Modern Data Warehousing + Business Intelligence

Great strides have been made in the world of big data technologies over the past decade. **Cloud infrastructure**, along with **new storage paradigms** like **Hadoop**, **graph databases** and **message buses**, have provided a new dimension to the arsenal of tools available to tackle the deluge of data the world is facing.

A **"modern" data warehouse architecture** that combines **SQL**, as well as **NO-SQL technologies**, is a preferred way to **store** and **enrich** enterprise data where "small data" is seen as the latest fad, placing the focus back on **quality** rather than quantity.

Our view is that a well-structured data warehouse, with **SQL at its core**, still provides a level of **technological maturity** that is lacking in these new-world tech solutions that remain relevant to financial service providers.

Data Warehouse High-Level Architecture

1. SOURCE SYSTEMS: Transactional systems used for day-to-day business operations.

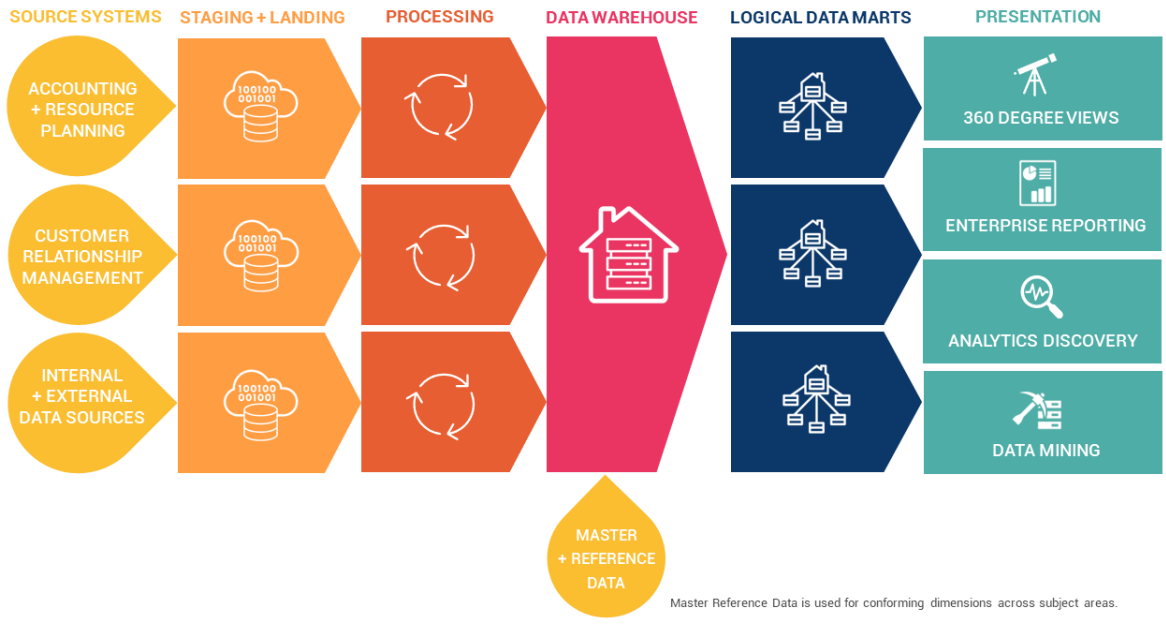
2. STAGING + LANDING: Data is received from the source systems, either by extracting data or by having the source systems push the data into a landing area.

3. PROCESSING: Applying business logic and rules in one central location.

4. DATA WAREHOUSE: Location where the processed data is published and conformed to enterprise entities.

5. LOGICAL DATA MARTS: Logical or virtual subject-orientated data markets (i.e. sales, marketing, finance, etc.) that feed off the data warehouse.

6. PRESENTATION: Presentation layer where the business users will consume the information.



The following **high-level stages** are generally prescribed when implementing a data warehouse that will provide a **return on investment** to the business:

STAGES	STEP	HIGH-LEVEL TASKS
1. ANALYSIS & DESIGN	**1. Understand reference and master data management challenges**	1. Hold brainstorming workshops with business stakeholders 2. Identify and analyze business unit's key business entities (i.e., customers, accounts, products, suppliers, etc.)
	2. Define and maintain the data architecture	1. Diagram out the current state of data landscape 2. Define a to-be integration architecture 3. Update and maintain the existing organisation data architecture (as required)
	3. Establish the 'golden record' definitions	Assess 'golden records' field requirements and meta data for each
	4. Identify matching rules between entities	Assess differing system reference formats that will be used for matching purposes
	5. Develop to-be data model	Assess any required hierarchies and relationships between entities
2. DEVELOP	**1. Establish infrastructure**	1. Provision server infrastructure

	2. Procure, install and license software. (Look to leverage existing software licences) 3. Implement data connectors to internal and external sources
2. Source and consolidate 'golden records'	1. Stage and load 'golden records' from disparate source system 2. Reconcile mismatched entities using data quality tools and through interactions with business users
3. Replicate and distribute reference and master data into a data repository	1. Implement updates to operational systems through data connections or API interfaces (as required) 2. Publish master data and reference data to analytical systems 3. Testing for both development and user acceptance testing

3. TRANSITION	**Manage changes to reference and master data**	1. Hand over the operation of the master data management system to business users 2. Provide for end-user training

Analytical Insight Benefits

Many **analytical tools** and **insights** are enabled as the data warehouse begins being populated with data that then emerges as a **reliable source of information**. From an exhaustive list of these tools, below are the ones we consider **most relevant**.

ANALYTICAL INSIGHTS		HIGH LEVEL DESCRIPTION	TYPICAL ATTRIBUTES
360° VIEWS	RELATIONAL	These types of reports address a subject area in the enterprise (i.e., the customer, a supplier, asset, product, etc.), typically the customer and their behavior.	1. Reports at the level of the subject 2. Many attributes

		These types of reports are useful for downstream advanced analytical discovery using advanced visualization, statistical and or machine learning.	(sometimes 100s of columns)
	GRAPH	True 360-degree view of a subject that allows for lateral exploration of connected attributes. Due to the unstructured nature of this type of view, entity information is stored in graph format requiring a semantic database engine. Allows transversal pattern discovery, specifically useful for forensics and root-cause analysis.	1. Requires a semantic graph database 2. Delivers on the promise of a single view
ENTERPRISE REPORTING	OPERATIONAL	Reports that address the real-time operational needs of the enterprise. This includes human and systems consumers. These reports generally retrieve information from an operational data store (ODS) or directly from the operational system	1. First order reporting 2. Build for purpose
	STRATEGIC / TACTICAL	Reports that address the higher-level requirements of the enterprise. These are tactical and strategic in nature and typically report off aggregate level data. Typically, the sources of these reports are derived from data warehouses having cleansed and conformed data. These reports generally source data from OLAP cubes.	1. Single-Version-of-the-Truth reporting 3. Typically, OLAP cube based 4. Drilldown capabilities
	OLAP	An analytical tool used for slice-and-dice operations typically deriving information from OLAP cubes. Has the ability to process aggregate level calculations through MDX calculations. Movement towards in-memory processing which offers speed advantages.	1. Power Users reporting 2. Allows Drill-through

ANALYTICAL DISCOVERY	**VISUALIZATION**	Visualization platforms that dynamically and visually represent a subject area. The goal of these types of tools is to expose high dimensionality to the user through visual means as the optical system of a human user has the highest interpretation throughput capability. High dimensionality refers to the ability to represent data series, category information, values on multiple axes, time, animation and data point attribute representation (i.e. colors, point size and shapes).	1. Discovery tools 2. Requires agile provisioning of data to be effective
DATA MINING	**PREDICTIVE MODELLING**	Business users can use of software to describe and predict outcomes by making use of statistical techniques.	1. Statistics 2. Machine Learning

Proposal / Call to action

Ilion is looking to help financial services organisations **optimise their return of investment** in **data**, centred around financial products and services. Our executive team have extensive experience in financial analysis for banks, credit risk and stress testing, and data management.

We see an overlap in these areas that provide a sweet spot for a return on investment made.

When you are working on your next financial analytical undertaking, please don't hesitate to call us. We would be happy to help you find a solution that is optimal for your organisation.

Peter Gross
Director: Data Management
Email: peter@ilion.co.za
Mobile: +27 (0)83 388 1674