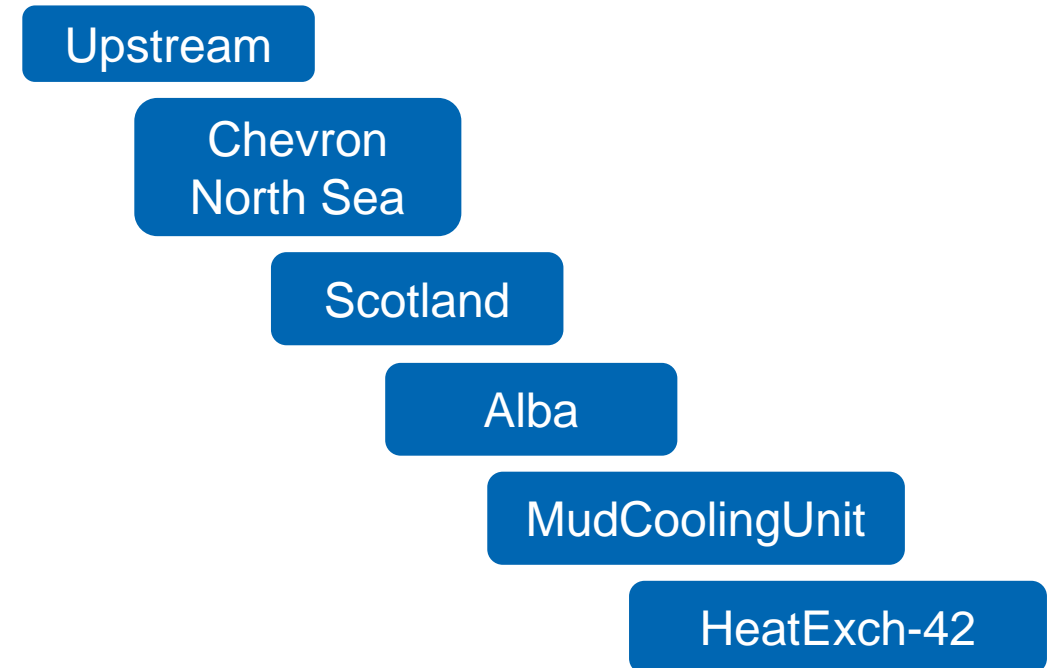
Inspections System



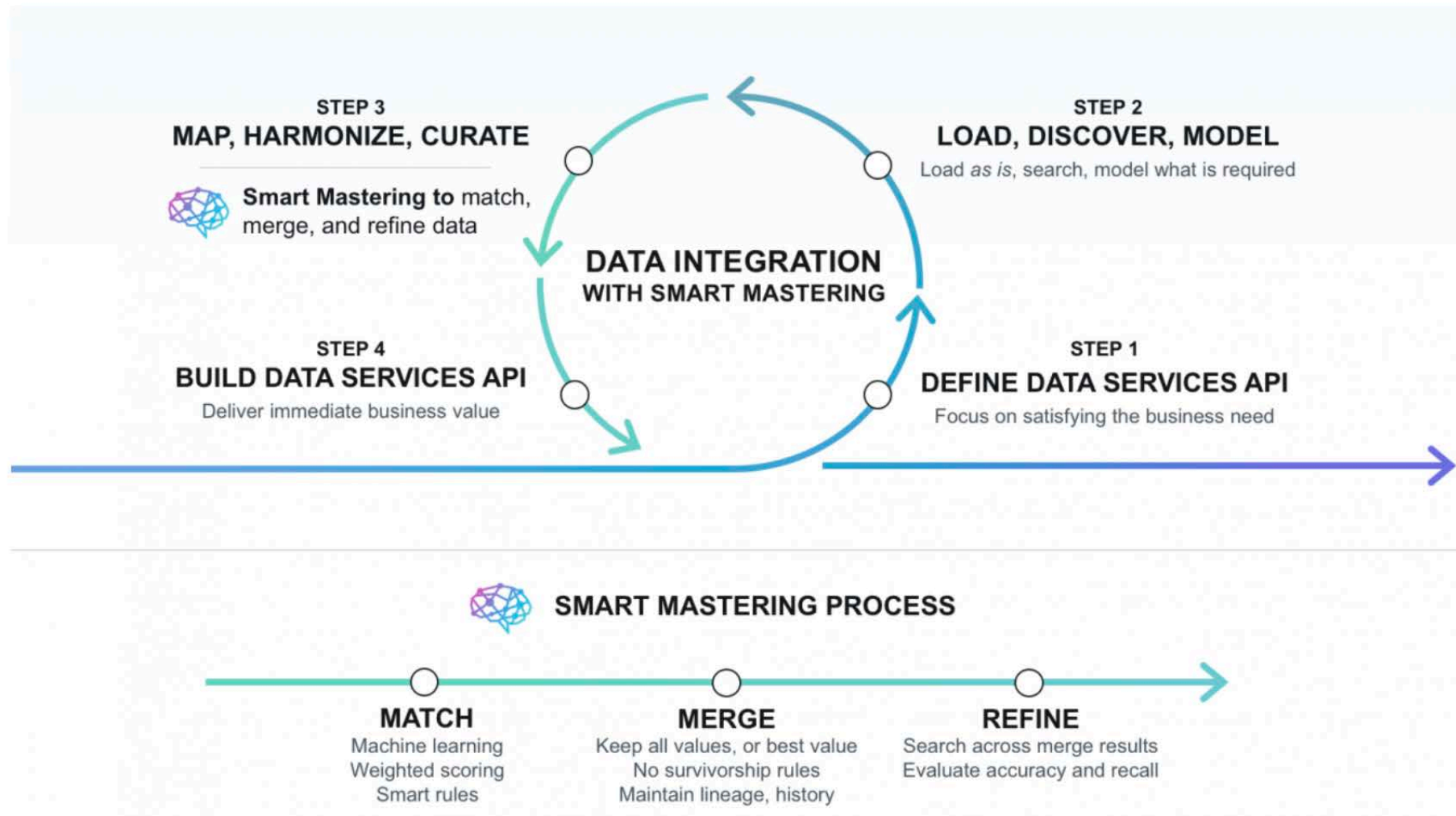
Maintenance System



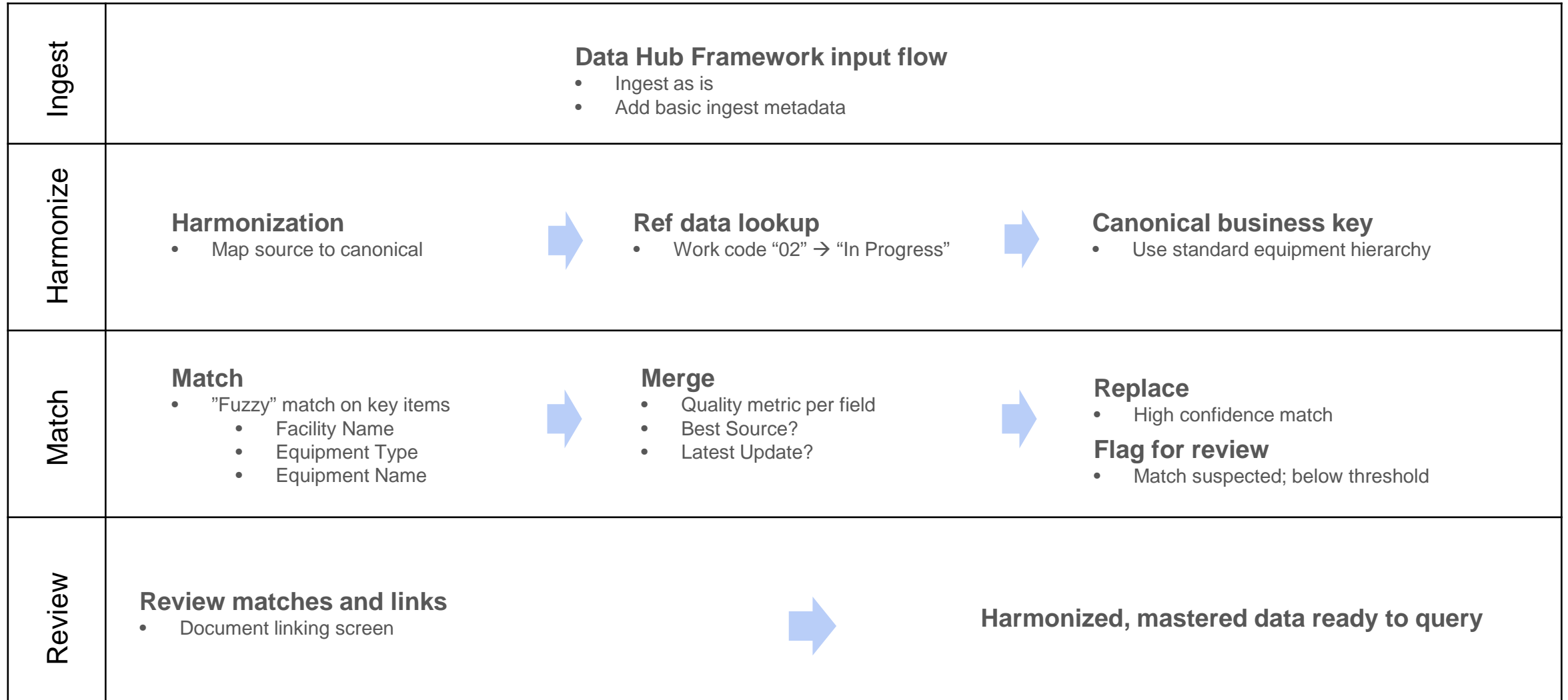
Enterprise View



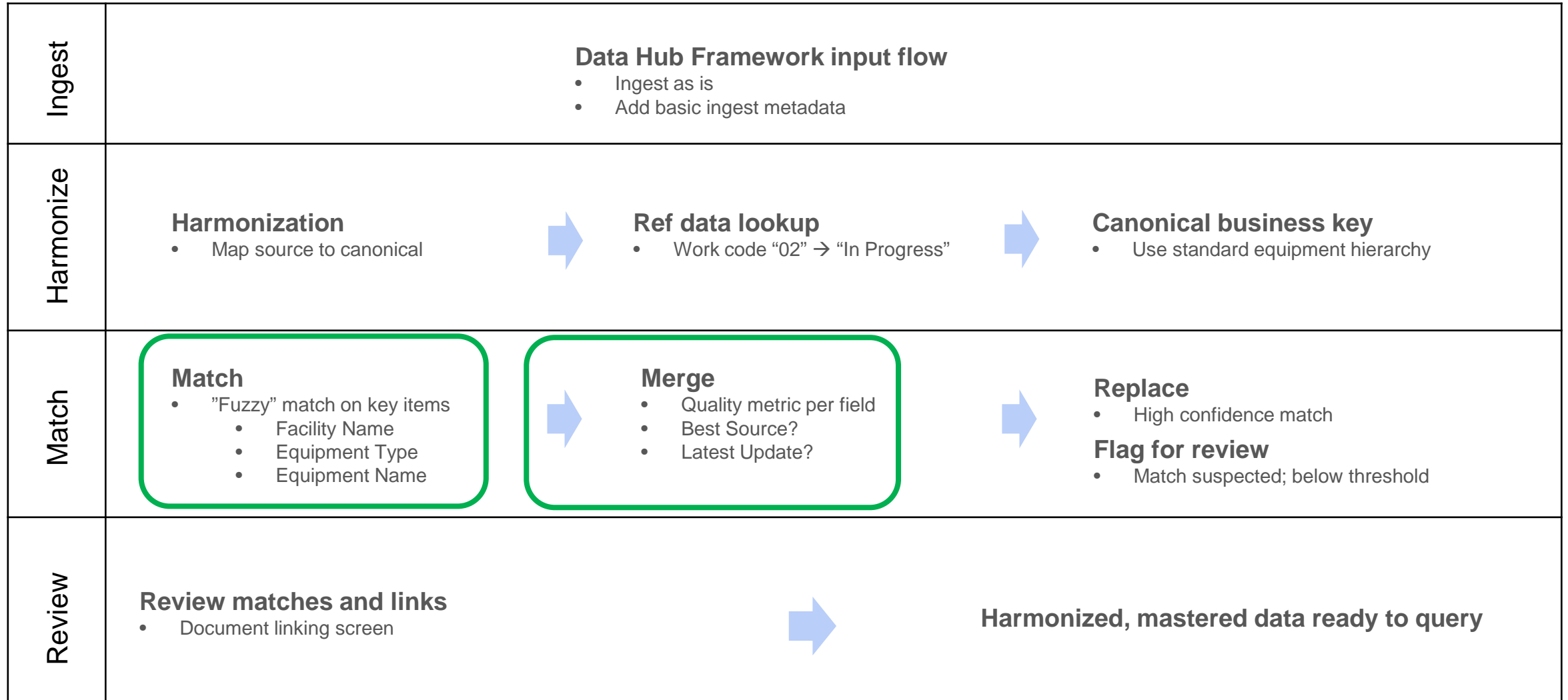
How implemented technically - MarkLogic Smart Mastering



Chevron data flow (Data Hub Framework and Smart Mastering)



Chevron data flow (Data Hub Framework and Smart Mastering)



The solution approach - matching (tech deep dive)

```
<equipment>
  <type>Heat Exchanger</type>
  <name>Heat Exchanger - Mud-cooling #42</name>
  <hierarchy>
    <OpCo>Upstream</OpCo>
    <businessUnit>CUE
    <Operating Area>Chevron North Sea</businessUnit>
    <Facility>Alba</Facility>
    <unit></unit>
    <equipmentID>HeatExch-42</equipmentID>
  </hierarchy>
[ . . . ]
```

```
<equipment>
  <type>Heat Exchanger</type>
  <name>Heat Exchanger - Mud-cooling #42</name>
  <hierarchy>
    <OpCo>Upstream</OpCo>
    <businessUnit>Chevron Upstream Europe</businessUnit>
    <Operating Area>Chevron N Sea LTD</businessUnit>
    <country>Scotland</country>
    <Facility>Alba</Facility>
    <unit>MudCooling</unit>
    <equipmentID>HeatExch-42</equipmentID>
  </hierarchy>
[ . . . ]
```

Match

- "Fuzzy" match on key items
 - Facility Name
 - Equipment Type
 - Equipment Name

Field	Matches?	Score	Confidence
OpCo	Yes	3 pts	100%
Operating Area	Close	3 pts	75% - close match
Facility	Yes	4 pts	100%
Unit	No	0 pts	
Equipment ID	Yes	20 pts	100%
		30 PTS	80% Match Confidence



The solution approach - MarkLogic Smart Mastering

Match

- "Fuzzy" match on key items
 - Facility Name
 - Equipment Type
 - Equipment Name

```
<equipment>
  <type>Heat Exchanger</type>
  <name>Heat Exchanger - Mud-cooling #42</name>
  <hierarchy>
    <OpCo src="CMMS" conf="100">Upstream</OpCo>
    <Business Unit src="CMMS" conf="100">CUE</Business Unit>
    <Operating Area src="CMMS" conf="75" Chevron North Sea</Operating Area>
    <Facility src="CMMS" conf="100">Alba</Facility>
    <unit src="CMMS" conf="100">Mud Cooling Unit</unit>
    <equipmentID src="IDMS" conf="75">HeatExch-42</equipmentID>
  </hierarchy>
  <country src="IDMS" conf="75" >Scotland</country>
[ . . . ]
```

- Fuzzy search for every record
 - Inspections, work orders, design documents
- Tracks provenance src="IDMS" src="CMMMS"
- Tracks confidence scores (Chevron extension) conf="75"



The solution approach - merging (tech deep dive)

- Different Values across systems

<businessUnit>Chevron N Sea LTD</businessUnit>

<businessUnit>Chevron North Sea</businessUnit>

- Different hierarchies and “natural keys”

- **Country** is captured in the inspection system
- **Unit** for is captured in the maintenance system

- Idiomatic conventions

– BusinessUnit = “North Sea / Scotland”

- Here a country prefix “shoe horns” another field into the Business Unit
- As these are stripped, we decrease the resulting data confidence

- Customizable confidence

– Additional rules examine other data to modify confidence scores, per source system

Merge

- Quality metric per field
- Best Source?
- Latest Update?



The benefits - core capabilities and goals

- Standardized data available to all consumers
- Alignment of critical data across functions
 - Facilities Engineering
 - Base Business
 - Health, Environmental, Safety
- Alignment of critical data across divisions and operating companies within Chevron
 - Upstream
 - Downstream
 - Midstream
 - Other (Capital Projects, Complex Process Facilities)
- Uniform identifiers
- Uniform data services and formats



The benefits - solution rollout - the “factory”

- Onboard new data sources in 2-4 weeks
 - Much of the hard work is done – we have invested for the future onboarding new entities and systems of record
 - Empowered the business analysts to map and configure
 - Minimized developer coding
- Provide data to downstream consumers and functions
 - Data Science analyses
 - Time series exports and analysis
- Future plans include
 - More equipment-related data sources
 - Regional federation for fast access
 - Migration to Azure internal cloud



MarkLogic Smart Mastering



Mastering, MDM and Smart Mastering

- **Mastering:** Creating and controlling a single, consolidated entity
- **MarkLogic Smart Mastering:**
 - Matching
 - Merging
 - Provenance
- **Enables a data hub**
 - Many sources naturally leads to overlaps and duplicates
- **Valid reporting and analytics**

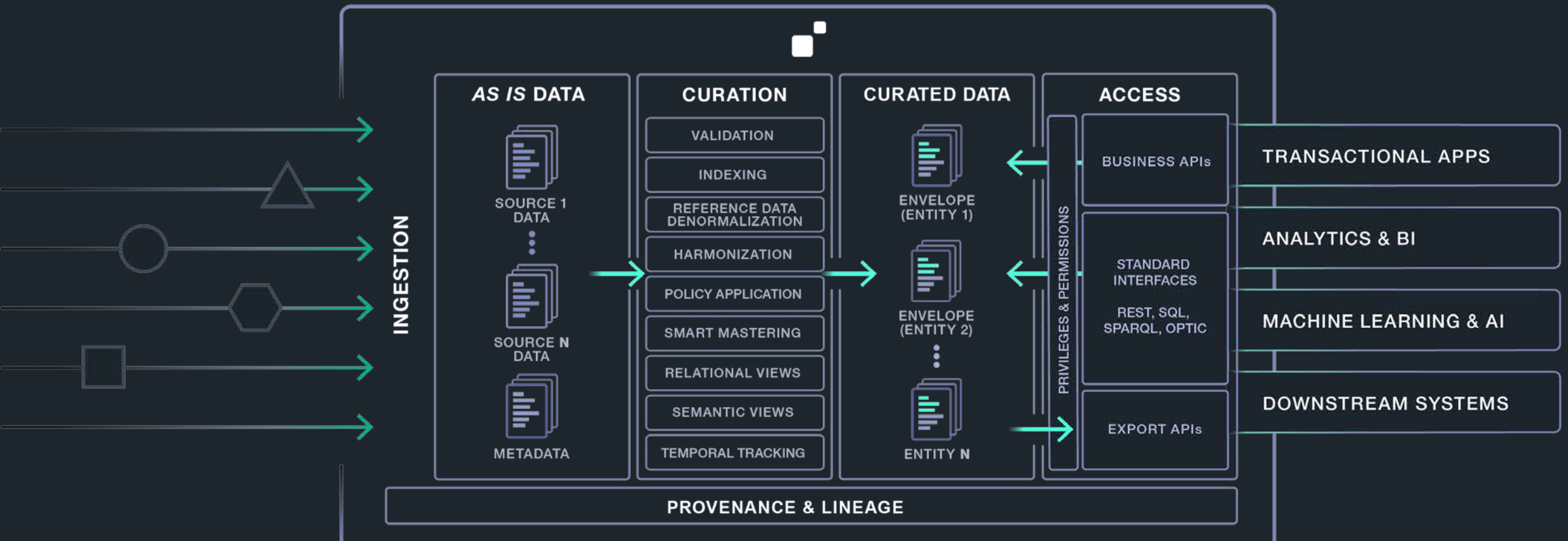


The Value of Mastering

- **Different Business Units** can use the same values for reporting in Apples to Apples Comparison (fleet analytics)
 - By equipment type or “class”
- **Deduplicated, harmonized data** is clean data
 - “How many repairs occurred in facilities with Acme Operations Corp overseeing maintenance activities?”
- **Provenance and Consolidation** = trust
 - I see this pump has had 10 repairs in the last year!
 - Who said that? Where did this data come from?



MarkLogic Data Hub and Smart Mastering



MarkLogic Data Hubs & Smart Mastering Summary

- **Bring** many data sources together
- **Specialization** in system X = data quality issue in the data hub
- **Deal** with variation and differences
 - Identify the same entities across sources
 - Globally-useful keys
- **Merge** and clean up data
- **Expose** consolidated, curated data globally, including real-time access





Thank you