The background of the slide is a light gray gradient with several realistic water droplets of various sizes scattered across it. The droplets have highlights and shadows, giving them a three-dimensional appearance. The main title is centered in the upper half of the slide.

# ENSEMBLE MODELING AND DATA VAULT

DATA ACTION NETWORK, 2022-12-09

The background of the slide is a light gray gradient with several realistic water droplets of various sizes scattered across it. The droplets have highlights and shadows, giving them a three-dimensional appearance. The main title is centered in the middle of the slide.

# WHY ENSEMBLE MODELING?

IT'S ABOUT TIME

The background of the slide is a light gray gradient. It is decorated with several realistic water droplets of various sizes and shapes, scattered across the top and right sides. The droplets have highlights and shadows, giving them a three-dimensional appearance.

# DATA VAULT DATA MODELING

A SHORT INTRODUCTION

# DATA VAULT

- WHO?
- WHAT?
- WHEN?
- WHERE?
- HOW?
- WHY?

# WHO?

- **DAN LINSTEDT**
- HANS HULTGREN
- KENT GRAZIANO
- MICHAEL OLSCHIMKE
- ROELANT VOS
- JOHN GILES
- PATRICK CUBA

# WHAT?

- DAN LINSTEDT'S DEFINITION

- DETAIL ORIENTED,
- HISTORICAL TRACKING AND
- UNIQUELY LINKED SET OF
- NORMALIZED TABLES THAT
- SUPPORT ONE OR MORE FUNCTIONAL AREAS OF BUSINESS

- UNIFIED DECOMPOSITION

- SEPARATION OF KEYS, RELATIONSHIPS AND DESCRIPTIVE ATTRIBUTES
  - BUSINESS KEY AS INTEGRATION POINT
- OPTIMIZED FOR DATA WAREHOUSING
    - CORE LAYER OF A 3-LAYER (INMON) ARCHITECTURE
    - SINGLE VERSION OF FACTS
    - 100 % OF THE DATA 100 % OF THE TIME

# WHEN?

1990'S DEVELOPMENT BY DAN LINSTEDT

2000 FIRST PUBLICATION

2007 SUPPORT FROM BILL INMON

2011 SUPER CHARGE YOUR DATA WAREHOUSE (LINSTEDT)

2012 MODELING THE AGILE DATA WAREHOUSE WITH DATA VAULT (HULTGREN)

2013 PRESENTATION OF DATA VAULT 2.0

2015 BUILDING A SCALABLE DATA WAREHOUSE WITH DATA VAULT 2.0 (LINSTEDT/OLSCHIMKE)

2019 THE ELEPHANT IN THE FRIDGE (JOHN GILES)

2020 THE DATA VAULT GURU (PATRICK CUBA)

# WHERE?

## GEOGRAPHY

- USA
- NETHERLANDS
- AUSTRALIA

## INDUSTRIES

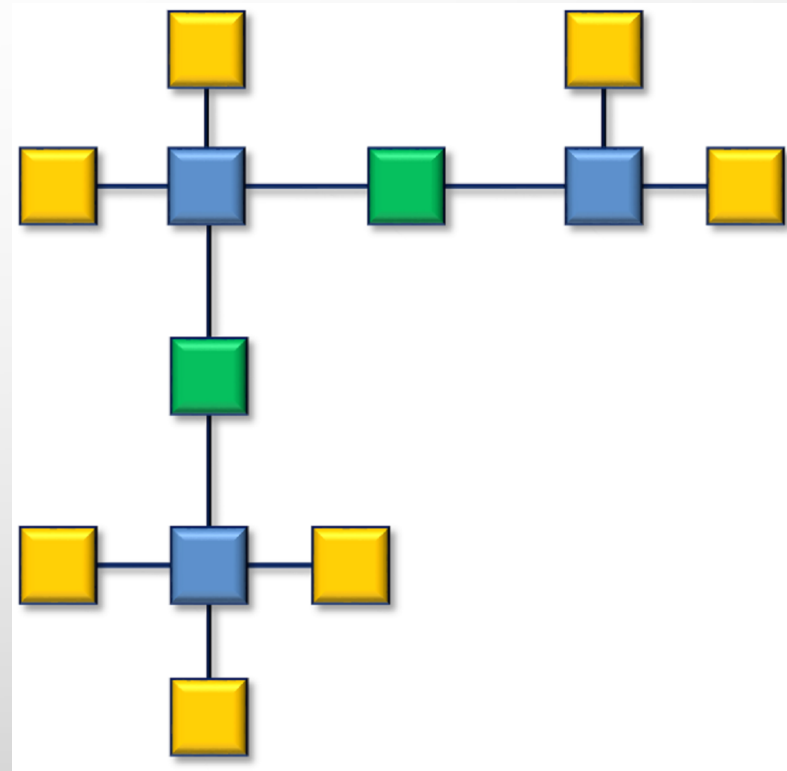
- BANKING & INSURANCE
  - ALLIANZ, COMMONWEALTH BANK, ING
- RETAIL
  - OTTO, MYTOYS, HAGEBAU
- MANUFACTURING
  - VOLKSWAGEN, LOCKHEED, MICRON
- NSA



# HOW?

## THREE BASIC TABLE TYPES

- **HUBS** FOR BUSINESS KEYS
- **LINKS** FOR RELATIONSHIPS
- **SATELLITES** FOR CONTEXT



# HOW?

## **HUB**

- CORE BUSINESS CONCEPT
- UNIQUE LIST OF BUSINESS KEYS
- INTEGRATION POINT FOR DATA FROM DIFFERENT SOURCES

## EXAMPLES

- BOOK – ISBN
- CAR – VIN
- CUSTOMER – CUSTOMER ID
- BILL – BILL NUMBER
- SHIP – IMO NUMBER

# HOW?

## LINK

- NATURAL BUSINESS RELATIONSHIP
- RELATIONSHIP BETWEEN HUBS
  - UNIT OF WORK
  - HIERARCHY
  - IDENTITY (SAME-AS LINK)
- ALWAYS MANY-TO-MANY

## EXAMPLES

- EMPLOYEE *REPORTS TO* MANAGER
- AUTHOR *WRITES* BOOK
- COMPONENT *IS PART OF* CAR MODEL
- DOCTOR *TREATS* PATIENT *IN* HOSPITAL
- REPRESENTATIVE *MAKES* SPEECH *ON* LAW

# HOW?

## SATELLITE

- CONTEXT AND HISTORY
- ATTRIBUTES OF A HUB OR LINK AT A SPECIFIC POINT IN TIME
- MULTIPLE SATELLITES PER HUB POSSIBLE
  - SOURCE SYSTEM
  - RATE OF CHANGE
  - DATA PROTECTION RULES (GDPR!)

## EXAMPLES

- *CUSTOMER*
  - PERSONAL DATA (NAME, DATE OF BIRTH)
  - CONTACT (ADDRESS, PHONE, E-MAIL)
  - DEMOGRAPHY (AGE, FAMILY, INCOME)
- *PRODUCT*
  - DESCRIPTIVE DATA (LOOK, PROPERTIES)
  - FINANCIAL DATA (LIST PRICE, COST)

# HOW?

## CONTROVERSIAL: **TRANSACTIONS**

- LINK WITH DEGENERATE FIELD(S)

*OR*

- ADDITIONAL TRANSACTION HUB

## CONTROVERSIAL: **LINK SATELLITES**

- YES, PLEASE

*OR*

- ONLY FOR EFFECTIVITY DATES

*OR*

- NO, THANKS

# WHY?

## STRENGTHS

- AUTOMATABILITY
- EXTENSIBILITY
- FLEXIBILITY
- AUDITABILITY
- PARALLELIZABILITY

## CHALLENGES

- HIGH NUMBER OF TABLES (AND JOINS)
- LIMITS OF AUTOMATABILITY
  - DETERMINING BUSINESS KEYS
  - PRESENTATION LAYER
- REDUNDANCIES AND INCONSISTENCIES
- PERFORMANCE

The background of the slide is a light gray gradient with several realistic water droplets of various sizes scattered across it. The droplets have highlights and shadows, giving them a three-dimensional appearance. The text is centered on the page.

# CINEMANIA

CASE STUDY

# CINEMANIA

- SUCCESSFUL MOVIE THEATER CHAIN
  - DIVERSE OFFERINGS
    - MULTIPLEX CINEMAS
    - 2D/3D
    - ENGLISH/ORIGINAL WITH SUBTITLES
    - DISCOUNT DAY, FAMILY SCREENINGS, ...
  - TICKETS AVAILABLE ONLINE, VIA PHONE OR AT BOX OFFICE
- POSSIBLE CORE BUSINESS CONCEPTS
    - CINEMA
    - AUDITORIUM
    - FILM
    - SCREENING
    - CUSTOMER
    - TICKET
    - EMPLOYEE



The background of the slide is a light gray gradient with several realistic water droplets of various sizes scattered across it. The droplets have highlights and shadows, giving them a three-dimensional appearance. The word "FEEDBACK" is centered in the middle of the slide in a large, bold, black, sans-serif font.

# FEEDBACK

CHRISTIAN.KAUL@OBAYSCH.NET

(OR IN DAN SLACK)