

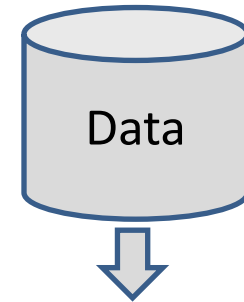
Business Intelligence and Data Warehouse

Knut Hinkelmann



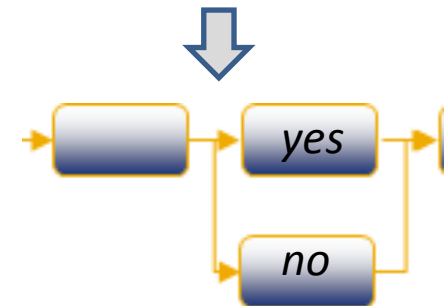
Business Intelligence – Definition(s)

- *Sabherwal (2011)*: «We define BI as providing decision makers with valuable information and knowledge by leveraging a variety of sources of data as well as structured and unstructured information. [...] The key intellectual output of BI is **knowledge that enables decision making with information and data being the inputs.**»
- *Howson (2007)*: Business Intelligence allows people at all levels of an organisation to **access, interact with and analyse data to manage the business, improve performance, discover opportunities, and operate efficiently.**

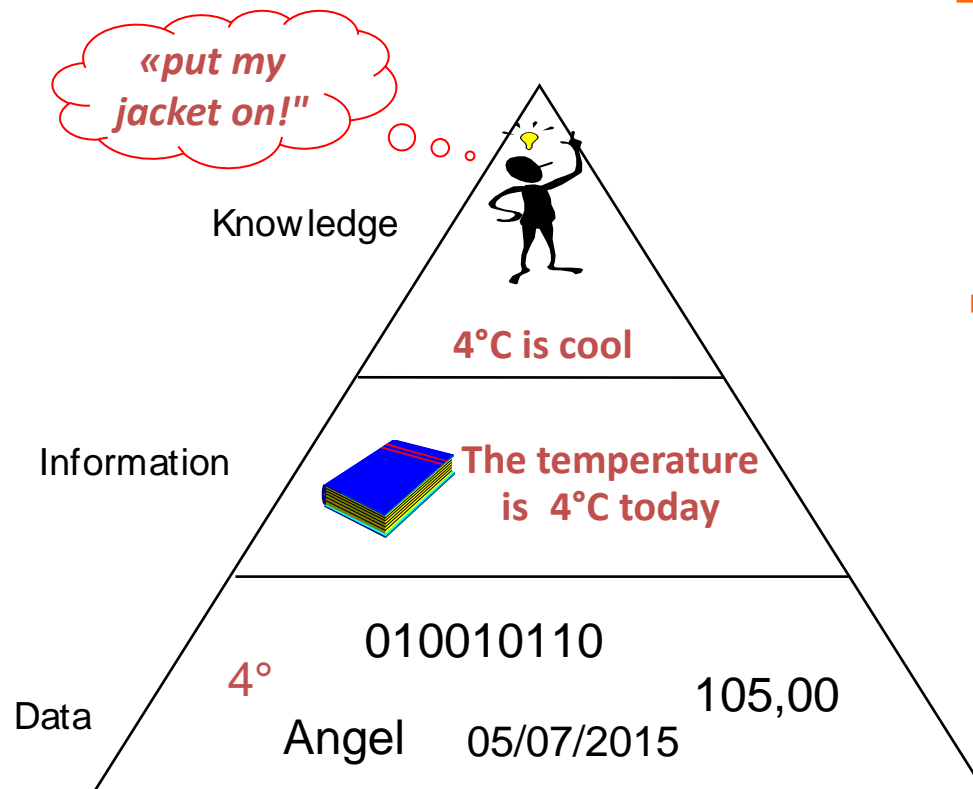


Alpha Corporation
Sales in EUR

	'10	'11	ΔPY
Germany	84	87	+3
Austria	19	17	-2
France	28	27	-1
Rest	36	39	+3
Europe	167	170	+3



Data, information and knowledge



- **Knowledge** enables decisions and actions
 - originates from messages (information), experience, insight
 - is embedded into the beliefs and opinions of its owner
- **Information** is an interpretation of data, often assembled in messages
 - influences the judgment and behaviour of the recipient and
 - that has a significance (relevance, purpose)
- **Data** is a set of facts and/or signals
 - Do not have meaning by itself
 - To understand data you need an interpretation

Analytic vs. transaction processing

- BI focuses on **analytic** processing instead of **transaction** processing
 - ◆ transaction processing supports *execution of core business processes*
 - *use knowledge*
 - ◆ analytic processing supports *insight and decision-making*
 - *create knowledge*



Perspectives on BI – pain points

MARKETING

*For targeted campaigns, we would urgently need a data basis that is harmonised with sales [...] ideally on an **integrated platform** where we can communicate with sales.*

MANAGEMENT

*I told my people that I wanted to retrieve some numbers myself from my laptop. I then got **access to various (!) systems** [...] I finally gave up and now have an employee who does nothing but **create reports** for me [...]*

SALES

*In most review meetings, we spend half the time discussing whose figures are the right ones because everyone **brings their own reporting**. I have the impression that for any key figure it is **possible to produce any value from the raw data**.*

ADVISORY BOARD

Why weren't you able to preview that trend? All our competitors seem to have reacted long before we did!



Why introduce BI? – primary motivations

■ Drive company strategy

- ◆ being able to connect strategising/planning to measuring of impact (do not manage «blindly»)

■ Growth and competitiveness:

- ◆ anticipate market trends and adapt R&D accordingly
- ◆ better customer relationships through better-targeted offers
- ◆ better leverage of customer potential (cross-/up-selling)
- ◆ optimise business processes

■ Single point of truth

- ◆ no by-pass reporting, consistent data

■ Cost reduction

- ◆ faster access to information
- ◆ automation of reports, self-service BI
- ◆ no interference of analytics with operational systems

business
drivers

technical
drivers



Decision making

- **Decision making** = *The action of selecting among alternatives to achieve a goal*
 - ◆ each alternative leads to a different future
 - ◆ what is needed is the ability to predict the futures
- **Options:**
 1. predict based on gut feeling
 - cheap in the first place
 - risk of low-quality decisions
 2. Experiment with real system (try out)
 - too risky
 - too time-consuming (only one set of conditions at a time)
 3. Predict based on the past:
 - Data collection is time-consuming
 - difficult to determine when to stop and make a decision
 - Too little or too much information



BI overview

Questions

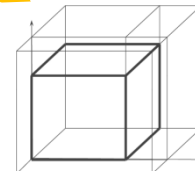
- strategic**
- What are our goals?
 - Are we reaching our goals?
 - If not, where is the problem?

- operative**
- Which customers are interested in the new product?
 - Which department has how many efforts that cannot be booked?

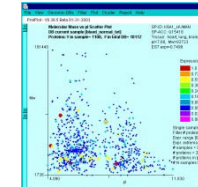
Analyses



measure, aggregate, visualise

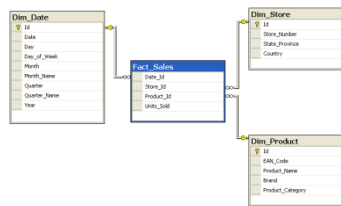


Ad hoc queries, OLAP



find patterns (data mining)

dimensional modelling

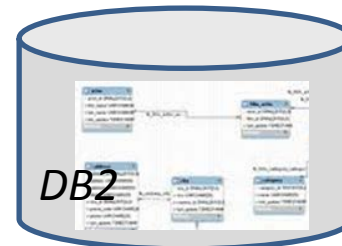


ETL

ETL

IE

raw data



Drivers for BI

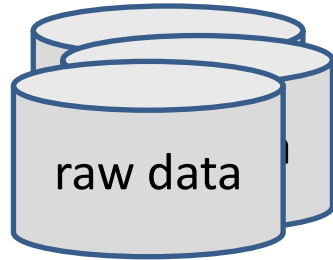


BI and fact-based decision making

- Fact-based decisions are based on information
- BI supports decision making by providing that information, usually in the following way:
 - ◆ the human decision maker (HDM) formulates the decision problem
 - ◆ the HDM manually identifies the questions that need to be answered in order to take an informed decision
 - ◆ the HDM consults a BI tool to get the answers, usually by querying or browsing (e.g. OLAP)
 - ◆ the HDM uses the answers to take an informed decision



Data-driven vs. business-driven BI

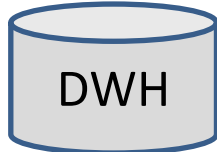


*We have data.
What can we do
with it?*

ETL



Consolidate and
integrate data



Analyze
data

*We have questions.
How do we get
answers?*



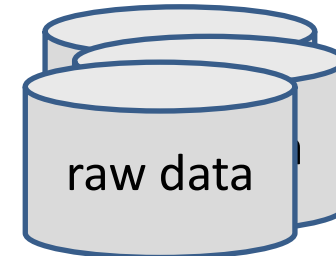
What data
do we need to
answer the
questions?

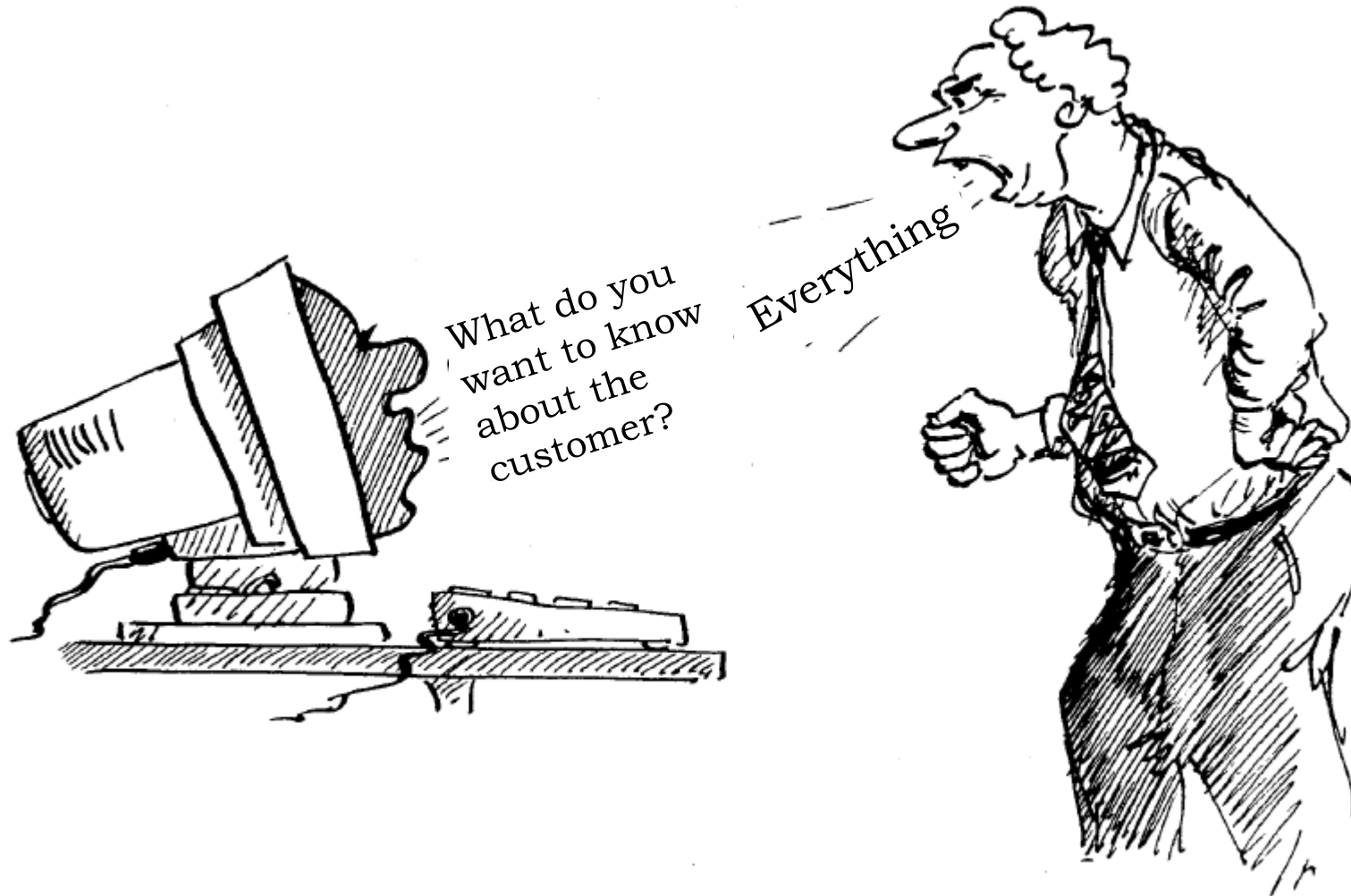


ETL



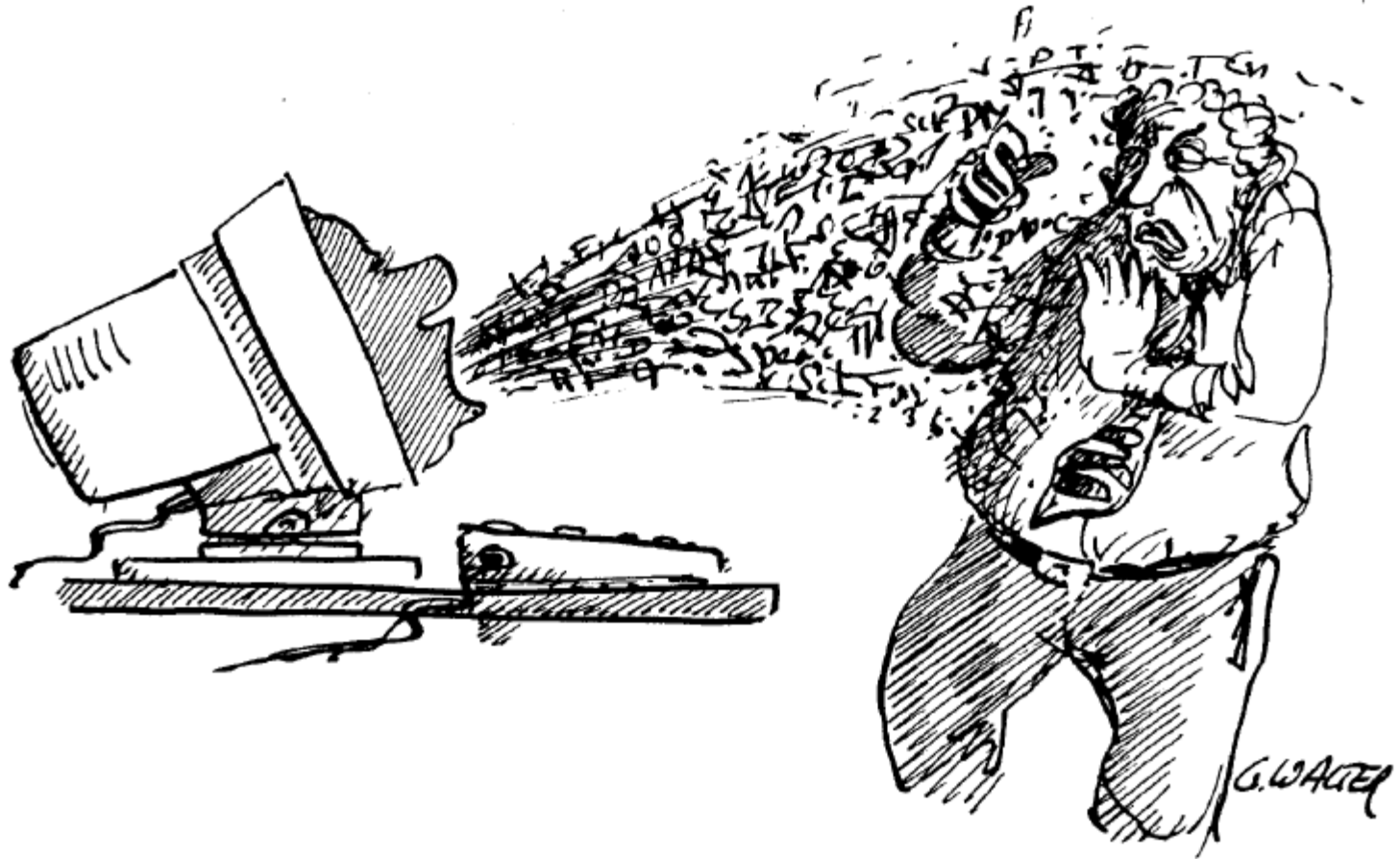
Collect and
consolidate data





adapted from slides by Dani Schneider





Where questions come from

- Generally speaking, companies need information to
 - ◆ monitor and improve performance
 - ◆ recognize and mitigate risks
 - ◆ recognize and seize opportunities

- All this can happen both on a strategic and an operative level



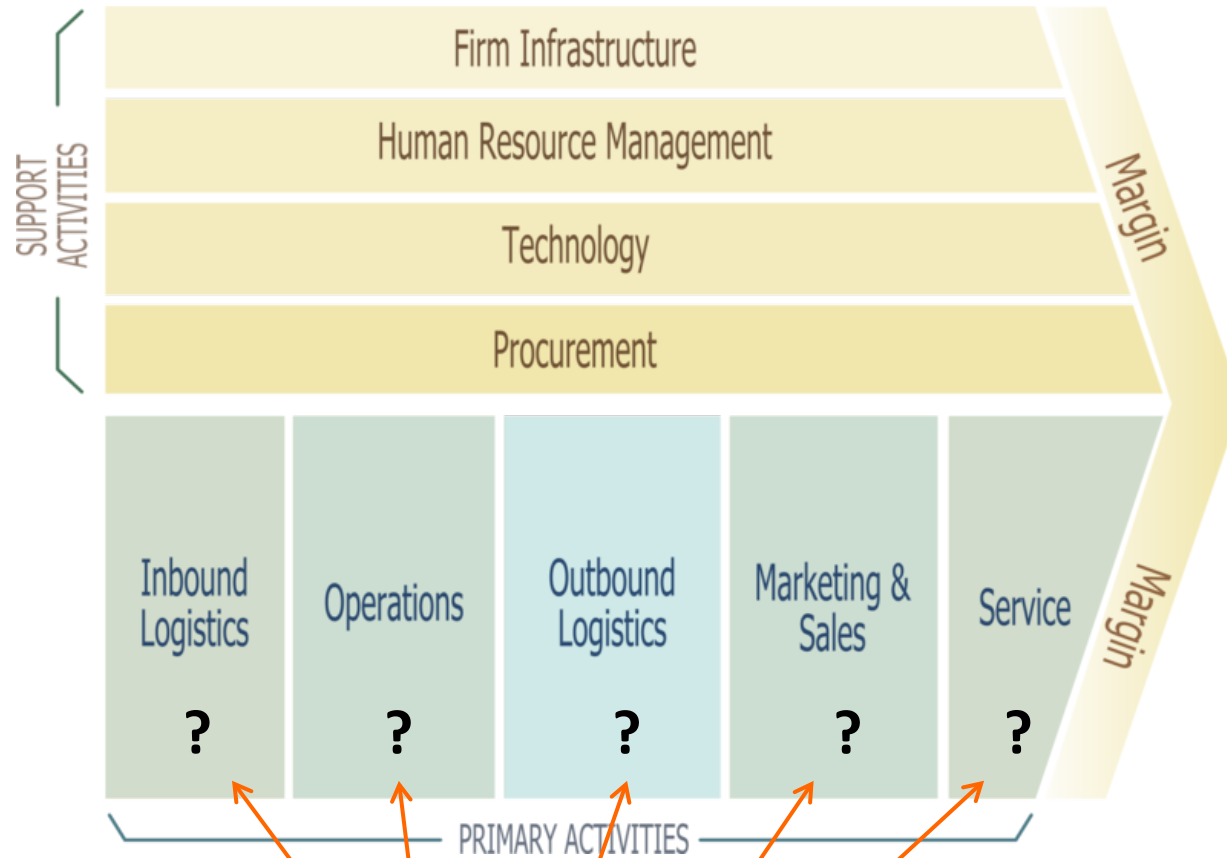
Strategic decisions...

■ Business Performance Management:

- ◆ “how to perform better as a company?”
- ◆ BI helps to achieve that by enabling measurement of achievement of strategic goals via KPIs
 1. Define **strategy**
 2. Define **goals**
 - e.g., identify key business processes to be improved, derive (concrete) strategic goals
 - for each goal, define KPIs and target values
 3. **Measure**
 - current values of KPIs (dashboard/cockpit)
 - analyse / compare current to targeted values
 4. **Decide...**
 - understand the (possible) deviation of KPI values from target!

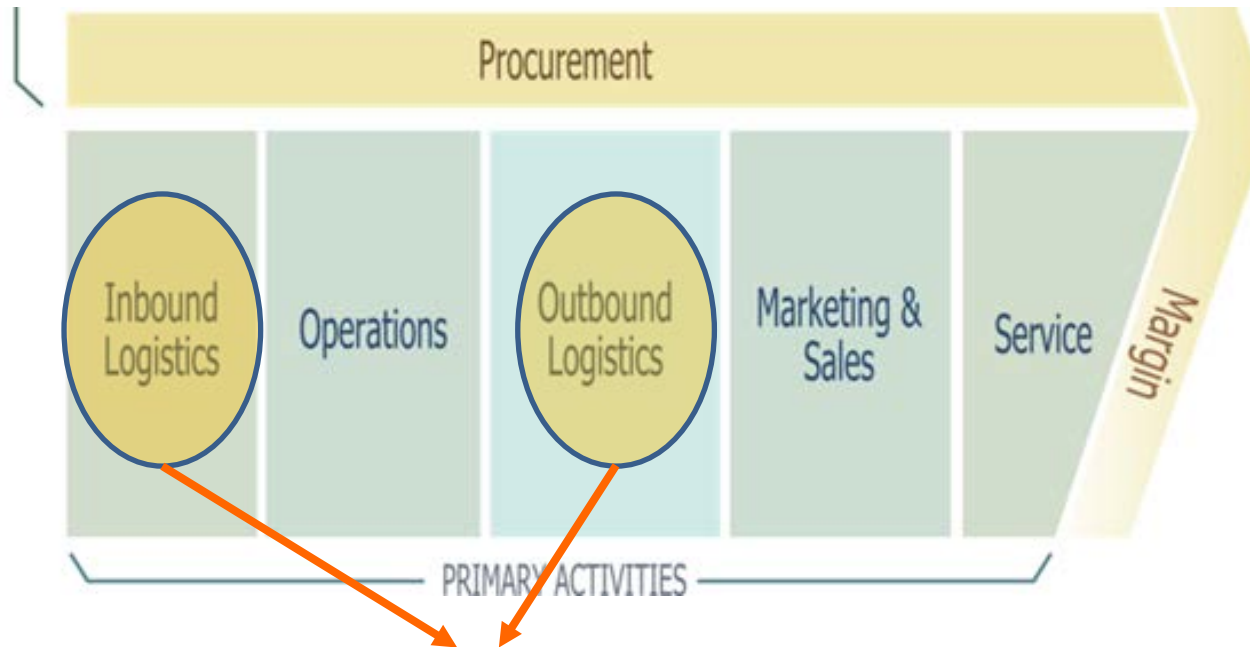


Operative decisions: where BI creates value...



decisions to be taken in corresponding business processes?

Operative decisions: where BI creates value...

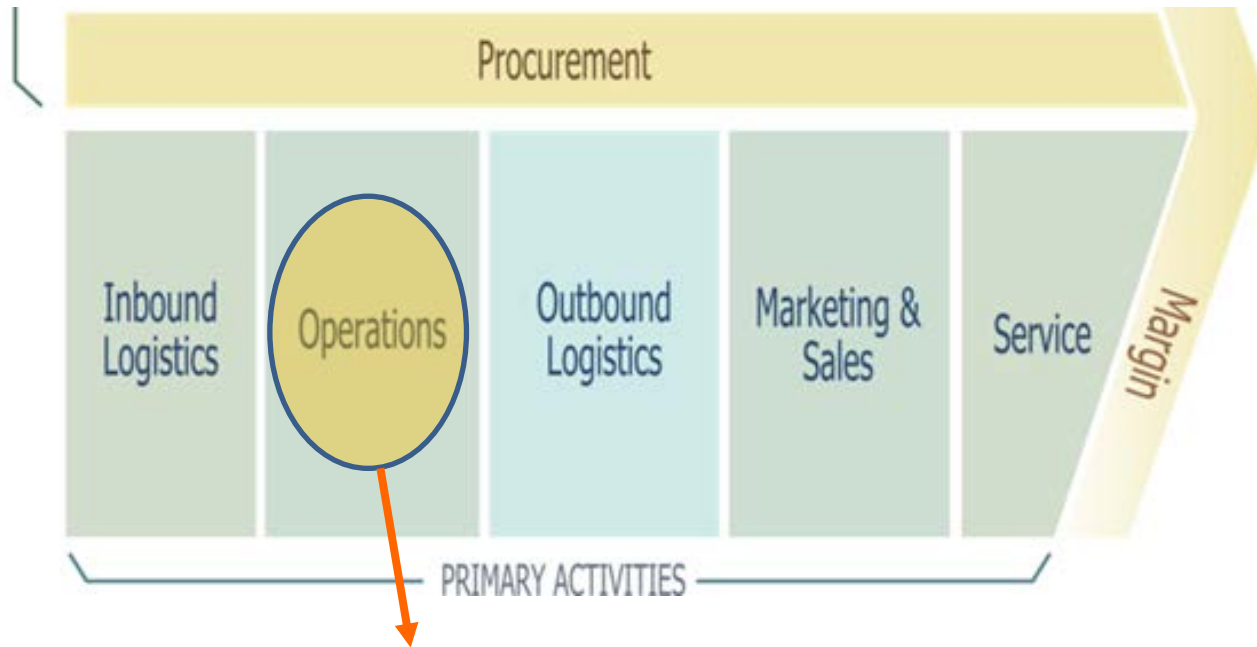


Logistics:

the process of planning, implementing and controlling the efficient, effective flow and storage of goods, services and related information from the point of origin to the point of consumption for the purpose of conforming to customer requirements

- **how to best use resources (inbound)?**
 - which parts to order, in which quantity, at what time, from which supplier?
- **how to optimise processes (outbound)?**
 - which route/channel to use, how to schedule deliveries?

Operative decisions: where BI creates value...

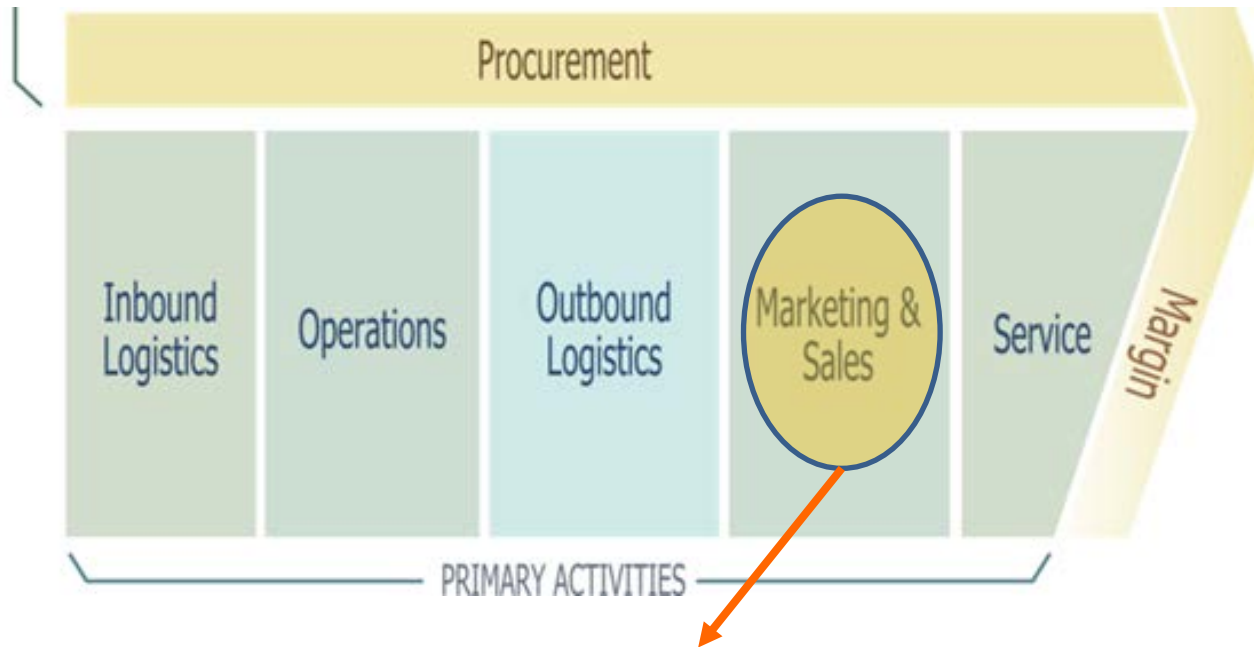


Operations:

activities associated with the functions of transforming inputs into the final product form, such as machining, packaging, assembly, equipment maintenance, testing, printing, and facility operations.

- ***how to improve efficiency and effectiveness of processes?***
 - which resources to allocate, in which quantity, ...

Operative decisions: where BI creates value...

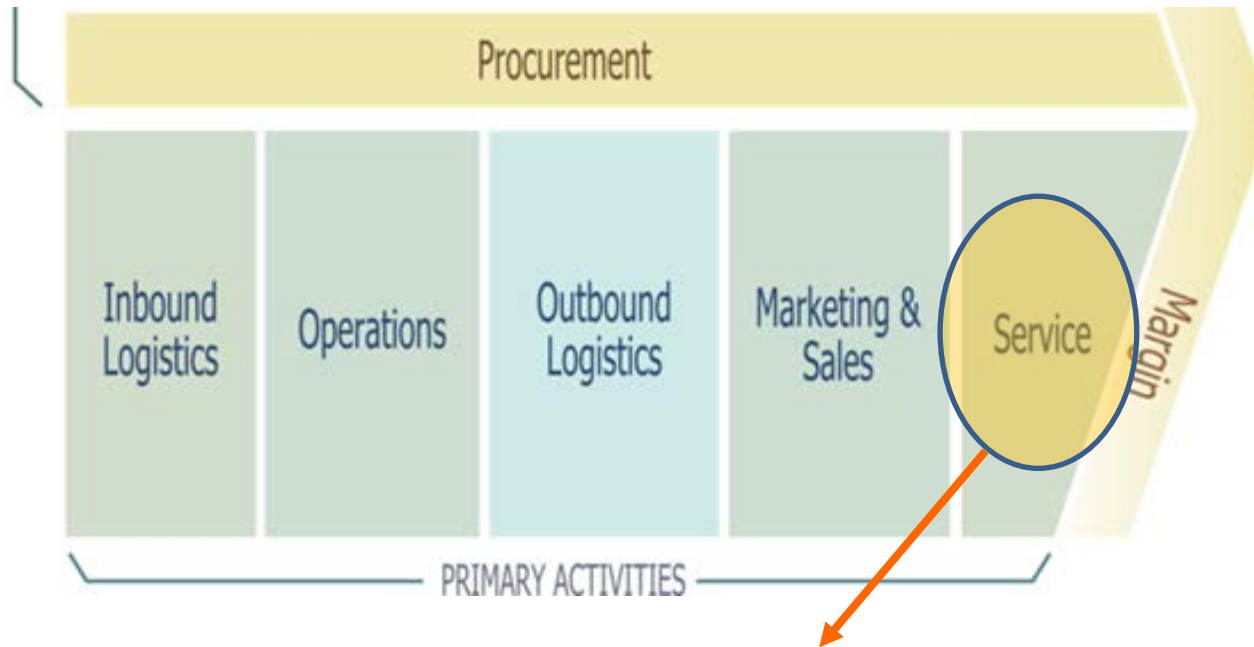


Marketing/Sales:

activities associated with the functions of providing the means by which buyers can purchase the product and inducing them to do so, such as advertising, promotion, quoting, pricing, channel and sales force management.

- **how to understand and best address the market?**
 - cross-selling: which offers to make?
 - where to place products in stores?
 - which customers to approach with a campaign?
 - client profitability: which customers to treat with special care?
 - pricing decisions

Operative decisions: where BI creates value...



Service:

activities associated with the functions of providing service to enhance or maintain the value of the product, such as installation, repair, training, parts supply, and product adjustment.

- ***how to meet customer requirements and anticipate problems?***
 - which distribution channels to use for service delivery?
 - which quality problems to address first?
 - Attrition prediction: which customers to retain with special offers?

Question types – summary

■ Types of questions identified:

- ◆ **query** for particular numbers or facts
 - *e.g. list of all policies that have been lost, list of all complaints, list of treatments that have been billed twice, list of high-value customers...*
- ◆ **compute a measure or KPI by aggregating numbers**
 - *e.g. cost, margin, turnover, profitability*
- ◆ **analyse KPIs / facts in different ways**
 - *e.g. sales/bookings by product/customer/sales rep/time*
 - *e.g. receipts/failures/stock by part/supplier*
 - *e.g. number of clicks/purchases by buyer/seller/page*
- ◆ **predict**
 - *e.g. predict fraudulent transactions/claims*
 - *e.g. predict if a customer will buy a product*
 - *e.g. detect types of customers or types of complaints*



Monitor and improve performance

- **Strategic level:** be able to measure if strategic goals are achieved
 - ◆ e.g. be able to measure the satisfaction of our customers over the last year
 - so that we can decide to change our customer service model

- **Operative level:** monitor performance within certain business processes, in small time intervals
 - ◆ e.g. find out that/why (individual) customers are not satisfied today
 - so that we can decide to call them and find a solution



Recognise and mitigate risks

- **Strategic level:** be able to recognise general threats to our business
 - ◆ e.g. become aware that sales in certain product category are dropping dramatically (which is threatening our whole business)
→ so that we can revise our product portfolio

- **Operative level:** be able to recognise risks related to individual processes, customers, suppliers, employees, ...
 - ◆ e.g. in telecommunications, be able to predict if a customer is going to cancel (or not renew) her contract
→ so that we can decide to make a special offer to that customer



Recognise and seize opportunities

- **Strategic level:** be able to recognise general opportunities for our business
 - ◆ e.g. become aware that (potential) customers are asking for a certain kind of product or product feature in social media
→ so that we can decide to develop such a product

- **Operative level:** be able to recognise opportunities related to individual process instances, customers, suppliers, employees...
 - ◆ e.g. recognise that we can cross-sell a certain product to an existing customer
→ so that we can decide to make the customer aware of that product



Specific requirements

When analysing requirements for a certain company's future BI solution, usually at least the following need to be fixed:

- ◆ Which are the strategic goals → which are our **KPIs**?
e.g. revenue, delivery time, profitability, ...
- ◆ By which criteria should KPI values be **grouped and/or filtered**?
e.g. by customer, by sales rep, by region, by date, ...
- ◆ Which **drill-down paths** should be possible?
e.g. date: year → quarter → month → ...
e.g. region: country → region → city

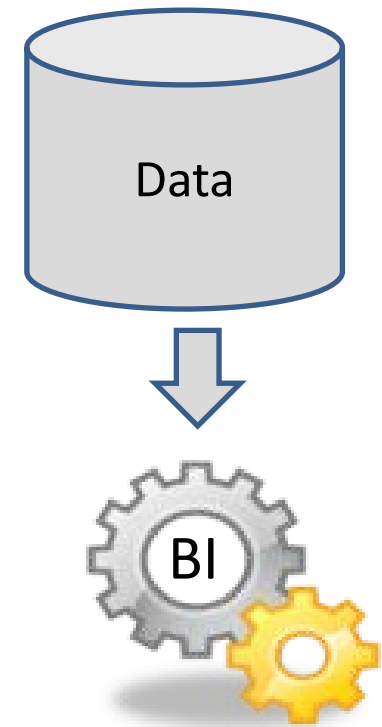


Data

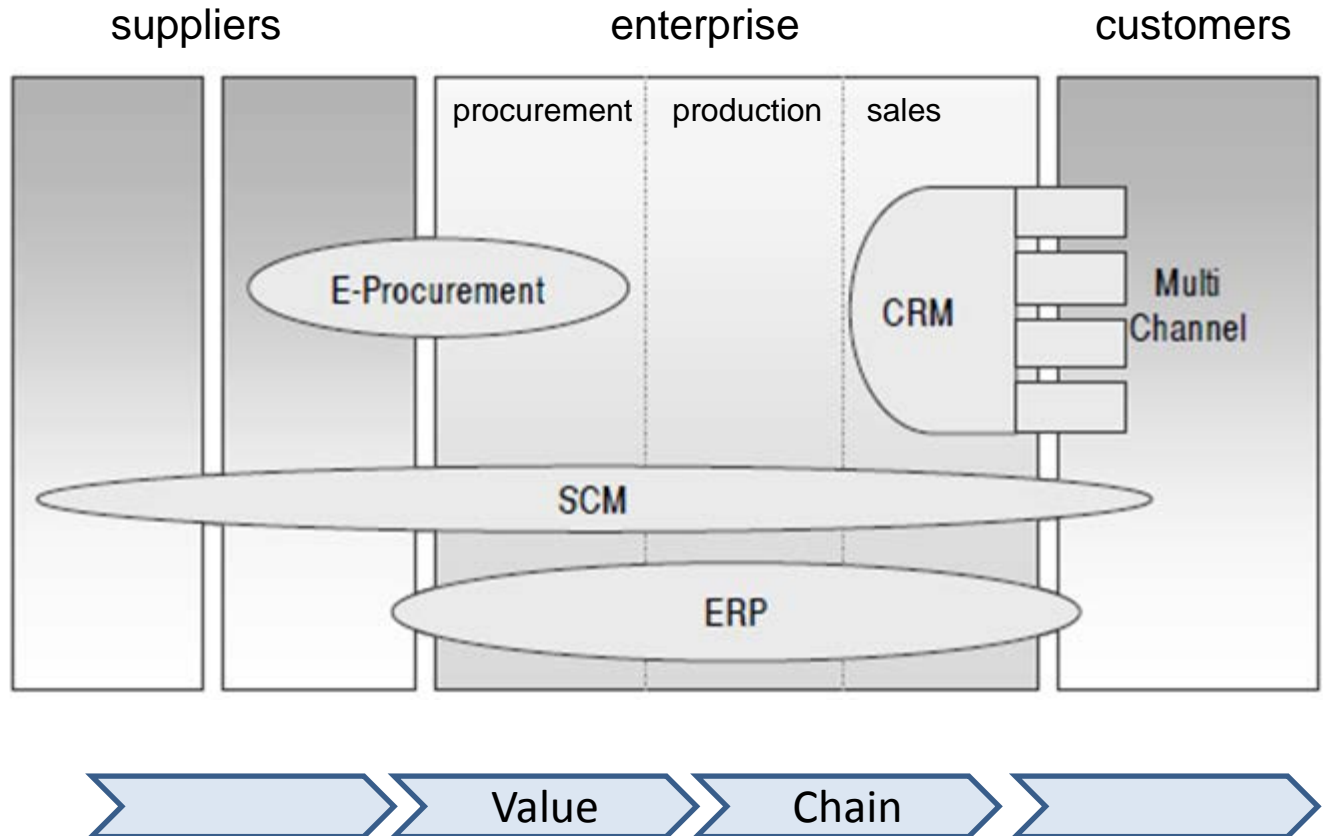


Remember...

- ... transform **raw data** into meaningful and useful **information**...
- **Raw data** is the starting point!



Where the data come from... (1)



CRM – Customer Relationship Management
SCM – Supply Chain Management
ERP – Enterprise Resource Planning

adapted from Kemper et al. 2004



Where the data comes from (2)

■ Internal data sources:

- ◆ (Transactional) standard business applications: SCM, ERP, CRM, ...
- ◆ Legacy databases
- ◆ Web data: clickstreams from server logs, application logs
- ◆ textual documents (from DMS, CMS, intranet, email,...)



*structured-
ness*

■ External data sources:

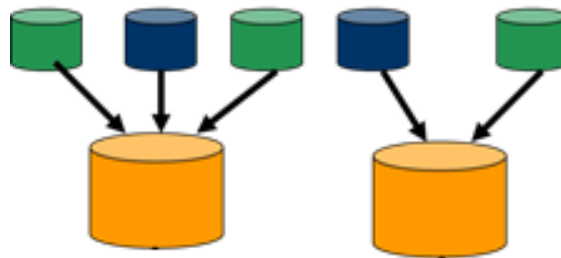
- ◆ Web and web 2.0



BI tools – backend

■ Observations:

- ◆ many questions involve multiple (types of) data
- ◆ sometimes the data can be expected to originate from more than one source system
- ◆ for answering the questions, data from various sources needs to be connected
 - example: «Which is the best way to distribute product XYZ to customers?» → involves information about customers (e.g. profitability, behaviour) as well as about channels (e.g. cost of each channel)



Planning Data vs. Operative Data (1)

- **operative data:** generated by and used in processing operational transactions (on-line transaction processing, OLTP)
 - ◆ many concurrent users access and modify the same data
 - ◆ focus on transactions
 - ◆ example: booking/reservation systems
- **planning data:** used for decision support
 - ◆ read-only data

following Kemper et al. ch 2.1



Planning Data vs. Operative Data(2)

	Operative data	Planning data
users	clerk, IT professional	knowledge worker, decision maker
Function/goal	Support day to day operations (value adding business processes)	decision support
DB design	application-oriented	subject-oriented
data	current, up-to-date detailed information on business events, flat relational	historical, summarized, multidimensional integrated, consolidated
usage	Continuous, repetitive, concurrent	ad-hoc
access	read/write index/hash on primary key	lots of scans
Queries	Static, transactions embedded in application code	Ad-hoc, for changing information needs
# records accessed	tens	millions
#users	thousands	hundreds
DB size	100MB-GB	100GB-TB
metric	transaction throughput	query throughput, response

adapted from <http://www.slideshare.net/idnats/data-warehousing-and-data-mining-presentation-725476>

