Applied A.I. Solutions

Foundations of Data Management

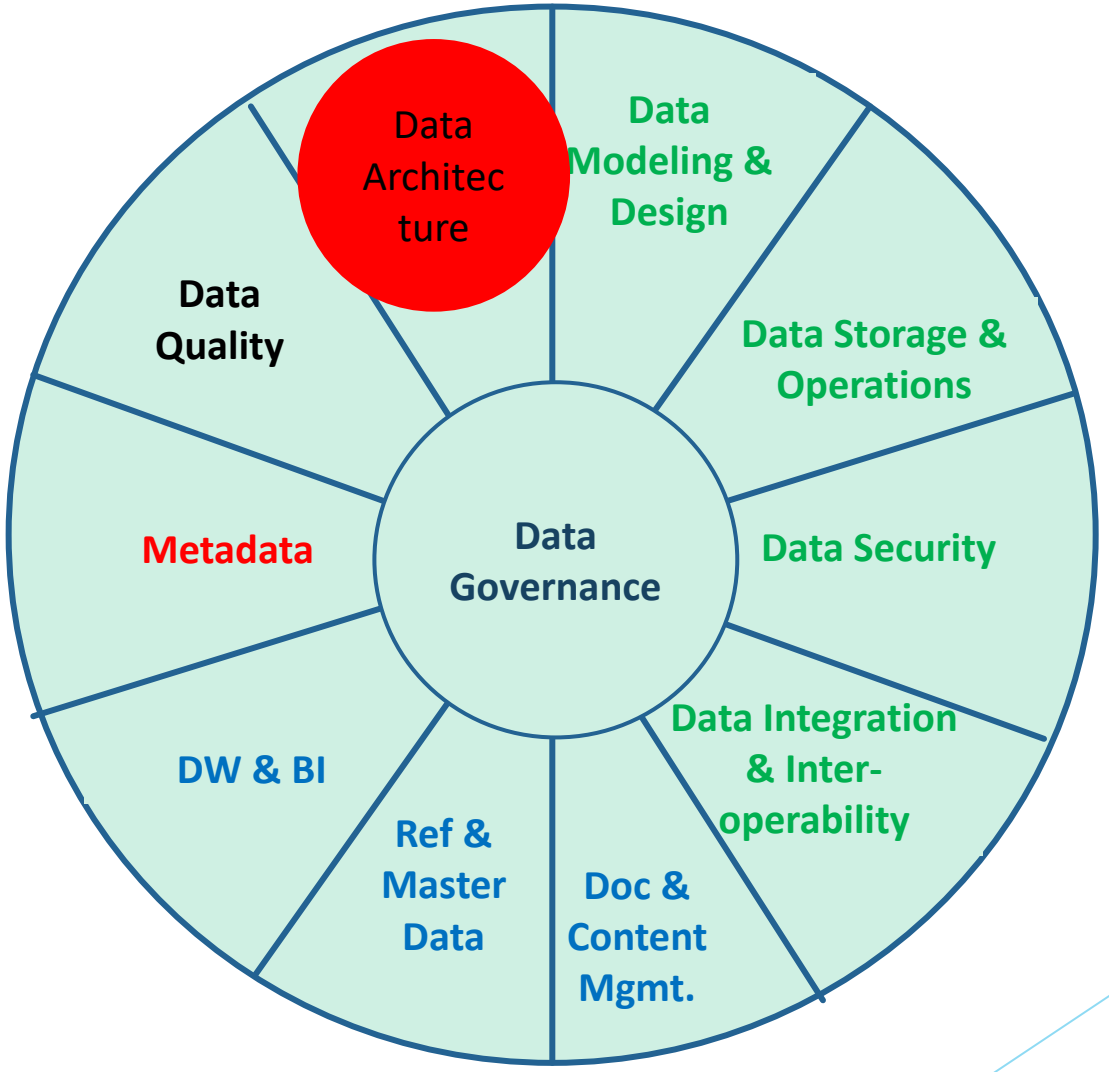
Professor

Daniel Vitaver-Bronstein, B.Sc., EMBA

daniel.vitaver-bronstein@georgebrown.ca

DATA ARCHITECTURE

The DAMA Wheel



[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Introduction

- Architecture refers to the art and science of building things and to its result

- Enterprise Architecture encompasses domain architecture, including business, data, application and technology

[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

- Data Architecture is fundamental to DM
- The master design documents represents data at different levels of abstraction
- Provides a formal data model, with data names, comprehensive data, Metadata definitions, conceptual and logical entities and relationships, and business rules
- Provides standards that governs how data is collected, stored, arranged, used and removed

[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Definition

- To design and maintain the master blueprints to meet identified enterprise data needs
- Using the master blueprints, to guide data integration, control data assets, and align data investments with business strategy

[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Goals

- Identify data storage and processing requirements
- Design structures and plans to meet current and future data requirements
- Strategically prepare the organization to quickly take advantage of new business opportunities, and technologies

[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Inputs

- Enterprise Architecture
- Business Architecture
- IT Standards and Goals
- Data Strategies

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Activities

1. Establish Enterprise Data Architecture
    - Evaluate existing data architecture specifications
    - Develop a Roadmap
    - Manage enterprise requirements within projects
2. Integrate with Enterprise Architecture

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

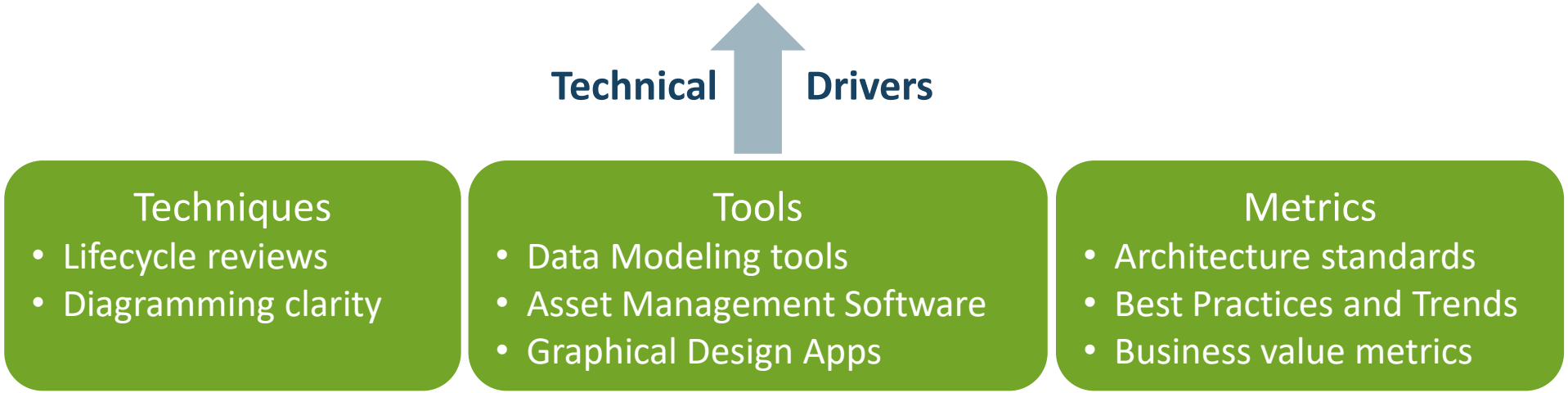
Deliverables

- Data Architecture Design
- Data Flows
- Data Value Chains
- Enterprise Data Model
- Implementation Roadmap

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Suppliers	Participants	Consumers
• Enterprise Architects • Data Stewards • SME • Data Analysts	• Enterprise Data Architects • Data Modelers	• Database administrators • Software developers • Project Managers • Support Teams

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA



[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Business Drivers

- Readiness to take advantage of new business opportunities and new technologies
- Fulfillment of business process requirements re data
- Manage data and information delivery in the enterprise

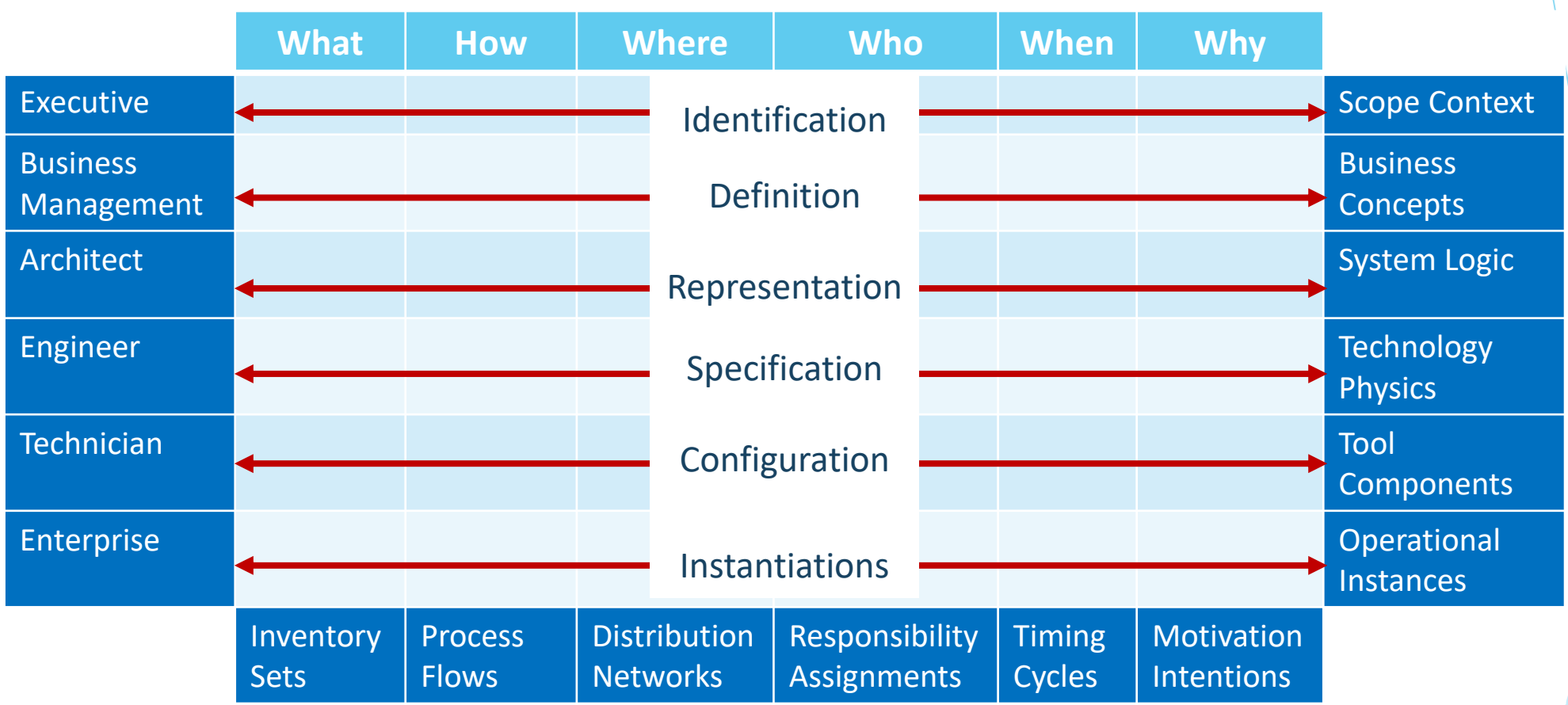
[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Domains

Domain	Business Architecture	Data Architecture	Application Architecture	Technical Architecture
Purpose	To create value	To describe how data should be organized, managed	To describe app structure and functionality	To describe the physical tech to deliver value
Elements	Business models, processes, strategies, capabilities, vocabulary	Data models, data definitions, data mapping, specs, dataflows, APIs	Business systems, SW packages, databases	Technical platforms, networks, security, integration tools
Dependencies	Establishes req for other domains	Manages data created and required by EBA	Acts on specified data according to business req	Hosts, executes the E App A
Roles	Business Architects, Analysts, Data Stewards	Business Architects, Modelers, Data Stewards	Application Architects	Infrastructure architects

[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Zachman Framework



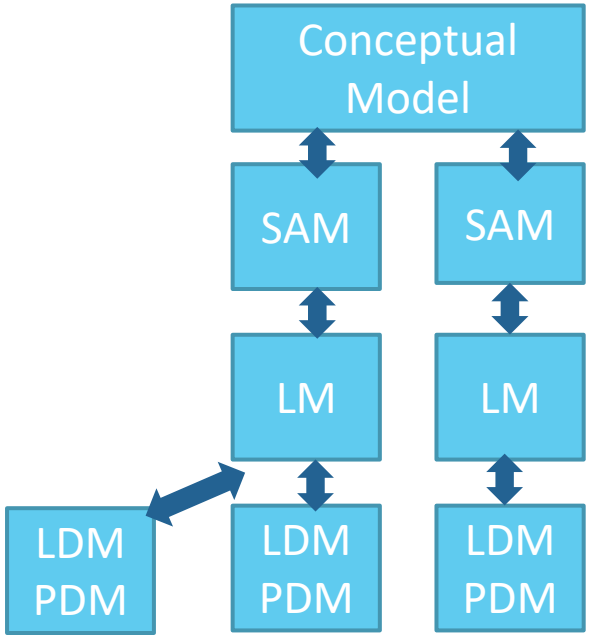
[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Enterprise Data Model (EDM)

- It is a conceptual or logical data model providing a common consistent view of data across the enterprise
- It includes key enterprise data entities, their relationship, mapping, specs, business rules and critical attributes
- It is the foundation for all data and data-related projects
- It can be built incrementally and iteratively, using layers

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

- A conceptual view over the enterprise's subject areas
- Views of entities and relationships for each subject area
- Detailed, partially attributed logical views of the same subject area
- Logical and physical models specific to an application or project



Application or Project-Specific

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Models

EDA deals with two complexities: a) Quality, and b) Innovation

EDA practice includes:

- Strategy (framework, roadmap)
  - Acceptance and culture (behaviour)
    - Organization (accountabilities)
      - Working methods
        - Results (DA)

[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Entities

- Entity is a thing that exists “per se”, separate from other things
- It is a thing about which the organization collects information
- They are referred to as nouns
- The entity represents the answer to:
    - Who
    - What
    - When, Where, Why
    - How
    - Measurement

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

- Entity definitions are core Metadata
- High quality data definition has three main characteristics:
    - Clarity
    - Accuracy
    - Completeness

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

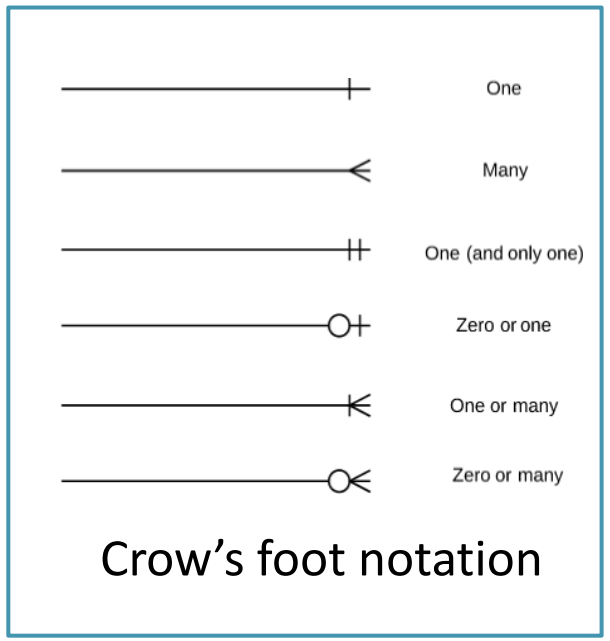
Relationship

- Association between entities (Chen, 1976)
    - It captures the high-level interactions between conceptual entities
    - The detailed interactions between logical entities
    - The constraints between physical entities

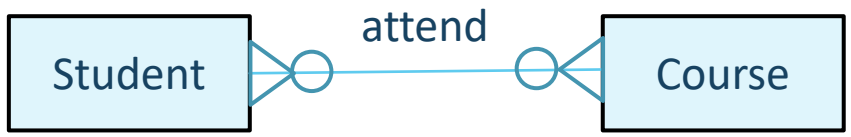
[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Relationship Cardinality

- It captures how many entity-instances participate in a relationship with how many of the other entity



Crow's foot notation

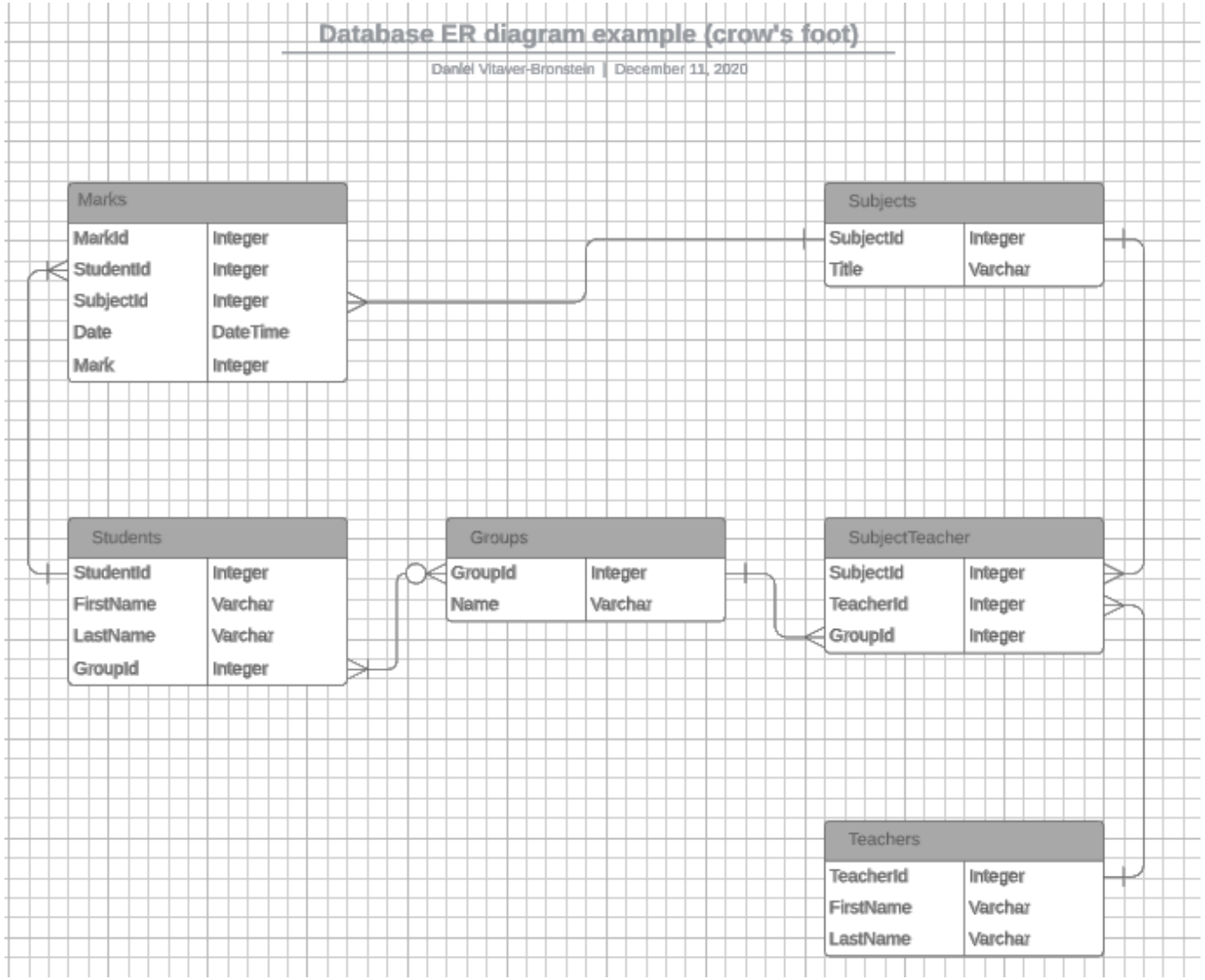


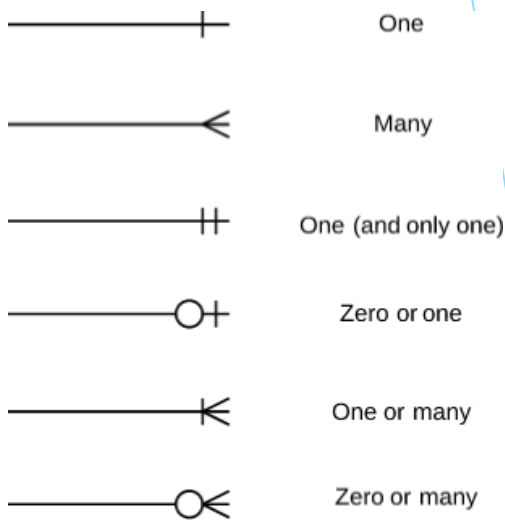
Business Rules

- Each student may attend one or many courses
- Each course may be attended by one or many students

[1] Source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

Entity-Relationship Diagram ERD

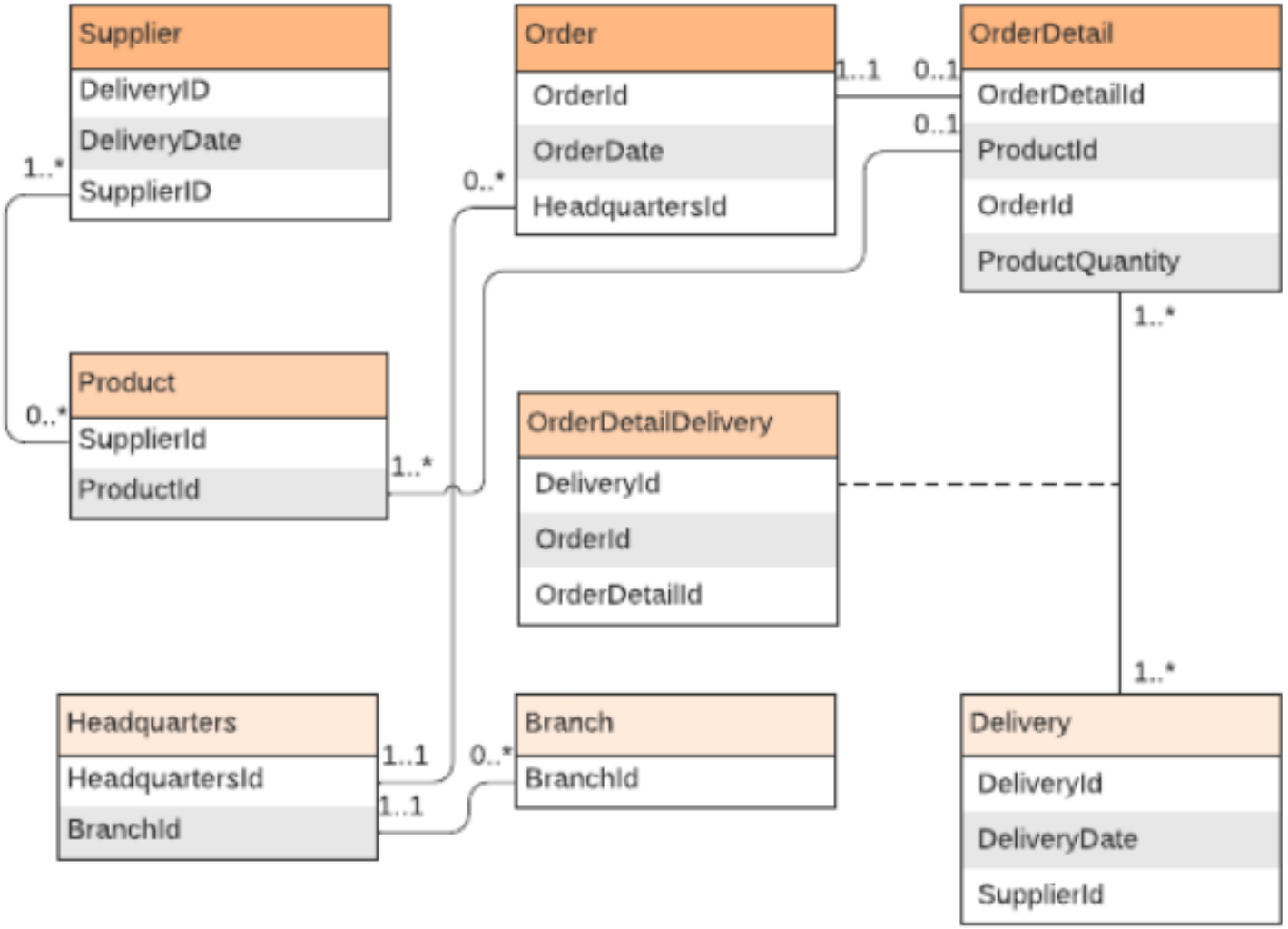




Crow's foot notation

ERD with colored entities example (UML notation)

System Templates | October 12, 2020



Traceability

All levels are part of the EDM, and linkages create paths to trace an entity from top to bottom and between models

- **Vertical**: models in each level map to models in other levels
- **Horizontal**: entities and relationships may appear in multiple models in the same level

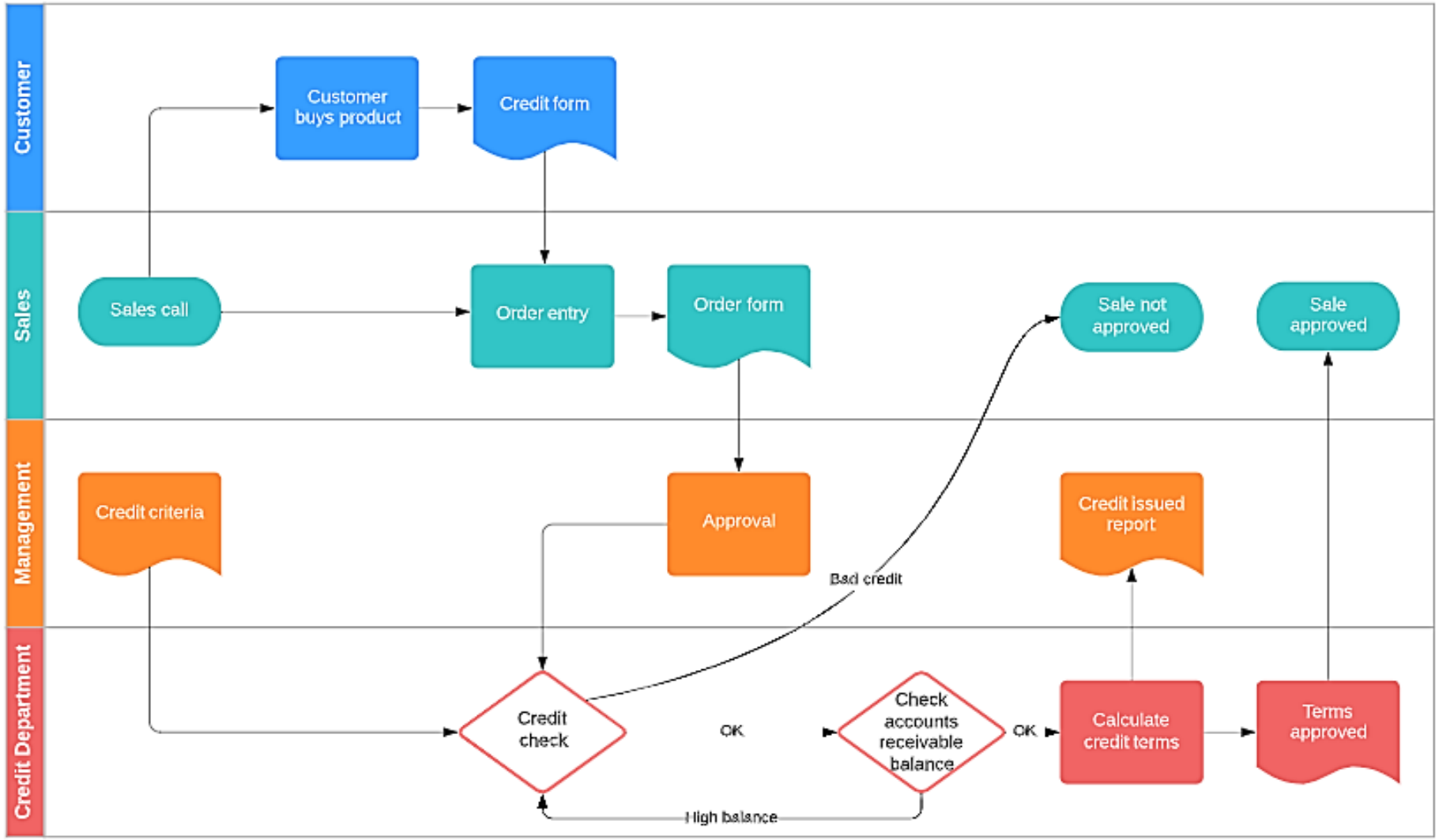
[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

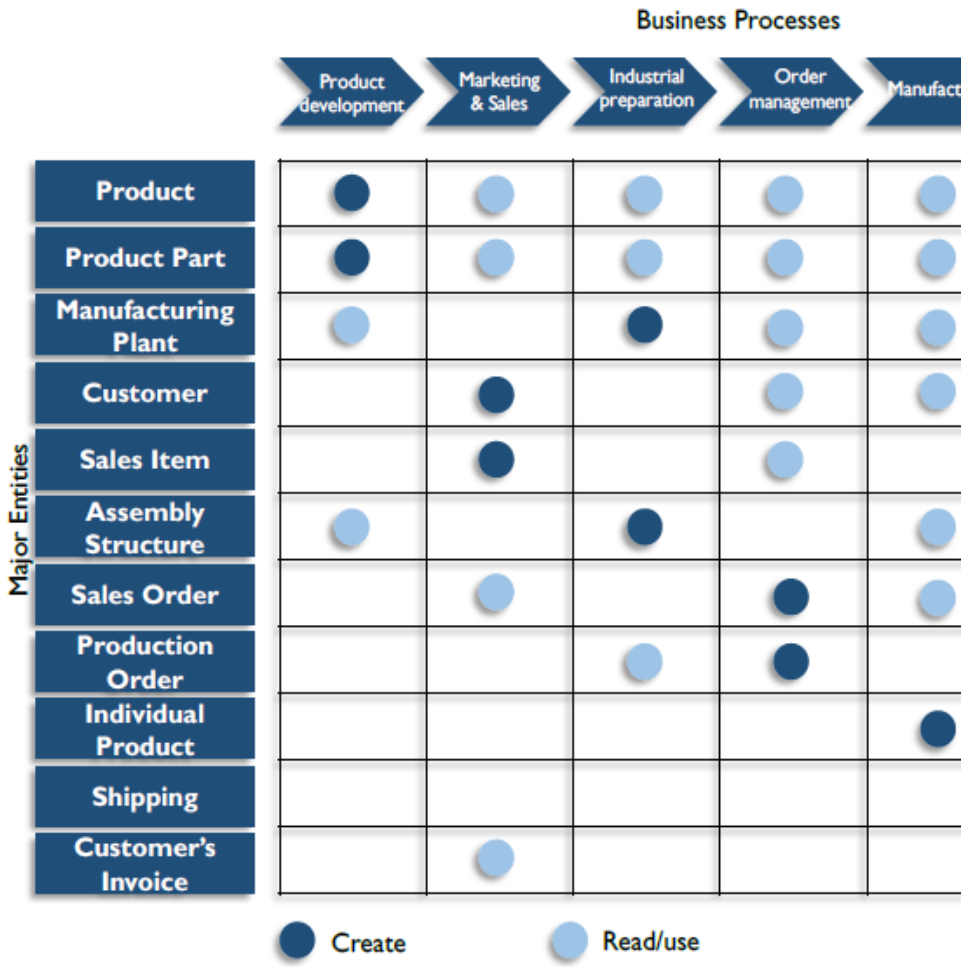
Data Flow Design

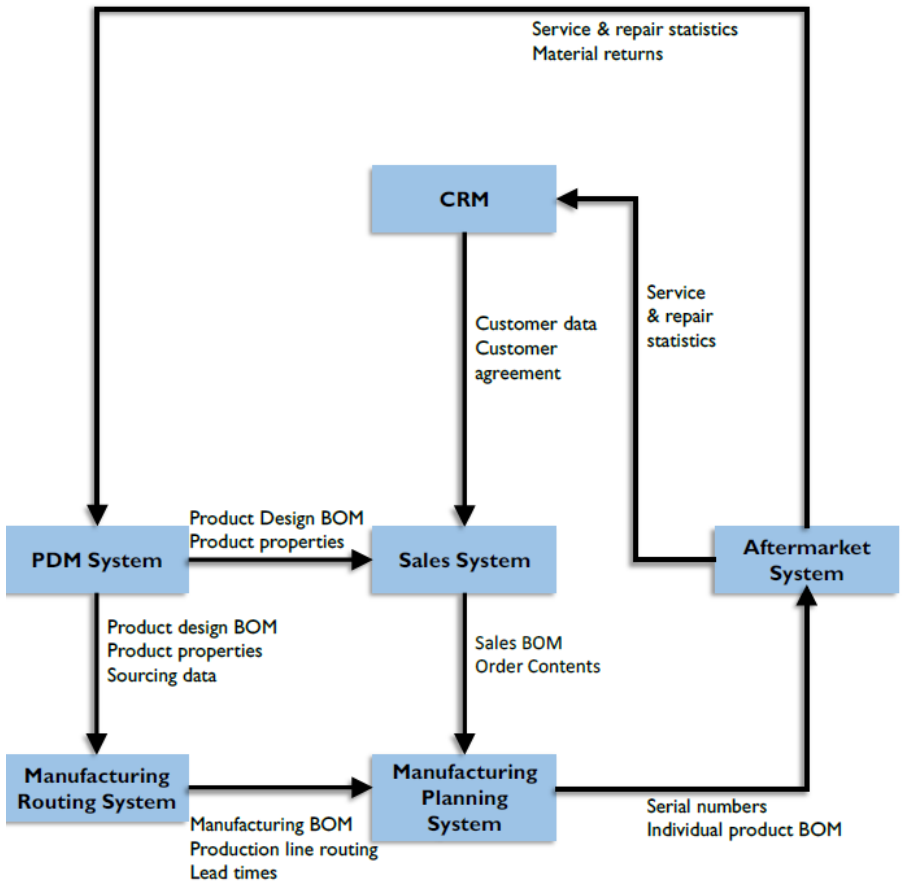
- Defines the requirements and master blueprint for data storage and processing across databases, applications, platforms, and networks

- Data flows **map the movement of data** to business processes, locations, business roles (CRUD) and technical components

[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA


Customer
Sales
Management
Credit Department
Customer buys product
Credit form
Sales call
Order entry
Order form
Sale not approved
Sale approved
Credit criteria
Approval
Credit issued report
Bad credit
Credit check
OK
Check accounts receivable balance
OK
Calculate credit terms
Terms approved
High balance





[1] Main source: Copyright © 2017 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

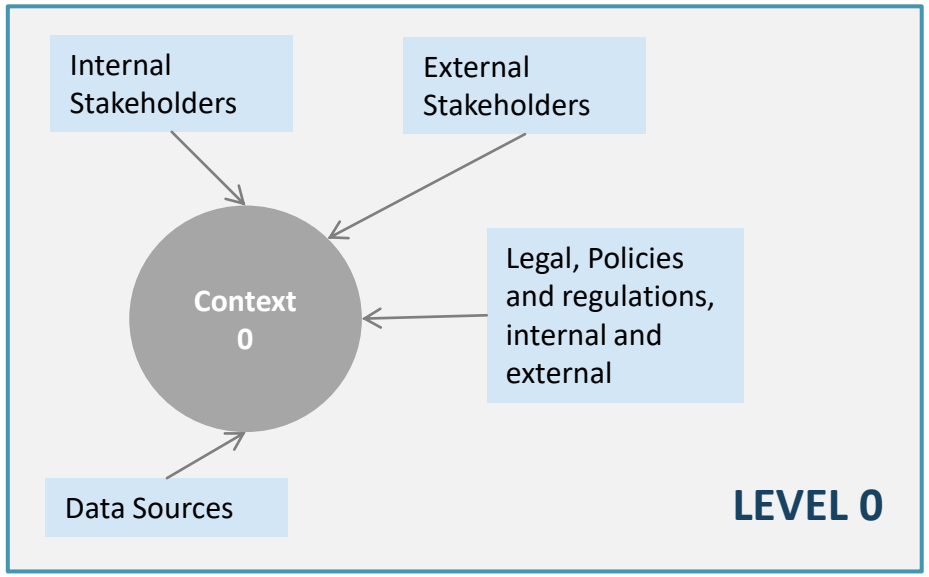
Business Processes

- A process is a set of interrelated actions and activities performed to achieve a pre-specified product, result, or service
- Each process is characterized by its input, the tools and techniques that can be applied, and the resulting outputs
- Each process can be broken down into other processes, and so on, until we reach a level in which simple activities are defined using measurable inputs and measurable outputs

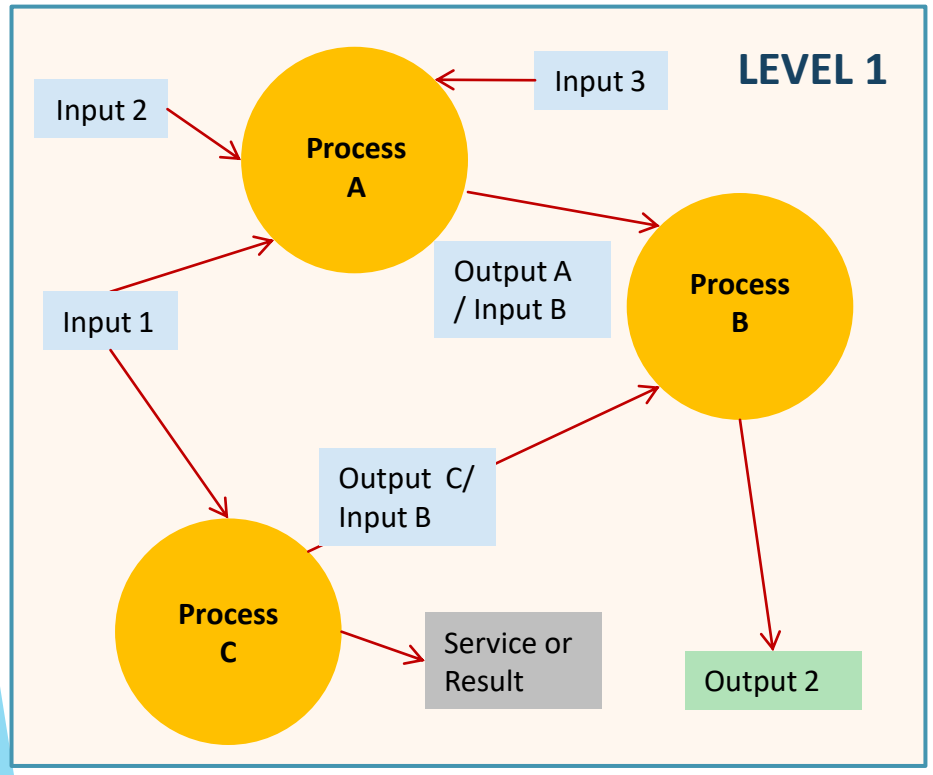
Business Processes

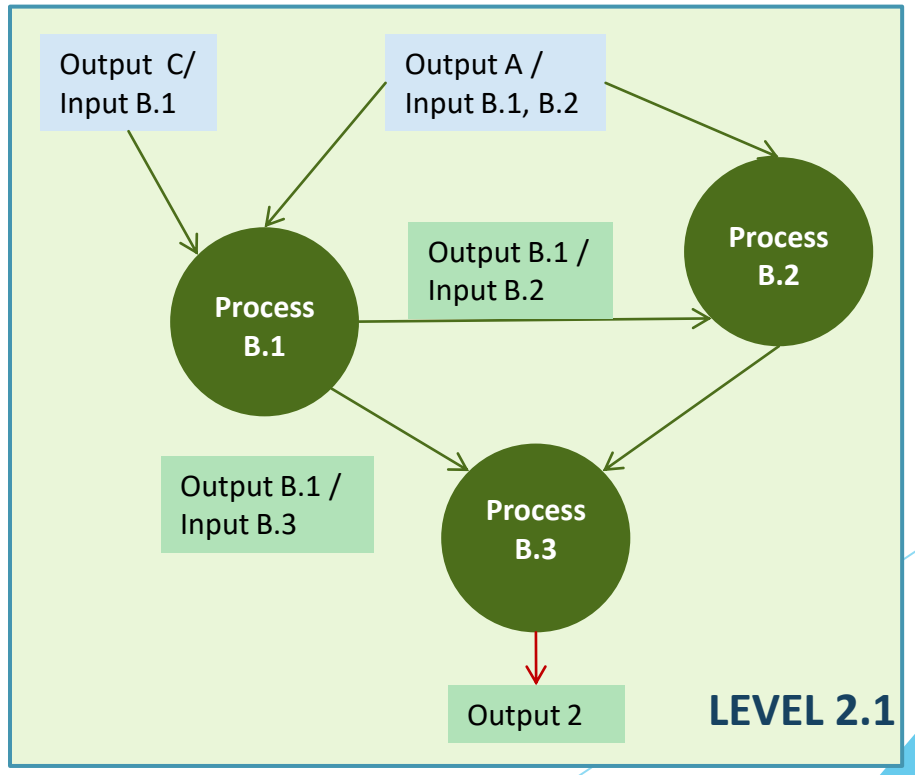
- Level 0 (“The Context”) shows the relationship between the main process and stakeholders
- Level 1: Level 0 is broken down into major processes
- Level 2: Level 1 is broken down into other processes/activities

  ::::::::::::
- Level n: level n-1 is broken down into measurable tasks

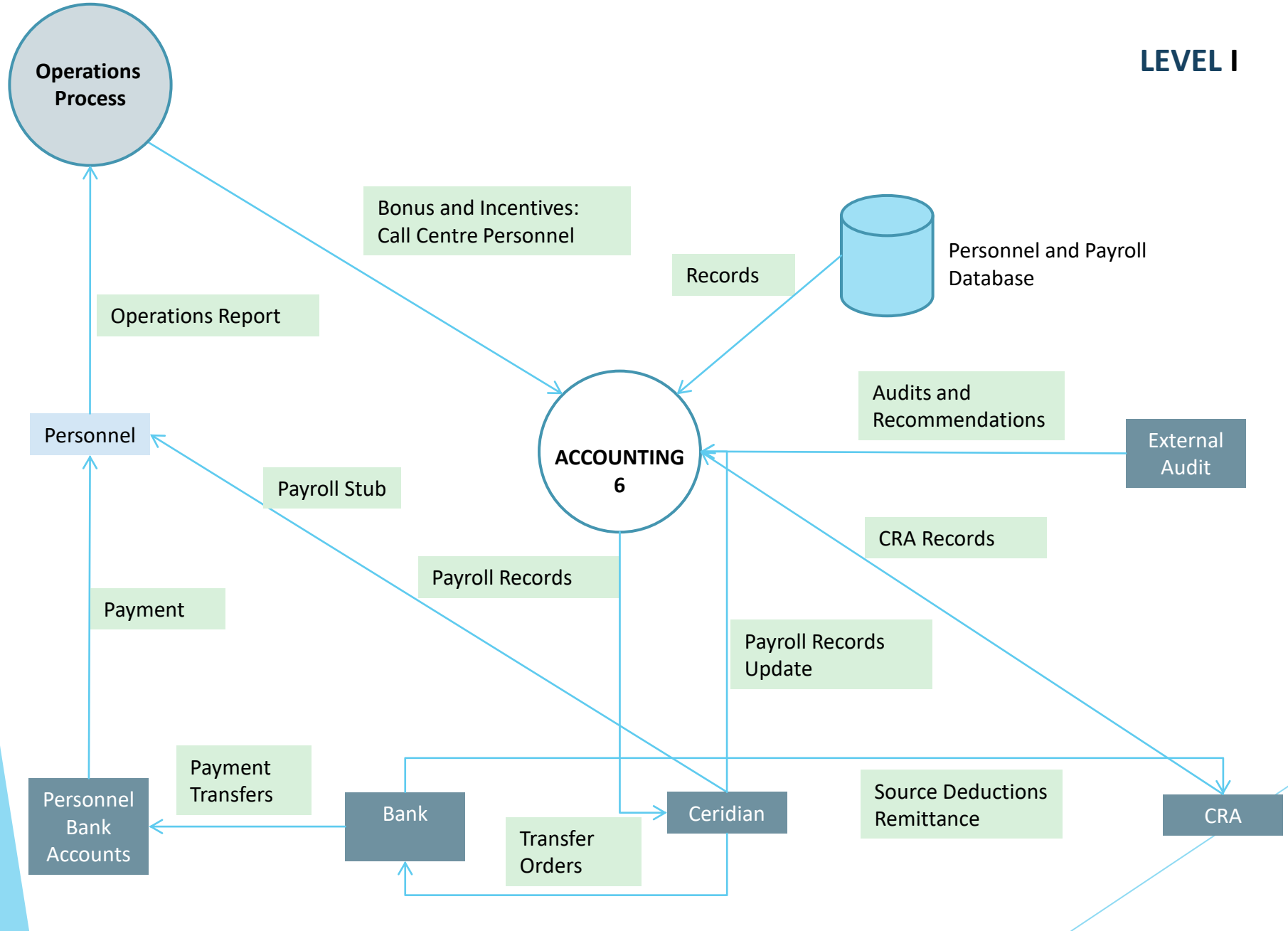

Internal Stakeholders
External Stakeholders
Context 0
Legal, Policies and regulations, internal and external
Data Sources
LEVEL 0


LEVEL 3


LEVEL 1
Input 2
Input 3
Process A
Process B
Output A / Input B
Input 1
Output C/ Input B
Process C
Service or Result
Output 2


Output C/ Input B.1
Output A / Input B.1, B.2
Process B.1
Output B.1 / Input B.2
Process B.2
Output B.1 / Input B.3
Process B.3
Output 2
LEVEL 2.1

LEVEL I


Operations Process
Bonus and Incentives: Call Centre Personnel
Records
Personnel and Payroll Database
Operations Report
Audits and Recommendations
Personnel
ACCOUNTING 6
External Audit
Payroll Stub
CRA Records
Payroll Records
Payment
Payroll Records Update
Payment Transfers
Personnel Bank Accounts
Bank
Ceridian
Source Deductions Remittance
CRA
Transfer Orders