1 Introduction and Basics

1.1 Motivation

Business Intelligence is a relatively fuzzy term which cannot be sharply separated from a set of other domains.

<u>Four attempts at defining BI</u>

- "Business intelligence (BI) is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance."
  (from: Gartner, Inc. a large US company for market research and analysis http://www.gartner.com/technology/it-glossary/business-intelligence.jsp, Oct. 2013)

- "Business intelligence (BI) is a set of theories, methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information for business purposes.
  [...]
  Common functions of business intelligence technologies are reporting, online analytical processing, analytics, data mining, process mining, complex event processing, business performance management, benchmarking, text mining, predictive analytics and prescriptive analytics."
  (Wikipedia, http://en.wikipedia.org/wiki/Business_intelligence, Oct. 2013)

- "Business Intelligence ist ein Begriff aus der IT (Informationstechnologie). Die geläufigste und einfachste Definition ist das Gleichsetzen von Business Intelligence mit dem Begriff Data Warehouse (DWH)." (Translation: BI is a term from IT. The most popular and simplest definition is setting Business Intelligence equal to the term Data Warehouse.)
  (Business Intelligence Wissensportal, http://www.business-intelligence24.com/, Oct. 2013)

- My own definiton:
  Business Intelligence has been a hype since about the beginning of this century, and its development is gaining speed with the advent of ever faster computers and the gathering of ever more data. It roughly comprises the following aspects:
  I. classical data warehousing (invented in the 90ies of the last century)
  II. data analysis:
  a) with OLAP (more traditional type of analysis, mostly SQL based) and
  b) data mining (a conglomerate of methods from different fields like machine learning/artificial intelligence, mathematics and visualization to find "intelligent" information from mere data)
  III. BigData architectures: more modern architectures for processing huge amounts of data in parallel, often using algorithms like map/reduce to speed up and facilitate

analysis.
IV. studying application areas where large amounts of data are used for gaining insight into and often optimizing social, medical, technical, commercial, or biological processes.

What is Big Data?

By "Big Data" people refer to very large collections of data. These occur more frequently with improving computer technology and collection of large amounts of data in the internet.
Big Data implies the need to store, search, and analyze this data which presents specific problems because of the size of the data sets and has hence prompted the development of new storage technologies, new search and analysis strategies. It also often invokes discussion about data privacy.

The specific problems that occur in the context of Big Data are often summed up as "The Four V's":
- Volume: the size of the data
- Variety: data often does not easily fit into predefined patterns, like in a relational table
- Velocity: speed. Data is generated very fast and needs to be analyzed quickly:
- Veracity: correctness of the data

Some authors add even a fifth V:
- Value: is the collection and the analysis useful? Does it generate some value for the collectors?

=> In summary, one could say:
Big Data is what you put into a data warehouse to be stored, retrieved, and analyzed.

Sometimes, not a data warehouse is used but recently different architectures have emerged, based on large data stores that employ not traditional database technology but newer concepts like NoSQL systems. Hadoop and Hive are examples of systems that are more often encountered in this context.

What is included in our class:
- Applications of BI with examples
- Data warehousing (concept, architecture, modelling)
- OLAP (online analytical processing)
- Data Mining / Data Analytics (concepts, algorithms, lab with Rapid Miner)

In particular, we will proceed through the materials on the following route:
- Introduction and general overview (discussion of concepts (in particular multi-dimensional model) and applications)
- Data Warehouse reference architecture
- Closer study of the components of the data warehouse reference architecture and their interactions
- OLAP
- Data Mining / Analytics (also a brief look into AI)

- Data warehouse modelling
- Student presentations about various BI topics

Has anyone here ever worked with a data warehouse?

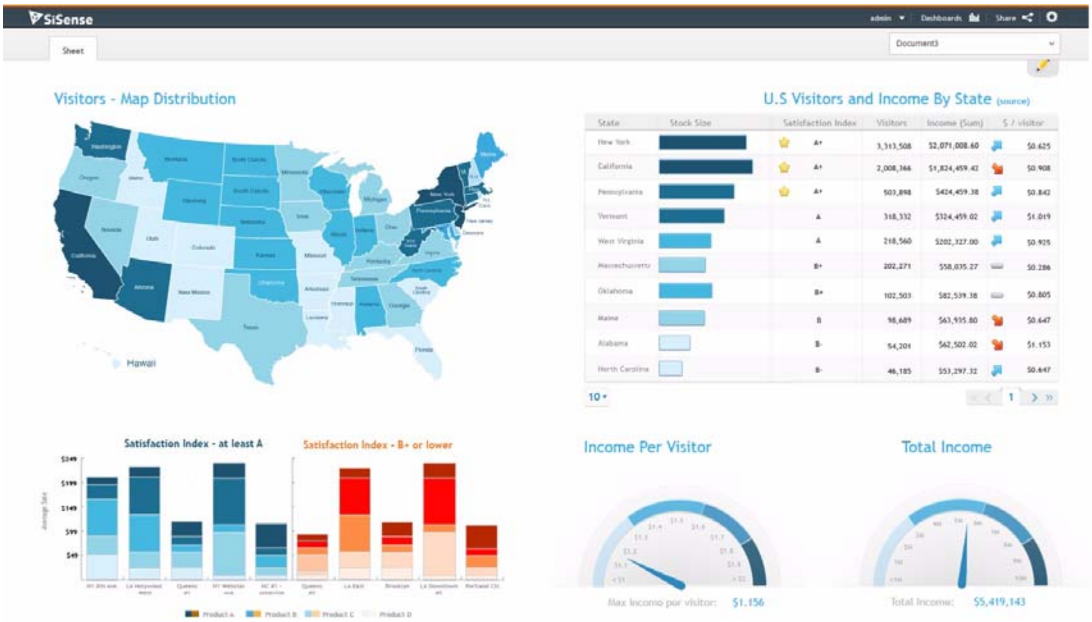
Language note: The word "warehouse" is not equal to "Warenhaus", but "Lagerhaus"!

Databases have been around for a long time. They are
- optimized for numbers of transactions and therefore often called OLTP (online transaction processing) systems.
- They are based on ER modelling and
- contain detailed, current data that changes rapidly.

Managers of enterprises have informational needs for decision making that are based on the company data but often cannot be readily satisfied by using OLTP systems.

Managers ask questions
- that are as such very simple ("In which regions have sales of this product group risen the fastest in the last year?") but whose answers often require complex queries that the managers cannot write themselves (because the model is too complex) and
- that join a large number of tables so that the response times are slow and badly affect the system performance for other users.
- Also, the data in OLTP systems reflects the current status, usually not historical data, which however, is often needed for decision making. (Example: How many cars are on stock today vs. last week?).

Managers want short (~1 page) reports that can be created by a single button click and extended easily as needed ("slice and dice").



Interactive dashboards are GUI functions that allow the easy display and interactive structuring of reports that are generated from data in source systems (e.g. a database or data warehouse).

Product	Region	Sales This Month	Growth in Sales vs. Last Month	Sales as % of Category	Change in Sales as % of Category vs. Last Month	Change in Sales as % of Category YTD vs. Last Year YTD
Framis	Central	110	12%	31%	3%	7%
Framis	Eastern	179	(3%)	28%	(1%)	3%
Framis	Western	55	5%	44%	1%	5%
Total Framis		344	6%	33%	1%	5%
Widget	Central	66	2%	18%	2%	10%
Widget	Eastern	102	4%	12%	5%	13%
Widget	Western	39	(9%)	9%	(1%)	8%
Total Widget		207	1%	13%	4%	11%
Grand Total		551	4%	20%	2%	8%

From: Kimball, The Data Warehouse Toolkit: "The ideal data warehouse report"

Another example (from Stuttgarter Zeitung, summer 2008):

Der Überschuss der Südwestbank hat sich l

Die Folgen der Finanzmarktkrise treffen auch die Stuttgarter Regionalbank – Filialnetz sol

STUTTGART. Die Südwestbank AG ist im vergangenen Jahr in allen strategischen Geschäftsfeldern gewachsen und hat neue Kunden hinzugewonnen. Negativ schlug 2007 der Einbruch beim Zinsüberschuss zu Buche.

Von Andrea Gregor

Die Stuttgarter Südwestbank ist zwar von den Verwerfungen auf dem US-Hypothekenmarkt nicht direkt betroffen. Aber indirekt sehr wohl. Die wichtigste Ertragsquelle, der Zinsüberschuss, hat sich um elf Millionen Euro verringert. Grund hierfür sei ein starker Anstieg der kurzfristigen Zinsen, sagte Vorstandssprecher Bernd Kiene bei der Bilanzpressekonferenz. Dadurch verteuert sich die Refinanzierung erheblich. Kiene: „Der Zinsüberschuss ist unsere Achillesferse."

Auch in diesem Jahr werde es bei den Zinsen keine echte Entspannung geben, solange die Finanzmärkte sich nicht beruhigten, prognostiziert der Südwestbank-Chef. Der Gewinn werde in diesem Jahr wohl nur leicht über dem Vorjahresniveau liegen. 2007 hat sich der Jahresüberschuss der Privatbank auf rund vier Millionen Euro halbiert, die Eigenkapitalrendite vor Steuern ist von 12,9 auf 3,4 Prozent gesunken.

Die Südwestbank ist bekannt für ihre günstigen Konditionen in der Baufinanzierung. Vorstandsmitglied Andreas Maurer deutete an, dass die Bank in diesem Segment fast nichts mehr verdient. Gegengesteuert hat das Institut mit einem einjährigen Festgeldangebot für neu angelegtes Geld. Das „Zinshoch" ist zunächst mit fünf Prozent ges wird jetzt jedoch niedriger verzinst. mit diesem Produkt habe man 3500 Kunden gewonnen, der Bank seien insg mehr als 200 Millionen Euro zugefl sagte Vorstandssprecher Kiene. Mit Einlagen stärke die Bank ihre Möglichk ausgelegte Kredite selbst zu refinan und sich so vom Kapitalmarkt unabhä zu machen. Die Kundeneinlagen sind u Prozent auf 3,38 Milliarden Euro gewa das Kreditvolumen wurde um 8,1 Proze 3,08 Milliarden Euro ausgeweitet.

In allen drei strategischen Geschä dern – der Mittelstandsfinanzierung, c nanzierung von Landwirtschaft und wirtschaft (zum Beispiel regenerative gieanlagen) sowie bei den vermögende vatkunden – sei die landesweit tätige gewachsen, betonte Kiene. Signifikant i Steigerung in der Vermögensverwa Mittlerweile betreut die ehemalige Genc schaftsbank Wertpapierdepots mit eine samtvolumen von 504 Millionen Euro 92 Millionen).

Kiene räumt ein, dass die Eigentüm Bank, die Biotech- und Pharmaunterne

KENNZAHLEN

SÜDWESTBANK AG	2007 Mio. Euro	2006 Mio. Euro	Veränd. in %
Gewinn[1]	4,02	8,67	– 53,6
Betriebsergebnis	4,33	15,19	– 71,5
Handelsergebnis	3,84	2,41	+ 58,9
Zinsüberschuss	46,37	57,40	– 19,2
Provisionsüberschuss	23,28	22,13	+ 5,2
Bilanzsumme	4 309	3 897	+ 10,6
Beschäftigte[2]	568	569	– 1
Dividende in Euro	–	–	–

[1] Jahresüberschuss [2] zum Stichtag 31.12.

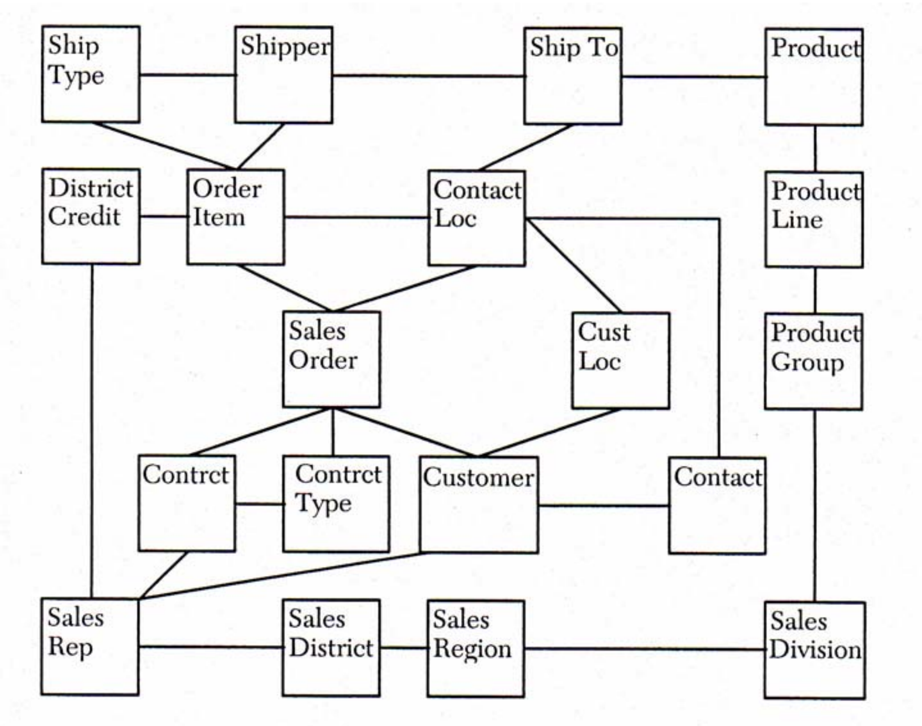
Kennzahlen = Key Performance Indicators (KPI)

=> For satisfying these needs, information systems (data warehouses)

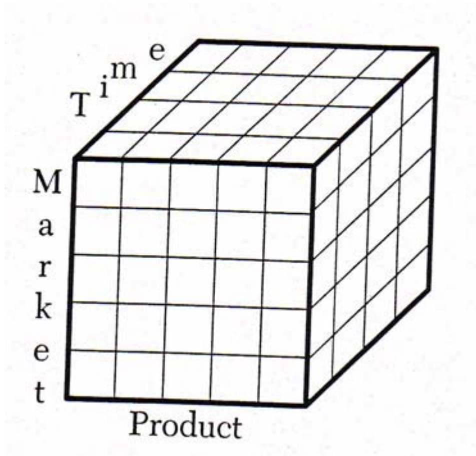
- are built on different hardware (so that the OLTP systems are not slowed down)
- with data structured as simple models
- with simple, fast analysis tools that can be handled by managers.
- The data in a data warehouse (DW) is built as an explicit time series by moving snapshots of the OLTP systems to the DW. The data in the DW thus changes only at loading times when the next snapshot is transferred.

The modelling for data warehouses takes a different approach from ER modelling; it is called **(multi-) dimensional modelling**.

Managers may state something like
"We sell products in various markets, and we measure our performance over time."



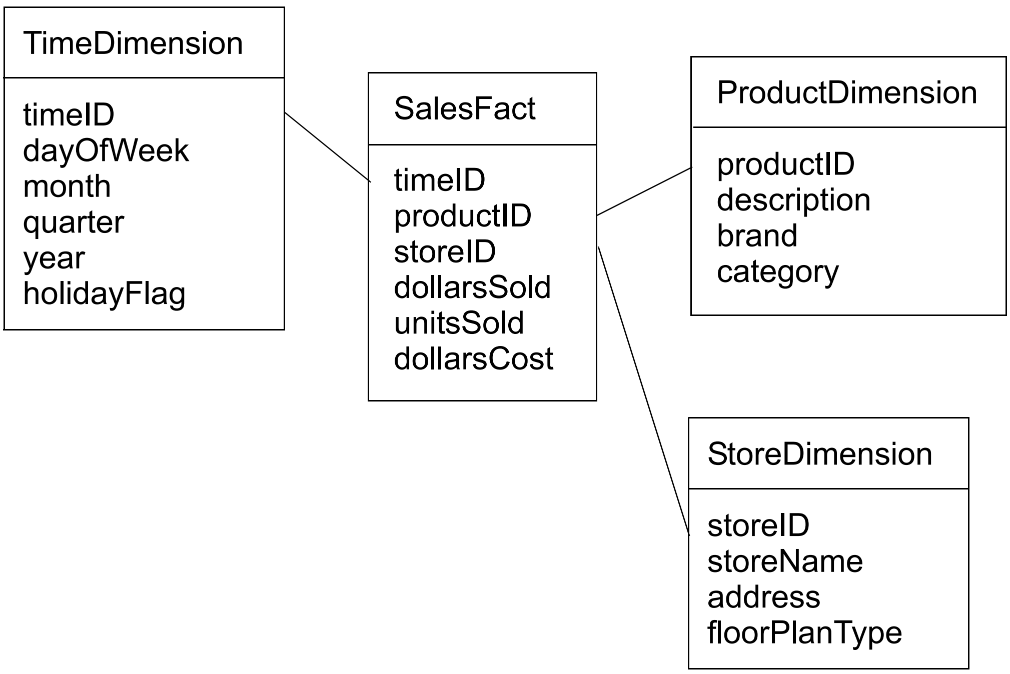
From: Kimball, The Datawarehouse Toolkit: "The data dependencies model of a business"



From: Kimball, The Datawarehouse Toolkit: "The dimensional model of a business. Each point in the cube contains measurements for a particular combination of a product, market, and time."

Dimensional models can be implemented in several ways:
- with a multidimensional database system
- A dimensional model can be expressed in terms of relations in a relational database system with a "star join schema". This name is derived from the form of the schema with one table in the middle that is connected via foreign keys to a number of schemas around it.

Example:



(From the same book by Kimball).

The table in the middle is called a fact table, the ones around it are called dimensional

tables because this is what their roles are.
"Brand" is the name for a collection of products that share some qualities. Example: Persil is a brand for a collection of laundry detergent products.

Facts

A fact is a measurable attribute that is not a natural, constant quality (like for instance a name) of an entity but a value that can be observed and measured, and which typically changes over time. Usually, facts are *numeric*, *additive*, and *continuously valued* (i.e. changing their value over time).
Additivity means that the values of a fact can be added up to a meaningful sum (e.g. cost or profit sums).
Examples of facts are quantities of money or units of products.
According to Kimball, most attributes whose values must be observed and measured should be modelled as facts.

Dimensions

Dimensions are described by attributes that are typically not numeric, but textual. Typical examples are product name, brand, description, size, and packaging type of products.
Dimension attributes are normally used in DW queries as attributes that are projected on (in the Select clause) and to place constraints on (in the Where clause).

Example (referring to the star join schema in the previous figure):

List the total $ sales and the number of sold units for all product brands sold in the first quarter of 1995.

```
SELECT Product.brand, sum(Sales.dollars), sum(Sales.units)
FROM     Sales, Product, Time
Where    Sales.productkey = Product.productkey
And      Sales.timekey = Time.timekey
And      Time.quarter = '1 Q 1995'
Group By Product.brand
Order By Product.brand
```

=> This is a join between the fact table "Sales" and the two dimension tables "Product" and "Time". Brand is a dimensional attribute that is used as a "row header". Quarter is a dimensional attribute that is used to place a constraint on.
By "row header" Kimball means dimension attributes for which aggregated values (from the fact table) are collected. He suggests to list those attributes first in the SELECT clause because the report generated this way will look more natural than if the non-aggregated attributes were interspersed with the aggregated ones.

Note:
The fact table can be considered the relation derived from a n-m-relationship between the dimensions. The keys of the dimension tables are used as foreign keys in the fact table. Typically, the key of the fact table is composed of the keys of all dimension

tables.

1.2 Definition of Data Warehouses

Datawarehouse:

- from a technical point of view: a database that integrates data from various (heterogeneous) sources
- from a business point of view: a system that provides analysis of data

Data warehouses were developed for decision support.

Definitions for data warehouses vary. Standardization efforts were done by OLAP Council (www.olapcouncil.org[1], but the Council does not exist anymore since 1999) and the Metadata Coalition which merged into the OMG (Object Management Group (www.omg.org).
The OMG is developing the CWM (Common Warehouse Metamodel), see http://www.omg.org/technology/cwm. It is a specification that describes metadata interchange among data warehousing, business intelligence, knowledge management, and portal technologies.

So far there is no agreed standard. Discussion takes place for instance on comp.databases.olap.
http://www.bi-verdict.com/ is a more recent site with Business Intelligence information. (It used to be called The OLAP Report).

Inmon was one of the first to define a data warehouse.

Definition of a data warehouse:
"A data warehouse is a subject oriented, integrated, non-volatile, and time variant collection of data in support of management's decisions."

- *Subject orientation*:
  Inmon opposes a system focus on the applications (e.g. as required by different departments) of a company to a focus on its subject areas (the more holistic interests of a company, subjects that relate to all departments, like profit, customers, products, etc.), whereas traditional database systems are developed for the applications and data warehouses for the subject areas.
- *Integration*:
  Data from various sources is represented in the data warehouse. Different sources often use different conventions in which their data is represented. This must be unified to be represented in a single format in the data warehouse.

1. The OLAP Council was established in January 1995 to serve as an industry guide and customer advocacy group. Participation in the OLAP Council is open to organizations that are interested in the advancement of on-line analytical processing (OLAP) server technology.
General Members: Applix/TM1 Incorporated, Business Objects, IBM, Hyperion Solutions (Formerly Arbor Software, OLAP Council founder), Management Science Associates, NCR, Oracle (OLAP Council founder), Platinum Technology, Inc., Sun Microsystems

Example: Application A uses "m" and "f" to denote gender, application B uses "1" and "0", application C uses "female" and "male". One of the conventions can be used for the data warehouse; the others must be converted.

- *Non-volatility*:
  Data that have been migrated into the data warehouse are not changed or deleted.
- *Time variance*:
  Processing of data warehouse data is stored in a way to allow comparisons of data loaded at different times (e.g. a company's profits of last year versus the profits of the year before that). => Data warehouse data is like a series of snapshots of the data of its different sources, taken at different times, over a long period of time (typically 5-10 years). (In comparison: The purpose of most databases is to present current, not historical data. Data in traditional databases is not always associated with a time, whereas data in a data warehouse always is).

From: Inmon: "Building the Data Warehouse"

A newer and more general definition is found in Bauer/Günzel:

Definition of a data warehouse:
A data warehouse is a physical database that allows an integrated view to arbitrary data.

Definition of a data warehouse system:
A data warehouse system is a system that provides functionality to manage a data warehouse. It provides all necessary components for integrating the necessary data, administrating them in a homogeneous database, and analyzing the data.

Definition of data warehousing:
Data warehousing (aka "the data warehouse process") is the dynamic process of acquiring, storing and analyzing the data, i.e. the processing of all necessary data from their origins to the final analysis result.
=> A data warehouse is only useful with data warehousing.

Since data warehouses support analysis, an appropriate modelling is required.
This is often achieved by a multidimensional data model.

OLAP (online analytical processing) is one method of retrieving information from a multidimensional data model.
Another analysis approach is Data Mining (the search for patterns or relationships in a data warehouse, often by statistical methods).

1.3 Comparison to other types of information systems

1.3.1 Comparison to transactional systems and traditional database systems

For many applications in a company, traditional database systems are used. They are

sometimes called *operational systems*, i.e. systems that are used to support the core business transactions of an enterprise (as opposed to the planning and disposition of the business).
A data warehouse, in contrast, is aimed at the dispositional and analytic aspects of a business. It contains a database whose data is based on other sources, often operational systems.

Data warehouses focus on integration and analysis. In contrast to that, transactional systems focus on classic database transactions (OLTP systems).

Differences between data warehouses and OLTP can be classified according to the categories Queries, Data, and Users:

Queries:

	Transactional (OLTP)	**Analytic (DWS)**
Focus	read, insert, update, delete (CRUD)	read, insert periodically
Duration and type of transaction	short read/write transactions	long read transactions
Frequency	high	low
Query structure	simple	complex
Volume of data	few records	many records
Data model	flexible for different queries	analysis oriented
Concurrency Control	necessary	not necessary

Data:

	transactional	**analytic**
Data sources	usually one or few	usually several / many
Qualities	(usually) not derived, detailed, current, autonomous, dynamic	derived, consolidated, historical, integrated, stable
Volume of data	Mega Byte - Giga Byte	Gigabyte - Tera Byte
Accesses	single records (tuples)	entire domains

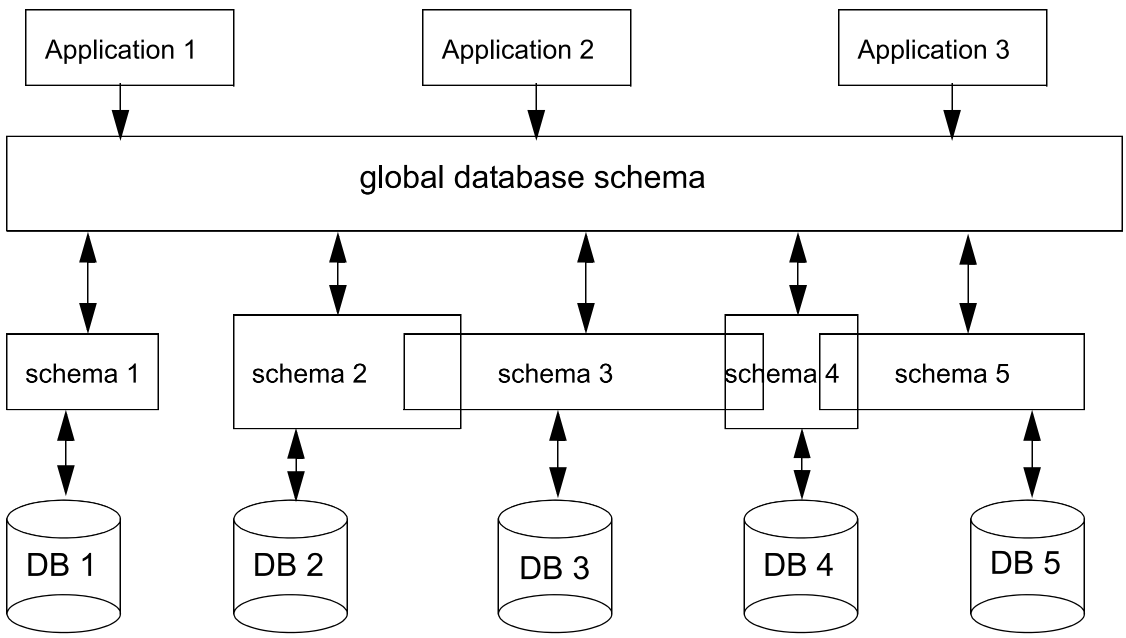
Users:

	transactional	analytic
User type	typical: insertion and retrieval by office clerks	evaluation and analysis by managers, controllers, analysts
Number of users	very many	few (up to a few hundred)
Response times	ms - s	s - min

1.3.2 Comparison to other database concepts

Distributed and federated databases have some similarities to data warehouses.

Distributed database:
a database with one common schema whose parts are physically distributed via a network. For a user, a distributed database appears like a central database, i.e. it is invisible to users where each data item is actually located.
It is centrally administered.



Federated database:
A FDB is a collection of distributed, (different), partial, redundant and partially autonomous databases which are coordinated by a federated database system with a new conceptual schema relating the subschemas.
Examples:
Websphere Information Integration provides functionality for federating different DBMSs.
A DB2 instance can be installed as a federated server and thus allow access to tables

from different DBMSs like Oracle, MS SQL-Server or in general data sources that can be accessed via ODBC.
MySQL includes a Federated Engine that allows accessing tables on a different MySQL server.

Autonomy means that each DB of a federation has a certain degree of autonomy regarding the design, the execution, and the communication with the other DBs. Therefore DBs may adopt different conceptual schemas (including domains, relations, naming conventions....), and certain operations are performed locally by the DBs, without interactions with the other DBs.

Data warehouses are somewhat similar to federated databases, with the differences that
- a data warehouse schema serves specific analytic purposes,
- the data in a DW are kept redundantly, but there is no copying of data from the sources to a federated system,
- write access to the source systems is not required for a DW system.

1.3.3 Comparison to other analytic systems

Already in the 60ies, the term management information systems (MIS) was coined and has been researched ever since.

Executive Information Systems (EIS), Decision Support Systems (DSS) and some others aim at more or less the same functionality as MIS: to support decision makers in enterprises on various hierarchical levels with the necessary information.

Typically, such systems nowadays are implemented as data warehouses.

1.4 Application Areas

Most application areas are in the business field.

- Business
  anything to support business success by providing the necessary information about an enterprise to the managers of this enterprise.
- E-Business
  DWs can be used for information purposes in E-Commerce via the Internet, intranets, and extranets. (Difference: Internet is accessible by anyone, intranet only by employees of a company, extranet includes also customers and suppliers (by password)).
  Enterprise information portals[1] (a portal for management information, a MIS) are often based on a DW.
  Online Marketing: E-Commerce companies use online-monitoring to collect information about the customers' surf, information, and transaction profiles, in order to personalize customer care. Online Sales: online auctions and shops use DWs to manage their information. In Electronic Procurement (purchases for enterprise

internal purposes), DWs are used to analyse procurement behaviour.

- Science (technical, medical, biological)
  Scientific and empiric studies often produce or use large amounts of data (for instance measurements). Examples: climate and weather research (meteorological data in TB sizes every day), Mars Mission, genome analysis, cancer studies
- Technical Applications
  Examples: water quality for environment control; power supply and consumption; production factory with a materials database to document ingredients or materials => in case of liability for erroneous products, the responsible lots or suppliers can be found quickly.

Specific application examples

- WalMart (a large American retailer):
  maintains a data warehouse sized ~ 25TB, ~20,000 queries every day, used for market basket analysis, customer classification (Question: Is 20,000 much or little? How may this number come to pass?)
- EOS (Earth Observing System), a project concerned with climate and environment research: receives ~ 1.9 TB of meteorological data every day, analysis with data mining tools

1.5 Usage areas

There are typically 4 types of usage of a data warehouse

1.5.1 Information

Providing information is the most important usage of a data warehouse, especially for preparing reports with key performance indicators (KPI, „Kennzahlen") and business performance measurement (BPM, examples: return on investment (ROI), turnover, cash flow, etc.). 70-80% of DW users are only looking for information.

1.5.2 Analysis

Methods from business adminstration are used to analyze the DW data according to application requirements.
Examples:

---

1. According to Christopher Shilakes and Julie Tylman, "Enterprise Information Portals are applications that enable companies to unlock internally and externally stored information, and provide users a single gateway to personalized information needed to make informed business decisions. " They are: ". . . an amalgamation of software applications that consolidate, manage, analyze and distribute information across and outside of an enterprise (including Business Intelligence, Content Management, Data Warehouse & Mart and Data Management applications.)"

1. Controlling
   Example: Would profits be greater with greater or smaller investments in marketing and sales offices? With greater or smaller variety of products? With larger or more focussed target groups? With larger or regionally more restricted sales areas?
2. Economic reference/ key performance indicators:
   Business performance indicators are often an important part of analytic applications. Well-known example: DuPont-System of Financial Control (with Return on Investment (ROI) as primary indicator).
   Balanced Scorecard: combines several indicator systems from different areas. Different views: **finance** (classic financial index numbers), **customer** (e.g. relation of old customers to new, number of customers, satisfaction indicators), **innovation** (e.g. employee qualification, number of new products), **business process** (internal company processes).
3. Cost Calculation and Management
   Cost centers, cost types, orders.

<u>People profiling:</u>

In recent years, a rising number of companies analyzes data about people in order to determine profiles of customers and prospective customers.
The goal: optimize profits to be had from customers.
=> Find out who are good and who are bad customers.
Build classifications built on the expected profit to be had from these customers. Often these are based on probablities. Example: "This customer comes with a 90% probability of paying his bills on time, his probability for being alive in 5 years is 85 %, the probability that he will buy one of our products in the near future is 43%".

1.5.3 Campaign Management

Campaigns are applications that are executed only in special situations to support strategic goals. Often used in sales and marketing.

Example from the company DeutschlandCard, a company that collects customer profile data for companies like supermarket chains.

Die DeutschlandCard GmbH ist eine Tochter der arvato AG: http://www.arvato-services.de/de/index.html

2) Diplomarbeitsthema: Kontaktoptimierung bei konkurrierenden Kampagnen im Kampagnenmanagement eines Multipartner-Kundenkartenprogrammes

Das Diplomarbeitsthema, das wir zu vergeben haben, beschäftigt sich mit einem Optimierungsproblem mit Nebenbedingungen, das bei uns praktische Anwendung erfahren soll.

Ich leite bei der DeutschlandCard GmbH das Kampagnenmanagement. Im Kampagnenmanagement führen wir auf Basis einer umfangreichen Teilnehmer-Datenbank Zielgruppen-Optimierungen für Mailings durch. In der Regel besteht die Fragestellung, welche Teilnehmer für ein gegebenes Angebot die höchste Responsewahrscheinlichkeit beseitzen. Diese Fragestellung wird über Prognoseverfahren abgebildet.

Darüber hinaus entsteht jedoch das Problem, dass verschiedene Angebote zur gleichen Zeit um die Kontakte der Teilnehmer konkurrieren.

Ein einfaches Beispiel:
Unser Partner EDEKA möchte den Verkauf von Teigwaren durch ein Mailing, das an 500.000 Teilnehmer gehen soll, bewerben. Zur gleichen Zeit möchte unser Partner porta den Verkauf einer bestimmten Möbelmarke mit einem Mailing an 200.000 Teilnehmer bewerben. Nehmen wir an, unser Datenbank würde nur aus 700.000 Teilnehmer bestehen, dann müssen wir Regeln finden, die die optimale Zuordnung der Teilnehmern steuern. Diese Regeln orientieren sich an den Zielen unserer Partner. Diese geben vor, dass der zu erwartende Umsatz aus den Kampagnen zu maximieren ist. Die Nebenbedingung dabei sind die Auflagen, die die Partner bestellen.

Ein komplexeres Beispiel:
Wir verschicken einmal im Quartal an jeden Teilnehmer ein Punktestandsmailing, das dem Teilnehmer, den Punktestand und die neu gesammelten Punkte kommuniziert (ähnlich einem Kontoauszug einer Bank). Bei diesem Punktestandsmailing schicken wir einen Bogen mit Coupons unserer Partner mit. Ein Coupon enthält eine Punkteincentivierung der Partner für einen Einkauf durch die Teilnehmer (z.B. 5-fach Punkte für einen Einkauf von min. € 20,- bei der EDEKA). Jedes Couponblatt, das ein Teilnehmer erhält, besteht aus 8 Couponplätzen, die frei vergeben werden können. Bei einer Million Teilnehmer wären das 8 Millionen Coupons, die die DeutschlandCard an die Partner vermarkten kann. Die Partner kaufen Teile dieser Gesamtauflage mit verschiedenen Angeboten. Nehmen wir an, die EDEKA würde 3 Millionen Coupons verkaufen. Die EDEKA möchte allerdings 2 Millionen Coupons mit 2-fach Punkten, 0,5 Million Coupons mit 5-fach Punkten und 0,5 Million Coupons mit 10-fach Punkten vergeben. Entsprechende Auflagen buchen die anderen Partner bis die 8 Millionen Coupons vergeben sind. Die Coupons sollen nun so auf die Teilnehmer verteilt werden, dass
1.) eine maximale Umsatzsteigerung durch Verteilung der Coupons über alle Partner erreicht wird (Ziel1)
2.) eine maximale Einlösequote erreicht wird
3.) kein Teilnehmer mehr als 2 Coupons von einem Partner erhält (Nebenbedingung)
4.) jeder Teilnehmer 8 Coupons erhält (Nebenbedingung)

Für die Einlösequote und die Umsatzsteigerung existieren prognostische Modelle (sog. Scoringmodelle), auf deren Werte und Konfidenzen zurück gegriffen werden kann.

Ich hoffe die beiden Beispiele machen das Thema transparent. Bitte zögern Sie nicht zu fragen, falls Sie genauere Angaben brauchen.

Wir planen für das geschilderte Problem eine Software einzusetzen. Wir haben hier Applikationen der Firma SAS im Einsatz. Für die Zielgruppenoptimierung bietet die Firma SAS die Software SAS Marketing Optimisation an. Ich schicke Ihnen anbei 2 kurze Beschreibungen dieser Software mit.

1.5.4 Planning

Enterprise goals must be expressed in numbers to be concrete and controllable.
DWs contain numbers reflecting the current and past status. In order to facilitate planning, sometimes projected numbers (planning results) can be inserted into the DW and compared to current values.