Stockholms universitet

Information Systems: Collection Management Systems, Business Intelligence, Big Data Analytics and IoT

Erik Perjons

Collection Management Systems

What is a Collection Management System?

- A collection management system (CMS) - is a system that organizes, controls, and manages collections objects (artefacts) by tracking all information related to and about those objects
- A collection management system is used in museums, archives, galleries and libraries.

[Sully, Perian (8 July 2006). "Inventory, Access, Interpretation: The Evolution of Museum Collection Software”]

What is a Collection Management System?

- Sometimes they are called collection information systems, or content management system
- Sometimes practitioners and researchers **see content management system and collection information systems as different type of systems**. Content management systems can be seen as a system that can be used to present information, such as text, images, documents, videos, on the web, for example using the tool WorldPress

[Sully, Perian (8 July 2006). "Inventory, Access, Interpretation: The Evolution of Museum Collection Software"]

Fuctionality of a CMS?

- Object entry
- Acquisition
- Inventory control
- Location and movement control
- Catalog description
- Conservation management
- Risk management
- Insurance management and valuation control
- Exhibition management
- Dispatch/shipping/transport
- Loaning and borrowing
- Deaccessioning and disposal

[Canadian Heritage Information Network (2012). Collections Management Software Criteria Checklist]

How to aquire an approriate CMS?

- What do you need to do to aquire an approriate CMS for your organization?
- How to carry out the requirement engineering?
- Buy or develop a CMS? Benefit and drawbacks?
- Software as a Service? Benefit and drawbacks?
- Web solution?

[Sully, Perian (8 July 2006). "Inventory, Access, Interpretation: The Evolution of Museum Collection Software”]

What other types of systems?

- What other types of systems are needed in a museum or a gallery?

Stockholms universitet

Business Intelligence

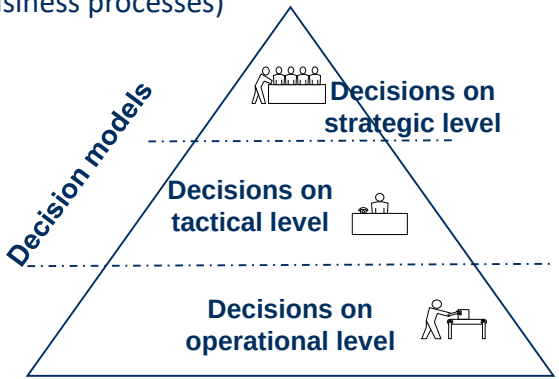
What is Business Intelligence?

- Business intelligence (BI) - is an umbrella term that is commonly used to describe the technologies, applications, and processes for gathering, storing, accessing, and analyzing data to help users make better decisions

Business intelligence – an overview

Goals
(vision, enterprise goals, objectives)

Means
(mission, strategies, tactics, business processes)



BI systems/tools

- Dashboard systems
- Analytical decision support tools
- Expert systems
- Data warehouse systems
- OLAP tools
- Data mining tools
- Reporting tools

System integration tools

- ETL tools

BI related methods

- Performance management
- Six sigma
- Balanced scorecard
- Lean
- Activity based costing
- Data and process mining
- Decision and risk analysis

Visualization

- Visualization techniques

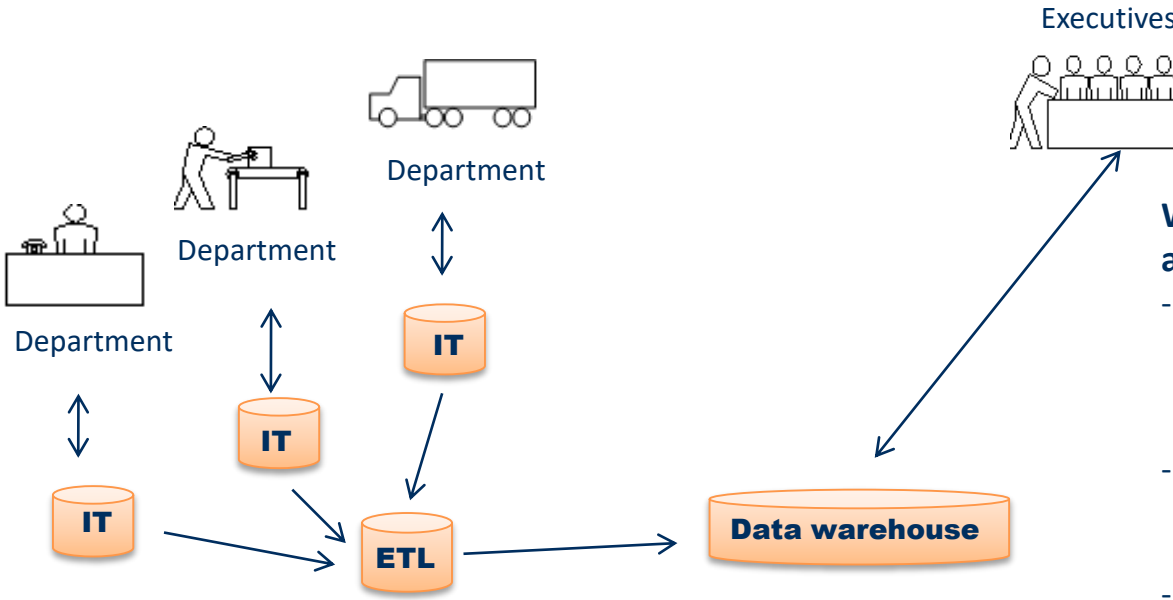
Other areas related

- Business processes
- Business rules
- Models for decision
- Data governance
- Master data management
- Meta data management
- Org structures
- IT architecture
- Data repository
- Cloud computing
- Security
- Business case/ROI

Systems supporting decision making and data sources

Operational systems: System that support the daily business, such as business systems (ERPs), BPM system, CRM system, etc

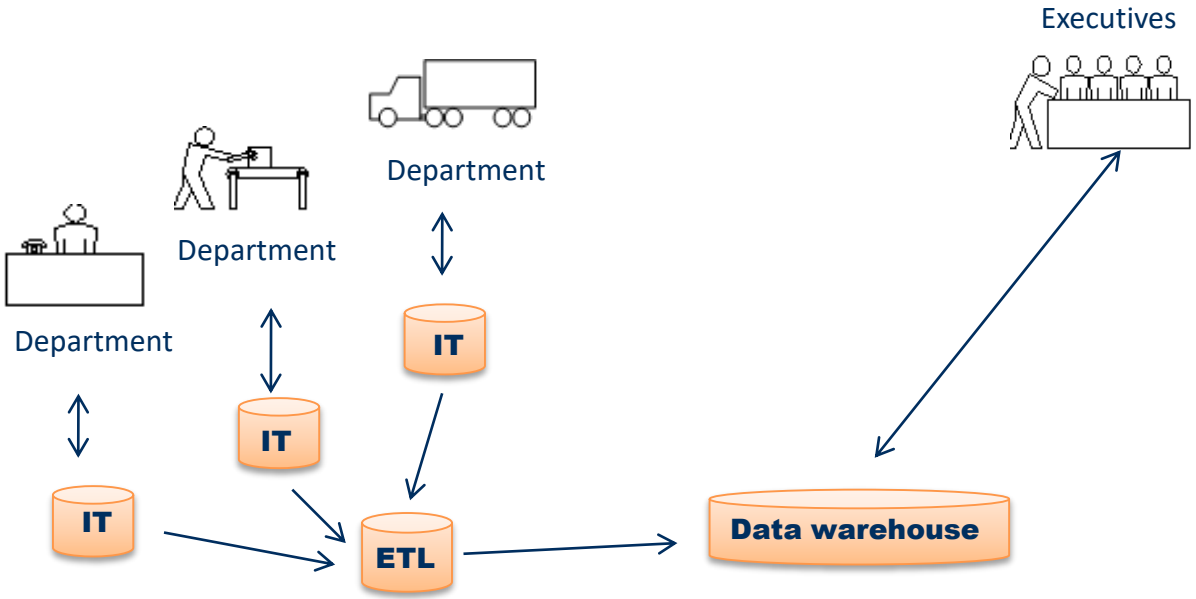
BI solution: Data Warehouse



What problem does DW address:

- Data about customers, artefacts/products, museusm are spread out in different IT systems
- Data in th different systems has different defintions
- The IT systems can be hard to query

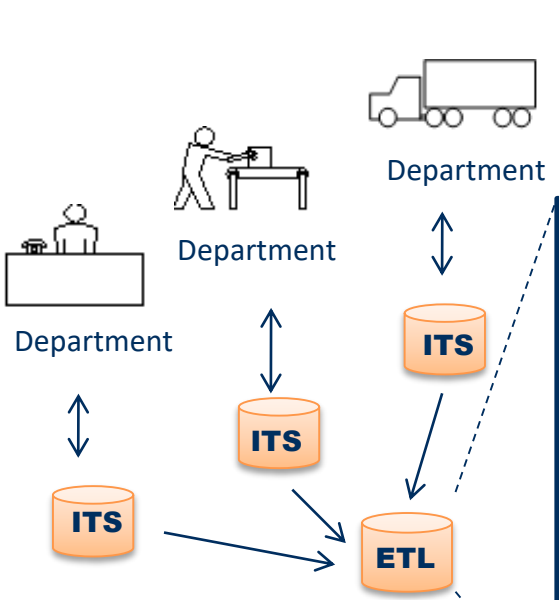
BI solution: Data Warehouse



Questions users can ask the DW system:

- Which museums have most visitors?
- Which parts of the museum have most vistors?
- Which time of the day do we have most visitors?
- Which campains results in most visitors?
- Which products are sold most?

BI solution: Data Warehouse and ETL



ETL – a tool / system that support "extract", "transform", and "load" processes (also called ETL processes), that is, a tool that supports extraction of data from the operational IT systems (ITS); supports transformation of data into a common information structure; and supports data loading into DW.

The ETL tool can also verify that data to be loaded into DW is correct by correcting the cleansing info. This can be done by linking to address registers (and, for example, correcting zip codes) and standards (such as correcting physicians coding of ICD-10 actions).

Incorrect data can be corrected in the tool.

Big Data Analytics/Data Science

What is Big Data?

- Big data - is a key enabler of a new discipline called data science that seeks to **leverage new sources of structured and unstructured data, coupled with predictive and prescriptive analytics**, to uncover new variables and metrics that are **better predictors of performance**

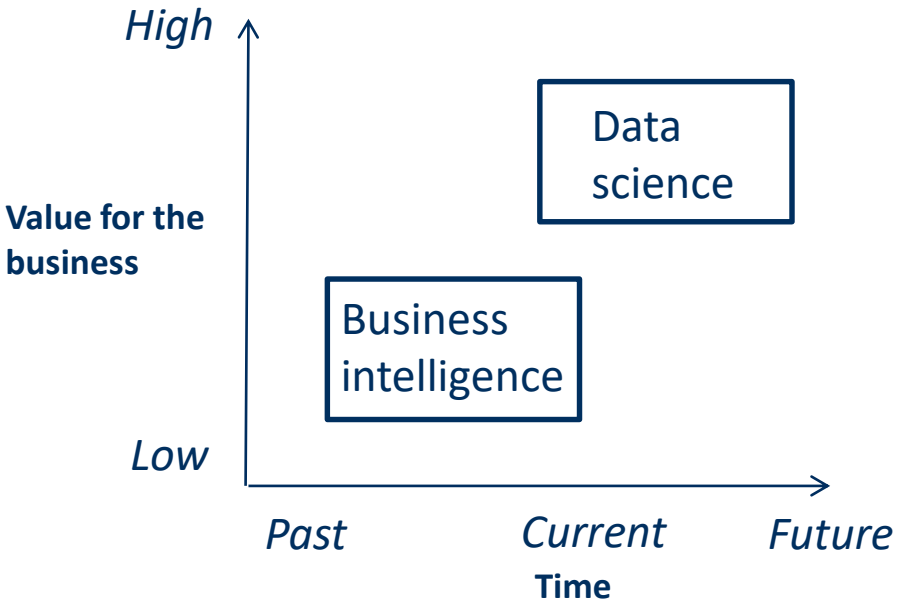
[Lewis (2004) Moneyball: The Art of Winning an Unfair Game]

What is Data Science?

Data science - is about finding new variables and metrics that are better predictors of performance

[Lewis (2004) Moneyball: The Art of Winning an Unfair Game]

Business intelligence vs. Data science



(Bill Schmarzo, Big Data MBA, Wiley, 2016)

BI vs. Data science: The questions are different

Business Intelligence

- ***Focus on descriptive analytics:*** ”What happend?” type of questions: How many units of products X did we sell in Jan 2017

Data Science

- ***Focus on predictive analytics:*** ”What is likely to happend? type of questions: How many units of products X will we sell in Jan 2018?
- ***Focus on prescriptive analytics:*** ”What should we do?” type of question: How many components A, B, C should I order to support the sales of product X?

BI vs. Data science: The views on business are different 1(2)

Business Intelligence

- Aggregated data on business entities, such as customers, products

Data Science

- Build **analytic profiles** on each business entity. Example of business entities are customers, partners/suppliers, devices, machines
- For exampel, analytic profiles for customers could be used for managing customer rentension/attrite rate

(Bill Schmarzo, Big Data MBA, Wiley, 2016)

BI vs. Data science: The views on business are different 2(2)

- For example a customer profile could include:
  - **Demografic information** (e.g. name, addresses, age, children, income level, value of home)
  - **Transactional metrics** (e.g. number of purchases, purchase amounts, product purchase)
  - **Social media metrics** (e.g. social media comments)
  - **Behaviour grouping** (e.g. favorite producs/services, frequency, length and recent store visits)
  - **Classifications** (e.g. lifestyle classification (heavy traveler, light gym visiter))
  - **Association rules** (e.g. usage patterns such as when buying x also buy y, when use x och buy y)
  - **Scores** (e.g. loyality score, product usage score)

(Bill Schmarzo, Big Data MBA, Wiley, 2016)

BI vs. Data science: The analytic approaches are different 1(2)

BI analytic approach

- Step 1: Pre-build a data model (Schema on load)
- Step 2: Make use of (visualisation) tools that automatically generated SQL commands from drag and drop using attributes/dimensions/facts
- Step 3: Make use of the generated SQL commands to generate reports automatically

Data science analytic approach

- Step 1: Define hypothesis (test/prediction)
- Step 2: Gather data (Data Lake, Hadoop)
- Step 3: Build data model (Schema on query)
- Step 4: Build analytic models (Data mining, Machine learning, SAS, R)
- Step 5: Evaluate model goodness of fit

(Bill Schmarzo, Big Data MBA, Wiley, 2016)

BI vs. Data science: The analytic approaches are different 2(2)

Schema on load

– a schema must be built prior to loading data into the data warehouse

Schema on query

– a schema is defined as needed based on data being used, and the data scientist will go through different versions of the schema until finding a schema that support the analytical model

(Bill Schmarzo, Big Data MBA, Wiley, 2016)

Stockholms universitet

How can you identify opportunities with big data analytics for an organization?

Method for identifying opportunities with big data och data analysis

1. Understand the main goals, strategies, activities and concepts – that make the organization successful

2. Identify central business initiatives

3. Brainstorm big data solutions for the business initiatives

4. Design use case that support the initiatives – and define requirements on data and data analysis

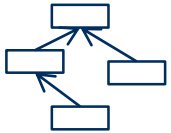
5. Make proof-of-concepts of prioritized use cases – analyze the business value (ROI) and make plans to manage data and analysis for each use case

6. Design och implement the solution

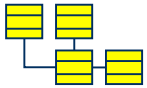
(Bill Schmarzo, Big Data. Understand How Data Powers Big Business, Wiley, 2013)

Method for identifying opportunities with big data och data analysis

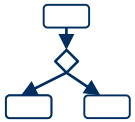
1. Understand the main goals, strategies, activities and concepts – that make the organization successful



Goal-means-strategy-model



Conceptual model – defining central business terms/concepts



Identify key activities

2. Identify central business initiatives

- Read business reports
- Read presentations by executives
- Interview key employees

(Bill Schmarzo, Big Data. Understand How Data Powers Big Business, Wiley, 2013)

Method for identifying opportunities with big data och data analysis

3. Brainstorm big data solutions for the business initiatives

4. Design use case that support the initiatives – and define requirements on data and data analysis

Four ways that big data and analytics cän impact business initiatives

- ”Mining” more detailed transaction data
- Integrate unstructured internal and external data
- Improve real time delivery of data
- Apply different forms of predictive analytics

Break down business initiatives in use cases (that is, functions that supports the business initiatives)

- Stakeholders
- Central decisions, questions that the use cases shall support – that is user requirements
- Requirements on data
- Requirements on data analysis - algorithms/models

(Bill Schmarzo, Big Data. Understand How Data Powers Big Business, Wiley, 2013)

Method for identifying opportunities with big data och data analysis

5. Make proof-of-concepts of prioritized use cases – analyze the business value (ROI) and make plans to manage data and analysis for each use case

6. Design och implement the solution

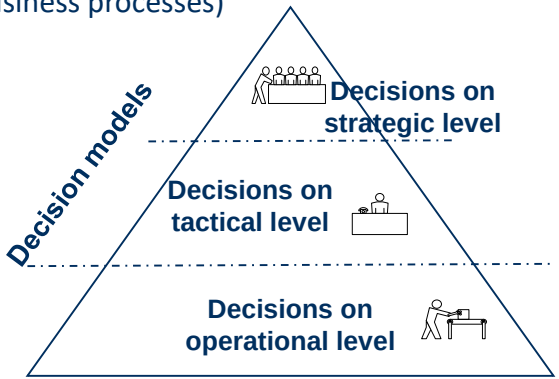
Proof-of-concept for each prioritized use case
Develop a business case (ROI/cost-benefit)
Make a plan to manage data – manage source systems, transformations, cleaning data, master data, etc
Make a plan to adapt and test analytical models

(Bill Schmarzo, Big Data. Understand How Data Powers Big Business, Wiley, 2013)

Big data/Data science – an overview

Goals
(vision, enterprise goals, objectives)

Means
(mission, strategies, tactics, business processes)



Big data related systems/techniques

- Hadoop
- MapReduce YARN/SPARK
- SAS/R

Additional data sources

- Social media
- Internet of Things
- Web
- Video/Camera
- Machines/Devices
- Maps
- Satellites

Big Data related methods for data analysis

- Data mining
- Machine learning
- Language technology
- Process mining
- Web mining

Processes for data analysis

- CRISP-DM
- Knowledge Discovery in Databases (KDD)

Visualization

- Visualization techniques

Other areas related

- Business processes
- Business rules
- Models for decision
- Data governance
- Master data management
- Meta data management
- Org structures
- IT architecture
- Data repository
- Cloud computing
- Security
- Business case/ROI

Systems supporting decision making and data sources

Operational systems: System that support the daily business, such as business systems (ERPs), BPM system, CRM system, etc

Stockholms universitet

Internet of Things

What is Internet of Things?

- Phycical Objects
  +
- Controllers, Sensors, Acentuations
  +
- Internet

What is Internet of Things?

Stockholms universitet

- “The Internet of Things (IoT) is the network of physical devices, vehicles, home appliances and other items embedded with electronics, software, sensors, actuators, and connectivity which enables these objects to connect and exchange data. Each thing is uniquely identifiable through its embedded computing system but is able to inter-operate within the existing Internet infrastructure”

[https://en.wikipedia.org/wiki/Internet_of_things]

How can IoT be used museums or galleries?

- How can you identify how IoT can be used in museums and/or galleries?
- How can IoT can be used in museums and/or galleries?