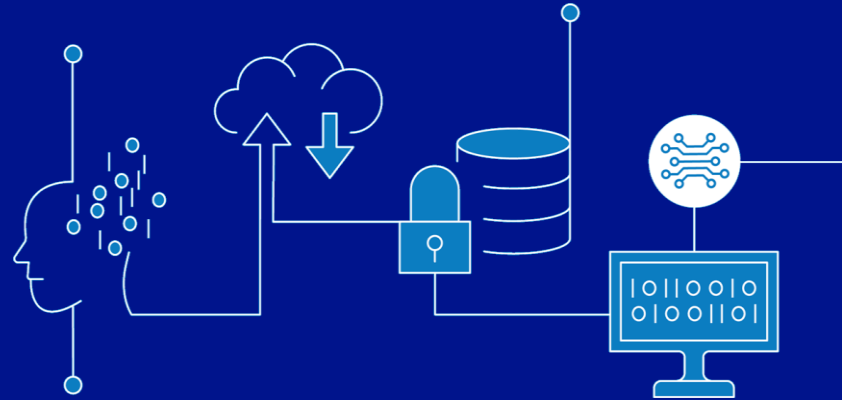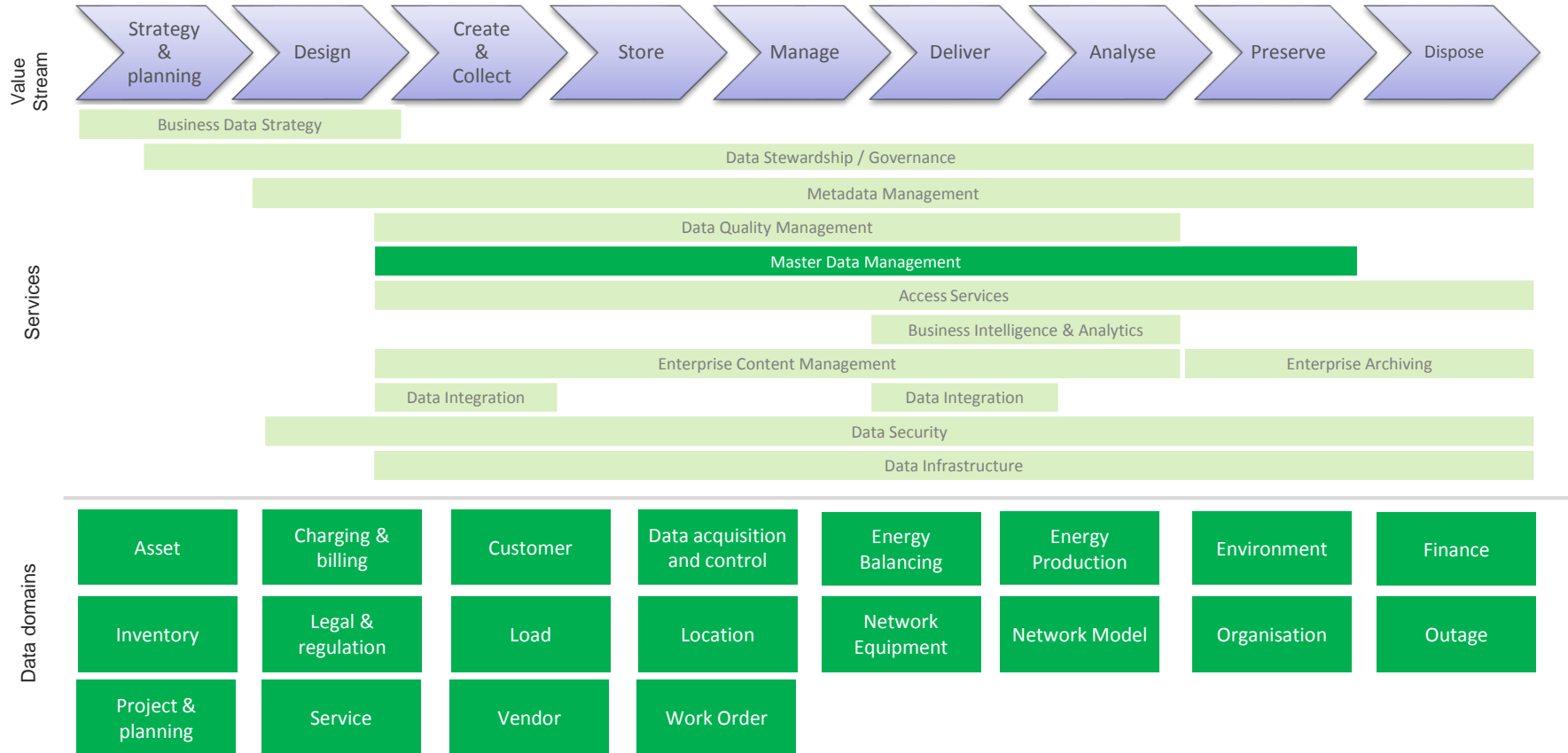# Master Data Management
Data Technology Services

Enterprise Architecture

Classification: Confidential
Author   Enterprise Architecture
Date   Q3 2016
Version   v0.3
Status   PROPOSED

Information Management Value Stream

Value Stream	Strategy & planning	Design	Create & Collect	Store	Manage	Deliver	Analyse	Preserve	Dispose

Services

- Business Data Strategy
- Data Stewardship / Governance
- Metadata Management
- Data Quality Management
- Master Data Management
- Access Services
- Business Intelligence & Analytics
- Enterprise Content Management
- Enterprise Archiving
- Data Integration
- Data Integration
- Data Security
- Data Infrastructure

Data domains

Asset	Charging & billing	Customer	Data acquisition and control
Inventory	Legal & regulation	Load	Location
Project & planning	Service	Vendor	Work Order

Energy Balancing	Energy Production	Environment	Finance
Network Equipment	Network Model	Organisation	Outage

nationalgrid | **information**technology

Master Data Management Strategy Map

Vision

" Bring under management and control, the 'golden version' of our key data that drives our core business processes, to enable improved performance and better process outcomes"

Principles

MDM is not 'one time' its an ongoing activity

Business Data Management Principles

EA / IT technology Principles

Objectives

Business
- Sustain operations
- Reduce data management risk
- Reduce data management costs
- Improve data quality and integration across multiple sources

Information
- Common vocabulary, glossary, terms ,definitions
- Master data 'hubs' established for priority data domains
- Business data entities standardized, held in a common repository (used across multiple business processes)
- Maintain match rules
- Right level of security
- Manage changes to reference and master data

Application
- Systems of record
- Leverage data integration, metadata management, data quality services
- Application adaptors and connectors
- Supports for batch and transaction level processing
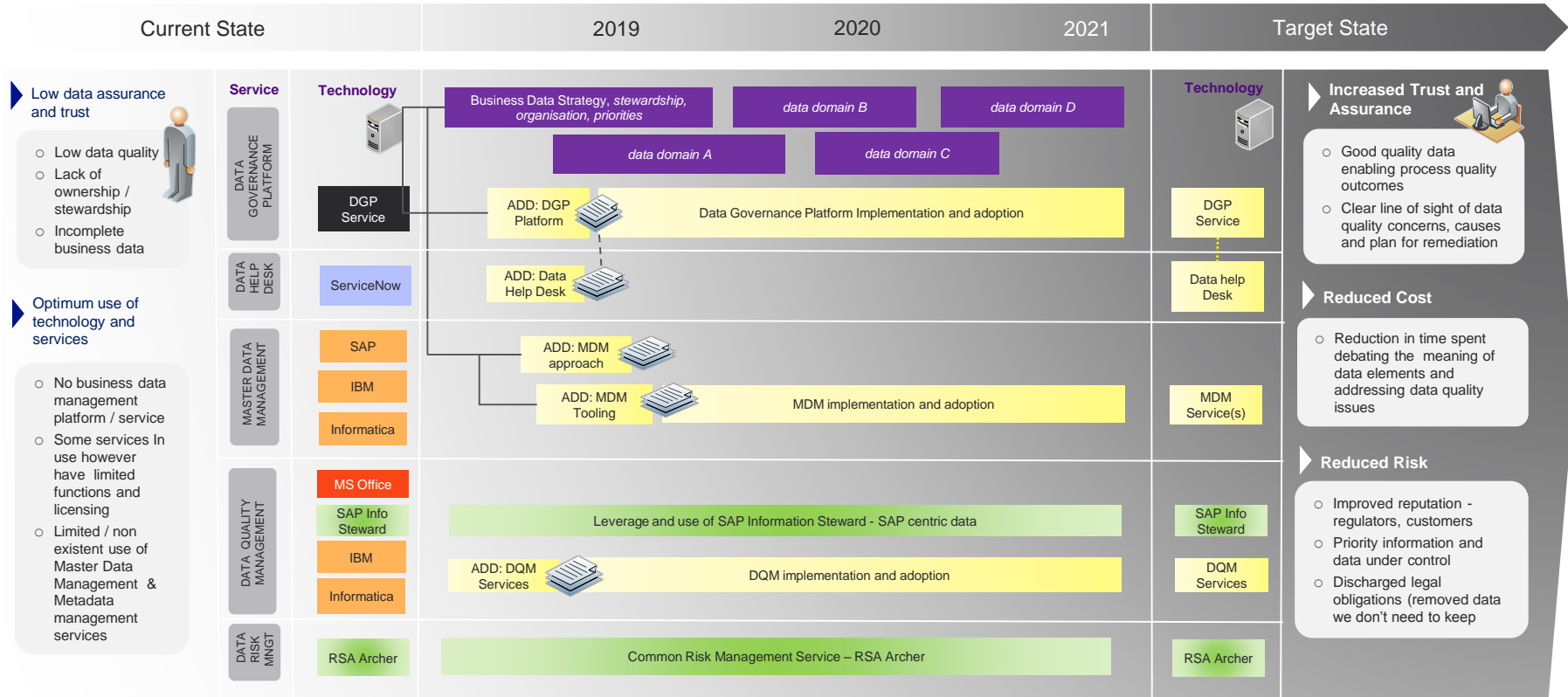- MDM monitoring and reporting services deployed

Technology
- Hybrid architecture relatively cost effective and time efficient to a full blown MDM Repository solution
- Service based Integration support
- Scalable solution
- Workflow and versioning support
- Archive service

,

Outcomes

- Master data 'hubs' established for priority data domains
- Business driven roadmap – master data priorities driven by key initiatives / demands (not IT centric)
- Data stewardship & data governance processes established for management and distribution of master data (simple at first, evolves over time)
- Enable consistent reference and master data across multiple consumers / applications
- Supports consolidated 360 degree view of information for effective reporting and analysis

- Common conceptual / logical MDM architecture
- Multiple applications maybe used for different master data objects

Data Management Technology Services (SOAP)

| Planned | GAP | Phase out | Standard | Containment | Architecture Design Decision |

| | Current State | | 2019 | 2020 | 2021 | Target State |

Low data assurance and trust
- Low data quality
- Lack of ownership / stewardship
- Incomplete business data

Optimum use of technology and services
- No business data management platform / service
- Some services In use however have limited functions and licensing
- Limited / non existent use of Master Data Management & Metadata management services

Service / **Technology**

DATA GOVERNANCE PLATFORM — DGP Service

Business Data Strategy, *stewardship, organisation, priorities*
data domain A
data domain B
data domain C
data domain D

ADD: DGP Platform — Data Governance Platform Implementation and adoption

DATA HELP DESK — ServiceNow — ADD: Data Help Desk

MASTER DATA MANAGEMENT — SAP, IBM, Informatica
ADD: MDM approach
ADD: MDM Tooling — MDM implementation and adoption

DATA QUALITY MANAGEMENT — MS Office, SAP Info Steward, IBM, Informatica
Leverage and use of SAP Information Steward - SAP centric data
ADD: DQM Services — DQM implementation and adoption

DATA RISK MNGT — RSA Archer
Common Risk Management Service – RSA Archer

Technology (Target State)
- DGP Service
- Data help Desk
- MDM Service(s)
- SAP Info Steward
- DQM Services
- RSA Archer

Increased Trust and Assurance
- Good quality data enabling process quality outcomes
- Clear line of sight of data quality concerns, causes and plan for remediation

Reduced Cost
- Reduction in time spent debating the meaning of data elements and addressing data quality issues

Reduced Risk
- Improved reputation - regulators, customers
- Priority information and data under control
- Discharged legal obligations (removed data we don't need to keep

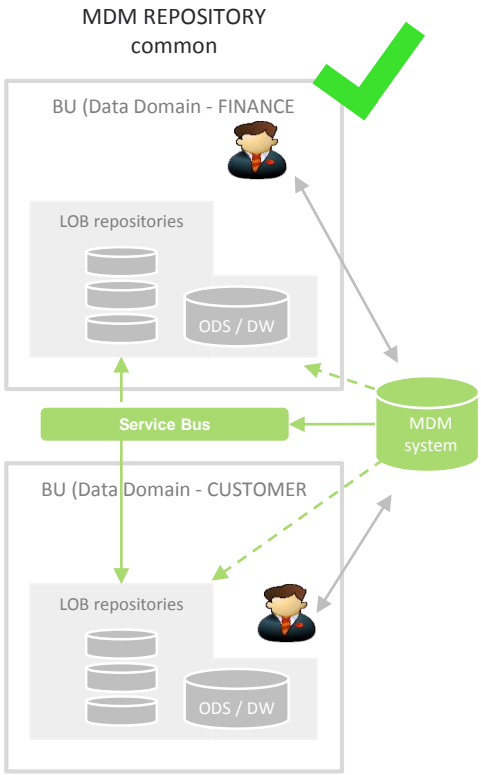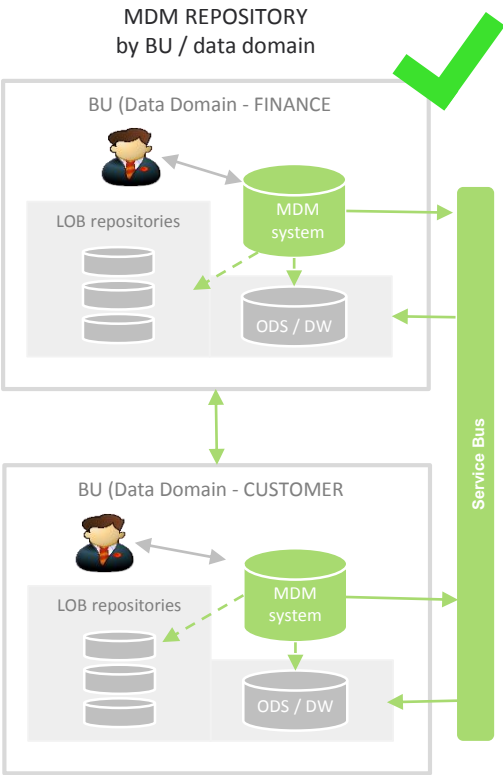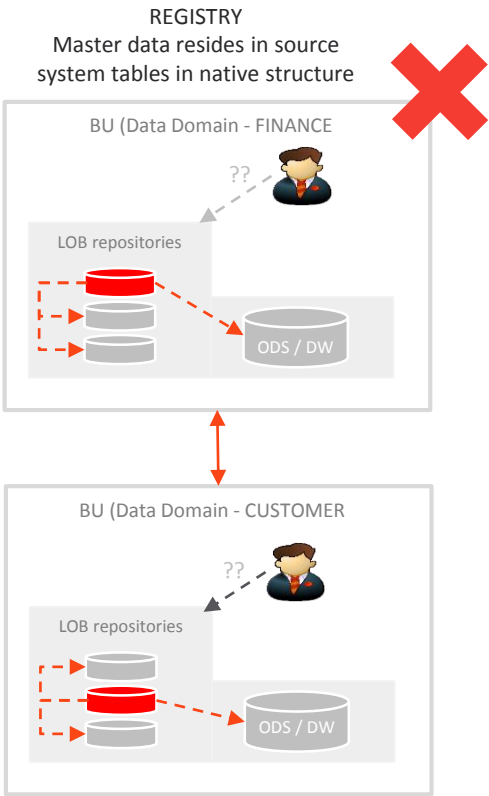# Master Management Architecture Considerations

- Recommended approach to Master Data Management (MDM) is key business data entities are standardized and held in a common repository, used across multiple business processes
- With this approach, there are opportunities to be flexible in both scope, timescale and deployment, thereby ensuring business demands are constantly reviewed and met within given timescales.
- Strong Data Governance structure and processes are pre-requisite to ensure data stewardship & trust
- Hybrid model approach below is recommended

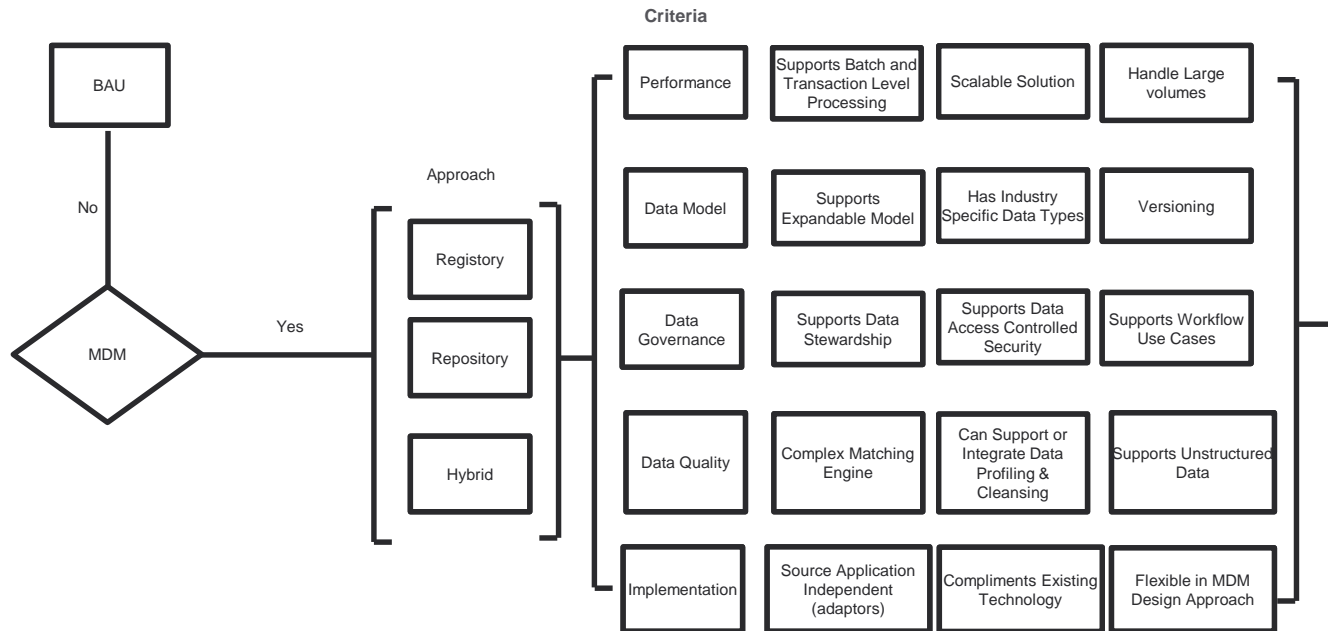| Approach	Description	Used	Pros	Con's
Registry Model	Source system IDs, foreign Keys and the key data values (needed for matching) are stored	Where key application owners cannot agree on a standard model	relatively quick to implement compared to the other approaches	• approach can lead to latency issues (where complex queries are required) to find the right entities and attributes across multiple applications • it does not improve the data quality as no data is updated on the source system until it is required.
Repository Model	Complete collection of master data is stored in a single repository / database	Each application is configured to use the master data repository for the key data entities	Ensures consistent, quality of data is maintained through out Prebuilt mater data models provided by enterprise vendors	• A corporate model / approach needs to agreed • All applications are modified to use master data management repository • Takes time to implement
Hybrid Model	Combination of above approaches		Supports a phased approached to MDM – start small	• Moving to repository may never happen (unless strong governance) • Latency and quality issues may persist

MDM Architecture Comparison

Master Data Management

DECISION CRITERIA

Criteria

Product options

```
BAU
 │
No│
 │
 ▼
◇ MDM ◇ ──Yes──► Approach
```

Approach:
- Registory
- Repository
- Hybrid

Performance	Supports Batch and Transaction Level Processing	Scalable Solution	Handle Large volumes
Data Model	Supports Expandable Model	Has Industry Specific Data Types	Versioning
Data Governance	Supports Data Stewardship	Supports Data Access Controlled Security	Supports Workflow Use Cases
Data Quality	Complex Matching Engine	Can Support or Integrate Data Profiling & Cleansing	Supports Unstructured Data
Implementation	Source Application Independent (adaptors)	Compliments Existing Technology	Flexible in MDM Design Approach

IBM
Infosphere MDM

- Integrates with exiting IBM footprint in UK Tx and US
- Data Stage, Data Quality, Information Analyser, Metadata Workbench

SAP MDG

- Integration with existing SAP footprint in UK and US
- SAP ERP holds majority of shared services master data

Informatica MDM

- Best of Breed MDM tool

nationalgrid | informationtechnology

Master Data Management

| Service	Principles	Rationale	Products	Guidance
Master Data Management	Data Management Principles EA principles Data Technology Services Architecture Principles	• national grid has a large number of application systems with overlapping data entities and attributes. • With limited data governance in place there are a number of issues in respect to consistency of data which leads to duplication, outdated, incomplete data and generally means poor data quality. • This has reduced confidence levels within the business in trusting the data they are receiving	SAP MDG IBM Infosphere Talend Informatica MDM	• MDM architecture approach must be considered before product • Registry based solution is relatively cost effective and time efficient to a full blown MDM Repository solution, but does not achieve all outcomes • Recommendation is Hybrid Model (cross between the two approaches and the approach). Hybrid model supports a phased approached to MDM, whereby we can start small (registry model) and build to a repository model • The IBM and SAP MDM offerings can be considered as both tools have an existing footprint (directly or indirectly) within the landscape and are considered industry leading. • An alternative option is to consider a SaaS based MDM solution vendor such as Talend. However care must be taken in considering overall security and integration elements, depending on which approach is the target end state.

Discussion / Design decisions

- MDM architecture option approach

  - Registry

  - Repository (multi / central hub)

  - Hybrid – combination of registry and repository options

- MDM product options

  - Domain based (business or data domain) - discreet products

  - Common enterprise product / standard

Appendix

Enterprise Data Model
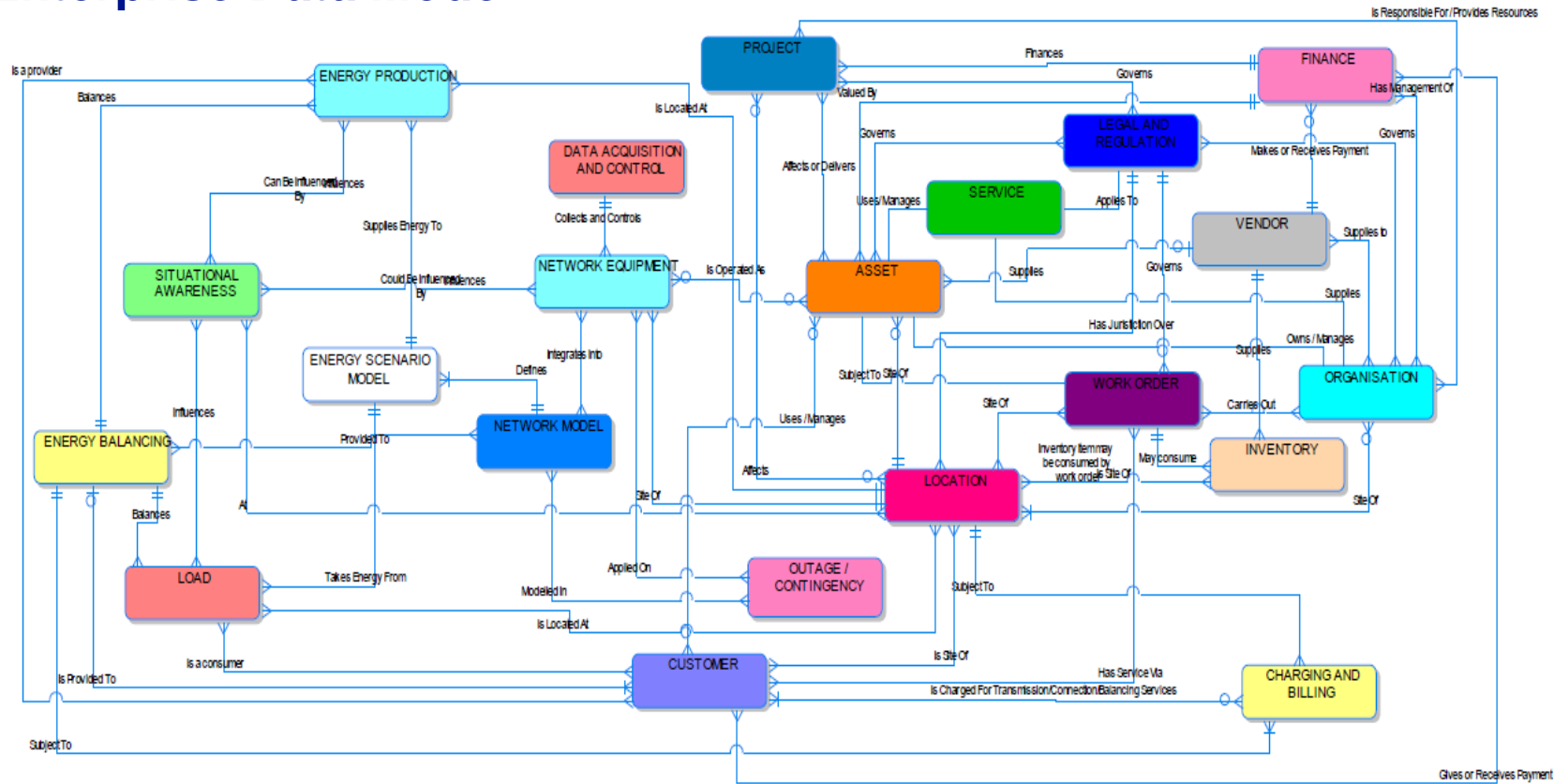
Data Technology Services Reference Architecture
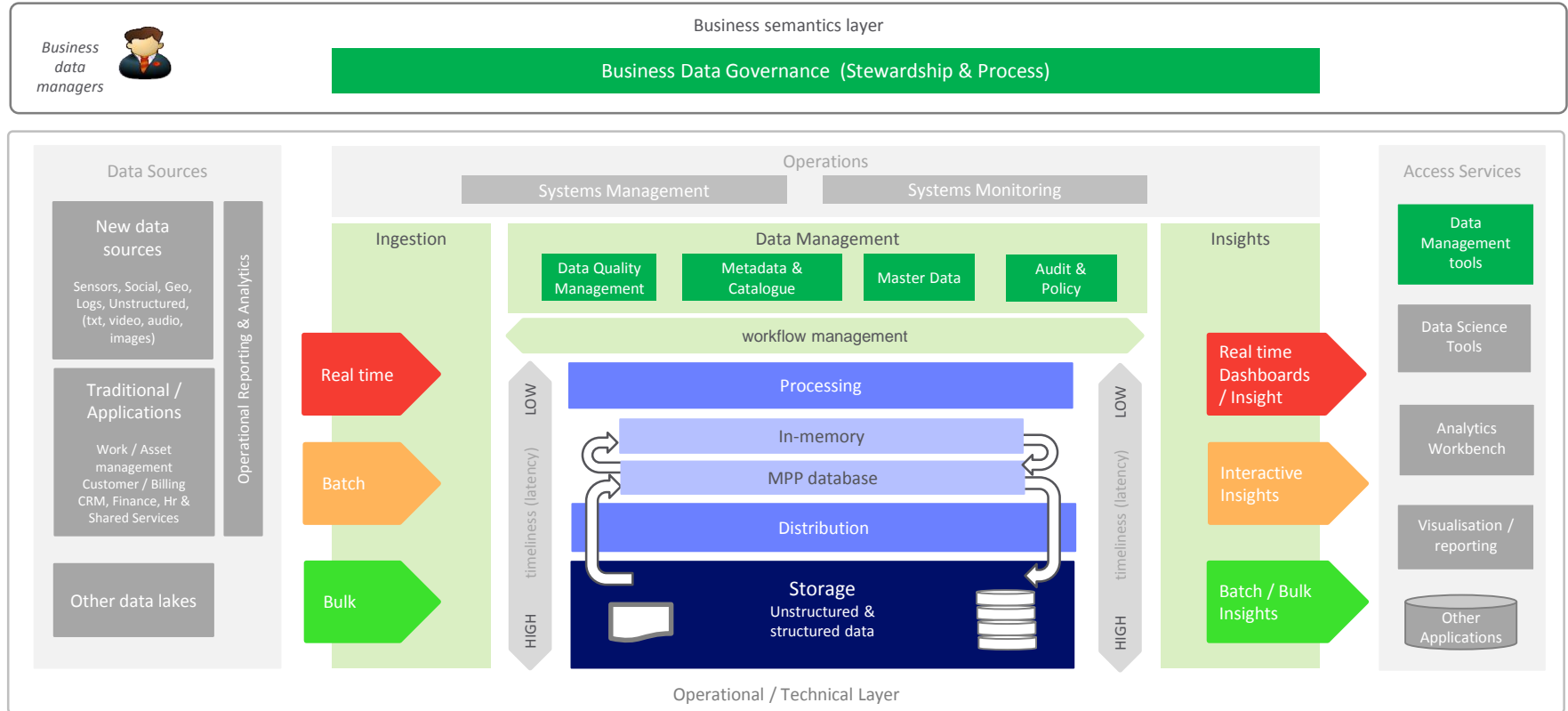
Architecture Positioning

Data Management Services

- Execution layer of data management services
- Improving data quality (monitoring, profiling and correction functions)
- Managing master (& reference) data
- Maintaining and making available data catalogue for all data assets in the 'data lake'

Data Sources

- Support for all types of data sources – structured and un structured, cloud, on premise sources etc.

Data Ingestion

- Capabilities applied within data ingestion processes and architectures supporting all levels of velocity and volume and data types
- Mechanisms to assure the quality of the data being integrated and delivered

Data Services (SOI)

- Service-oriented integration addresses problems with integrating legacy and inflexible heterogeneous systems by enabling functionality locked in existing applications as reusable services (API, Web)

Data Storage and Processing Services

- Scalable data storage, data processing and workflow management services
- File system, In memory and MPP storage services supported
- Enables storage of data of any size, shape and speed for

Data Governance Service

- Processes and tools for supporting the goals of a data governance initiative, and its associated roles and stakeholders (data stewards, etc.)
- Common vocabulary and semantics - establishes meaning, priority, and lineage of data, linkage to operational / technical metadata

Access Services

- Tools to support wide variety of use cases including, reporting, data visualisation and dashboards, analytics and advanced analytics and data science tools
- Other applications and data consumers

Data Preparation Services

- Pre-processing services often used to merge different data sources with different structures and different levels of data quality into a clean, consistent format
- Enables self service and onward processing

nationalgrid | informationtechnology
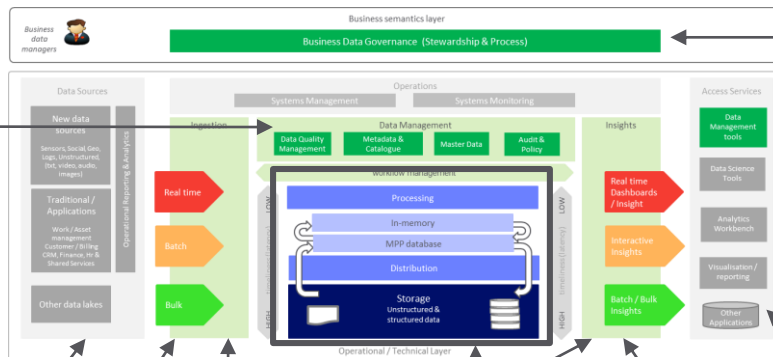
# Master Data Management Architecture Pattern
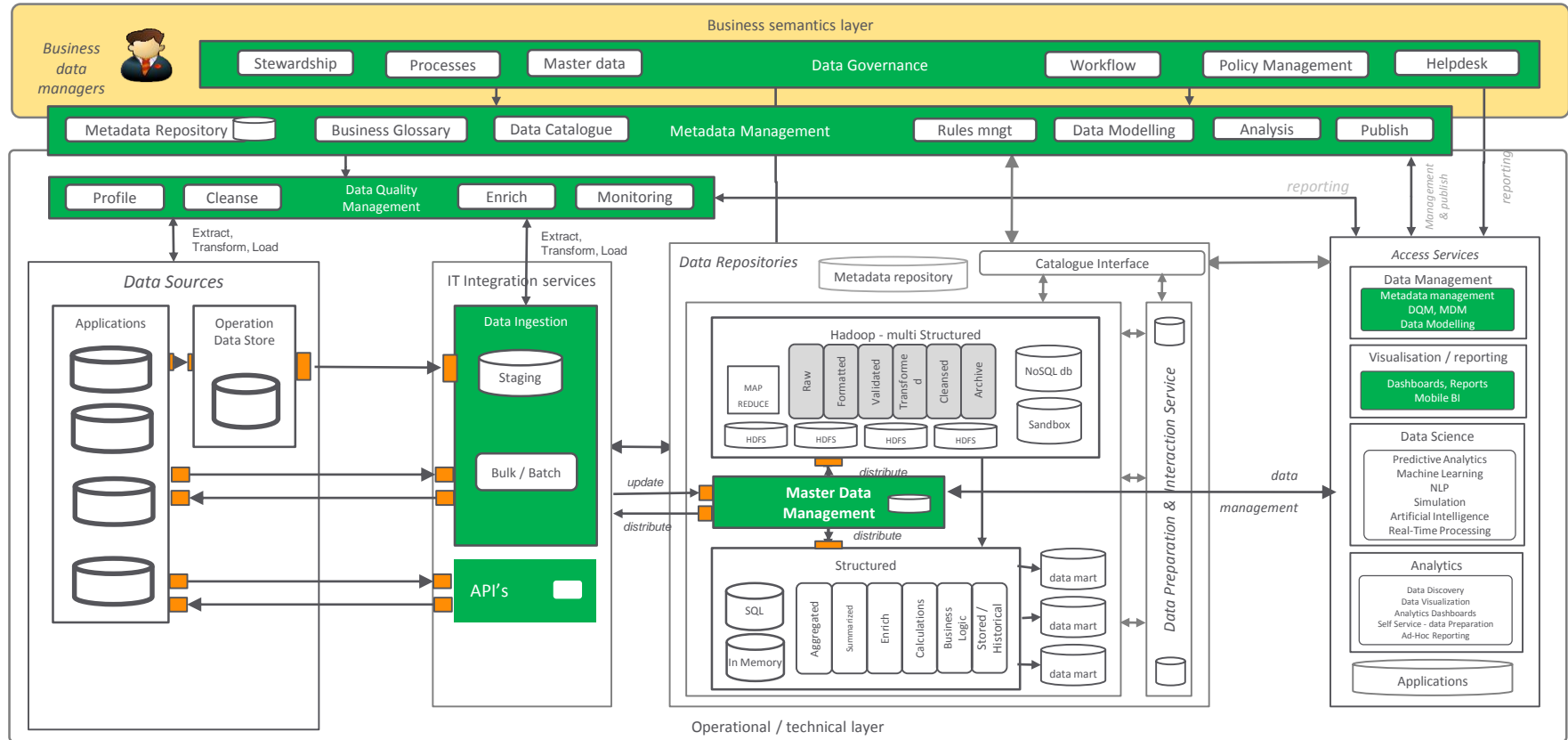
Principle – Importance of Semantic Alignment

Business Semantics

■ Business semantics are (business) metadata that describes the data & information that exists within the organization.

Semantic alignment

■ The ability to efficiently exchange data between the business and technical layers

■ To correctly interpret the meaning of the data that has been exchanged in the right context and within a reasonable time
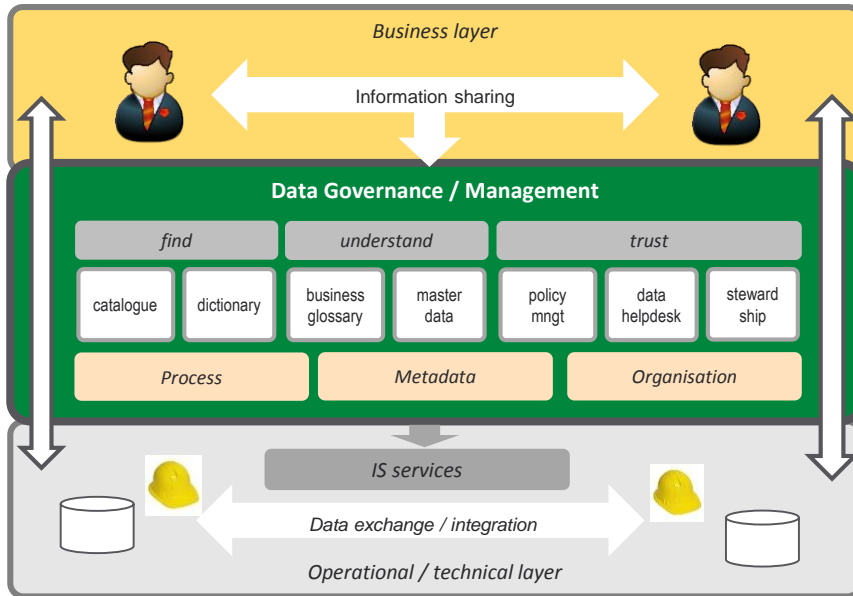
Enables

■ Alignment of business and IT
  ■ explicating the meaning, usage and whereabouts of all data assets.

■ Knowledge sharing
  ■ Between business stakeholders through more accurate information delivery
  ■ Enables better search, navigation, discovery, content management, web-sites, and many other knowledge-intensive applications

■ Data exchange
  ■ Between disparate systems
  ■ Automatic transformation between data formats.
  ■ Provides a better foundation for many other initiatives such as master data management, business intelligence, SOA, BPM, etc.