# Data Warehousing Reimagined:

## Fueling AI-powered BI at Scale

peter.aiken@anythingawesome.com +1.804.382.5957



### Peter Aiken, Ph.D.

- I've been doing this a long time
- My work is recognized as useful
- Associate Professor of IS (vcu.edu)
- DAMA International (dama.org)
  - IDEAL Affinity Group
  - GenAI in Data Affinity Group
- Anything Awesome (anythingawesome.com)
- Experienced w/ 500+ data management practices worldwide
- 13 books and dozens of articles
- Multi-year immersions
  - US DoD (DISA/Army/Marines/DLA)
  - Nokia
  - Deutsche Bank
  - Wells Fargo
  - Walmart
  - HUD
  - NGA …

**$1,500,000,000.00 USD**

# Program Overview

- Definitions
  - Data Warehousing
  - DM BoK guidance
  - Legacy to digital conversion (cloud) + AI
- Integration
  - Incorporate leading not bleeding edge
  - Requires an adaptive rather than a prescriptive approach
  - Emphasize engineering talent (Forward Deployed Engineer)
  - AI enables further structural formalization
- Preparation
  - Emphasis on storytelling first and visualization second
  - Analytics is both ubiquitous and not well understood
  - Keep improvements practically focused by strategy
- Best Practices
  - Cannot use what is not understood-understand what you have
  - PDCA (plan, do, check, act)
  - Iteratively refine/cull with respect to strategic direction
- Take Aways/References/Q&A

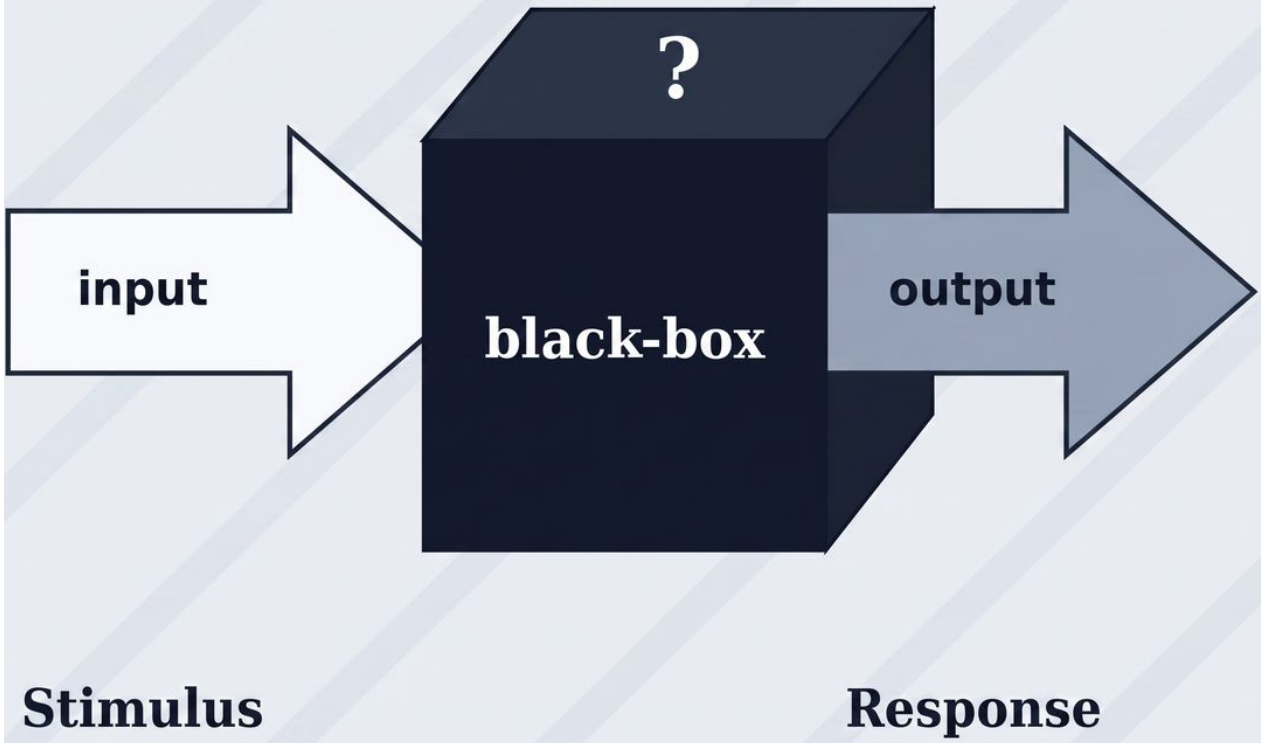

**Data Warehousing Reimagined:**

**Fueling AI-powered BI at Scale**

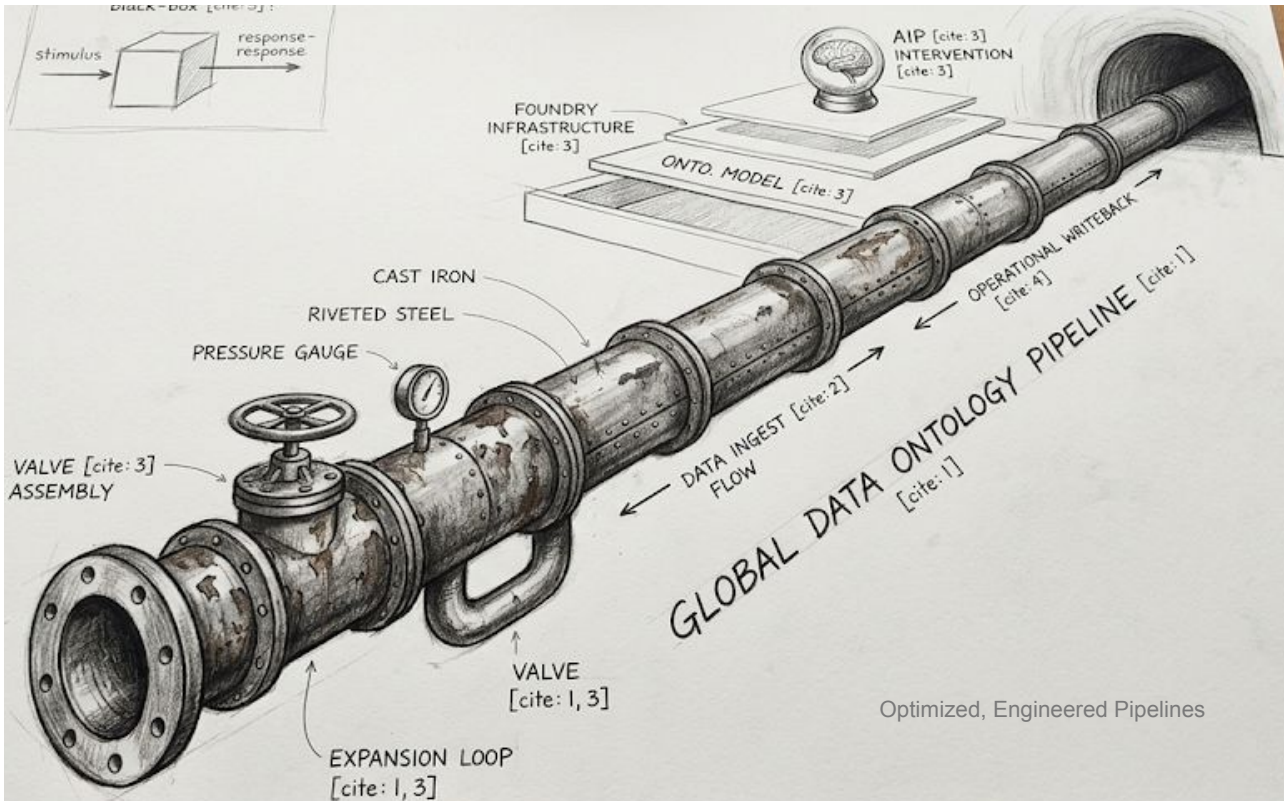
DATA-ED ONLINE
DATAVERSITY



## Data Management Practices are typically 'Technical'



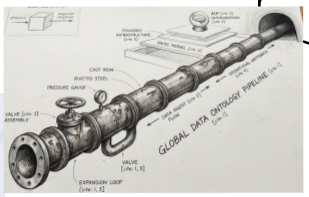
# Introducing Warehousing

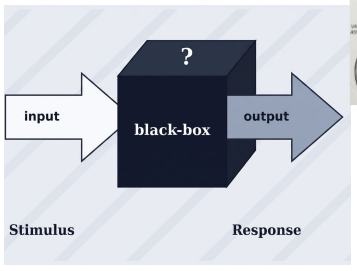





# Improving Warehousing

Data Management Practices

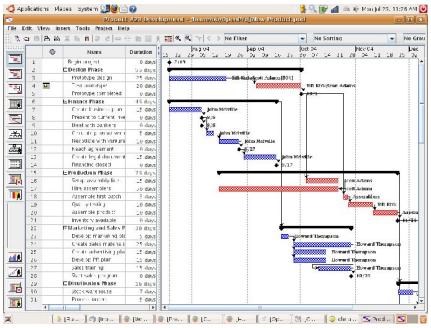



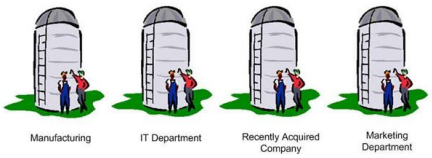



Optimized, Engineered Pipelines

# Data Warehousing is not a Project → Data Product

- Durable asset
  - An asset that has a usable life more than one year
- Reasonable project deliverables
  - 90 day increments
  - Data evolution is measured in years
- Data
  - Evolves - it is not created
  - Significantly more stable
- Readymade data architectural components
  - Prerequisite to agile development
- Only alternative is to create additional data siloes!








## Data Management Body of Knowledge (DM BoK V2)


11 Practice Areas

from *The DAMA Guide to the Data Management Body of Knowledge* 2E © 2017 by DAMA International

# Data Warehousing & Business Intelligence Management

**Definition**: Planning, implementation, and control processes to provide decision support data and support knowledge workers engaged in reporting, query and analysis.

**Goals**:
1. To support and enable effective business analysis and decision making by knowledge workers.
2. To build and maintain the environment / infrastructure to support business intelligence activity, specifically leveraging all the other data management functions to cost effectively deliver consistent integrated data for all BI activity.

**Inputs**:
- Business Drivers
- BI Data and Access Requirements
- Data Quality Requirements
- Data Security Requirements
- Data Architecture
- Technical Architecture
- Data Modeling Standards and Guidelines
- Transactional Data
- Master and Reference Data
- Industry and External Data

**Suppliers**:
- Executives and Managers
- Subject Matter Experts
- Data Governance Council
- Information Consumers (Internal and External)
- Data Producers
- Data Architects and Analysts

**Activities:**
1. **Understand Business Intelligence Information Needs (P)**
2. **Define and Maintain the DW / BI Architecture (P)**
3. **Implement Data Warehouses and Data Marts (D)**
4. **Implement BI Tools and User Interfaces (D)**
5. **Process Data for Business Intelligence (O)**
6. **Monitor and Tune Data Warehousing Processes (C)**
7. **Monitor and Tune BI Activity and Performance (C)**

**Participants**:
- Business Executives and Managers
- DM Execs and Other IT Mgmt
- BI Program Manager
- SMEs and Other Info Consumers
- Data Stewards
- Project Managers
- Data Architects and Analysts
- Data Integration (ETL) Specialists
- BI Specialists
- Database Administrators
- Data Security Administrators
- Data Quality Analysts

**Tools:**
- Database Management Systems
- Data Profiling Tools
- Data Integration Tools
- Data Cleansing Tools
- Business Intelligence Tools
- Analytic Applications
- Data Modeling Tools
- Performance Management Tools
- Meta-data Repository
- Data Quality Tools
- Data Security Tools

**Primary Deliverables**:
- DW/BI Architecture
- Data Warehouses
- Data Marts and OLAP Cubes
- Dashboards and Scorecards
- Analytic Applications
- File Extracts (for Data Mining/Stat. Tools)
- BI Tools and User Environments
- Data Quality Feedback Mechanism/Loop

**Consumers:**
- Knowledge Workers
- Managers and Executives
- External Customers and Systems
- Internal Customers and Systems
- Data Professionals
- Other IT Professionals

**Metrics:**
- Usage Metrics
- Customer/User Satisfaction
- Subject Area Coverage %s
- Response/Performance Metrics

**Activities: (P) – Planning (C) – Control (D) – Development (O) - Operational**




# Health Care Provider Data Warehouse

- 1.8 million members
- 1.4 million providers
- 800,000 providers no key
- 29% prov_ssn ≠ 9 digits
- 2.2% prov_number = 9 digits (required)
- 1 User
- $30 million

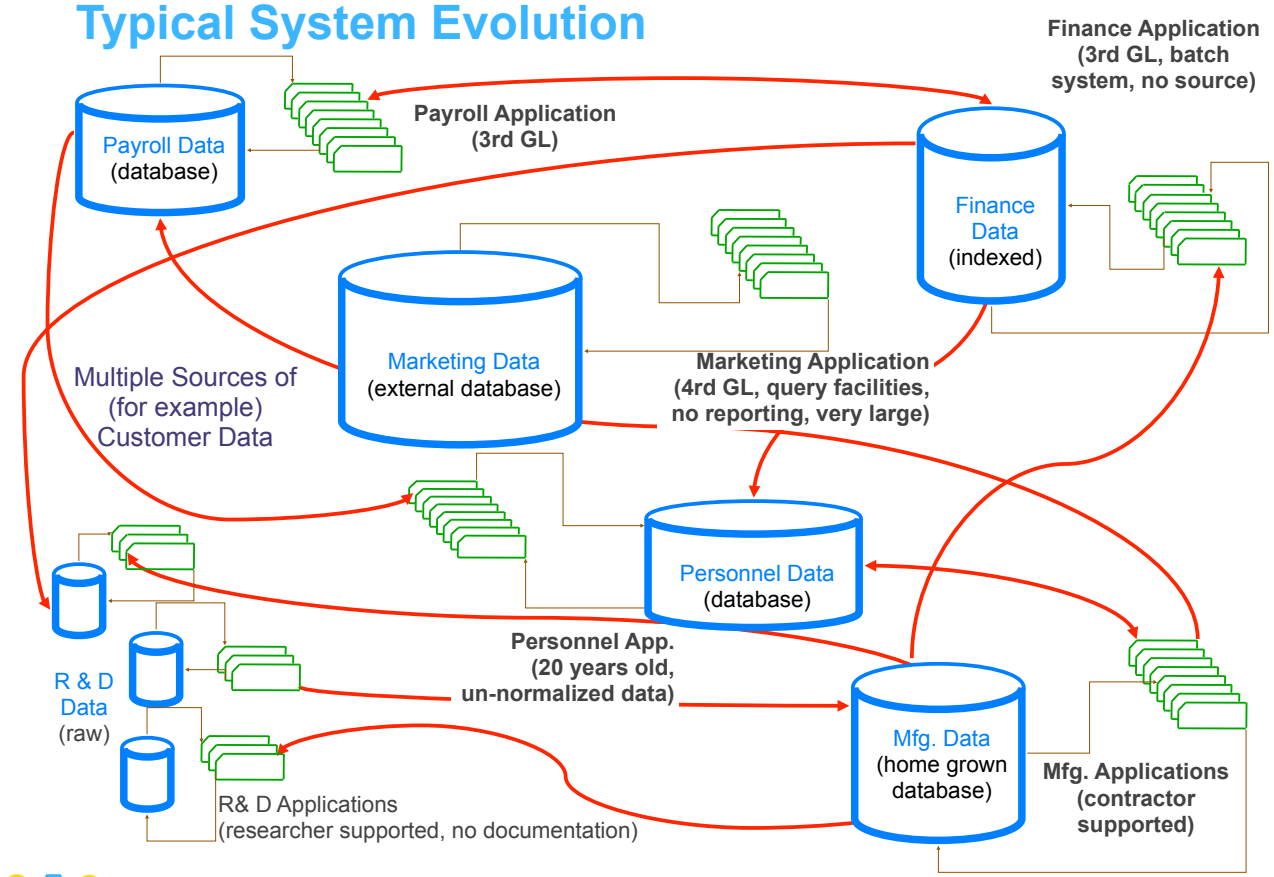

| | | |
|---|---|---|
| ,335 | $ | 4,812 |
| ,280 | $ | 55,445 |
| ,130 | $ | (118,059) |
| ,016 | $ | (18,016) |
| ,238 | $ | (70,786) |
| 4.39 | $ | (12.90) |
| per Visit | | |

# Defining Data Warehousing, BI/Analytics

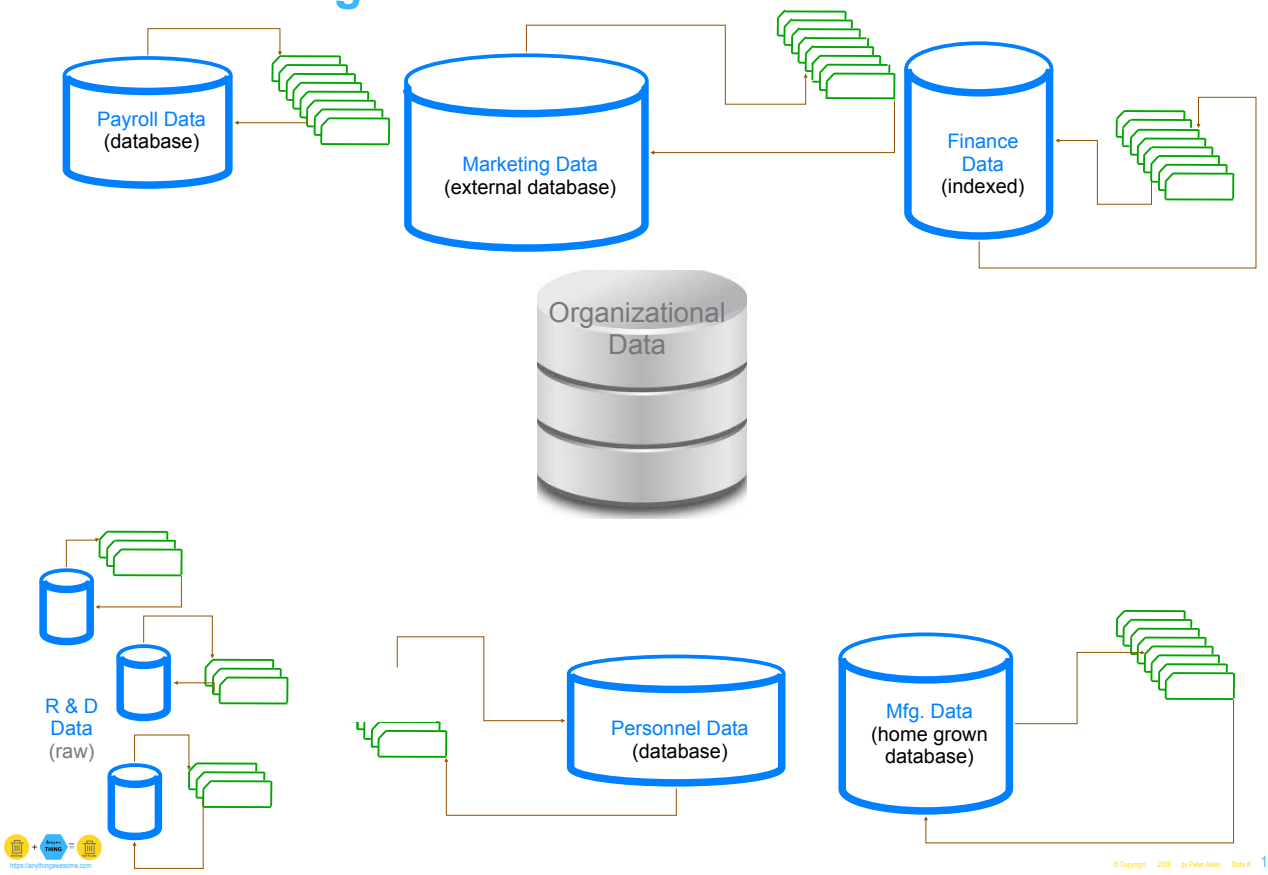

- Data Warehousing
  - A technology solution supporting … business capabilities such as: query, analysis, reporting and development of these capabilities
  - Analysis of information not previously integrated
  - Another, often new, set of organizational capabilities
- Business Intelligence (aka. decision support)
  - Dates at least to 1958
  - Support better business decision making
  - Technologies, applications and practices for the collection, integration, analysis, and presentation of business information
  - Understanding historical patterns in data to improve future performance
  - Use of mathematics in business
- Analytics (aka.) enterprise decision management, marketing analytics, predictive science, strategy science, credit risk analysis. fraud analytics - often based on computational modeling
- Reframing the question …
  - From: what data warehouse should we build?
  - To: how can data warehouse-based integration address challenges?



# Typical System Evolution



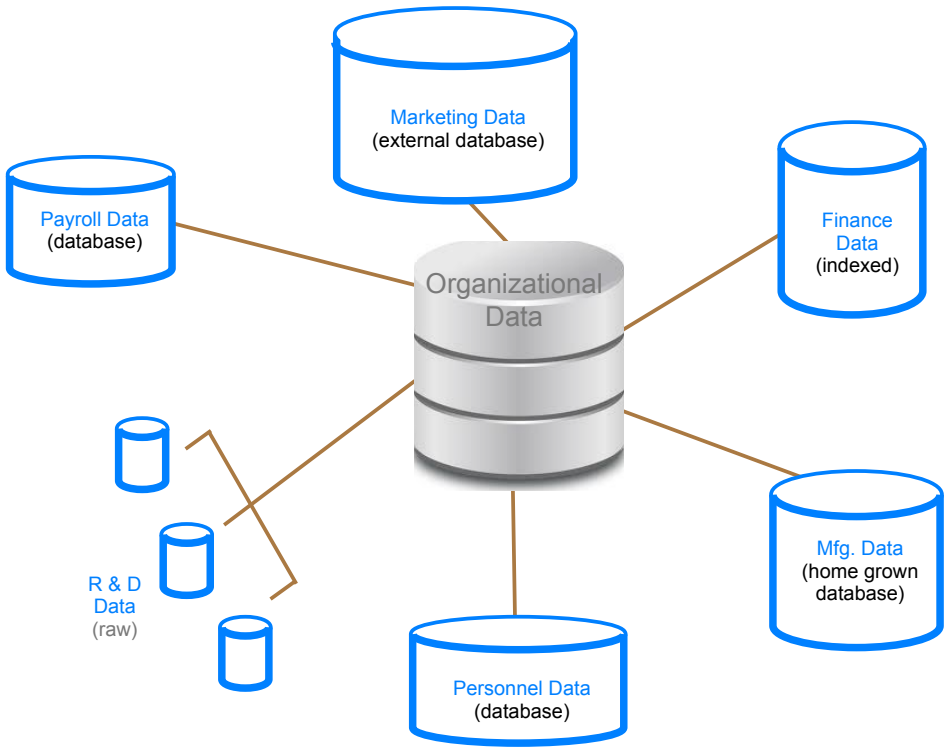
# ... Then Integrate





# ... Then Re-architect





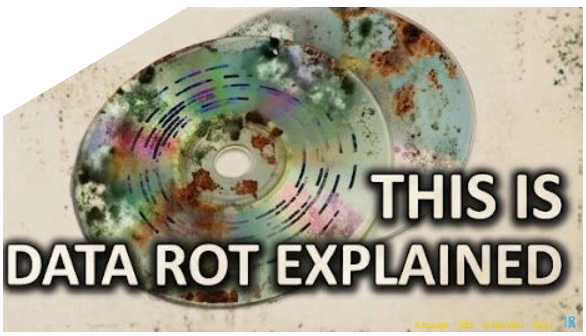
## send some of all to the cloud


Not Cloud | Cloud



## Organizing the Wheat Separated from the Chaff

- Better organized data increases in value
- Poor data management practices are costing organizations money/time/effort
- 80% of organizational data is **ROT**
  - **R**edundant
  - **O**bsolete
  - **T**rivial
- The question is which data to eliminate?
  - Most enterprise data is never analyzed



## … Fixing Data in the Cloud is Like Using A Glovebox



Without requisite training …



## Data in the cloud should have three attributes that data outside the cloud/warehouse should not have. It should be:

Cleaner

Smaller

Sharable-er

SHAREABLE

# Making —Cloud—— Successful

Problems with forklifting

1. no basis for decisions made
2. no inclusion of architecture/ engineering concepts
3. no idea that these concepts are missing from the process
4. 80% of organizational data is ROT

Data Branding

Transform

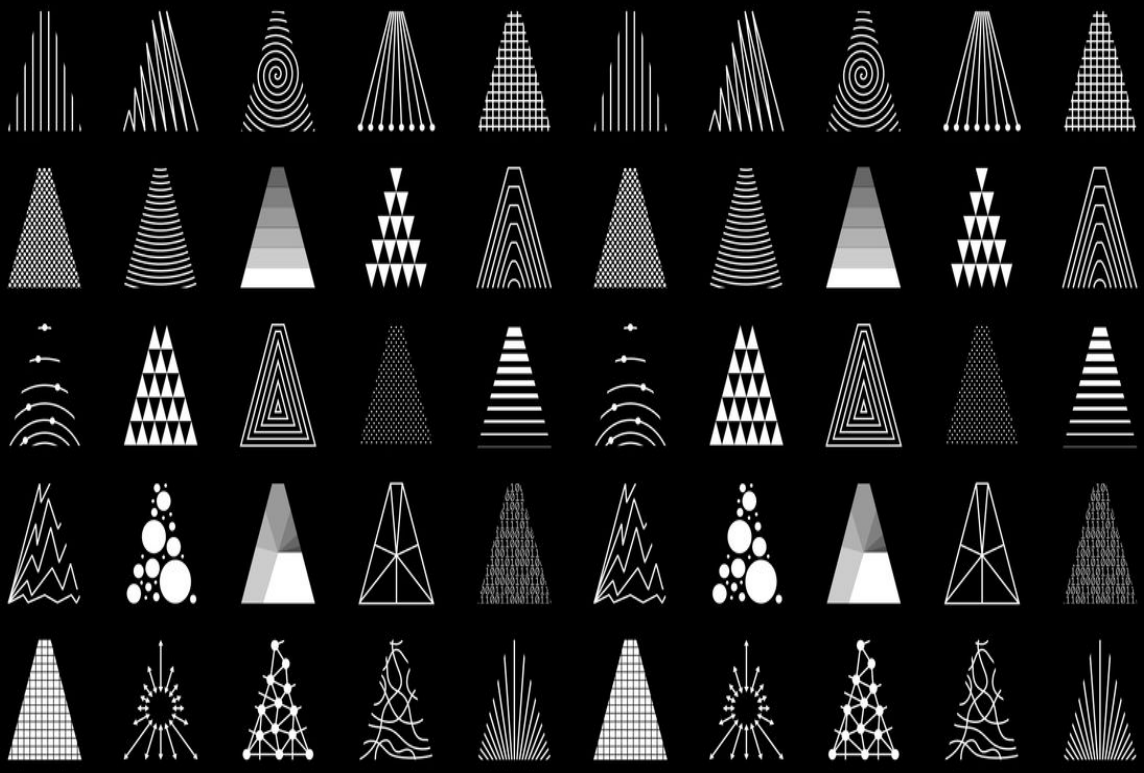
Less

Cleaner

More shareable ... data



# Data Branding



https://mindsparklemag.com/design/arcadia-data/


## Program Overview

- Definitions
  - Data Warehousing
  - DM BoK guidance
  - Legacy to digital conversion (cloud) + AI
- Integration
  - Incorporate leading not bleeding edge
  - Requires an adaptive rather than a prescriptive approach
  - Emphasize engineering talent (Forward Deployed Engineer)
  - AI enables further structural formalization
- Preparation
  - Emphasis on storytelling first and visualization second
  - Analytics is both ubiquitous and not well understood
  - Keep improvements practically focused by strategy
- Best Practices
  - Cannot use what is not understood-understand what you have
  - PDCA (plan, do, check, act)
  - Iteratively refine/cull with respect to strategic direction
- Take Aways/References/Q&A


**Data Warehousing Reimagined:**

**Fueling AI-powered BI at Scale**

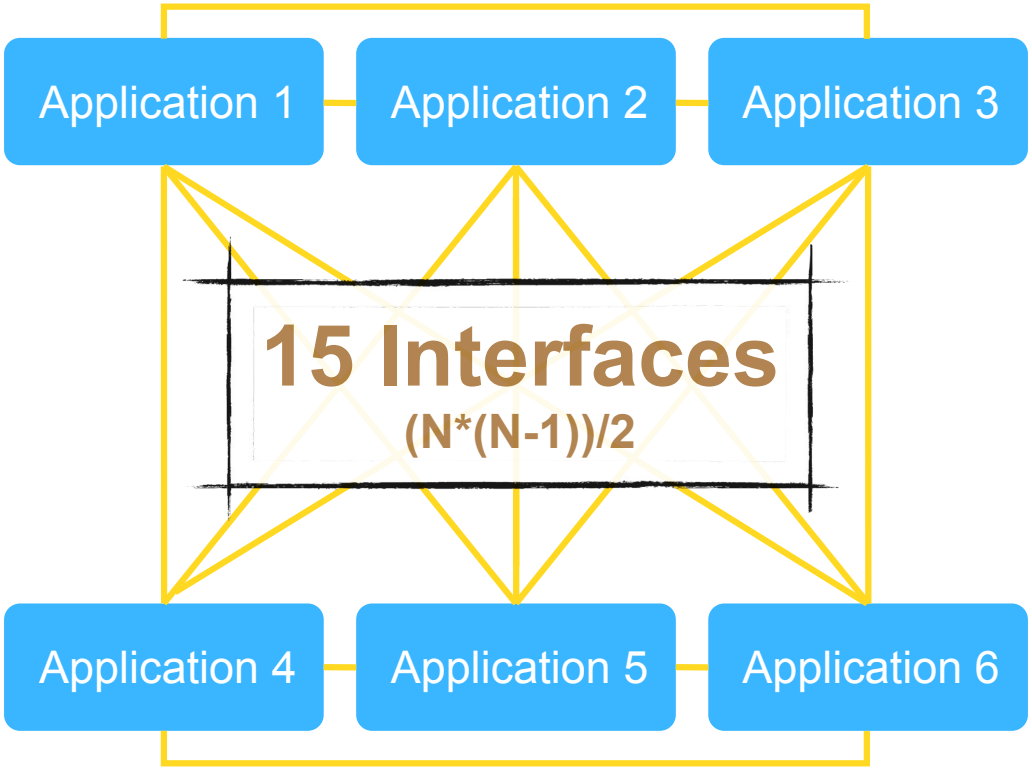
DATA-ED ONLINE
DATAVERSITY

## Two Basic Warehouse Purposes

- Integration
  - of disparate data sources for purpose of potential subsequent analyses
  - Most organization data is never analyzed
  - Same type of inputs as output
  - Downstream knowledge is incorporated upstream
- Preparation
  - Seen as the last mile (preparation) of the data before presentation as part of decision making activities
  - Closed ended activities
  - Final possible application of programmatic quality measures




## How many interfaces are required to solve this integration problem?

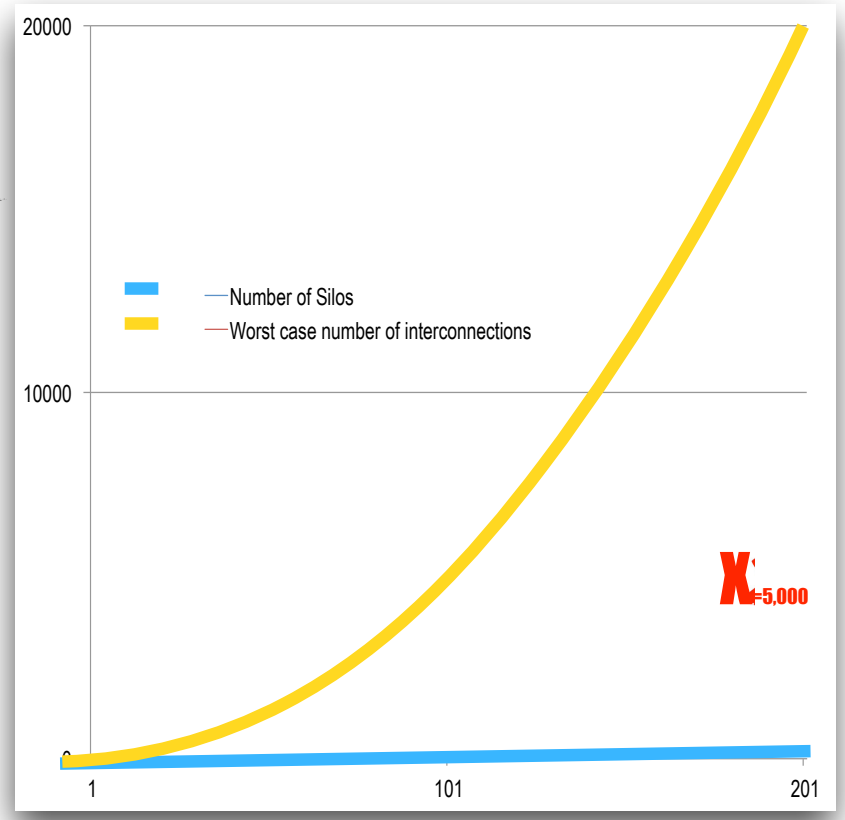



RBC: 200 applications - 4900 batch interfaces

# The rapidly increasing cost of complexity

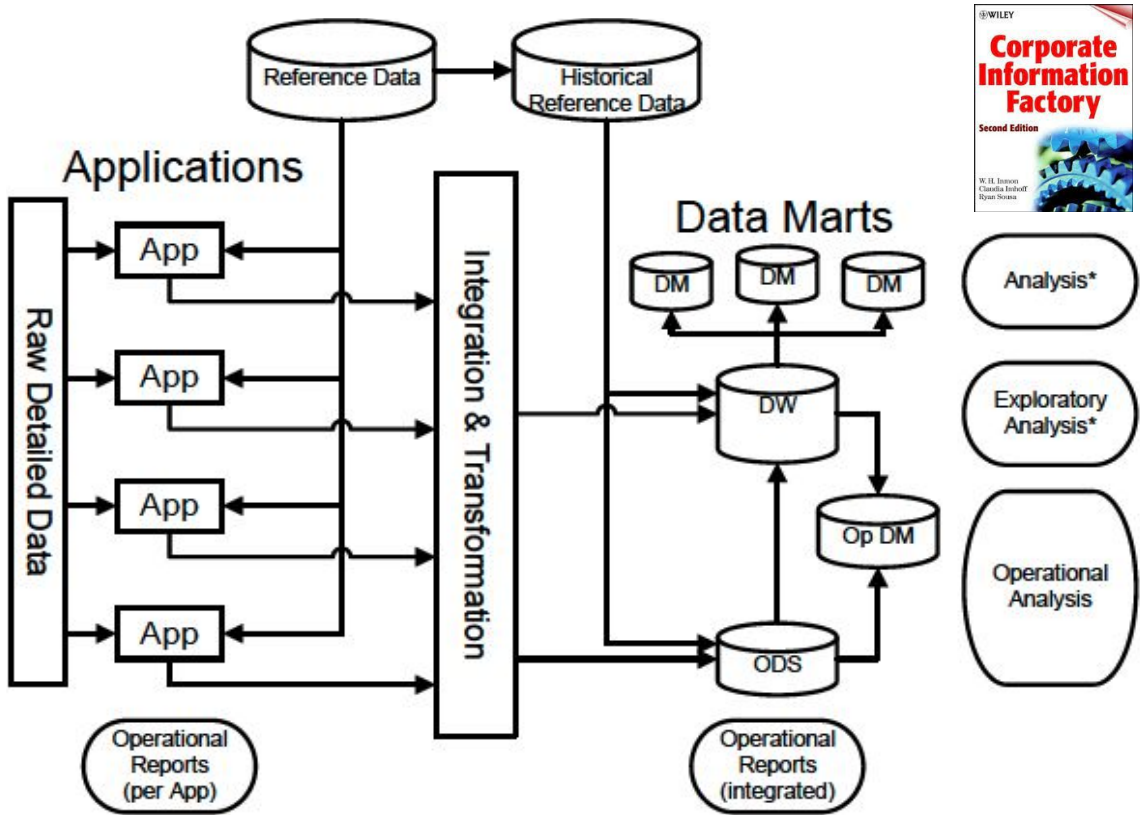
**Silos / Interconnections**

- 6 / 15
- 60 / 1,770
- 600 / 179,700
- 200 / 19,900
- 200 / 5,000 (actual)






# Corporate Information Factory Architecture

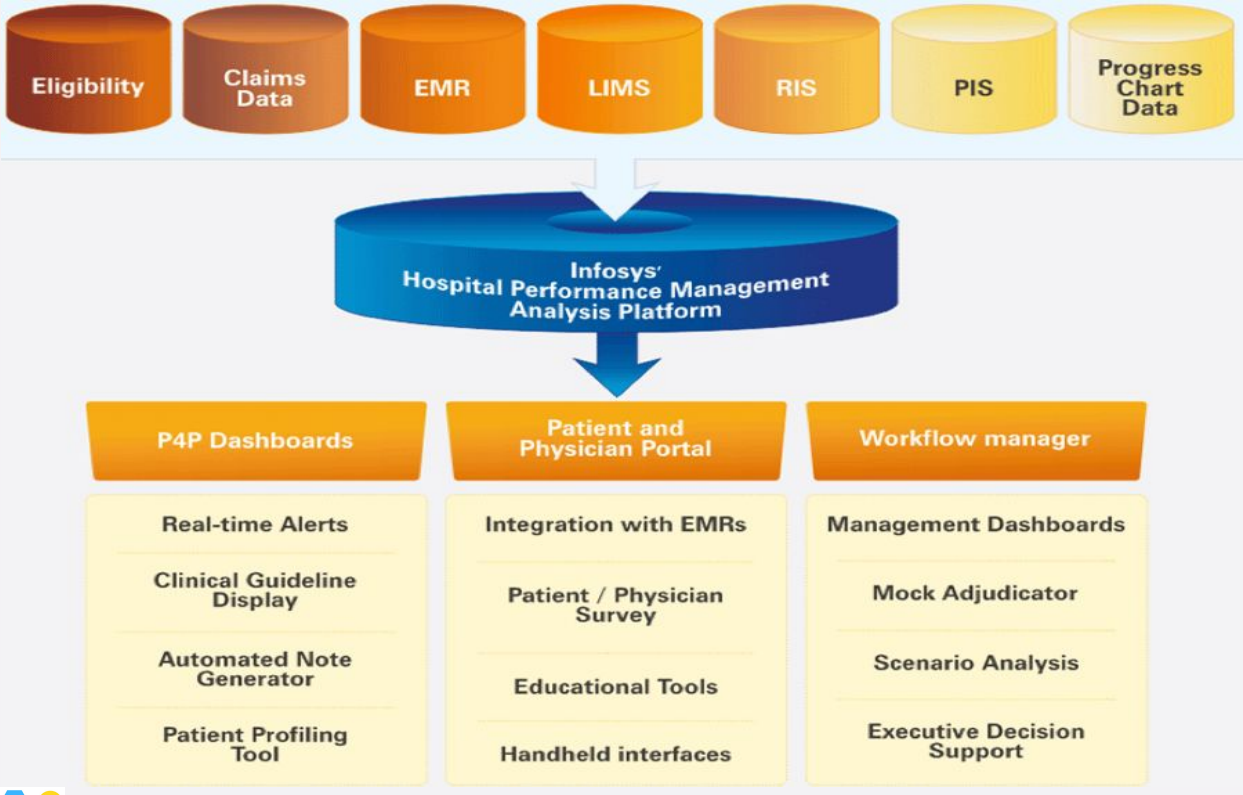






from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International

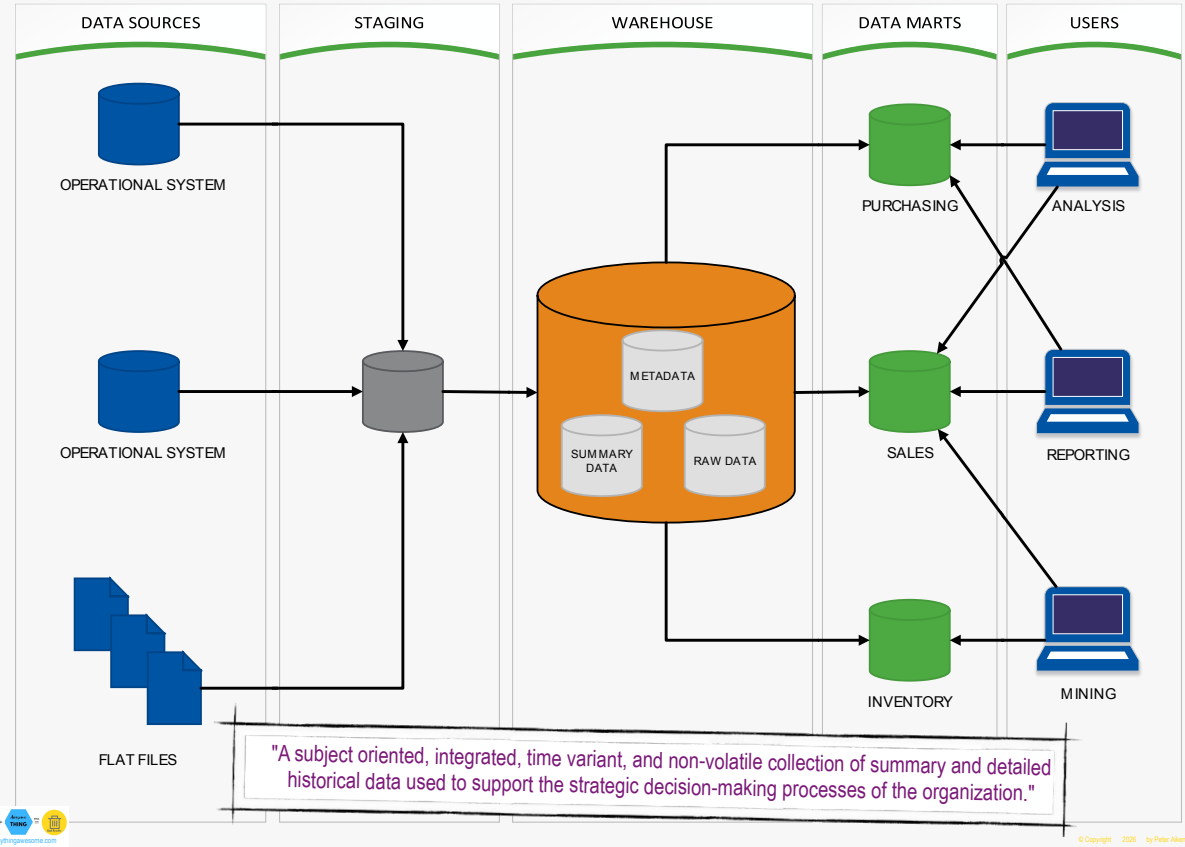
# Warehousing applied to a specific challenge



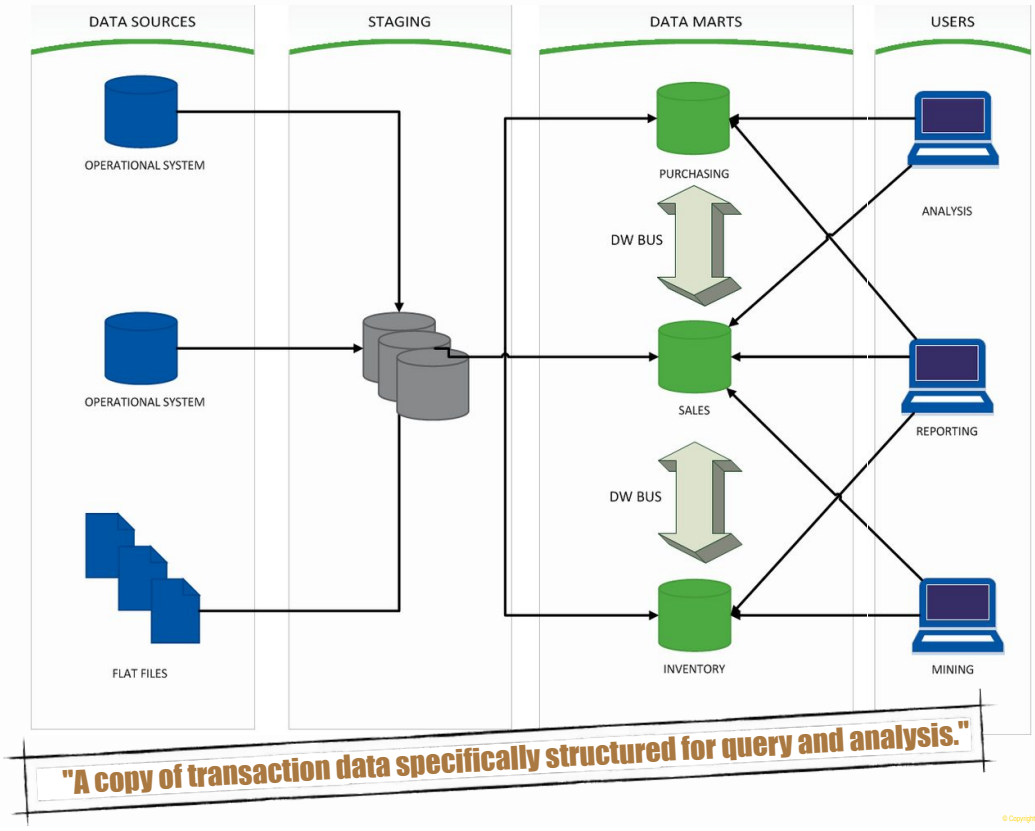

*http://www.infosys.com/industries/healthcare/industryofferings/Pages/healthcare-data-warehousing.aspx*



# Inmon Implementation/3NF

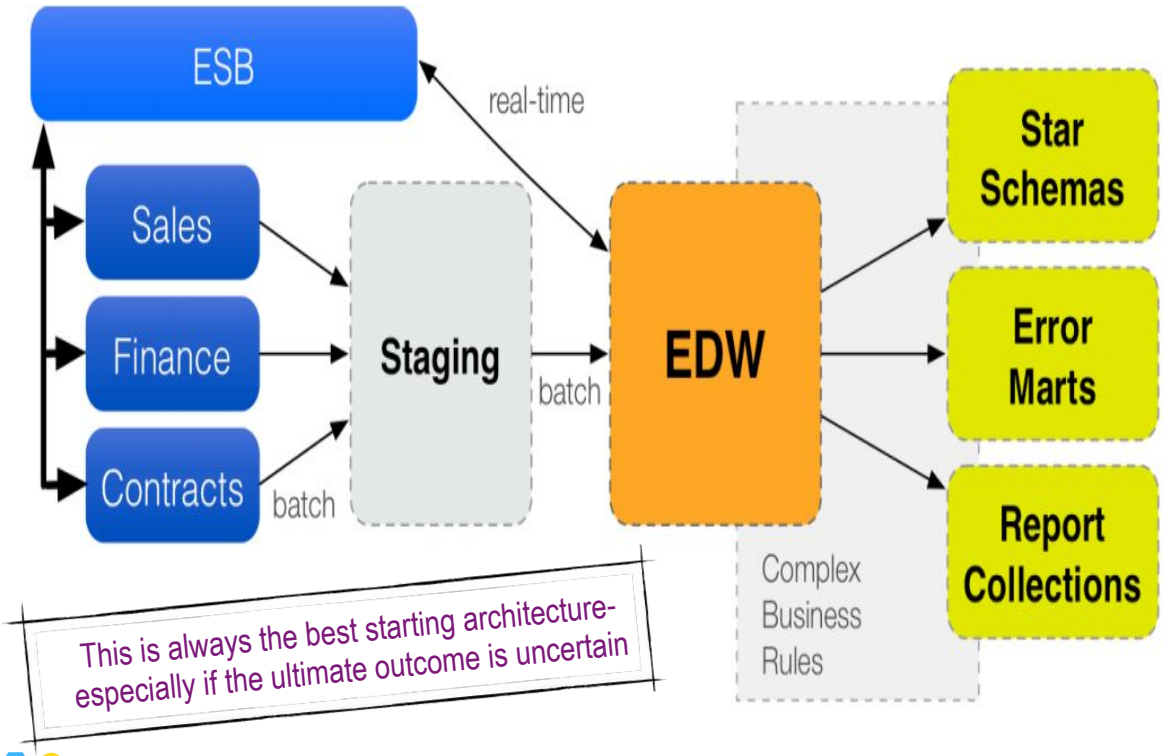


# Kimball Implementation/Dimensional



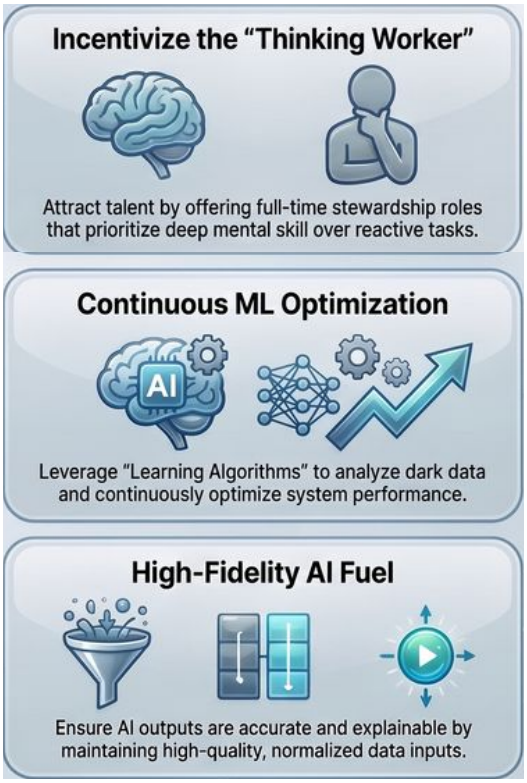



# Data Vault Implementation



## Attracting/Retaining Good Design & Engineering Talent

- Solution design is not based on semantic understanding
- Focus is often critical due to speed issues
- AI/ML focus on learning how the system performs and how to improve it

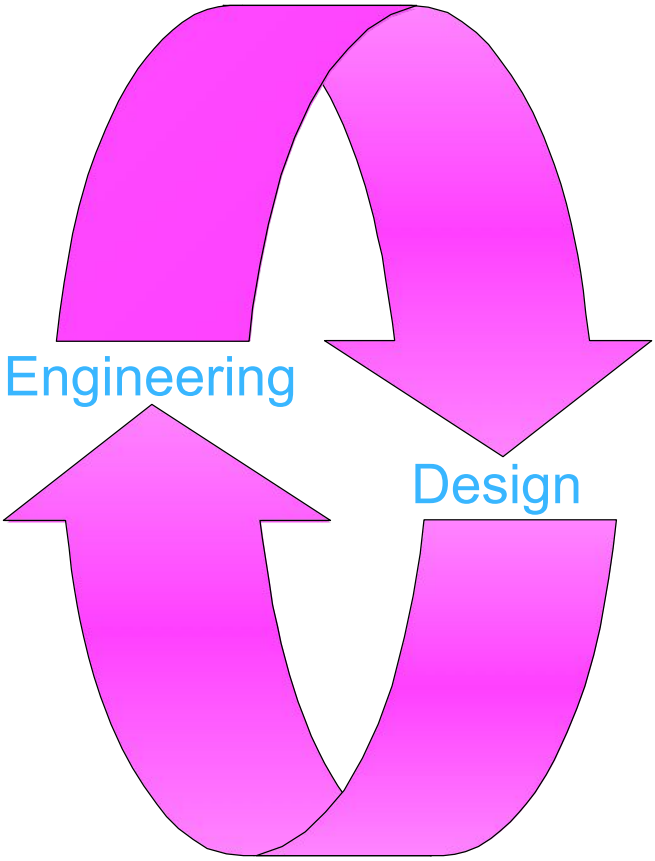







## Engineering/Design Relationship

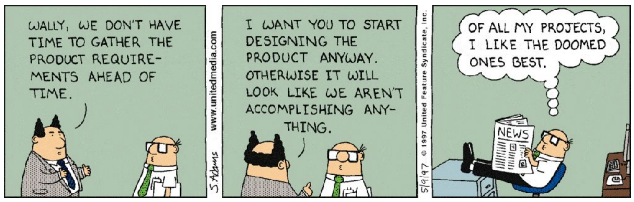
- Design is used to create and build systems too complex to be treated by engineering analysis alone
  - Require technical details as the exception
- Engineers develop the technical designs
  - Engineering/Crafts-persons deliver components supervised by:
    - Manufacturer
    - Building Contractor



## sales collect requirements → engineers develop

**Someone must directly go to the customer site, explore together, and use it while making changes**

- Forward Deployed Engineering (FDE) originated as a proprietary operational strategy




Palantir



## Basic tenants of FDE

Living with the users

To prevent disasters, engineers must undergo a profound, often uncomfortable mental transformation

1. From Outputs to Outcomes
   - In traditional engineering, jobs are defined by outputs: writing code, fixing Jira tickets, passing tests, and hitting deployment deadlines. Traditional roles are structured to insulate developers from customer reality, allowing them to ship code and immediately move to the next ticket.
   - Shift to owning outcomes. Work products are not judged by the elegance of their codebase, but by whether the customer actually achieves their business goals and "wins" with the solution. **You must feel the direct, personal weight of a customer depending on your work.**
2. Becoming a User Before a Builder
   - Traditional developers build software based on abstract product specifications. An FDE must become a user before they become a builder.
   - By rotating into the field, engineers witness how their systems actually collide with messy, real-world workflows.
3. Calibrating Engineering to Reality
   - Traditional training teaches engineers to always build the most robust, scalable, and beautifully designed system possible. In the field, this is often a recipe for failure.
   - FDEs must learn to calibrate their engineering to the exact situation.
4. Navigating the "Crucible" of Ambiguity
   - In a centralized headquarters, developers thrive on deep focus time, clear requirements, and stable release processes. The field is chaotic, fragmented, and highly interrupt-driven.

# Example Drivers

## Signal

Top modeling companies are cutting off the middleman role of traditional system integrators (SIs) and taking control of the "last mile" of implementation themselves.

**Salesforce plans to recruit 1000 FDEs**

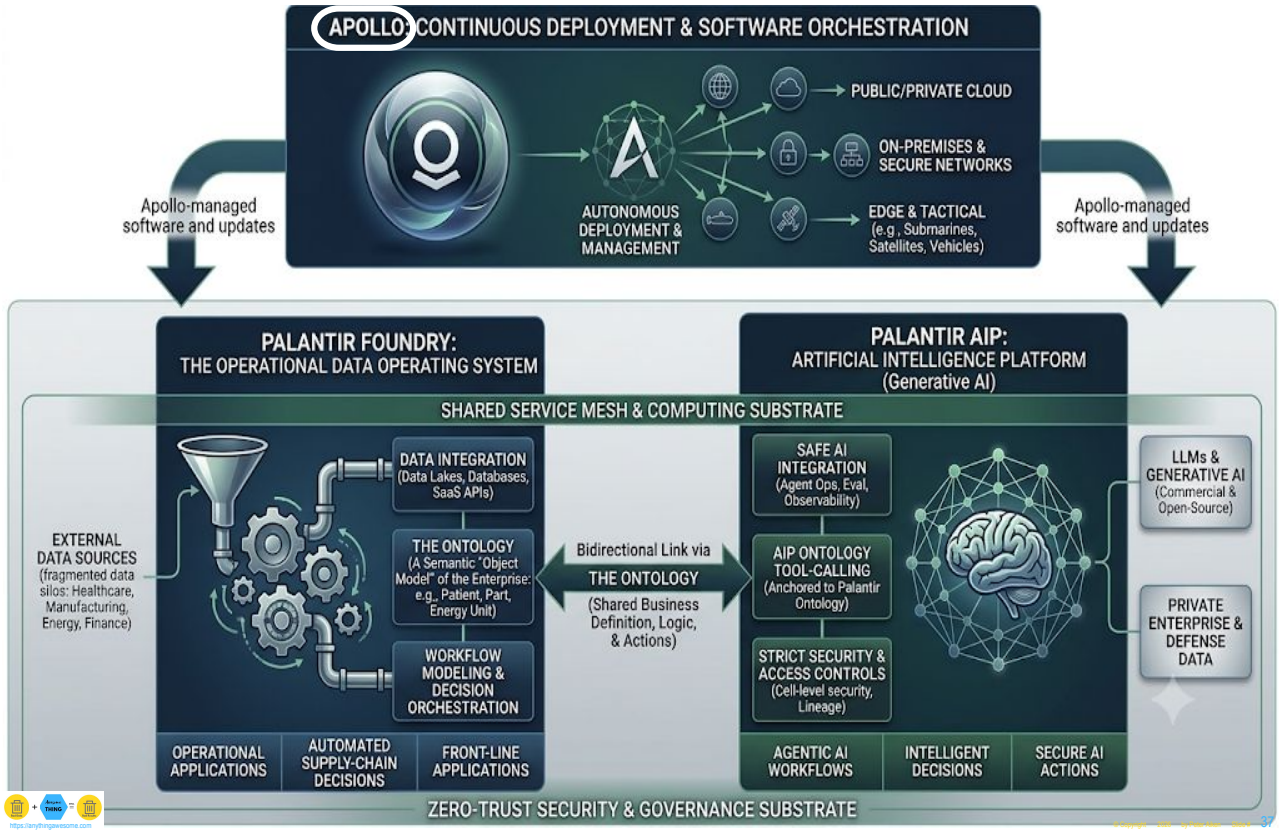



# Core Primitives of the Ontology

Palantir structures enterprise twins using 2 complementary primitives:

- Semantic Elements (Nouns & Relationships):
  - **Object Types**: Instead of an orders or customers table, data maps to an entity with typed properties, identity resolution, and a primary key. A Vessel or Patient object aggregates attributes from multiple disparate systems (e.g., SAP, Salesforce, IoT feeds).
  - **Link Types**: Explicit, bidirectional relationships connecting objects (e.g., Patient → Admitted To → Bed). Users and AI models traverse these links at runtime as an operational graph rather than constructing multi-table SQL JOIN statements.
- Kinetic Elements (Verbs & State Transitions):
  - **Action Types**: Governed transactions that mutate data (e.g., Reroute Shipment, Schedule Maintenance). Actions enforce business logic, schema constraints, and audit logs before writing updates back to Foundry object indexes and downstream ERPs/source systems.
  - **Functions & Models**: Server-side logic and machine learning models registered directly to objects, providing dynamic computations (such as calculating risk scores on an asset).

# Palantir Core Platforms






# Palantir Core Platforms

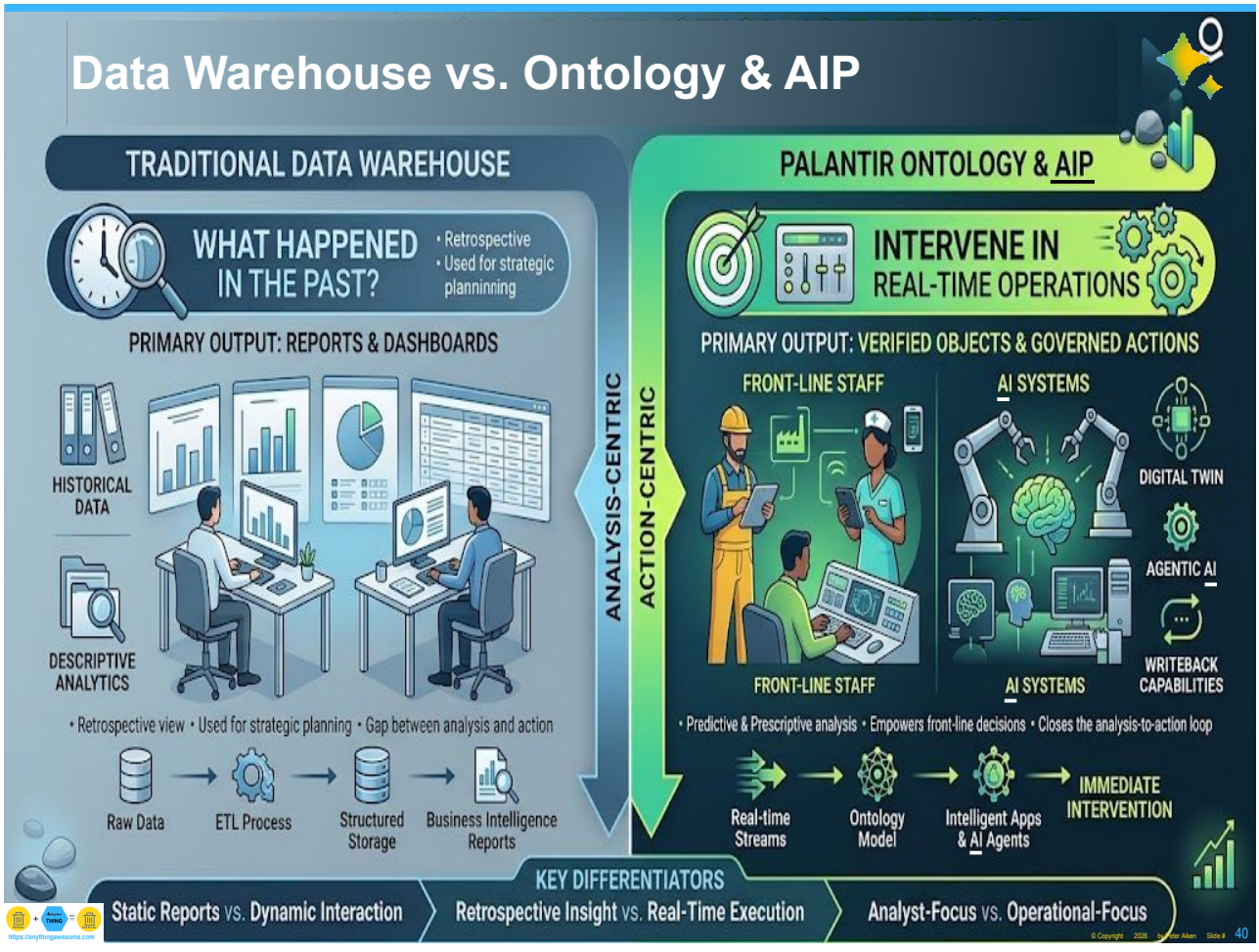
- Palantir Gotham: Originally built for defense and intelligence agencies (such as the DoD, CIA, and FBI). It allows analysts to map relationships, track targets, detect patterns, and coordinate tactical intelligence across battlefield operations.
- Palantir Foundry: The commercial enterprise platform. It serves as an operational data operating system, allowing corporations in healthcare, manufacturing, energy, and finance to integrate fragmented data silos, model workflows, and automate supply-chain decisions.
- Palantir AIP (Artificial Intelligence Platform): Designed to safely integrate large language models (LLMs) and generative AI directly into private enterprise and defense networks, anchoring AI actions to clear ontology rules and strict security access controls.
- Apollo: A continuous deployment and software orchestration platform that manages updates and deployment across disconnected, secure cloud, and edge environments (including submarines, satellites, and tactical ground vehicles).

# Comparing Styles

- Where a data warehouse organizes information into static relational schemas (tables, rows, and foreign keys) meant for descriptive analytics, Palantir Foundry and AIP position the Ontology as an operational layer that binds semantic definitions, computational logic, and transactional actions into a single graph.

| Architectural Dimension | Traditional Data Warehouse / Lakehouse | Palantir Foundry & AIP Ontology |
|---|---|---|
| **Core Abstraction** | Tables, views, columns, dimensions, and facts | Objects, links, properties, interfaces, and actions |
| **Primary Workload** | Read-heavy aggregations, SQL analytics, BI dashboards | Read/write operational workflows, decision loops, agentic AI |
| **Writeback & Execution** | Read-only target; writebacks require external ETL/reverse-ETL | Native **Action Types** with validation, side effects, and sync |
| **AI Integration (LLMs)** | RAG via text/vector search or text-to-SQL over tables | Grounded agent tool-calling bounded by object types and actions |
| **Data Governance** | Role-based / row-and-column access control on tables | Fine-grained object, link, and action-level policies with lineage |





# Program Overview

- Definitions
  - Data Warehousing
  - DM BoK guidance
  - Legacy to digital conversion (cloud) + AI
- Integration
  - Incorporate leading not bleeding edge
  - Requires an adaptive rather than a prescriptive approach
  - Emphasize engineering talent (Forward Deployed Engineer)
  - AI enables further structural formalization
- Preparation
  - Emphasis on storytelling first and visualization second
  - Analytics is both ubiquitous and not well understood
  - Keep improvements practically focused by strategy
- Best Practices
  - Cannot use what is not understood-understand what you have
  - PDCA (plan, do, check, act)
  - Iteratively refine/cull with respect to strategic direction
- Take Aways/References/Q&A

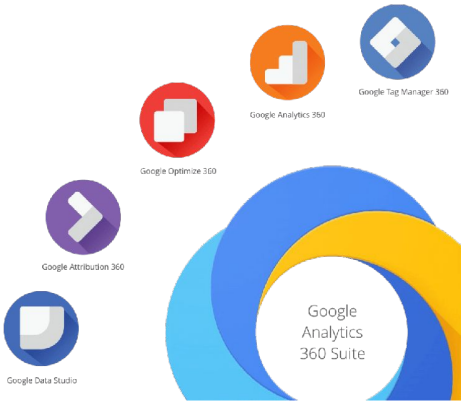

**Data Warehousing Reimagined:**

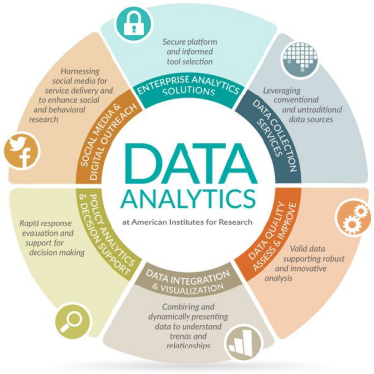
**Fueling AI-powered BI at Scale**

DATA-ED ONLINE
DATAVERSITY

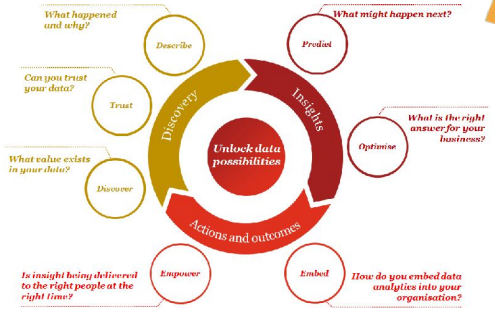


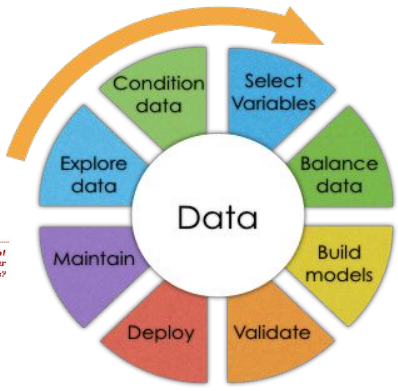




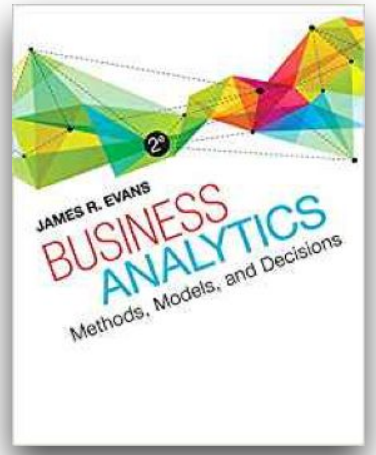

**Analytics**



## Business Analytics, or simply analytics, is:

- is the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their business operations and make better, fact-based decisions.
- is "a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving."
- supported by various tools such as Microsoft Excel and various Excel add-ins, commercial statistical software package such as SAS or Minitab and more complete business intelligence suites that integrate data with analytical software

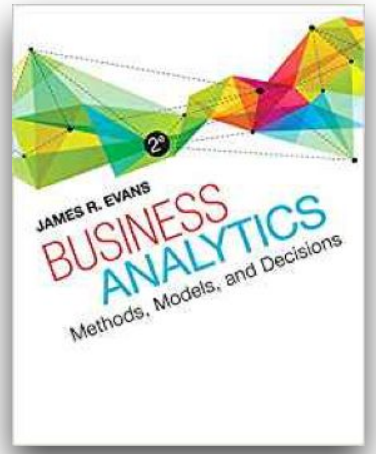



*Business Analytics: Methods, Models, and Decisions*, Second Edition by James R. Evans, page 4.



## Analytics

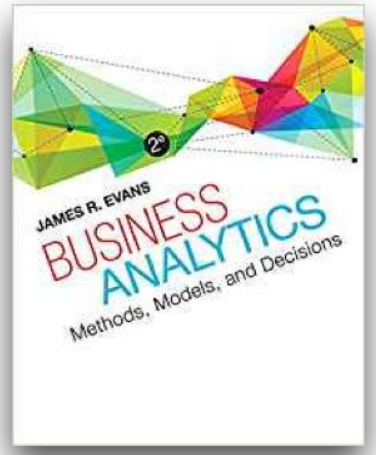
- is the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their business operations and make better, fact-based decisions.
- is "a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving."
- supported by various tools such as Microsoft Excel and various Excel add-ins, commercial statistical software package such as SAS or Minitab and more complete business intelligence suites that integrate data with analytical software




*Business Analytics: Methods, Models, and Decisions*, Second Edition by James R. Evans, page 4.

# Analytics

- is the use of data, hardware/software, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their business operations and make better, fact-based decisions.
- is “a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving.”
- supported by various tools such as Microsoft Excel and various Excel add-ins, commercial statistical software package such as SAS or Minitab and more complete business intelligence suites that integrate data with analytical software

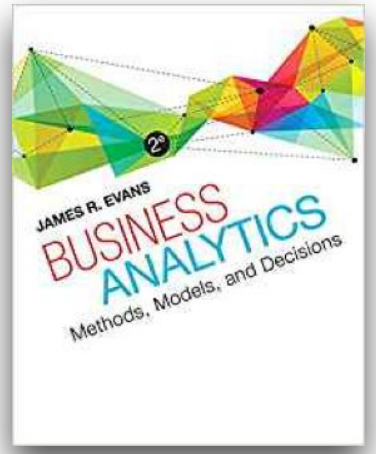



*Business Analytics: Methods, Models, and Decisions*, Second Edition by James R. Evans, page 4.



# Analytics

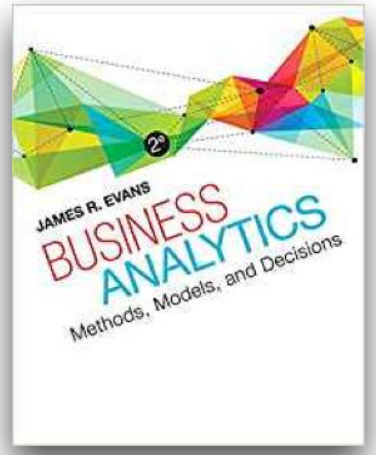
- is the use of data, hardware/software, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their business operations and make better, fact-based decisions.
- is “a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving.”
- supported by various tools such as Microsoft Excel and various Excel add-ins, commercial statistical software package such as SAS or Minitab and more complete business intelligence suites that integrate data with analytical software




*Business Analytics: Methods, Models, and Decisions*, Second Edition by James R. Evans, page 4.

# Analytics

- is the use of data, hardware/software, quantitative methods, and models to help managers gain improved insight about their business operations and make better, fact-based decisions.
- is “a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving.”
- supported by various tools such as Microsoft Excel and various Excel add-ins, commercial statistical software package such as SAS or Minitab and more complete business intelligence suites that integrate data with analytical software

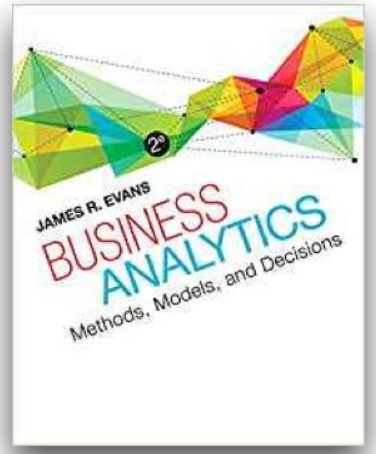



*Business Analytics: Methods, Models, and Decisions*, Second Edition by James R. Evans, page 4.



# Analytics

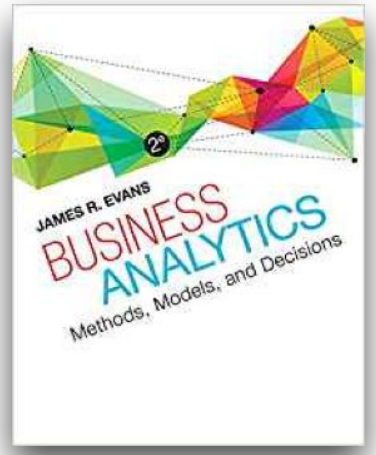
- is the use of data, hardware/software, quantitative methods, and models to help managers make better decisions.
- is “a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving.”
- supported by various tools such as Microsoft Excel and various Excel add-ins, commercial statistical software package such as SAS or Minitab and more complete business intelligence suites that integrate data with analytical software




*Business Analytics: Methods, Models, and Decisions*, Second Edition by James R. Evans, page 4.

# Analytics

- is the use of data, hardware/software, quantitative methods, and models to help managers make better decisions.
- the process of transforming data into actions through analysis to solve problems
- supported by various tools such as Microsoft Excel and various Excel add-ins, commercial statistical software package such as SAS or Minitab and more complete business intelligence suites that integrate data with analytical software

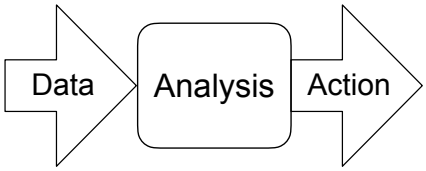



*Business Analytics: Methods, Models, and Decisions*, Second Edition by James R. Evans, page 4.



# Analytics (31 words almost ⅓ of original)

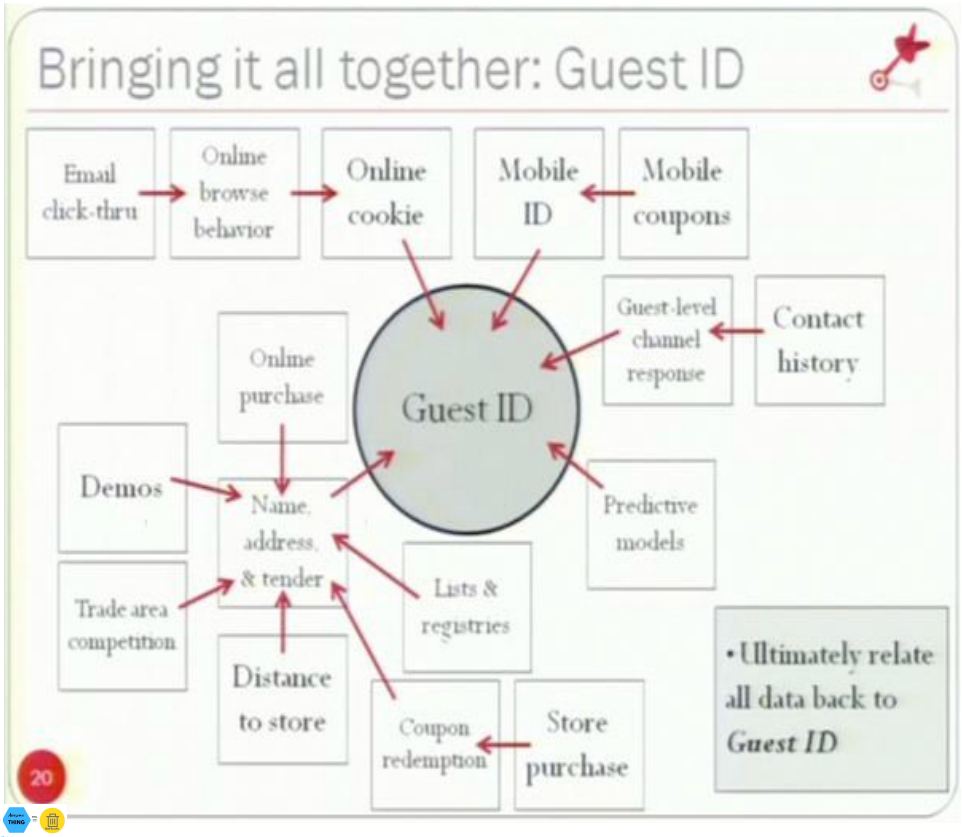
- is the use of data, hardware/software, quantitative methods, and models to help managers make better decisions- the process of transforming data into actions through analysis to solve problems using **tools**





http://rmportal.performedia.com/node/1373 and http://www.predictiveanalyticsworld.com/patimes/target-really-predict-teens-pregnancy-inside-story/ http://rmportal.performedia.com/rm/paw10/gallery_01#1373

# Target Isn't Just Predicting Pregnancies

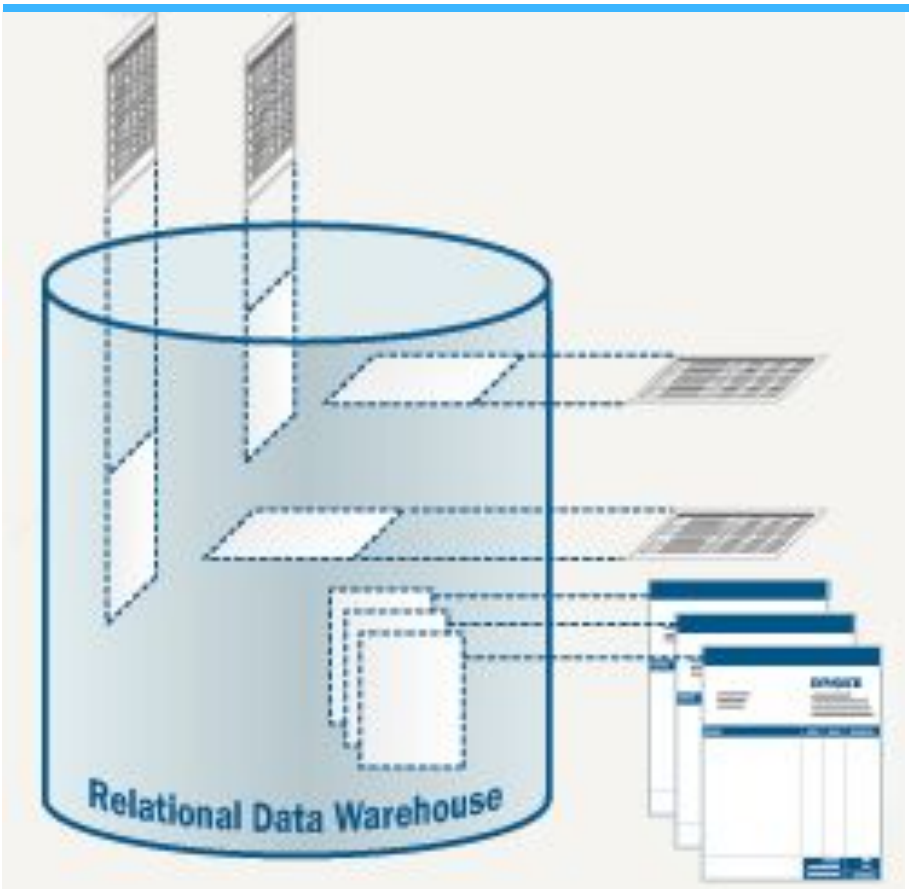





# Mission

- *Advance change in America by ensuring equitable access to nutritious food for all in partnership with food banks, policymakers, supporters, and the communities we serve*
- Great understanding and skill executing rapid response to local needs

- What else could be done with their data?
- Rerouting busses = eliminating food deserts for thousands of residents

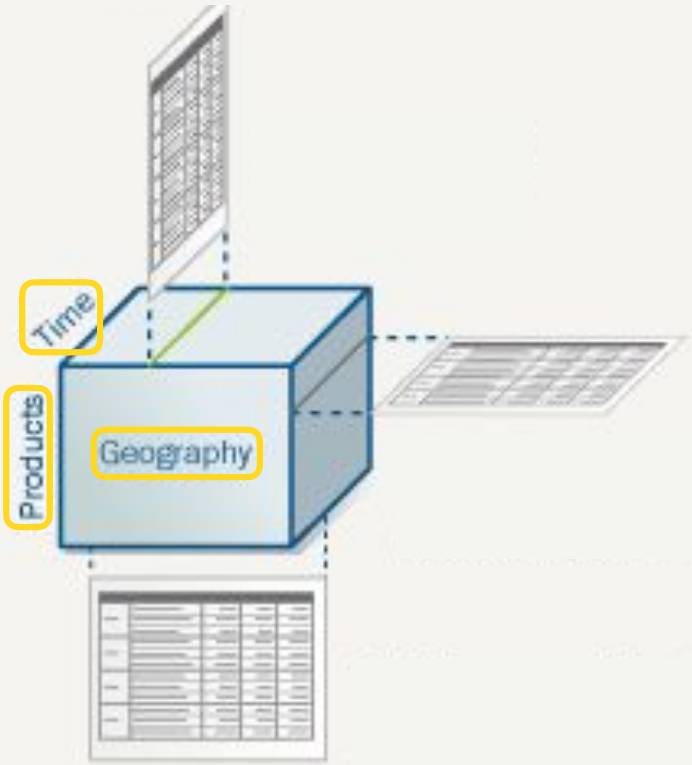




## Basics

- Users can "drill" anywhere
- Entire collection "cube" is accessible
- Summaries to transaction-level detail

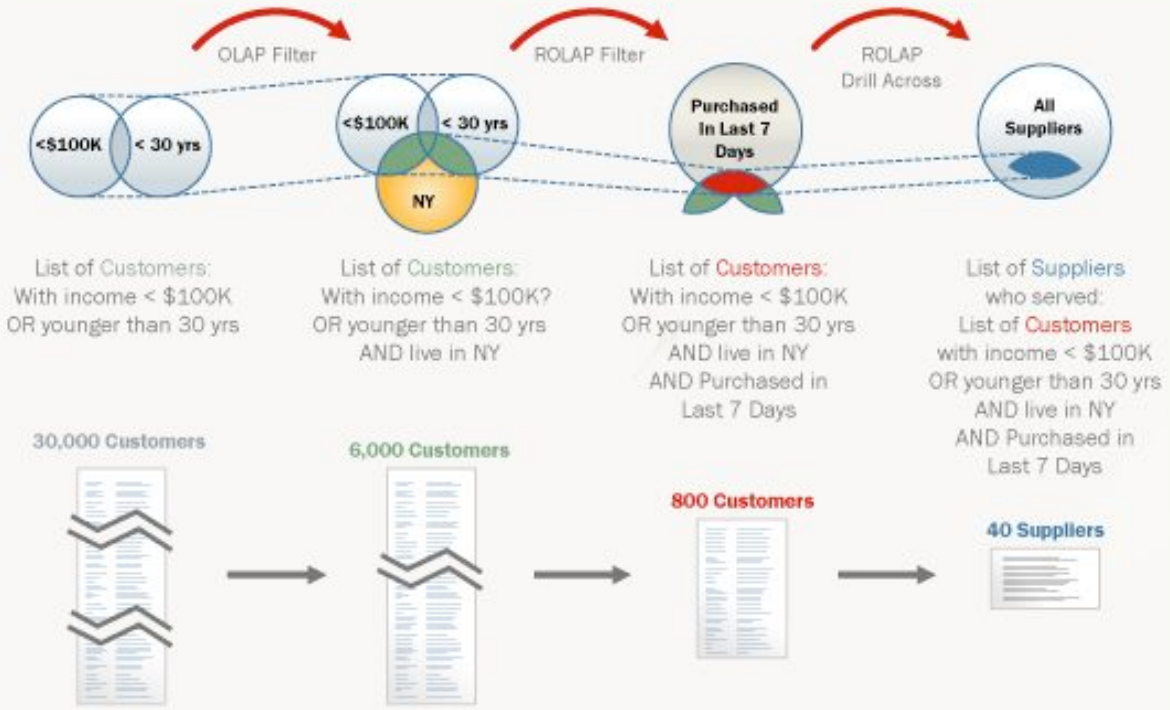


## Sample questions …



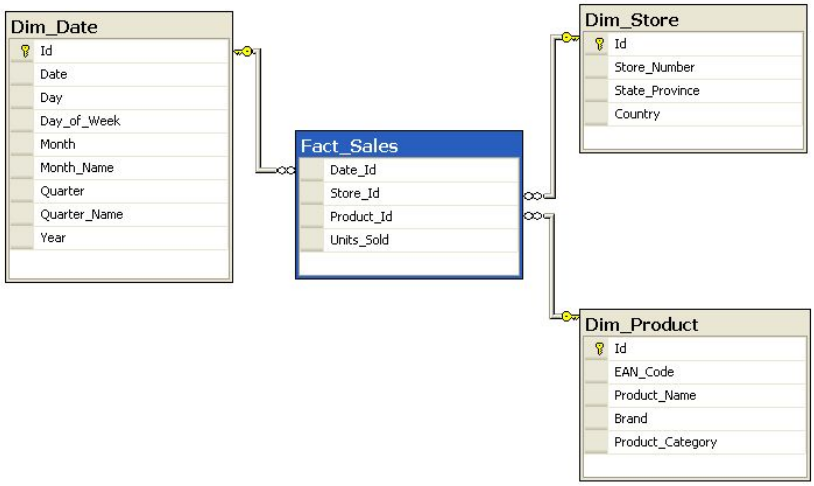

- Emphasis on the "cube"
- 'N' dimensions
- Permits different users to "slice and dice" subsets of data
- Viewing from different perspectives

# Example: Set Analysis





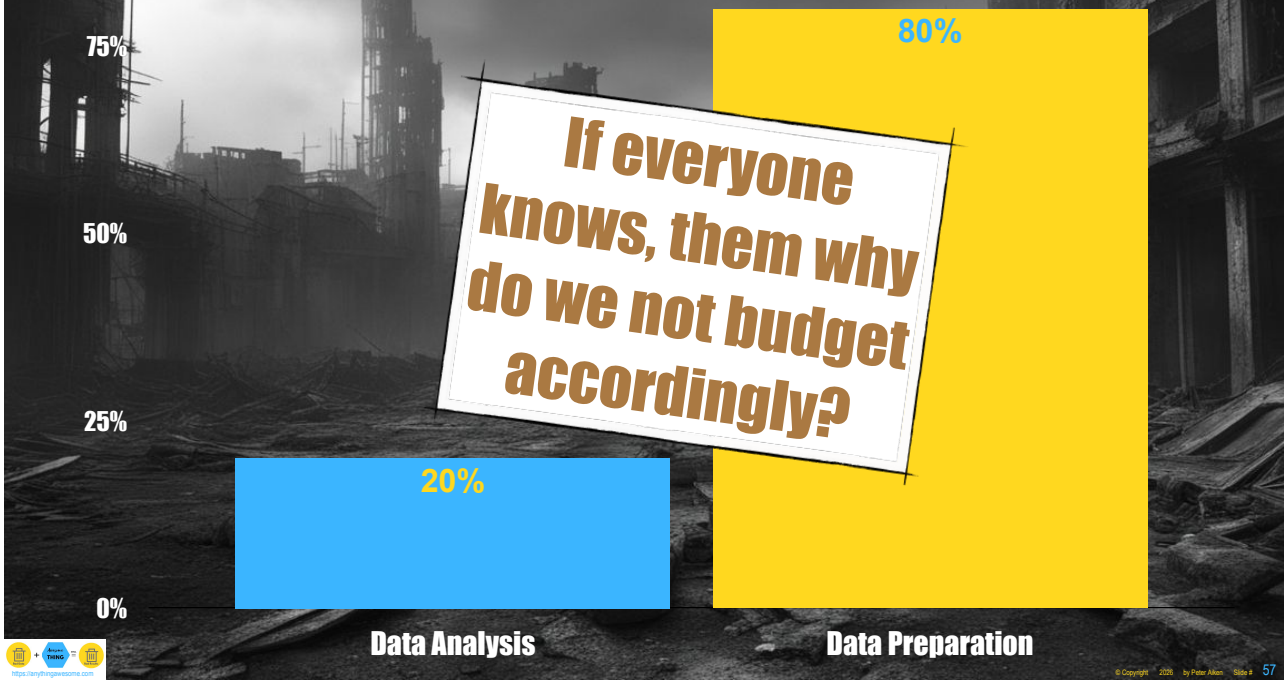

# Star Schema




- Comprised of "fact tables" that contain quantitative data, and any number of adjoining "dimension" tables
- Optimized for business reporting
- Use Cases
  - OLAP (Online Analytic Processing)
  - BI

# Everyone wants to do better data analysis …

- Some data preparation is inevitable
- What would a 'good' ratio be?
- "Everyone knows"

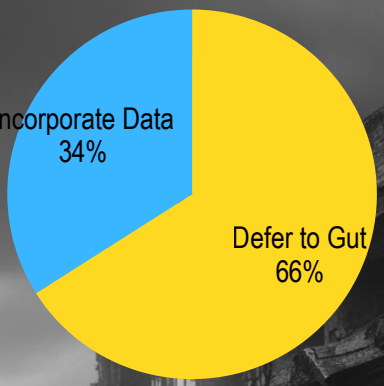




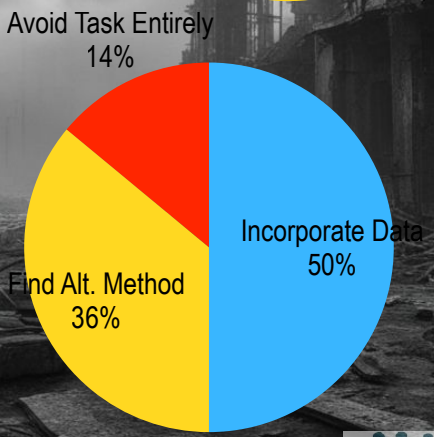

http://TheDataLiteracyProject.org

# When asked to incorporate data

Too many organizations have simply put data in the hands of employees and expected them to make a success of it

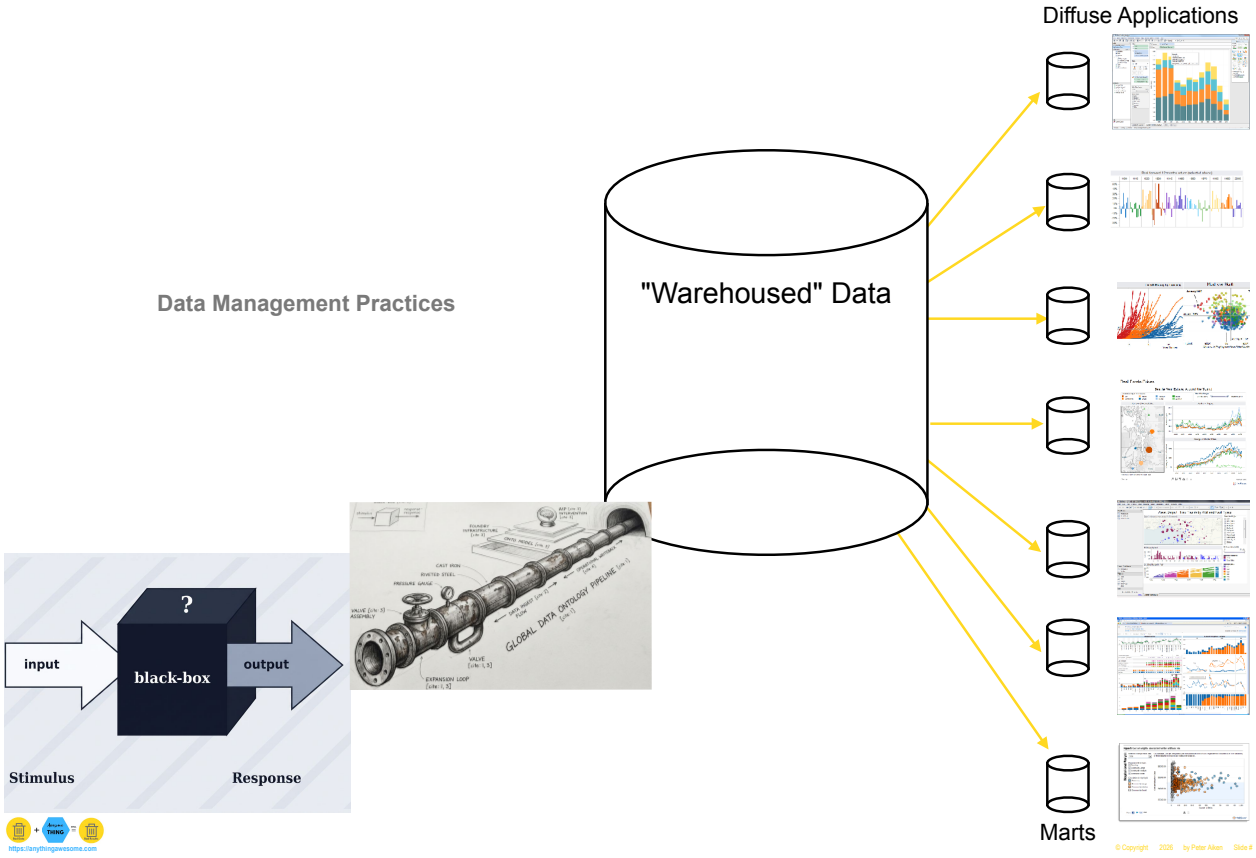



- Data appreciation isn't translating into employee adoption
  - 48% frequently make gut decisions
  - 66% for C-suite executives
- Lack of data skills is limiting workplace productivity
  - 36% said they would find an alternative method to complete the task without using data
  - 14 percent avoid the task entirely




http://TheDataLiteracyProject.org

# Where To Focus Limited Resources?





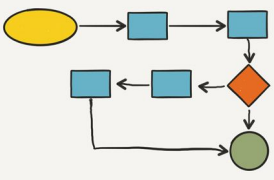

# What is Strategy?


strat·e·gy
/'stratəjē/

*noun*

1. a plan of action or policy designed to achieve a major or overall aim.
   "time to develop a coherent economic strategy"
   *synonyms:* master plan, grand design, game plan, plan (of action), action plan, policy, program; More


- Current use derived from military
  - **a pattern in a stream of decisions**
    [Henry Mintzberg]



## Former Walmart Business Strategy

# Every Day Low Price

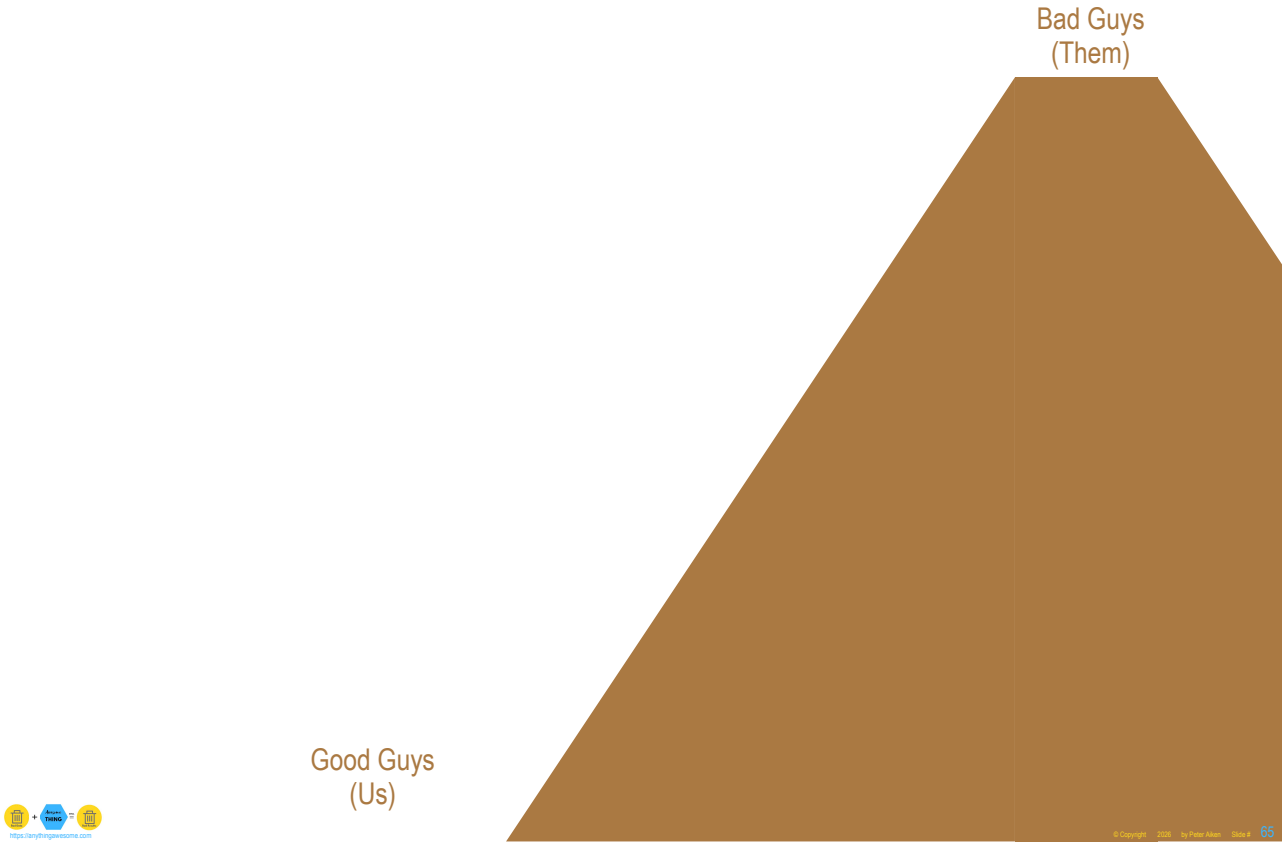


# Strategy Example 3




# Strategy Example 3

## Strategy Example 3



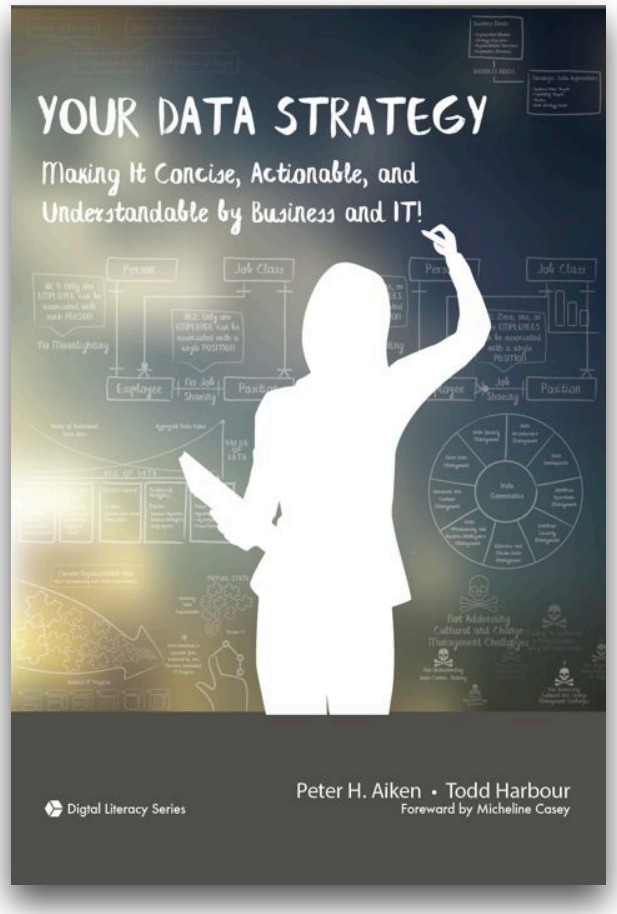



## Strategy Guides Workgroup Activities

# A pattern in a stream of decisions

## Your Data Strategy

- Highest level data guidance available ...
- Focusing data activities on business-goal achievement ...
- Providing guidance when faced with a stream of decisions or uncertainties
- Data strategy most usefully articulates how data can be best used to support organizational strategy
- This usually involves a balance of remediation and proactive measures






## Program Overview

- Definitions
  - Data Warehousing
  - DM BoK guidance
  - Legacy to digital conversion (cloud) + AI
- Integration
  - Incorporate leading not bleeding edge
  - Requires an adaptive rather than a prescriptive approach
  - Emphasize engineering talent (Forward Deployed Engineer)
  - AI enables further structural formalization
- Preparation
  - Emphasis on storytelling first and visualization second
  - Analytics is both ubiquitous and not well understood
  - Keep improvements practically focused by strategy
- Best Practices
  - Cannot use what is not understood-understand what you have
  - PDCA (plan, do, check, act)
  - Iteratively refine/cull with respect to strategic direction
- Take Aways/References/Q&A


**Data Warehousing Reimagined:**

**Fueling AI-powered BI at Scale**

DATA-ED ONLINE DATAVERSITY

# Reframing the question

From: How shall we build this data warehouse?

- (or worse) … What should go into this warehouse?

To: How can warehousing capabilities solve this business challenge?

- (better still) … How can warehousing capabilities solve this class of business challenges?

Other examples

- Are you ready for warehousing?
  - Foundational practices
  - Project deliverables
- Will you get it right the first time?
  - Is the business environment constantly evolving?
  - Do you have an agreed upon enterprise-wide vocabulary?
- Is your data warehouse intended to be the enterprise audit-able system of record?
  - Extract, transform and load requirements
  - Data transformation requirements
  - How fast do you need results?
  - Performance of inserts vs reads





https://aiinstitute.hbs.edu/platform-digit/submission/skywise-bringing-data-analysis-to-the-airline-industry/

# Skywise

DIGITAL INNOVATION AND TRANSFORMATION
MBA STUDENT PERSPECTIVES

COURSE ASSIGNMENTS

Assignment: Competing with Data

## Skywise: Bringing Data Analysis to the Airline Industry

By Crimson Ranger
Alumni

POSTED NOV 13, 2019

Airbus launched its Skywise system to allow its airline customers to collect and better analyze their aircraft data with the goal of improving operations and reducing costs.

The global airline industry is constantly looking for ways to manage controllable costs (maintenance, fleet optimization, aircraft performance) as it responds to cost trends outside of its control (fuel prices, taxes, global growth trends, macroeconomic shocks). Against this backdrop, aircraft manufacturers seek to differentiate themselves by offering superior after sale services to further reduce the operating costs of their aircraft. In June 2017, Airbus launched its open aviation data platform, Skywise, "to support the digital transformation of the industry and add value for our customers' operations"[1].

# Data Fragmentation, Scale, and Complexity

**In aviation, data isn't just large—it's disconnected. Five million parts across thousands of global suppliers.**

### Fragmented Ecosystem

Data was siloed across flight ops, maintenance logs, engineering specs, and supply chain ERPs. Decision-making required weeks of manual ETL and cross-departmental alignment.

### High-Frequency Telemetry

Modern aircraft like the A350 generate massive volumes of sensor data that legacy data warehouses struggle to operationalize in real-time.

Calculation for per-flight telemetry volume (Up to 1TB/day).



# Kinetic Primitives of Skywise

### Objects

The digital twins of physical assets. Every engine part is a distinct, governed object in Skywise.

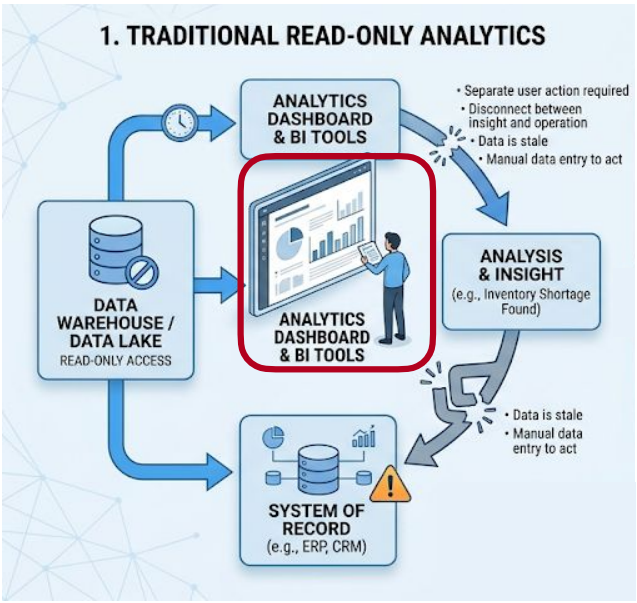
### Links

The relationships between objects. Connecting sensor telemetry directly to specific tail numbers.

### Actions

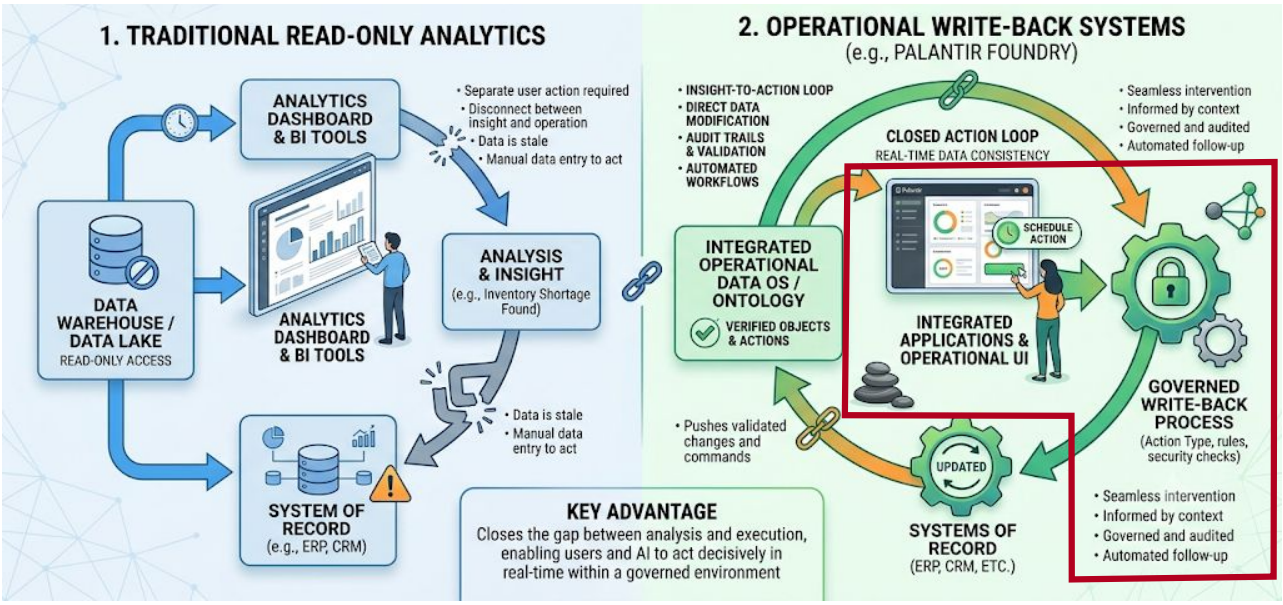
Governed Write-Back that allow shop-floor engineers to trigger procurement directly from analytics.

# Write-Backs (e.g., Palantir Foundry, BI tools)






# Write-Backs (e.g., Palantir Foundry, BI tools)




- The dashboard or application layer allows users—or automated agents—to click a button, edit a table row, or trigger a workflow that immediately writes changes back to the production databases or source tools.
- An engineer or AI agent doesn't just read an alert saying an aircraft engine valve is worn out; they execute a "Schedule Inspection" or "Reroute Part" action.
- The platform runs validation rules, checks user permissions, logs an audit trail, updates the internal ontology state, and pushes the change out via API directly to the underlying transactional systems (like an ERP).

## The Ontology: Mapping Reality

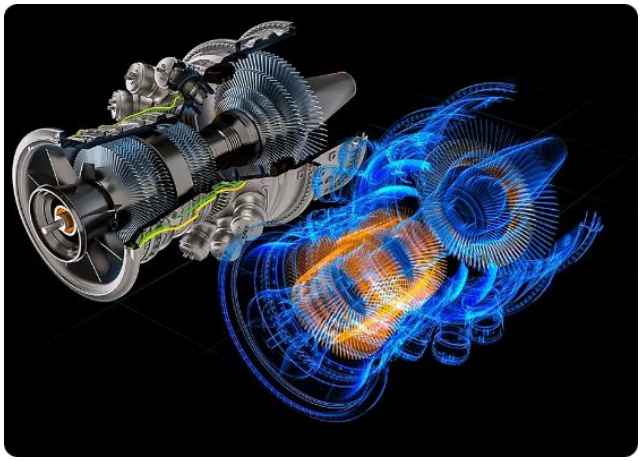
- Extensions beyond rows and columns

- Foundry maps relational data into a Digital Twin. Instead of querying tables, data professionals interact with high-level Objects:

**Aircraft**: Aggregates flight history, sensor feeds, and age.

**Components**: Links parts to maintenance cycles and serials.

**Maintenance**: Connects labor availability to parts inventory.





## Production Acceleration by Synchronizing the Supply Chain

Reduction in A350 Lead Time

**33%**

AIRBUS

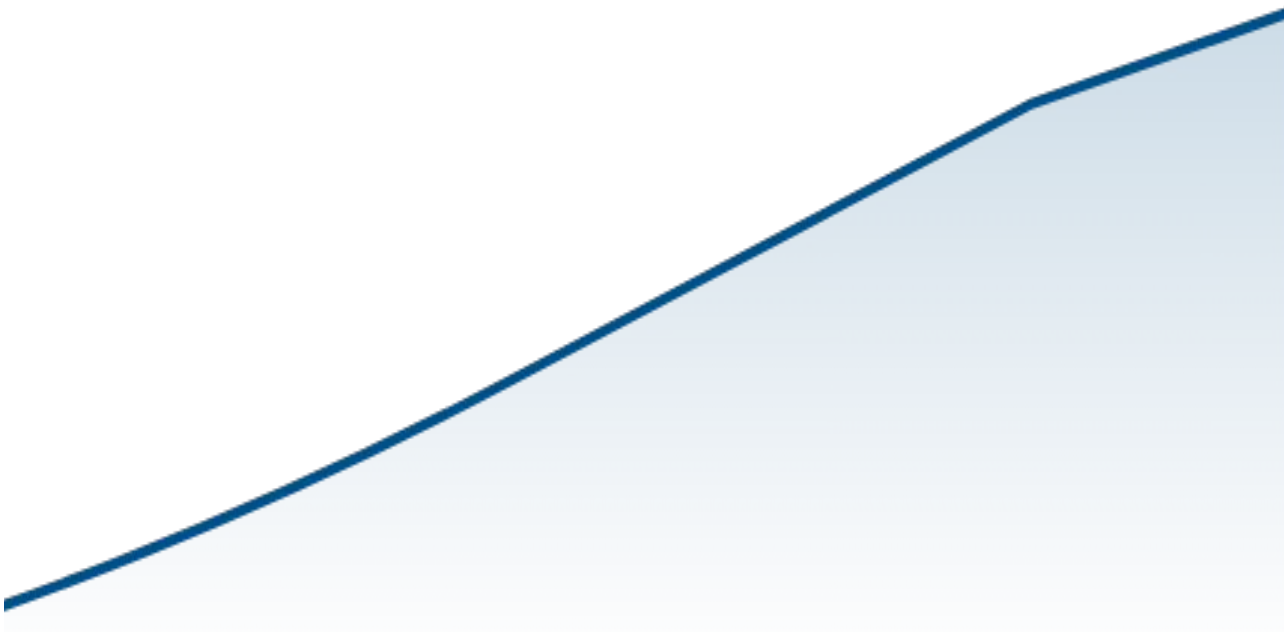
Eliminated supply chain bottlenecks—integrating five million parts across thousands of global suppliers.

# Operational ROI: Predictive Maintenance





# Skywise Growth-Total Aircraft Connected

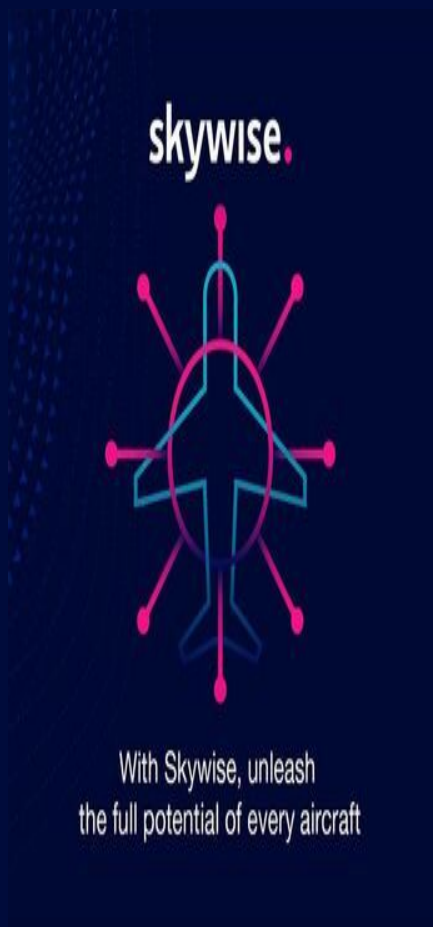


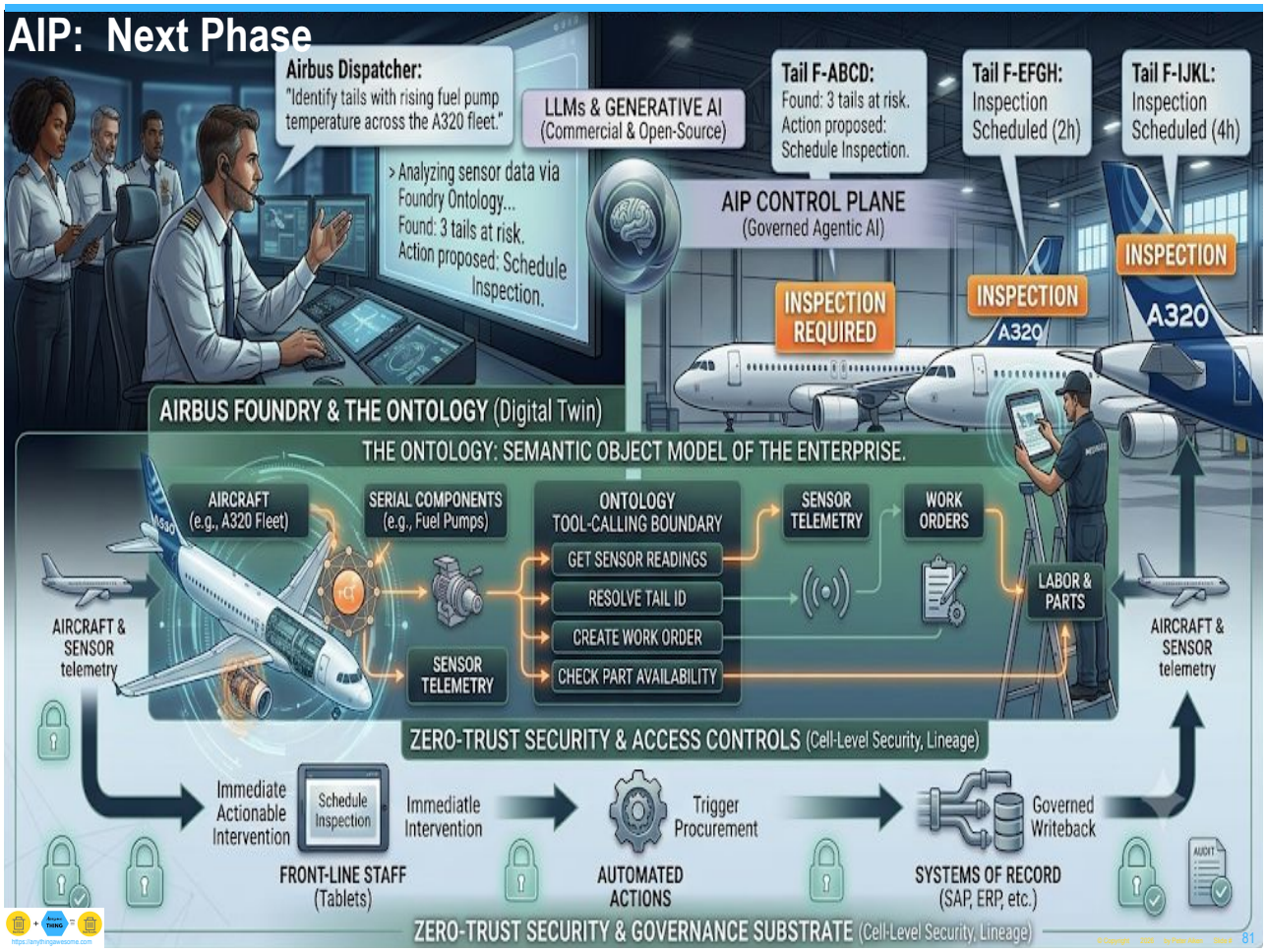
Launch 2017

500 Airlines and 12k Aircraft

# By the Numbers

| Capability | Status / Metric | Impact |
|---|---|---|
| Fleet Connectivity | 12,300+ Aircraft | Total global visibility. |
| Daily Data Ingest | ~1 Terabyte / A350 | Petabyte-scale time series management. |
| Data Throughput | Petabyte Scale | Industrial-grade processing. |
| Global Users | 55,000+ Personnel | Democratized data access. |
| Partner Airlines | 150+ Companies | Global industry standard. |
| Certified App Partners | 20+ | Extensible API-first ecosystem. |






# Program Overview

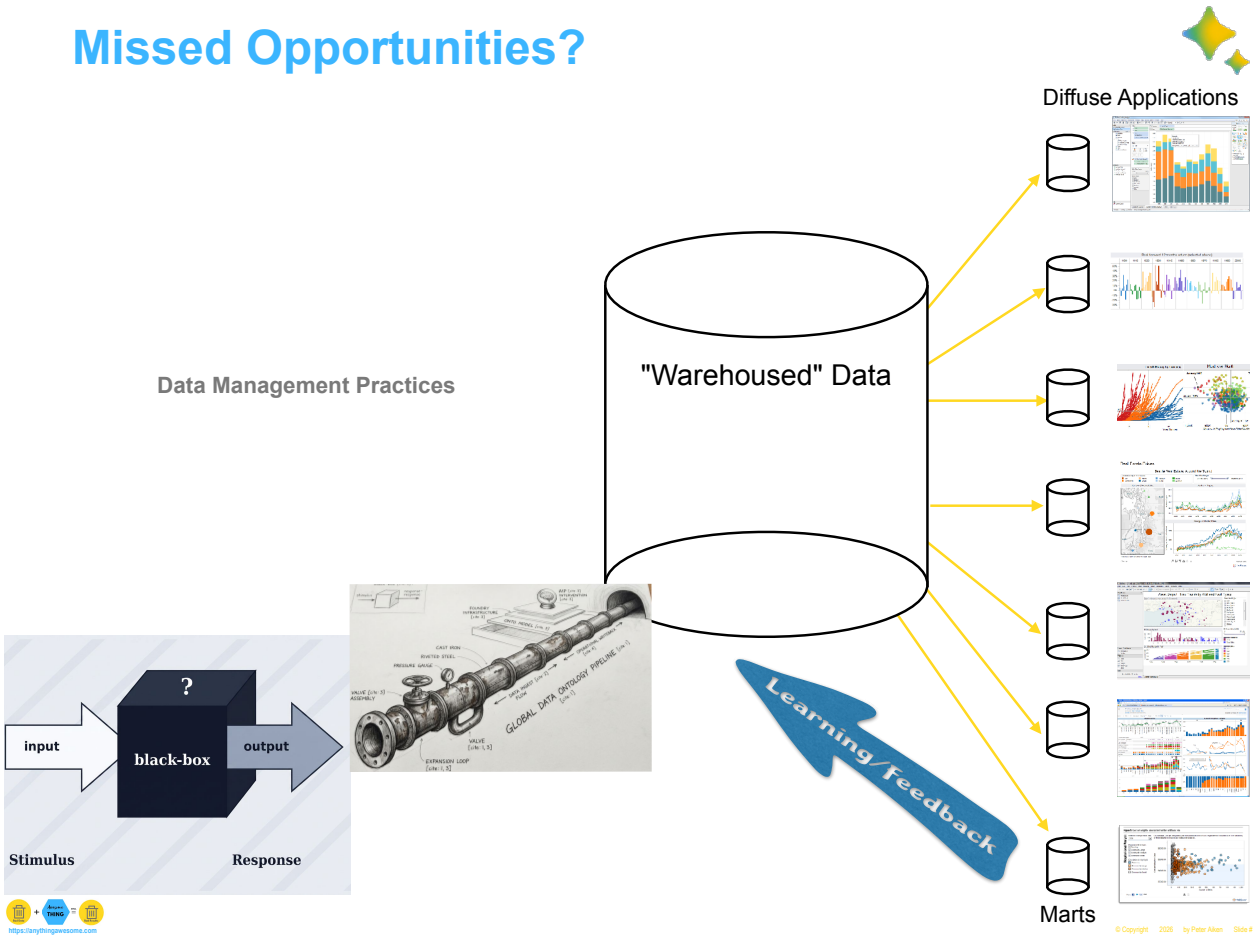
- Definitions
  - Data Warehousing
  - DM BoK guidance
  - Legacy to digital conversion (cloud) +AI
- Integration
  - Incorporate leading not bleeding edge
  - Requires an adaptive rather than a prescriptive approach
  - Emphasize engineering talent (Forward Deployed Engineer)
  - AI enables further structural formalization
- Preparation
  - Emphasis on storytelling first and visualization second
  - Analytics is both ubiquitous and not well understood
  - Keep improvements practically focused by strategy
- Best Practices
  - Cannot use what is not understood-understand what you have
  - PDCA (plan, do, check, act)
  - Iteratively refine/cull with respect to strategic direction
- Take Aways/References/Q&A


**Data Warehousing Reimagined:**

**Fueling AI-powered BI at Scale**

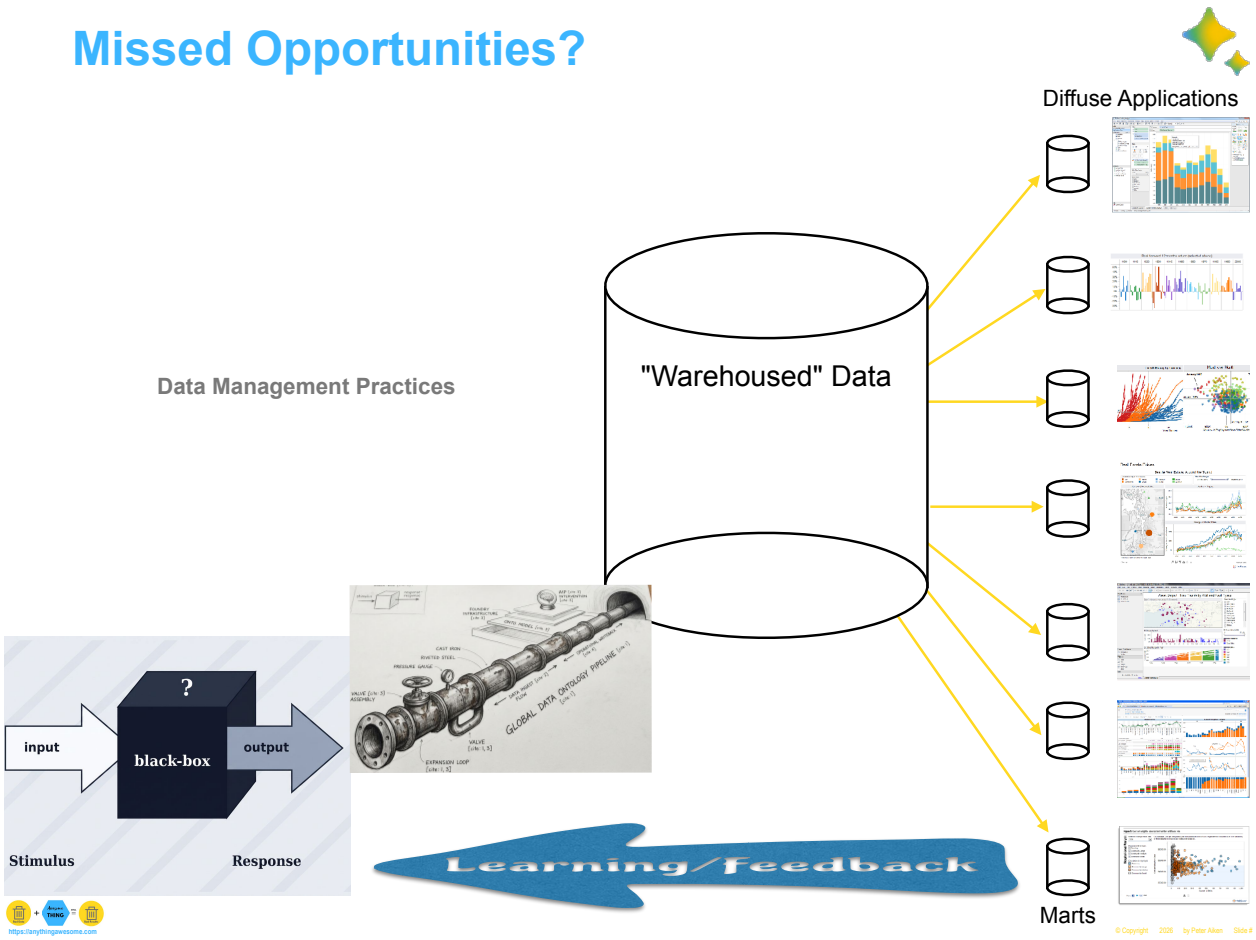
DATA-ED ONLINE
DATAVERSITY

# Missed Opportunities?



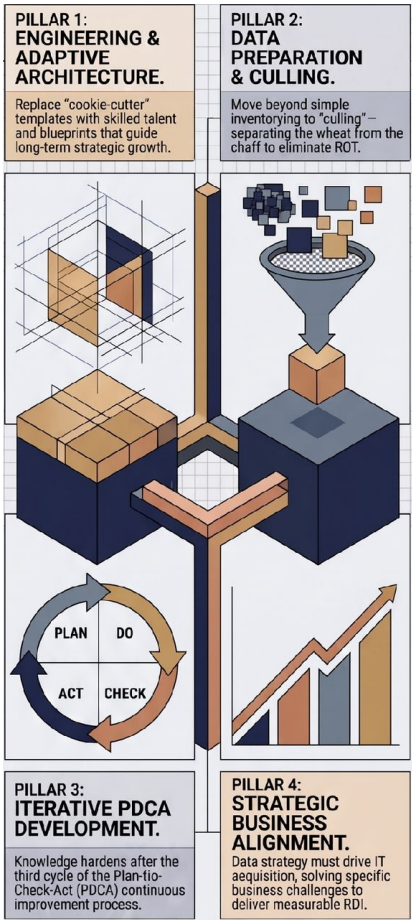



# Missed Opportunities?



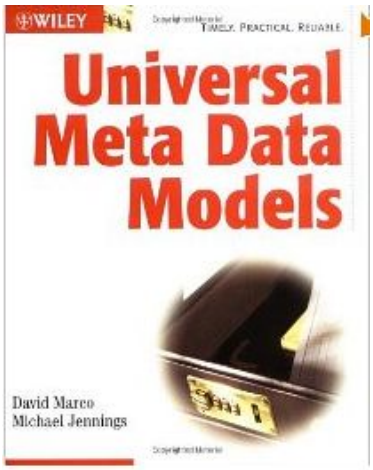
# In Sum

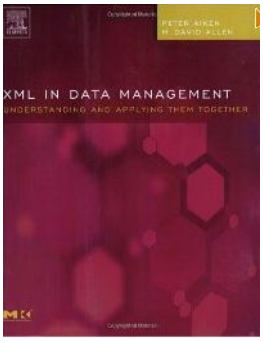
- Engineering and Architecture
  - Effective data warehousing requires skilled engineering talent and a holistic, adaptive architectural approach, rather than a prescriptive, cookie-cutter method.
- Data Preparation and Quality
  - The absolute the need to separate "the wheat from the chaff," with a focus on eliminating redundant, obsolete, or trivial (ROT) data before it enters the warehouse.
- Iterative Development
  - Continuous improvement cycle (plan, do, check, act) iteratively refines the data warehouse capabilities in alignment with strategic business direction.
- Business Alignment
  - Stressing that data warehousing must be driven by business needs and a clear strategy, with a focus on using warehousing capabilities to solve specific business challenges rather than just building a large repository.

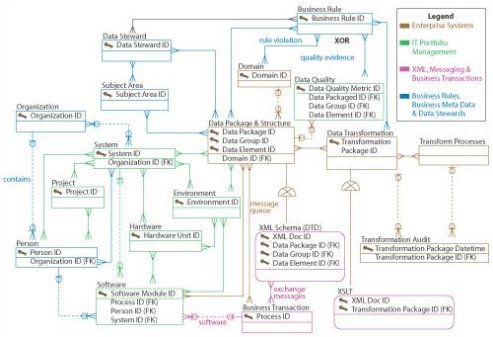


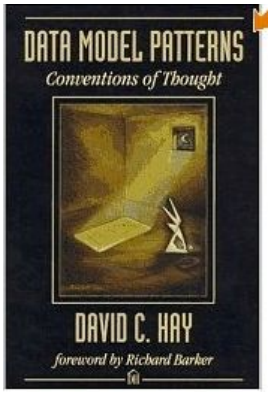


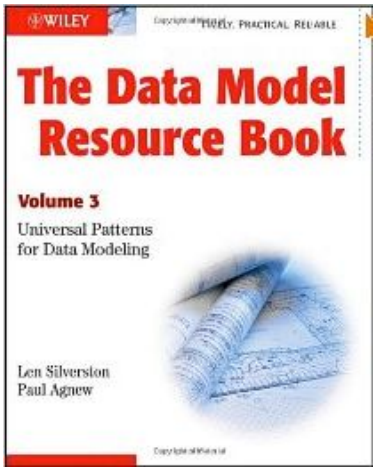

# Meta Data Models




Source:http://dmreview.com/article_sub.cfm?articleID=1000941

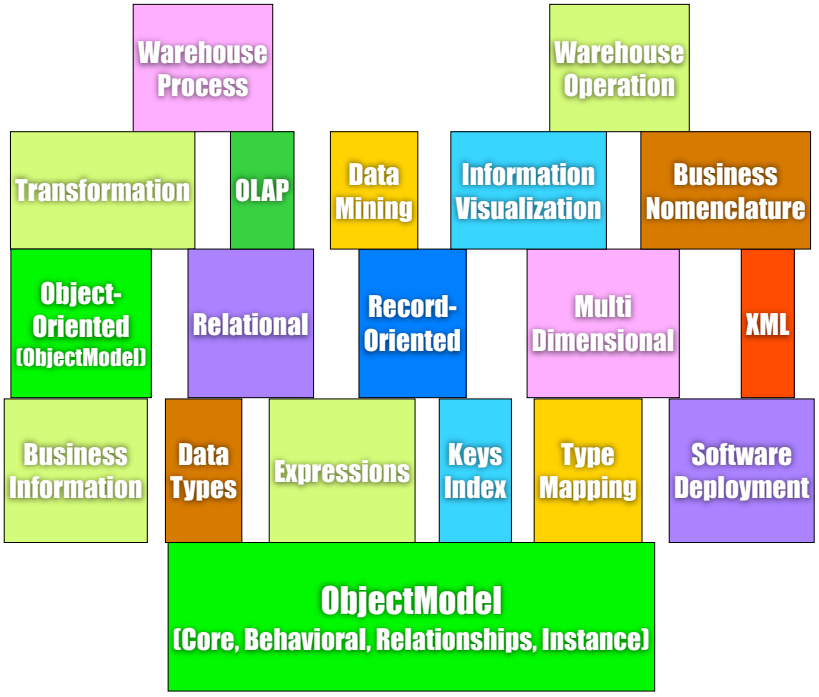
# Overview of CWM Metamodel

Warehouse Management

Analysis

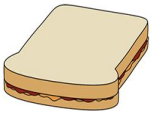
Resources

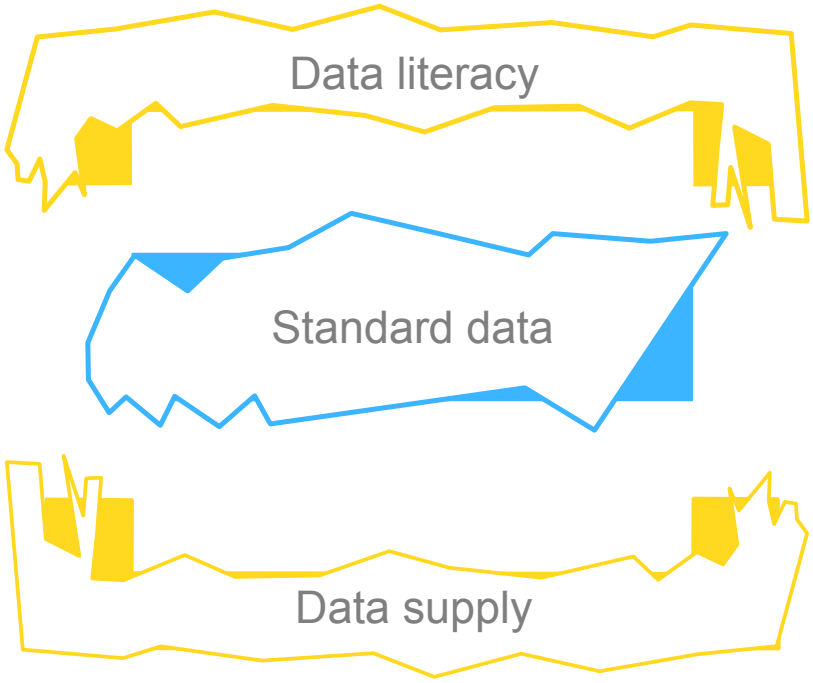
Foundation



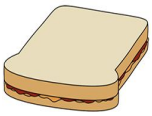

http://www.omg.org/technology/documents/modeling_spec_catalog.htm

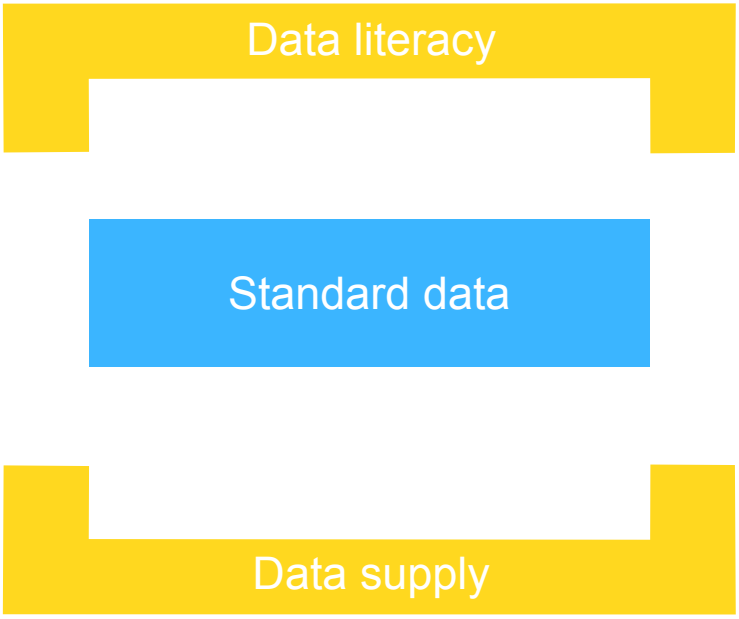


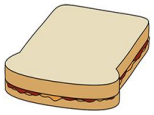



Data literacy
Standard data
Data supply






Data literacy
Standard data
Data supply



*This cannot happen without investments in engineering and architecture!*

Data literacy

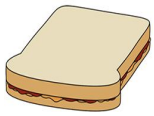
Standard data

Data supply

**Quality engineering/ architecture work products do not happen accidentally!**



## Leverage point - high performance automation



*This cannot happen without investments in data engineering and architecture!*

Data literacy

Standard data

Data supply

**These are necessary but insufficient prerequisites to successful AI augmentation**

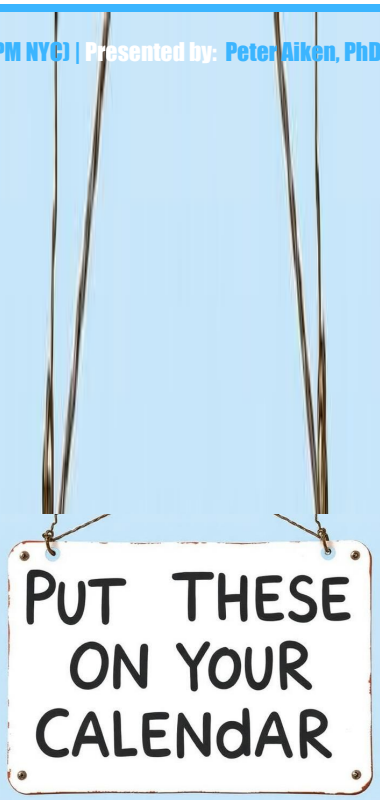
# Upcoming Events

Time: 19:00 UTC (2:00 PM NYC) | Presented by: Peter Aiken, PhD

## Metadata Matters: Powering AI Data Discovery and Context
13 *October 2026*

## Data Quality by Design: Ensuring Reliable Inputs for AI Systems
*10 November 2026*

## Data Governance/Stewardship in Action: Governing AI's Data for Long-Term Value
*8 December 2026*



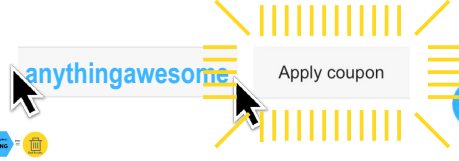

Brought to you by: DATA-ED ONLINE, DATAVERSITY

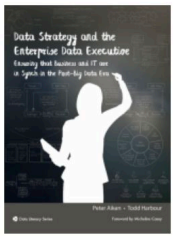
Note: In this .pdf, clicking any webinar title opens the registration link



# Event Pricing on Peter's Books

- 20% off directly from the publisher on select titles
- My 'Book Store' @ https://anythingawesome.com/books-overview.html
- Enter the code "anythingawesome" at the Technics bookstore checkout where it says to "Apply Coupon"

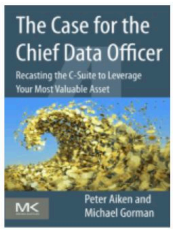



**Data Strategy and the Enterprise Data Executive**

**Ensuring** that Business and IT are in Synch in the Post-Big Data Era

Learn More of Data Strategy




**The Case for the Chief Data Officer**

Recasting the C-Suite to Leverage Your Most Valuable Asset

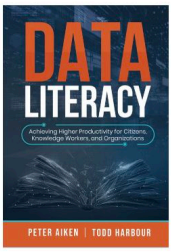
(The Chinese Translation Title is: Chief Data Officer Combat)

Learn More of the Case for Data Leadership


**Monetizing Data Management**

17 Case Studies Illustrating How Data Leveraging (Big and Small) Can Produce Quantifiable Results That Are of Keen Interest to C-Suite Occupants

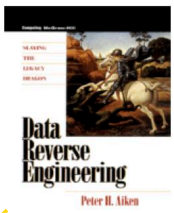
Learn More of Monetizing Data




**Data Literacy: Achieving Higher Productivity for Citizens, Knowledge Workers, and Organizations**

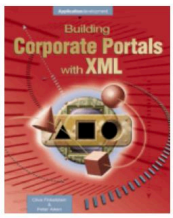
Citizens and organizations need to improve their data literacy to 'do more with data'

Learn More of Data Literacy



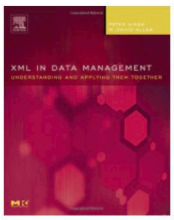

**Data Reverse Engineering**

Learn More of Data Reverse Engineering



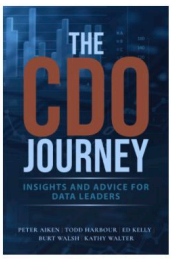

**Building Corporate Portals with SML**

Learn More Corporate Portals (& XML)




**XML in Data Management**

Learn More of XML and Data Management




**The CDO Journey: Insights and Advice for Data Leaders**

Learn More of the CDO Journey

Critical Design Review?


Executive Data Literacy Training?

Collaboration?


Peter.Aiken@AnythingAwesome.com +1.804.382.5957

Independent Verification & Validation

Reverse Engineering Expertise?

Hiring Assistance?

# Thank You!

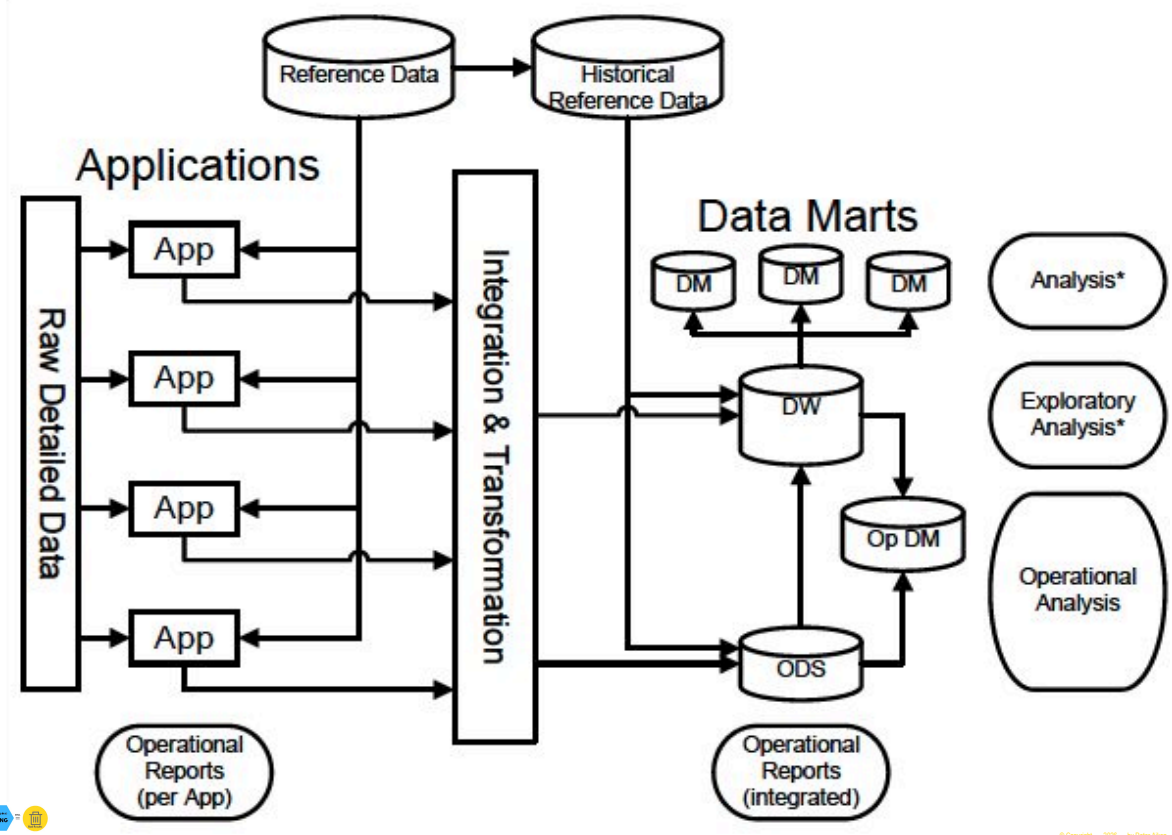
Use your data more strategically?

Tool/automation evaluation?

Book a call with Peter to discuss anything - https://anythingawesome.com/OfficeHours.html



## Corporate Information Factory Architecture



# Corporate Information Factory Architecture

| | |
|---|---|
| Raw Detailed data | Operational / Transactional Application data of the enterprise. The raw detailed data provides the source data to be integrated into the Operational Data Store (ODS) and DW components. They can also be in database or other storage or file format. |
| Integration and Transformation | This layer of the architecture is where the un-integrated data from the various application sources stores is combined / integrated and transformed into the corporate representation in the DW. |
| Reference Data | Reference data was a precursor to what is currently referred to as Master Data Management. The purpose was to allow common storage and access for important and frequently used common data. Focus and shared understanding on data upstream of the Data Warehouse simplifies the integration task in the DW. |
| Historical Reference Data | When current valued reference data is necessary for transactional applications, and at the same time it is critical to have accurate integration and presentation of historical data, it is necessary to capture the reference data that was in place at any point in time. For more discussion on reference data, see Chapter 8 Master and Reference Data Management. |

from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International



from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International

# Corporate Information Factory Architecture

| | |
|---|---|
| Operational Data Store (ODS) | The focus of data integration is meeting operating and classically operational reporting needs that require data from multiple operational systems. The main distinguishing data characteristics of an ODS compared to a DW include current-valued vs. DW historical data and volatile vs. DW non-volatile data.<br><br>Note: ODS is an optional portion of the overall CIF architecture, dependent upon specific operational needs, and acknowledged as a component that many businesses omit. |
| Operational Data Mart (Oper-Mart) | A data mart focuses on tactical decision support. Distinguishing characteristics include current-valued vs. DW historical data, tactical vs. DW strategic analysis, and sourcing of data from an ODS rather than just the DW. The Oper-Mart was a later addition to the CIF architecture. |
| Data Warehouse (DW) | The DW is a large, comprehensive corporate resource, whose primary purpose is to provide a single integration point for corporate data in order to serve management decision, and strategic analysis and planning.<br><br>The data flows into a DW from the application systems and ODS ,and flows out to the data marts, usually in one direction only. Data that needs correction is rejected, corrected at its source, and re-fed through the system. |
| Data Marts (DM) | The purpose of the data marts is to provide for DSS / information processing and access that is customized and tailored for the needs of a particular department or common analytic need. |

# Corporate Information Factory Architecture

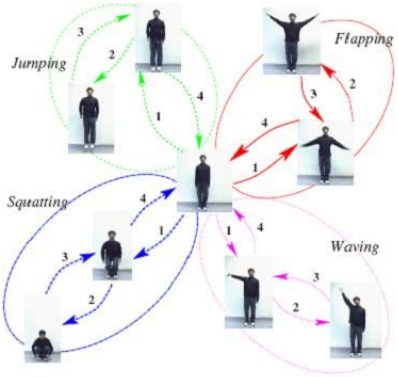
| | | |
|---|---|---|
| Applications | Isolated Operational Reports | Limited to data within one application instance |
| ODS | Integrated Operational Reports | Reports requiring data from multiple source systems. Typically, they have more operational than analytical orientation, with little historical data. |
| DW | Exploratory Analysis | The complete set of corporate data allows for discovery of new relationships and information. Many BI data mining tools work with flat-file extracts from the DW, which can also offload the processing burden from the DW. |
| Oper-Mart | Tactical Analytics | Analytic reporting based on current-values with a tactical focus. Dimensional data modeling techniques employed. |
| Data Mart | Analytics – classical management decision support, and Strategic Analytics | Inmon's early focus was on "departmental analysis", which was experientially true for real-world organizational issues, such as political and funding expediency. Later work expanded concepts to common-analytic needs crossing departmental boundaries. |



# Goals and Principles

1. To support and enable effective business analysis and decision making by knowledgeable workers
2. To build and maintain the environment/infrastructure to support business intelligence activities, specifically leveraging all the other data management functions to cost effectively deliver consistent integrated data for all BI activities

# Activities




- Understand BI information needs
- Define and maintain the DW/BI architecture
- Process data for BI
- Implement data warehouse/data marts
- Implement BI tools and user interfaces
- Monitor and tune DW processes
- Monitor and tune BI activities and performance

from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International



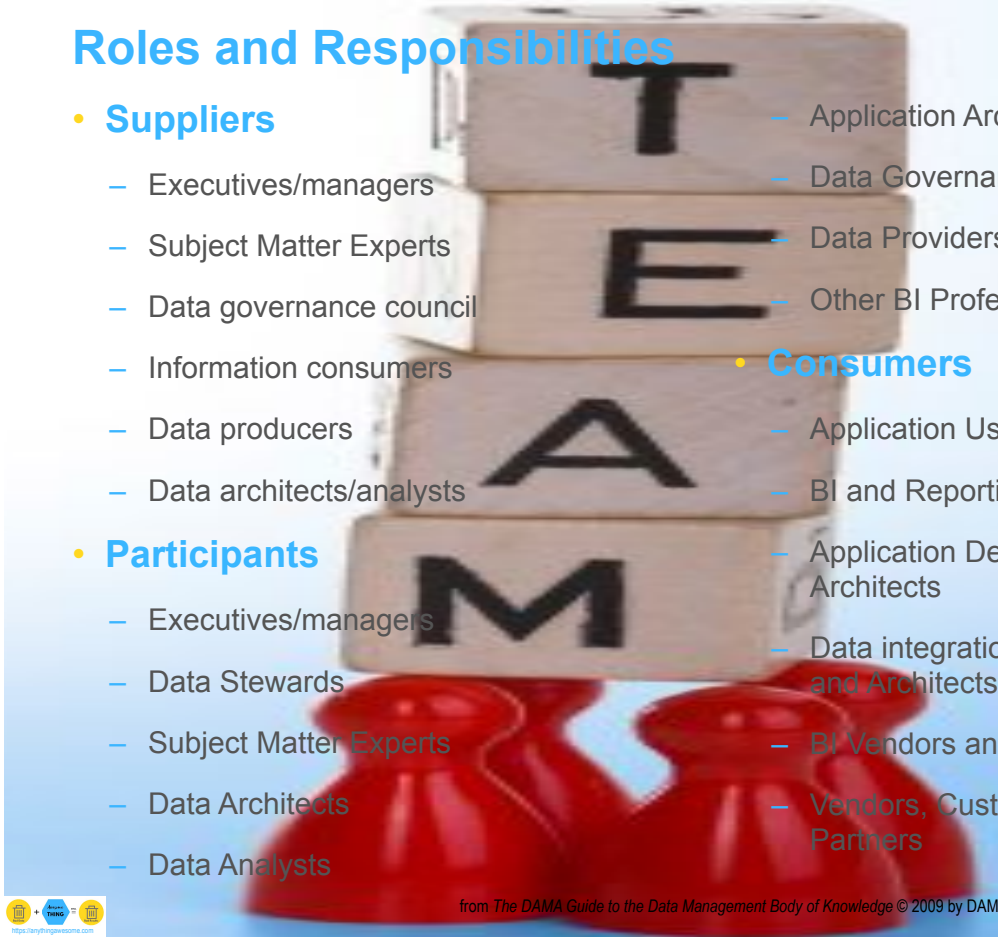
# Primary Deliverables

- DW/BI Architecture
- Data warehouses, marts, cubes etc.
- Dashboards-scorecards
- Analytic applications
- Files extracts (for data mining, etc.)
- BI tools and user environments
- Data quality feedback mechanism/loop

from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International

# Roles and Responsibilities

- **Suppliers**
  - Executives/managers
  - Subject Matter Experts
  - Data governance council
  - Information consumers
  - Data producers
  - Data architects/analysts
- **Participants**
  - Executives/managers
  - Data Stewards
  - Subject Matter Experts
  - Data Architects
  - Data Analysts
  - Application Architects
  - Data Governance Council
  - Data Providers
  - Other BI Professionals
- **Consumers**
  - Application Users
  - BI and Reporting Users
  - Application Developers and Architects
  - Data integration Developers and Architects
  - BI Vendors and Architects
  - Vendors, Customers and Partners




from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International



from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International

# Technology

- ETL
- Change Management Tools
- Data Modeling Tools
- Data Profiling Tools
- Data Cleansing Tools
- Data Integration Tools
- Reference Data Management Applications
- Master Data Management Applications
- Process Modeling Tools
- Meta-data Repositories
- Business Process and Rule Engines

## Guiding Principles


- Obtain executive commitment and support.
- Secure business SMEs.
- Be business focused and driven. Let the business drive the prioritization.
- Demonstrate data quality is essential.
- Provide incremental value.
- Transparency and self service.
- One size does not fit all: Find the right tools and products for each of your segments.
- Think and architect globally, act and build locally.
- Collaborate with and integrate all other data initiatives, especially those for data governance, data quality and metadata.
- Start with the end in mind.
- Summarize and optimize last, not first.





## References

Adamson, Christopher. Mastering Data Warehouse Aggregates: Solutions for Star Schema Performance. John Wiley & Sons, 2006. ISBN 0-471-77709-9. 345 pages.

Adamson, Christopher and Michael Venerable. Data Warehouse Design Solutions. John Wiley & Sons, 1998. ISBN 0-471-25195-X. 544 pages.

Adelman, Sid and Larissa T. Moss. Data Warehouse Project Management. Addison-Wesley Professional, 2000. ISBN 0-201-61635-1. 448 pages.

Adelman, Sid and others. Impossible Data Warehouse Situations: Solutions from the Experts. Addison-Wesley, 2002. ISBN 0-201-76033-9. 432 pages.

Brackett, Michael. The Data Warehouse Challenge: Taming Data Chaos. New York: John Wiley & Sons, 1996. ISBN 0-471-12744-2. 579 pages.

Caserta, Joe and Ralph Kimball. The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming and Delivering Data. John Wiley & Sons, 2004. ISBN 0-764-56757-8. 525 pages.

# References


Correy, Michael J. and Michael Abby. Oracle Data Warehousing: A Practical Guide to Successful Data Warehouse Analysis, Build and Roll-Out. TATA McGraw-Hill, 1997. ISBN 0-074-63069-5.

Covey, Stephen R. The 7 Habits of Highly Effective People. Free Press, 2004. ISBN 0743269519. 384 Pages.

Dyche, Jill. E-Data: Turning Data Into Information With Data Warehousing. Addison-Wesley, 2000. ISBN 0-201-65780-5. 384 pages.

Gill, Harjinder S. and Prekash C. Rao. The Official Guide To Data Warehousing. Que, 1996. ISBN 0-789-70714-4. 382 pages.

Hackney, Douglas. Understanding and Implementing Successful Data Marts. Addison Wesley, 1997. ISBN 0-201-18380-3. 464 pages.

Imhoff, Claudia, Nicholas Galemmo and Jonathan G. Geiger. Mastering Data Warehouse Design: Relational and Dimensional Techniques. John Wiley & Sons, 2003. ISBN 0-471-32421-3. 456 pages.

Imhoff, Claudia, Lisa Loftis and Jonathan G. Geiger. Building the Customer-Centric Enterprise: Data Warehousing Techniques for Supporting Customer Relationship Management. John Wiley & Sons, 2001. ISBN 0-471-31981-3. 512 pages.

Inmon, W. H. Building the Data Warehouse, 4th Edition. John Wiley & Sons, 2005. ISBN 0-764-59944-5. 543 pages.

Inmon, W. H. Building the Operational Data Store, 2nd edition. John Wiley & Sons, 1999. ISBN 0-471-32888-X. 336 pages.

Inmon, W. H., Claudia Imhoff and Ryan Sousa. The Corporate Information Factory, 2nd edition. John Wiley & Sons, 2000. ISBN 0-471-39961-2. 400 pages.

Inmon, W. H. and Richard D. Hackathorn. Using the Data Warehouse. Wiley-QED, 1994. ISBN 0-471-05966-8. 305 pages.

Inmon, William H., John A. Zachman and Jonathan G. Geiger. Data Stores, Data Warehousing and the Zachman Framework. McGraw-Hill, 1997. ISBN 0-070-31429-2. 358 pages.

Kimball, Ralph and Margy Ross. The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling, 2nd edition. New York: John Wiley & Sons, 2002. ISBN 0-471-20024-7. 464 pages.

Kimball, Ralph, Laura Reeves, Margy Ross and Warren Thornwaite. The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing and Deploying Data Warehouses. John Wiley & Sons, 1998. ISBN 0-471-25547-5. 800 pages.

Kimball, Ralph and Richard Merz. The Data Webhouse Toolkit: Building the Web-Enabled Data Warehouse. John Wiley & Sons, 2000. ISBN 0-471-37680-9. 416 pages.

Mattison, Rob, Web Warehousing & Knowledge Management. McGraw Hill, 1999. ISBN 0-070-41103-4. 576 pages.

Morris, Henry. *Analytic Applications and Business Performance Management*. DM Review Magazine, March, 1999. www.dmreview.com. Note: www.dmreview.com is now www.information-management.com.

Moss, Larissa T. and Shaku Atre. Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications. Addison-Wesley, 2003. ISBN 0-201-78420-3. 576 pages.

Poe, Vidette, Patricia Klauer and Stephen Brobst. Building A Data Warehouse for Decision Support, 2nd edition. Prentice-Hall, 1997. ISBN 0-137-69639-6. 285 pages.

Ponniah, Paulraj. Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals. John Wiley & Sons – Interscience, 2001. ISBN 0-471-41254-6. 528 pages.

Westerman, Paul. Data Warehousing: Using the Wal-Mart Model. Morgan Kaufman, 2000. ISBN 155860684X. 297 pages.

### 9.5.2 Business Intelligence

Biere, Mike. Business Intelligence for the Enterprise. IBM Press, 2003. ISBN 0-131-41303-1. 240 pages.

Eckerson, Wayne W. Performance Dashboards: MEassuring, Monitoring, and Managing Your Business. Wiley, 2005. ISBN-10: 0471724173. 320 pages.

Bischoff, Joyce and Ted Alexander. Data Warehouse: Practical Advice from the Experts. Prentice-Hall, 1997. ISBN 0-135-77370-9. 428 pages.

Howson, Cindi. "The Business Intelligence Market". http://www.biscorecard.com/. Requires annual subscription to this website.

Malik, Shadan. Enterprise Dashboards: Design and Best Practices for IT. Wiley, 2005. ISBN 0471738069. 240 pages.

Moss, Larissa T., and Shaku Atre. Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications. Addison-Wesley, 2003. ISBN 0-201-78420-3. 576 pages.

Vitt, Elizabeth, Michael Luckevich and Stacia Misner. **Business Intelligence**. Microsoft Press, 2008. ISBN 073562660X. 220 pages.

### 9.5.3 Data Mining

Cabena, Peter, Hadjnian, Stadler, Verhees and Zanasi. Discovering Data Mining: From Concept to Implementation. Prentice Hall, 1997. ISBN-10: 0137439806




# Additional References

- http://www.information-management.com/infodirect/20050909/1036703-1.html
- http://www.agiledata.org/essays/dataWarehousingBestPractices.html
- http://www.cio.com/article/150450/Five_Key_Business_Intelligence_Trends_You_Need_to_Know?page=2&taxonomyId=3002
- http://www.computerworld.com/s/article/9228736/Business_Intelligence_and_analytics_Conquering_Big_Data?taxonomyId=9
- http://www.enterpriseirregulars.com/5706/the-top-10-trends-for-2010-in-analytics-business-intelligence-and-performance-management/
- http://www.itbusinessedge.com/cm/blogs/vizard/taking-the-analytics-pressure-off-the-data-warehouse/?cs=50698
- http://www.informationweek.com/news/software/bi/240001922

# References & Recommended Reading

### 11.4.1 General Reading


Brathwaite, Ken S. Analysis, Design, and Implementation of Data Dictionaries. McGraw-Hill Inc., 1988. ISBN 0-07-007248-5. 214 pages.

Collier, Ken. Executive Report, Business Intelligence Advisory Service, *Finding the Value in Metadata Management* (Vol. 4, No. 1), 2004. Available only to Cutter Consortium Clients, http://www.cutter.com/bia/fulltext/reports/2004/01/index.html.

Hay, David C. Data Model Patterns: A Metadata Map. Morgan Kaufmann, 2006. ISBN 0-120-88798-3. 432 pages.

Hillmann, Diane I. and Elaine L. Westbrooks, editors. Metadata in Practice. American Library Association, 2004. ISBN 0-838-90882-9. 285 pages.

Inmon, William H., Bonnie O'Neil and Lowell Fryman. Business Metadata: Capturing Enterprise Knowledge. 2008. Morgan Kaufmann ISBN 978-0-12-373726-7. 314 pages.

Marco, David, Building and Managing the Meta Data Repository: A Full Life-Cycle Guide. John Wiley & Sons, 2000. ISBN 0-471-35523-2. 416 pages.

Marco, David and Michael Jennings. Universal Meta Data Models. John Wiley & Sons, 2004. ISBN 0-471-08177-9. 478 pages.

Poole, John, Dan Change, Douglas Tolbert and David Mellor. Common Warehouse Metamodel: An Introduction to the Standard for Data Warehouse Integration. John Wiley & Sons, 2001. ISBN 0-471-20052-2. 208 pages.




# References, cont'd


Poole, John, Dan Change, Douglas Tolbert and David Mellor. Common Warehouse Metamodel Developer's Guide. John Wiley & Sons, 2003. ISBN 0-471-20243-6. 704 pages.

Ross, Ronald. Data Dictionaries And Data Administration: Concepts and Practices for Data Resource Management. New York: AMACOM Books, 1981. ISN 0-814-45596-4. 454 pages.

Tannenbaum, Adrienne. Implementing a Corporate Repository, John Wiley & Sons, 1994. ISBN 0-471-58537-8. 441 pages.

Tannenbaum, Adrienne. Metadata Solutions: Using Metamodels, Repositories, XML, And Enterprise Portals to Generate Information on Demand. Addison Wesley, 2001. ISBN 0-201-71976-2. 528 pages.

Wertz, Charles J. The Data Dictionary: Concepts and Uses, 2nd edition. John Wiley & Sons, 1993. ISBN 0-471-60308-2. 390 pages.

# References, cont'd


### 11.4.2 Meta-data in Library Science

Baca, Murtha, editor. Introduction to Metadata: Pathways to Digital Information. Getty Information Institute, 2000. ISBN 0-892-36533-1. 48 pages.

Hillmann, Diane I., and Elaine L. Westbrooks. Metadata in Practice. American Library Association, 2004. ISBN 0-838-90882-9. 285 pages.

Karpuk, Deborah. METADATA: From Resource Discovery to Knowledge Management. Libraries Unlimited, 2007. ISBN 1-591-58070-6. 275 pages.

Liu, Jia. Metadata and Its Applications in the Digital Library. Libraries Unlimited, 2007. ISBN 1-291-58306-6. 250 pages.

### 11.4.3 Geospatial Meta-data Standards

http://www.fgdc.gov/metadata/geospatial-metadata-standards.




# References, cont'd


### 11.4.4 ISO Meta-data Standards

ISO Standards Handbook 10, Data Processing—Vocabulary, 1982.

ISO 704:1987, Principles and methods of terminology.

ISO 1087, Terminology—Vocabulary.

ISO 2382-4:1987, Information processing systems—Vocabulary part 4.

ISO/IEC 10241:1992, International Terminology Standards—Preparation and layout.

FCD 11179-2, Information technology—Specification and standardization of data elements - Part 2: Classification for data elements.

ISO/IEC 11179-3:1994, Information technology—Specification and standardization of data elements - Part 3: Basic attributes of data elements.

ISO/IEC 11179-4:1995, Information technology—Specification and standardization of data elements - Part 4: Rules and guidelines for the formulation of data definitions.

ISO/IEC 11179-5:1995, Information technology—Specification and standardization of data elements - Part 5: Naming and identification principles for data elements.

ISO/IEC 11179-6:1997, Information technology—Specification and standardization of data elements - Part 6: Registration of data elements.

# More Analytical Applications

### Typical Analytical Applications for Internal Processes

**Activity-based costing (ABC).** The first step in activity-based management is to allocate costs accurately to aspects of the business such as customers, processes, or distribution channels; models incorporating activities, materials, resources, and product-offering components then allow optimization based on cost and prediction of capacity needs.

**Bayesian inference (e.g., to predict revenues).** A numerical estimate of the degree of belief in a hypothesis before and after evidence has been observed.

**Biosimulation (e.g., in pharmaceutical "in silico" research).** Manipulation of biological parameters using mathematics and/or rule bases to model how cells or other living entities react to chemical or other interventions.

**Combinatorial optimization (e.g., for optimizing a product portfolio).** The efficient allocation of limited resources to yield the best solution to particular objectives when the values of some or all of the variables (e.g., a given number of people) must be integers (because people can't be split into fractions) and there are many possible combinations. Also called *integer programming*.



# More Analytical Applications (continued)

**Constraint analysis (e.g., for product configuration).** The use of one or more constraint satisfaction algorithms to specify the set of feasible solutions. Constraints are programmed in rules or procedures that produce solutions to particular configuration and design problems using one or more constraint satisfaction algorithms.

**Experimental design (e.g., for Web site analysis).** In the simplest type of experiment, participants are randomly assigned to two groups that are equivalent to each other. One group (the program or treatment group) gets the program and the other group (the comparison or control group) does not. If the program results in statistically significant differences in the outcome variable, it is assumed to have the hypothesized effect.

**Future-value analysis.** The decomposition of market capitalization into current value (extrapolation of existing monetary returns) and future value, or expectations of future growth.

**Monte Carlo simulation (e.g., for R&D project valuation).** A computerized technique used to assess the probability of certain outcomes or risks by mathematically modeling a hypothetical event over multiple trials and comparing the outcome with predefined probability distributions.

# More Analytical Applications (continued)

**Multiple regression analysis (e.g., to determine how nonfinancial factors affect financial performance).** A statistical technique whereby the influence of a set of independent variables on a single dependent variable is determined.

**Neural network analysis (e.g., to predict the onset of disease).** Systems modeled on the structure and operation of the brain, in which the state of the system is modified by training until it can discriminate between the classes of inputs; used on large databases. Typically, a neural network is initially "trained," or fed large amounts of data and rules about data relationships—for example, "A grandfather is older than a person's father."

**Textual analysis (e.g., to assess intangible capabilities).** Analysis of the frequency, semantic relationships, and relative importance of particular terms, phrases, and documents in online text.

**Yield analysis (e.g., in semiconductor manufacturing).** Employing basic statistics (mean, median, standard deviation, etc.) to understand yield volume and quality, and to compare one batch of items with another—often displayed visually.