USING DATA AUTOMATION TO MIGRATE YOUR DATA WAREHOUSE TO THE CLOUD

quest for knowledge

q4k.com

FRANK MARTENS

Leads the data automation initiative within Quest for Knowledge

Over 15 years of experience, designing and implementing BI and analytics solutions with a strong sense of enthusiasm, combining traditional BI frameworks with modern data technologies

Data Warehousing expertise including Data Quality and Profiling and Data Governance

Business intelligence expertise including data visualization, dashboarding, reporting, self-service BI, and KPI’s

Significant experience in information analysis, designing data architectures, and deploying end-user applications

Worked in agile/scrum teams in various roles, including product owner and developer

Certified and seasoned trainer in SAP Data management & Analytical tools

AGENDA

- Data Warehouse Automation
- Traditional vs automated approach
- Why data automation?
- Use cases
- Data Warehouse Automation tasks
- Cloud migration - What to migrated?
- Migration options
- ETL migration
- Key Phases of a lift and shift migration
- Demo

DATA WAREHOUSE AUTOMATION

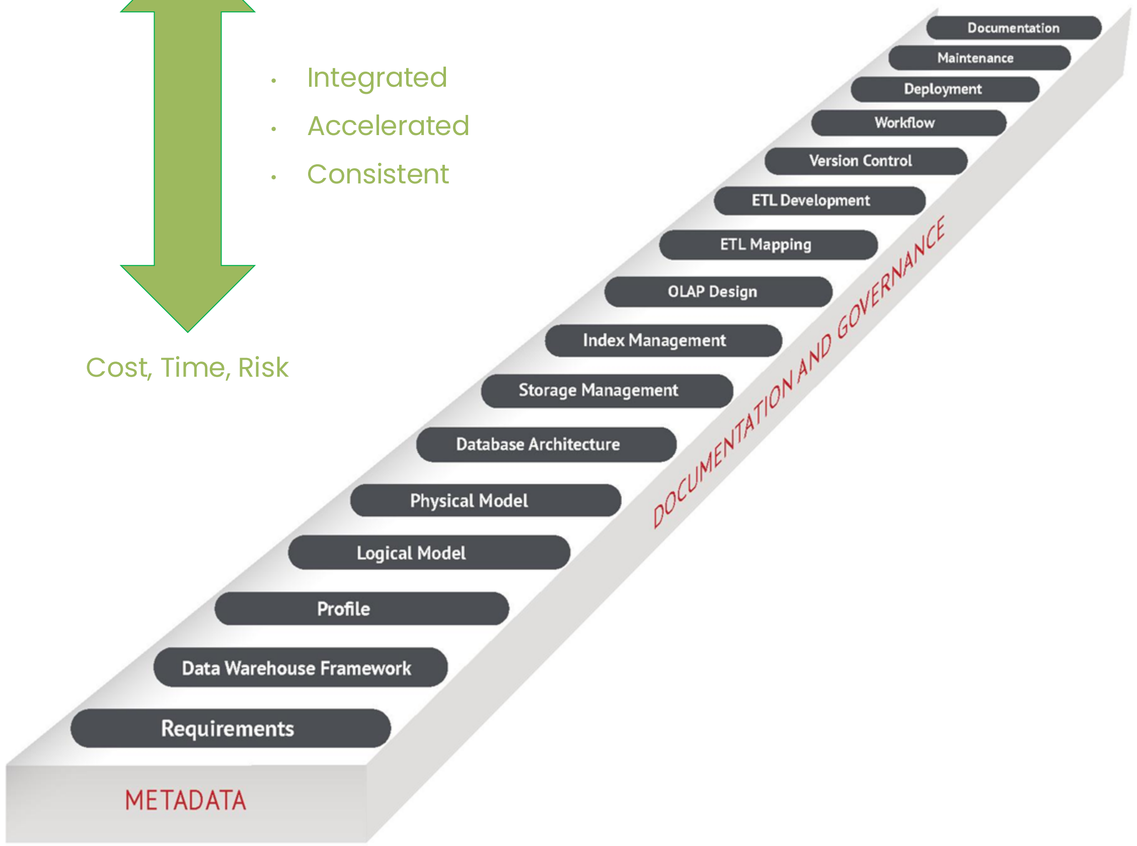
Data Warehouse Automation (DWA) refers to the use of software tools and technologies to **automate the design, development, deployment, and management** of data warehouses and related processes such as data integration (ETL/ELT), testing, documentation, and updates using **Meta Data** and **templates**

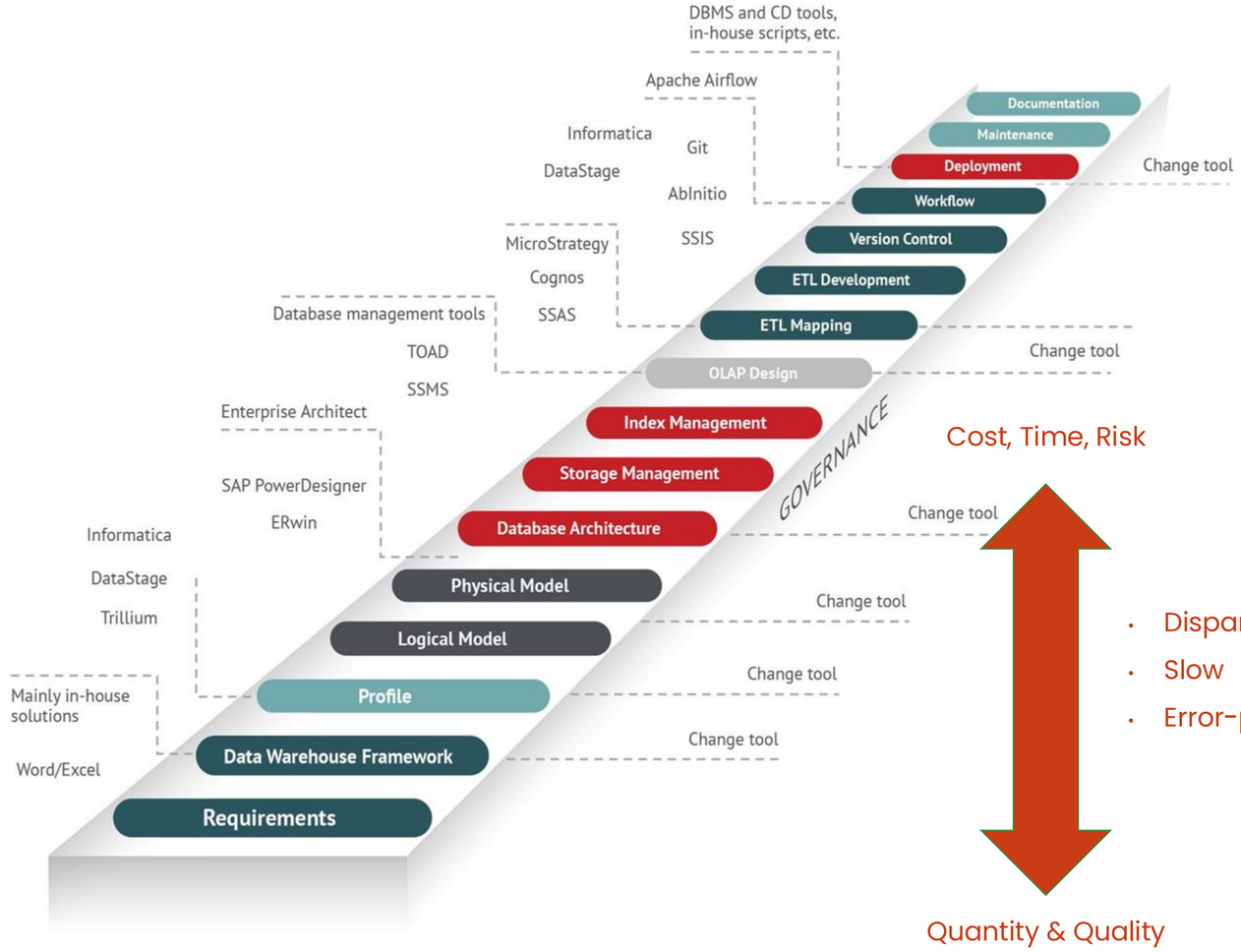
TRADITIONAL VS AUTOMATED APPROACH

Quality & Quantity

- Integrated
- Accelerated
- Consistent

Cost, Time, Risk





Cost, Time, Risk

- Disparated
- Slow
- Error-prone

Quantity & Quality

WHY USE DATA AUTOMATION?

Cost Reduction

Risk Mitigation

Complexity Management

Urgency

Agility

USE CASES

Data Warehousing

Data Warehouse (cloud) migration

Data Products

Big Data Integration

Data Vault

Data Marts

WHAT CAN YOU DO WITH DW AUTOMATION?

Generate different DW data model designs – Data Vault, E/R and Dimensional

ELT/ELT job generation

Semantic layer generation

Processing streaming data for near real-time DW

Integrate cloud data warehouses with cloud-based lakehouses / data lakes

Integration with data catalogues

WHAT EXACTLY MUST BE MIGRATED?

- DW and DataMart schema
- Data
  - Volume
  - Performance on transfer
  - Compliance (GDPR)
- ETL jobs
- Metadata
- Security (users, roles and priviledges)
- BI Front-end

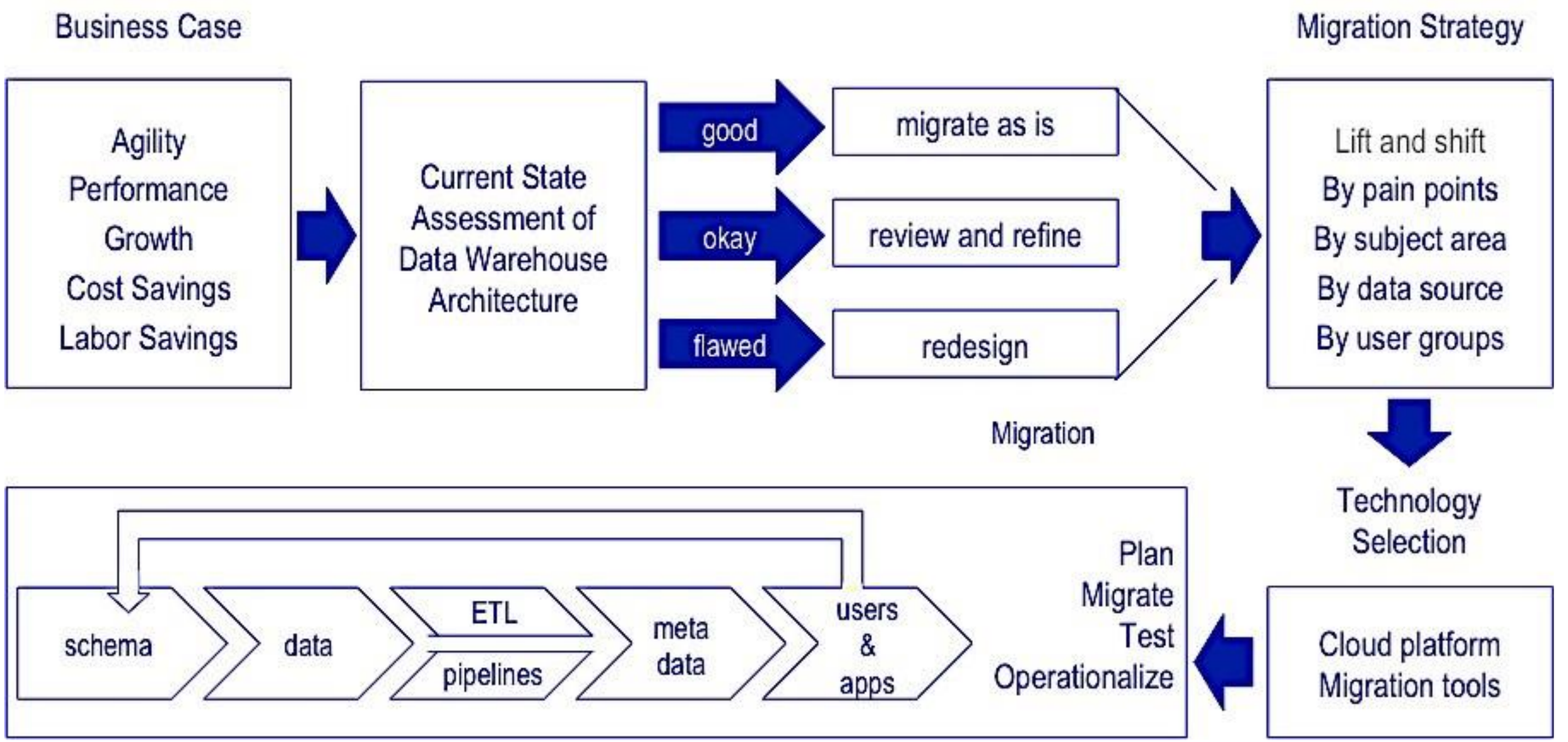
MIGRATION OPTIONS [1]

Lift and shift your existing data warehouse as-is

Simplify your existing data warehouse and then Lift and shift

Complete re-design and then migrate your data

MIGRATION OPTIONS [2]



Eckerson

ETL MIGRATION

Current state	Risks/Benefit	Migration Options
Custom 3GL code and scripts	• Code provides no metadata lineage • Hard to re-engineer if authors have gone	• Plan to re-develop these using Data Warehouse Automation of cloud data integration software
Stored procedures that run in your existing data warehouse DBMS	• Likely to be significant differences between existing DW & Target DW • No metadata lineage	• Plan to re-develop these using Data Warehouse Automation of cloud data integration software
Graphical ETL tool	• Avoids re-development • Minimizes risk and Quicker to migrate • Quick to migrate, easy to maintain	• Continue using your existing ETL tool and switch the target to cloud DW • Possibly move to a could version of your existing ETL tool and port the metadata to run ETL jobs on the cloud. • Plan to re-develop these using Data Warehouse Automation of cloud data integration software
Data Warehouse Automation software	• Avoids re-development • Minimizes risk and quicker to migrate	• Continue using your existing Data Warehouse automation tool switching the target and staging to cloud Data Warehouse

KEY PHASES OF A LIFT AND SHIFT MIGRATION[1]

1. Discovery & Design
    - Auto-reverse-engineer existing model (logical data model)
    - Generate a physical target data model (Star, Snowflake schema’s)
2. ETL/ELT Metadata Mapping & Transformation
    - Drag-and-drop mapping; autosuggest transformations
    - Parameterize Incremental loads vs full loads
3. Code Generation & Deployment
    - Data Model deployment
    - ETL One-Click generation of SQL/Script or orchestration workflows
    - Auto-deploy pipelines into you cloud environment

KEY PHASES OF A LIFT AND SHIFT MIGRATION[2]

4. Testing & Validation
    - Out-of-the-box data-reconciliation checks
    - Auto-gen data quality reports
    - Testing reports
5. Cut-Over & Continuous Synchronization
    - Switch from full to Incremental CDC-driven delta loads
    - Built-in scheduling and monitoring dashboards

DEMO

THANK YOU

quest for knowledge

q4k.com