# DATA ARCHITECTURE IN VIETNAMESE BANKS

IT4350 - Information Systems Architecture and Applications

Dr. Nguyen Binh Minh

1

## TOPICS

1. **Data in Banking**
   Highlight main data source of a bank Challenge when working with banking data Why Banks need Data Architecture
2. **Data Architecture components**
   Data Delivery Architecture Enterprise Data Model
   Information Value Chain
3. **How to become a Data Architect** Data Architect as a professional Knowledge, skills, experience
   Sample career path

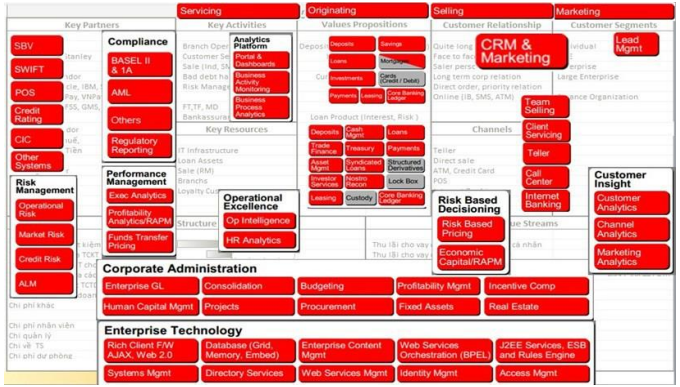
2

| Key Partners | Key Activities | Values Propositions | Customer Relationship | Customer Segments |
|---|---|---|---|---|
| Investors | Branch Operation | Deposit (Interest Rate, Branch HO) | Quite long term (years) | Individual |
| SVN, Morgan Stanley | Customer Services | Flexible withdrawal | Face to face (service consult) | SME |
| | Sale (Ind, SME, Corp) | | Saler personal relationship | Enterprise |
| Technology Vendor | Bad debt handling | Current (Better FT service) | Long term corp relation | Large Enterprise |
| Temenos, Oracle, IBM, SAP, Misys | Risk Management | | Direct order, priority relation | |
| Banknet, OnePay, VNPay, Incom | | | Online (IB, SMS, ATM) | Finance Organization |
| FPT,.. | FT,TF, MD | | | |
| | Bankassurance | Loan Product (Interest, Risk ) | | |
| Regulatory Vendor | **Key Resources** | | **Channels** | |
| SVN, CIC, TC Thuế, | | Fee (Finance Services) | | |
| BHTG, PC Rửa Tiền | IT Infrastructure | | Teller | |
| | Loan Assets | Security Trading | Direct sale | |
| Banking Partner | Sale (RM) | | ATM, Credit Card | |
| BIDV, SHB, | Branchs | Trusted Investment (UTDT) | POS | |
| Vietcombank | Loyalty Cust | | Internet Banking | |
| Visa | | | Mobile, SMS | |

| Cost Structure | Revenue Streams |
|---|---|
| Trả lãi tiền gửi tiết kiệm | Thu lãi cho vay các tổ chức kinh tế và cá nhân |
| Trả lãi tiền gửi của TCKT và cá nhân | Thu lãi cho vay đồng tài trợ qua TCTD |
| Trả lãi vốn nhận UT cho vay của các TCKT & Cá | Thu lãi cho vay bằng vốn tài trợ, ủy thác đầu tư c |
| Trả lãi tiền gửi của các Tổ chức tín dụng khá | HĐ Kinh doanh tiền tệ |
| Trả lãi tiền vay các TCTD khác | HĐ Đầu tư |
| Chi mua bán kinh doanh chứng khoán | Thu lợi nhuận, cổ tức đầu tư dài hạn |
| Chi phí khác | Hoạt động khác (phí, hoàn nhập) |
| Chi phí nhân viên | |
| Chi quản lý | |
| Chi về TS | |
| Chi phí dự phòng | |

3

## IT solution in Banking Business (sample)



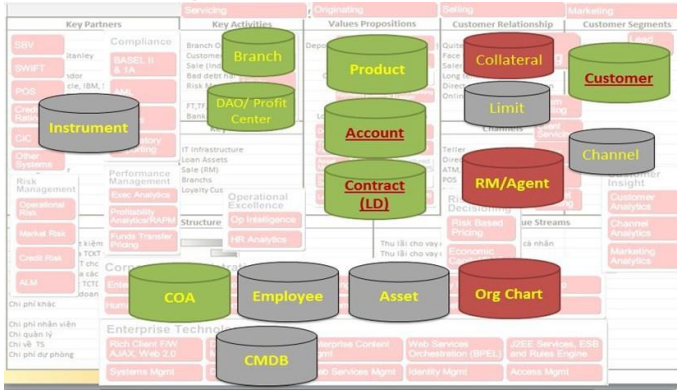

4

# Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path

5

## Data in Banking Business (logical master data)




6

## Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path

7

## CHALLENGE WHEN WORKING WITH BANKING DATA

- Lot of data
  - T24 Core Bank has 30 Banking Ops, 300 modules, 3000 applications
  - Flexcube Corebank DB has 10 Banking Ops, 1500 Tables
  - Oracle ERP Finance DB has 10 Modules, 3000 Tables
- Complicate Business Logic
  - T24 Core Bank userguide cost 3.6GB docs, 66000 files in 4600 folders
  - Flexcube Corebanking user manual has 4000 pages in 30 docs
  - Oracle ERP Userguide for 1 module (AR, AP, GL, PO, OP, INV…) has around 1400 pages.
- Hard to Access
- Lot of Legacies

10

8

## Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path

9

## WHY HUMANKIND NEED ARCHITECTURE

*If you get really honest and search all of history, seven thousand years of known history of humankind, to find out how humanity has learned to cope with two things,* ***complexity*** *and* ***change****, there is one game in town:* ***ARCHITECTURE****.*

*John Zachman*

Architecture help human deal with **Complexity** and **Change**.

10

## WHY BANK NEED DATA ARCHITECTURE

- To deal with Data Complexity
  - Build massive data constraction from smaller parts
  - Make components work together
  - Ensure people talk the samething
- To make change became possible
  - Change with minimum time, disruption, risk and cost
  - Improvement became realistic
  - Connect data value to business value

11

## Agenda

1. **Data in Banking**
   Highlight main data source of a bank
   Challenge when working with banking data
   Why Banks need Data Architecture
2. **Data Architecture components**
   Data Delivery Architecture
   Enterprise Data Model
   Information Value Chain
3. **How to become a Data Architect**
   Data Architect as a professional
   Knowledge, skills, experience
   Sample career path

12

## Information Architecture

**ar· chi· tect** (är'ki-tekt') n. (Abbr. arch., archt.). One who designs and supervises the construction of buildings or other large structures.

**in· for· ma· tion** (in'fər-ma'shən) n. (Abbr. inf.)
Knowledge derived from study, experience, or instruction. Knowledge of a specific event or situation; intelligence.
See synonyms at knowledge. A collection of facts or data; statistical information. The act of informing or the condition of being informed; communication of knowledge.
Example: Safety instructions are provided for the information of our passenger one who designs and supervises the construction of knowledge derived from study, experience, or instruction, or knowledge of a specific event or situation, or a collection of facts or data."

The information architect's job is to define the nature and scope of an information service before the first line of code is put into place. Information architecture is the plan for the data, much like an architect's drawing.

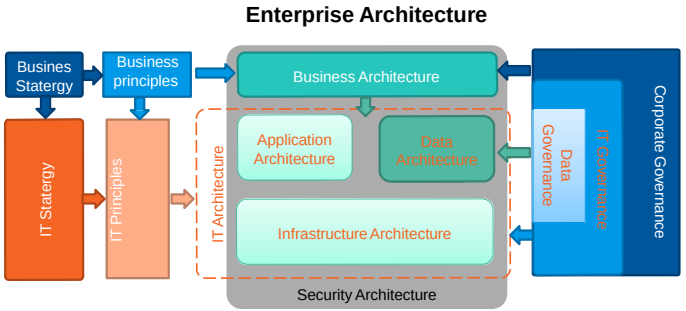
13

## Data Architecture

**Identify the data needs** of the enterprises (regardless of structure) and designing and maintaining the master blueprints to meet those needs. Using master blueprints to guide **data integration**, control **data assets**, and **align data investments with business strategy**.

14

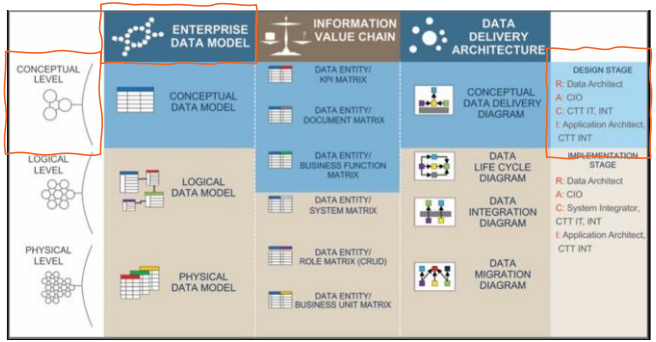
## DATA ARCHITECTURE IN ENTERPRISE ARCHITECTURE




*Source: my data architecture design for a bank*

15

## DATA ARCHITECTURE COMPONENTS




16

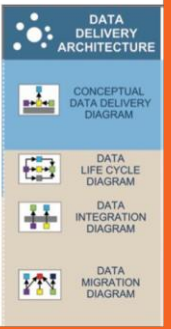
## Components' Levels

| | | |
|---|---|---|
| CONCEPTUAL LEVEL | **Mức khái niệm (Conceptual Level)** | các cấu phần được mô tả một cách tổng quan nhất, các thành phần chi tiết được khái quát hóa thành những khái niệm rộng hơn; các thành phần không quan trọng bị loại bỏ khỏi mô hình. Mục tiêu là để giúp người đọc hình dung tổng quan nội dung các thành phần có trong mức thấp hơn đồng thời thống nhất các "khái niệm" được sử dụng trong Ngân Hàng. |
| LOGICAL LEVEL | **Mức logic (Logical Level)** | các cấu phần được mô tả một cách chi tiết và rõ ràng hơn nhưng vẫn sử dụng các từ ngữ và khái niệm kinh doanh. Logical Data Model có thể bao gồm cả tên thực thể thông tin, thuộc tính và mối quan hệ của chúng với các thực thể thông tin khác. Mục đích của mức trình bày này là để giúp kết nối giữa nghiệp vụ/người sử dụng và công nghệ/người xây dựng. Đặc điểm của các mô hình ở lớp này là nó đủ chi tiết để Người xây dựng biết cách thực hiện và đủ đơn giản để Người sử dụng hình dung được kết quả sau quá trình xây dựng. |
| PHYSICAL LEVEL | **Mức vật lý/triển khai (Physical Level)** | ở mức vật lý/triển khai, các cấu phần được thiết kế rất chi tiết, đầy đủ các thành phần lớn nhỏ cùng với các chỉ dẫn và kí hiệu mà thông thường chỉ có các chuyên gia công nghệ mới hiểu. Mô hình ở mức này là mô hình phản ánh chính xác nhất hiện trạng và tương lai của các cấu phần. |

17

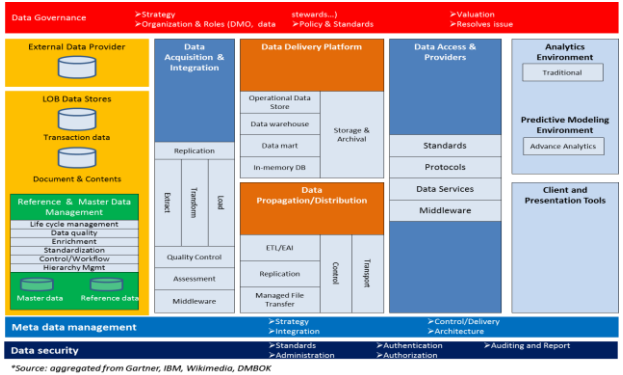
## Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path




18

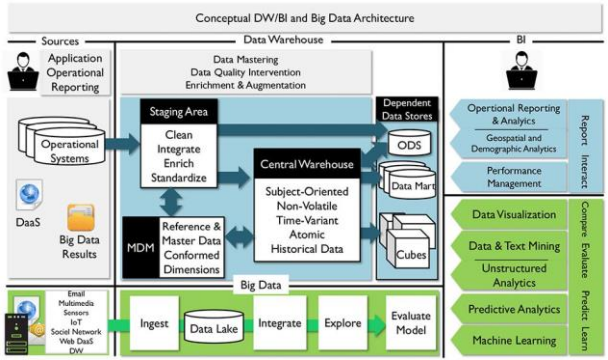
## **Reference Data Delivery Architecture** (conceptual level)




*Source: aggregated from Gartner, IBM, Wikimedia, DMBOK

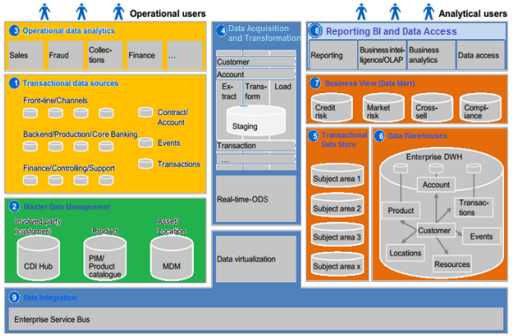
19

## Actual **Data Delivery Architecture** (conceptual level)




20

## Actual **Data Delivery Architecture** (logical level)




From Business User perspective



## DATA DELIVERY ARCHITECTURE HOW TO DESIGN

1. Choose a **Reference Architecture**
2. Business Strategy + IT Strategy = **Targeted Data Maturity** Level
3. Design Targeted **Data Delivery Architecture**
4. Discover current status => **Current Data Maturity** Level
5. Identify **Gaps** = Targeted Data Maturity Level - Current Data Maturity Level
6. Define **Data Business requirements** for all components
7. Choose the **Key Success Factors** for all component
8. Develop **Road map** to fullfil the Gaps and Requirements

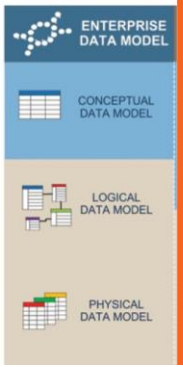
## DATA DELIVERY ARCHITECTURE HOW TO AVOID FAILURE

1. Design with **Architecture Level** in mind
2. **Diagram** just for mind **not enough** for implementation
3. **Right technology** for right problems. New technology for new data problem.
4. Lack of **supporting Data Delivery components**: e.g. data life cycle, data migration, security, metadata management, data governance…

23

## Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path

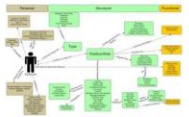



24

## 3 LEVELS OF DATA MODEL

**1 CONCEPTUAL DATA MODELS**

- Key concepts and high level relationships to meet functional need



**2 LOGICAL DATA MODELS**

- Detailed definitions and attributes (characteristics) that are more granular than a CDM

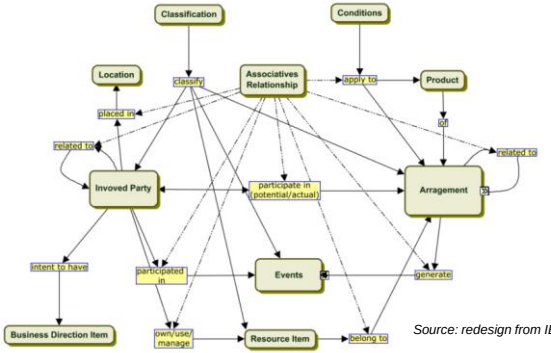
**3 PHYSICAL DATA MODELS**

- Definition of required tables, columns and database properties against the physical application and infrastructure architecture

*Source: PwC*

25

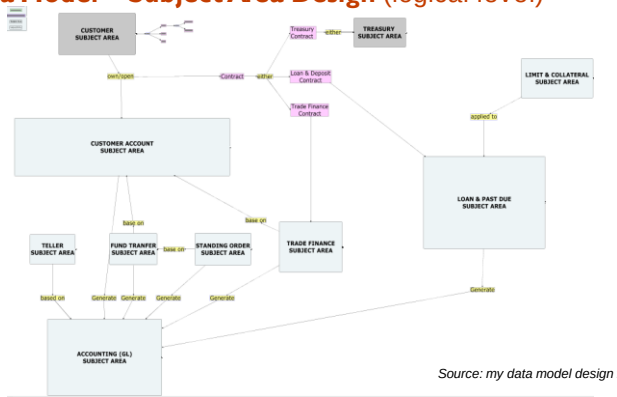
## **Banking Data Model** (sample conceptual level)




*Source: redesign from IBM*

26

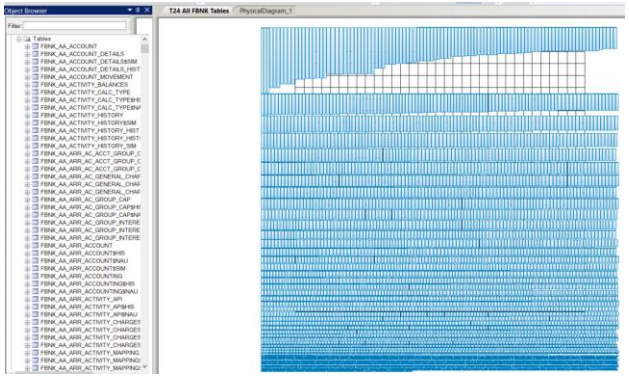
## Data Model - Subject Area Design (logical level)




*Source: my data model design for a bank*

27

## T24 CORE BANKING DATA MODEL (PHYSICAL LEVEL)



28

## ENTERPRISE DATA MODEL HOW TO DESIGN

1. Learn **Data Model Design Mindset**: Relational, Multidimensional, Subject Area
2. From **Data Delivery Architecture choose design mindset** for each component
3. Collect all input information: data dictionary, old data model, reverse engineering, data design patterns
4. Collect and define **non-information data requirements**: data lifecycle, security, retention, volume, velocity, availability...
5. Design Data Model level by level (Conceptual, Logical, Physical)
6. **Test Data Model**

29

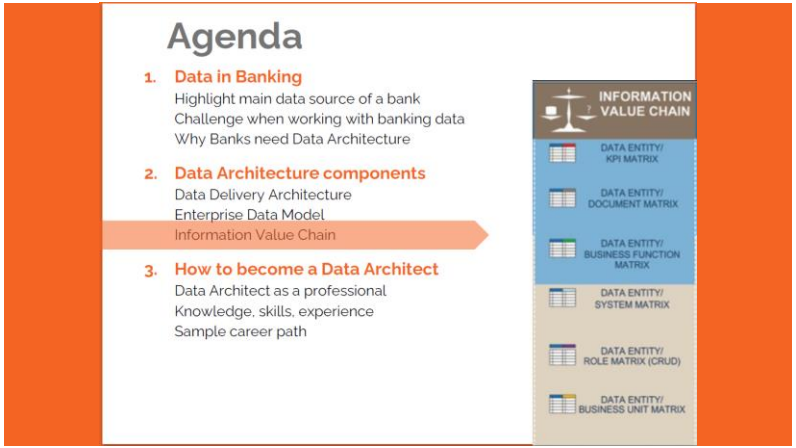
| | | Mindset | | |
|---|---|---|---|---|
| | | **Relational** | **Dimensional** | **Subject Area Oriented** |
| **Types of model** | **CDM** | Key concepts and their business rules, such as an "Each `Customer` may place one or many `Orders`." | Key concepts focused around one or more measures, such as "I want to see Gross Sales Amount by Customer." | Key concepts (subject Area) and their logical relationships, such as "`Credit Contract` generates `Accounting Transaction`" |
| | **LDM** | All attributes required for a given application or business process, neatly organized into entities according to strict business rules and independent of technlogy such as "Each `Customer ID` value must return at most one `Customer Last Name`". | All attributes required for a given reporting application, focused on measures and independent of technology such as "I want to see `Gross Sales Amount` by `Customer` and view the customer's first and last name." | All entites inside a certain Subject Area. Key concepts specify Subjects and supporting concepts around (area) with some simple business rules such as "`Credit Contract` belong to `Customer Segment` and `Product Line`. |

30

## ENTERPRISE DATA MODEL HOW TO AVOID FAILURE

1. Learn **design mindset** to design right data model
2. Apply **best practice** in design
3. Careful, meticulous, perfectionist
4. **Specific** vs **Abtract** (TAB_CAR vs TAB_VEHICLE)
5. **Test** before use

31




32

## INFORMATION VALUE CHAIN

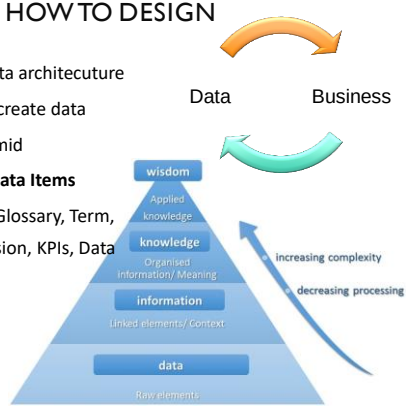
1. Link Business to Data in term of strategy, architect, goal, KPI.
2. Define relationship between:
    1. Business Strategy and Data Strategy
    2. Business Architeture and Data architecture
    3. Business Investment and Data Values
    4. Business Unit and Data Mart
    5. Business Process and Data Subject Area
    6. Business Domain and Data Domain
    7. Business object and data entity
    8. Business terms and data items
    9. KPIs and Measures

| | |
|---|---|
| **Information Value Chain analysis** | A process to link conceptual and logical data models to process models, applications, organizations, roles and/or goals, to provide context, relevance, and timeframes. |

33

## INFORMATION VALUE CHAIN HOW TO DESIGN

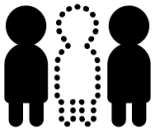
1. Keep **business mind set** when design data architecuture
2. **Cycle** of data for business and business create data
3. The chain **links stages** in the DIKW pyramid
4. The chain link **Business Concepts** with **Data Items**
5. Some keywords to learn: Data Catalog, Glossary, Term, Metadata, Data Entity, Measure, Dimension, KPIs, Data Domain, Subject Area, Data Quality




34

## INFORMATION VALUE CHAIN HOW TO AVOID FAILURE

1. Plan and **Spend enough resource** for Information Value Chain
2. Practice habit: “Only **use** data **when understanding** it”
3. Data without **context** has no meaning
4. Beautiful number is a trap: you need to be known how it was calculated



35

## Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path


36

# Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path


37

## DATA ARCHITECT AS A PROFESSIONAL

The data architect **defines** a **set of rules, policies, standards, and models** that govern the type of data that's used and collected in the organization database. They **design, create, and manage the data architecture** in an organization. A data architect **develops data models and data lake** designs to capture business's key performance indicators (KPIs) and **enable data transformation**. They **ensure consistent data performance** and **data quality** across the organization.

*Source: Solutions Architect's Handbook*

Data architects **create blueprints for data management systems**. After assessing a company's potential data sources (internal and external), **architects design** a plan to integrate, centralize, protect and maintain them. This **allows employees to access critical information** in the right place, at the right time.

41

*Source: http://www.mastersindatascience.org/careers/data-architect/*

38

## DATA ARCHITECT RESPONSIBILITY

1. Selection of database technology
2. Data warehousing for data analysis and BI tools
3. Data lake as the centralized datastore
4. Design Relational and Dimensional data models
5. Develop data models for database structures
6. Machine learning tools
7. Data security and encryption
8. Data compliance
9. Collaborate with IT teams and management to devise a data strategy that addresses industry requirements
10. Build an inventory of data needed to implement the architecture
11. Research new opportunities for data acquisition
12. Identify and evaluate current data management technologies
13. Create a fluid, end-to-end vision for how data will flow through an organization
14. Design, document, construct and deploy database architectures and applications (e.g. large relational databases)
15. Integrate technical functionality (e.g. scalability, security, performance, data recovery, reliability, etc.)
16. Implement measures to ensure data accuracy and accessibility
17. Constantly monitor, refine and report on the performance of data management systems
18. Meld new systems with existing warehouse structures
19. Produce and enforce database development standards
20. Maintain a corporate repository of all data architecture artifacts and procedures

42

39

## Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path


40

## KNOWLEDGE, SKILLS AND EXPERIENCES

**Technical Skills** for Data Architects

- Application server software (e.g. Oracle)
- Database management system software (e.g. Microsoft SQL Server)
- User interface and query software (e.g. IBM DB2)
- Enterprise application integration software (e.g. XML)
- Development environment software
- Backup/archival software
- Agile methodologies and ERP implementation
- Predictive modeling, NLP and text analysis
- Data modeling tools (e.g. ERWin, Enterprise Architect and Visio)
- Data mining
- UML
- ETL tools
- Python, C/C++ Java, Perl
- UNIX, Linux, Solaris and MS Windows
- Hadoop and NoSQL databases
- Machine learning
- Data visualization

*Source: http://www.mastersindatascience.org/careers/data-architect/*



## KNOWLEDGE, SKILLS AND EXPERIENCES

**Business Skills for Data Architects**

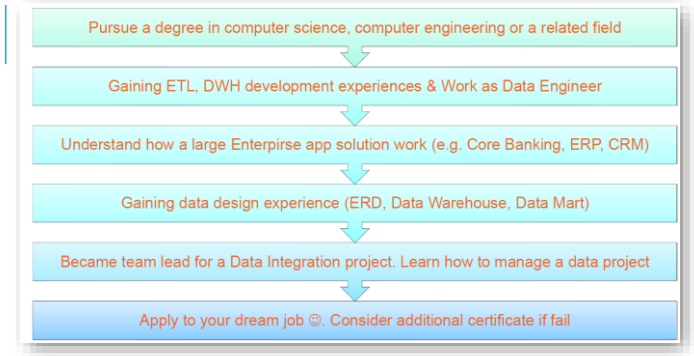
- **Analytical Problem-Solving**: Approaching high-level data challenges with a clear eye on what is important; employing the right approach/methods to make the maximum use of time and human resources.
- **Effective Communication**: Carefully listening to management, data analysts and relevant staff to come up with the best data design; explaining complex concepts to non-technical colleagues.
- **Expert Management**: Effectively directing and advising a team of data modelers, data engineers, database administrators and junior architects.
- **Industry Knowledge**: Understanding the way your chosen industry functions and how data are collected, analyzed and utilized; maintaining flexibility in the face of big data developments.

*Source: http://www.mastersindatascience.org/careers/data-architect/*

# Agenda

1. **Data in Banking**
   - Highlight main data source of a bank
   - Challenge when working with banking data
   - Why Banks need Data Architecture
2. **Data Architecture components**
   - Data Delivery Architecture
   - Enterprise Data Model
   - Information Value Chain
3. **How to become a Data Architect**
   - Data Architect as a professional
   - Knowledge, skills, experience
   - Sample career path


43




44

## SALARY BENCHMARK (GLASSDOOR & PAYSCALE)

| Jobs | Data Architect | Data Analysts | Data Scientist | BI Analyst |
|---|---|---|---|---|
| Entry level | | $83,750 | | |
| Standard | $123,680 | $100,250 | $118,370 | $80,154 |
| Expert | $102,000-$175,000 | $118,750-$142,500 | $147,000-$200,000 | $101,000 |

45

## BONUS SLIDES: A DAY LIFE OF A DATA ARCHITECT

Q: Ưu và nhược điểm của công việc này?

- Khởi đầu mỗi ngày với một bàn làm việc sạch sẽ. Bạn phục vụ tất cả mọi người trong doanh nghiệp không chỉ IT. Yêu cầu sẽ đến đầy bàn chỉ trong buổi chiều. Chúng đều được đánh dấu to đùng **"Ngay và Luôn"**. Chẳng có chỗ cho những "**Ý tưởng lớn**" bạn cần trả lời chúng.
- Vấn đề là rất nhiều câu hỏi tương tự nhau cứ đến với bạn hàng ngày. Việc này không thể kéo dài mãi. Bạn cần có "**Thiết kế Kiến trúc"** để phản hồi nhanh hơn với những yêu cầu lặp đi lặp lại.
- Công việc của bạn bị giằng xé giữa một bên là yêu cầu nhanh và đơn giản một bên là kế hoạch dài hạn và chuẩn chỉnh. Bạn cần thời gian.

46

## BONUS SLIDES: A DAY LIFE OF A DATA ARCHITECT

Q: Tầm ảnh hưởng của DA ở IBM?

- Nhu cầu về dữ liệu càng ngày càng lớn với nhiều phương thức khai thác dữ liệu khác nhau. Sự thật là nó sẽ không diễn ra tốt đẹp như mong đợi. Sẽ có những thất vọng và không hài long về kết quả ban đầu nhưng một ngày nào đó, có thể là hai năm nữa bạn sẽ có một tổ chức hoạt động nhanh hơn, trả lời nhiều câu hỏi hơn về doanh nghiệp, khách hàng, hiểu rõ hơn cách để làm mọi thứ nhanh hơn và chính xác hơn hiện tại.
- Nhiều dữ liệu không làm chúng ta "thông minh" hơn. Nếu không biết cách tiêu dùng chúng, dữ liệu sẽ làm chúng ta trở nên kém hiểu biết. VD khi ta có nhiều dữ liệu hơn nhưng lại không thể xử lý thì vô hình chung mức độ hiểu biết của bạn đã giảm đi. Dữ liệu nếu không được tổ chức và cung cấp tốt, chúng sẽ trở nên rắm rối, khó hiểu


Waiting to do the second coat was the worst part. It was like watching paint dry.

50

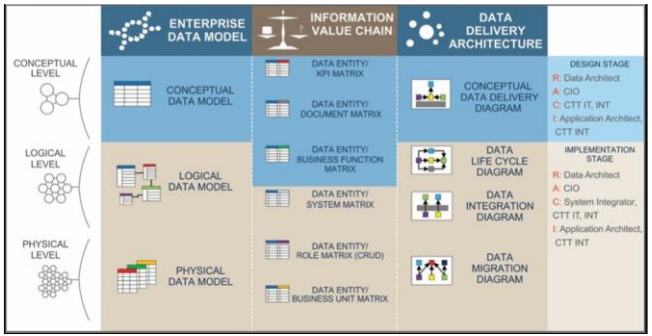
47

## BONUS SLIDES: A DAY LIFE OF A DATA ARCHITECT

Q: Lời khuyên cho các bạn sinh viên muốn trở thành Data Architecture?

- Lời khuyên tốt nhất là hãy trở thành một chuyên gia. Hãy trở thành người giỏi nhất trong lĩnh vực bạn chọn cho dù là lĩnh vực gì: kế toán, quản trị dữ liệu, lập trình. 5 năm nữa công việc của bạn sẽ hoàn toàn khác so với hiện tại.
- Và bạn cần học kỹ năng mới lại từ đầu và trở thành chuyên gia trong lĩnh vực mới đó. Với mỗi vài năm bạn cần làm việc đó một lần.
- Khả năng chuyển đổi công việc theo xu hướng mới cho dù đó là gì là vô cùng quan trọng. Nếu bạn làm được bạn sẽ giúp được người khác nhiều hơn bằng cách trở thành chuyên gia mà mọi người đều tin tưởng.


48


ENTERPRISE DATA MODEL
INFORMATION VALUE CHAIN
DATA DELIVERY ARCHITECTURE
CONCEPTUAL LEVEL
LOGICAL LEVEL
PHYSICAL LEVEL
CONCEPTUAL DATA MODEL
LOGICAL DATA MODEL
PHYSICAL DATA MODEL
DATA ENTITY/ KPI MATRIX
DATA ENTITY/ DOCUMENT MATRIX
DATA ENTITY/ BUSINESS FUNCTION MATRIX
DATA ENTITY/ SYSTEM MATRIX
DATA ENTITY/ ROLE MATRIX (CRUD)
DATA ENTITY/ BUSINESS UNIT MATRIX
CONCEPTUAL DATA DELIVERY DIAGRAM
DATA LIFE CYCLE DIAGRAM
DATA INTEGRATION DIAGRAM
DATA MIGRATION DIAGRAM
DESIGN STAGE
R: Data Architect
A: CIO
C: CTT IT, INT
I: Application Architect, CTT INT
IMPLEMENTATION STAGE
R: Data Architect
A: CIO
C: System Integrator, CTT IT, INT
I: Application Architect, CTT INT

49