Research Approach for Measuring Information Quality in Data Warehousing

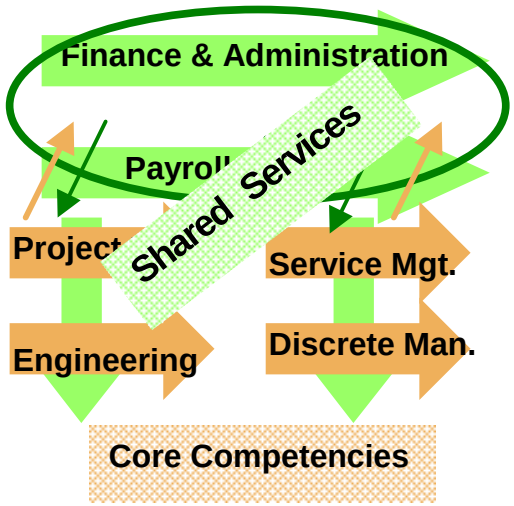
Markus Helfert
Institute of Information Management
University of St.Gallen

Markus.Helfert@unisg.ch
datawarehouse.iwi.unisg.ch

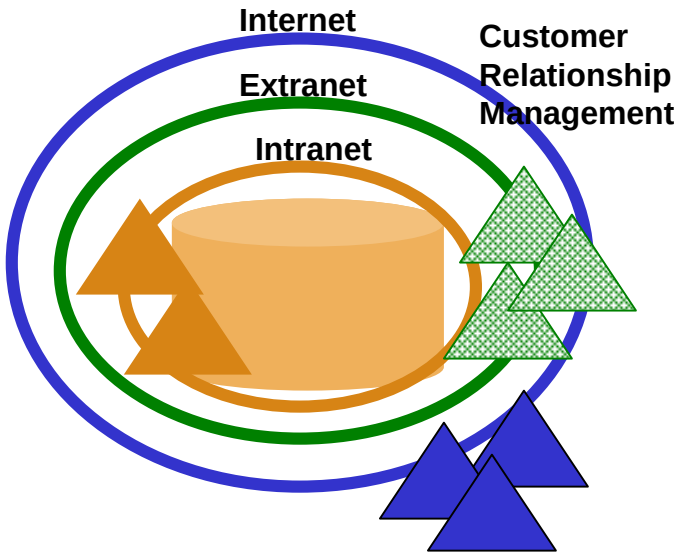
Renate Radon
Epicur GmbH
Germany

Renate.Radon@epicur.net
www.epicur.net

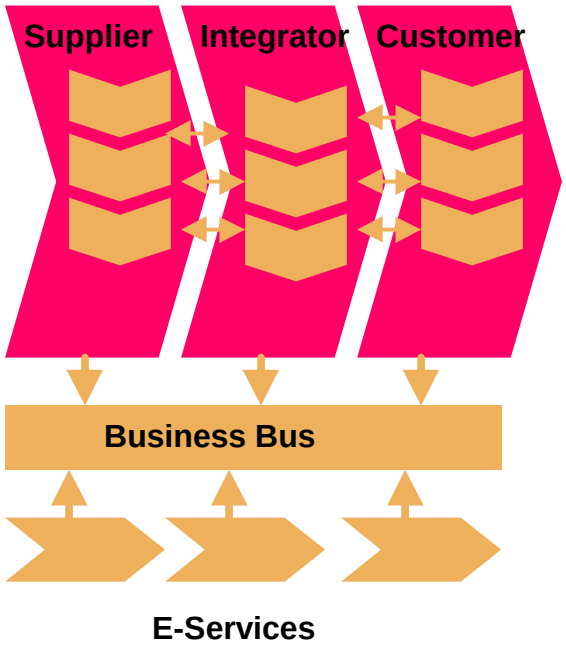
EPICUR – Competencies



Adding Economic Value



Business Intelligence



Business Networking

Competence Center Data Warehousing Strategy

sanitas

Württembergische VERSICHERUNG

ARAG

Rentenanstalt Swiss Life

winterthur

AGI IT SERVICES AG

VERTEIDIGUNG · BEVÖLKERUNGSSCHUTZ · SPORT
DEFENSE · PROTECTION DE LA POPULATION · SPORTS
DIFESA · PROTEZIONE DELLA POPOLAZIONE · SPORT
DEFENSIUN · PROTECZIUN DA LA POPULAZIUN · SPORT

swisscom

plenum AG

CREDIT SUISSE

UBS

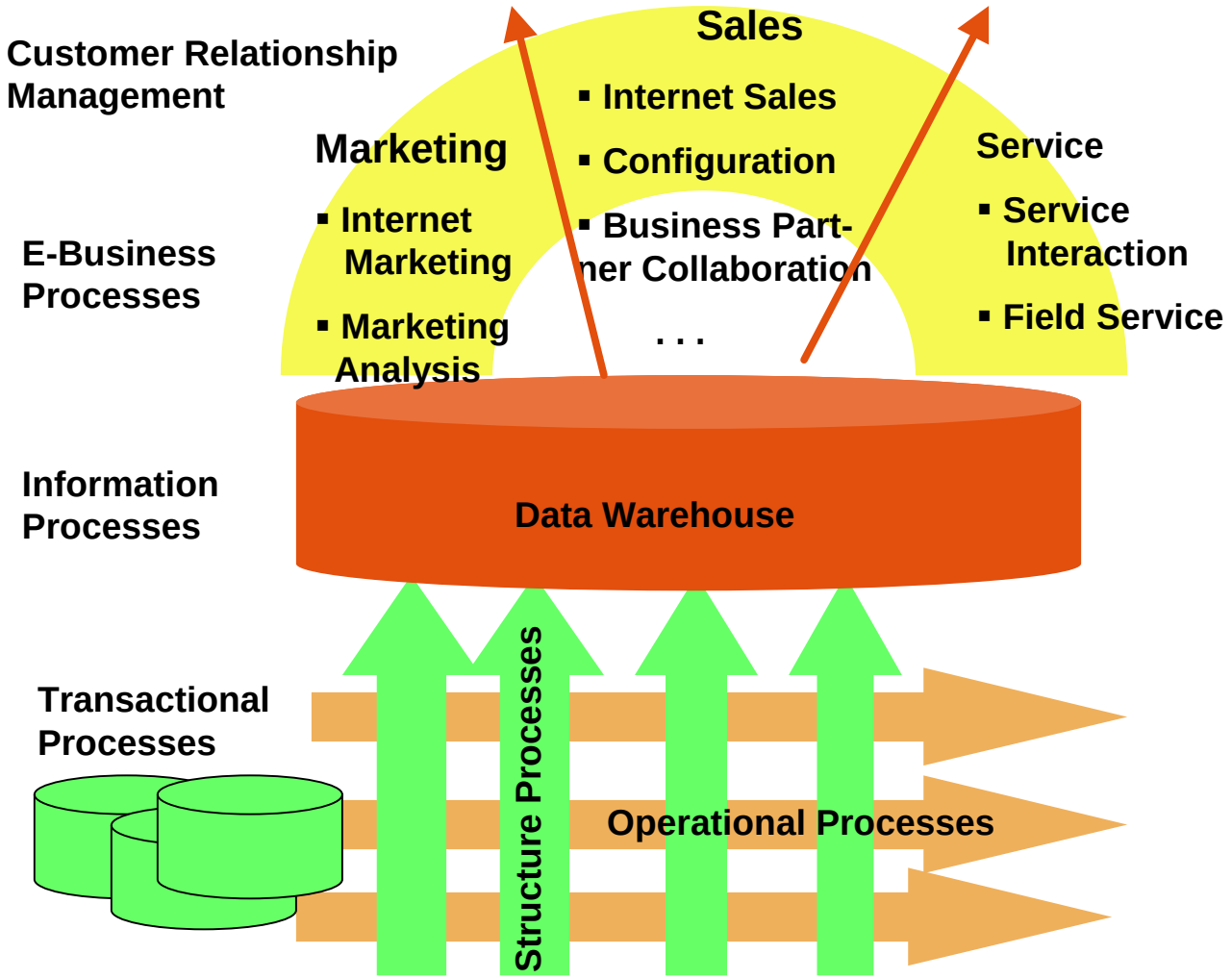
Research Project covering areas of

- Business Case / Cost Benefit Analysis
- DWH Introduction
- Organizational Concepts
- Information Requirement Analysis
- Data and Information Quality
- Operational Data Warehouse

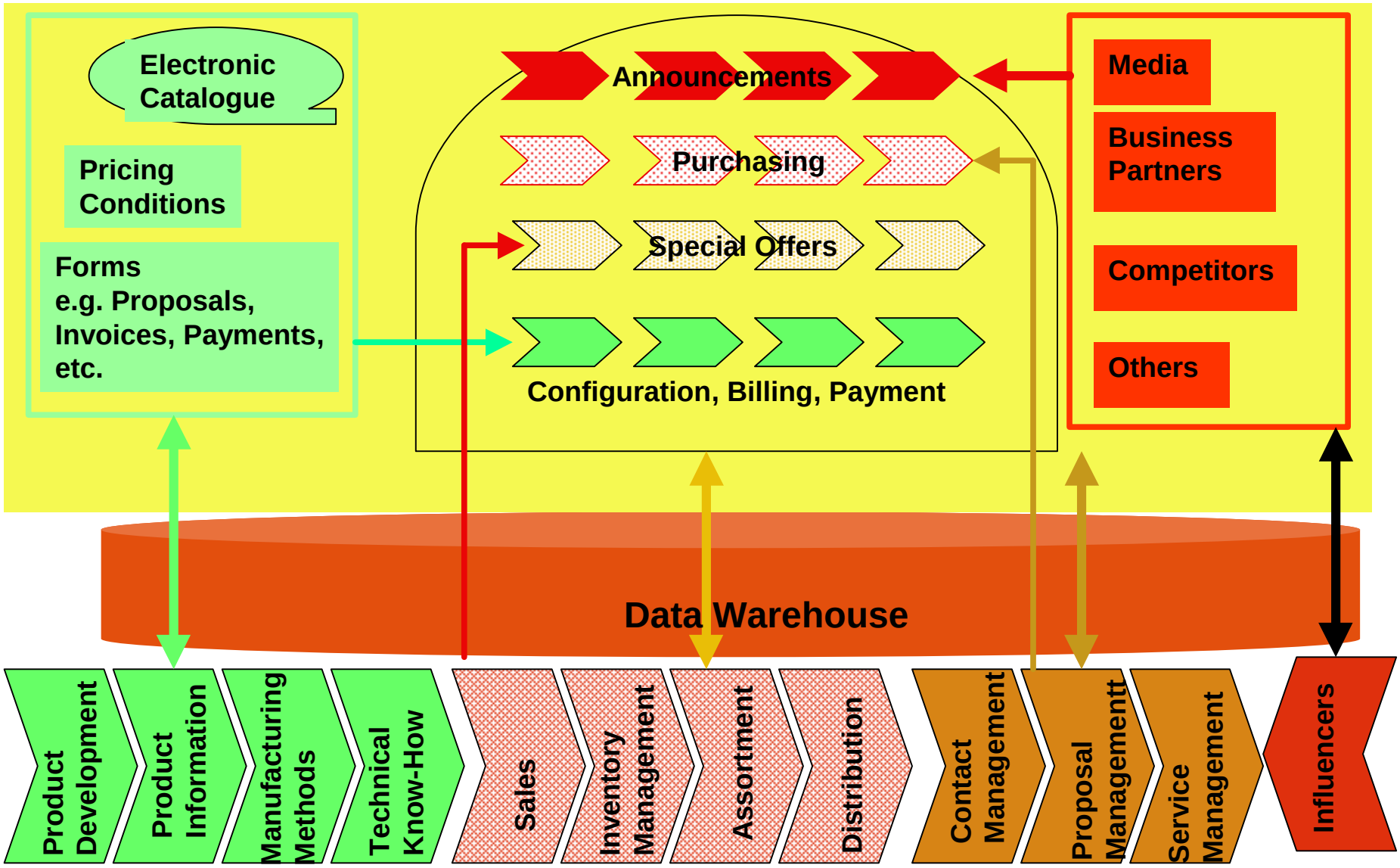
Content

Turntable for End-To-End Business Process Integration

- Provide value-adding information at any time, any location, any level of detail, any individual request
- Decompose traditional value chains and recompose according to information chains
- Add value to your business partners by providing reliable, high-quality information



Traditional Products + Information = Solution

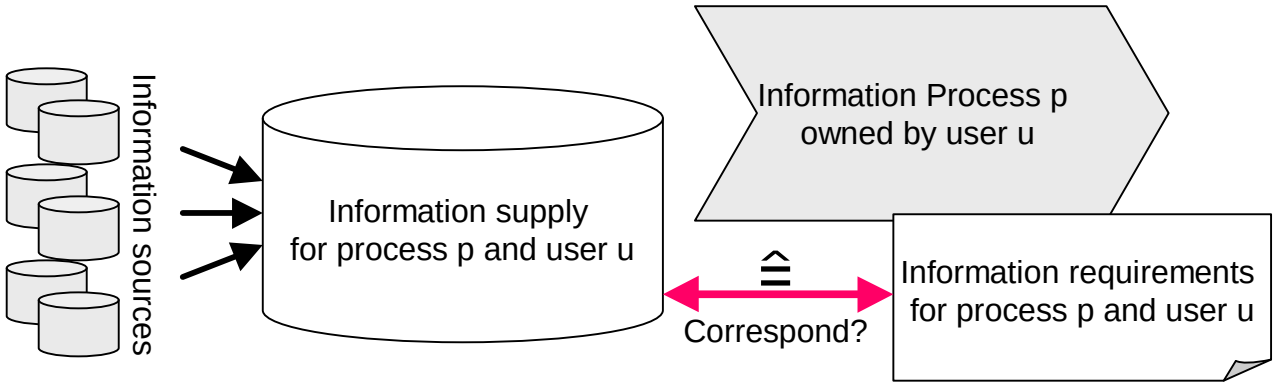


Information Quality

One accepted approach is focused on user and product fitness for use (Juran 1998)

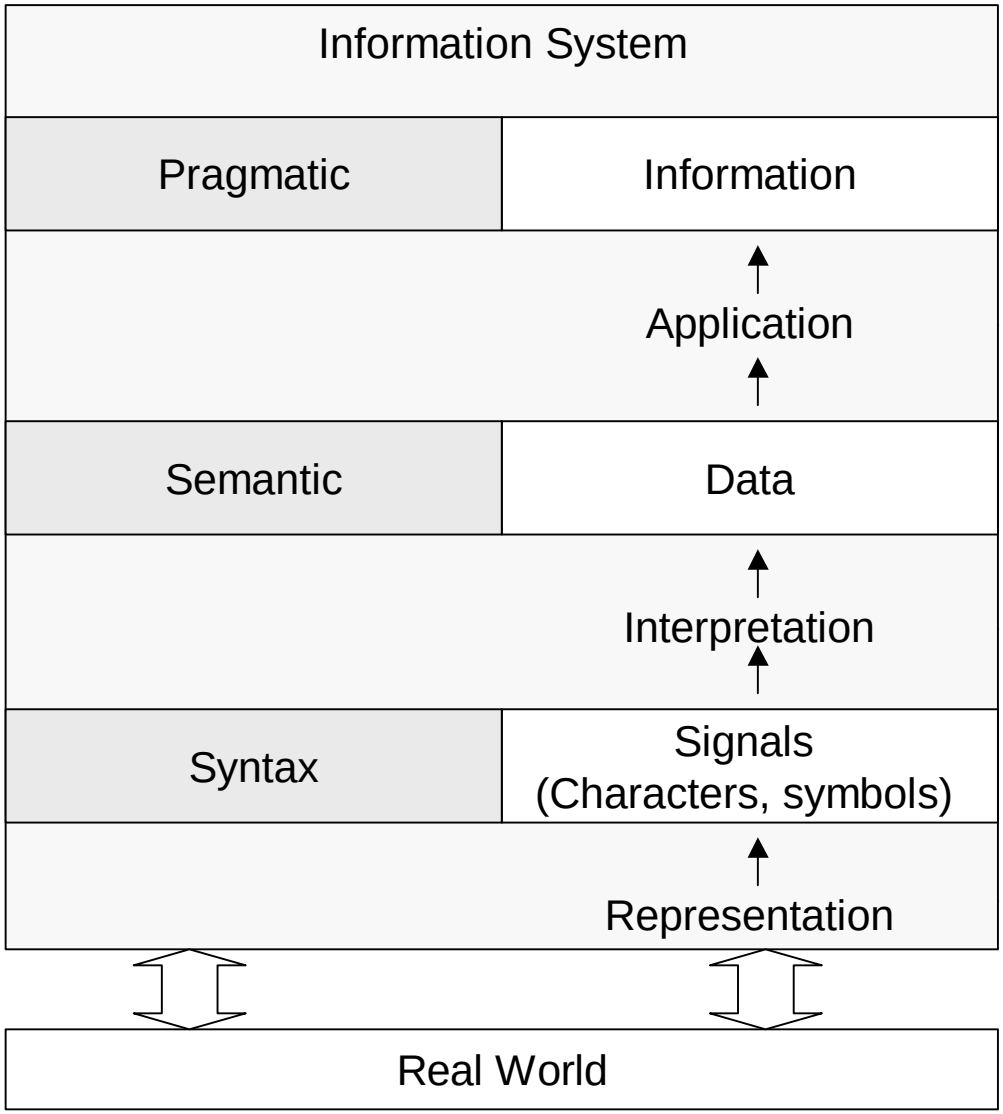
Characteristics of information **which meet user requirements** and thereby provide user satisfaction

Absence from information **deficiencies** that result in user dissatisfaction



⇨ Framework to specify information requirements
and measure information supply
on a uniform and consistent base (set of relevant information characteristics)

Information, Data and Signals



Example

Product configuration

09/03/00 DD/MM/YY

„Expected Date, when the product can be delivered to the customer“

09/03/00

String containing numbers NN/NN/NN

Expected Delivery Date

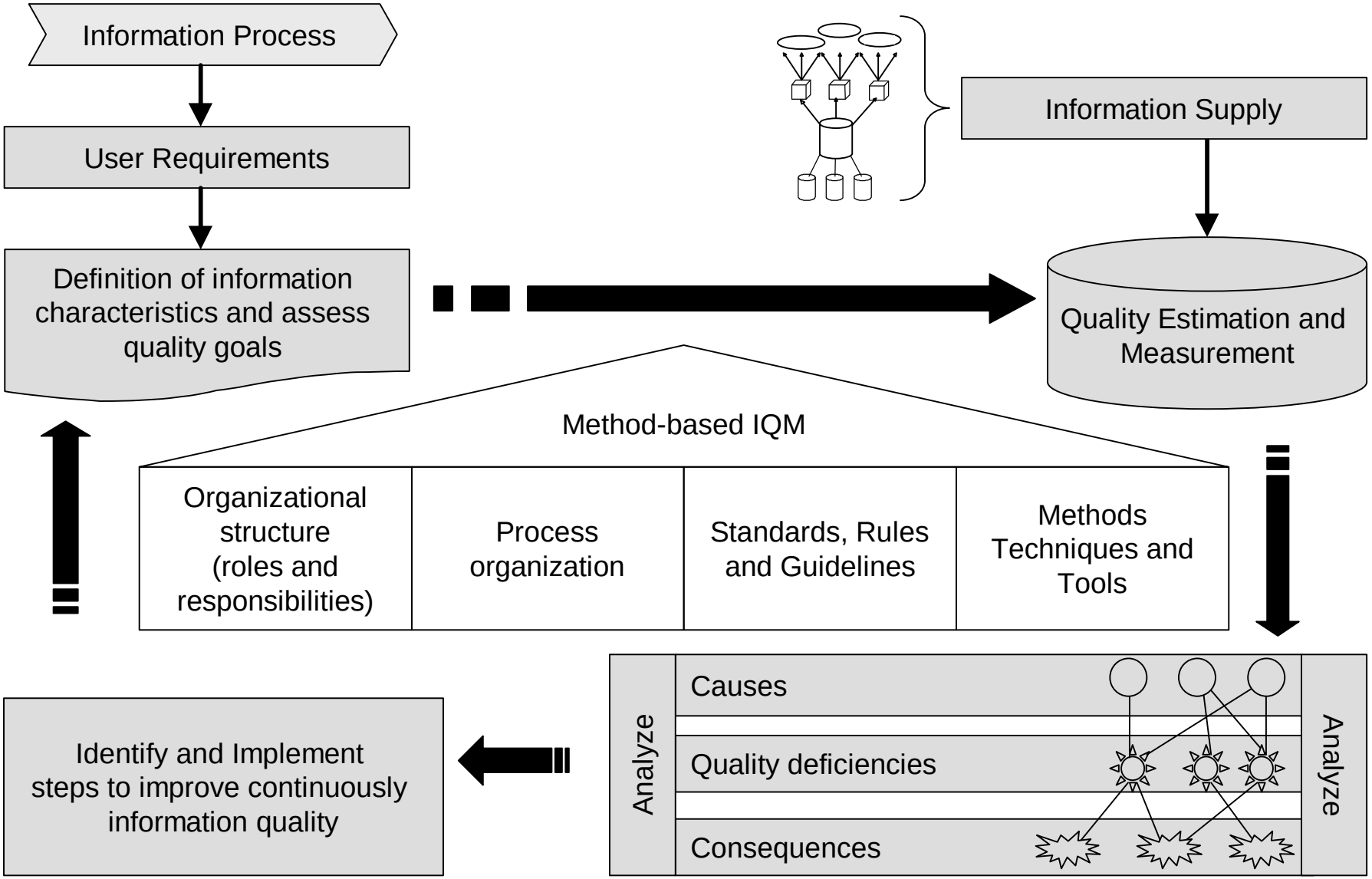
Information Quality Characteristics

Semiotic Level	Information Quality Characteristics	Measurement
Pragmatic	Relevance, completeness, timeliness	Information Process (Usage)
Semantic	Interpretability, accuracy, consistent data values, complete data values, precise data definitions, objective, believability, reliable, easy to understand	Real World (Comp. with experience and historical data)
Syntax	Syntactical correctness, consistent representation, security, accessibility	Syntactical Standards and Agreements

Research Questions

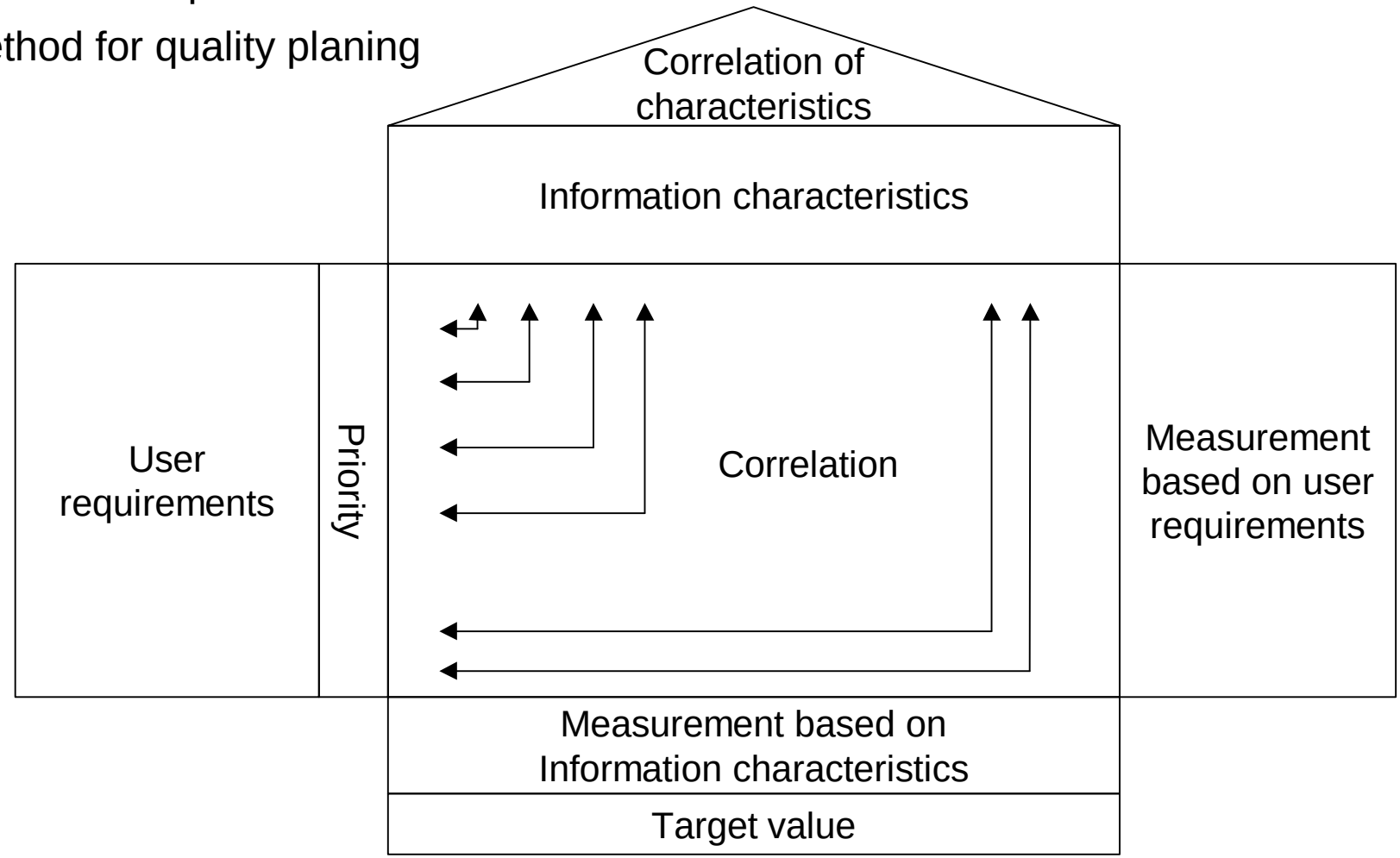
- Precise Definition of information characteristics
- Correlation between user requirement and information characteristics
- Methods and techniques for measuring information characteristics

Method-based IQM

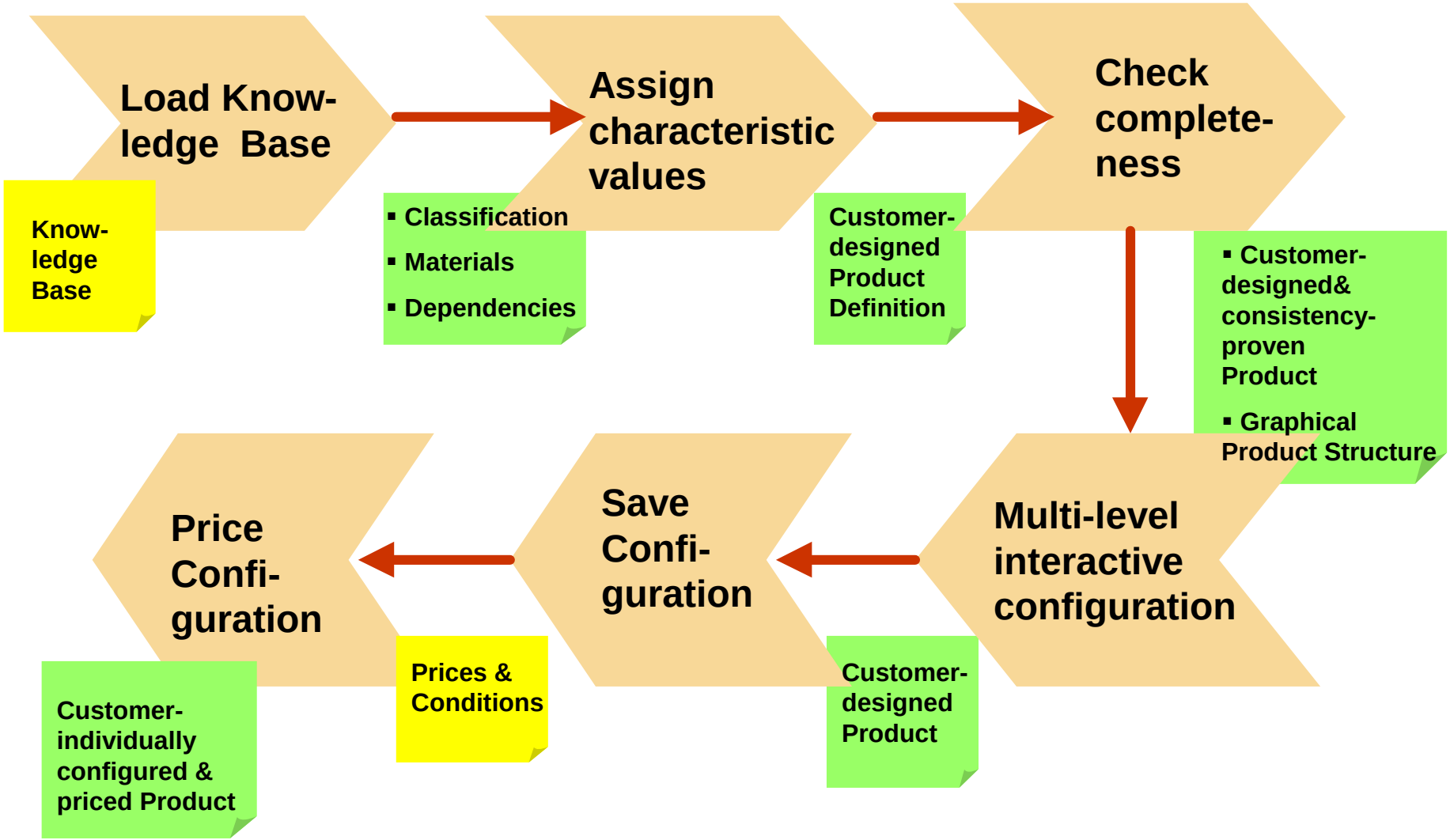


Quality Function Deployment

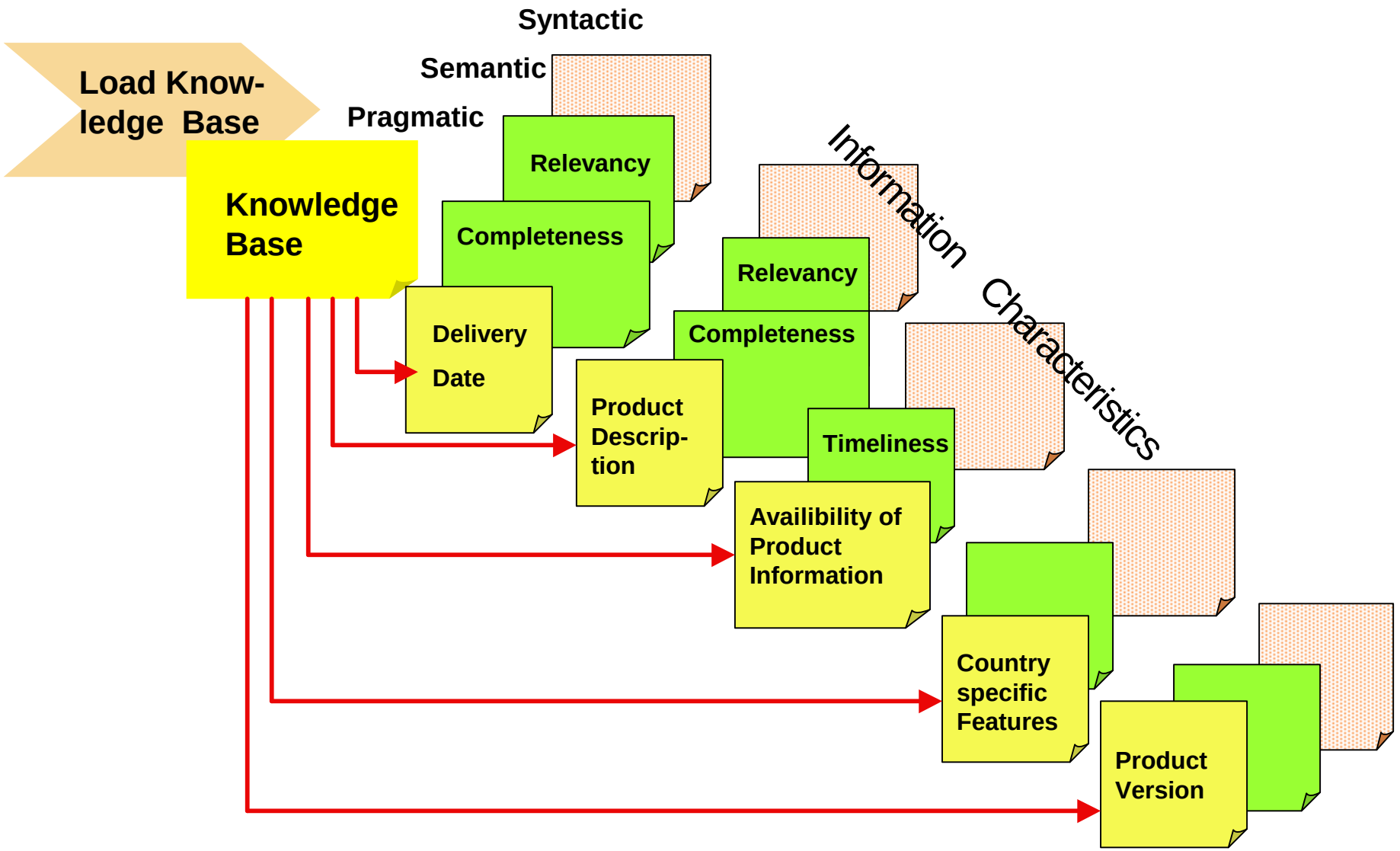
- Introduced by Akao in Japan in 1966 (Manufacturing Sector)
- Customer requirements as focus
- Method for quality planing



Configuration as an example for Information Quality Measurement



Configuration as an example for Information Quality Measurement



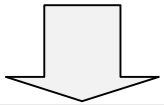
Example "Product Configuration"

Pragmatic	Relevancy	Completeness	Timeliness
Delivery Date (always)		X	
Product description (if possible)	X	X	
Availability of product information			X
...			
Measurement % of	unused information	additional information requests	not timely available information

Semantic	Interpretability	Accuracy	Complete data Values	Consistent data Values
Precise Definition for Delivery Date	X	X		X
Correct Delivery Date (no default value)		X	X	X
One Delivery Date for a product		X		
...				
Measurement % of	Non interpretable data values	Incorrect data values	Empty values	Inconsistent data values

Further Research

- Generic “House of Quality” for each semiotic level
- Integration of level specific “Houses of Quality” in a “Model of Information Quality”
- Relations between requirements, characteristics and semiotic levels



‚Model of Information Quality‘ as basic framework

- Methods and techniques for measuring information characteristics
- Methods to apply the Model for specific information processes (e. g. Product Config.)
- Validation of research results in practical projects (e. g. Customer Relationship Management)