



---

# Data Warehouse & Business Intelligence Fundamentals

---

Todor Kichukov

Faculty of Mathematics and Informatics

Sofia University

2022



# Todor Kichukov

## DW/BI Solution Architect

---

### BACKGROUND

- IT expert with 25+ years professional experience including 8 years at Oracle, 6 years experience in banking and 11+ years in IT companies
- 20+ years of experience in Data Warehousing, Business Intelligence, and Dimensional Modeling
- 7+ years of experience in Data Vault 2.0 Modeling

### ROLES

- Enterprise Architect
- Solution Architect, Solution Designer
- Practice Manager, Project Manager, Team Leader
- Principal / Senior Consultant
- Business Analyst
- Developer
- Trainer

### EXPERTISE

- Data Governance, Enterprise Architecture
- DW / BI solutions – Solution Architecture, Data Modeling, Implementation, Training
- Best practices of Kimball, Inmon, Linstedt

### MAIN ACHIEVEMENTS

- **Postbank, Bulgaria** – Enterprise Architecture, Data Architecture, Data Governance (ongoing)
- **Raiffeisenbank, Bulgaria** – Data / Application Enterprise Architecture, Data Governance, Data Vault solution, DWH Optimization;
- **Société Générale Expressbank, Bulgaria** – DW/BI solution architecture, data modeling, implementation + RDM (Reference Data Management) + custom-built FTP (Funds Transfer Pricing) and PM (Profitability Management) as solution architect;
- **OTP DSK Bank, Bulgaria** - DW solution architecture, data modeling, implementation + custom-built FTP / PM - project manager and solution architect;
- **UnicreditBulbank, Bulgaria** - DW solution architecture, data modeling, implementation + custom-built FTP / PM - project manager and solution architect
- **Isbank, Turkey** – prebuilt FTP/PM solution implementation;
- **Botswana Examinations Council (BEC)** – DW expert services;
- **Vipnet, Croatia** – ODS solution;
- **OTE, Greece** – Data Mart implementation

### NOW

- [www.bipartner.biz](http://www.bipartner.biz), [todor.kichukov@bipartner.biz](mailto:todor.kichukov@bipartner.biz)

# **Data Warehouse & Business Intelligence Fundamentals**

## **Course Scope**

- DW Concept
- DW Architecture
- DW Data Modeling
- Data Integration
- Gathering and Analyzing Requirements
- Business Intelligence
- Deployment, Support and Maintenance

## Course Setup

- Semester (15W)
  - Curricular Employment (45h)
    - Lectures + Exercises (Project) (30h+15h)
    - Wednesday, 18:15, room 306
    - All lectures available on PDF
  - Extracurricular Employment (90h)
    - Project (30h) – Team working on DW implementation
    - Individual work in a library or with internet resources (30h)
    - Exam preparation (30h)
- Exam
- Final Grade = Project 50% + Exam 50%
- [Sofia University - Data Warehouse & Business Intelligence Fundamentals 2022](#)  
[www.facebook.com/groups/SUDWBI2022/](http://www.facebook.com/groups/SUDWBI2022/)

# Literature

## Main

- Kimball, Ralph; Ross, Margy. The Kimball Group Reader: **Relentlessly Practical Tools for Data Warehousing and Business Intelligence** Remastered Collection, 2016
- Kimball, Ralph; Ross, Margy. **The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling**, 3<sup>rd</sup> Edition, 2013
- Inmon, W.H. **Building the Data Warehouse**, Fourth Edition, 2005
- Linstedt, Dan; Olschimke, Michael. **Building a Scalable Data Warehouse with Data Vault 2.0**, 2015

## Additional

- Kimball, Ralph; Ross, Margy. **The Data Warehouse Lifecycle Toolkit: Practical Techniques for Building Data Warehouse and Business Intelligence Systems**, 2nd Edition, 2012
- Linstedt, Dan. **The Official Data Vault Standards Document (Version 1.0)**, 2012

# Course Approach



*"Образованието не е просто учене на факти, а трениране на ума да мисли."*

*"Интелектът е това, което остава, когато забравим всичко, което сме научили в училище."*

*Алберт Айнщайн*

- The choice - why we are all here
- Expectations, Challenges
- Discussion, Workshop *vs* Lecture
- Theory *vs* Practice
- Team *vs* Individual

# **Data Warehouse Concept**

- Data Warehouse Definition
- How Data Warehouses Appeared
- OLTP vs OLAP
- Main DW Principles
- Business vs IT
- DW Camps

# Data Warehouse Definition

---

- **Bill Inmon**

- A data warehouse is a **subject-oriented, integrated, nonvolatile, and time-variant** collection of data in support of management's decisions.

- **Ralph Kimball**

- The mission of the data warehouse is to publish the organization's data assets to most effectively support decision making.
- The whole point of the data warehouse is to deliver data in a simple, actionable format for the benefit of end users and their analytic applications.
- A data warehouse is a **system that extracts, cleans, conforms, and delivers source data into a dimensional data store** and then supports and implements **querying and analysis for the purpose of decision making**.

- **Oracle**

- A data warehouse is a database designed to enable business intelligence activities: it exists to help users understand and enhance their organization's performance. **It is designed for query and analysis rather than for transaction processing.** Data warehouses separate analysis workload from transaction workload and enable an organization to consolidate data from several sources.



# What the Data Warehouse Is Not – Top 5

---

1. **A product.** Contrary to many vendor claims, you cannot buy a data warehouse. A data warehouse includes system and requirements analysis, data manipulation and cleansing, data movement, and finally dimensional modeling and data access. No single product can achieve all of the tasks involved in building a data warehouse.
2. **A language.** One cannot learn to *code* a data warehouse in the way you learn to implement XML, SQL, VB, or any other programming language. The data warehouse is composed of several components, each likely to require one or more programming or data-specification languages.
3. **A project.** A properly deployed data warehouse consists of many projects (and phases of projects). Any attempt to deploy a data warehouse as a single project will almost certainly fail. Successful data warehouses plan at the enterprise level yet deploy manageable dimensional data marts. Each data mart is typically considered a separate project with its own timeline and budget. A crucial factor is that each data mart contains conformed dimensions and standardized facts so that each integrates into a single cohesive unit—the enterprise data warehouse. The enterprise data warehouse evolves and grows as each data mart project is completed. A better way to think of a data warehouse is as a **process**, not as a project.
4. **A data model.** A data model alone does not make a data warehouse. Recall that the data warehouse is a comprehensive process that, by definition, must include the ETL process. After all, without data, even the best-designed data model is useless.
5. **A copy of your transaction system.** A common mistake is to believe copying your operational system into a separate reporting system creates a data warehouse. Just as the data model alone does not create a data warehouse, neither does executing the data movement process without restructuring the data store.

# How Data Warehouses Appeared


---

- 1960-1980, terms Data Warehouse, Data Mart, Dimensions, Facts mentioned
- 1985, Barry Devlin, An Architecture for a Business and Information System
- 1992, Bill Inmon, Building the Data Warehouse
- 1992, Wal-Mart DW, the first commercial DW that reached 1TB
- 1996, Ralph Kimball, The Data Warehouse Toolkit
- 1997, Ralph Kimball, A Dimensional Modeling Manifesto
- 2000, Dan Linstedt, Data Vault 1.0 modeling released
- 2013, Dan Linstedt, Data Vault 2.0
  
- Now & Future – Hybrids, Real-Time, Big Data

# OLTP vs OLAP

	OLTP	OLAP
Typical operations	Operational work - Withdrawing money from customer account	Analysis - Sales amount for last month by region, product and customer type
Workload, data modifications	Transactions, find specific record, insert, update, delete	Querying (select) across huge data volumes, aggregating, group by; ad-hoc or predefined analyses; insert/update/delete via ETL
Data integration	No integration, every OLTP is separate application	DW integrate data from many sources.
Historical data	Keeps only current state; history in archive; exception – transactions	Keeps history of all changes using different mechanisms
Data modeling	3NF, normalized	Denormalized, Multidimensional, 3NF
Index usage	Few, b-tree	Many, bitmap, b-tree
Joins	Many	Some
Derived data	Rare	Common

# Main Data Warehouse Principles



The most important slide!

Bill Inmon

- **Subject-oriented** – built around subjects (like customer, product, order) or subject area (aka business process - like sales, campaigns, accounting)
- **Integrated** – all identical entities from different sources go to one DW entity in consistent format
- **Nonvolatile** - once entered into the data warehouse, specific data should not change – same report all the time
- **Time-variant** – data content change over time is important – trends, patterns

Todor Kichukov

- **Atomic** – keep the most detail data!
- **Enterprise level** – there isn't a successful DW on any other lower level!
- **Designed for querying and analyses** – do not forget why DW appeared!

# Business vs IT

---

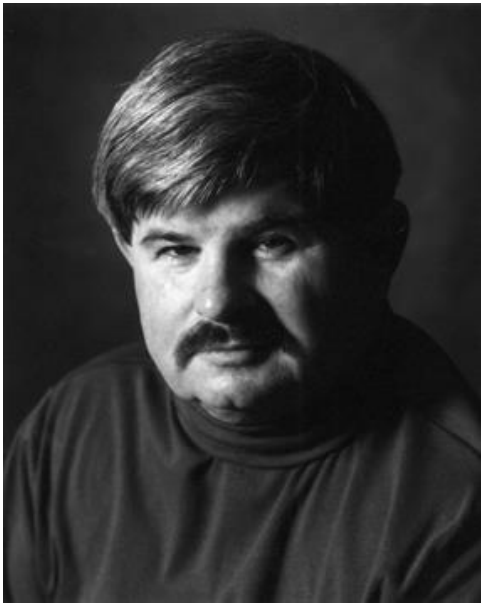
- DW/BI is a business solution, not IT initiative
- The DW/BI solution serves the needs of the business
- The DW/BI solution must be build by Business / IT / Vendor together

# Data Warehouse Camps / Religions / Wars

---

*Architecture*  
*Data Modeling*  
*Implementation*

**Bill Inmon**



**Dan Linstedt**



**Ralph Kimball**



# Grand Slam Achievement

---

***2019, Hannover  
WWDVC***



***with Bill Inmon***



***with Dan Linstedt***

***2008, Stockholm  
Dimensional Modeling In Depth***



***with Ralph Kimball***