COURSEWARE

DATA INTEGRATION AND INTEROPERABILITY

BASED ON CDMP

Van Haren PUBLISHING

MICHEL DEKKER

Data Integration and Interoperability Based on CDMP

Colophon

Title:	Data Integration and Interoperability
Subtitle:	Based on CDMP
Authors:	Michel Dekker
Publisher:	Van Haren Publishing, ‘s-Hertogenbosch
ISBN Hard Copy:	978 94 018 1384 6
Edition:	First edition, first print, October, 2025
Design:	Van Haren Publishing, ‘s-Hertogenbosch
Copyright:	© Van Haren Publishing 2025

For further information about Van Haren Publishing please e-mail us at: info@vanharen.net or visit our website: www.vanharen.net

Publisher about the Courseware

The Courseware was created by experts from the industry who served as the author(s) for this publication. The input for the material is based on existing publications and the experience and expertise of the author(s). The material has been revised by trainers who also have experience working with the material. Close attention was also paid to the key learning points to ensure what needs to be mastered.

The objective of the courseware is to provide maximum support to the trainer and to the student, during his or her training. The material has a modular structure and according to the author(s) has the highest success rate should the student opt for examination. The Courseware is also accredited for this reason, wherever applicable.

In order to satisfy the requirements for accreditation the material must meet certain quality standards. The structure, the use of certain terms, diagrams and references are all part of this accreditation. Additionally, the material must be made available to each student in order to obtain full accreditation. To optimally support the trainer and the participant of the training assignments, practice exams and results are provided with the material.

Direct reference to advised literature is also regularly covered in the sheets so that students can find additional information concerning a particular topic. The decision to leave out notes pages from the Courseware was to encourage students to take notes throughout the material. Although the courseware is complete, the possibility that the trainer deviates from the structure of the sheets or chooses to not refer to all the sheets or commands does exist. The student always has the possibility to cover these topics and go through them on their own time. It is recommended to follow the structure of the courseware and publications for maximum exam preparation.

The courseware and the recommended literature are the perfect combination to learn and understand the theory.

-- Van Haren Publishing

Other publications BY Van Haren Publishing

Van Haren Publishing (VHP) specializes in titles on Best Practices, methods, frameworks and standards within four domains:
- IT Management
- Architecture (Enterprise and IT)
- Business Management and
- Project Management

Van Haren Publishing is publishing on behalf of leading organizations and companies: Agile Consortium, World Commerce and Contracting, IAOP, IPMA World, KNVI, PMI-NL, NLAIC and The Open Group.

Van Haren Publishing is part of the Van Haren Group and additional to the book publishing also provides the following services: accredited training materials and e-learning through Van Haren Learning Solutions, as well as independent professional certification via examination through Van Haren Certify.

Topics are (per domain):

IT Management	IT Service Management	FitSM, ISM®, ISO/IEC20000, IT4IT®, ITIL®, VerISM®, SAF, TRIM, XLA®
	Data Management	Data literacy, Data visualization, DMBOK
	IT Asset Management	HAM, ITAM, SAM
	IT Security Management	BIO, ISO/IEC27001, NIS2
	Test Management	CTAP
	Application Management	ASL
	Other	eCF, IT-CMF, Scrum
Project Management	Project Management	Half Double, ICB, ISO/IEC21500, P3.express, PM2, PMBOK Guide, Praxis, PRINCE2
	Agile	Agile, Agile PM
	Other	PMO
Business Management	Operations Management	Lean, Lean Six Sigma, OBM, OMC, RASCI
	Contract Management	CATS CM, CATS RVM, IACCM World
	Business Information Management	BiSL, DID
	Artificial Intelligence	AI, Generative AI
	Outsourcing	OPBOK
Enterprise Architecture	Enterprise Architecture	BIAN, TOGAF
	Modeling	ArchiMate, BPMN
	Software Architecture	ISAQB
	Other	Open Agile Architecture

For the latest information on VHP publications, visit our website: www.vanharen.net.

Table of Content

Self-Reflection of understanding Diagram

'What you do not measure, you cannot control.'' – Tom Peters

Fill in this diagram to self-evaluate your understanding of the material. This is an evaluation of how well you know the material and how well you understand it. In order to pass the exam successfully you should be aiming to reach the higher end of Level 3. If you really want to become a pro, then you should be aiming for Level 4. Your overall level of understanding will naturally follow the learning curve. So, it's important to keep track of where you are at each point of the training and address any areas of difficulty.

Based on where you are within the Self-Reflection of Understanding diagram you can evaluate the progress of your own training.

Level of Understanding	Before Training (Pre-knowledge)	Training Part 1 (1st Half)	raining Part 2 (2nd Half)	After studying / reading the book	After exercises and the Practice exam
Level 4 I can explain the content and apply it .					
Level 3 I get it! I am right where I am supposed to be.					Ready for the exam!
Level 2 I almost have it but could use more practice.					
Level 1 I am learning but don't quite get it yet.					

(Self-Reflection of Understanding Diagram)

Write down the problem areas that you are still having difficulty with so that you can consolidate them yourself, or with your trainer. After you have had a look at these, then you should evaluate to see if you now have a better understanding of where you actually are on the learning curve.

Troubleshooting

Problem areas:	*Topic:*
Part 1	
Part 2	
You have gone through the book and studied.	
You have answered the questions and done the practice exam.	

Data Integration and Interoperability

COURSEWARE

Effective Data Foundation

Not-for-profit collective, which enables **professionals** to **leverage data** to make **sustainable business decisions**

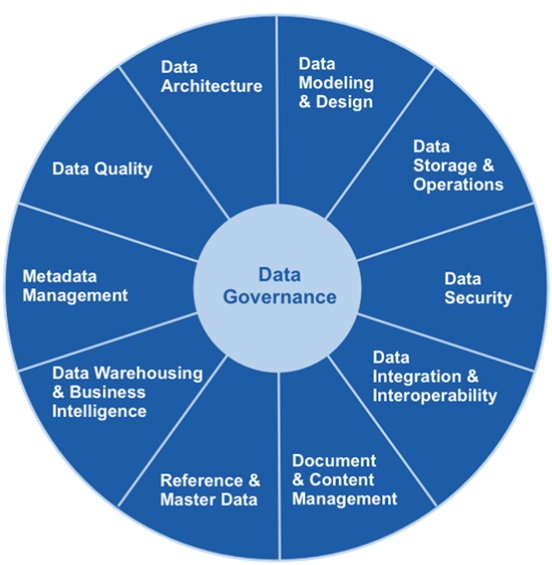
 2

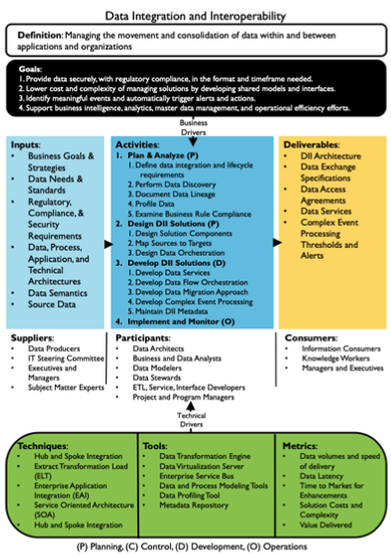
Effective Data Foundation

DATA

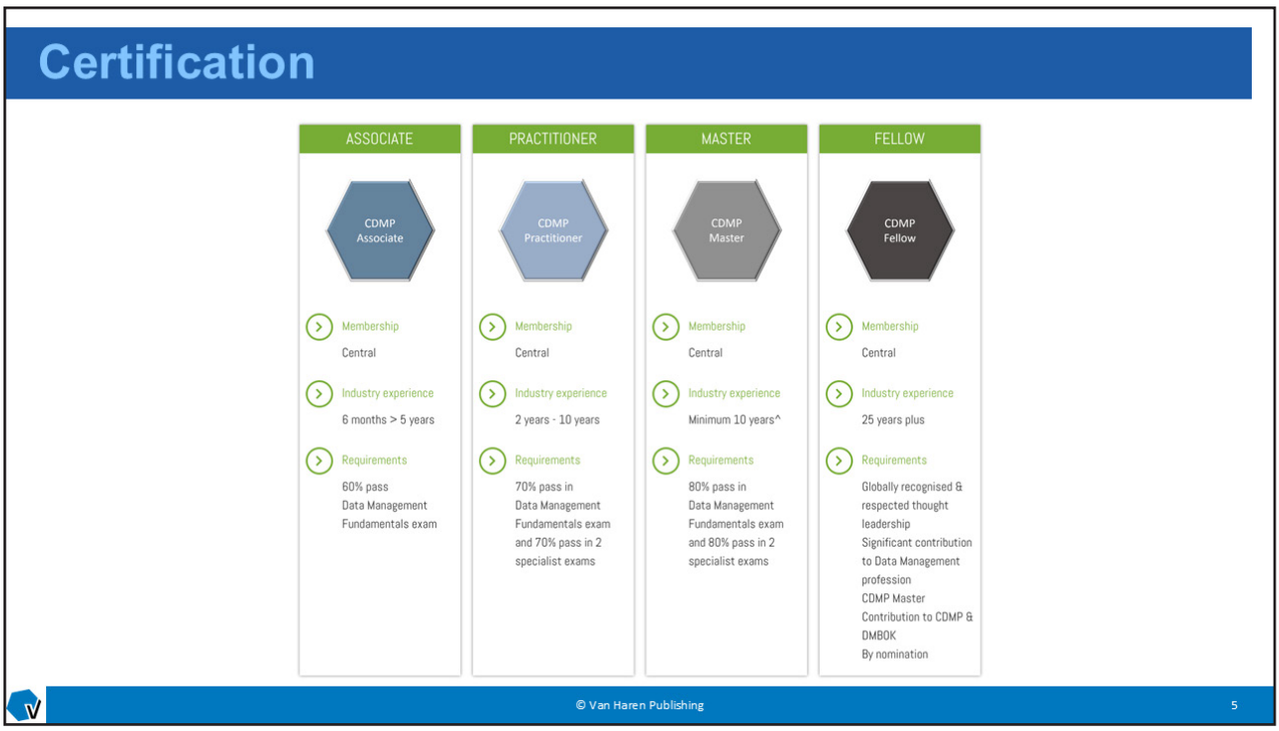
Analysis / Literacy / Management / Visualization

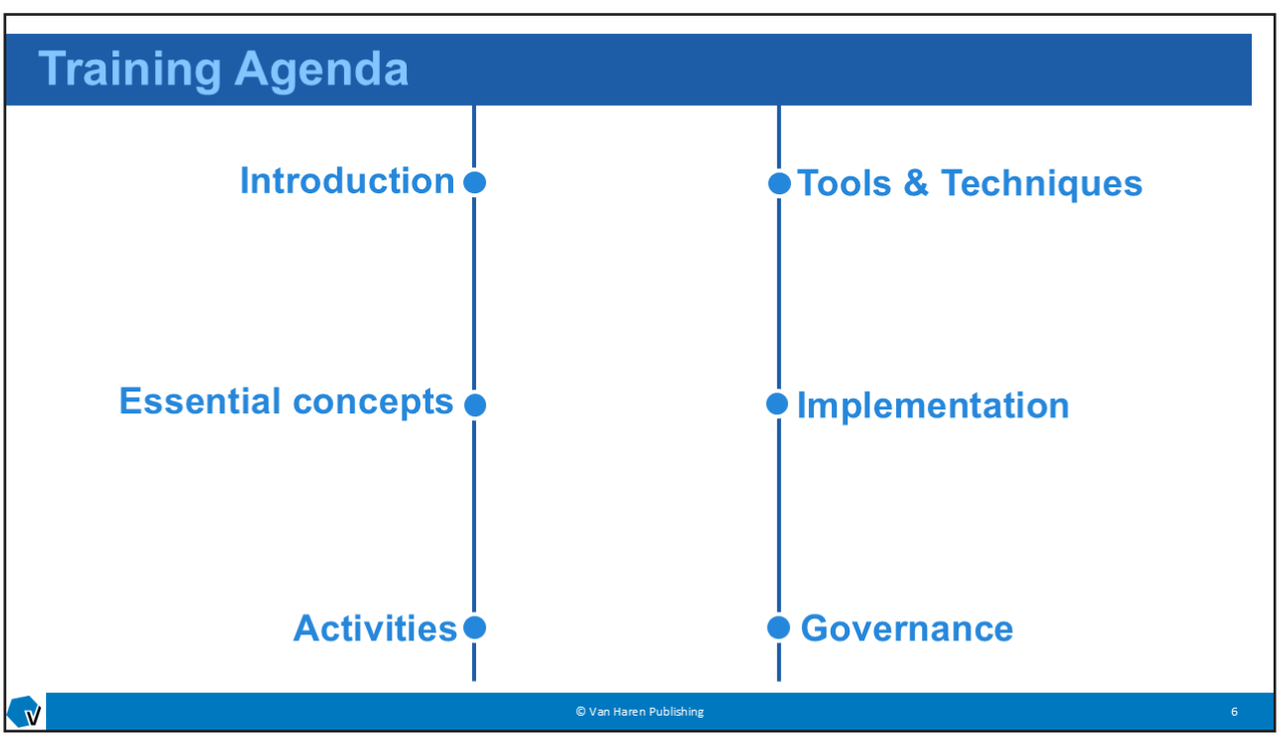
DAMA Wheel





Source: Dama.org


Certification
ASSOCIATE
CDMP Associate
Membership
Central
Industry experience
6 months > 5 years
Requirements
60% pass Data Management Fundamentals exam
PRACTITIONER
CDMP Practitioner
Membership
Central
Industry experience
2 years - 10 years
Requirements
70% pass in Data Management Fundamentals exam and 70% pass in 2 specialist exams
MASTER
CDMP Master
Membership
Central
Industry experience
Minimum 10 years^
Requirements
80% pass in Data Management Fundamentals exam and 80% pass in 2 specialist exams
FELLOW
CDMP Fellow
Membership
Central
Industry experience
25 years plus
Requirements
Globally recognised & respected thought leadership
Significant contribution to Data Management profession
CDMP Master
Contribution to CDMP & DMBOK
By nomination
© Van Haren Publishing
5


Training Agenda
Introduction
Essential concepts
Activities
Tools & Techniques
Implementation
Governance
© Van Haren Publishing
6

Introduction

Data integration is the combination of technical and business processes used to combine data from ***disparate sources into meaningful and valuable information***

William McKnight

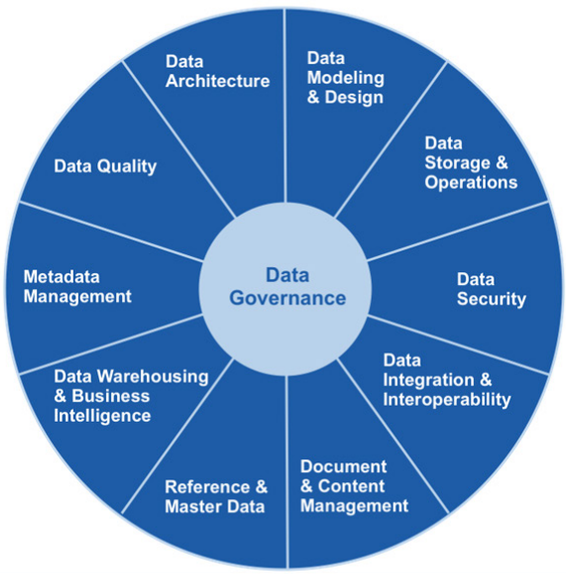
Definition

Managing the **movement** and **consolidation** of data **within** and **between** applications and organizations

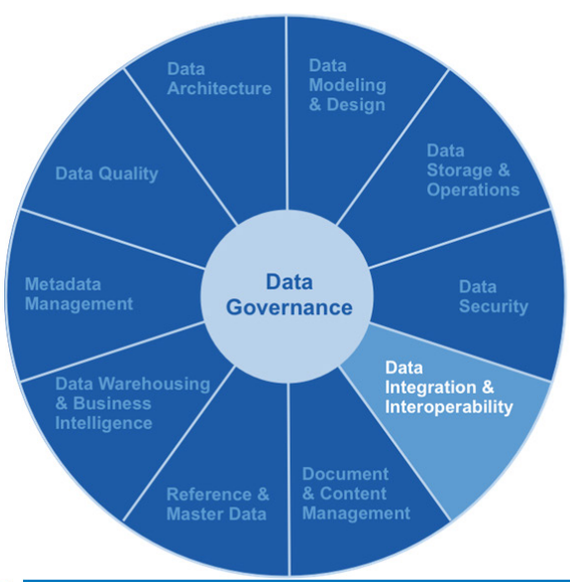
Data Integration
consolidates data into consistent forms, either physical or virtual

Data Interoperability
ability for multiple systems to communicate

Relations with other knowledge areas



Relation with Data Governance

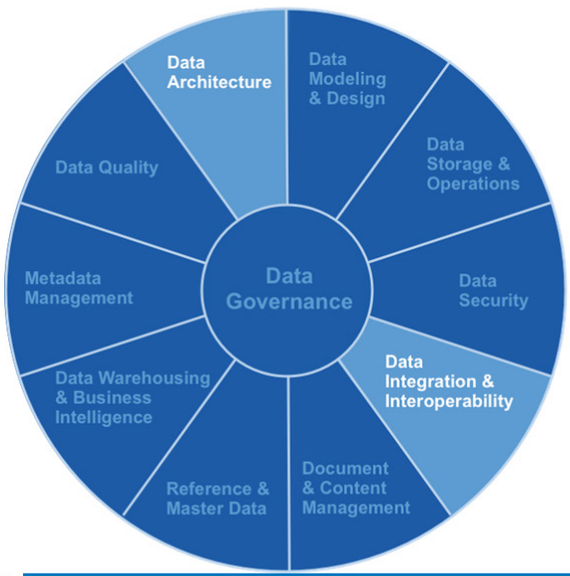


Alignment

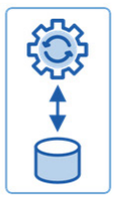
Quality & Consistency

Collaboration

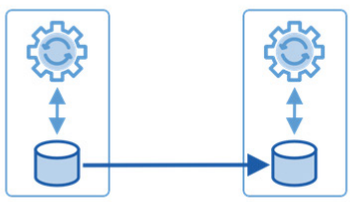
Relation with Data Architecture



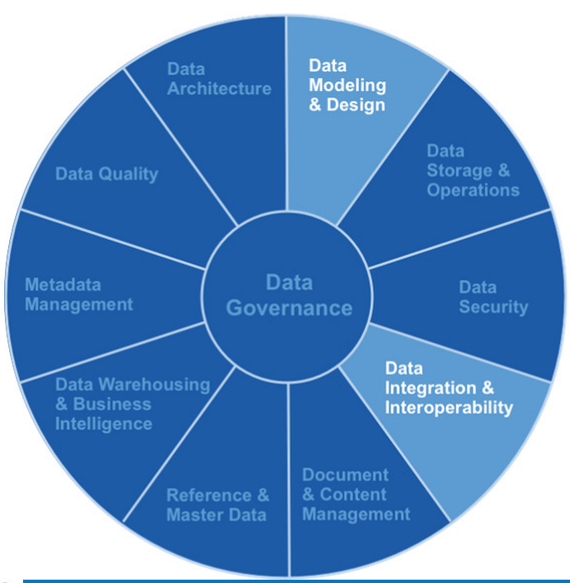
Data at rest



Data in motion



Relation with Data Modeling & Design

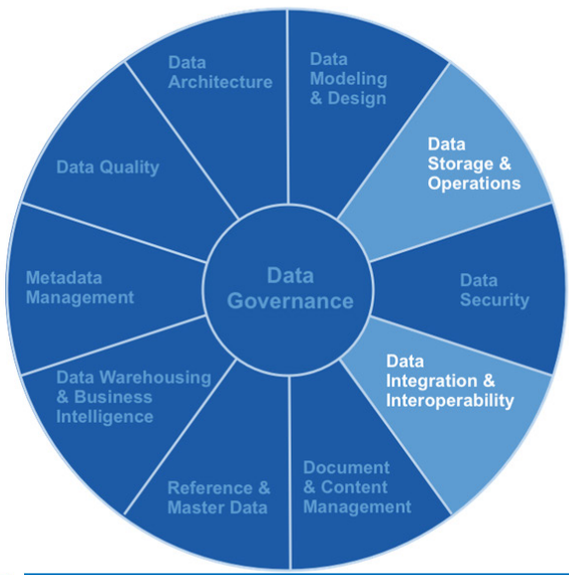


Structure and semantics

Relationships and business context

Consistent, accurate, and meaningful

Relation with Data Storage & Operations



Critical enabler of efficient

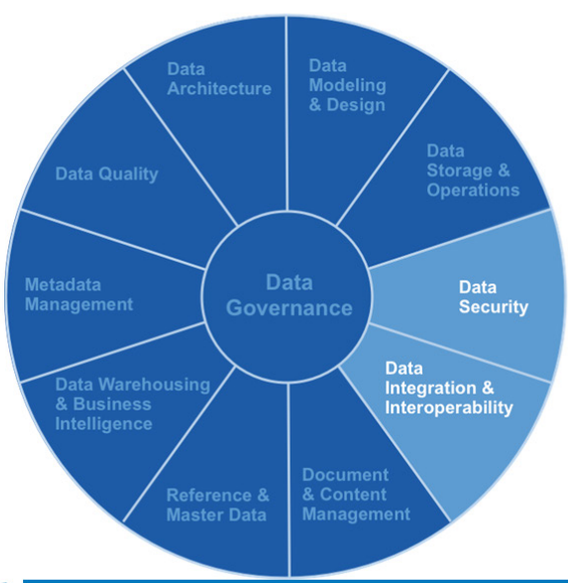
data storage

data management

data operation

 13

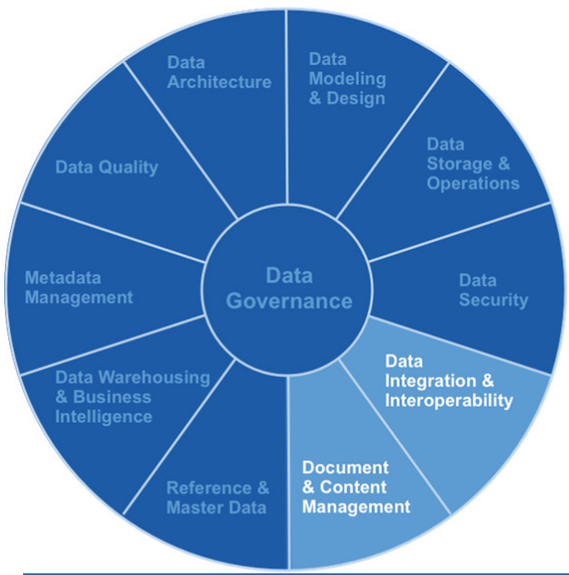
Relation with Data Security



Ensure that the integrated data remains **secure and protected** against potential threats and vulnerabilities.

 14

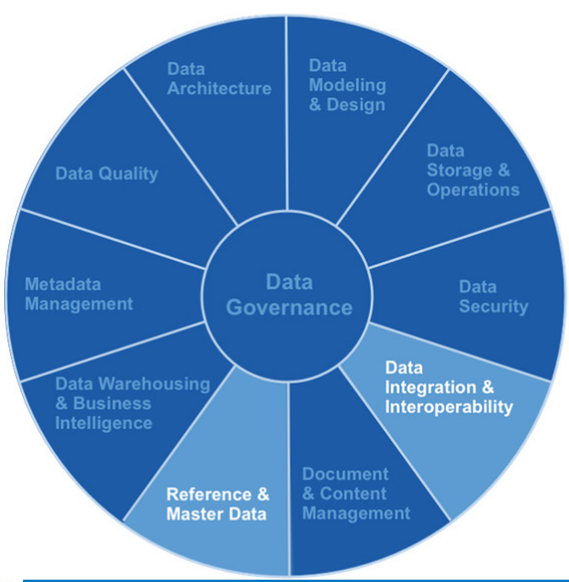
Relation with Document & Content Management



integrate **structured** and **unstructured** data

Lifecycle management

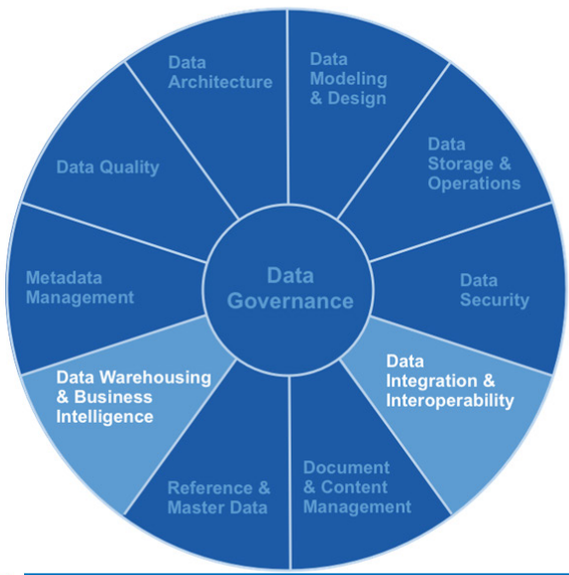
Relation with Reference & Master Data



Without DII any form of Reference or Master Data is **challenging**

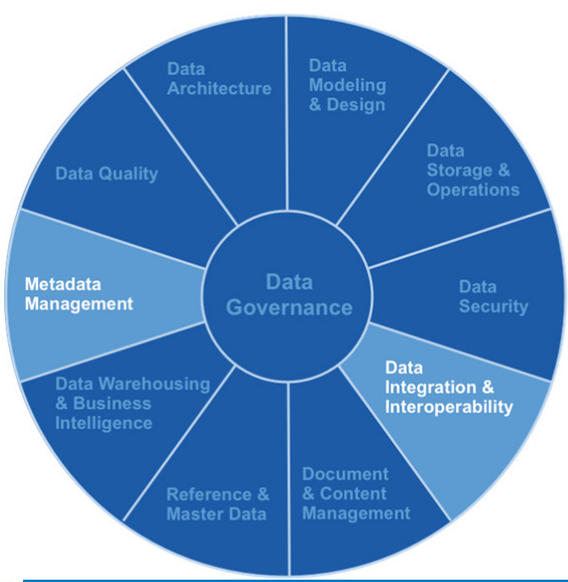
Robust Reference & Master Data **facilitates** data integration efforts

Relation with Data Warehousing & Business Intelligence



DW&BI is all about **collecting** data from different sources and **integration** of this data

Relation with Metadata Management



Metadata Management provides the necessary **context and structure**, which is fundamental for successful data integration and interoperability initiatives.

Relation with Data Quality

data quality management is crucial for ensuring that the integrated data is **trustworthy** and **fit for** its intended **use**.

Business Drivers

Manage data movement **efficiently**

Manage the **cost** of support of complex solutions

Comply with standards and regulations

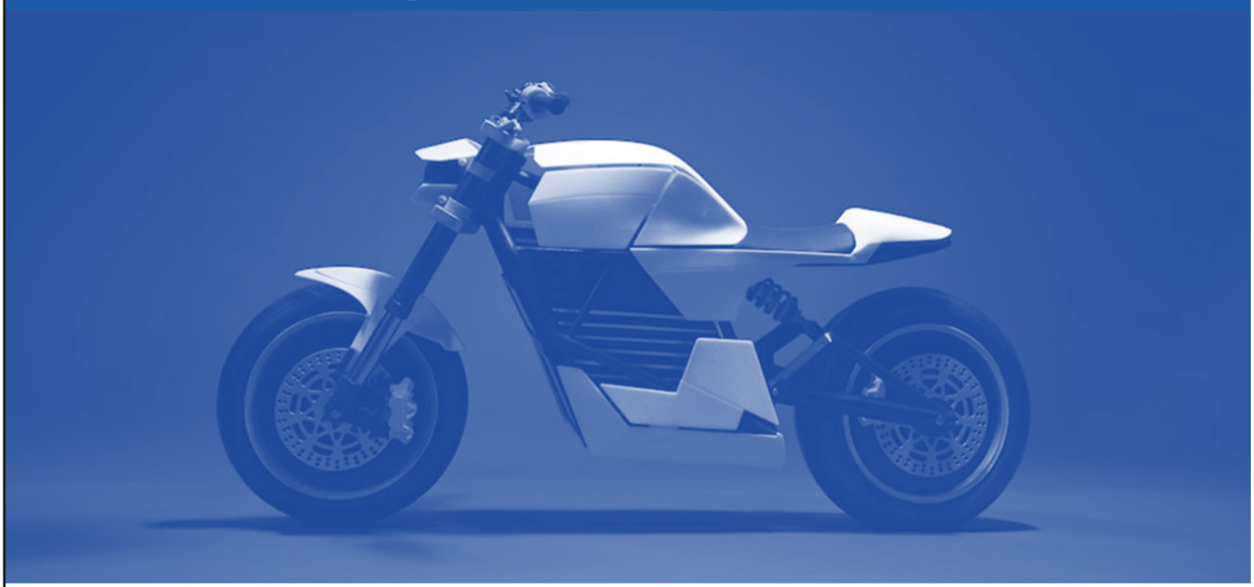
Goals & Principles

- Make data available in the format and timeframe **needed** by data **consumers**
- **Consolidate** data physically and virtually into data hubs
- Lower cost and complexity by developing **shared models and interfaces**
- Identify **meaningful events** and **automatically trigger** alerts and actions
- Support Business Intelligence, **analytics**, Master Data Management, and **operational efficiency** efforts

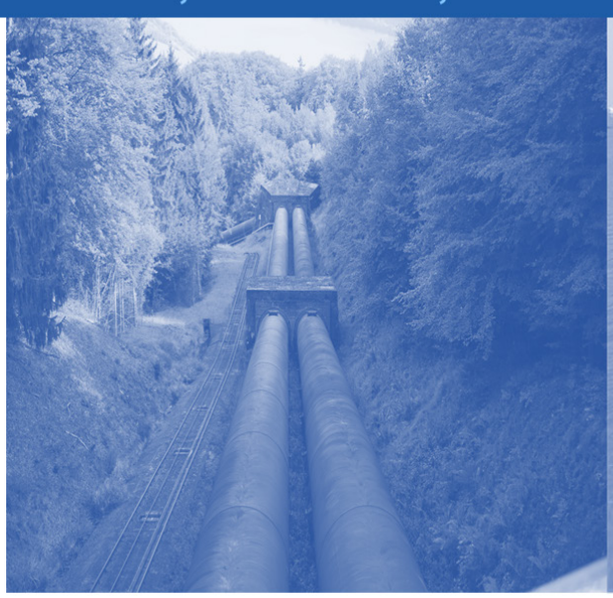
Goals & Principles

- Enterprise perspective in design
- Balance local data needs with enterprise data needs
- Ensure business accountability

Essential Concepts



Extract, Transform, Load



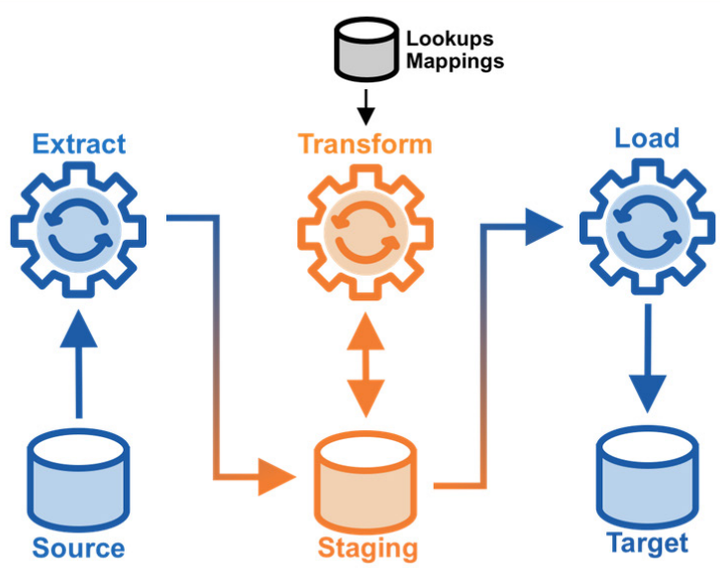
Extract, Transform, Load

Physical staging

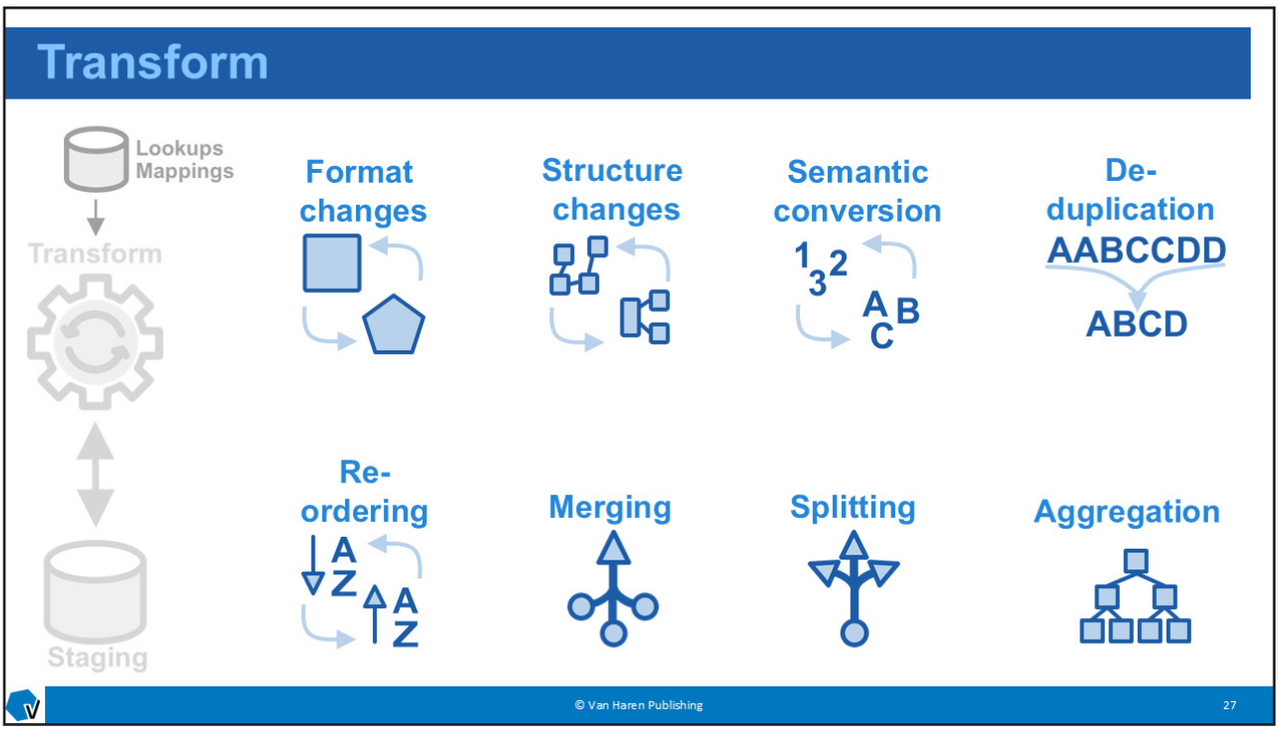
Virtual staging

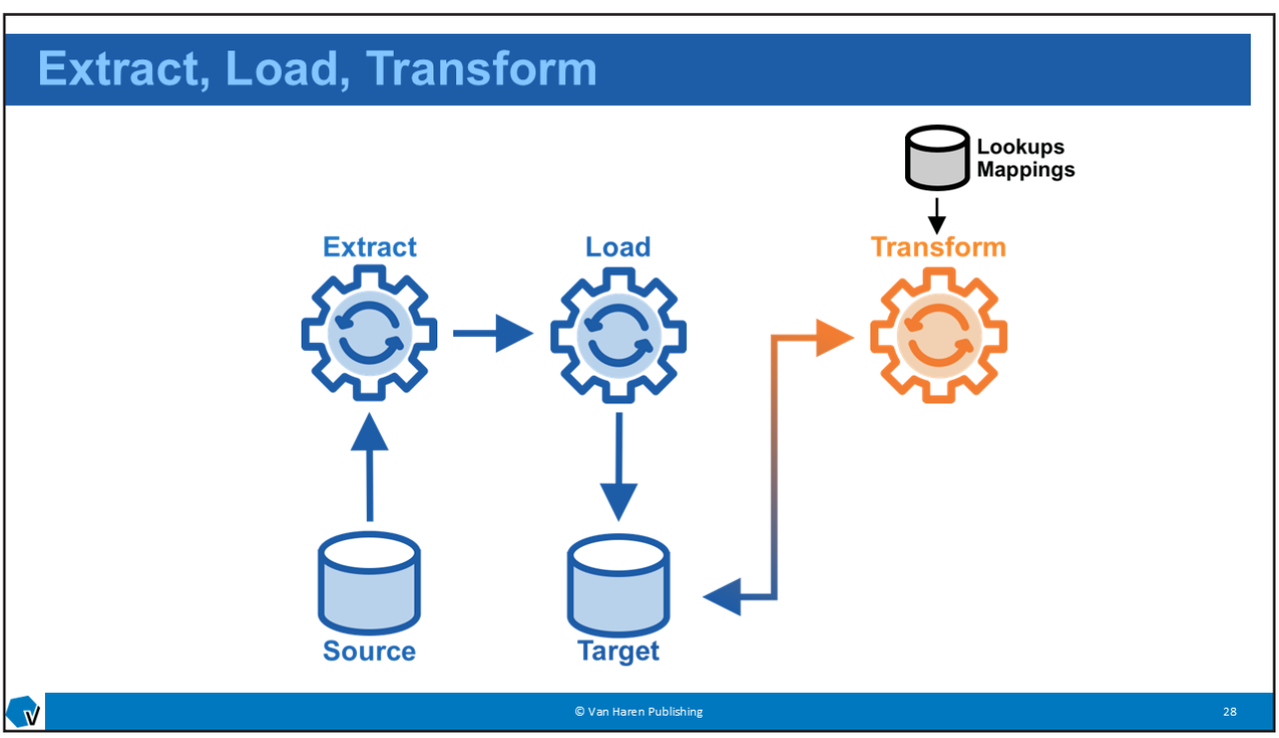
 25

Extract, Transform, Load



 26


Transform
Lookups
Mappings
Transform
Staging
Format changes
Structure changes
Semantic conversion
1 3 2
A B C
De-duplication
AABCCDD
ABCD
Re-ordering
A Z
A Z
Merging
Splitting
Aggregation
© Van Haren Publishing
27


Extract, Load, Transform
Lookups
Mappings
Extract
Load
Transform
Source
Target
© Van Haren Publishing
28

ETL and ELT considerations

	ETL	ELT
Pros	Structured Processing	Scalabilty
	Data Quality	Real-time Data
	Historical Data Transformation	Simple Architecture
Cons	Latency	Raw Data Load
	Complexity	Data Quality
	Mainly Batch Processing	Dependency on DWH

Processing types

Batch

Change Data Capture

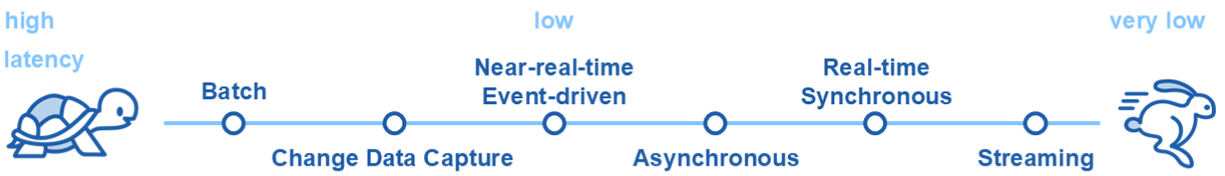
Near-real-time
Event-driven

Asynchronous

Real-time
Synchronous

Streaming

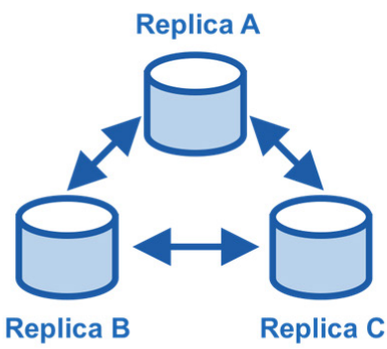
Latency



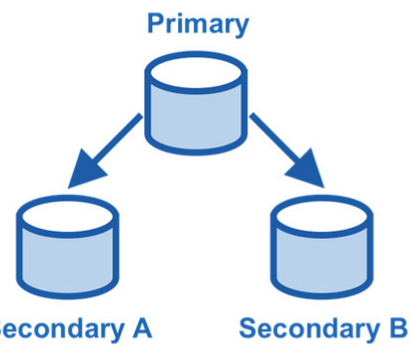
Replication

Active and Passive Replication

Active Replication

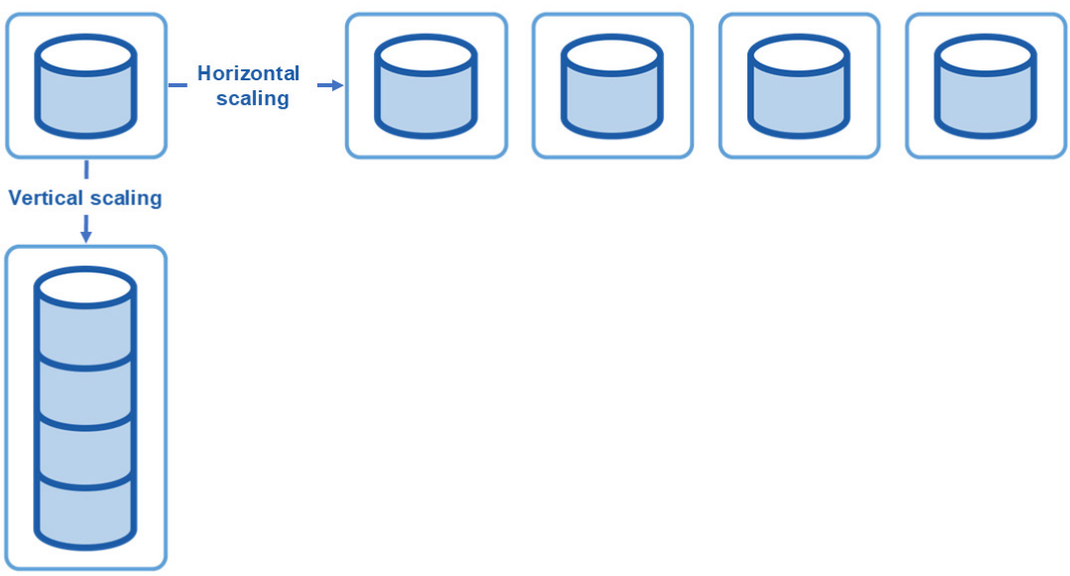


Passive Replication



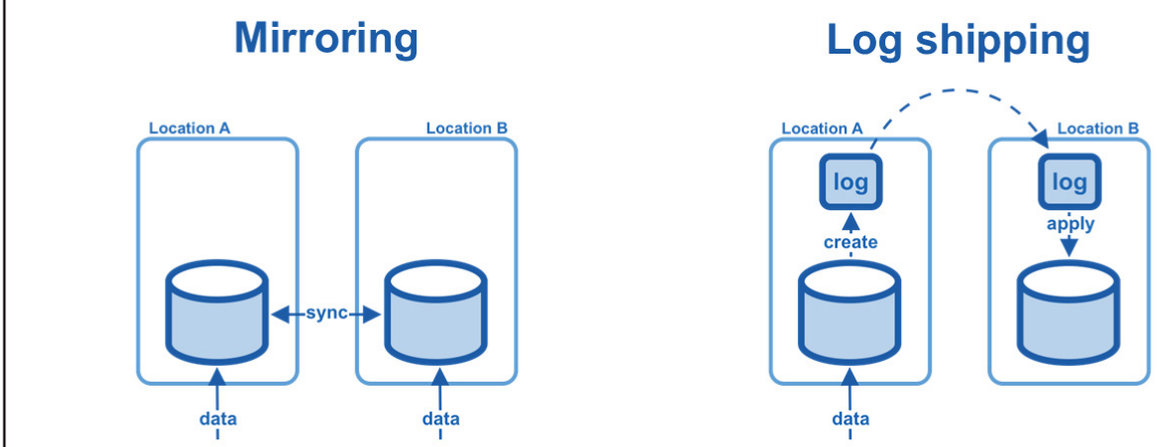
 33

Horizontal and Vertical data scaling

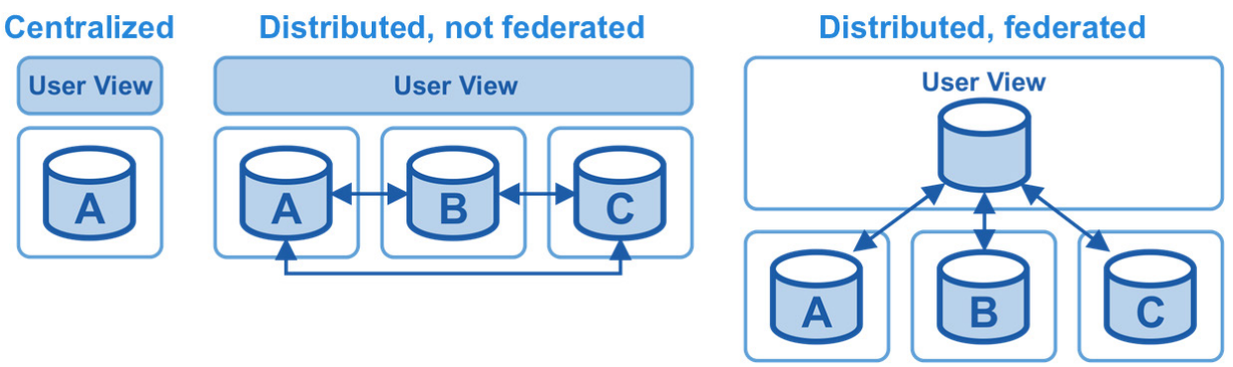


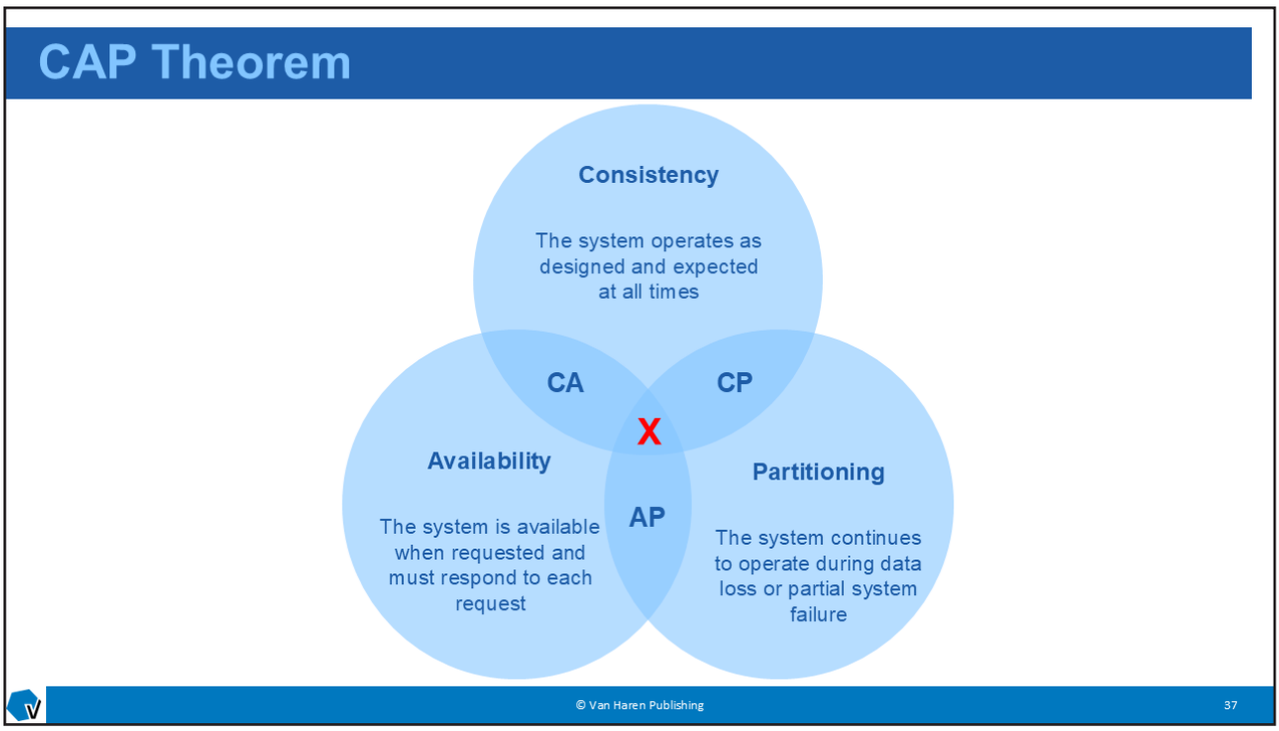
 34

Replication patterns



Distributed databases




CAP Theorem
Consistency
The system operates as designed and expected at all times
CA
CP
X
Availability
The system is available when requested and must respond to each request
AP
Partitioning
The system continues to operate during data loss or partial system failure
© Van Haren Publishing
37


Archiving
Cost Efficiency
Compliance & Regulations
Resource Optimization
Disaster Recovery
© Van Haren Publishing
38

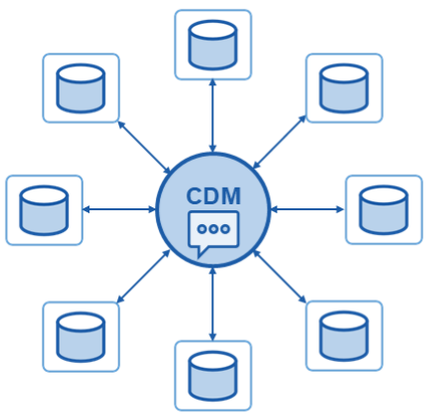
Canonical Model

Easy overview

One single language

Scalable

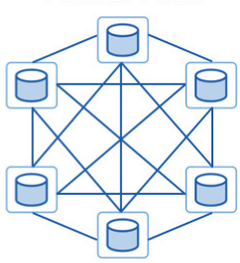
Speed of Development



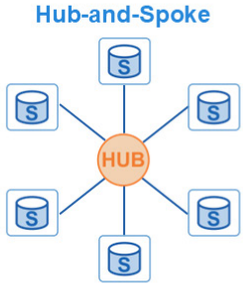
 39

Interaction models

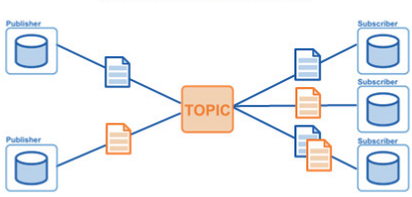
Point-to-Point



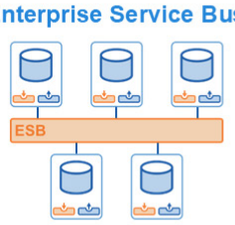
Hub-and-Spoke



Publish-Subscribe



Enterprise Service Bus

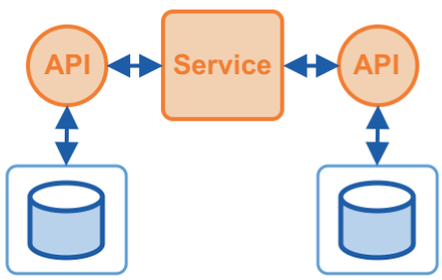


 40

Architecture: Application coupling

Tight coupling

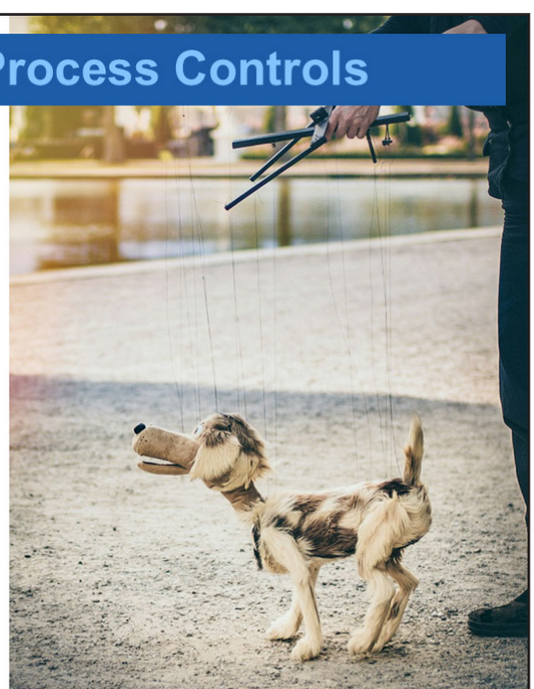
Loose coupling



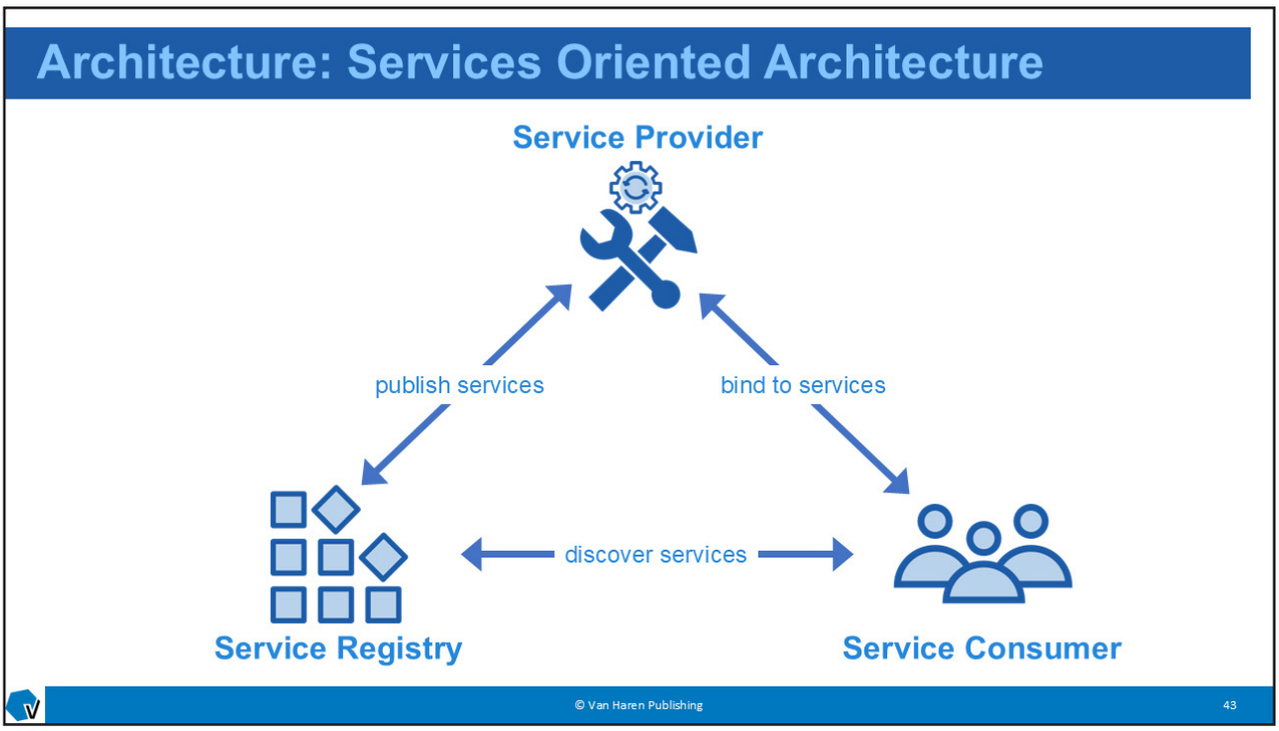
 41

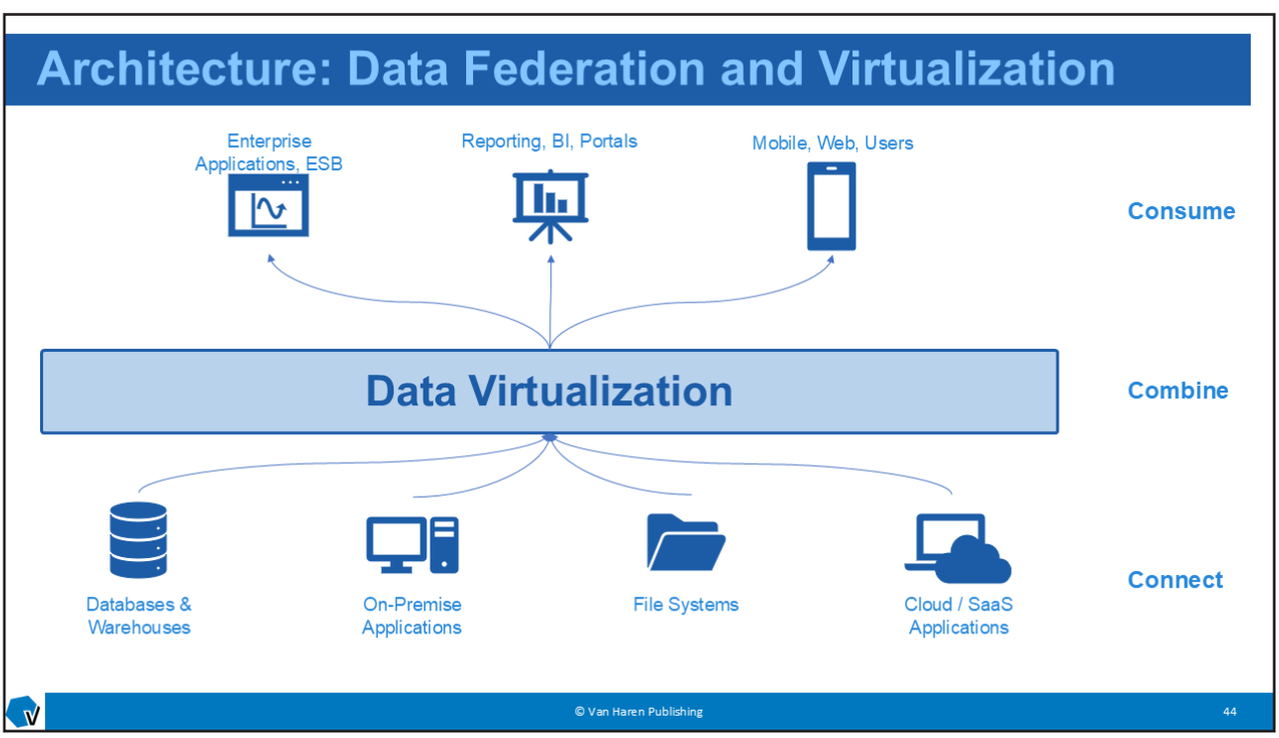
Architecture: Orchestration & Process Controls

- Database activity logs
- Batch job logs
- Alerts
- Exception logs
- Job dependence charts
- Job 'clock' information



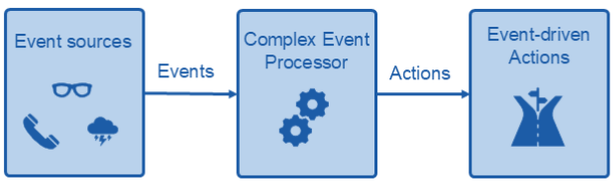
 42


Architecture: Services Oriented Architecture
Service Provider
publish services
bind to services
discover services
Service Registry
Service Consumer
© Van Haren Publishing
43

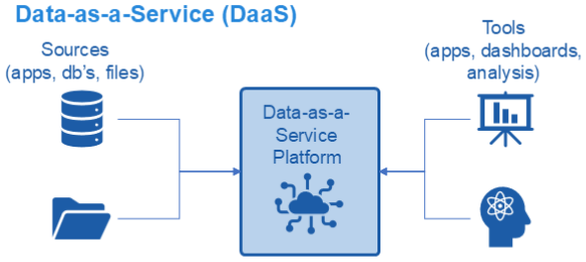

Architecture: Data Federation and Virtualization
Enterprise Applications, ESB
Reporting, BI, Portals
Mobile, Web, Users
Consume
Data Virtualization
Combine
Databases & Warehouses
On-Premise Applications
File Systems
Cloud / SaaS Applications
Connect
© Van Haren Publishing
44

But wait, there's more!

Complex Event Processing



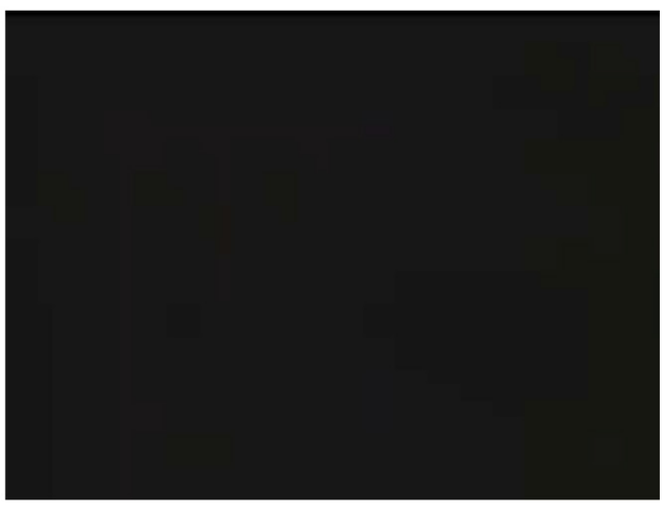
Data-as-a-Service (DaaS)



Cloud-based Integration (iPaaS)

 45

Universal Business Adapter (IBM)



 46

Data Exchange Standards

 47

SDLC

 48

Activities: Plan & Analyze

Define Data Integration and Lifecycle **Requirements**

Perform Data **Discovery**

Document **Data Lineage**

Profile Data

Collect **Business Rules**

 49

Activities: Design

Design Data Integration **Architecture**

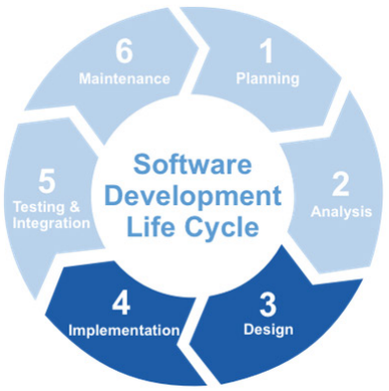
Model Data Hubs, Interfaces, Messages, and Data Services

Map Data Sources to Targets

Design Data **Orchestration**

 50

Activities: Develop



Develop Data **Services**

Develop Data **Flows**

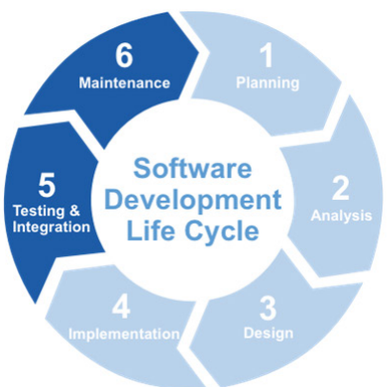
Develop Data **Migration** Approach

Develop **Publication** Approach

Develop **Complex Event Processing** Flows

Maintain DII **Metadata**

Activities: Implement and Monitor



Activate data services

Real-time **monitoring** of processes

Establish parameters to indicate **potential** processing **problems**

Monitor data interaction capabilities

Tools

It is essential to have good tools, but it is also essential that the ***tools should be used in the right way****.*

Wallace D. Wattles, American writer

 53

Tools

Data Transformation Engine / ETL tool

Data Virtualization Server

Enterprise Service Bus

Business Rules Engine

Data & Process modeling tools

Data Profiling tools

Metadata Repository

 54

Implementation guidelines

Ideas are easy
Implementation is hard

Guy Kawasaki

 55

Implementation guidelines

Readiness Assessment
Risk Assessment

Organization and
Cultural Change

 56