

# Lösung

## Übungsblatt 4 Wissensmanagement

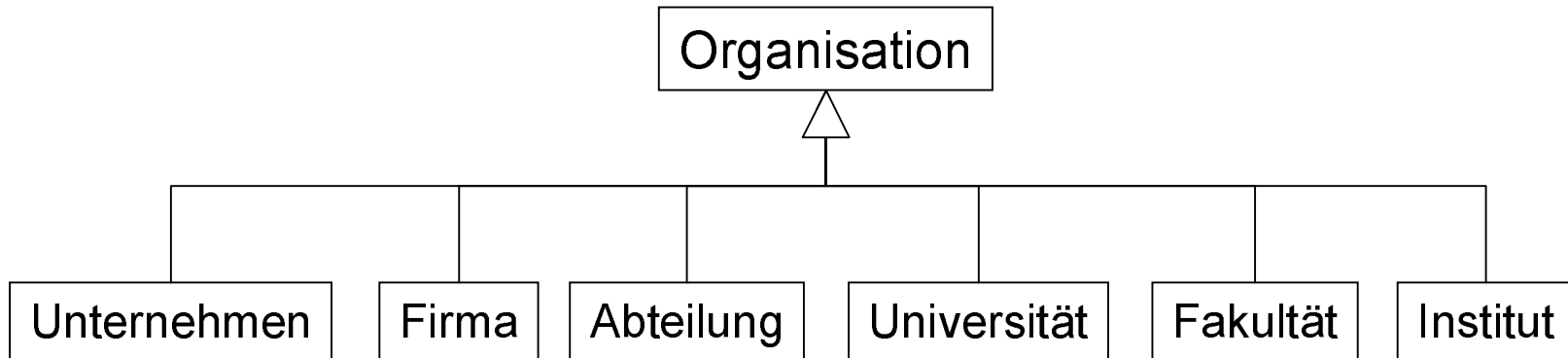
**Termin: 11. 07. 2001**

# 1. WORDNET – Semantische Relationen

- Synonymie: (gleichbedeutend)
  - Unternehmen, Firma
  - Doktorand, PhD-Student
- Antonymie: (gegenteilig)
  - Hardware, Software
- Homonymie: (gleiche Schreibweise, unterschiedliche Bedeutung)
  - Drucker (Hardware Peripherie)
  - Drucker (Person, (Gutenberg))

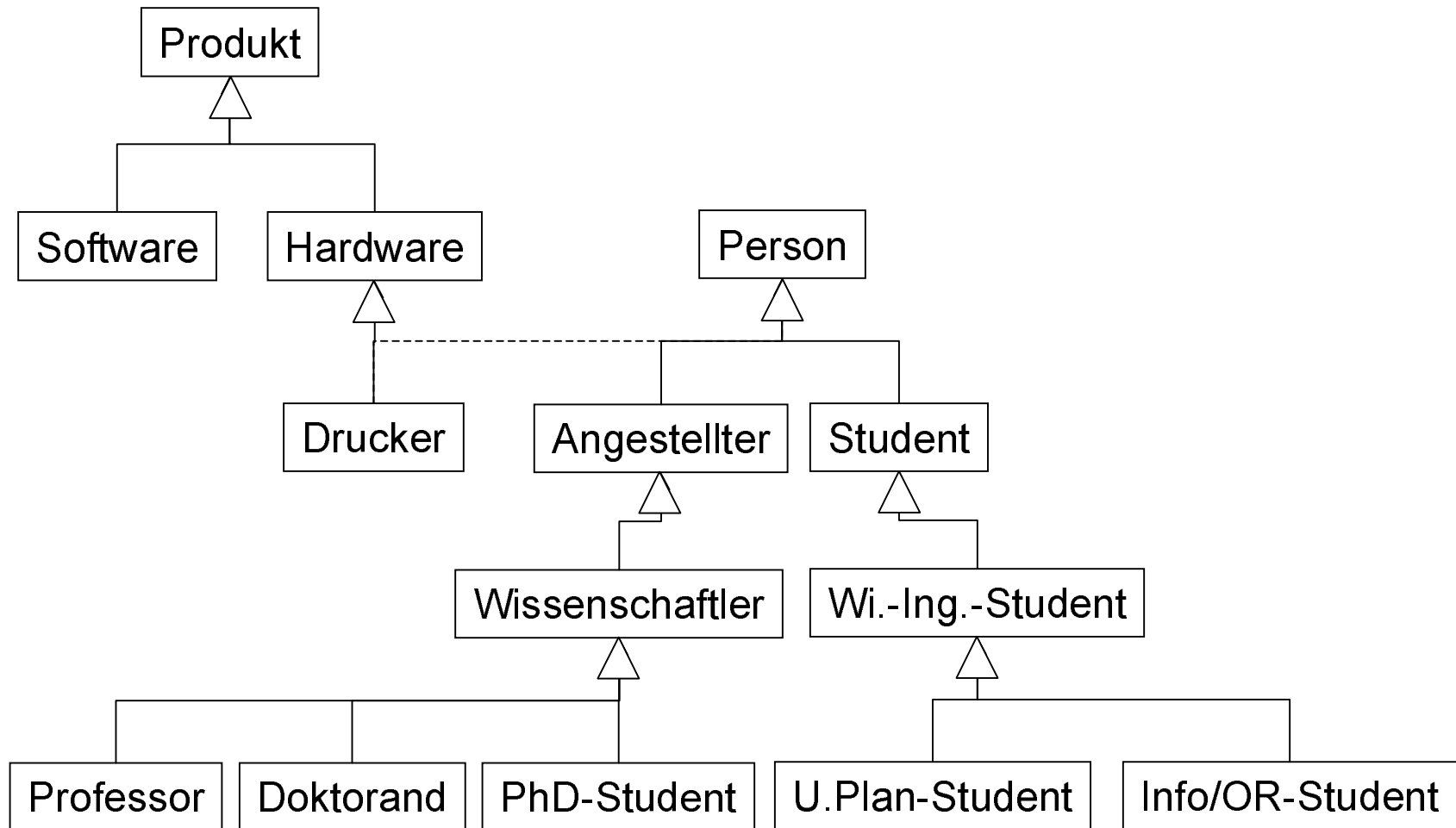
# 1. WORDNET – Semantische Relationen

- Hyponymie/Hyperonymie: (is-a, Unter-/Oberbegriff)



# 1. WORDNET – Semantische Relationen

- Hyponymie/Hyperonymie: (is-a, Unter-/Oberbegriff)



# 1. WORDNET – Semantische Relationen

- Meronomie: (part-of, Teil-von)
  - Abteilung, Unternehmen
  - Abteilung, Firma
  - Institut, Universität
  - Institut, Fakultät
  - Fakultät, Universität

evtl. auch:


  - Professor, Universität
  - Student, Universität
  - Angestellter, Abteilung
- Troponomie: (analog Hyponomie für Verben)
  - keine Verben!

## 2. Case-based reasoning

The most important elements in a case-based system development are:

- case description
  - including indexed and unindexed features
- similarity measure
  - including retrieving mechanism

## 2. Case-based reasoning - case description

**Patient Ref #:** 1024  
**Patient Name:** *John Doe*  
**Address:** 12 Elm Street  
**Next of Kin:** *Jane Doe*  
**Photo:** 

**Age:** 53  
**Sex:** Male  
**Weight:** 225 lbs  
**Height:** 5' 11"  
**Blood Type:** A neg.  
 ...

### unindexed features

Not predictive & not used for retrieval, they provide background information to users  
provide valuable contextual information and lessons learned

### indexed features

Predictive and used for retrieval

---

Case vocabulary - the features that describe a case

## 2a. Case-based reasoning

unindexed features

What are the name, telefon and e-mail of researcher  
from research group EA,  
that has research interest in research topic Knowledge  
Management and  
work on project with the name Ontobroker ?

indexed features

**Features are extracted from  
Competency Questionnaires**

CASE DESCRIPTION

type:Person

name:STRING

telefon:STRING

email:STRING

forschungsgruppe:Forschungsgruppe

forscht\_in\_gebiet:Forschungsgebiet

arbeitet\_in\_projekt:Projekt



## 2a. Case-based reasoning - case base (instances)

### CASE BASE:

type:Professor  
name:Taylor  
telefon:111  
email:jim@mail  
forschungsgruppe:E1  
forscht\_in\_gebiet:RA5  
arbeitet\_in\_projekt:Edu

type:Professor  
name:Taylor1  
telefon:1121  
email:jim@mail  
forschungsgruppe:EA  
forscht\_in\_gebiet:KDD  
arbeitet\_in\_projekt:Edu

type:Wissenschaftlicher\_Mitarebieter  
name:Taylor  
telefon:111  
email:jim@mail  
forschungsgruppe:E2  
forscht\_in\_gebiet:RA3  
arbeitet\_in\_projekt:Edu

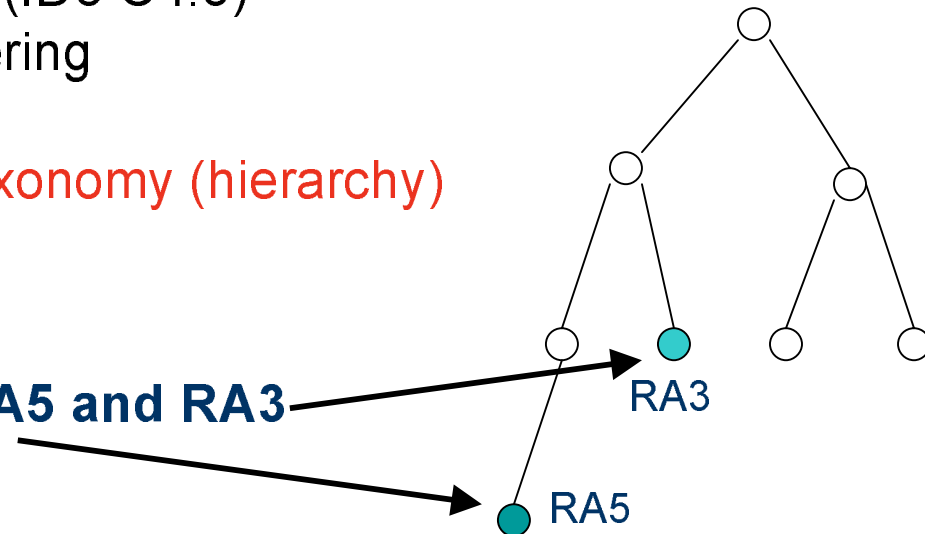
type:Professor  
name:Taylor2  
telefon:1131  
email:jim@mail  
forschungsgruppe:EA  
forscht\_in\_gebiet:RT  
arbeitet\_in\_projekt:Edu

## 2b. Case-based reasoning - similarity

- Degree of similarity = utility or reusability of the solution
- Goal of similarity modelling:
  - provide a good approximation
  - close to real reusability
  - and easy to compute
- Different approaches depending on case representation
  - k-nearest neighbour
  - ML – classifiers (ID3 C4.5)
  - Statistical clustering
  - fuzzy sets/logic

similarity in a taxonomy (hierarchy)

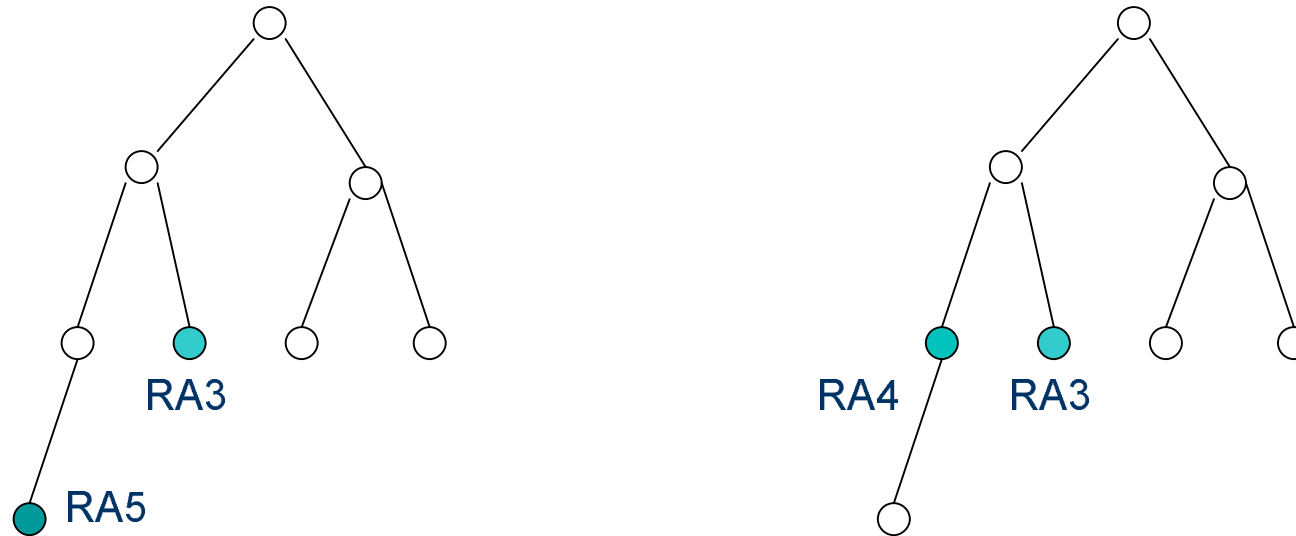
**How similar are RA5 and RA3**



Wissensmanagement SS2001

## 2b. Case-based reasoning - similarity

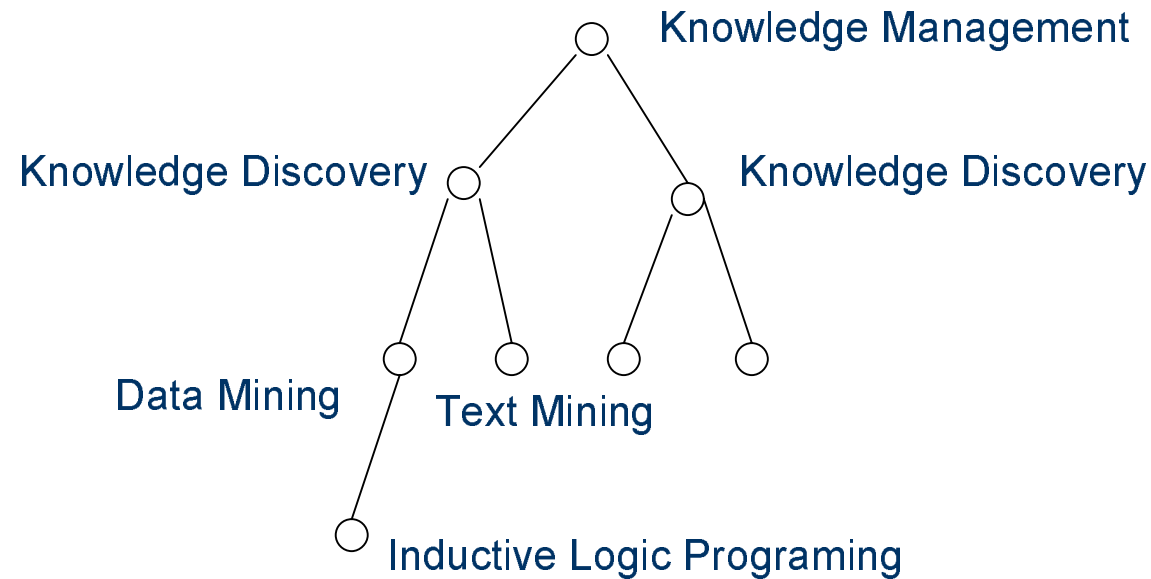
- Princip: Similarity in a hierarchy is based on the distance between two nodes



**RA3 is more similar to RA4 than to RA5**

## 2b. Case-based reasoning - similarity

- Research areas are organised in a hierarchy:



Inductive Logic Programming is more similar to Data Mining than to Text Mining

## 2c. Case-based reasoning - process

- Process-model:
  1. Retrieve
    - find all similar cases
  2. Reuse
    - adapt to the new case-problem
  3. Revise
    - evaluate solution for the new problem
  4. Retain
    - store the acquired experience

## 2c. Case-based reasoning - retrieve

### Case base

C1:

type:Professor

name:Taylor

telefon:111

email:jim@mail

forschungsgruppe:E1

forscht\_in\_gebiet:Data Mining

arbeitet\_in\_projekt:Edu

C2:

type:Wissenschaftlicher\_Mitarebieter

name:Dilan

telefon:111

email:jim@mail

forschungsgruppe:E2

forscht\_in\_gebiet:Text Mining

arbeitet\_in\_projekt:Edu

sim2

sim1

### New problem:

**Who does research in the  
research area Inductive  
Logic Programing?**

type:Person

name:x

telefon:x

email: x

forschungsgruppe:x

forscht\_in\_gebiet: Inductive

Logic Programing

arbeitet\_in\_projekt:x

**=> C1 and C2**

## 2c. Case-based reasoning - reuse

### Case base

C1:

type:Professor

name:Taylor

telefon:111

email: jim@mail

forschungsgruppe:E1

forscht\_in\_gebiet: Data Mining

arbeitet\_in\_projekt:Edu

C2:

type: Wissenschaftlicher\_Mitarebieter

name:Dilan

telefon:111

email: jim@mail

forschungsgruppe:E2

forscht\_in\_gebiet: Text Mining

arbeitet\_in\_projekt:Edu

New problem:

Who does research in the research  
area Inductive Logic Programing?

**sim2> sim1**

**=> Solution is:**

type: Professor

name: Taylor

telefon: 111

email: jim@mail

forschungsgruppe: E1

forscht\_in\_gebiet: Inductive  
Logic Programing

arbeitet\_in\_projekt: Edu

## 2c. Case-based reasoning - revise

### Case base

C1:

type:Professor

name:Taylor

telefon:111

email: jim@mail

forschungsgruppe:E1

forscht\_in\_gebiet: Data Mining

arbeitet\_in\_projekt:Edu

C2:

type: Wissenschaftlicher\_Mitarebieter

name:Dilan

telefon:111

email: jim@mail

forschungsgruppe:E2

forscht\_in\_gebiet: Text Mining

arbeitet\_in\_projekt:Edu

New problem:

Who does research in the research  
area Inductive Logic Programing?

### Is the new solution correct

type: Professor

name: Taylor

telefon: 111

email: jim@mail

forschungsgruppe: E1

forscht\_in\_gebiet: Inductive  
Logic Programing

arbeitet\_in\_projekt: Edu



## 2c. Case-based reasoning - retain

### Case base

C1:

type:Professor

name:Taylor

telefon:111

email: jim@mail

forschungsgruppe:E1

forscht\_in\_gebiet: Data Mining

arbeitet\_in\_projekt:Edu

C2:

type: Wissenschaftlicher\_Mitarebieter

name:Dilan

telefon:111

email: jim@mail

forschungsgruppe:E2

forscht\_in\_gebiet: Text Mining

arbeitet\_in\_projekt:Edu

New problem:

Who does research in the research  
area Inductive Logic Programing?

### New case

C3:

type: Professor

name: Taylor

telefon: 111

email: jim@mail

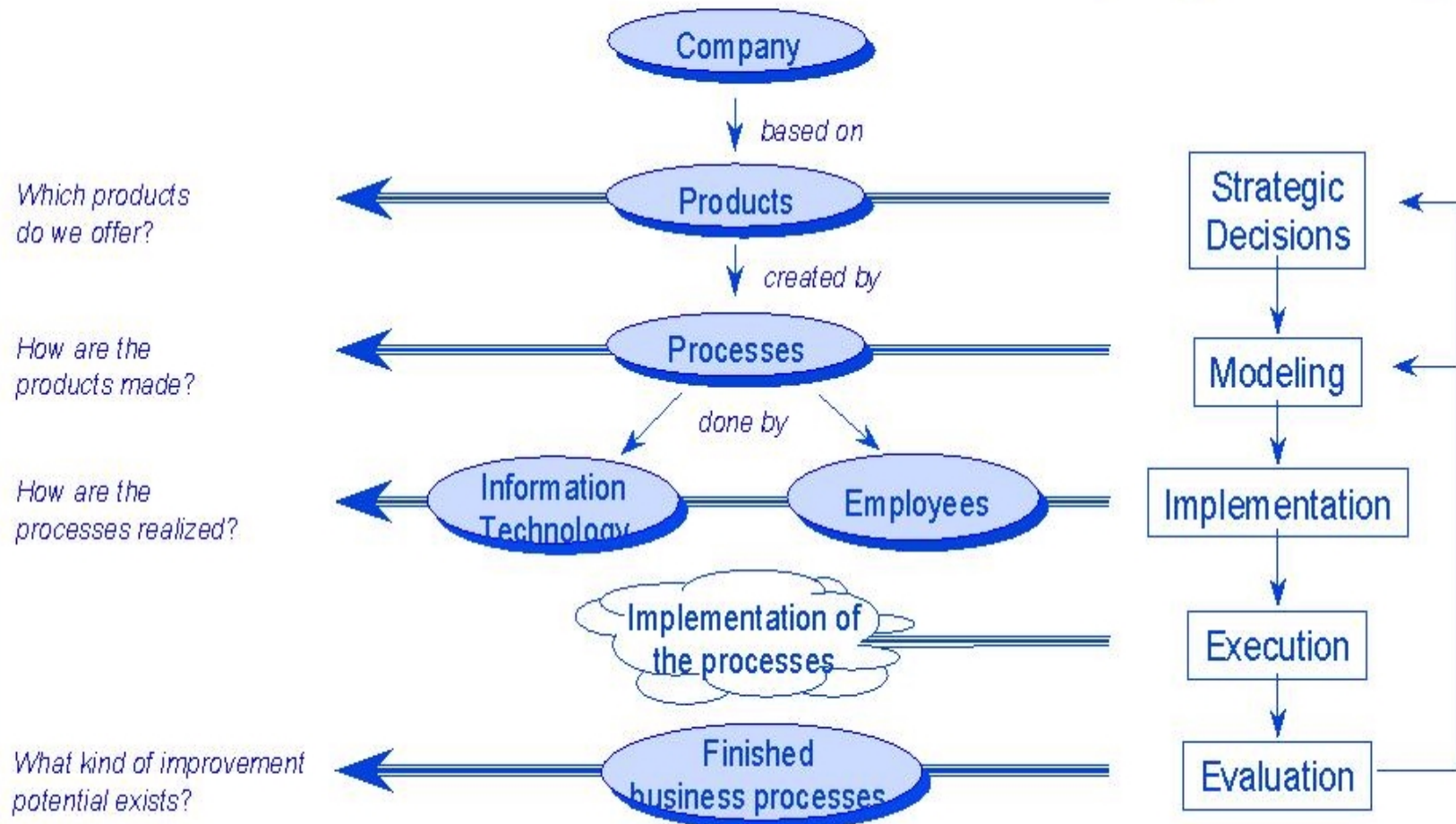
forschungsgruppe: E1

forscht\_in\_gebiet: Inductive  
Logic Programing

arbeitet\_in\_projekt: Edu

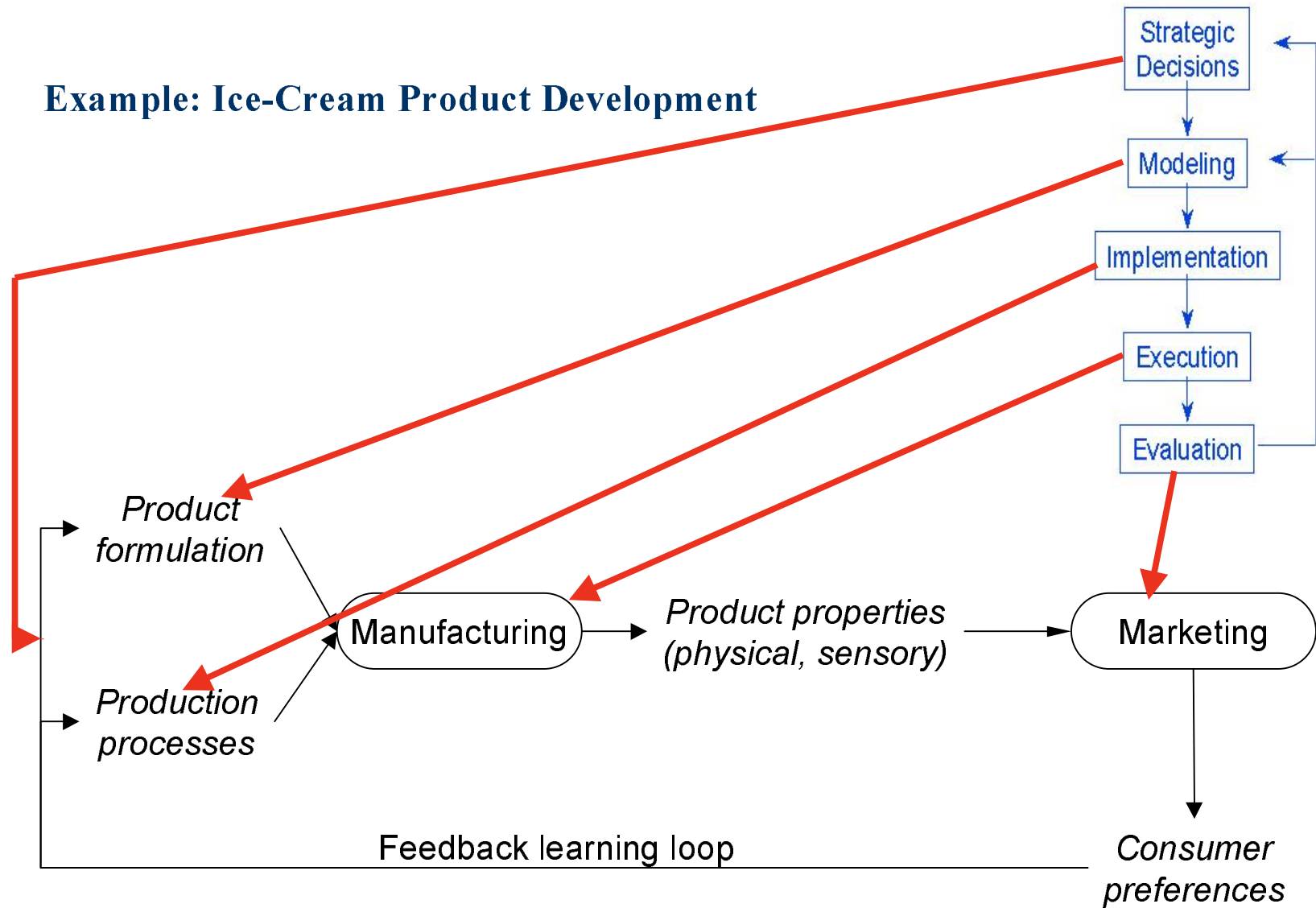
### 3. Business process management

#### BPMS-METHODOLOGY: PRODUCT/PROCESS-PHILOSOPHY [Karagiannis, 1994]



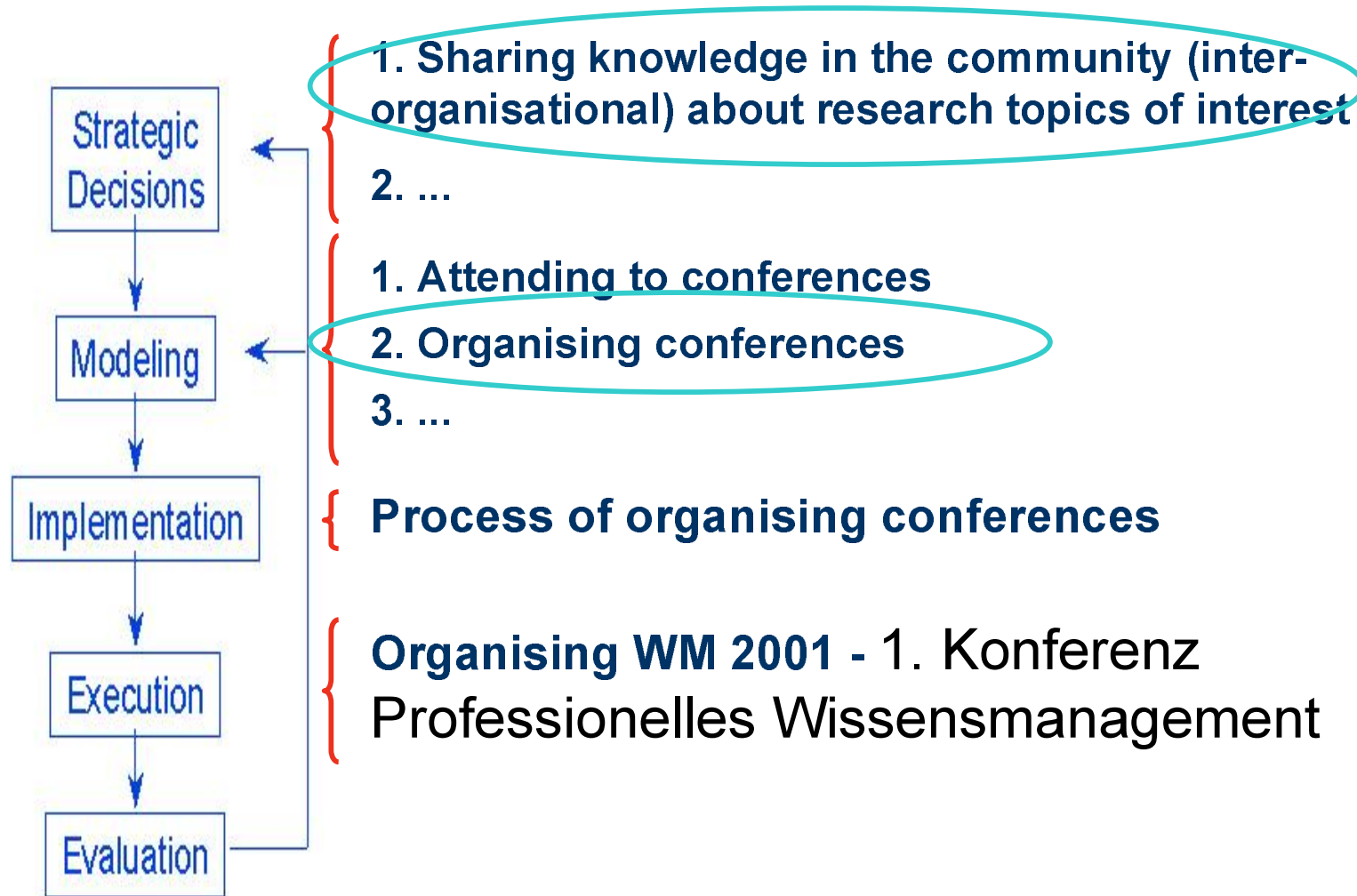
## 3a. Business process management

### Example: Ice-Cream Product Development



## 3a. Business process management

Real example: Research organisation

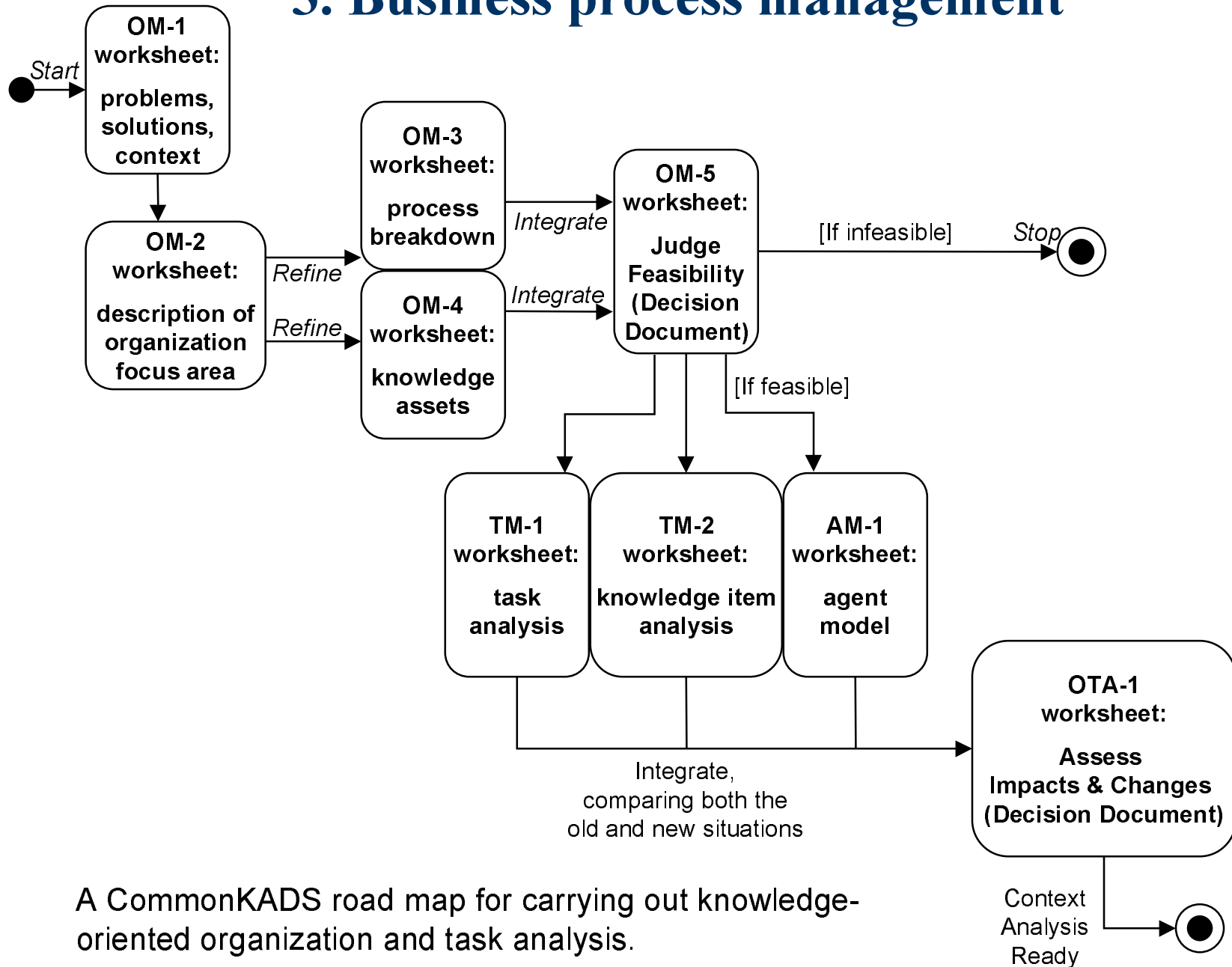


## 3b. Business process management

**Real example:** Sharing knowledge in a research community

**CommonKADS feasibility analysis**

### 3. Business process management



A CommonKADS road map for carrying out knowledge-oriented organization and task analysis.

## Research in a research organisation

Organization Model	Problems and Opportunities Worksheet OM-1
PROBLEMS AND OPPORTUNITIES	<ul style="list-style-type: none"> <li>* To enable more efficient sharing of knowledge in the research community</li> <li>* Searching for competent persons is most critical</li> </ul>
ORGANIZATIONAL CONTEXT	<p>Vision and strategy:</p> <ul style="list-style-type: none"> <li>* Achieve a situation as depicted later</li> </ul> <p>External factors:</p> <ul style="list-style-type: none"> <li>* A lot of information are on web pages (projects and personal pages)</li> <li>* Providing of information is decentralised (each community member can provide)</li> <li>* Searching for information should be focused -</li> <li>* Inter-discipline projects</li> </ul> <p>Major value drivers:</p>
SOLUTIONS	<p>Solution 1: Improve sharing through communication (conference, social events)</p> <p>Solution 2: Improve sharing through Information Portal (informal)</p> <p>Solution 2: Improve sharing through Semantic Portal (formal, structured information)</p> <p>Solution 3: Improve sharing using skill ontology</p>

Worksheet OM-1: Problems, organizational context and possible solutions

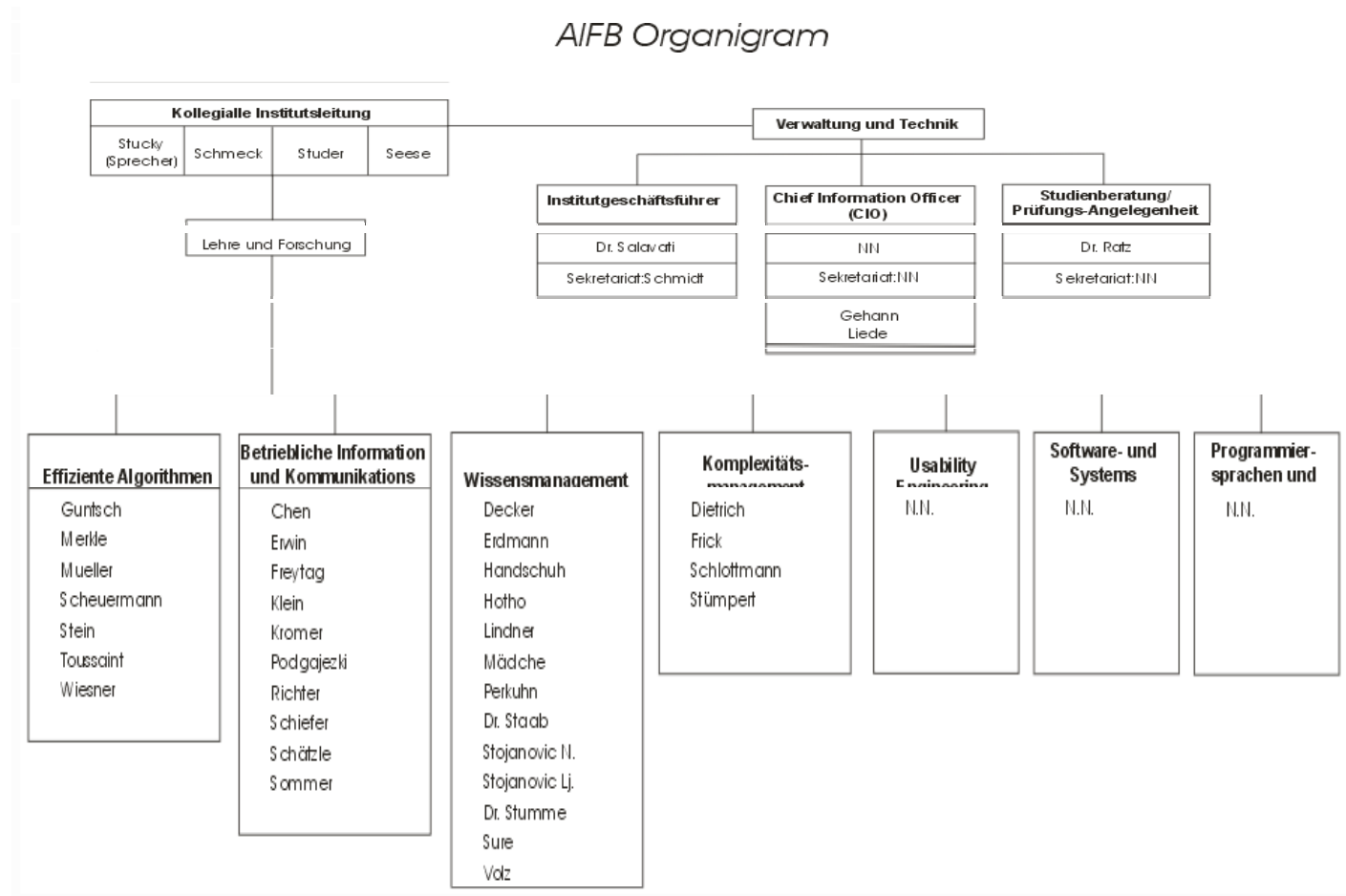
## Research in a research organisation

Organization Model	Variant Aspects Worksheet OM-2
STRUCTURE	See next slide
PROCESS	Sharing knowledge globally involves three major phases. See high-level breakdown in worksheet OM-3
PEOPLE	Mainly research staff is involved in sharing knowledge. Could be divided in knowledge provider, organizer, sharer and user
RESOURCES	<ul style="list-style-type: none"> <li>* 1. Information systems and other computing resources</li> <li>* 2. Equipment and materials</li> <li>* ...</li> </ul>
KNOWLEDGE	Linked to the different functional roles listed above
CULTURE & POWER	* Knowledge-sharing environment

Worksheet OM-2: Description of organizational aspects that have an impact on and/or are affected by chosen knowledge solutions



## Research in a research organisation

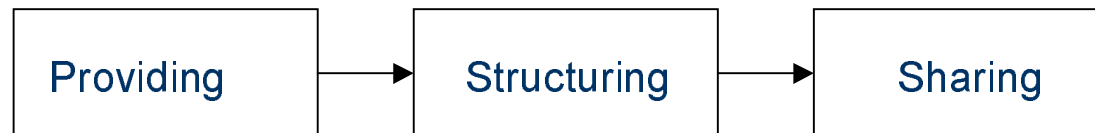


A typical organization structure of a research organisation

## Research in a research organisation

### Breakdown of Business Processes

- Business process is decomposed into several tasks
  - Provide knowledge: knowledge is acquired in a suitable form
  - Structure knowledge: knowledge is stored
  - Share knowledge: access to knowledge



Process of sharing-knowledge (high-level view)

## Research in a research organisation

Organization Model		Process Breakdown Worksheet OM-3				
NO.	TASK	PER-FORMED BY	WHERE?	KNOWL-EDGE ASSET	KNOWL-EDGE INTENSIVE	SIGNIFICANCE
1	Provide knowledge	Researchers	-	Research topics Research projects Person	Yes	Any organizing must subsequently pass all listed tasks successfully
2	Structure knowledge	Researchers, mainly	-		Medium	See above
3	Share knowledge	Researchers,	-	Research topics Research projects Person	High	See above

Worksheet OM-3: Top-level task breakdown for sharing-knowledge process

## Research in a research organisation

Organization Model		Knowledge Assets Worksheet OM-4				
KNOWL- EDGE ASSET	POSSESSED BY	USED IN	RIGHT FORM?	RIGHT PLACE?	RIGHT TIME?	RIGHT QUALITY?
Research topics	Researchers	Sharing knowledge	No (does not exist)	No	No	No
Person	Researchers	Sharing	No	Yes (web)	Yes	No
Projects	Researchers	Sharing knowledge	No (not structured, machine non- processabl e)	Yes (web)	Yes	No (informal, redundant, inconsistent )

Worksheet OM-4: An excerpt from the knowledge assets analysis.

## Research in a research organisation

### OM-5: First Decision Document: Knowledge-Improvement Scenarios

- Major knowledge bottlenecks have been identified
  - how to access to information (How to find useful knowledge) - used in TM2
- Different opportunities for knowledge improvement have been identified
- Sharing phase is the most crucial knowledge intensive task
  - all the information about projects and research topics are stored on web pages
  - how to find
    - whose research interest is topic X
    - who has experience in topic X (work on project)

## Research in a research organisation

- Different knowledge-improvement scenarios
  - Portal scenario
    - Information are in the form of personal, project data on the web and standard search is performed
  - Semantic Portal scenario
    - Information are in the form of personal, project data on the web, but in machine processible form
  - Skill-management scenario
    - Skill-ontology is developed and used in searching

## Research in a research organisation

Organization Model	Checklist for Feasibility Decision Document: Worksheet OM-5
BUSINESS FEASIBILITY	Based on the organization-model analysis, the most important knowledge bottlenecks in sharing knowledge process - how to access to information have been clarified. In addition, a number of different scenarios for knowledge improvement opportunities have been identified: Portal, Semantic Portal, Skill-management scenario. Each scenario represents a promising, feasible (to varying degrees), but different solution direction.
TECHNICAL FEASIBILITY	IT infrastructure is available For Semantic Portal Semantic Web infrastructure is needed
PROJECT FEASIBILITY	In ontology-based solutions Ontology-engineering has to be performed
PROPOSED ACTIONS	<ul style="list-style-type: none"> <li>* Further rank and prioritize the knowledge improvement scenarios, by detailed task / agent / knowledge item analysis, leading to both short-term and mid-term recommendations and actions.</li> <li>* Consider how the current IT architecture can be gradually extended to a broader and more knowledge-intensive support environment.</li> <li>* Select a first system module with a high potential impact that can be developed relatively quickly.</li> </ul>

Worksheet OM-5: First decision document, comprising various feasible improvement scenarios for sharing-knowledge process

## Research in a research organisation

Task Model	Task Analysis Worksheet TM-1
ORGANIZATION	Part of the knowledge-sharing process
GOAL AND VALUE	This task aims to distribute knowledge in a pull or push fashion to corresponding work-places
DEPENDENCY AND FLOW	Input tasks: Structuring knowledge Output tasks: Knowledge usage
OBJECTS HANDLED	Input objects: Structured knowledge Output objects: Pull/Push distribution of knowledge
TIMING AND CONTROL	

Worksheet TM-1: Analysis of the “sharing” task within the knowledge-sharing process (Part I)



## Research in a research organisation

Task Model	Task Analysis Worksheet TM-1 (continued)
AGENTS	From various functional areas; cf. OM-2 and OM-3 for examples
KNOWLEDGE AND COMPETENCE	Variety of domains; cf. OM-4 and TM-2 for examples
RESOURCES	Time is a resource of prime importance here
QUALITY AND PERFORMANCE	

Worksheet TM-1: Analysis of the “sharing” task within the knowledge-sharing process (Part II)

## Research in a research organisation

### TM-2: Detailed Knowledge Bottleneck Analysis

- Each knowledge item (knowledge assets of smaller grain size) is described in a separate worksheet

## Research in a research organisation

Task Model	Knowledge Item Worksheet TM-2	
NAME POSSESSED BY USED IN DOMAIN	How to find useful knowledge (from OM-5) research 3. Sharing knowledge-sharing process	
Nature of the knowledge		Bottleneck / to be improved?
Formal, rigorous		
Empirical, quantitative	X	
Heuristic, rules of thumb	X	
Highly specialized, domain-specific	X	
Experience-based	X	
Action-based		
Incomplete	X	X
Uncertain, may be incorrect	X	
Quickly changing	X	
Hard to verify	X	X
Tacit, hard to transfer	X	X

Worksheet TM-2: Characterization of the “sharing” knowledge item. (Part I)

## Research in a research organisation

Task Model		Knowledge Item Worksheet TM-2 (continued)	
NAME POSSESSED BY USED IN DOMAIN	How to find useful knowledge research 3. Sharing knowledge-sharing process		
<b>Form of the knowledge</b>		<b>Bottleneck / to be improved?</b>	
Mind	X	X	
Paper	X		
Electronic			
Action skill			
Other			
<b>Availability of knowledge</b>			
Limitations in time			
Limitations in space	X	X	
Limitations in access			
Limitations in quality	X	X	
Limitations in form			
<i>Remarks:</i>			

Worksheet TM-2: Characterization of the “sharing” knowledge item. (Part II)

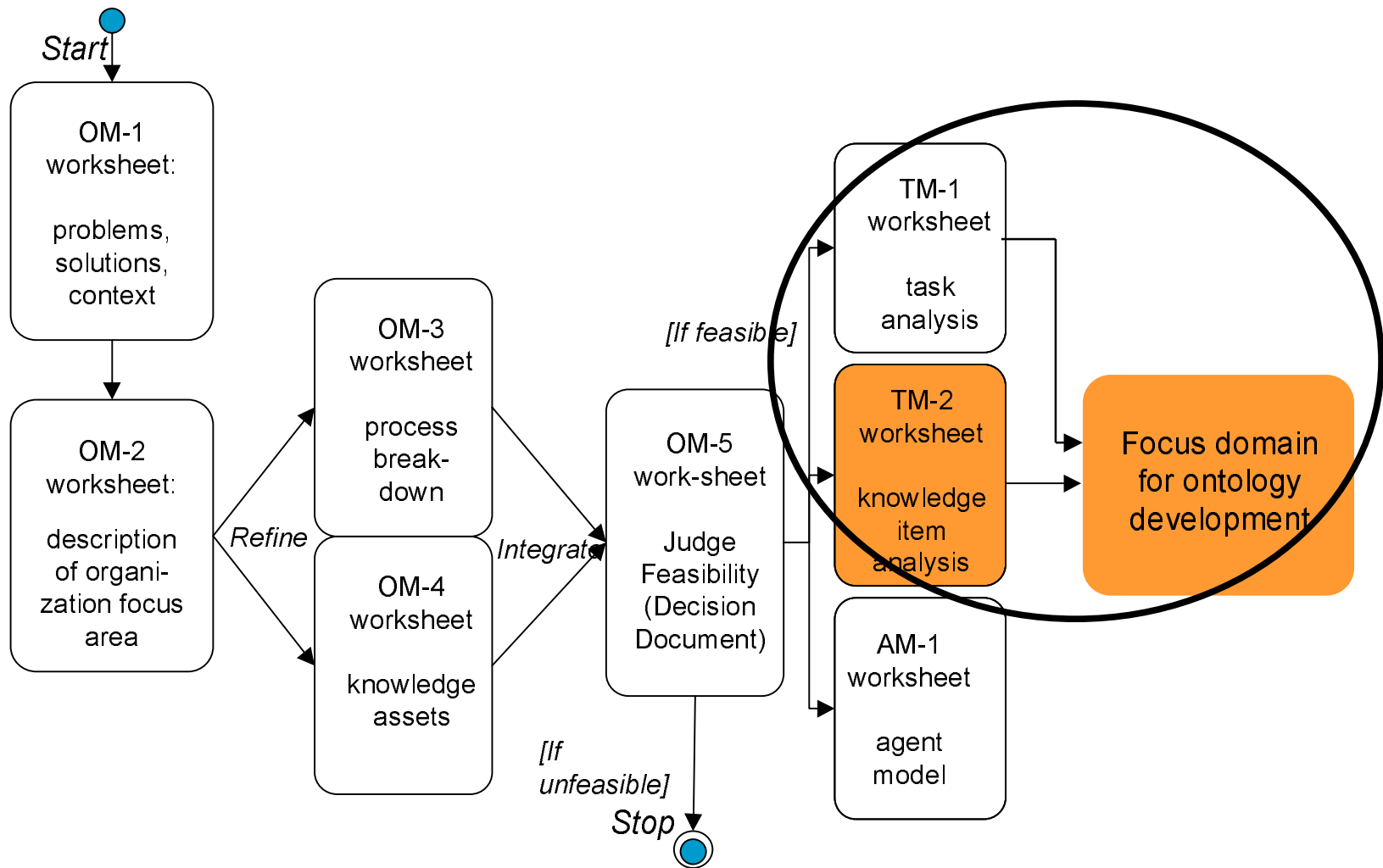
## Research in a research organisation

- OTA-1: Decision Summary of Recommendation and Actions
- Semantic Portal scenario is the most promising one

First module opportunity	Available knowledge	Technical feasibility	Potential benefits	Costs	Risks
Portal scenario	Good	Good	Good	Low	Low
...	...	...	...		
Semantic portal scenario	Good	Good	High	Medium	Medium
...	...	...	...	...	...
Skill-management	Poor	Medium	High	High	Medium

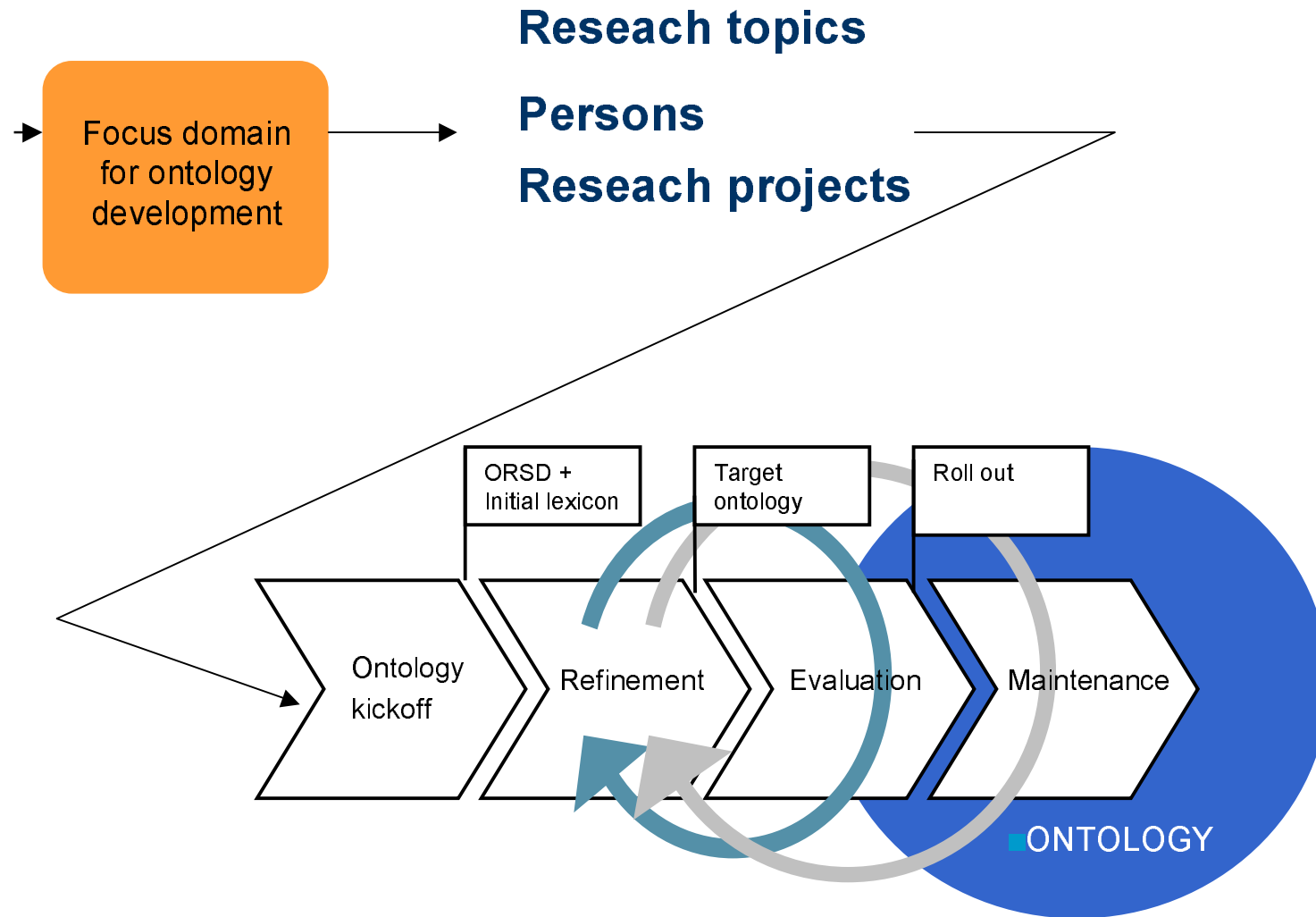
Comparison of knowledge-improvement scenarios based on task and knowledge asset analysis.

# CommonKADS - Context Modeling Road Map



## Research in a research organisation

# CommonKADS - Context Modeling Road Map



## Research in a research organisation

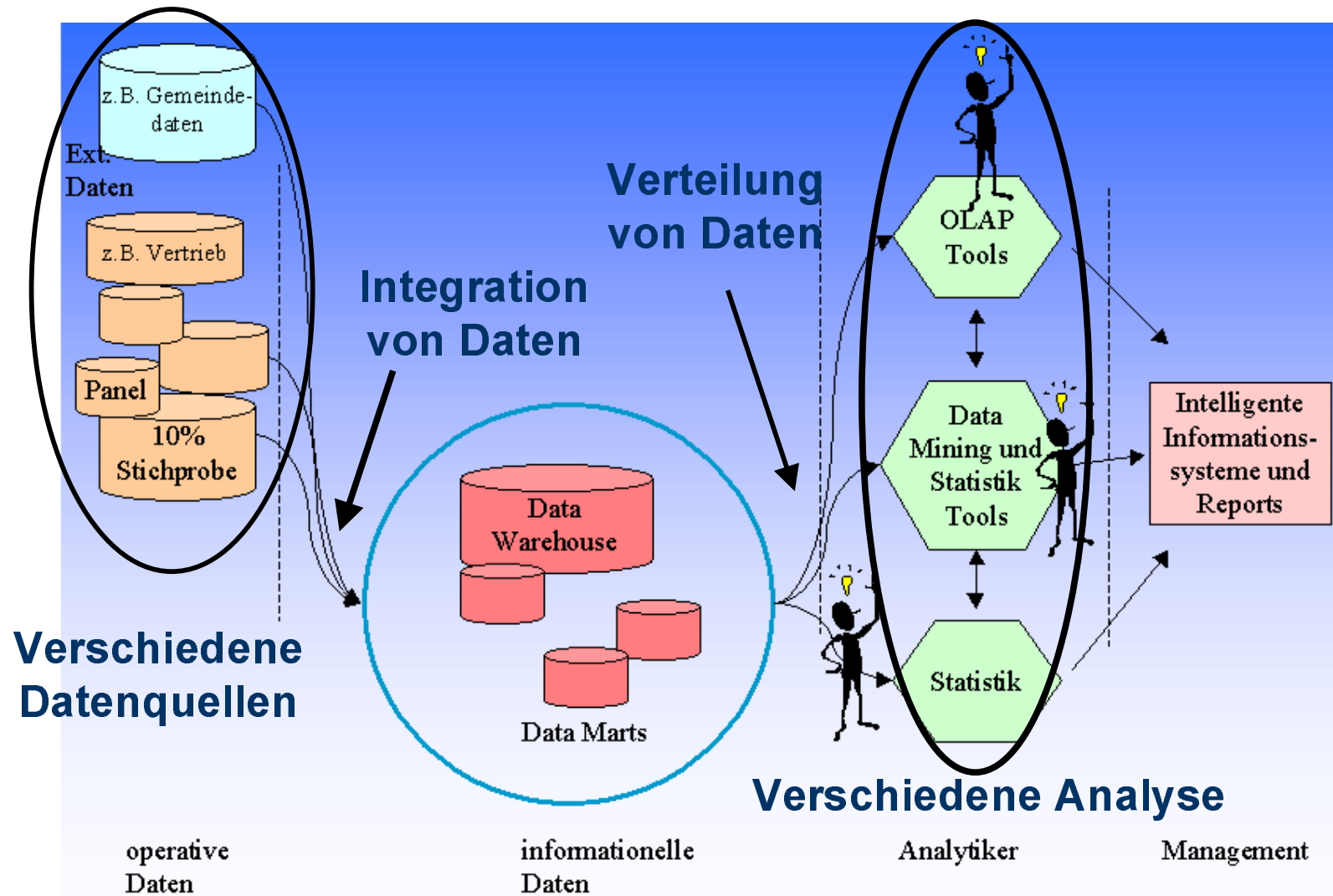
### Ontology development (Übung 3.)

=>

### Ontology Requirements Specification Document (ORSD) - Competency Questions



## 4. Data warehouse - environment



## 4b. Analysefragestellungen

Wann finden die wenigsten Gespräche statt

Wie viel wird Sonntags telefoniert

Welche Tarifzone sind meistens benutzt

Was für eine Verbindung ist zwischen Einkommen von Anrufer  
und Zeitpunkt des Anrufs

Welche Alter telefonieren am meisten

## 4b. Analysefragestellungen

Wann finden die wenigsten Gespräche statt

*Zeitpunkt*

*Gebühren*

Wie viel wird Sonntags telefoniert

*Dauer*

*Zeitpunkt*

Welche Tarifzone sind meistens benutzt

*Tarifzone*

**Verbindung**

Was für eine Verbindung ist zwischen Einkommen von Anrufer und Zeitpunkt des Anrufs

*Einkommen*

*Anrufer*

Welche Alter telefonieren am meisten

*Alter*

**Kunde**

## 4b. Subjects

Dimensionen: Verbindung, Kunde

- Verbindungen
  - Art des Dienstes (ISDN, Fax, Voice, Internet ...)
  - Kunde(Anrufer)
  - Kunde(Angerufener)
  - Ortsnetz
  - Zielortsnetz
  - TarifZone (Ort, Nah, Fern, Euro, Welt)
  - Zeitpunkt (Wochentag, Uhrzeit)
  - Dauer
  - Gebühren

## 4b. Subjects

- Kunde
  - Geschlecht
  - Familienstand (ledig, verheiratet)
  - Art des Anschluss (Analog, ISDN)
  - Alter (<18, 19-27, 27-39, 40-65, >65)
  - Einkommen (<10.000, 10.000-25.000, 25.000-100.000, >=100.000)
  - Beruf (Selbstständig, Angestellter, Beamter)
  - Wohnort (Wohngebiet, Ortsteil, Ort, Region, Bundes-Land)
  - Haushaltsgröße