AIFB

Intelligente Techniken für das Wissensmanagement

Steffen Staab

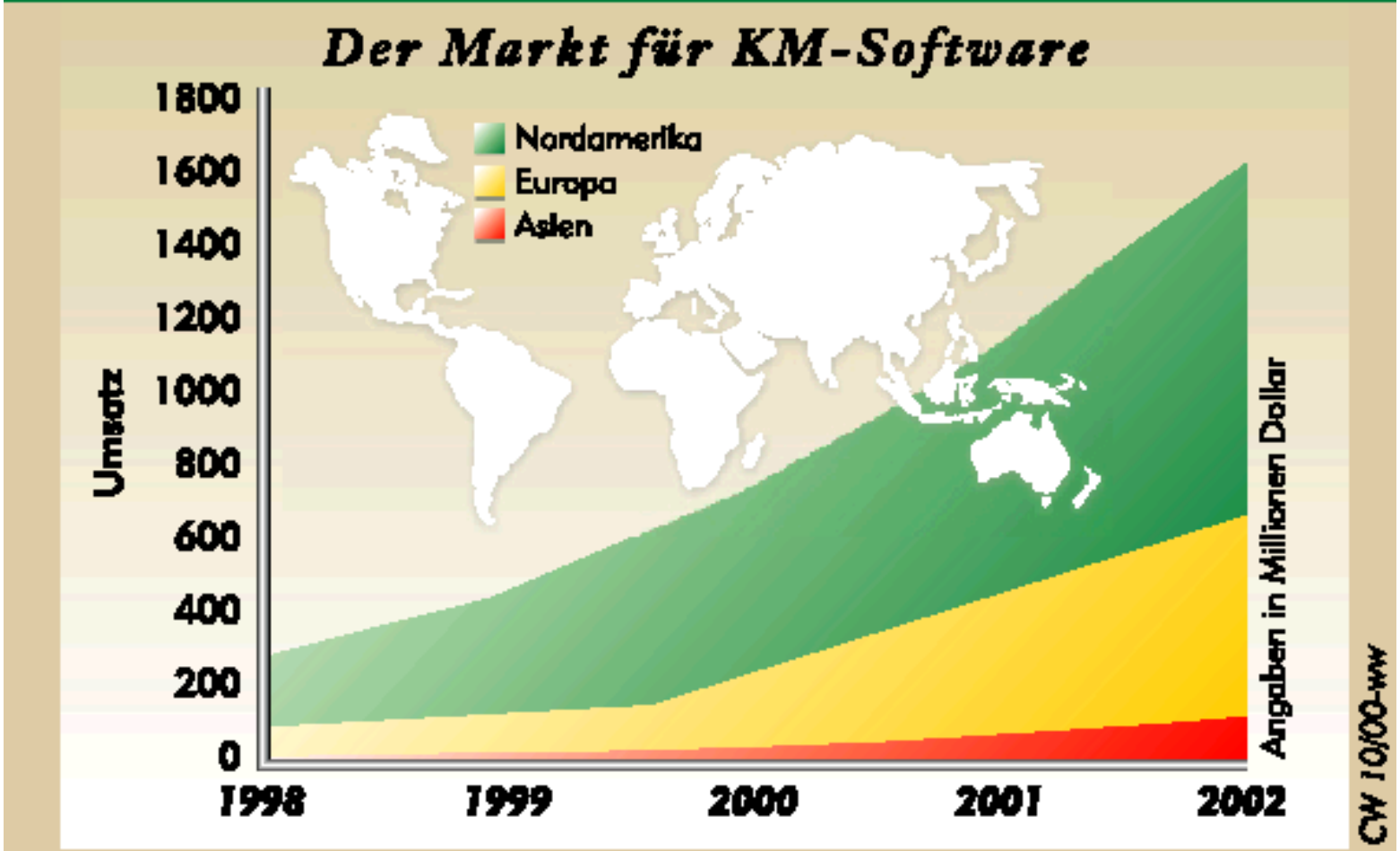
http://www.aifb.uni-karlsruhe.de/~sst
http://www.ontoprise.de
mailto:sst@aifb.uni-karlsruhe.de,staab@ontoprise.de

Institut AIFB, Univ. Karlsruhe (TH) & Ontoprise GmbH, Karlsruhe

Tutorium auf der WM-2001
14.3.2001 in Baden-Baden

AIFB

STETIGES WACHSTUM



DER BEDARF AN PROGRAMMEN zur Wissensverwaltung im Unternehmen steigt auch in den nächsten Jahren weiter an.

Quelle: IT-Research

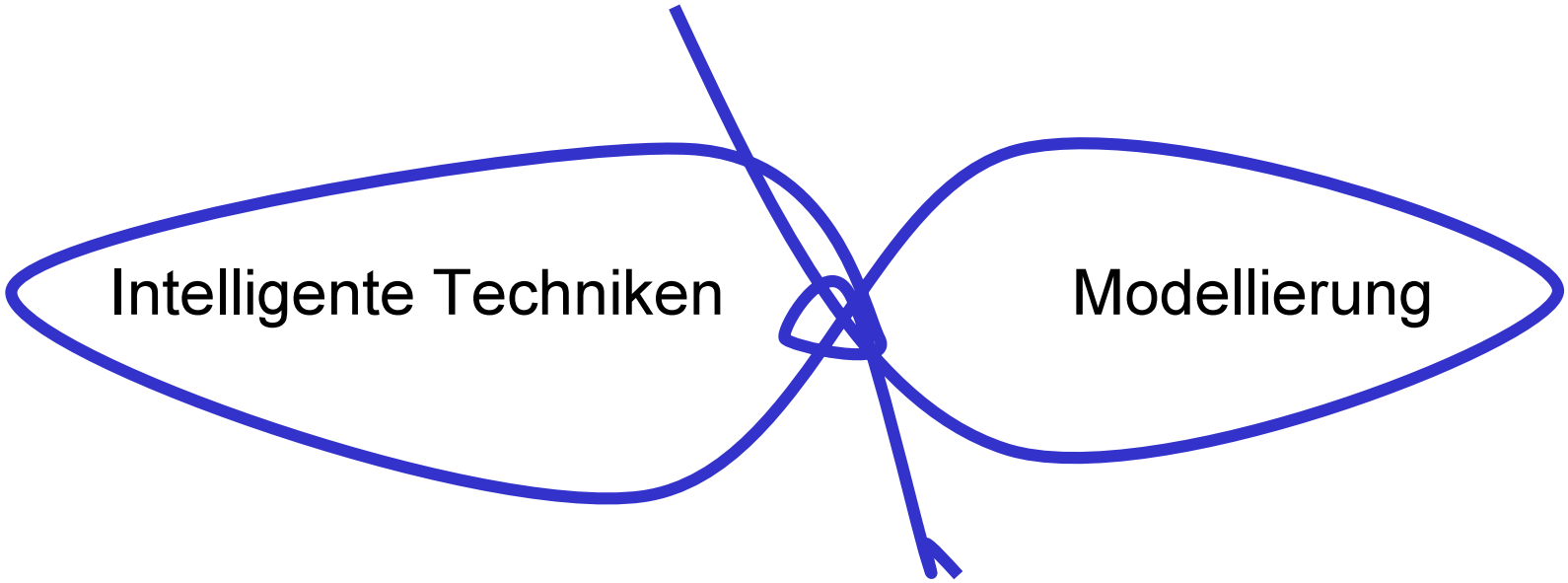

Intelligente Techniken
Modellierung

Table of Contents

AIFB

1.1 About the Lecturer

Steffen Staab, Dr. rer. nat.
http://www.aifb.uni-karlsruhe.de/~sst

Studium der Informatik in Erlangen, Philadelphia, Freiburg

- Consultant bei Fraunhofer IAO
- Mitgründer von Ontoprise GmbH
- Wissenschaftler bei British Telecom
- Wissenschaftlicher Assistent am AIFB, Universität Karlsruhe (TH)

Forschungsinteressen:

- Wissensmanagement
- Informationsextraktion
- Ontologien
- Knowledge Discovery,
- Web-Anwendungen
- Semantic Web

AIFB

1.2 Introduction to Knowledge Management

"Today's most technologically advanced economies are truly knowledge based. And as they generate new wealth from their innovations, they are creating millions of knowledge-related jobs in an array of disciplines that have emerged overnight: knowledge engineers, knowledge managers, knowledge coordinators."

[World Development Report 98/99]

AIFB

Why interest in Knowledge?

- Increased speed of Knowledge Creation (Research)
- shorter development cycles with more intelligent, sophisticated products
- increasingly complex regulations (environmental standards, suit risks...)
- Globalization of Economy (more competitors, complex markets with different cultures)
- Knowledge and Information is an economic asset itself
- Increasing fraction of Knowledge Intensive Work

Knowledge Work

- Knowledge Work is based on the Creation and Application of Knowledge - usually no fixed workflow and lot of Exceptions:
  - Research
  - Product Development
  - Medicine
  - Law
  - Diagnosis and Maintenance of complex Machine
- Shorter development cycles with more intelligent, sophisticated products

- Single activities in conventional processes are knowledge work
  - Loan approvement
  - Risk assessment of Insurance Policies

AIFB

Knowledge needs to be maintained

- Employees are the primary knowledge sources (but change employer)
- Knowledge is a power factor and is usually not shared
- Lessons learned are not recorded and reviewed (costly!)
- Knowledge Worker are using 50-80% of their work time for information search

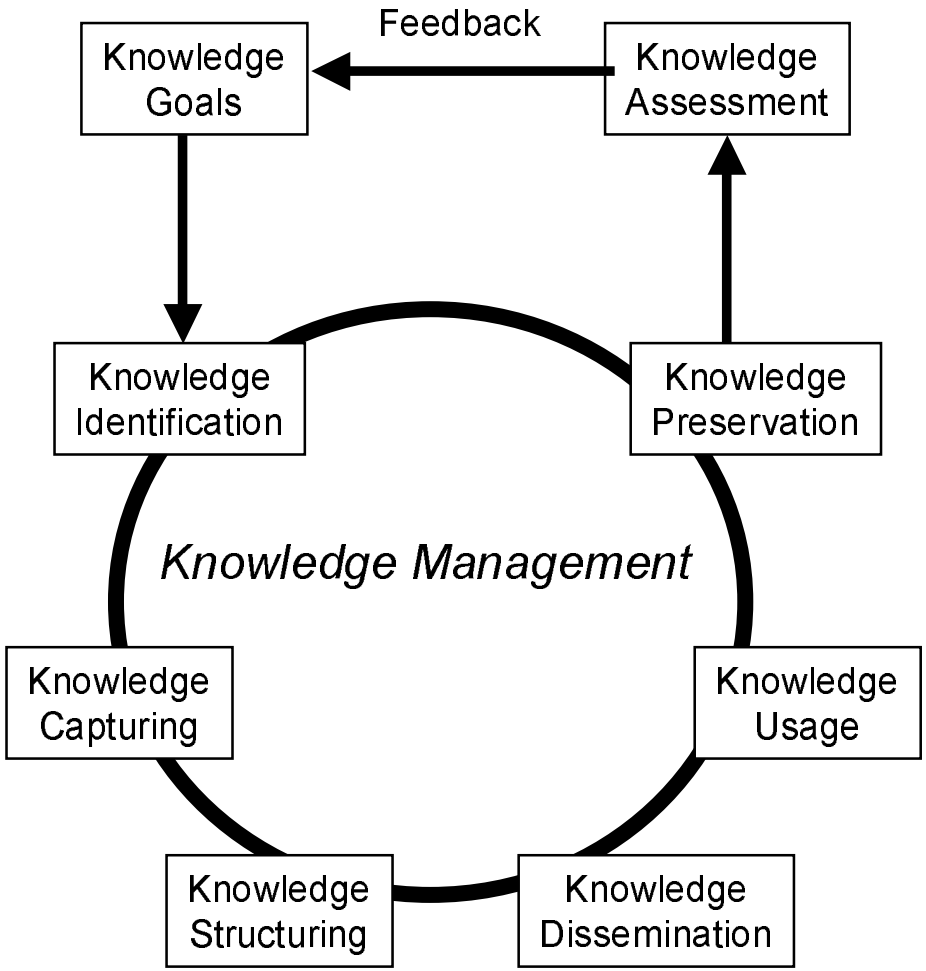
AIFB

Goals of Knowledge Management

- Effective Utilization of the available Knowledge
- Knowledge Sharing and Reuse
- Accessibility of Knowledge
- Embedding of Knowledge in the Work Context

Knowledge Management Process

[Probst et al. 1999]



Knowedge Goals: Determine Goals for KM Activities

Knowledge Identification: Create Overview about available Knowledge

Knowledge Structuring: Structuring and Integration of Knowledge

Knowledge Capturing: Acquisition of Knowledge

Knowledge Dissemination: goal oriented dissemination of Knowledge

Knowledge Usage: productive Usage of Knowledge for the Company

Knowledge Preservation: Storage and Maintence of Knowledge

Knowledge Assessment: Assesment of current Knowledge and compliance with goals

AIFB

Viewpoints on KM

- Management
  - Define Knowledge Goal
  - Asses Knowledge
  - Hire Employee
  - Change corporate culture
  - Employee Skill Management…
- Information Technology
  - Organizational Memory Information Systems
  - Intranets
  - Information Retrieval
  - Data Warehouse / Data Mining
  - Information Filtering/Agents
  - …

AIFB

Several Views on Knowledge (I)

- Tacit Knowledge
  - Personal, Created by Experience
  - Intuition, Mental Models, not documented
- Explicit Knowledge
  - Documented
  - Reconstructable

Several Views on Knowledge (II)

human readable

machine readable

informal:

- E.g. ASCII-Text, Word-Document, Powerpoint-Presentation

semi-structured:

- informal representation is enrichted with attributes
- Examples: XML, SGML, HTML, email

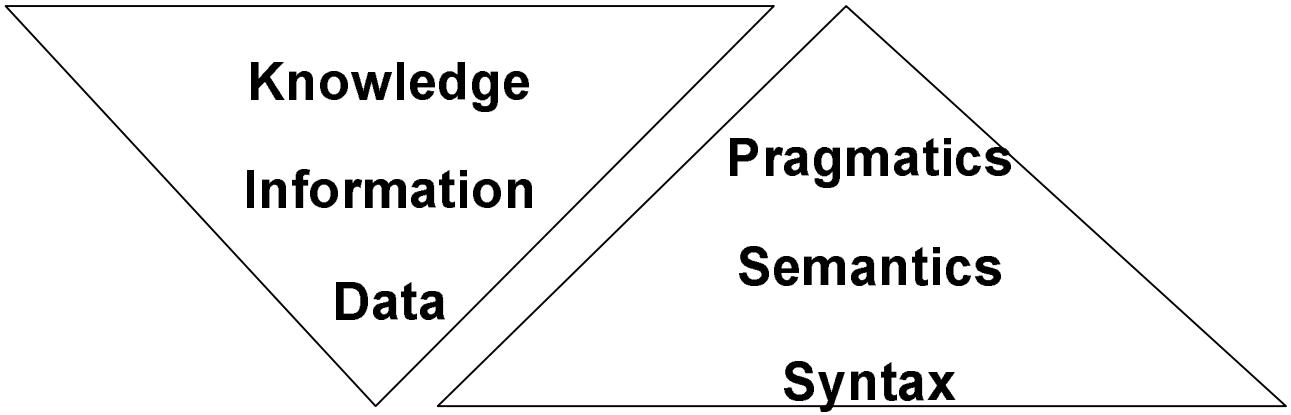
structured:

- Structured according to fixed set of attributes
- Example: (Text-)Database

formal:

- Examples: Frames, Production Rules, relational Database, Programm Code

Several Views on Knowledge (III)



Knowledge Management as
Information Management in Organizational Context

AIFB

Organizational Memory

- Knowledge Dissemination, Knowledge Preservation
- Combination of several Techniques:
- Usual Motivation:
  - Document writing is (relativly) easy.
  - Knowlege often available in documents (manual, internal memos etc.)
- Principle:
  - instead of formalization of Knowledge, administer Organizational Memories
    - **Documents containing Knowledge in a human readable and understandable form.**
    - **Knowlege Maps (Skills etc.)**

    e.g. (Text-)Databases, Document-Mgmt Systems, Intranet (Hypermedia)
  - Strucuring of Knowledge simplifies search und usage, e.g. by
    - **Classfication of Documents/ Indexing, Case Based Reasoning**
- Problem: Link to applications

From Objectives to Intelligent Techniques

Organization & People

- Determine objectives
- Identify Knowledge
- Capture Knowledge
- Structure Knowledge
- Use Knowledge
- Preserve Knowledge
- Disseminate Knowl.
- Assess Knowledge
- Preserve Knowledge

General Data Processing

- Capture Data
- Maintain Data
- Process Data
- Integrate Data
- Search for Data
- Use Data

AI Techniques

- Information Retrieval & Extraction
- Visualization Techniques
- Case-based Reasoning
- Ontology-based KM
- Meta data-based KM
- Knowledge Discovery
- (Knowledge Acquisition)

Everything is Knowledge Management??

AIFB

„My car is my favorite KM tool, because I always drive to my colleagues and my clients in order to exchange knowledge!“

AIFB

KM = Content Management + Volltextsuche?

KNOWLEDGE-MANAGEMENT-STUDIE

Anbieter	Produkt	Funktionale Bewertung (max. 19 Punkte)	Gesamt-ergebnis (max. 54 Punkte)	Web-Adresse
Arcplan	Insight/Dynasight 2.4	13	33	www.arcplan.de
Autonomy	KM-Suite 1.9	19	37	www.autonomy.com
Blue Angel	Metastar 2.1.3	16	30	www.blueangeltech.com
Dataware	Dataware II KM-Suite 2.0.4	19	46	www.dataware.com
Dr. Materna	Knowledge-Bridge/ Knowledge-Architect	13	43	www.materna.de
Excalibur	Retrievalware 6.7	16	35	www.excalib.com
Gauss Interprise	VIP 3.0	15	42	www.gauss.de
Grapevine	Grapevine for Compass Server 3.01B	15	34	www.grapevine.com
IDS Scheer	Enterprise Knowledge Portal (Betaversion)	12	37	www.ids-scheer.de
Opentext	Livelink 8.1.3	19	48	www.opentext.com
Pironet	Pirobase 4	16	41	www.pironet.com
SAP	Knowledge Warehouse	16	49	www.sap.de
Schema	Schematext	16	36	www.schema.de
USU	Knowledgeminer 2.2	16	35	www.usu.de
Verity	Knowledge Product Suite 3.6	15	39	www.verity.com
ZAP	Ucone 1.0.1	13	23	www.zapnet.de

Quelle: IT-Research (www.it-research.net)

A shot at categorization (subjective & highly incomplete!!!)

- *Information Retrieval*: Verity™, Connex™, Excalibur™, Eurospider™, Google™, Fulcrum™
- *Collaborative Filtering*: Grapevine™
- Intranet Portal: Intraspect™ , Open Text™ , Autonomy™, Ontoprise™
- *Groupware*: Lotus Notes™, MS Exchange™
- *Document Management*: PCDOCS™, InQuery™, Filenet™, Documentum™
- *Text Summarization*: Prosum
- *Database solutions*: Wincite™, Dataware™, Agentware™
- *Experience Factories*: at A.D.Little™, at Xerox™
- *Skill Management*: Loga HRMS (P&I)™, proprietary solutions
- *Semantic Nets-based*: USU™, Knowledge Park™
- *Visualization*: Inxight™, AIdministrator™
- *Knowledge Discovery*: Clementine™, IBM™, SAS™

AIFB

Techniques for Knowledge Management

3 Techniques for Knowledge Management

AIFB

Organization & People

- Determine objectives
- Identify Knowledge
- Capture Knowledge
- Structure Knowledge
- Use Knowledge
- Preserve Knowledge
- Disseminate Knowl.
- Assess Knowledge
- Preserve Knowledge

General Data Processing

- Capture Data
- Maintain Data
- Process Data
- Integrate Data
- Search for Data
- Use Data

Intelligent Techniques

- Information Retrieval & Extraction
- Visualization Techniques
- Case-based Reasoning
- Ontology-based KM
- Meta data-based KM
- Knowledge Discovery
- (Knowledge Acquisition)

3.1 Information Retrieval & Extraction

AIFB

Organization & People

- Determine objectives
- Identify Knowledge
- Capture Knowledge
- Structure Knowledge
- Use Knowledge
- Preserve Knowledge
- Disseminate Knowl.
- Assess Knowledge
- Preserve Knowledge

General Data Processing

- Capture Data
- Maintain Data
- Process Data
- Integrate Data
- Search for Data
- Use Data

Intelligent Technique

- Information Retrieval & Extraction

Main Objectives

- Find Knowledge
  - In unstructured/ semi-structured text documents
- Capture Knowledge
  - Facts
  - Knowledge Structures (Concepts, Relations)
- Use Knowledge

Information Retrieval

- Find in repository
  - Keyword search
  - Keyword search with Thesaurus
  - Find similar documents / documents indexed by some label
    - **Vector Space Model (Vector of TF/IDF weights per doc)**
    - **Probabilistic Model (Binary Vector with Bayes)**
    - **Latent Semantic Analysis/Indexing (Deerwester et al.)**
  - Topic Spotting (in particular for Audio! Cf. Wiener et al.)
  - Summarization

(Sparck-Jones & Willett, 1997)

Thesaurus - Objectives

- Map for a given field of knowledge
- Standard vocabulary for retrieval
- Unique terms for reference
- Locate new concepts in a scheme of relationships
- Broaden / narrow search through hierarchy

- Standardization of term usage

- Examples:
  - Roget's thesaurus,
  - WordNet / GermaNet / EuroWordnet
  - TEST (Thesaurus of Engineering and Scientific Terms), …

AIFB

Thesaurus (Foskett 1980) - „A treasury“ of words

- Improve effectiveness of communication between people
- Constantly developing / Permanent revisioning

- Contents
  - Guidelines for form of terms (e.g. singular/plural)
  - Guidelines for relationships (BT, NT, RT)

- Administration:
  - Check consistency (dangling links?)
  - Maintain statistics (keep it as small as possible!!)
  - Acceptability of terms
  - Maintain records of term history

AIFB

IR in Context - Proactive Delivery

- Deliver similar web documents proactively
  - build index on metainformation of documents
  - metainformation: automatically extracted keywords, summary, document title, URL and date and time of access
  - Retrieval Modes:
    - **keyword**
    - **what's new?**
    - **comparison with user profile**
    - **comparison with group**
  - Retrieval Technique:
    - **Keyword**
    - **Possibilities: Vector Space Model, ...**

(*Jasper;* Davies et al. 1995)

Capture Knowledge Contents

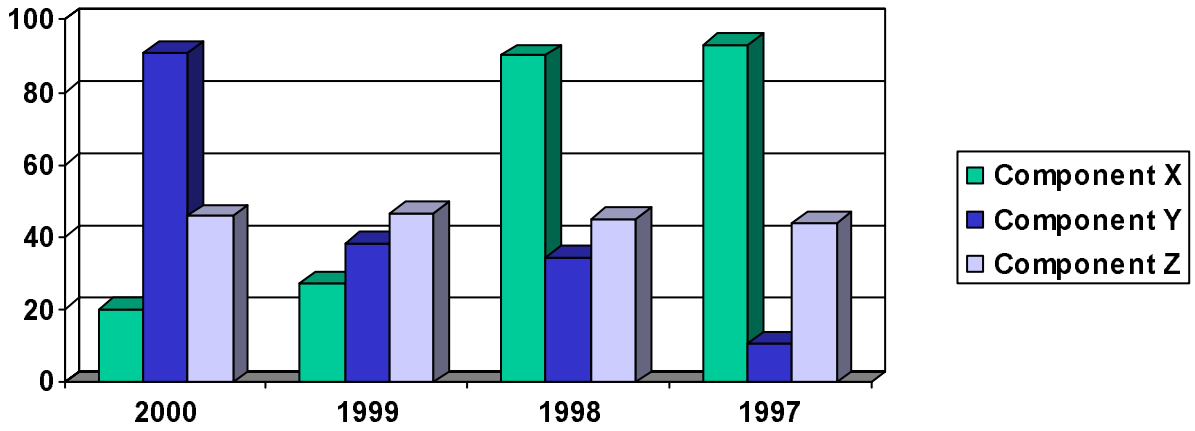
- Fact Extraction
  - Annotation/Metadata by hand
    - **Labeling with Keywords**
    - **Semantic Annotation** (Buitelaar et al. 2000; Decker et al. 1999)
      → Section on Metadata
  - Automated
    - **Automatic labeling of documents**
      - clustering
      - Latent Semantic indexing (problems with efficiency)
    - **Scisor (Rau 1988); ParseTalk (Hahn et al. 1999; Staab 1999)**
  - Semi-automated
    - **Information extraction proposing semantics** (Erdmann et al.)

Information Extraction

Case: Dow Chemicals

- Patent Analysis based on Information Extraction
- Extract degree expressions “10 tons of sulfur”
- Use OLAP/knowledge discovery (section to follow) to determine trends in use of chemicals

Qualitative Diagram:



AIFB

Capture Knowledge Structures

- Learning of hierarchical structures
  - Relevant Terms:
    - **Justeson & Katz**
    - **Terminological Engineering/Learning from Text**
  - Relevant Relations
    - **Faure et al.**
    - **Mädche & Staab**

Importance of Lessons Learned

„Human experts are not systems of rules, they are libraries of experiences.“

Riesbeck & Schank 1989

Lessons Learned

Organization & People

- Determine objectives
- Identify Knowledge
- Capture Knowledge
- Structure Knowledge
- Use Knowledge
- Preserve Knowledge
- Disseminate Knowl.
- Assess Knowledge
- Preserve Knowledge

General Data Processing

- Capture Data
- Maintain Data
- Process Data
- Integrate Data
- Search for Data
- Use Data

Intelligent Technique

- Case-based Reasoning

Motivation & Process

Motivation:

- Knowledge Documents (e.g. Project Report)

General CBR Process:

- RETRIEVE cases similar to current problem
- REUSE retrieved cases
- REVISION proposed solution (improve/correct)
- RETAIN new knowledge

Knowledge Containers

- Case base (collection)
- Vocabulary used to describe cases
- Similarity measure
- Adaption model for revision

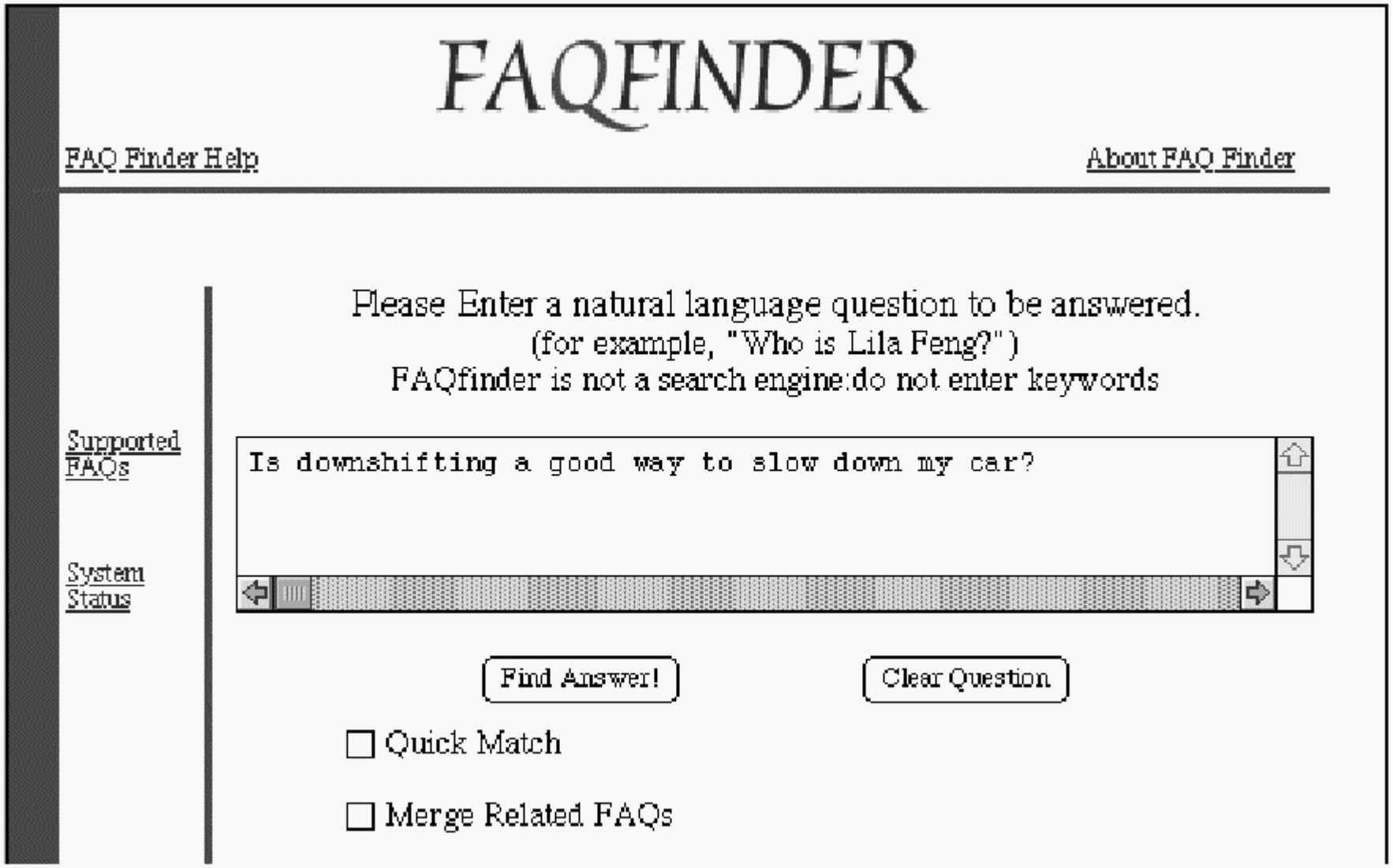
(Kolodner 1993),
(Lenz et al. 1998)

Textual Case-based Reasoning

"Know How Documents"

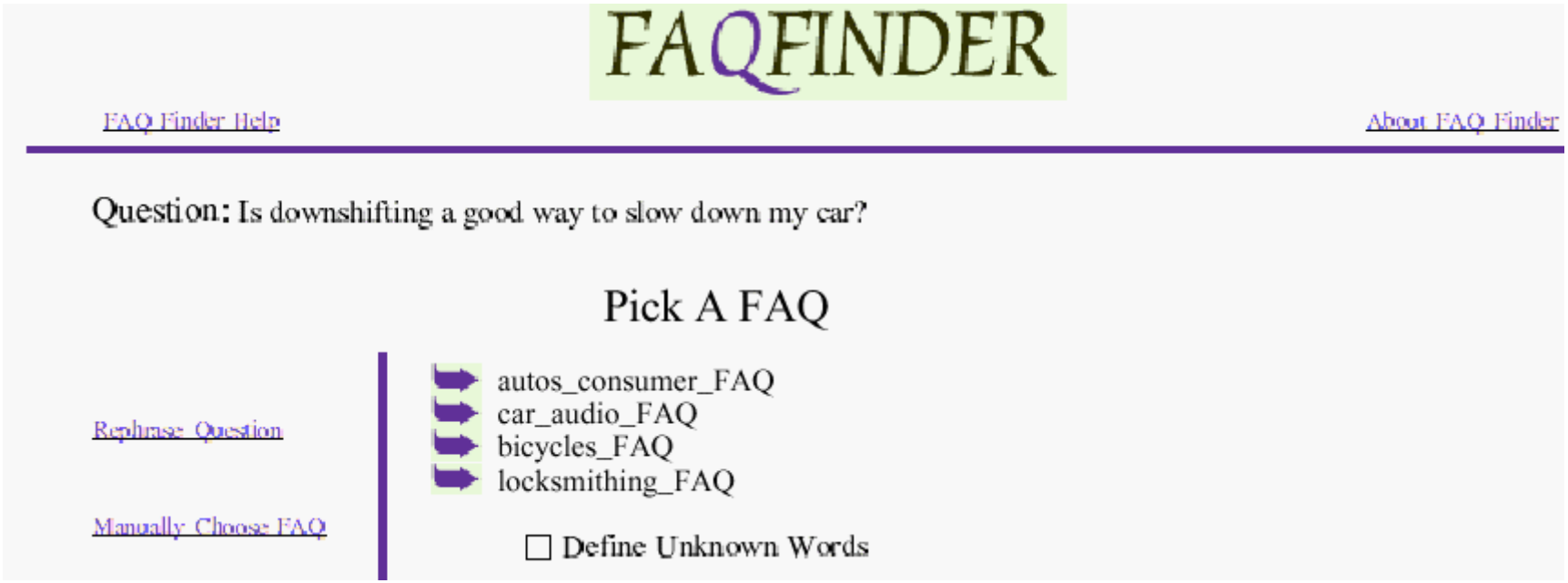
- FAQ Finder: Burke et al. 1997
- Automatic hotline for Siemens technicians (Lenz 1998)
  (human hotline as backup)
- In-house configuration management of LHS AG (Lenz 1998)
- Aircraft maintenance at British Airways (Magaldi 1999)

3.2 Case-based Reasoning



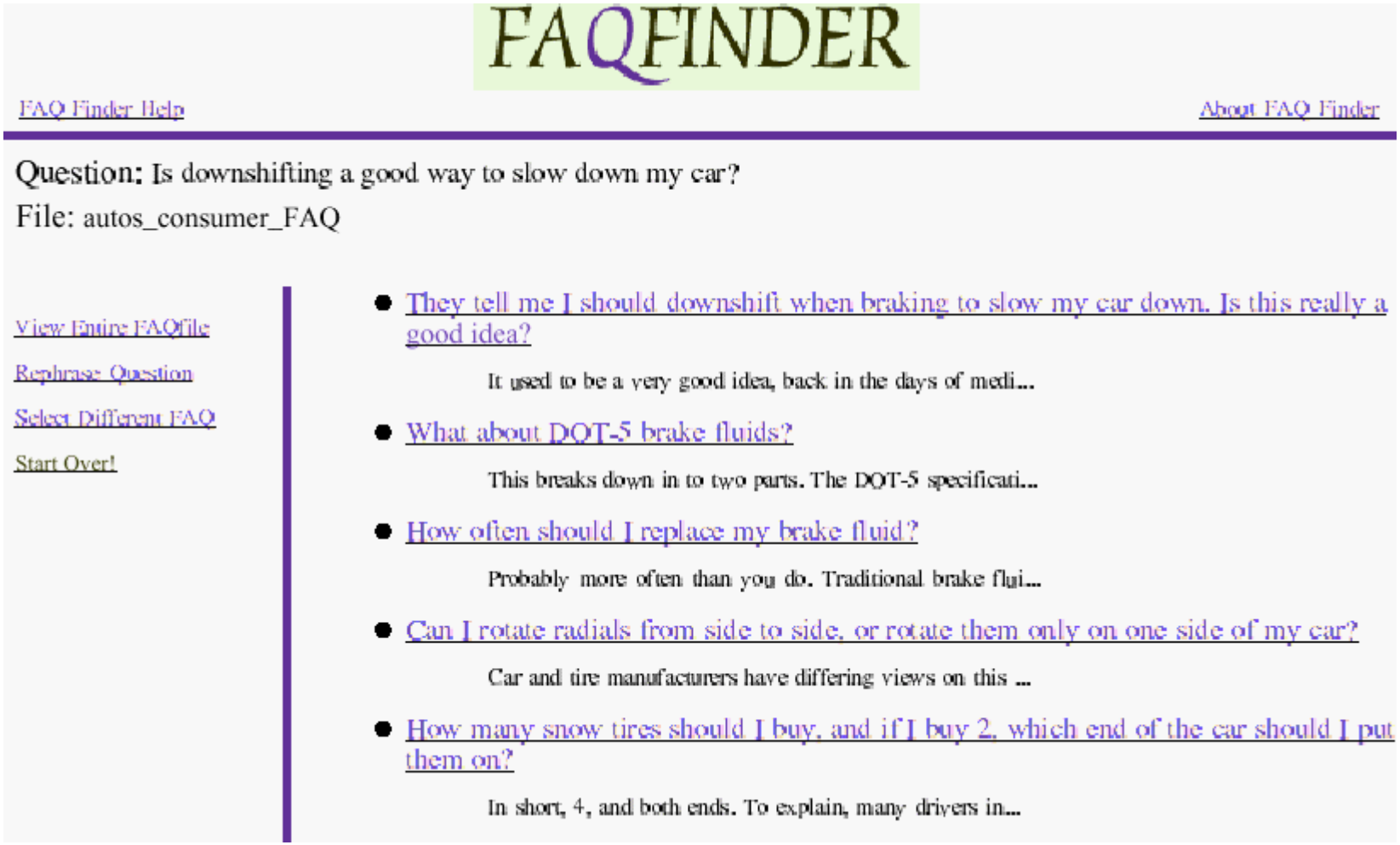
(© Burke et al. 1997)

3.2 Case-based Reasoning



(© Burke et al. 1997)

3.2 Case-based Reasoning



(© Burke et al. 1997)

3.2 Case-based Reasoning

FAQFINDER

FAQ Finder Help | About FAQ Finder

Previous Question | Next Question | View Entire FAQfile | Rephrase Question | Start Over!

File: autos_consumer_FAQ

They tell me I should downshift when braking to slow my car down. Is this really a good idea?

It used to be a very good idea, back in the days of mediocre, fade prone drum brakes. In modern disc brake equipped cars, use of downshifting to slow the car is not really necessary, except in cases of long, steep downhill runs. Otherwise, modern disc brakes are more than adequate to stop a passenger car in all circumstances, and they are much cheaper to repair than clutch linings.

On the other hand, many standard driver's license tests in the USA still specify that the driver being tested downshift under braking; I suggest that before taking a US driver's test, you either 1) learn to do this smoothly (which takes some time and practice) or 2) borrow a car with an automatic to take the test.

(© Burke et al. 1997)

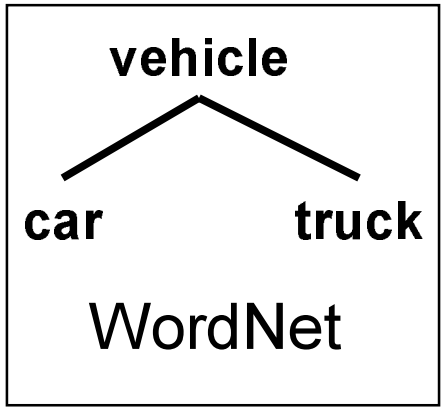
FAQ Finder Techniques

Matching User Query with QA pairs in FAQ file:

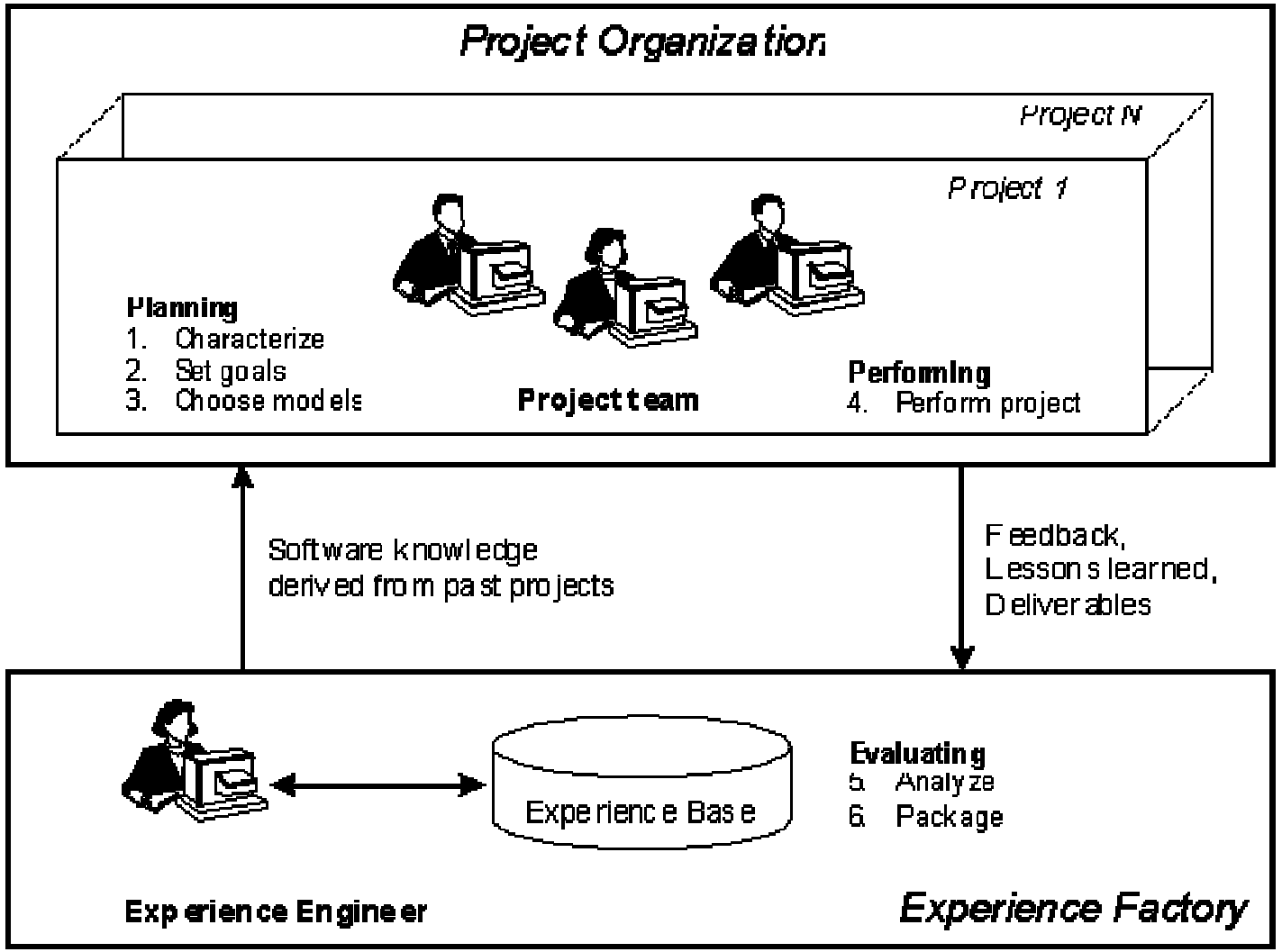
- Statistics model
  - Vector space model: term vectors with *tfidf* values
  - tfidf = n * log(M/m)
    (term frequency inverse document frequency)

- Semantic model
  - Word by word comparison of
    user question and FAQ questions
  - using marker passing in WordNet
  - punish for words that are not matched

- No syntactic model

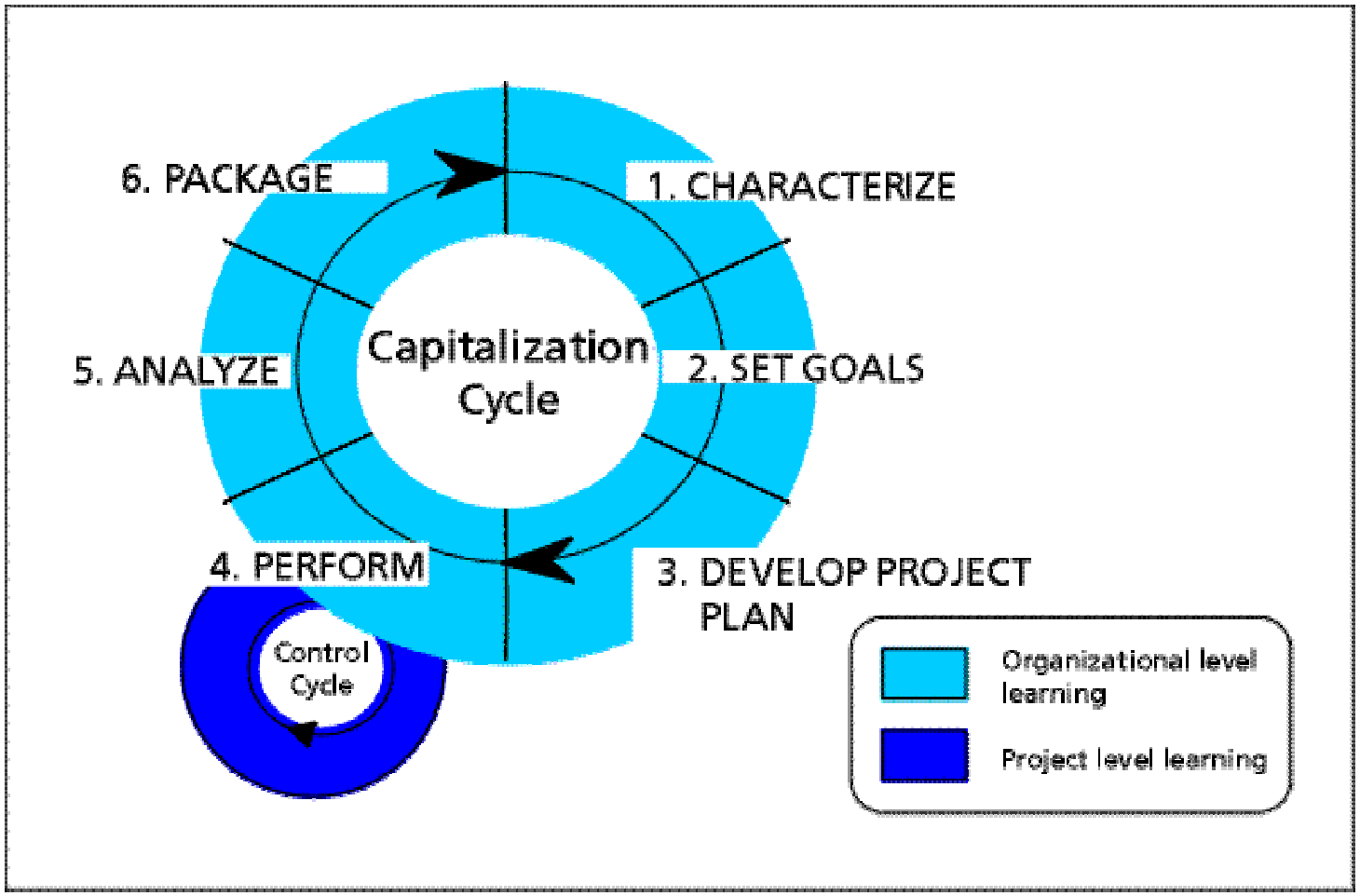


Software Experience Factory (Basili et al. 1992)



(© Althoff et al. 1998)

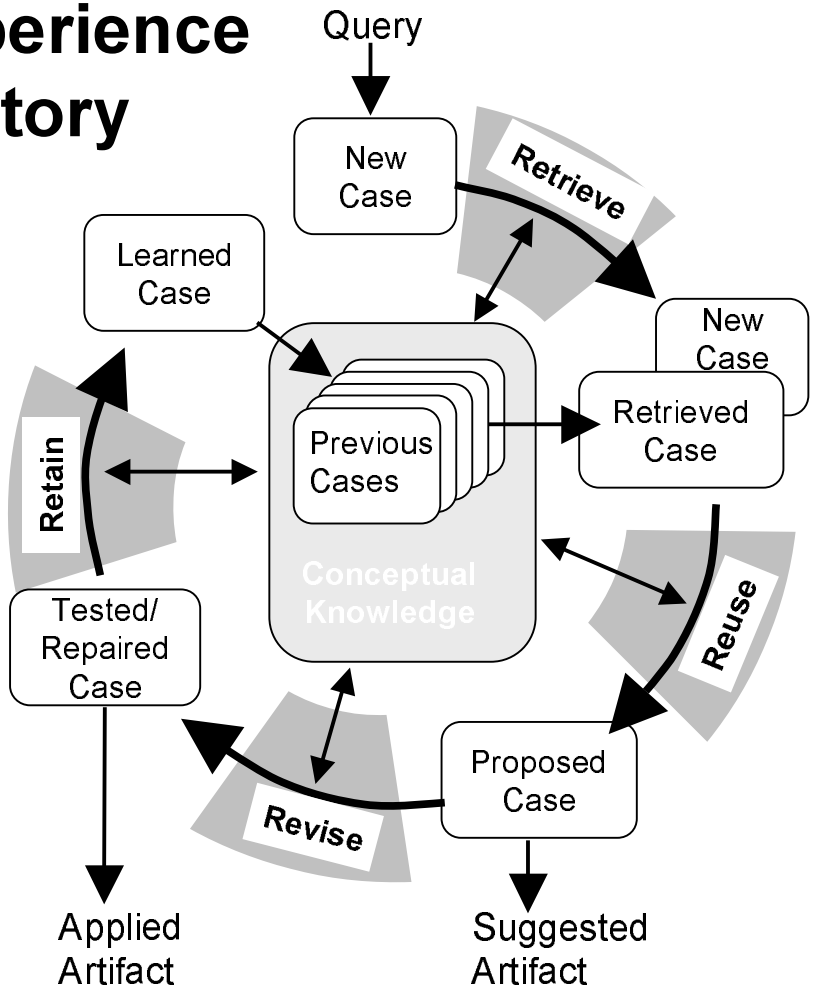
Software Experience Factory



(© Althoff et al. 1998)

5 Case-based Reasoning

Experience Factory



(© Althoff et al. 1998)

Case := Problem (characterization)/solution (artifact) pair

Query: Query at hand defines new case (problem without solution)

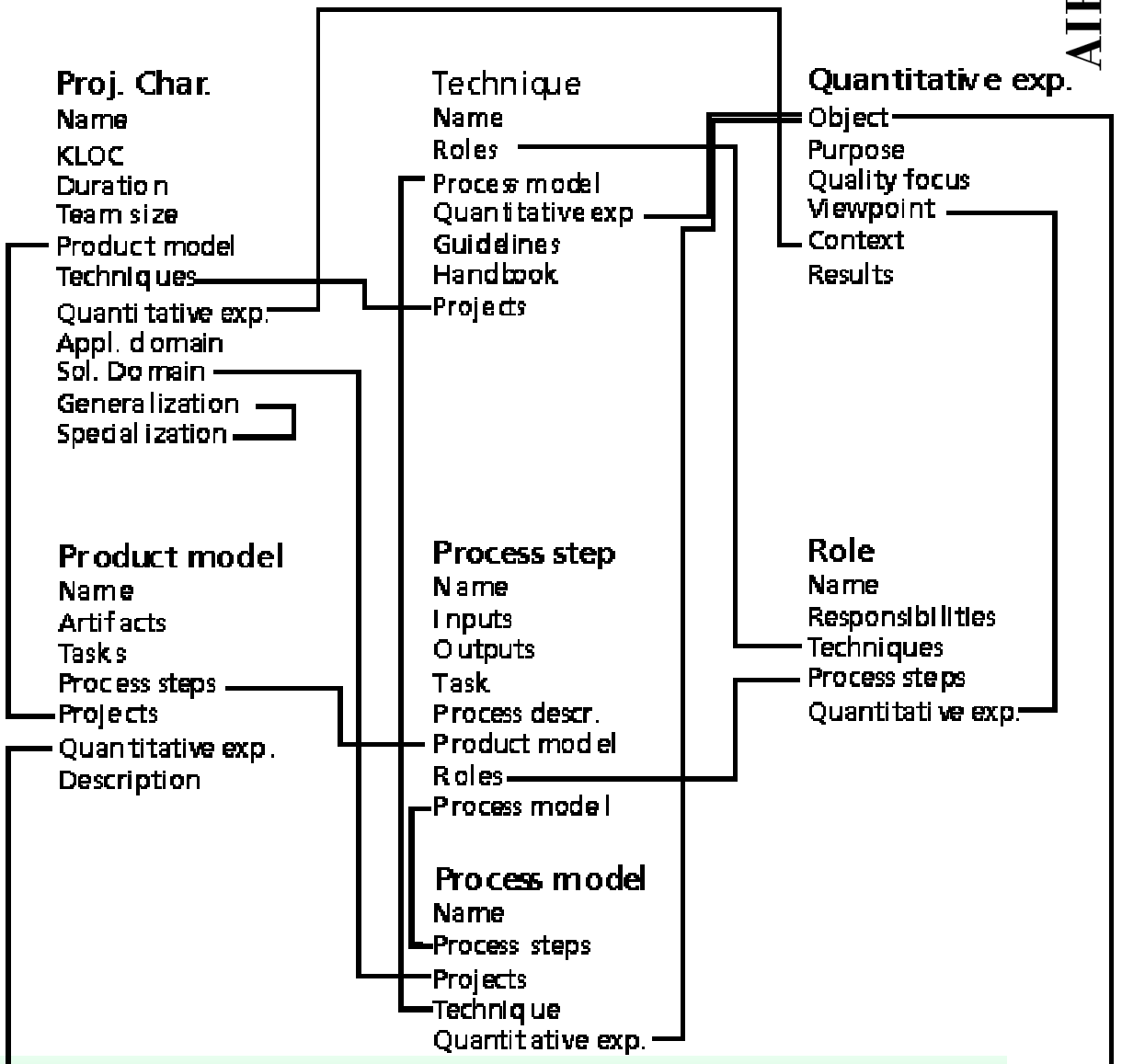
Retrieve: New case is used to find most similar case among the known (previous) cases

Reuse: New and retrieved case are combined to a proposed case including the suggested artifact

Revise: Suggested artifact is applied and evaluated

Retain: Useful experiences from applying the artifact are retained by adapting the case base and the conceptual knowledge

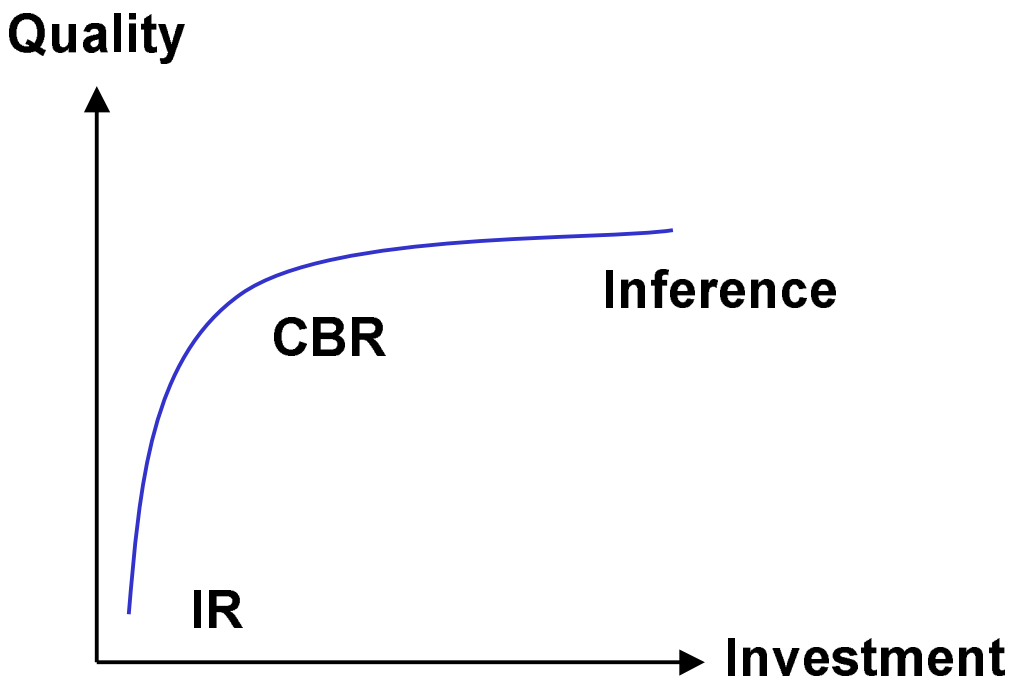
Software Experience Factory



(© Althoff et al. 1998)

Conclusion on CBR

- Domain modeling important, but also expensive
- Commercially successful
- Claim



Aggregating Knowledge out of Data

- Creating knowledge out of data: KDD, Machine Learning *per se* (Tutorial by Fayyad & Simoudis 97)

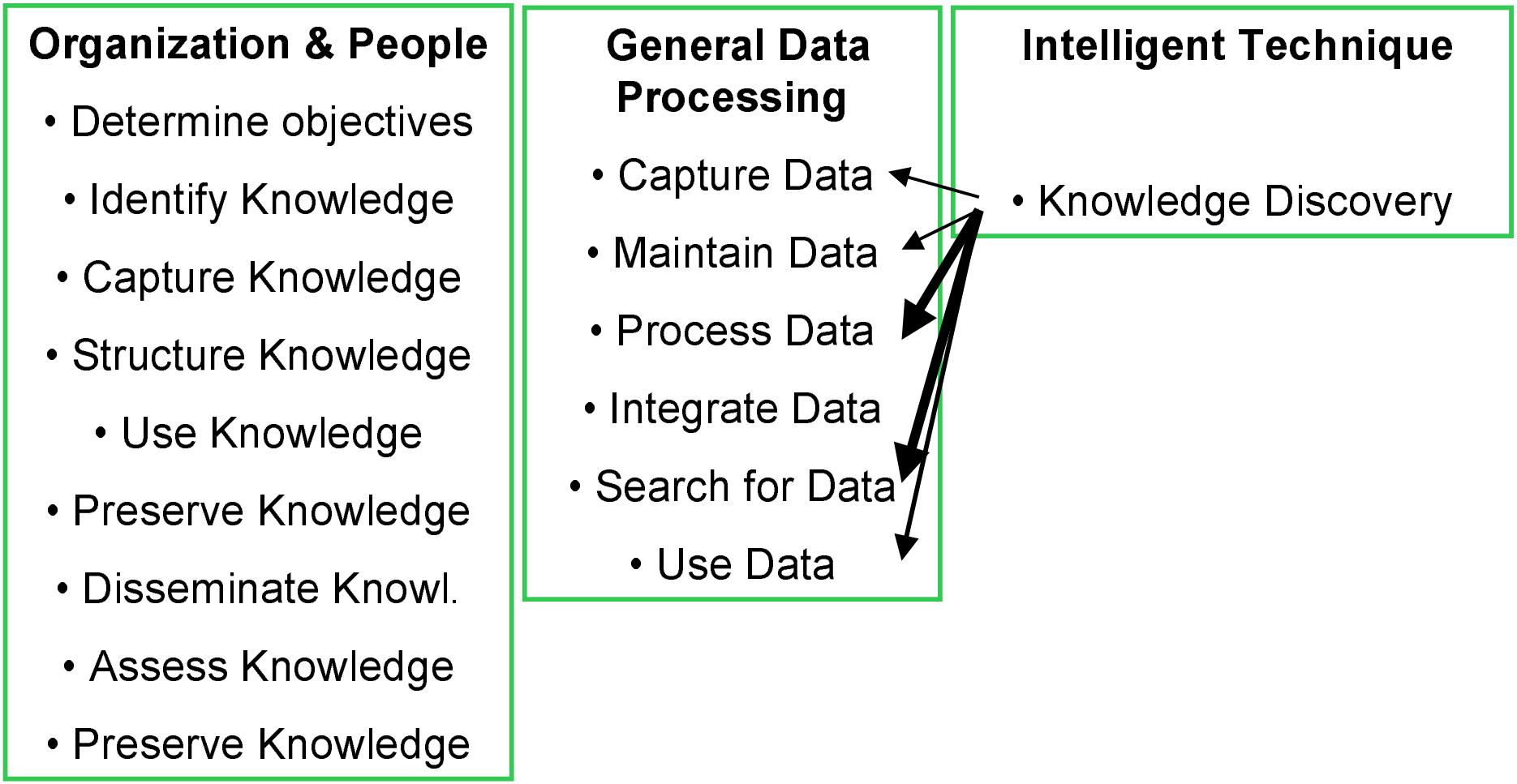
- Creating knowledge out of knowledge
  - Learn T-Box from A-Box: Kietz & Morik 94
  - Learn terminology from texts: Staab et al. (eds.) 2000

- Knowledge discovery/Data analysis as a **collaborative process**
  - Collaboration: Ackermann & Mandel 99; Staudt et al. 98
  - CRISP-DM Process Model: Chapman et al. 99

- Directing attention / User interface issues:
  - Collaborative filtering (e.g. Goldberg et al.1992, Resnick et al. 94)
  - User adaptation (e.g. Syskill & Webert, Pazzani et al. 96, 97)

3.3 Knowledge Discovery

AIFB



CRISP - KD as a process

Business Understanding	Data Understanding	Data Preparation	Modeling	Evaluation	Deployment
Determine Business Objectives *Background* *Business Objectives* *Business Success Criteria*	**Collect Initial Data** *Initial Data Collection Report*	*Data Set* *Data Set Description*	**Select Modeling Technique** *Modeling Technique* *Modeling Assumptions*	**Evaluate Results** *Assessment of Data Mining Results w.r.t. Business Success Criteria* *Approved Models*	**Plan Deployment** *Deployment Plan*
Assess Situation *Inventory of Resources* *Requirements, Assumptions, and Constraints* *Risks and Contingencies* *Terminology* *Costs and Benefits*	**Describe Data** *Data Description Report*	**Select Data** *Rationale for Inclusion / Exclusion*	**Generate Test Design** *Test Design*	**Review Process** *Review of Process*	**Plan Monitoring and Maintenance** *Monitoring and Maintenance Plan*
Determine Data Mining Goals *Data Mining Goals* *Data Mining Success Criteria*	**Explore Data** *Data Exploration Report*	**Clean Data** *Data Cleaning Report*	**Build Model** *Parameter Settings* *Models* *Model Description*	**Determine Next Steps** *List of Possible Actions* *Decision*	**Produce Final Report** *Final Report* *Final Presentation*
Produce Project Plan *Project Plan* *Initial Assessment of Tools and Techniques*	**Verify Data Quality** *Data Quality Report*	**Construct Data** *Derived Attributes* *Generated Records*	**Assess Model** *Model Assessment* *Revised Parameter Settings*		**Review Project** *Experience Documentation*
		Integrate Data *Merged Data*			
		Format Data *Reformatted Data*			

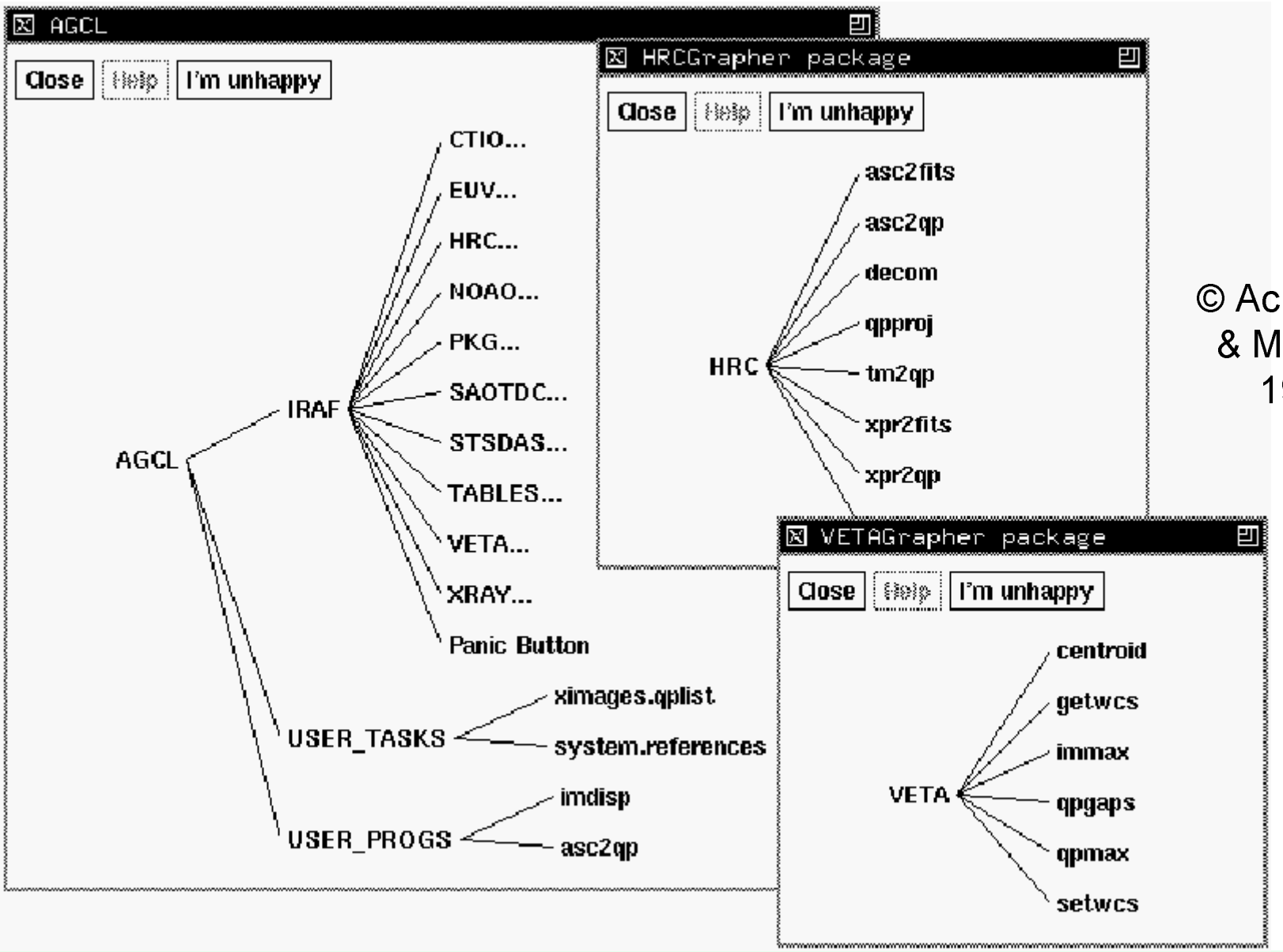
Generic Tasks (bold) and Outputs (italic) of the CRISP-DM Reference Model

Task-Based Organizational Memory

- ASSIST (OM for task of analysing astronomical data)
- Problem description
  - given: lots and lots of raw data in different formats
  - given: plenty of analysis tools building on different formats
- Objective: reuse tools and methodological knowledge
- Solution: OM centered around
  - data, software, other information
  - uniform, flexible, extensible interface

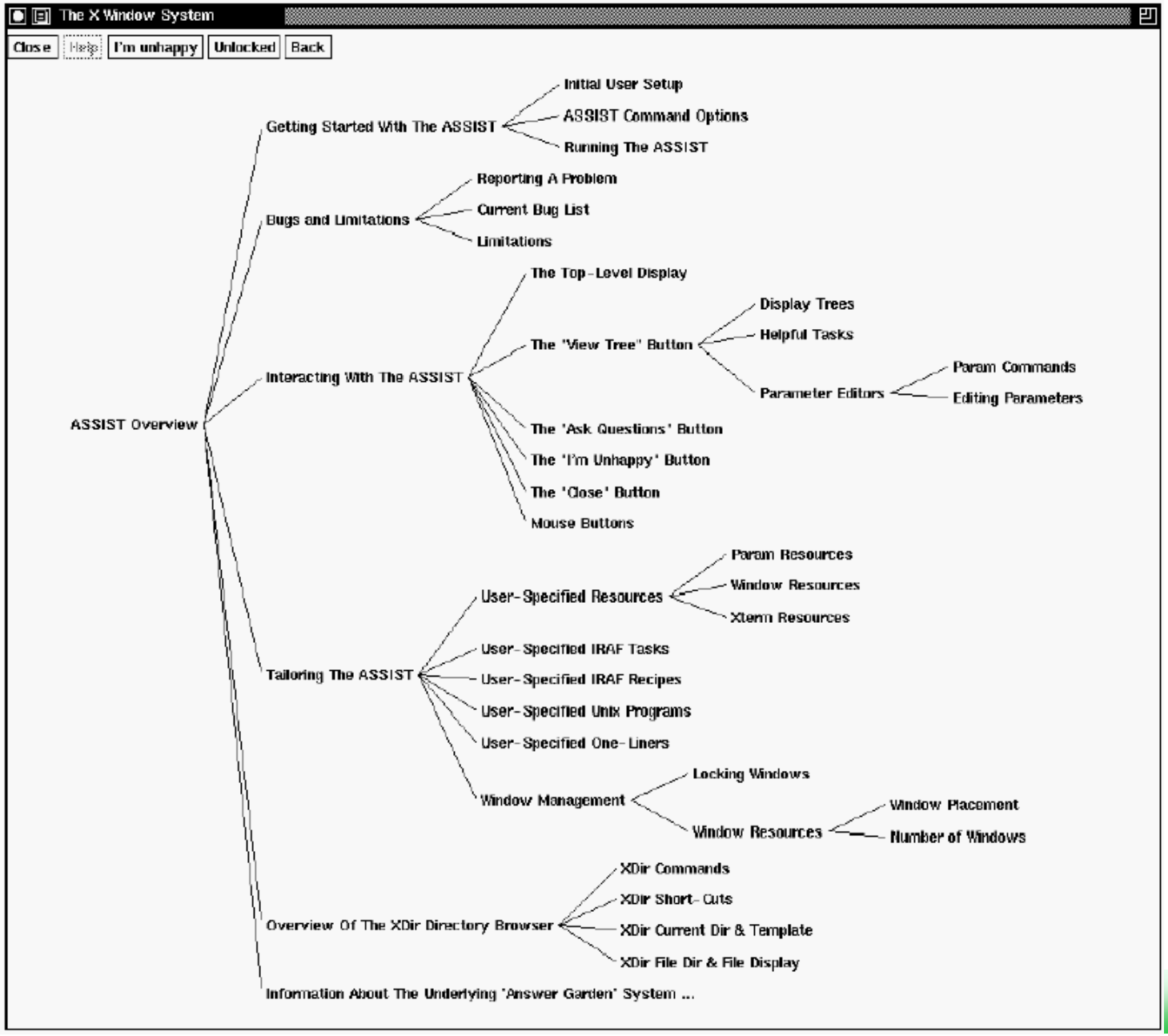
(Ackerman & Mandel 99)

3.3 Knowledge Discovery - Task-based OM



© Ackerman & Mandel, 1999

3.3 Knowledge Discovery - Task-based OM



© Ackerman & Mandel, 1999

Improve Sharing of Knowledge

Collaborative filtering of usenet news

- Let users score articles
- Predict scores of an unread articles depending on how your profile matches with other people who rated this article (no interference with content!)
- Example:

message	Ken	Lee	Meg	Nan
#1	1	4	2	2
#2	5	2	4	4
#3			3	
#4	2	5		5
#5	4	1		1
#6	?	2	5	?

- Technical gimmick: exploit usenet structure to distribute ratings (Resnick et al. 1994; http://www.movielens.umn.edu/)

AIFB

Predict Scores

- Determine Correlation Coefficient
  - **+1 perfect agreement**
  - **0 no correlation**
  - **-1 perfect disagreement**
  - χ_{XL} **:= cov(X,L) / (σ_X σ_L)**

  Estimate rating:
  - **Combine scores according to χ_{XL} (Resnick et al. 1994)**

- Use Singular Value Decomposition (LSI)

  (Billsus & Pazzani, 1998)

- RankBoost:

  Determine ranks for ratings

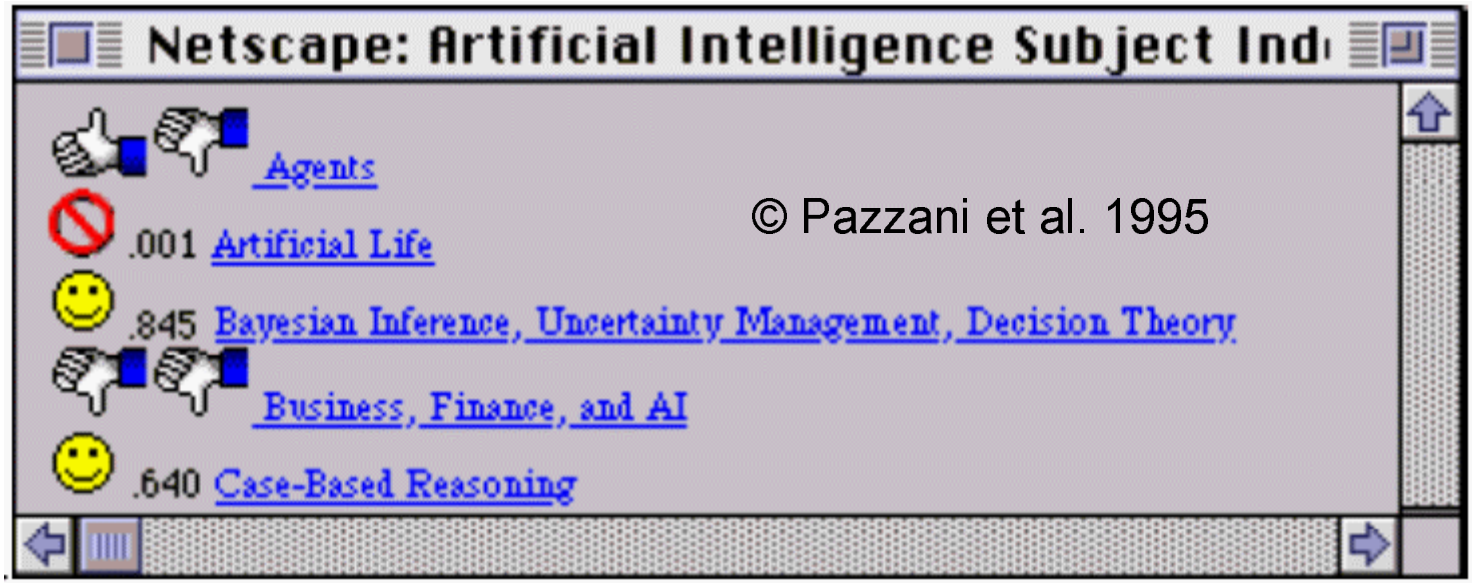
  Combine many weak learners into a strong learner by changing the distribution along the way in order to improve on wrong choices

  Each weak learner is simply a single user's opinion that does the best ranking given the actual distribution

  (Freund et al. 1998)

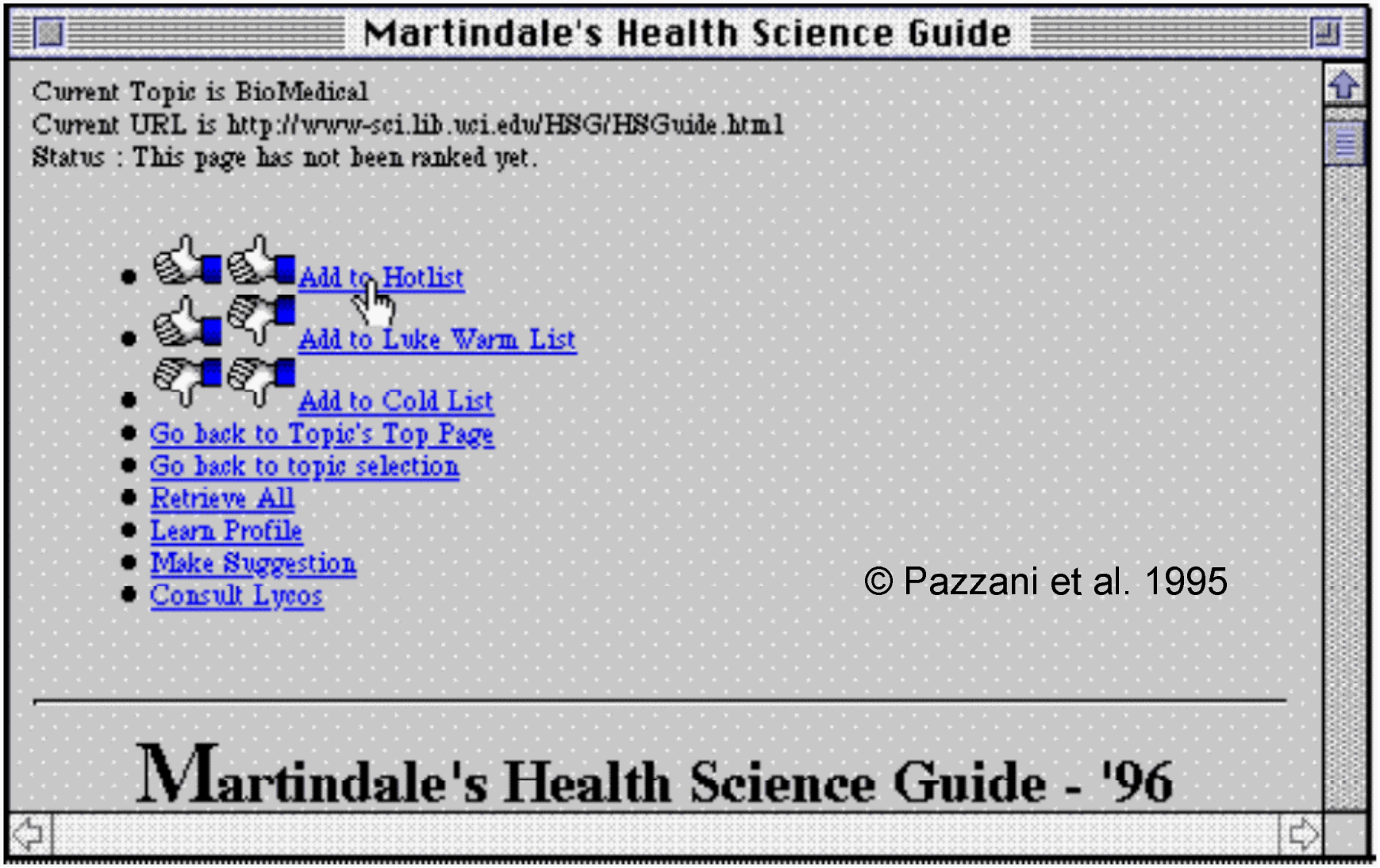
Syskill & Webert

- Rank pages accessible from current page

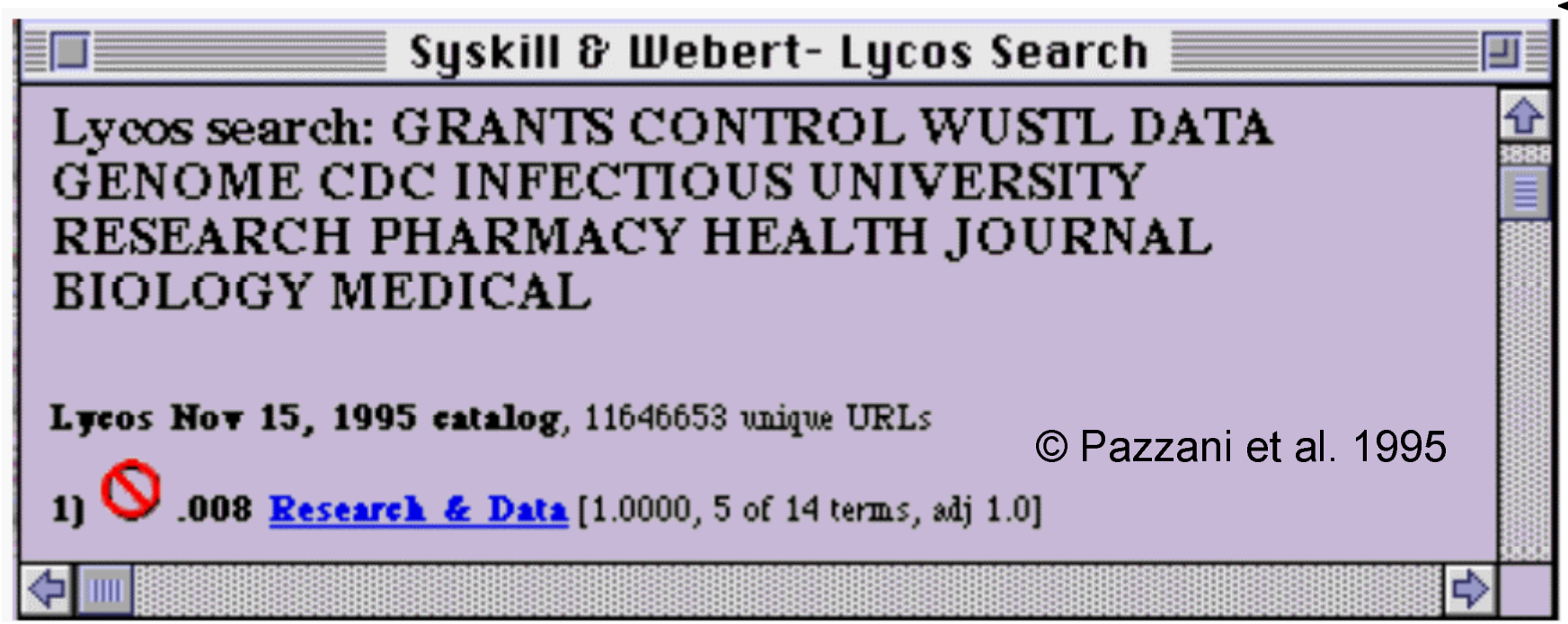


- Learn relevance from input:
  - Word vectors (selection of relevant words)
  - user-defined words indicating (un-)interestingness,

Syskill & Webert



Syskill & Webert



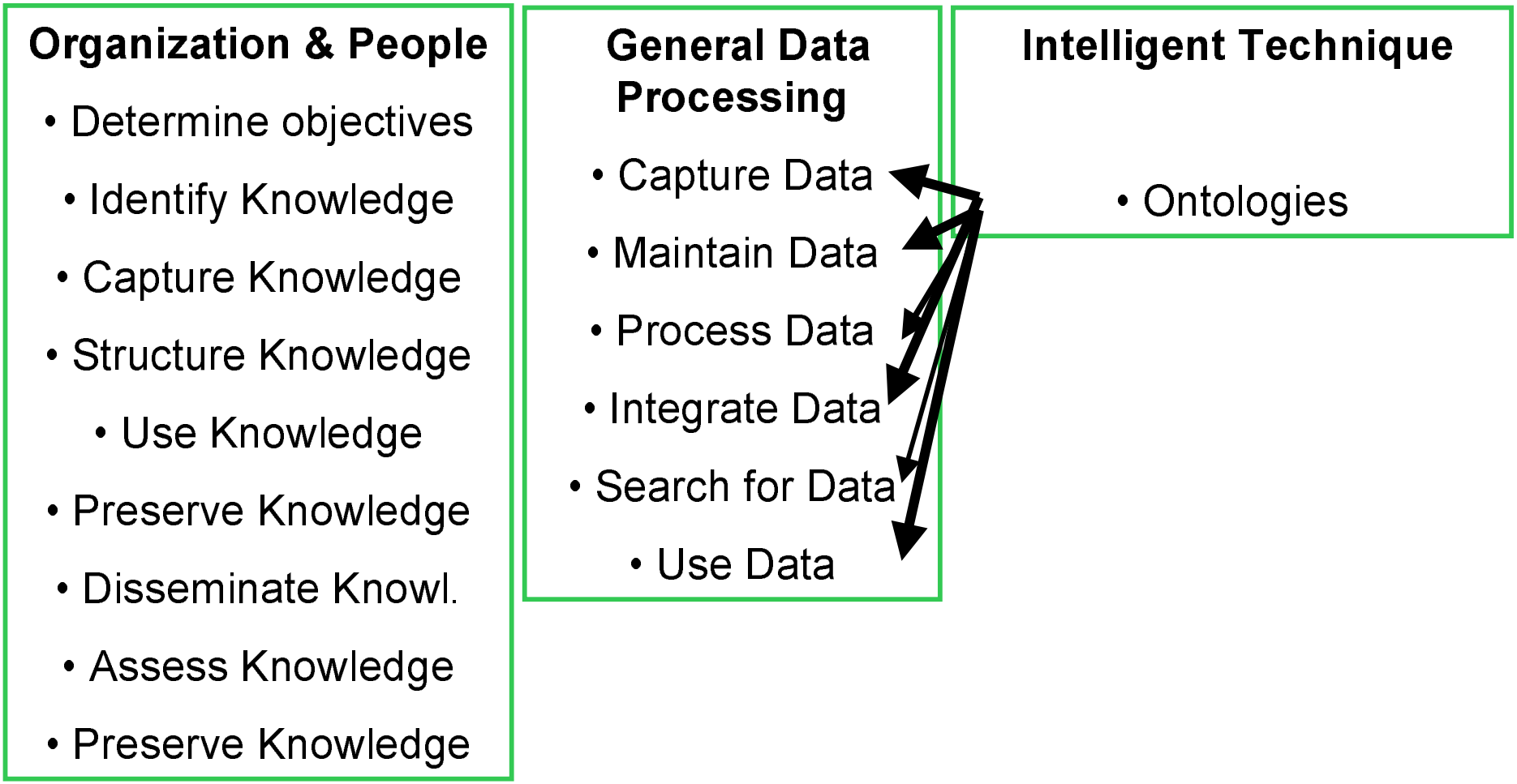
- Evaluation of Learning Algorithms
  - Bayes & Rocchio pretty good

AIFB

Ontology-based Knowledge Management

3.4 Ontology-based Knowledge Management

AIFB



Plan

- The Shape: Motivation
- The Skeleton: Framework
- The Meat: Ontology-based Tools for KM

Factors of Production

Capital

Land

Effective and Efficient Use!

Labour

Knowledge

AIFB

Eff-Use of Knowledge

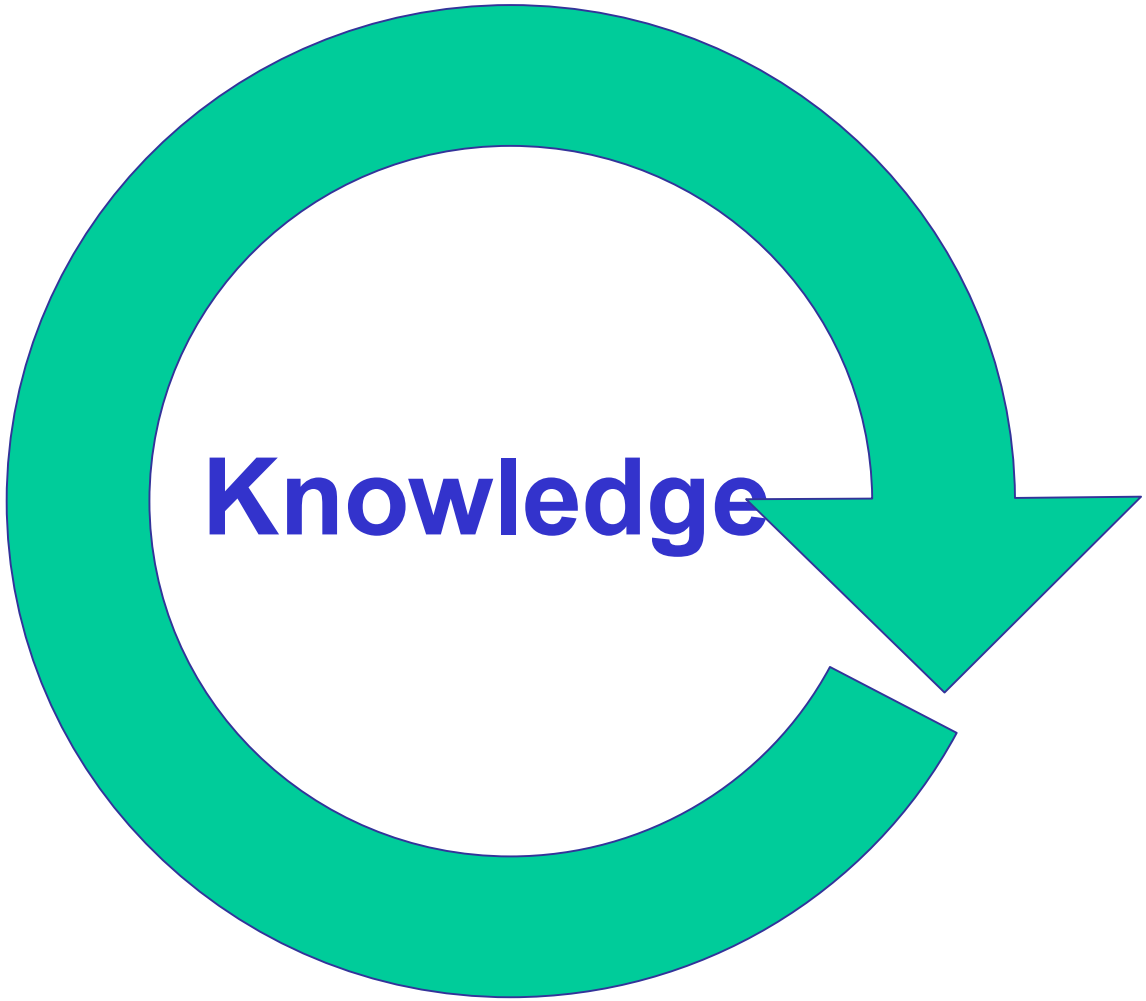
Capital

Land

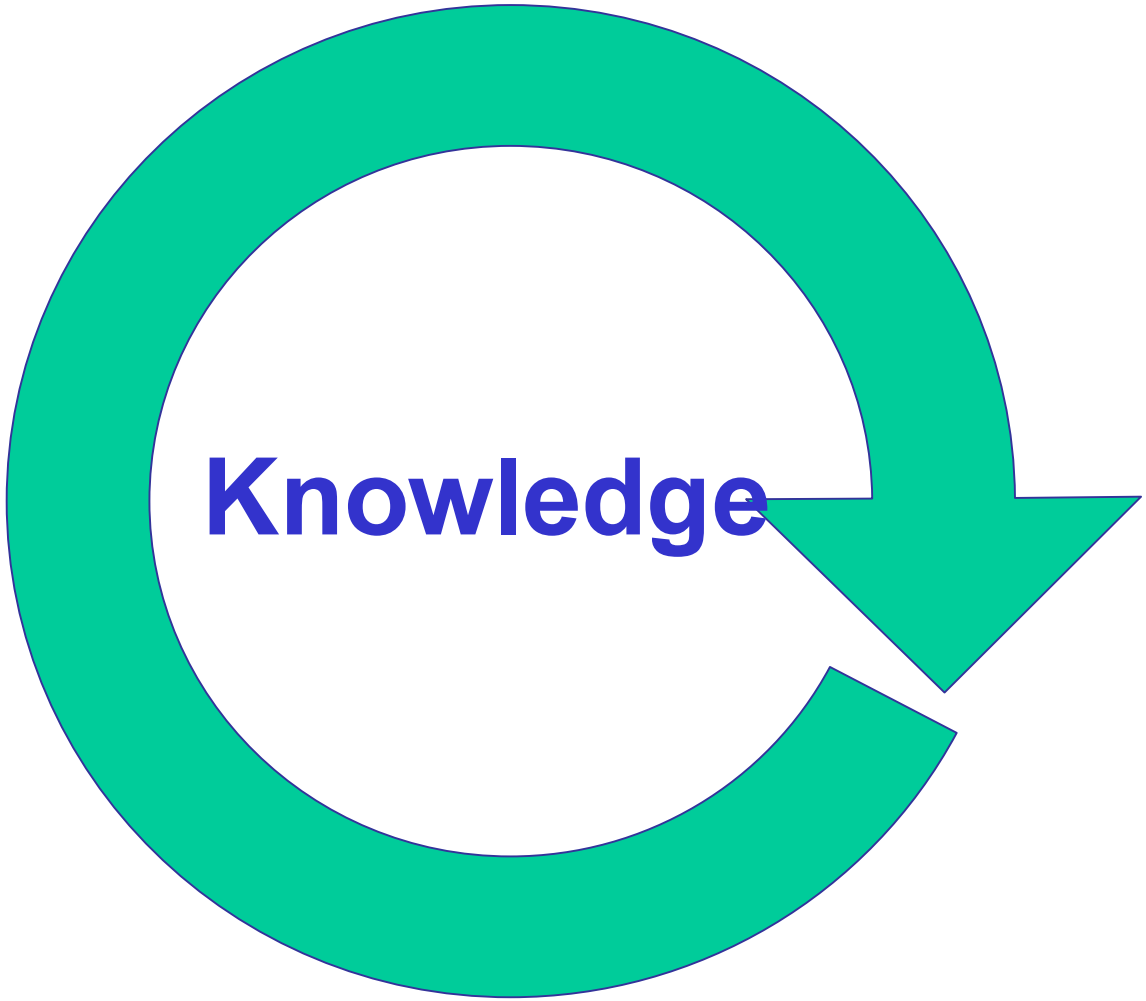
Knowledge

Labour

Eff-Use of Knowledge

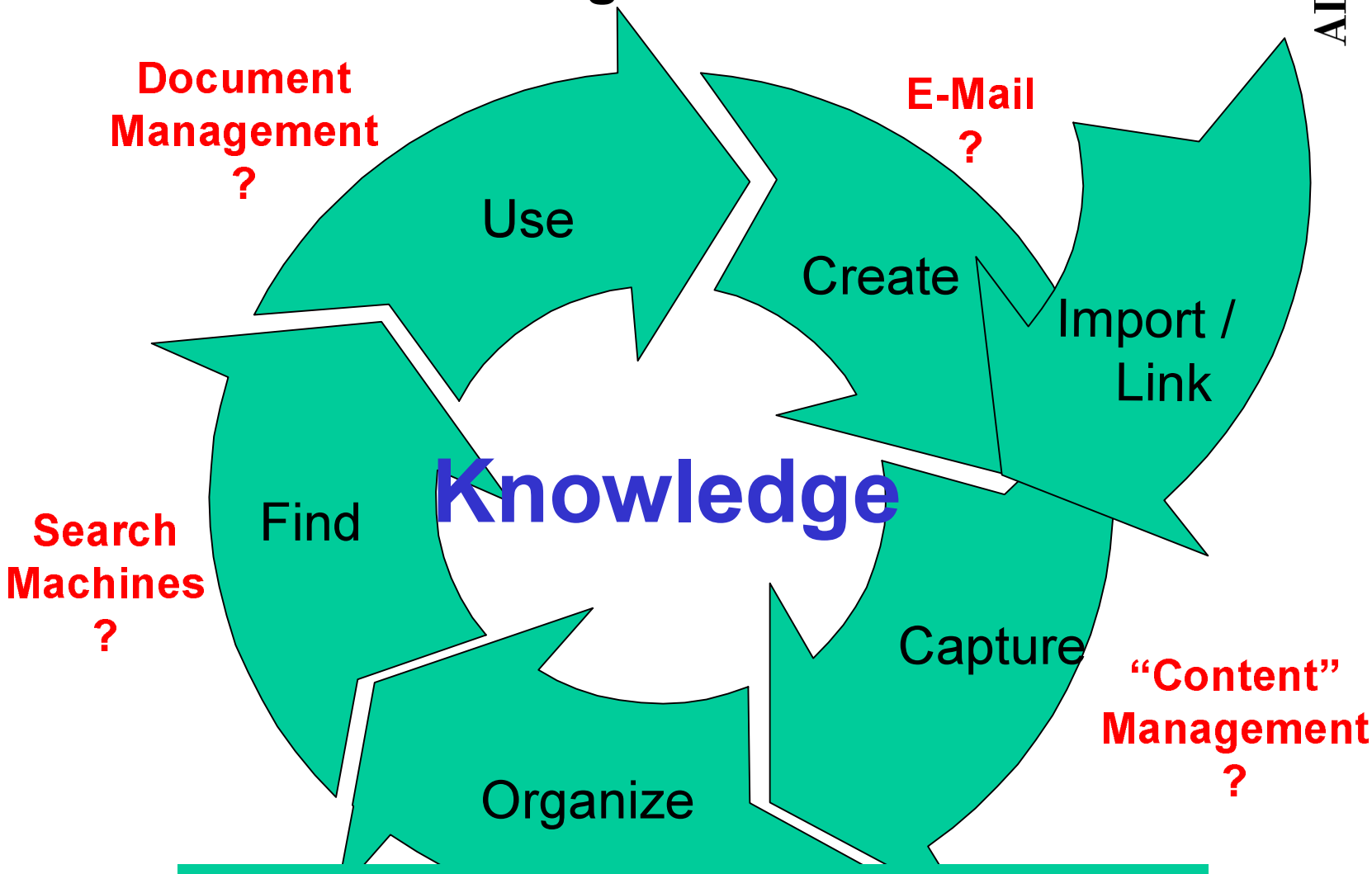


Improve Knowledge Life Cycle



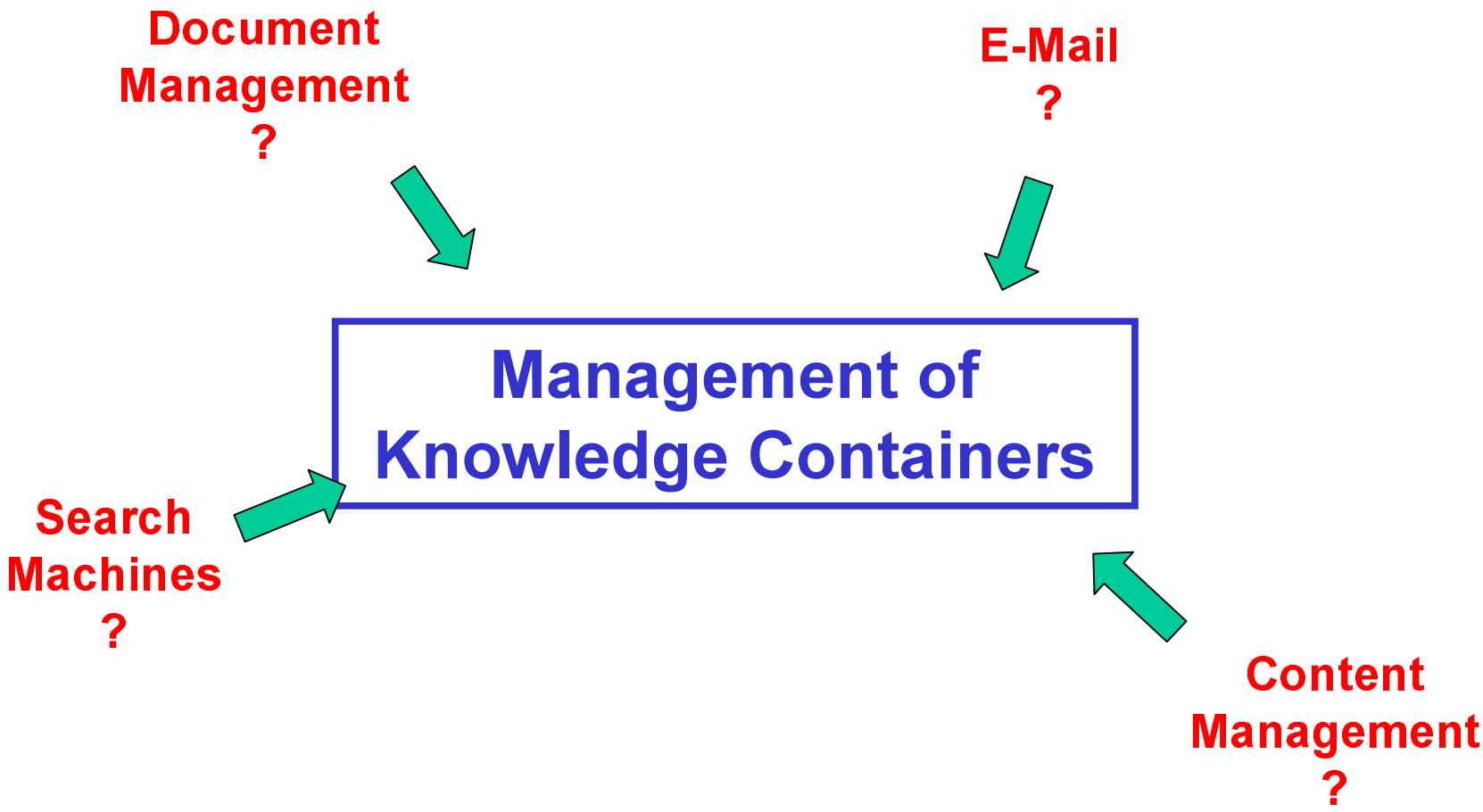
AIFB

Subtasks of Knowledge Processes



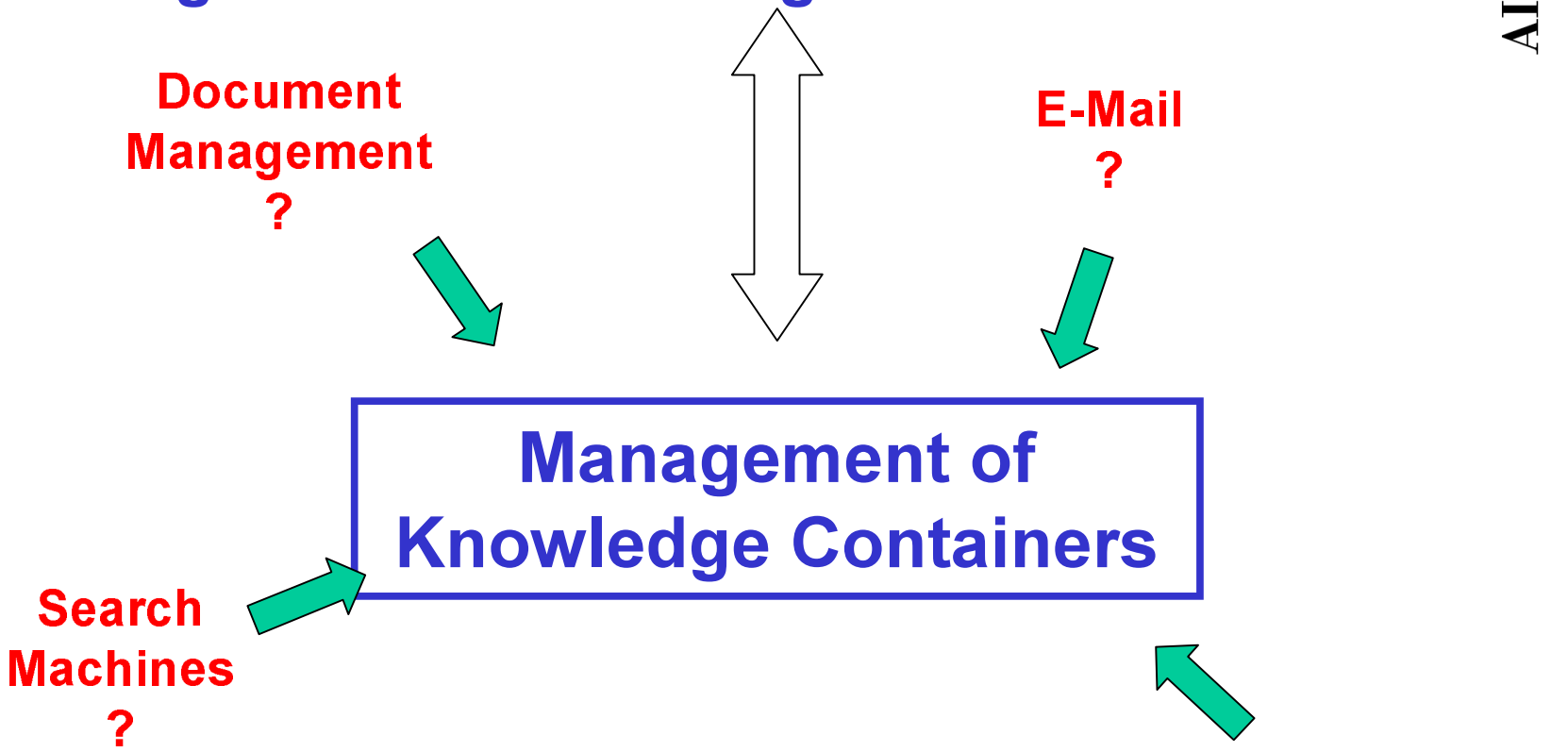
Don't forget the overall process!

Subtasks of Knowledge Processes



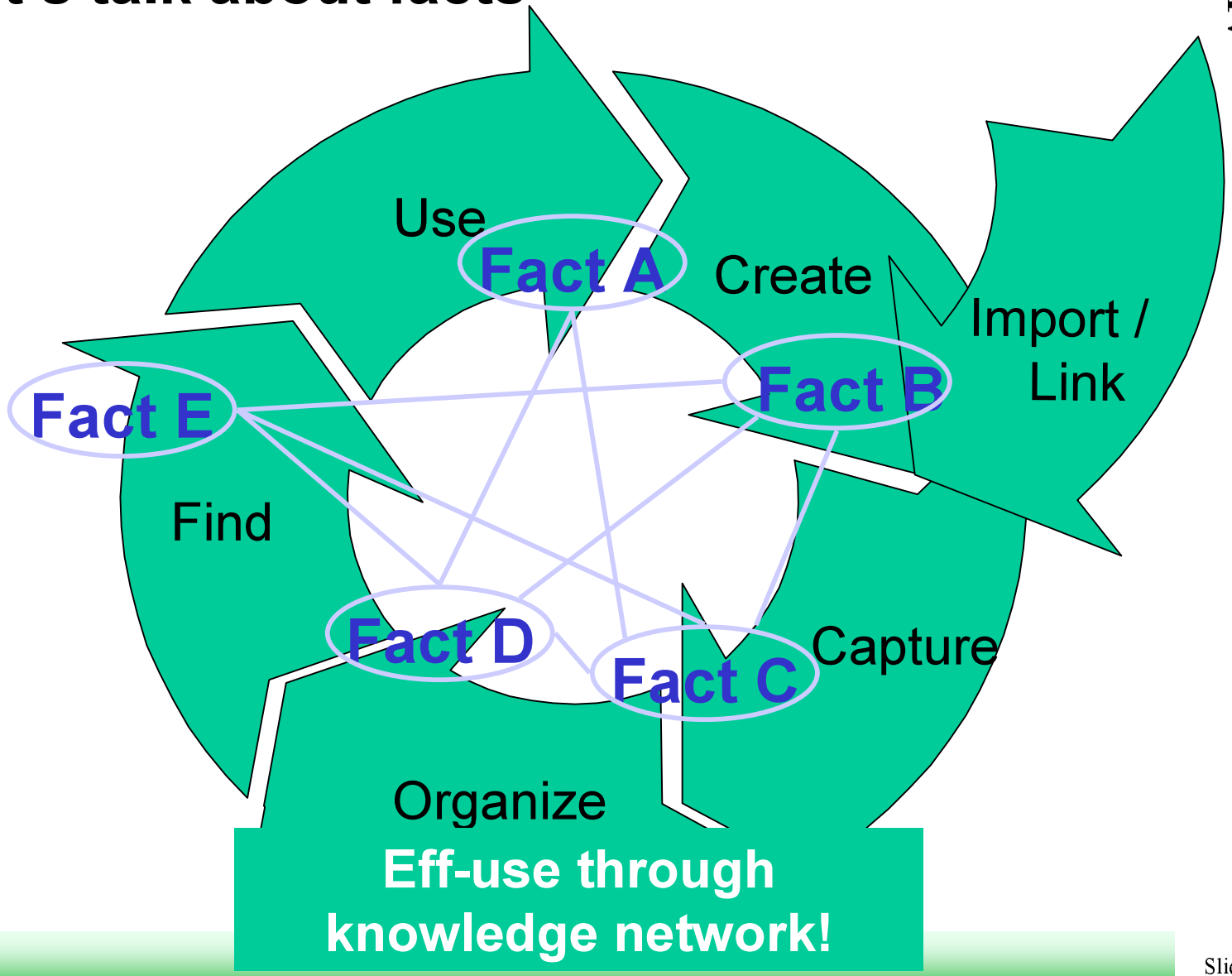
AIFB

Management of Knowledge Contents



containers may change - contents remain important!

Let‘s talk about facts



Ontology

+ Axioms

+ Other Relations

+ Taxonomy

+ Concepts

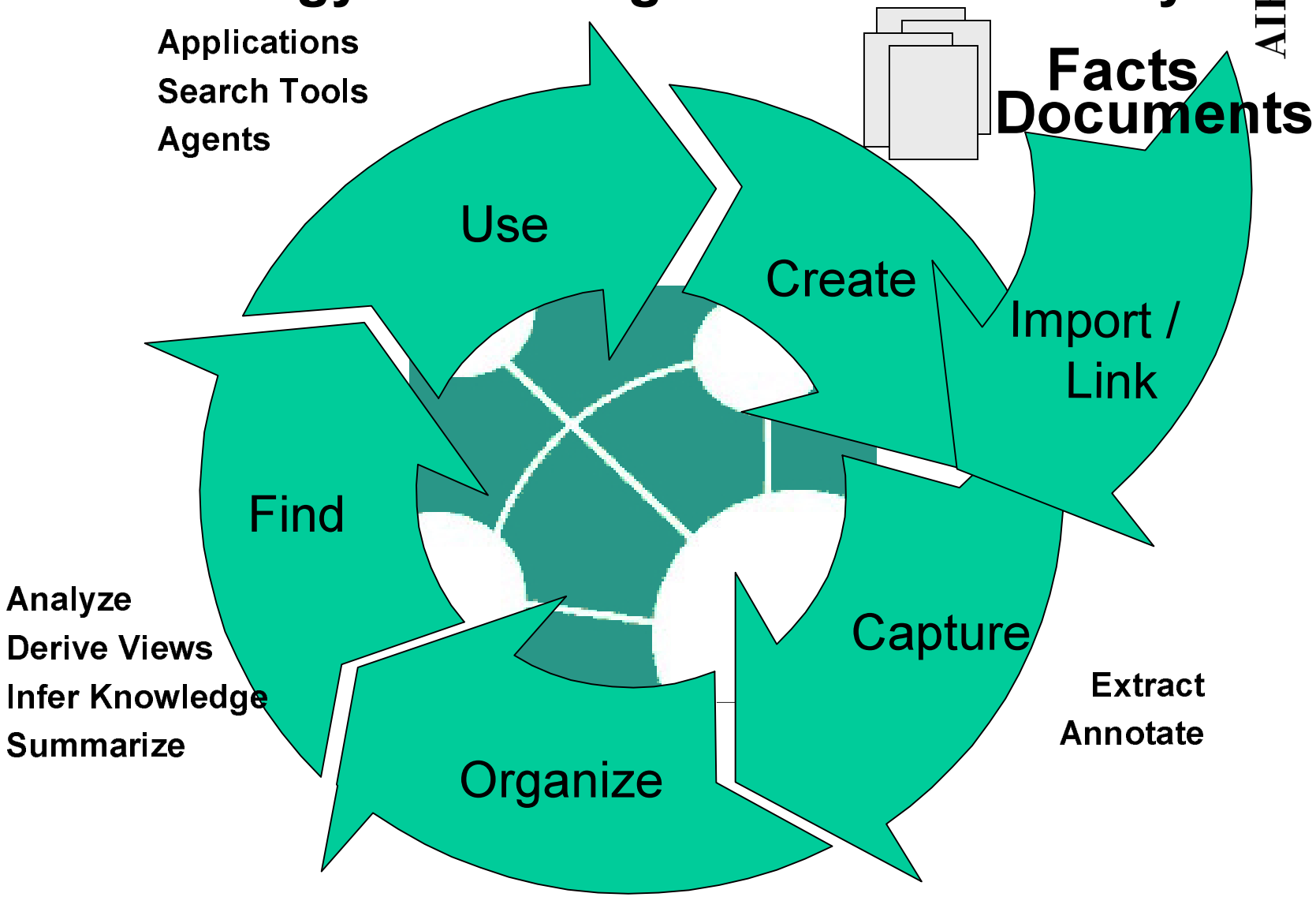
An ontology (in our sense) is ...

...a formal specification ⇒ executable

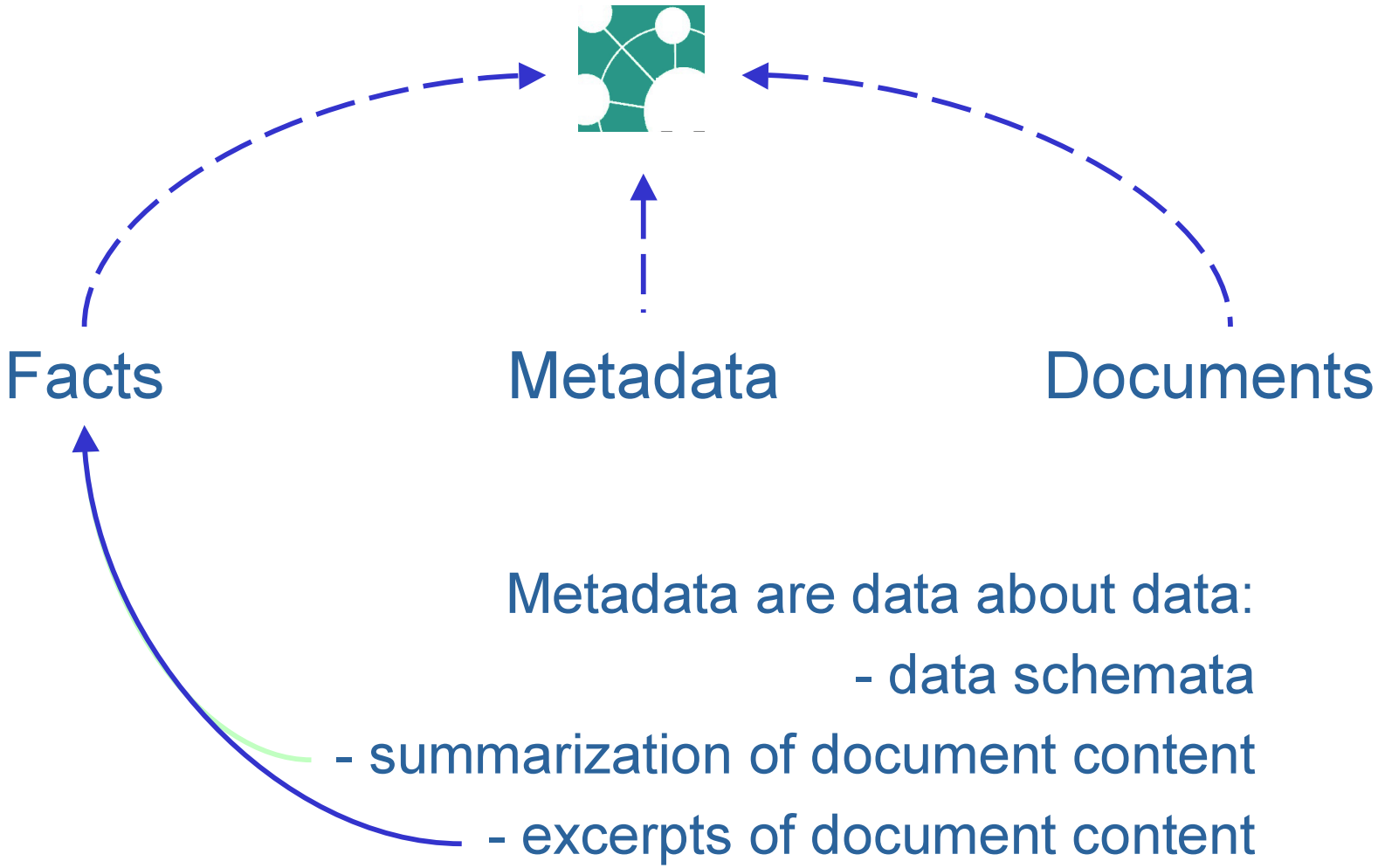
...of a shared conceptualization ⇒ group of people

...of some part of the world that is of interest ⇒ application

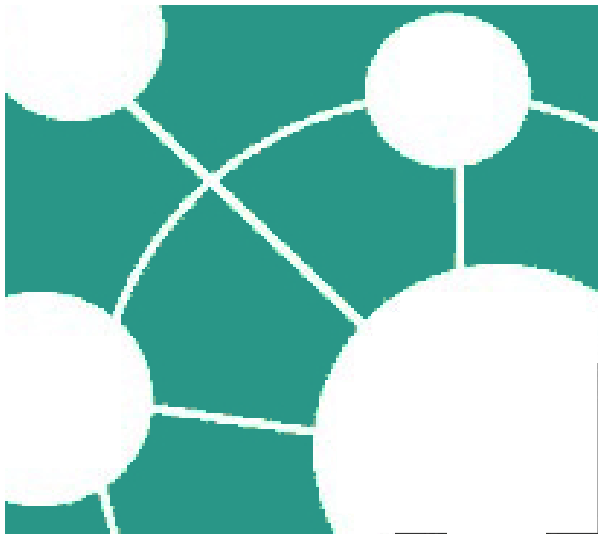
An Ontology-based Organizational Memory



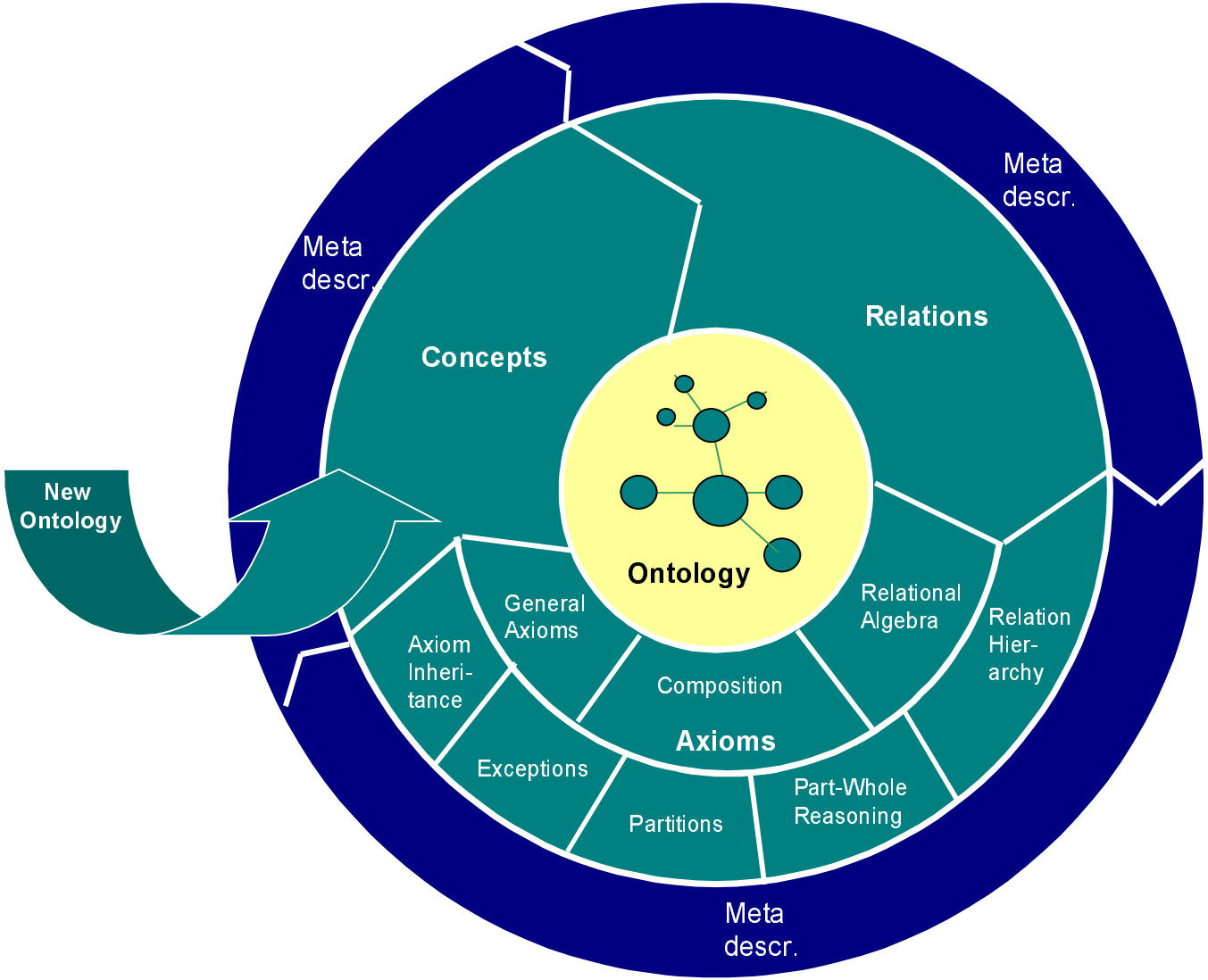
What about Documents?



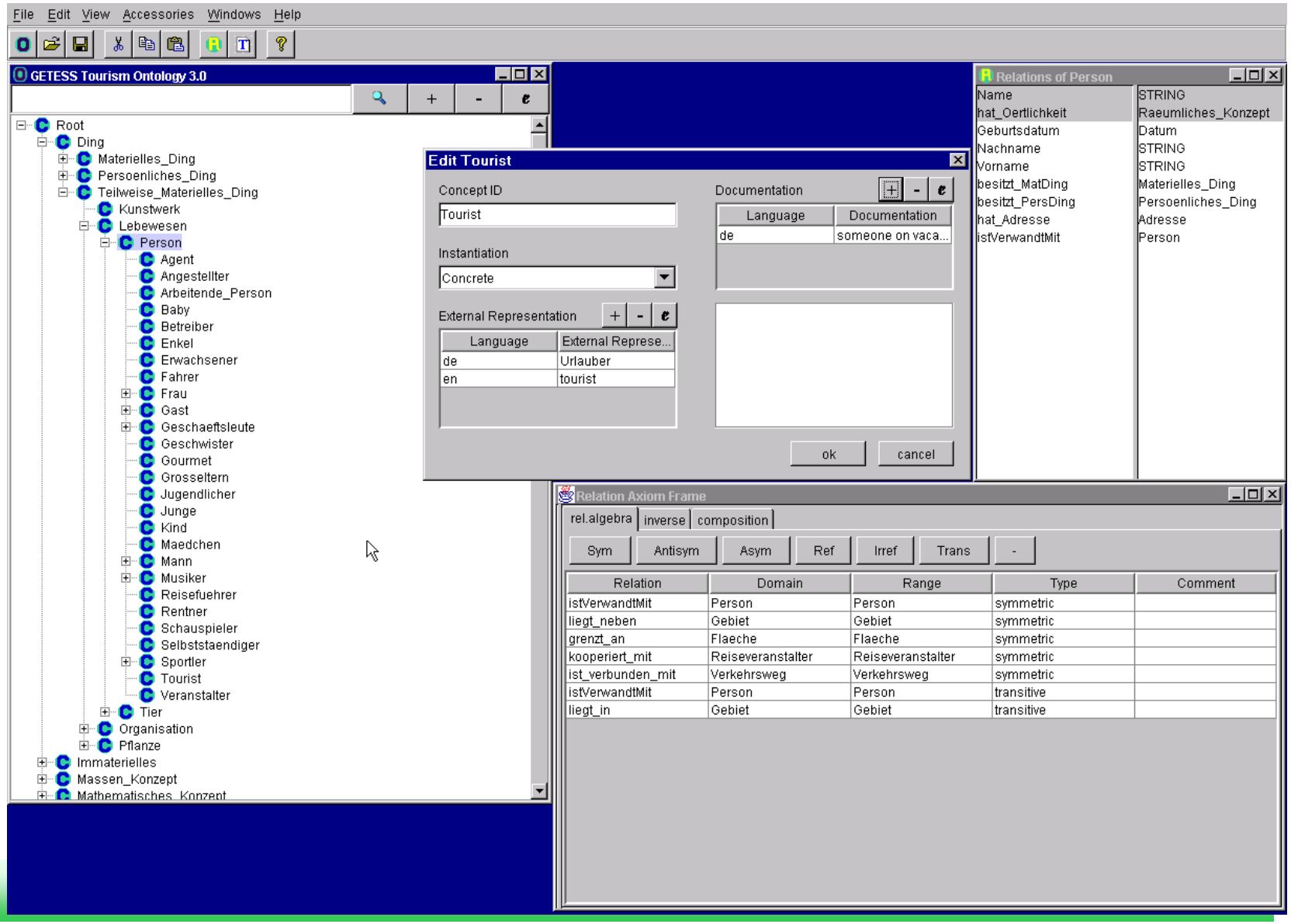
Our Ontology-based OM (or „meat to the skeleton“)



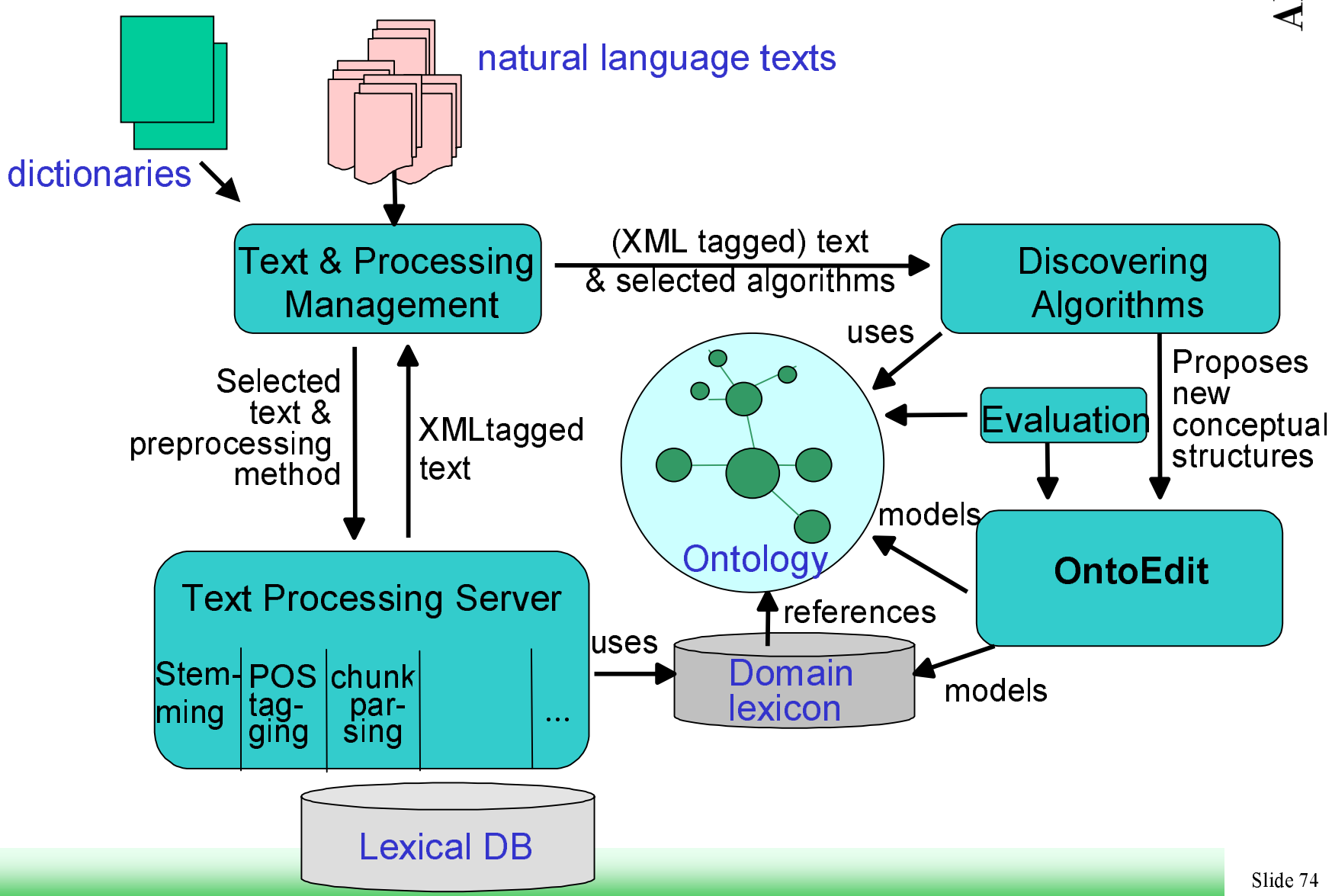
Structure - Ontology Engineering



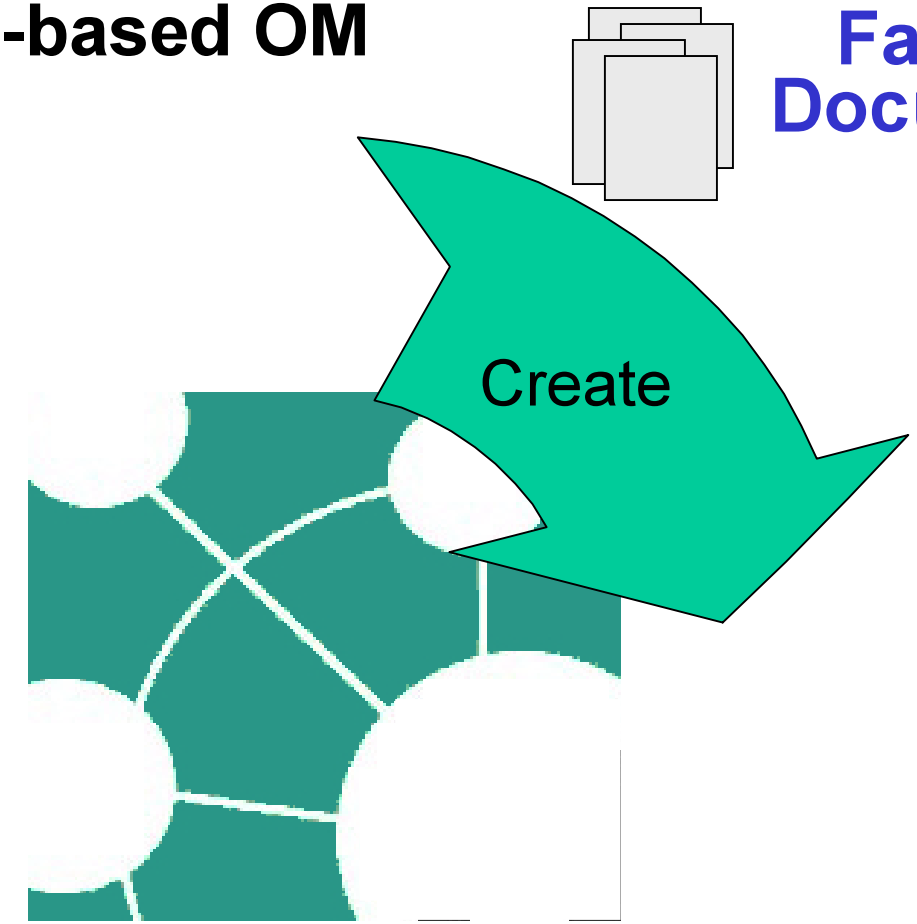
Structure - OntoEdit



Structure - Ontology Learning



Our Ontology-based OM



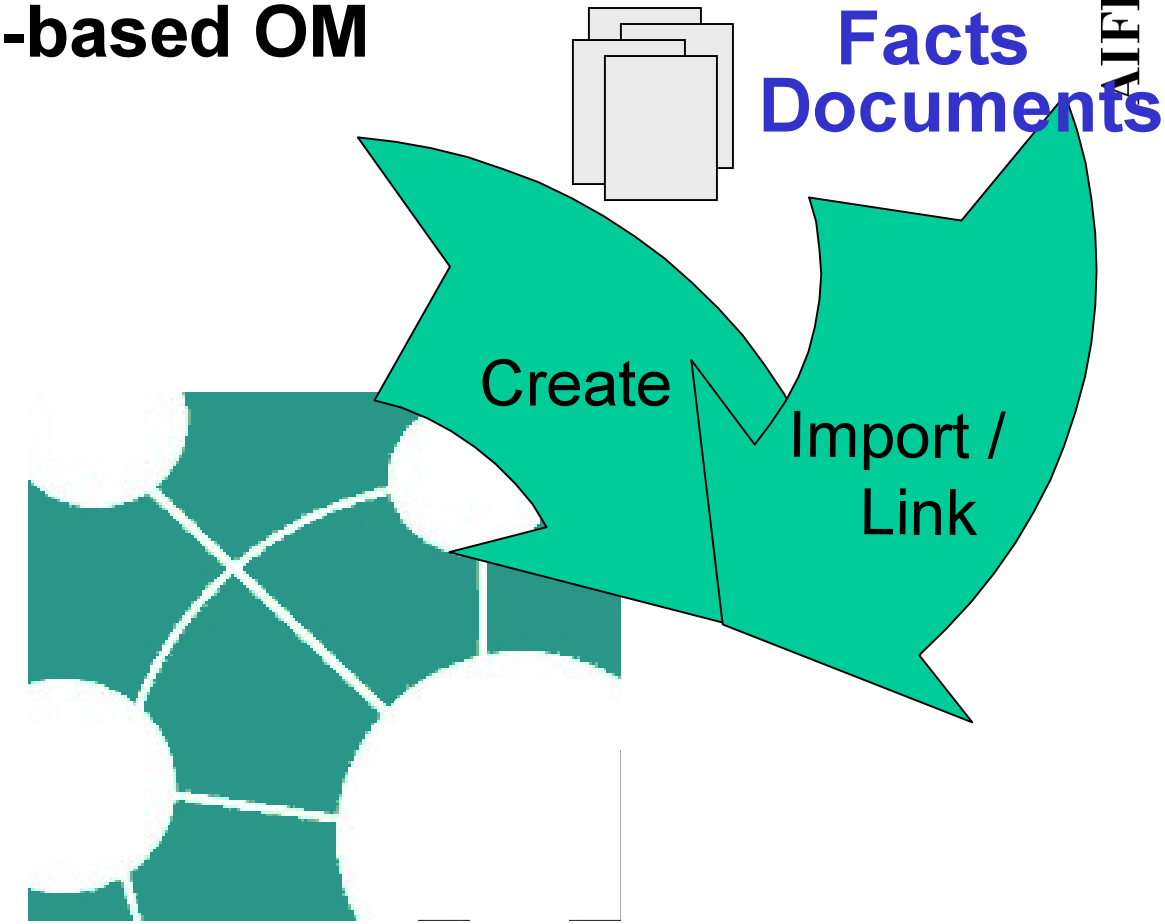
Create - Using Templates

```
<project>
<author>                </author>
<plandate>              </plandate>
<participants>
                        <member>            </member>
</participants>
<Ganttchart>            </Ganttchart>
<tasks>
                        <task>              </task>
</tasks>
</project>
```

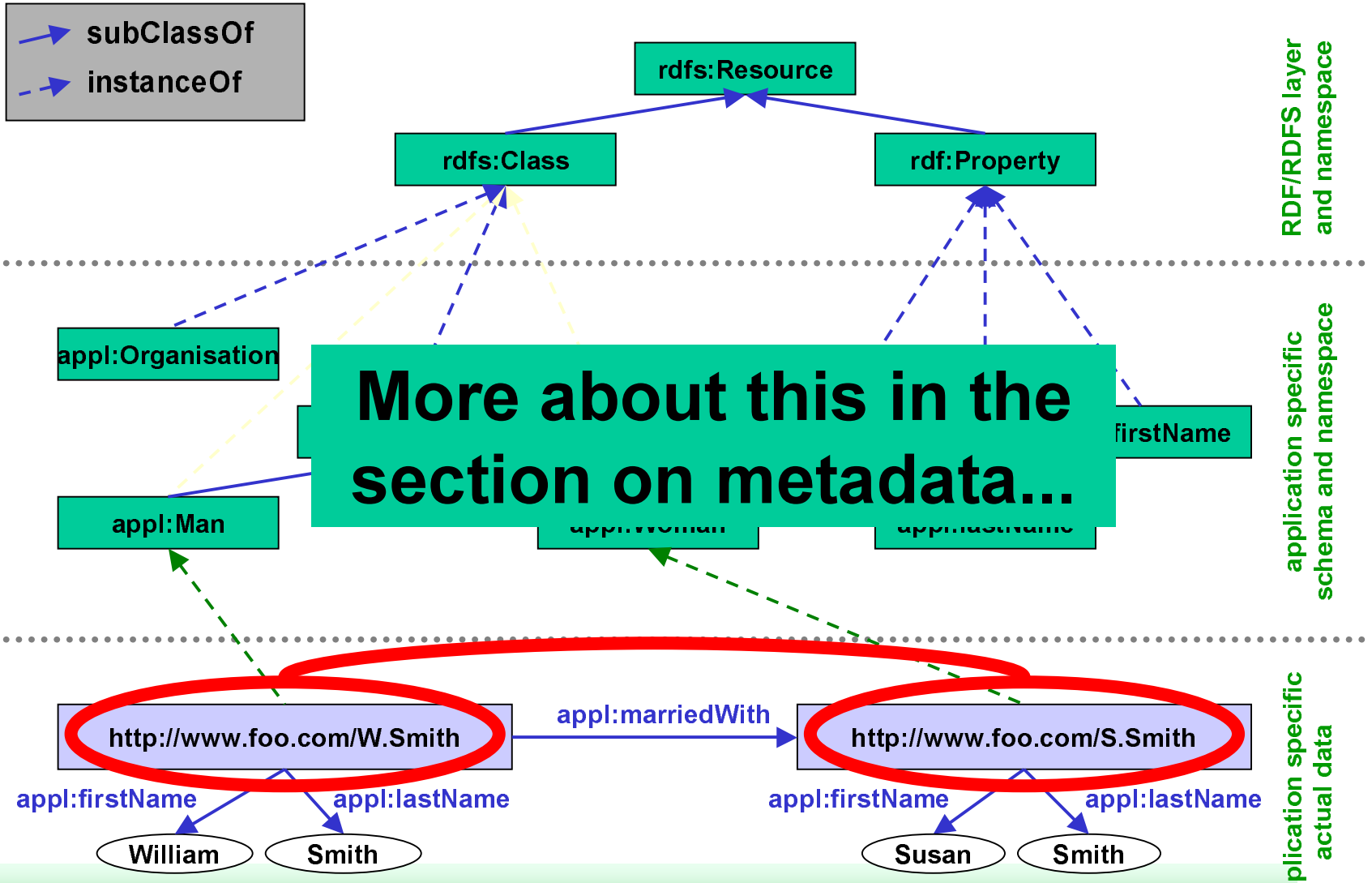
Create - Using Templates


```
<project>
<author>                          Jill Dole                          </author>
<plandate>                        October 18th, 1999                 </plandate>
<participants>
                  <member>   Jill Dole                               </member>
                  <member>   Hans-Peter Schnurr                      </member>
                  <member>   Steffen Staab                           </member>
</participants>
<Ganttchart>                      here goes the table                </Ganttchart>
<tasks>
                  <task>     Analysis of Nordic Life Business Processes  </task>
                  <task>     Analysis of Nordic Life IT environment      </task>
</tasks>
</project>
```

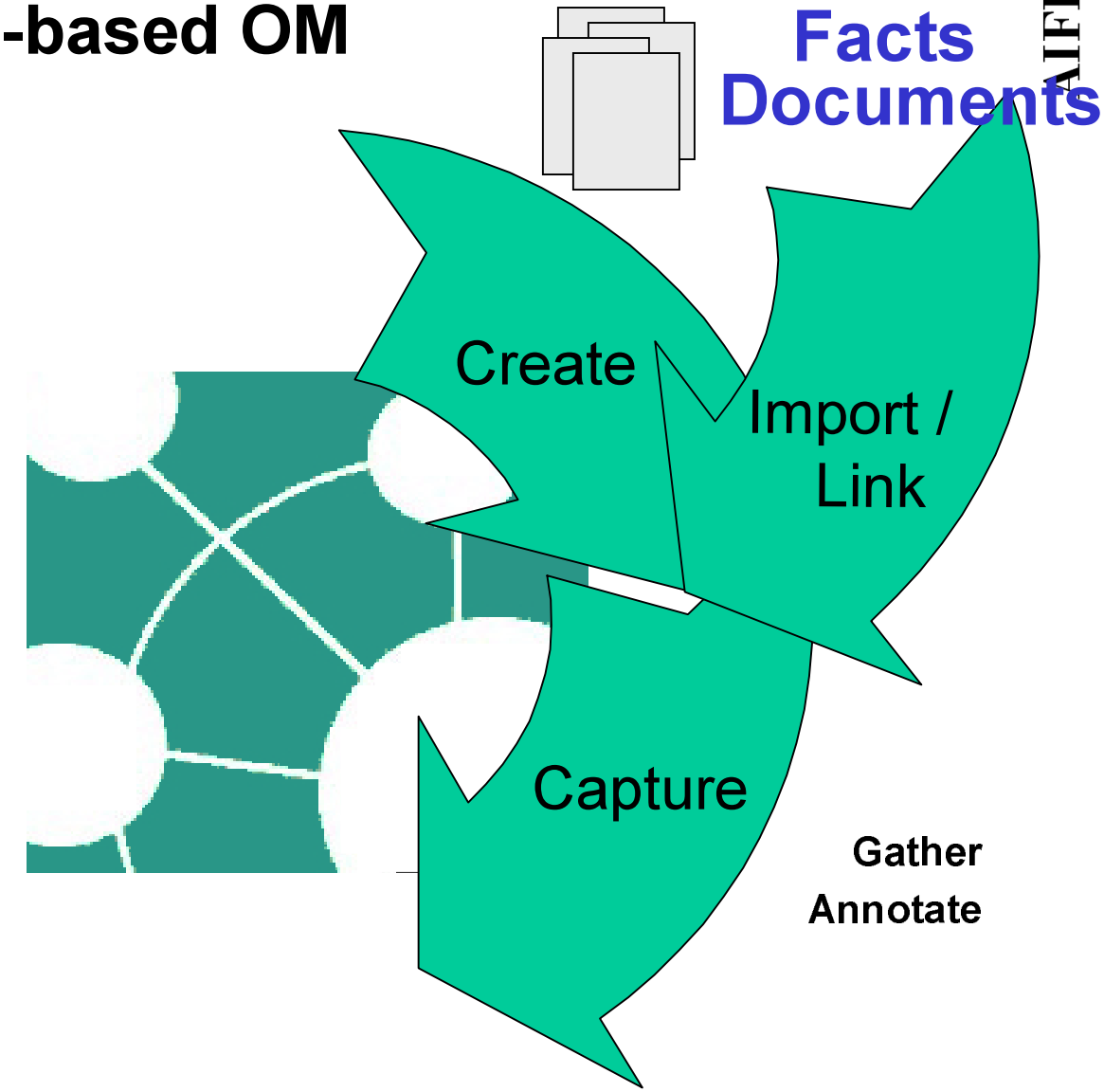
Our Ontology-based OM



Import / Link - RDF(S)

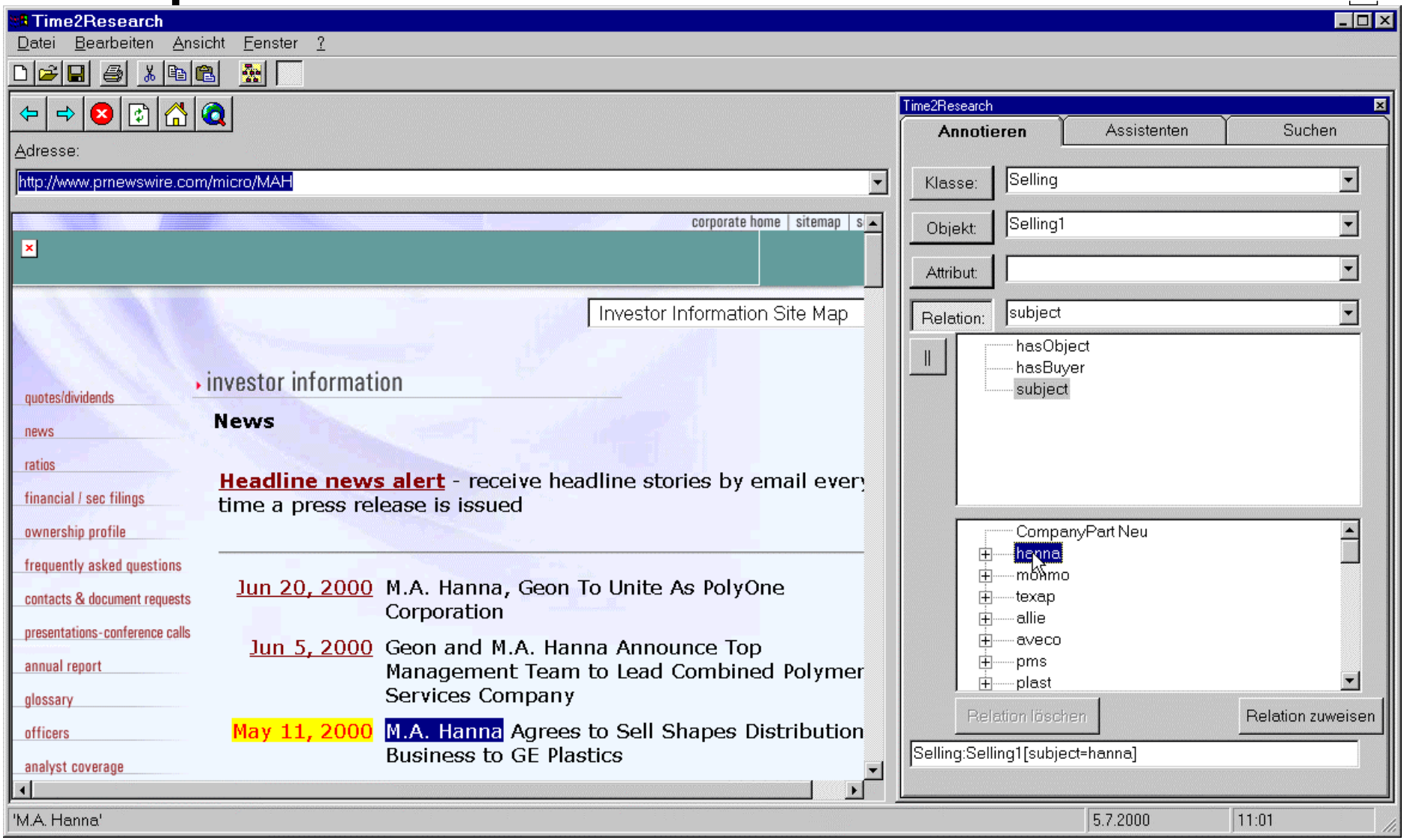


Our Ontology-based OM

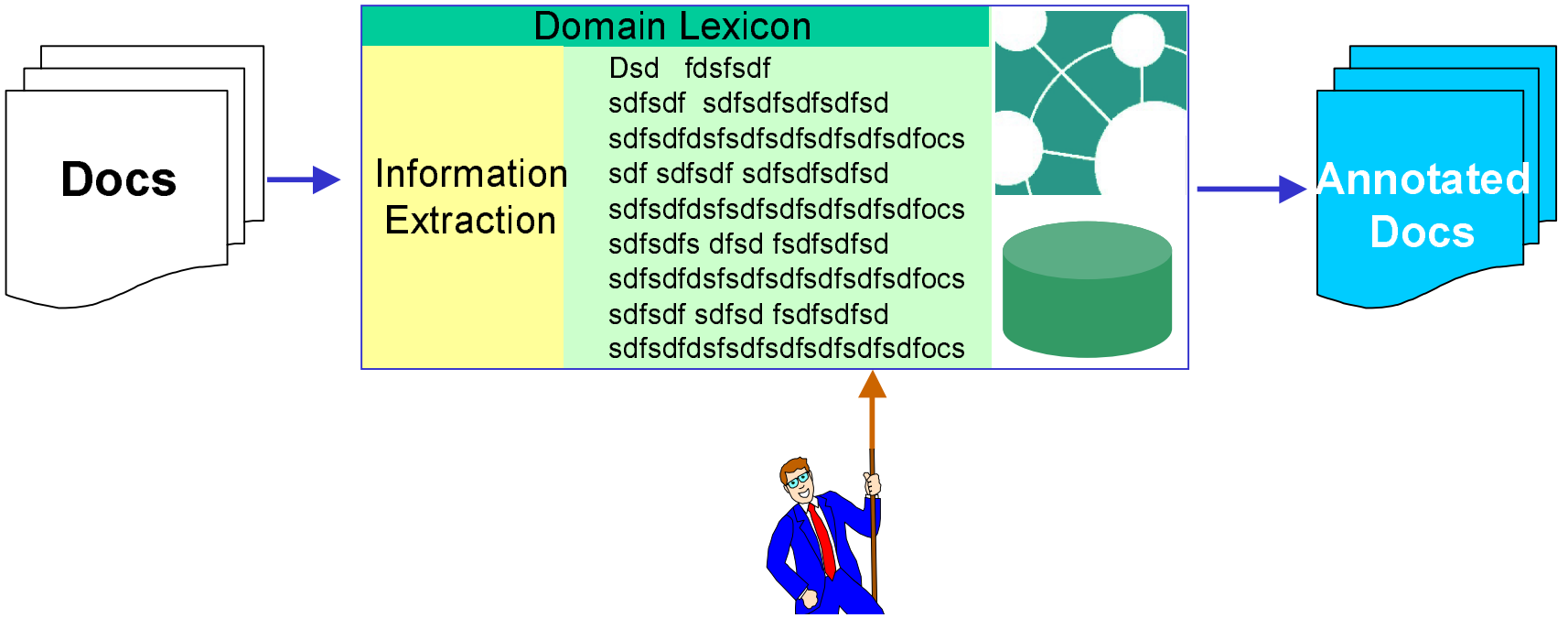


3.4 Ontology-based Knowledge Management

Capture - Manual Annotation

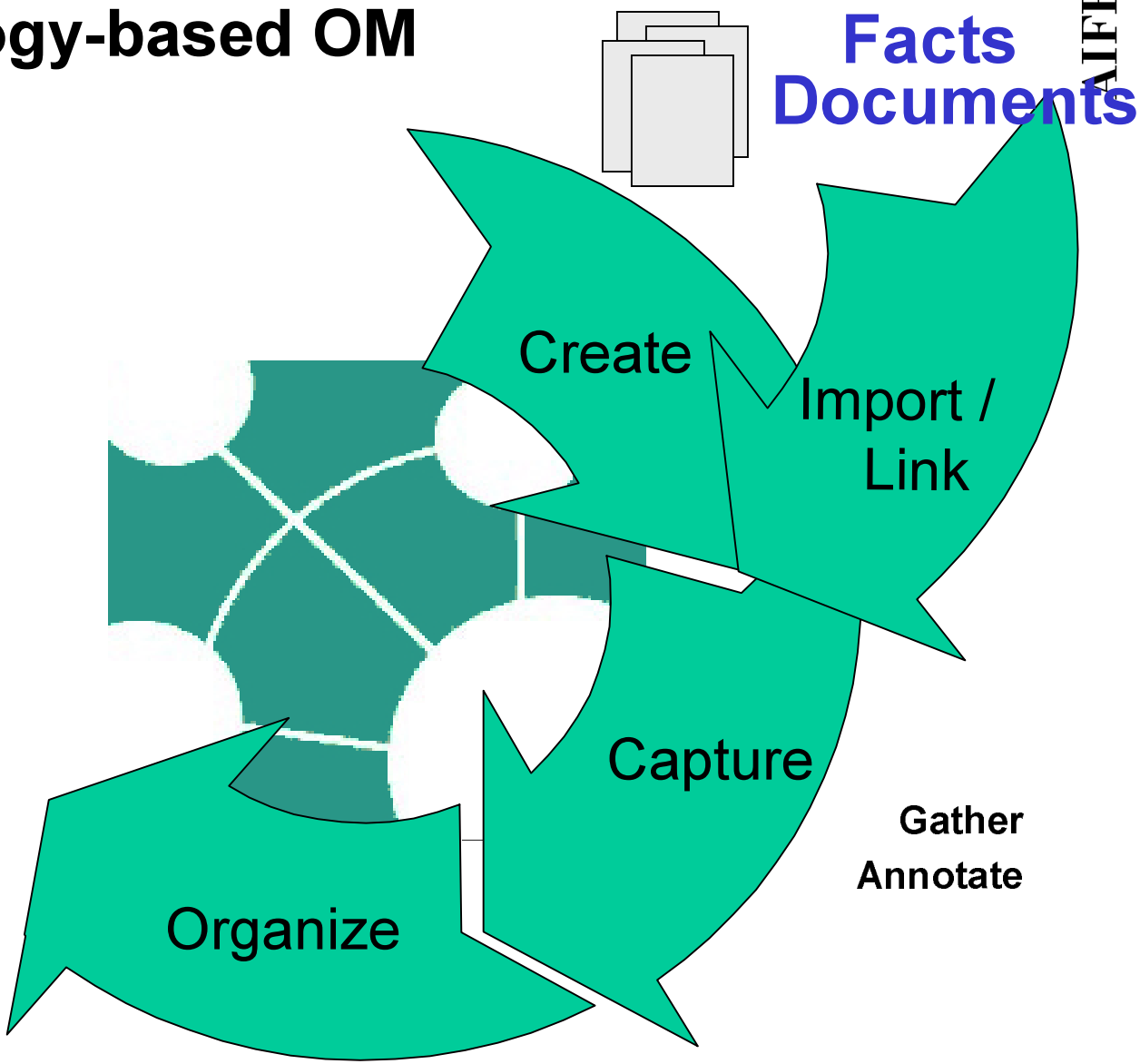


Capture - Semi-automatic Annotation

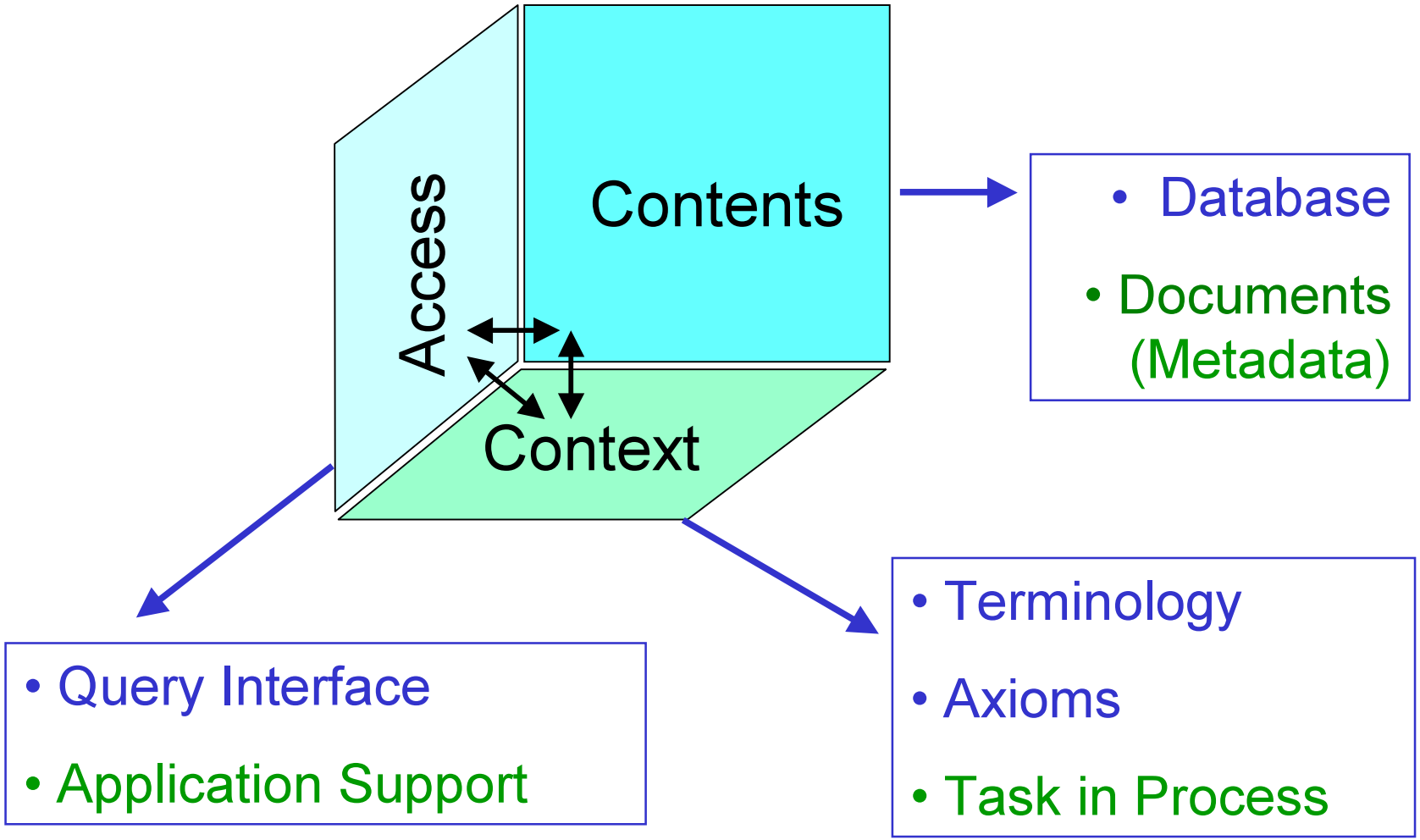


⇒ DAML Project

Our Ontology-based OM

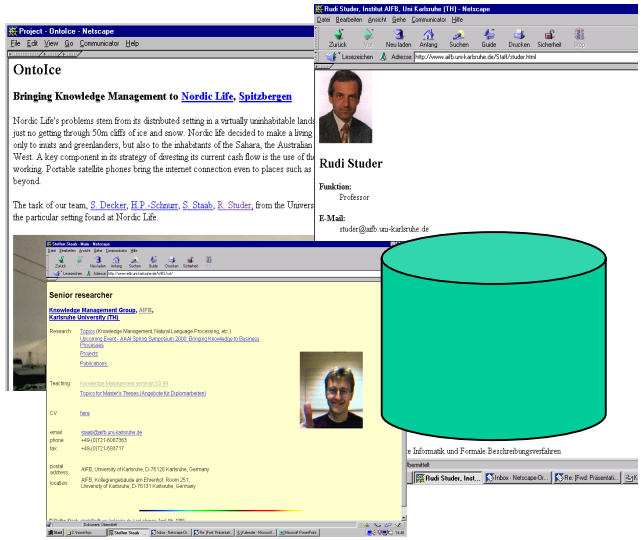


Organize - Process Support (,...)



Organize - Process Support

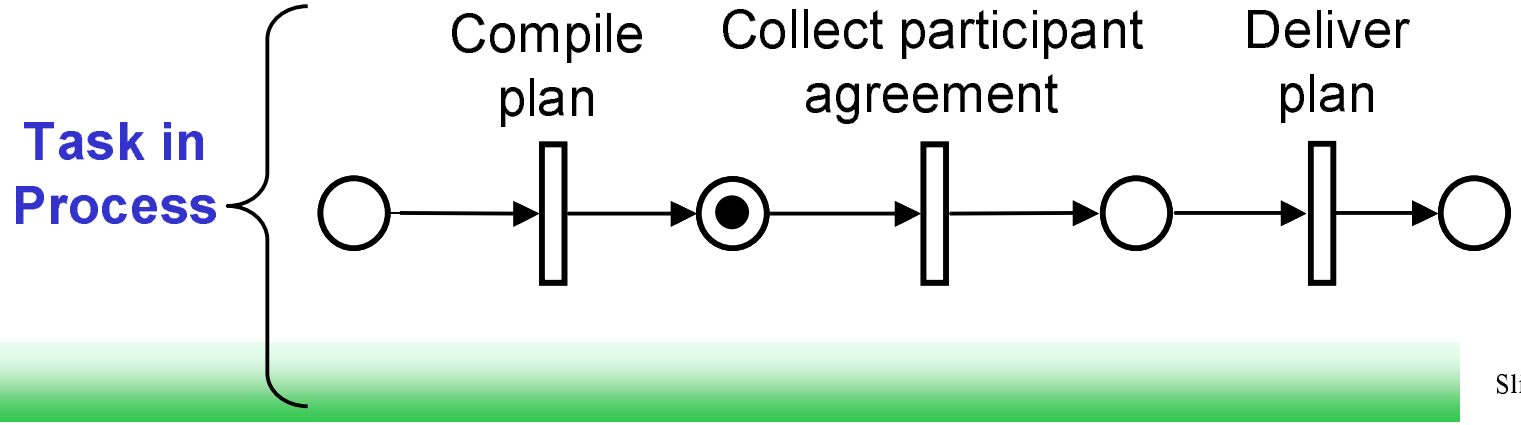
Background Knowledge



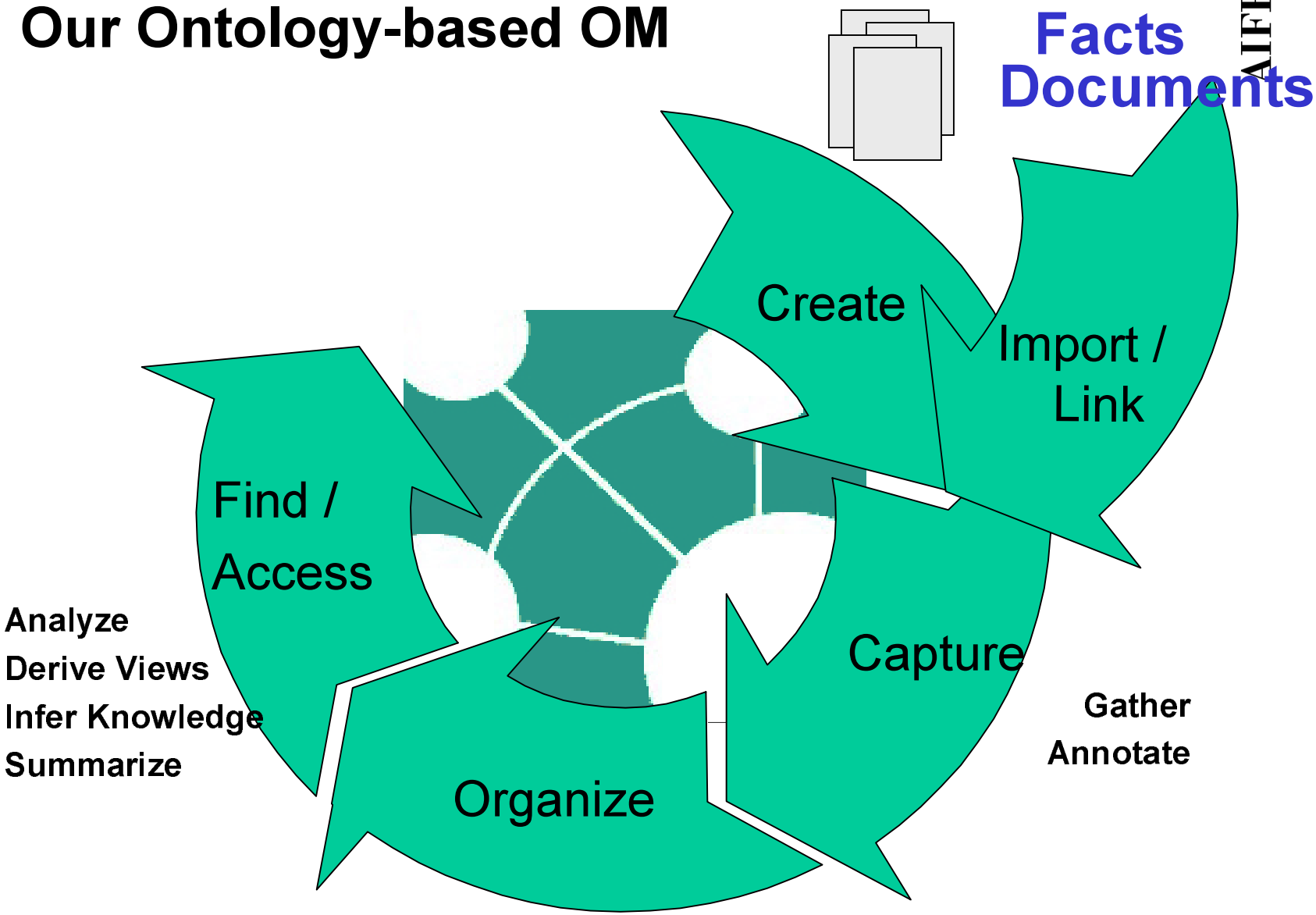
Document Template

partially filled

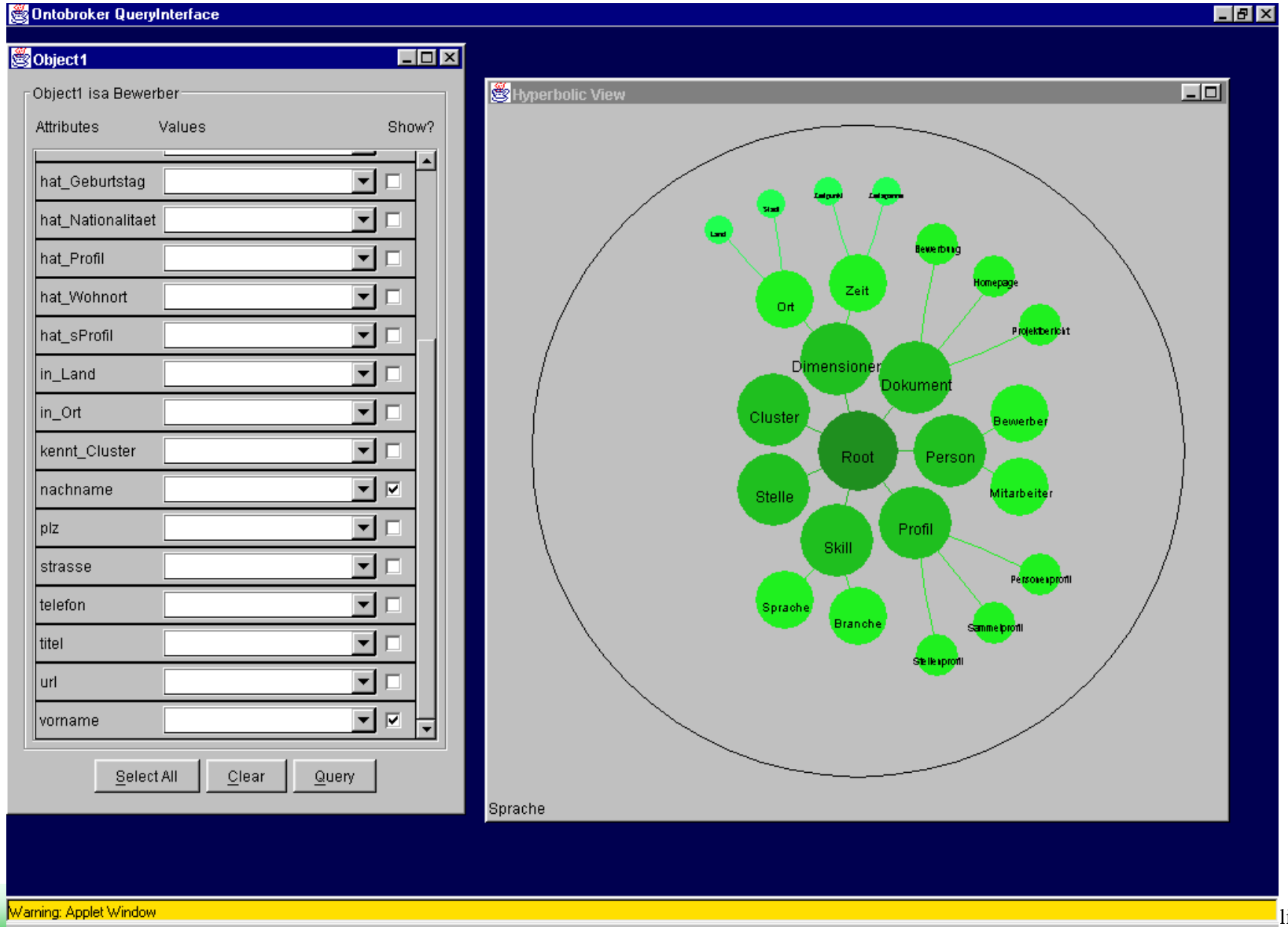
```
<employee>
<name>Rudi Studer</name>
<position>Professor</position>
<email>???????????<email>
 .....
</employee>
```



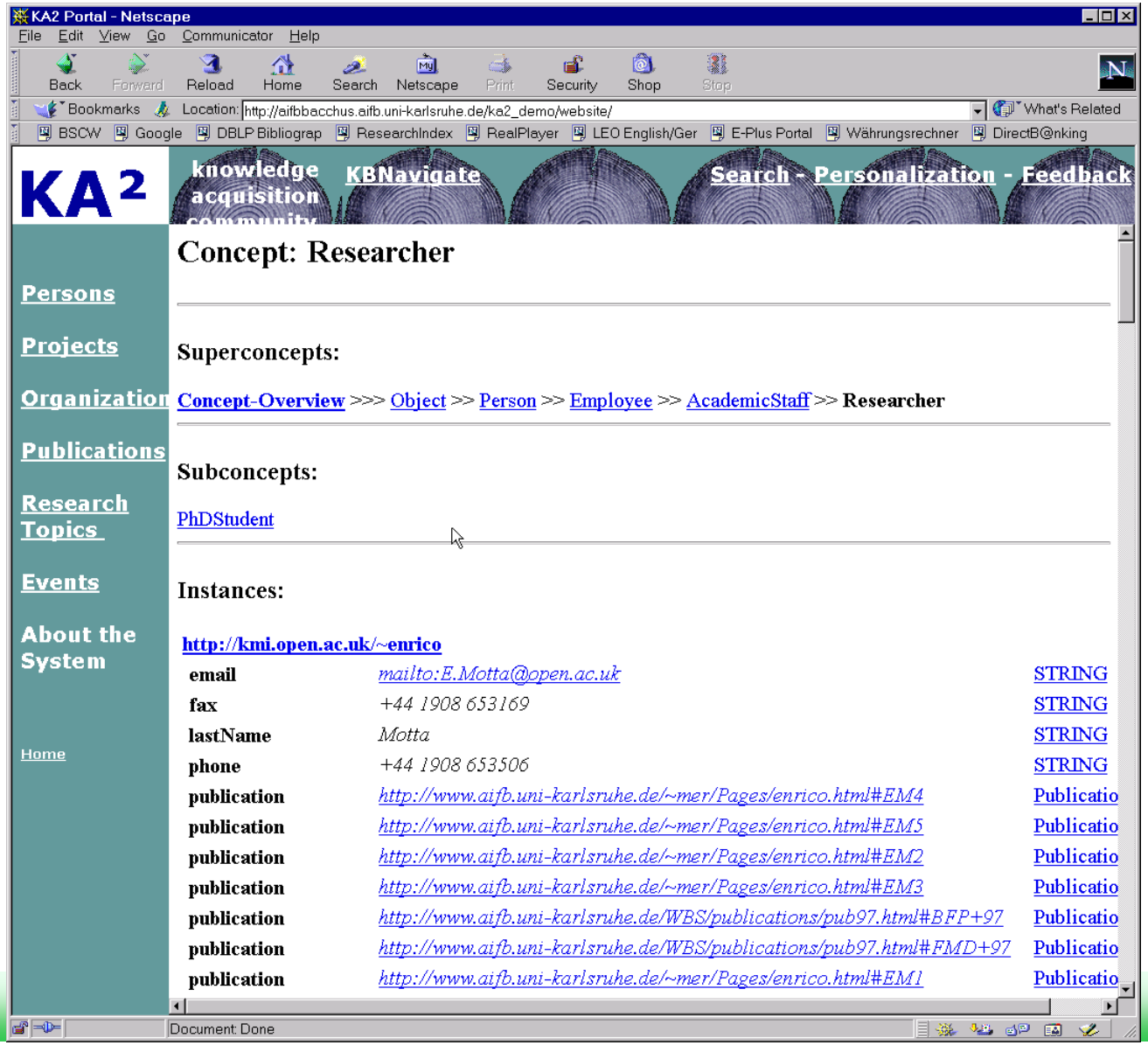
Our Ontology-based OM



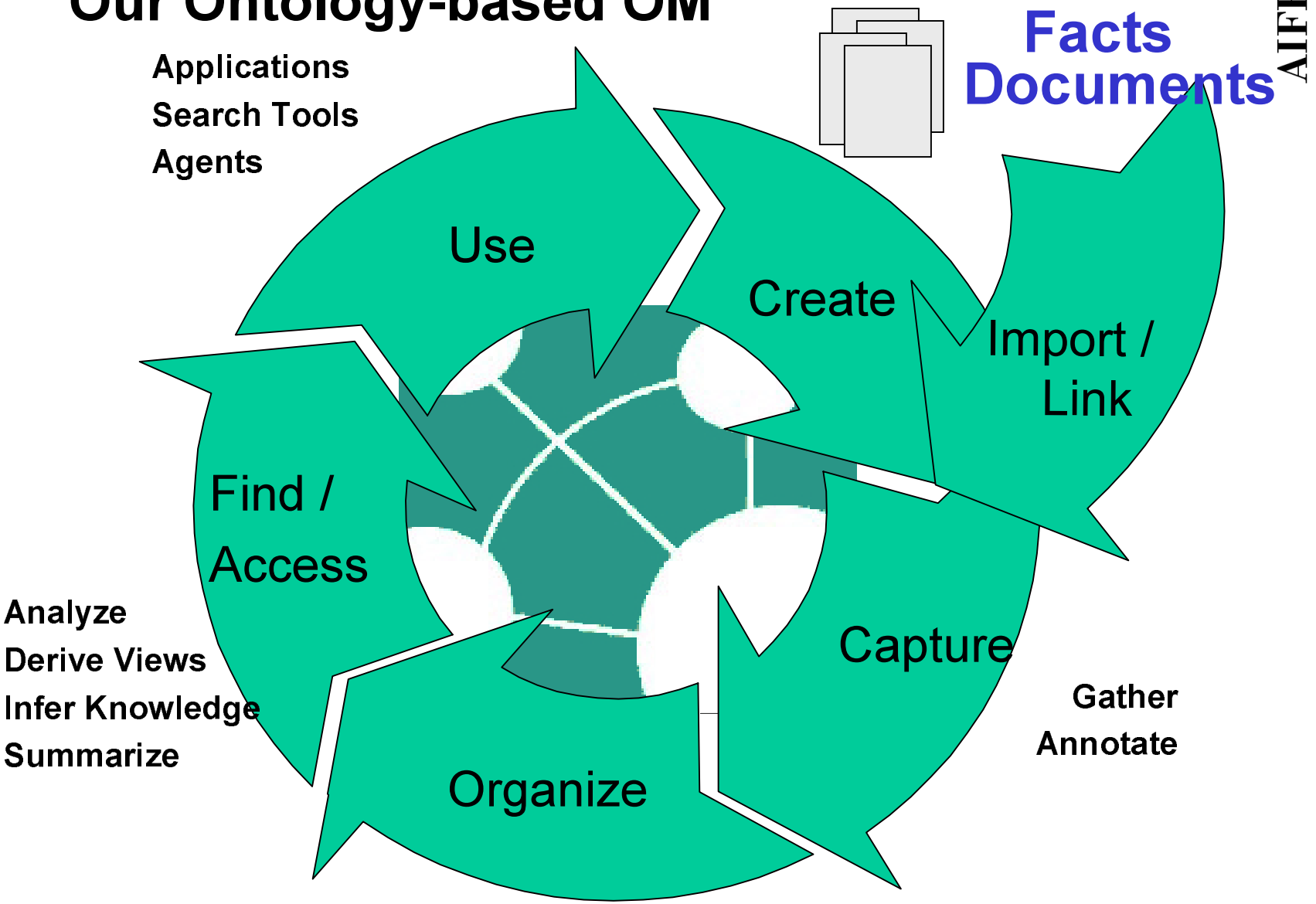
Find - Semantic Search



Find - Semantic Search



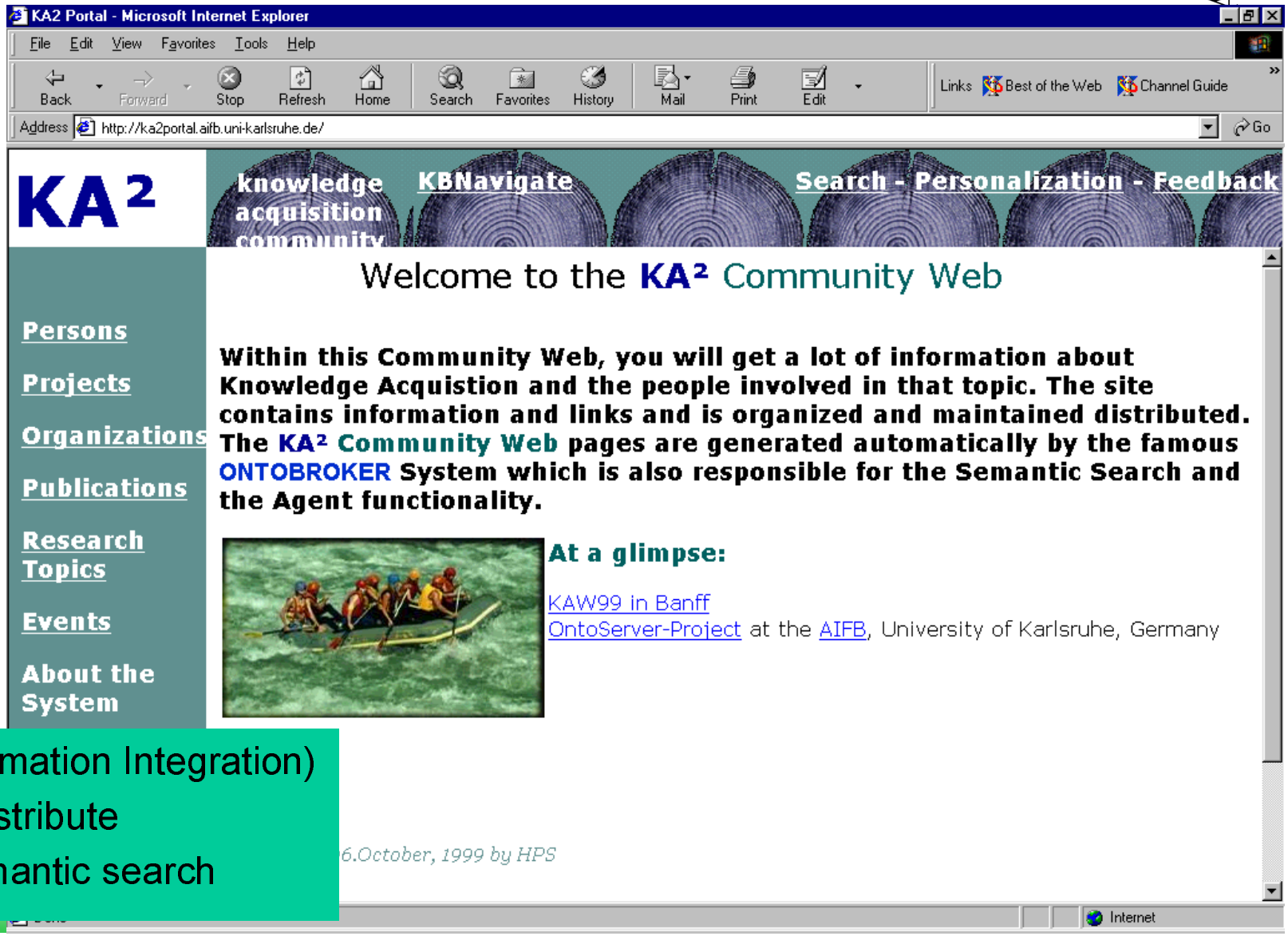
Our Ontology-based OM



3.4 Ontology-based Knowledge Management

Use

Semantic Community Web Portal - KA2Portal



Gather (Information Integration)

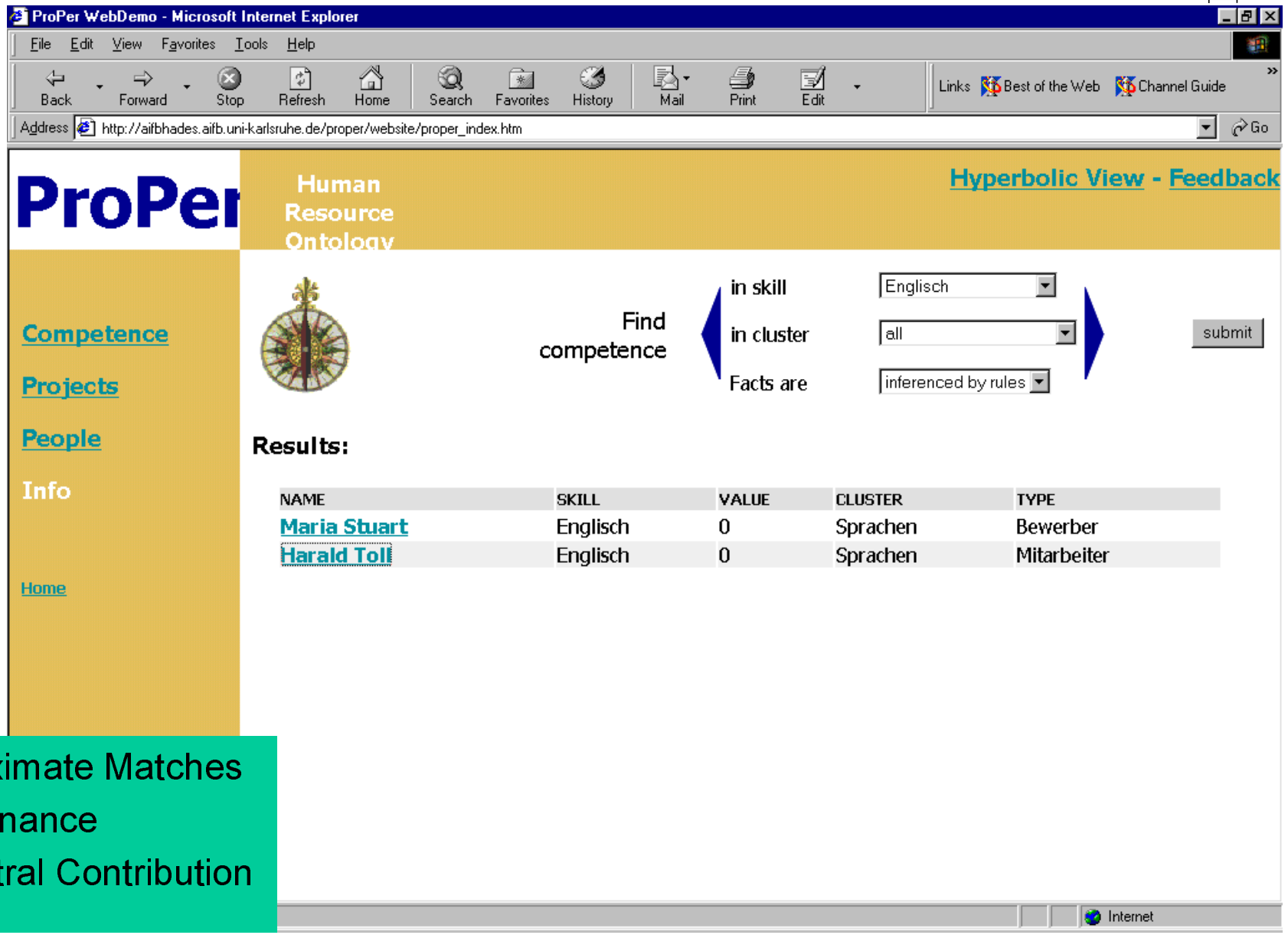
Structure, Distribute

Allow for semantic search

3.4 Ontology-based Knowledge Management

Use

Management of Skill Knowledge - OntoProPer



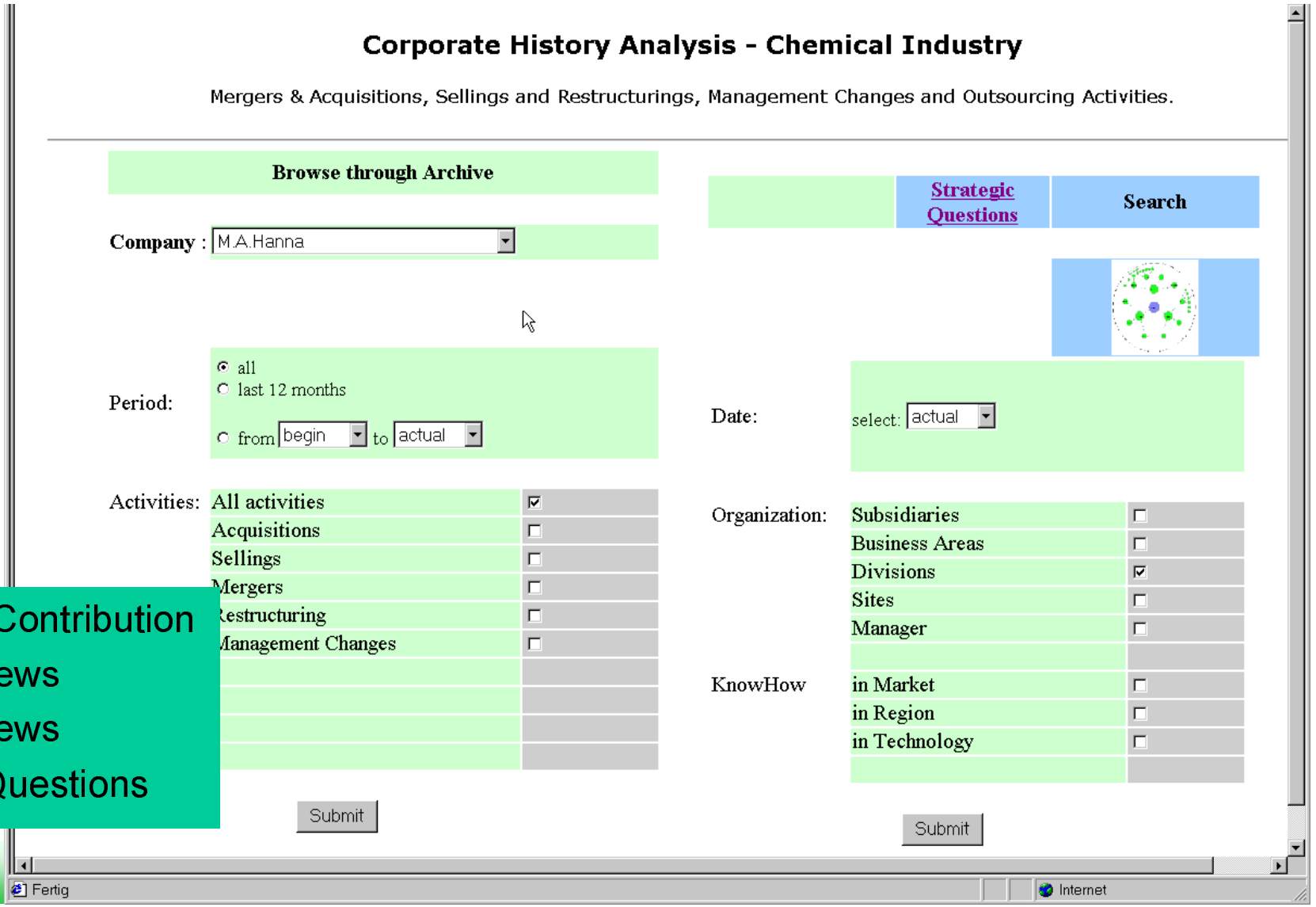
Approximate Matches
Maintenance
Decentral Contribution

3.4 Ontology-based Knowledge Management

Use

Management of Corporate Research - CHAR

AIFB



Decentral Contribution
Multiple Views
Derived Views
Strategic Questions

3.4 Ontology-based Knowledge Management

Use

Management of Corporate Research - CHAR

AIFB

Business Areas and Divisions of M.A.Hanna at 01.10.93

Division	Texapol Corporation
Division	Allied Color
Division	Avecor
Division	PMS Consolidated
Division	Plasticos Polisol
Division	Southwest Chemical
Division	Global Processing Corp.
Division	Burton Rubber
Division	Day International
Division	Bruck Plastics
Division	Gulf Colour
Division	Synthecolor
Division	Fiberchem
Division	Wilson Colorants
Division	Plastics Distribution Corp.
BusinessArea	Plastic Compounding
	rants
	ber Compounding
	n Distribution
	es Distribution
	neered Materials Group
	nial Rubber
	Hanna de Mexico
	Hanna Color
	Hanna Rubber
	Hanna Resin Distribution

Decentral Contribution
Multiple Views
Derived Views
Strategic Questions

Business Areas and Divisions of M.A.Hanna at 01.04.97

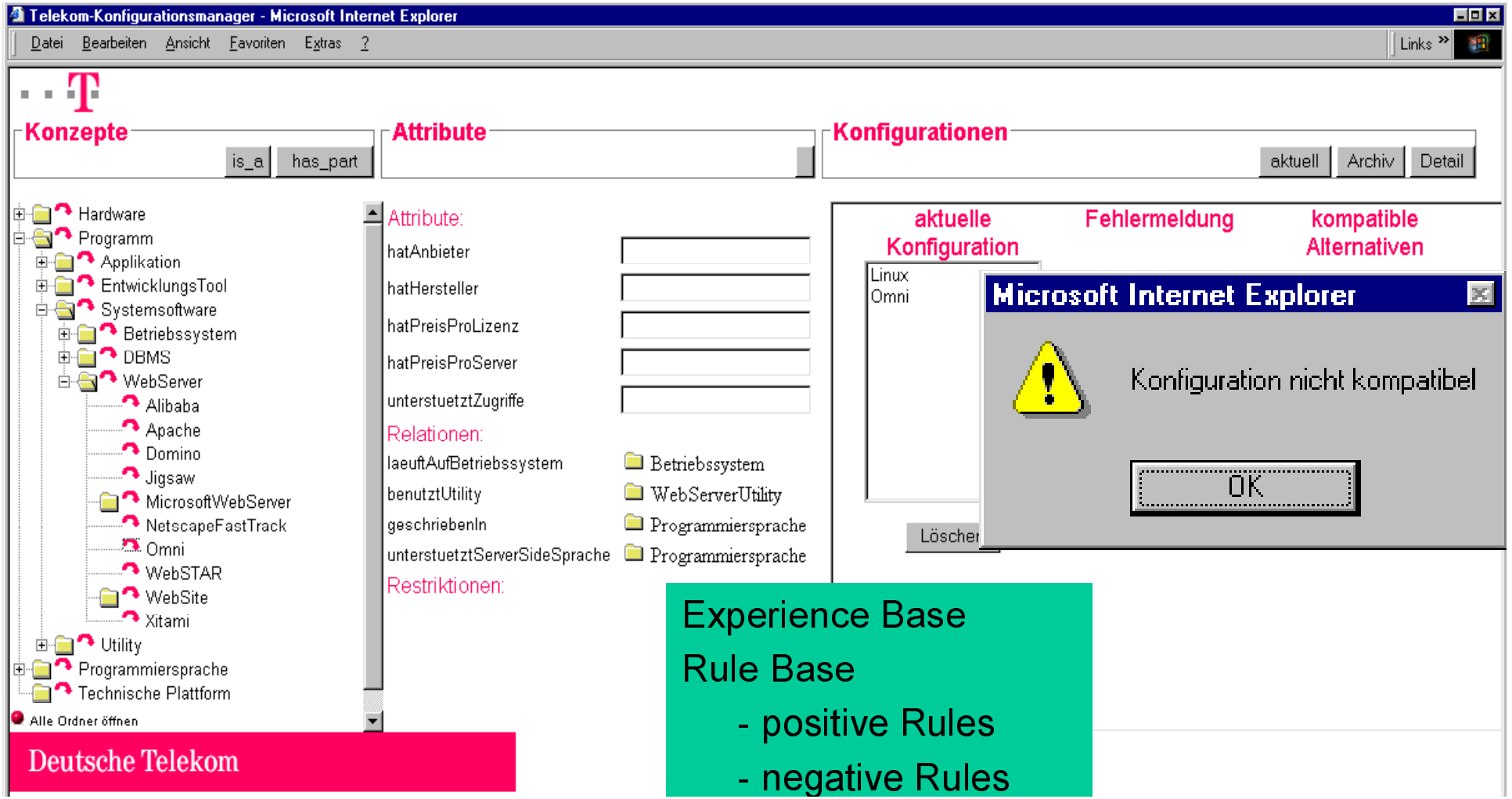
Division	Plasticos Polisol	
Division	EnviroCare Compounds (ECC)	
Division	North Coast Compounding	
Division	Day International	
Division	Bruck Plastics	
Division	Gulf Colour	
Division	Synthecolor	
Division	Fiberchem	
Division	Plastics Distribution Corp.	
BusinessArea	Plastic Compounding	
	Division	Compounding Technology, Inc. (CTi)
	Division	Southwest Chemical
	Division	Bergmann GmbH
BusinessArea	Colorants	
	Division	Victor International
	Division	Wilson Colorants
BusinessArea	Rubber Compounding	
BusinessArea	Resin Distribution	
BusinessArea	Shapes Distribution	
Division	Engineered Materials Group	
	Division	Texapol Corporation
Division	M.A.Hanna de Mexico	
Division	M.A.Hanna Color	
	Division	Allied Color
	Division	Avecor
	Division	PMS Consolidated

3.4 Ontology-based Knowledge Management

Use

Management of Configuration Knowledge - ONKO

AIFB

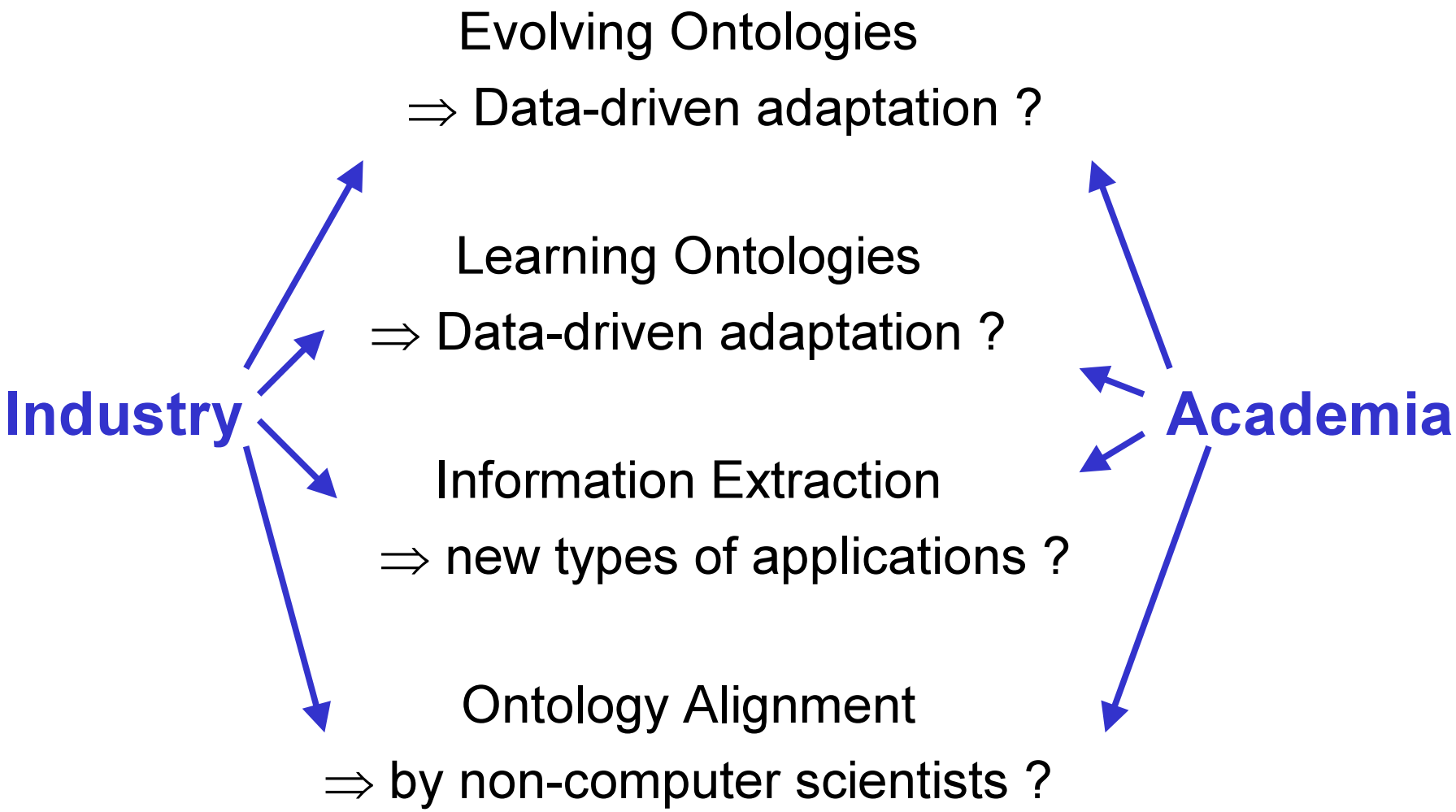


Experience Base

Rule Base

- positive Rules
- negative Rules

Open Topics


Evolving Ontologies
⇒ Data-driven adaptation ?
Learning Ontologies
⇒ Data-driven adaptation ?
Industry
Academia
Information Extraction
⇒ new types of applications ?
Ontology Alignment
⇒ by non-computer scientists ?

Wrap-up for Onto-based OMs

Don’t forget the overall process!

From containers to contents!

AIFB

Meta-Data based Knowledge Management

Objective

Tight integration:

In Knowledge Management

What is Meta-Data?

- Structured data about data
- provides basic information about resources (e.g documents in a company)
  - such as the author of a work, the date of creation, links to any related works
  - enables more effective search
- Example: meta-data is the card index catalogue in a library (meta-data about books)
- Meta-Data needs Standards
  - Which fields are available?
  - What to fill in?
  - Interoperability
  - Tool Development

Meta-Data Standard: Dublin Core

- (Simple!) Ontology for Metadata: Dublin Core
- 15 element metadata set
- resource discovery
- Web-based document-like objects
- emphasis on semantics
- widespread consensus
- several syntaxes
- http://purl.oclc.org/dc

AIFB

Meta-Data Standard: Dublin Core

- Title
- Creator
- Subject
- Description
- Publisher
- Contributor
- Date
- Type
- Format
- Identifier
- Source
- Language
- Relation
- Coverage
- Rights

Meta-Data Representation on the Inter-/Intranet

- Standards for Meta-Data Representation enable:
  - Interoperability
  - cost-effective development of tools
- W3C Recommendation for representing Dublin Core:
  - Resource Description FrameWork [O. Lassila, 1999]
- RDF is able to represent more than Dublin Core
  - World Wide accepted Ontology Representation Standard (?)

Introduction to RDF

- RDF (Resource Description Framework)
  - Beyond Machine readable to *Machine understandable*
- RDF unites a wide variety of stakeholders:
  - Digital librarians, content-raters, privacy advocates, B2B industries, AI...
  - Significant (but less than XML) industrial momentum, lead by W3C
- RDF consists of two parts
  - RDF Model (a set of triples)
  - RDF Syntax (different XML serialization syntaxes)
- RDF Schema for definition of Vocabularies (simple Ontologies) for RDF (and in RDF)

Design Goal for RDF

1.) Knowledge on Networks is distributed -
link Knowledge

2.) There is no universal truth and many opinions (Knowledge on the Web is biased)
it must be possible to dispute statements

3.) Many different user communities (one can't know what they want to represent)
Extensibility and Simplicity

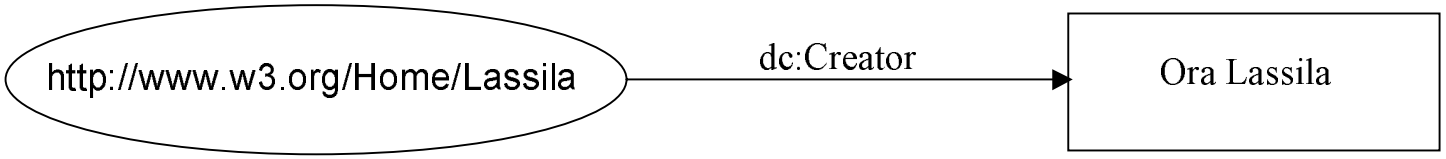
AIFB

RDF Data Model

- **Resources**
  - A resource is a thing you talk about (can reference)
  - Resources have URI's
  - RDF definitions are itself Resources (linkage)
- **Properties**
  - slots, define relationship to other resources or atomic values
- **Statements**
  - "Resource has Property with Value"
  - (Values can be resources or atomic XML data)
- **Similar to Frame Systems**

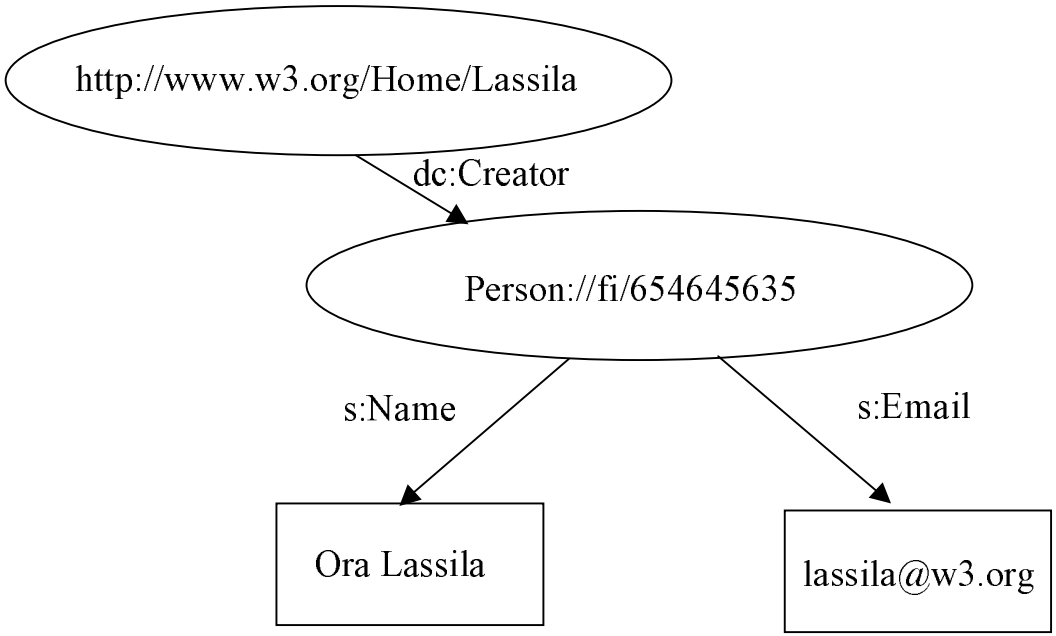
A Simple Example

- **Statement**
  - "Ora Lassila is the creator of the resource http://www.w3.org/Home/Lassila"
- **Structure**
  - Resource (subject) http://www.w3.org/Home/Lassila
  - Property (predicate) http://purl.org/dc/elements/1.1/creator
  - Value (object) "Ora Lassila"
- **Directed graph**



Another Example

- To add properties to Creator, point through an intermediate Resource.

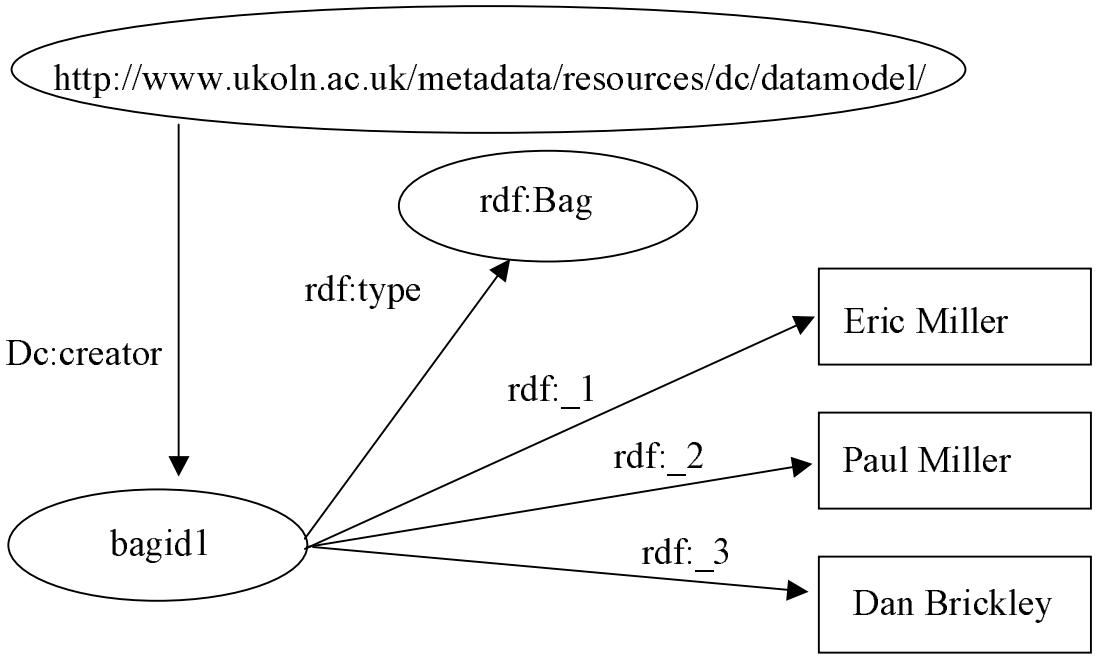


Collection Containers

- Multiple occurrences of the same PropertyType don't establish a relation between the values
  - Employee Miller knows Java, Perl and Python
  - This talk requires knowledge in Java or Python
  - (Molina, Widom, Ullman) are working at the project X
- RDF defines three special Resources:
  - **Bag** unordered values rdf:Bag
  - **Sequence** ordered values rdf:Seq
  - **Alternative** single value rdf:Alt
    - **Core RDF does not enforce 'set' semantics amongst values**

Example: **Bag**

- *The creators of the document* http://www.ukoln.ac.uk/metadata/resources/dc/datamodel *are Eric Miller, Paul Miller, and Dan Brickley.*



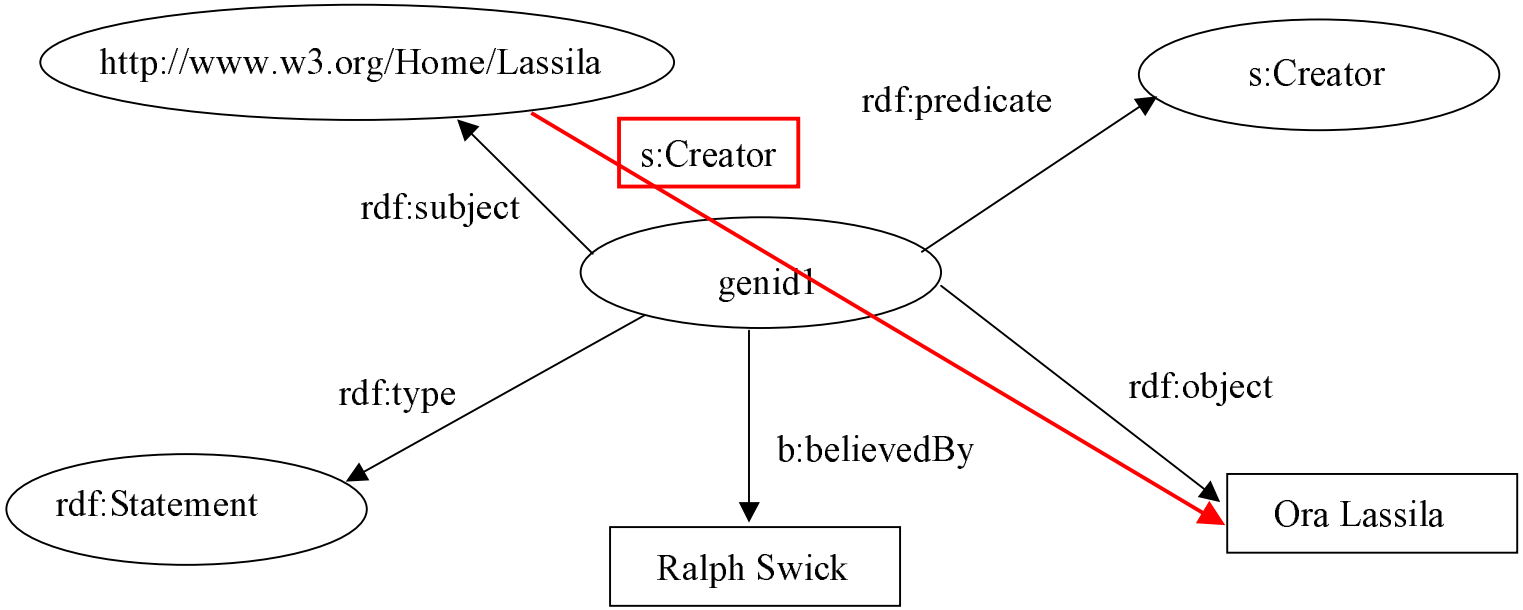
AIFB

Statements about Statements

- Making statements about *statements* requires a process for transforming statements into Resources
  - **subject** the original referent
  - **predicate** the original property type
  - **object** the original value
  - **type** rdf:Statement

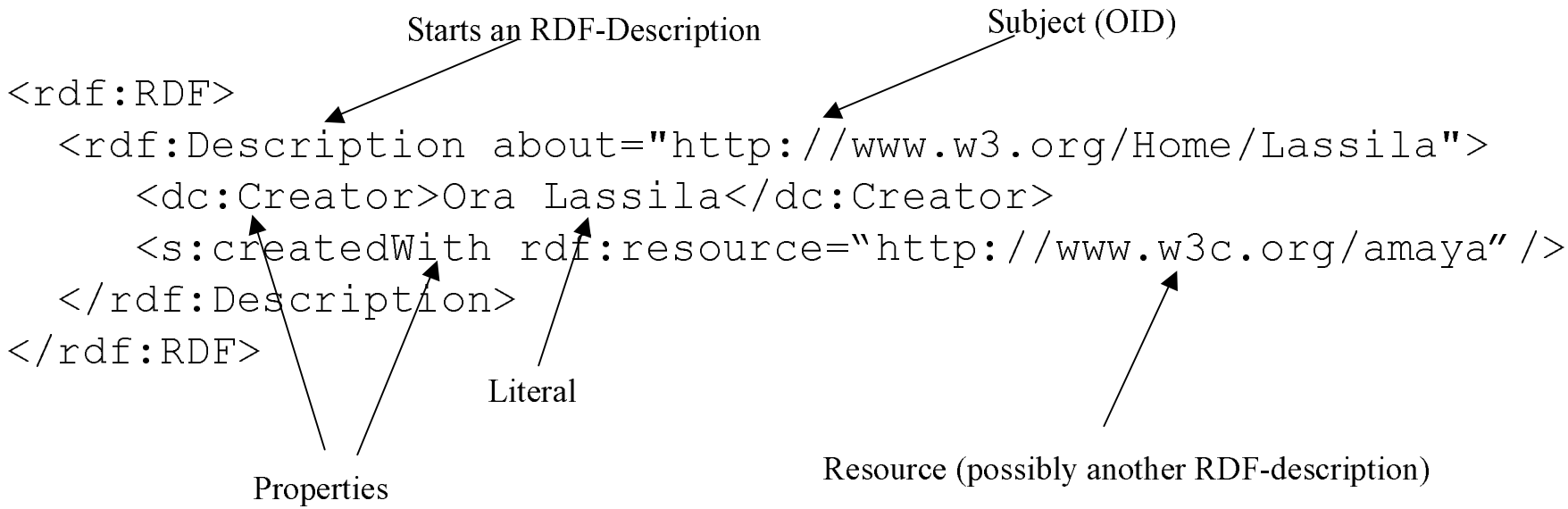
Example: **Reification**

- *Ralph Swick believes that*
  - *the creator of the resource http://www.w3.org/Home/Lassila is Ora Lassila*

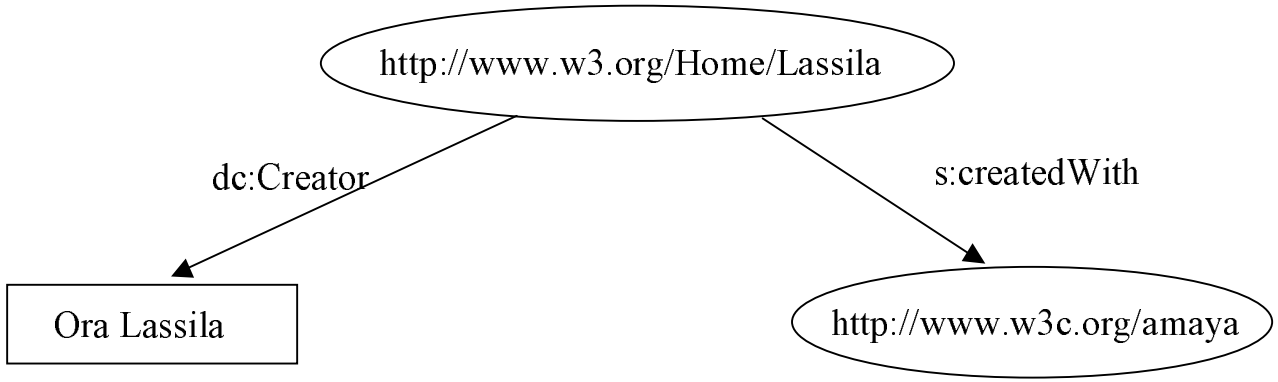


RDF Syntax I

- Datamodel does not enforce particular syntax
- Specification suggests many different syntaxes based on XML
- General form (Namespace-definitions are omitted):

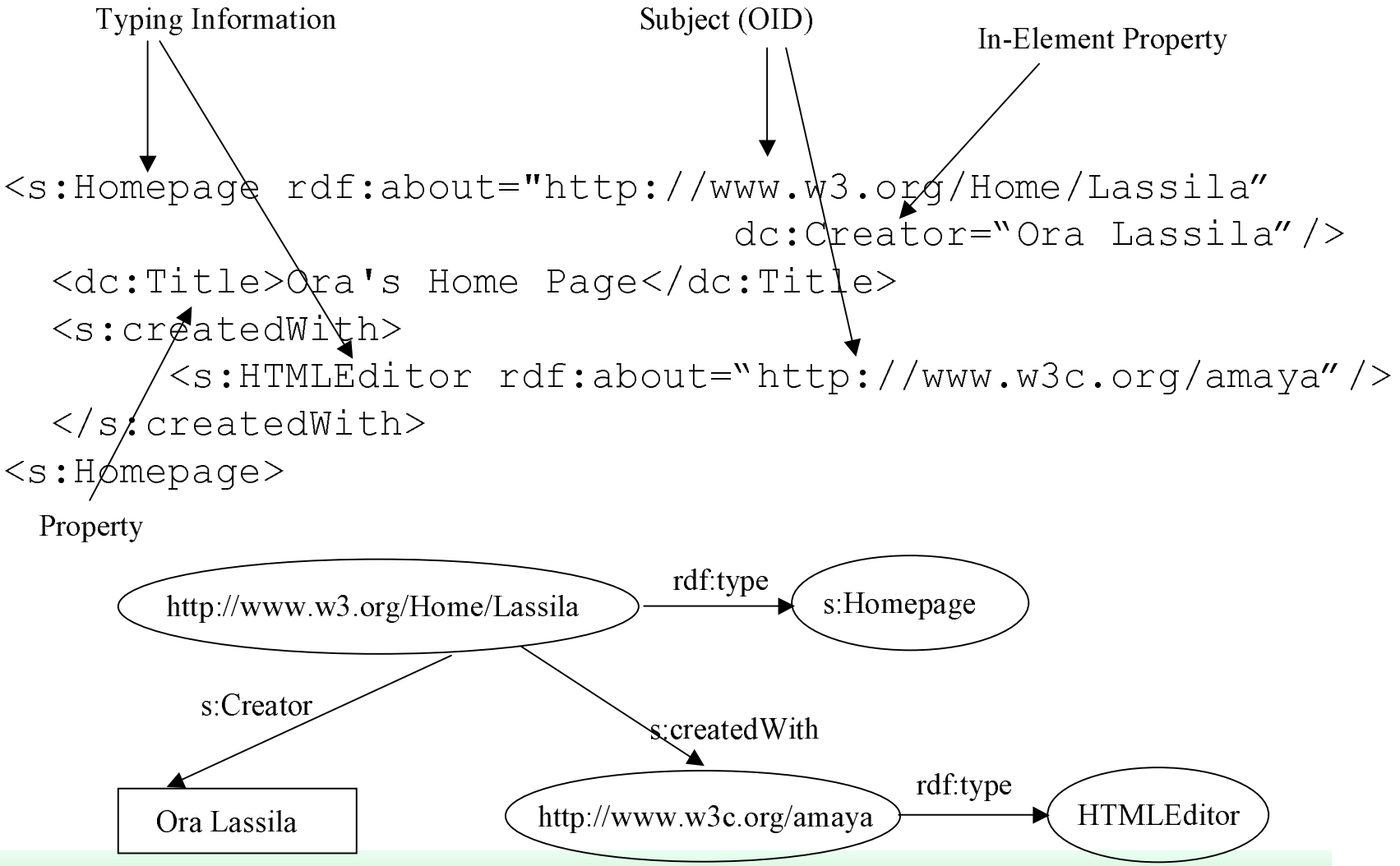


Resulting Graph




```
<rdf:RDF>
  <rdf:Description about="http://www.w3.org/Home/Lassila">
     <dc:Creator>Ora Lassila</dc:Creator>
     <s:createdWith rdf:resource="http://www.w3c.org/amaya"/>
  </rdf:Description>
</rdf:RDF>
```

RDF Syntax II: Syntactic Varieties



RDF Schema (RDFS)

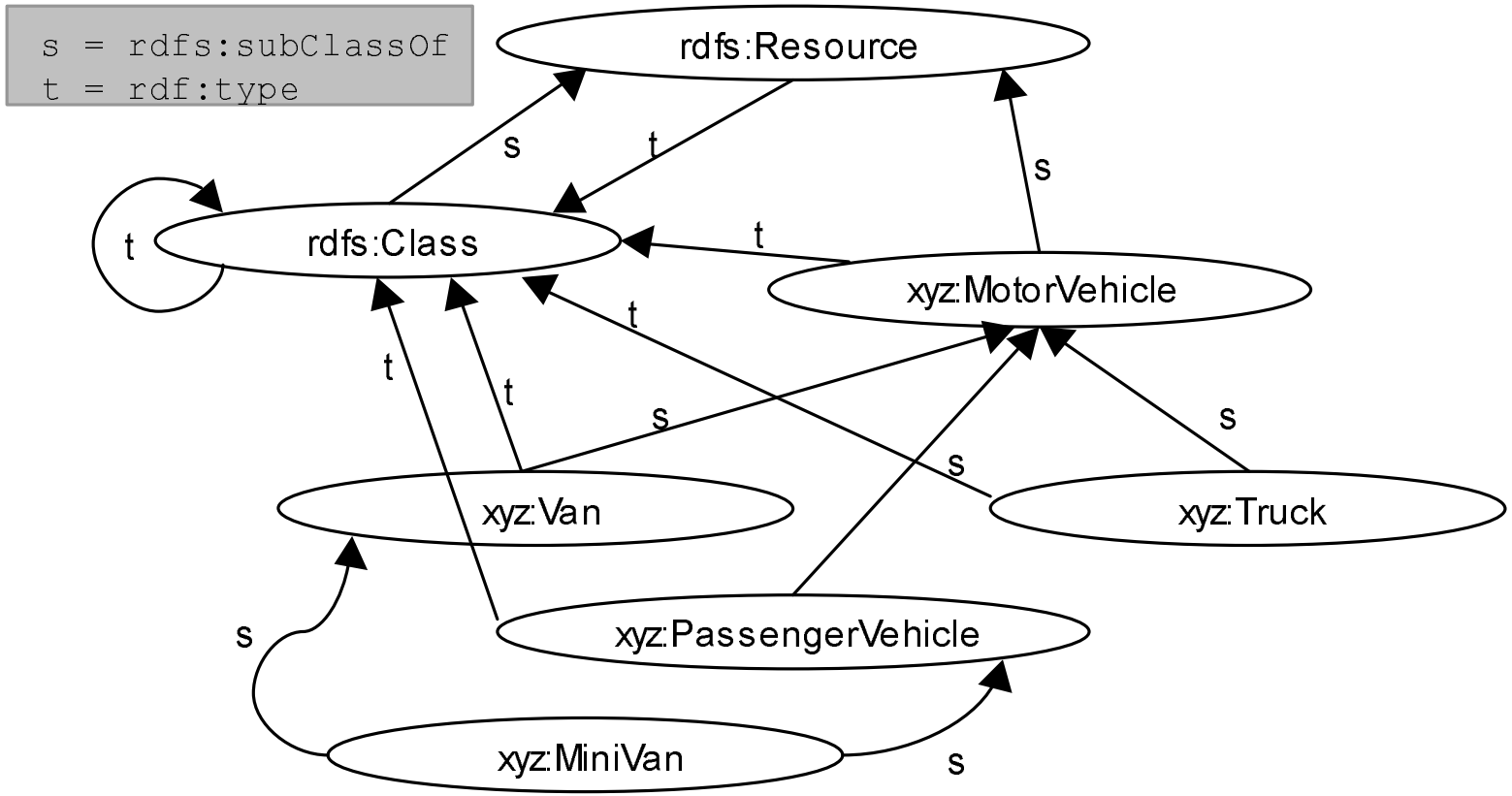
- RDF just defines the datamodel
- Need for definition of vocabularies for the datamodel - an Ontology Language!
- RDF schemas are Web resources (and have URIs) and can be described using RDF

[D. Brickley, 2000]

Most Important Modeling Primitives

- Core Classes
  - Root-Class `rdfs:Resource`
  - MetaClass `rdfs:Class`
  - Literals `rdfs:Literal`
- `rdfs:subclassOf`-property
- Inherited from RDF: properties (slots)
- `rdfs:domain` & `rdfs:range`
- `rdfs:label`, `rdfs:comment`, etc.
- Inherited from RDF: InstanceOf (`rdf:type`)

RDF-Schema: Example



Dublin Core in RDF-Schema


```
<? xml version='1.0'?>
 <rdf:RDF
    xmlns:rdf="http://www.w3.org/TR/REC-rdf-syntax#"
    xmlns:rdfs="http://www.w3.org/TR/WD-rdf-schema#"
    xmlns:dc="">

<rdf:Description ID=”Creator">
   <rdf:type rdf:resource="http://www.w3.org/TR/REC-rdf-syntax#Property"/>
   <rdfs:label>Author/Creator</rdfs:label>
   <rdfs:comment>The person or organization primarily responsible for
   creating the intellectual content of the resource. For example,
   authors in the case of written documents, artists, photographers, or
   illustrators in the case of visual resources.</rdfs:comment>
 </rdf:Description>
. . . . . .
</rdf:RDF>
```

Meta-Data & RDF-Conclusions

- The larger the document corpus the more important are meta-data
- RDF is a very general (purpose) representation format
- Basis for the “Semantic Web” (eg. for automated Information Agents)
- The first widely deployed Knowledge Representation Language (?)

Topic Navigation Maps

- Aiming at User Support for classifying and navigating large corpora of resources
- Topic Navigation Maps generalize:
  - indexes
  - glossaries
  - thesauri
  - catalogs
  - cross-references
- Allow browsing as well as querying for Information
- ISO Standard (ISO13250) (defined by the SGML/XML Community)

Topic Navigation Maps in a Nutshell I

- Topics (e.g. Germany)
  - have a topic-type (e.g. Country)
  - at least one base-name (Germany)
  - a multiple display names (e.g. Federal Republic of Germany)
  - have occurrences (in external resources)

    (eg. In CIA-World Fact Book: http://www.odci.gov/cia/publications/factbook/gm.html)
    - occurrences have a role: eg. “MENTION” (again a Topic)

Topic Navigation Maps in a Nutshell II

- Topics can related to each other via associations (eg. “Rau is_head_of Germany”)
- Association-Types (eg. “is_head_of”) are again a topics.
- Association-Roles define the role of a topic in a association (“President” for “Rau”, and “Country” for “Germany”). Association-Roles are again topics.
- Any assignment of a characteristic has a topic has a scope (eg. “Rau is_head_of Germany” has scope “1999-now”)
- Axioms (eg. Transitivity or Symmetry) are considered useful, but not part of the ISO-Standard

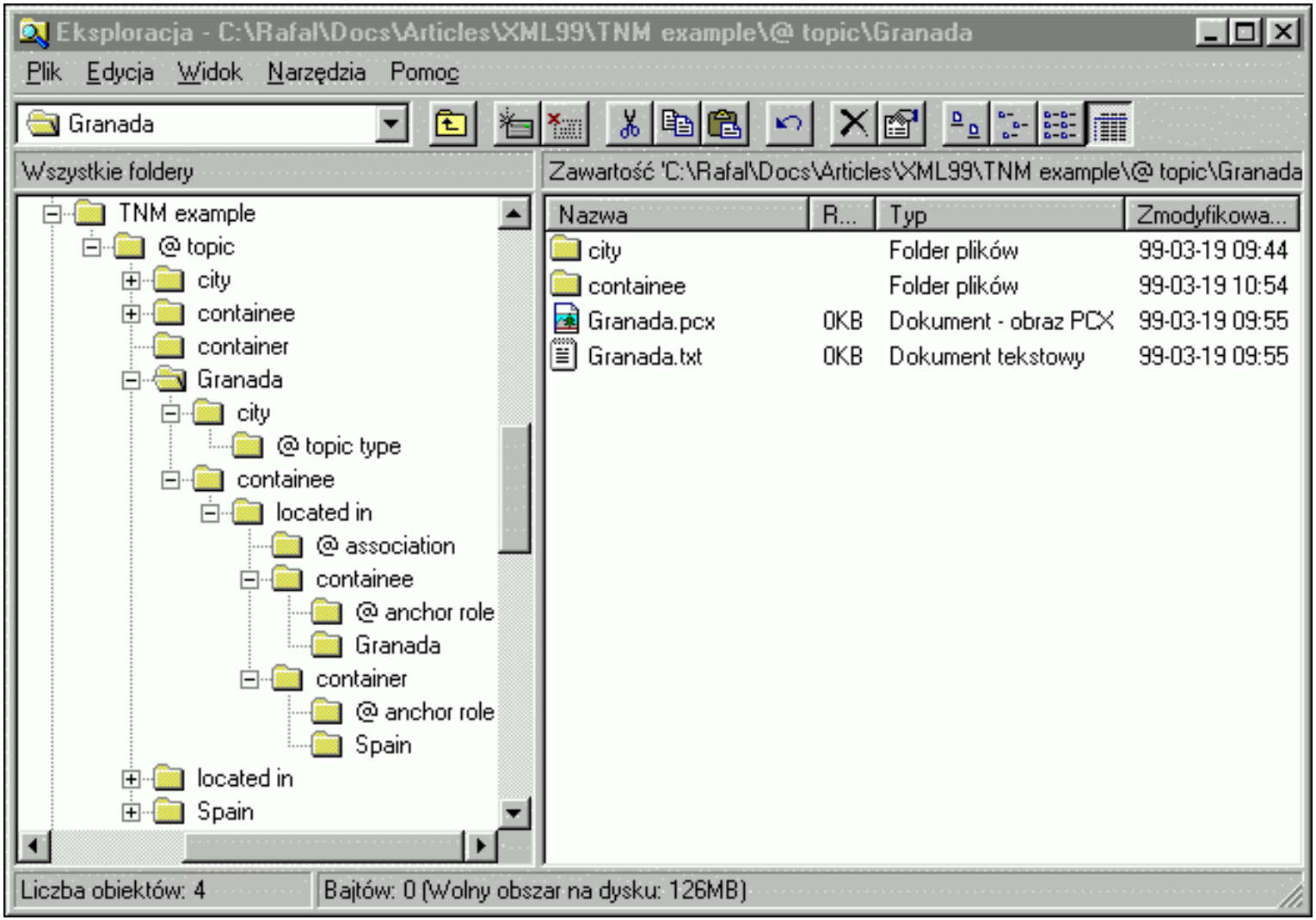
AIFB

Representation

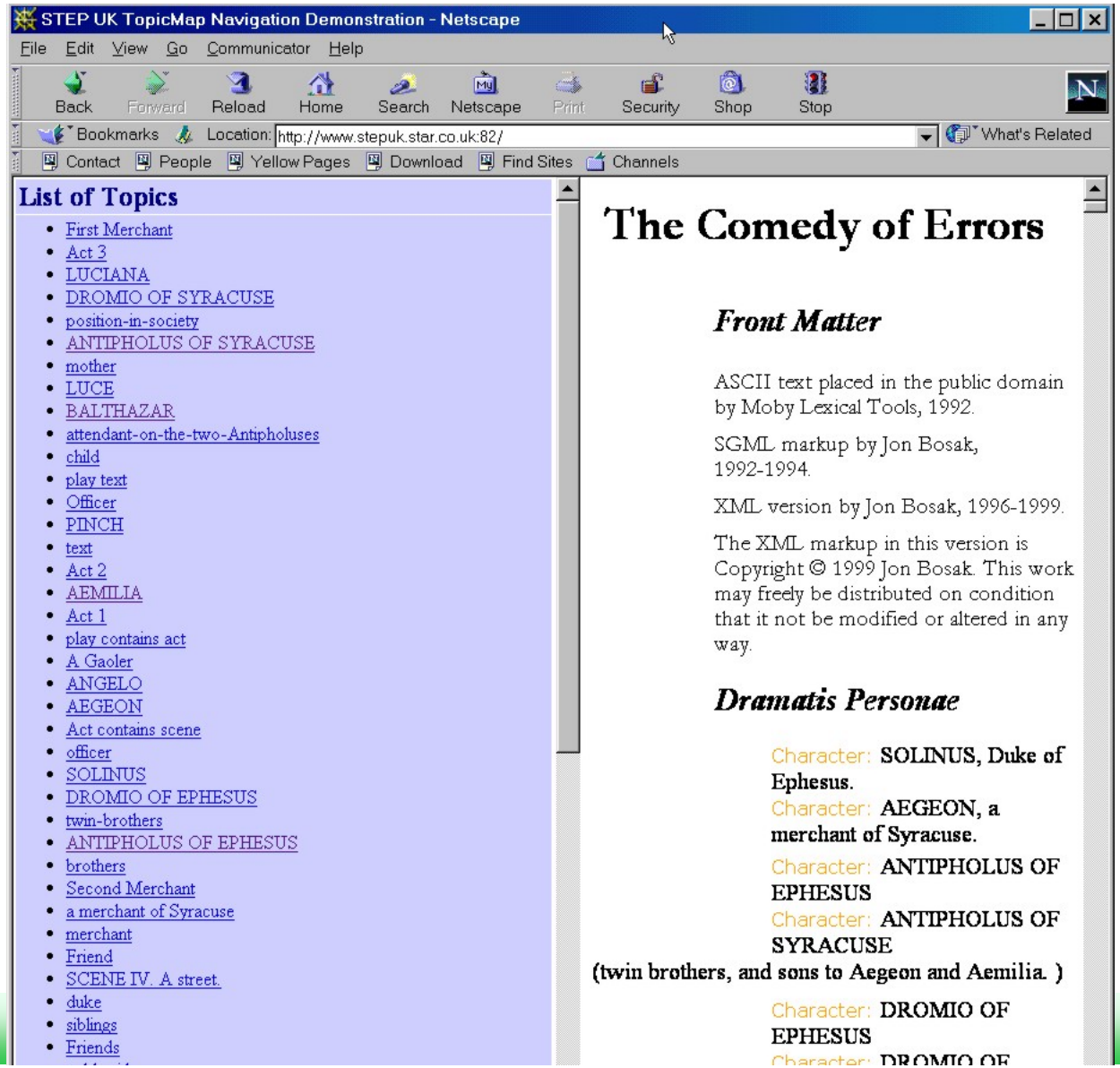
- ISO13250 defines a XML-Meta-DTD
- simplified version:

```
<topicmap>
<topic id="t1" types="COUNTRY">
  <topname>
  <basename>Germany</basename>
  <dispname>Federal Republic of Germany</dispname>
  <sortname>GERMANY</sortname>
  </topname>
  <occurs>
     <locator role="MENTION"
              href="http://www.odci.gov/cia/publications/factbook/gm.html " />
  </occurs>
</topic>
....
<assoc types="IS_HEAD">
  <assocrl role="PRESIDENT" href="#t123" />
  <assocrl role="COUNTRY" href="#t1" />
</assoc>
<topicmap>
```

Example of an Browsing Interface

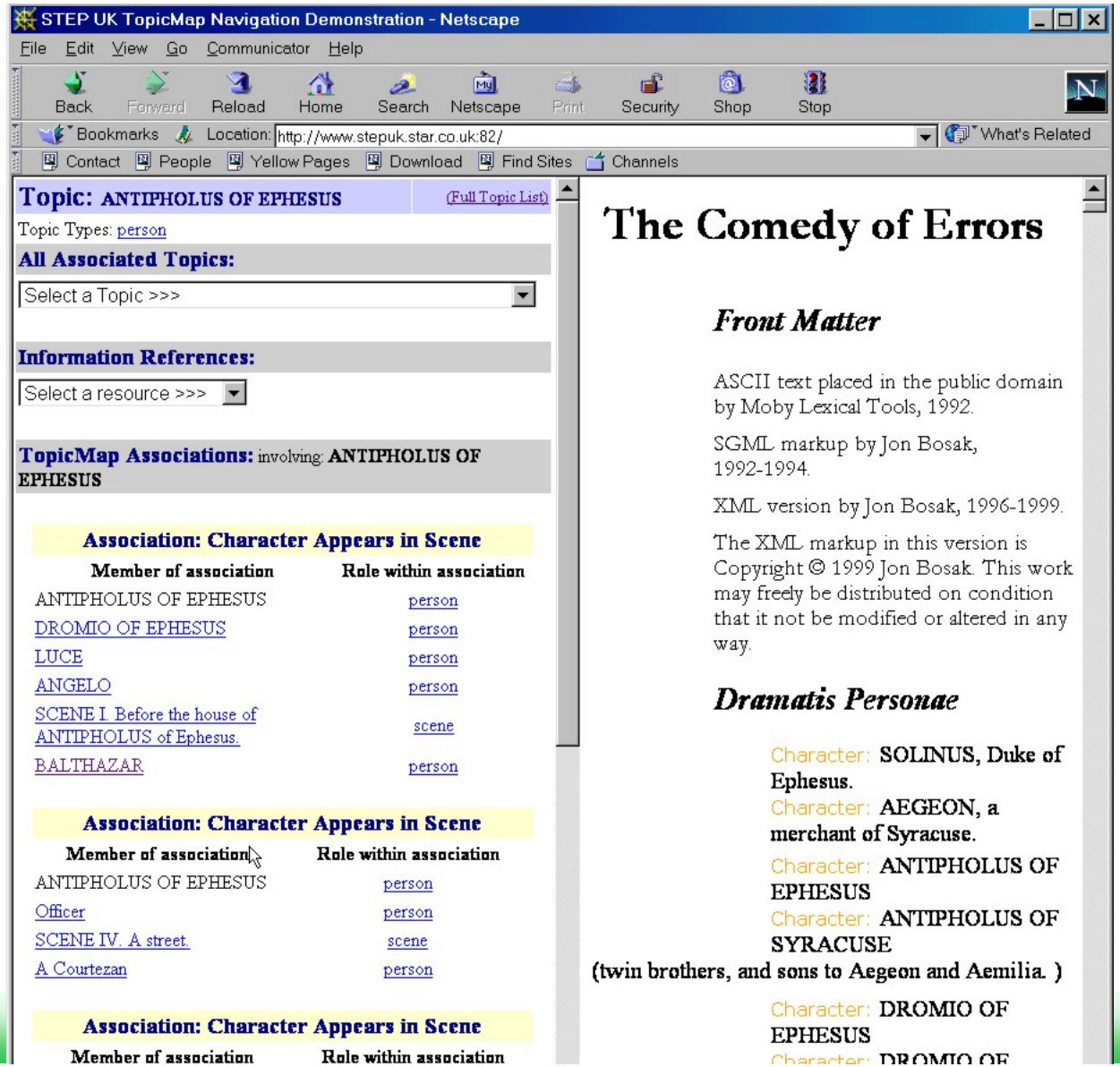


Demonstration: The Comedy of Errors



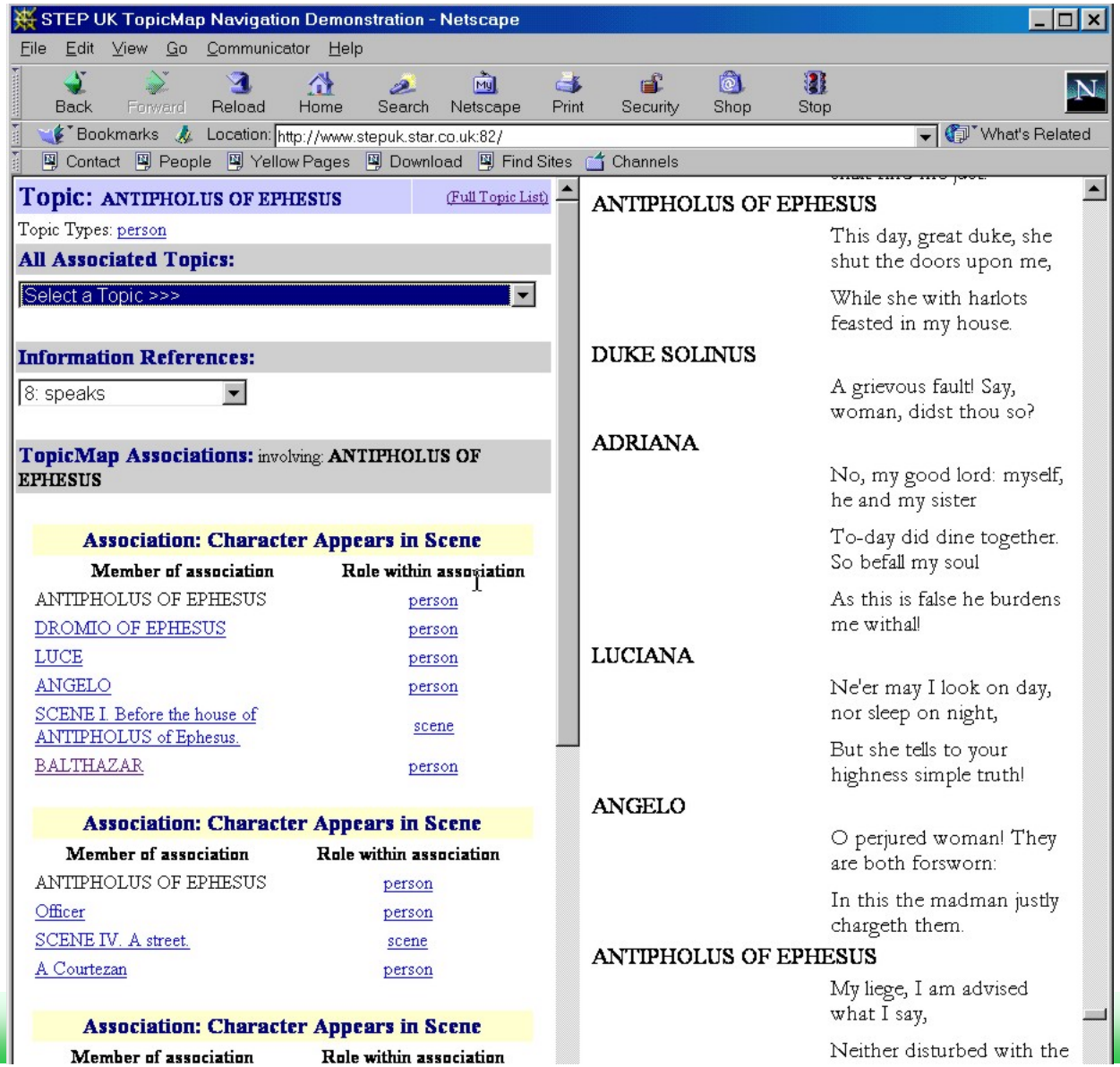
http://www.stepuk.star.co.uk:82/

The Comedy of Errors: Select Topic



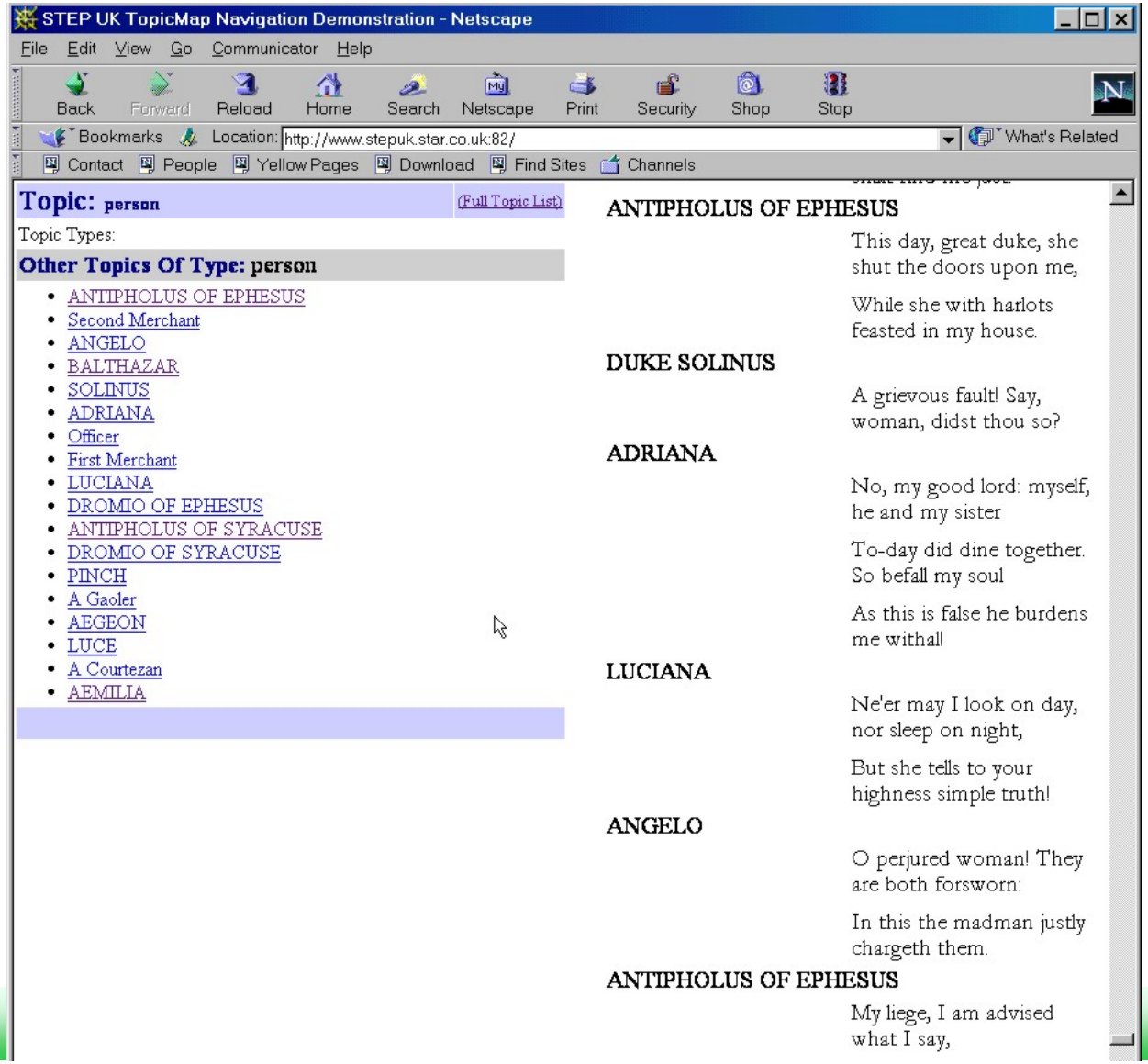
http://www.stepuk.star.co.uk:82/

Browse to External Reference



http://www.stepuk.star.co.uk:82/

Browse to other Persons



http://www.stepuk.star.co.uk:82/

Conclusion

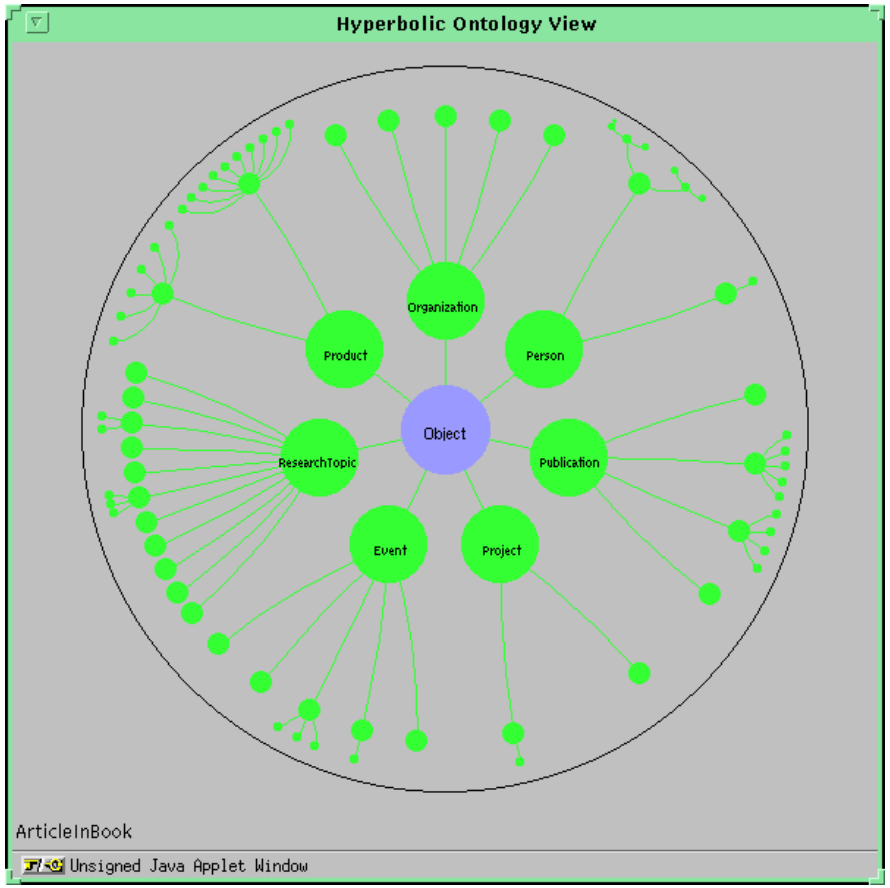
- Topic Navigation Maps define a simple Ontology Language (created by a non-AI community)
- Aiming at User Interaction (Browsing/Querying)
- Similar to RDF
- Need for Extensions (eg. axioms, background knowledge)

AIFB

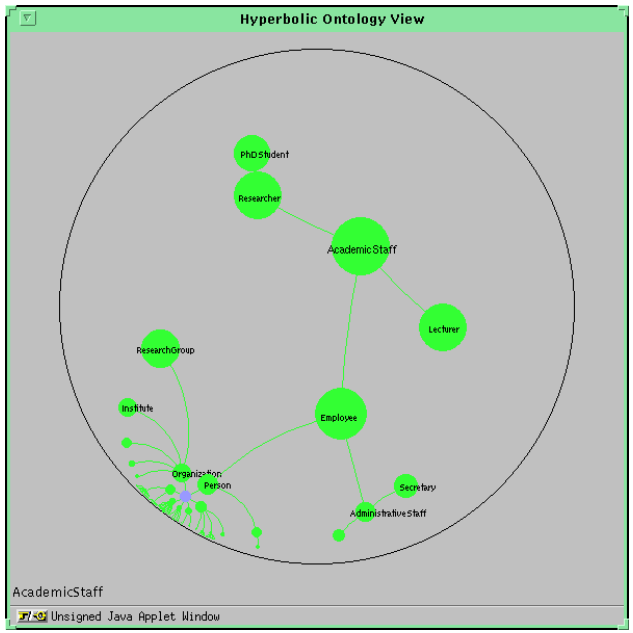
Visualization in Knowledge Management

- Knowledge Management involves browsing large complex data sets (eg. Ontologies)
- Common Visualization Techniques are not sufficient
- “lost in the Ontology” syndrom

Visualizing Hierachies: 2D Hyperbolic Viewer

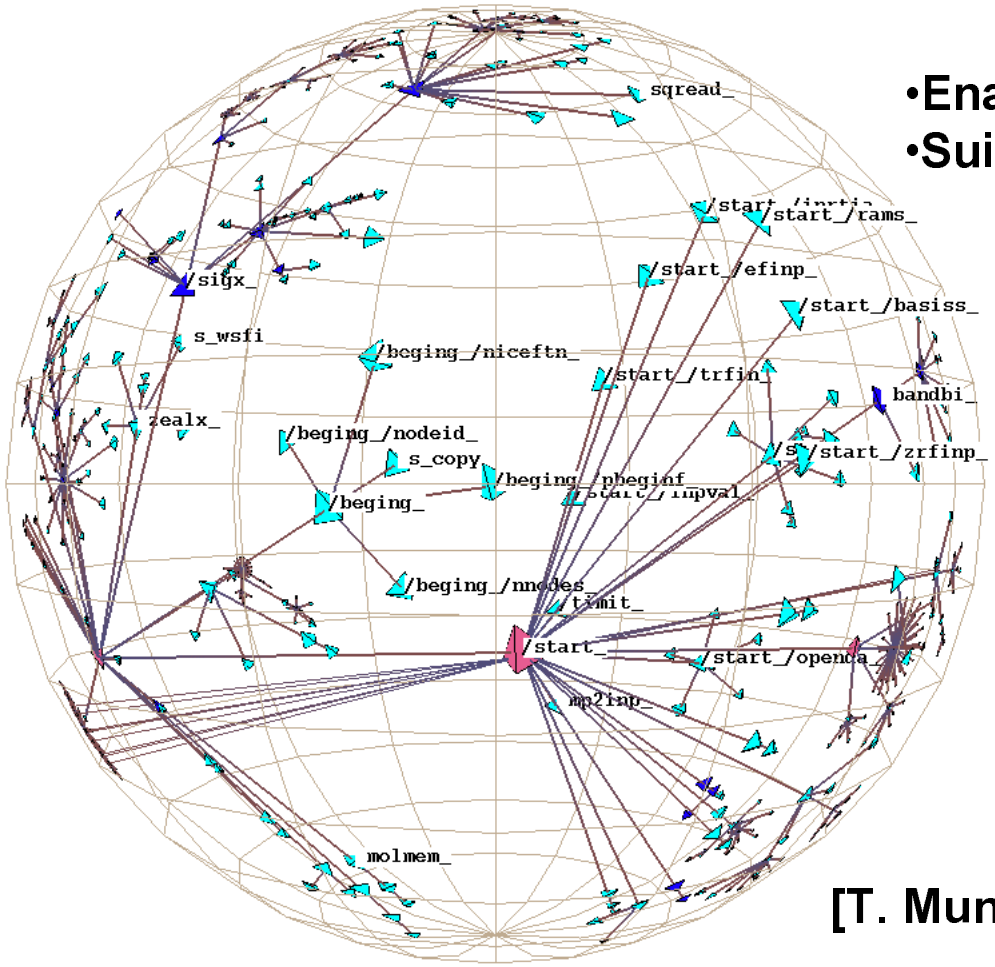


- "Focus & Context"
- Smooth Navigation



[J. Lamping 1996]

Visualizing Hierachies: 3D Hyperbolic View

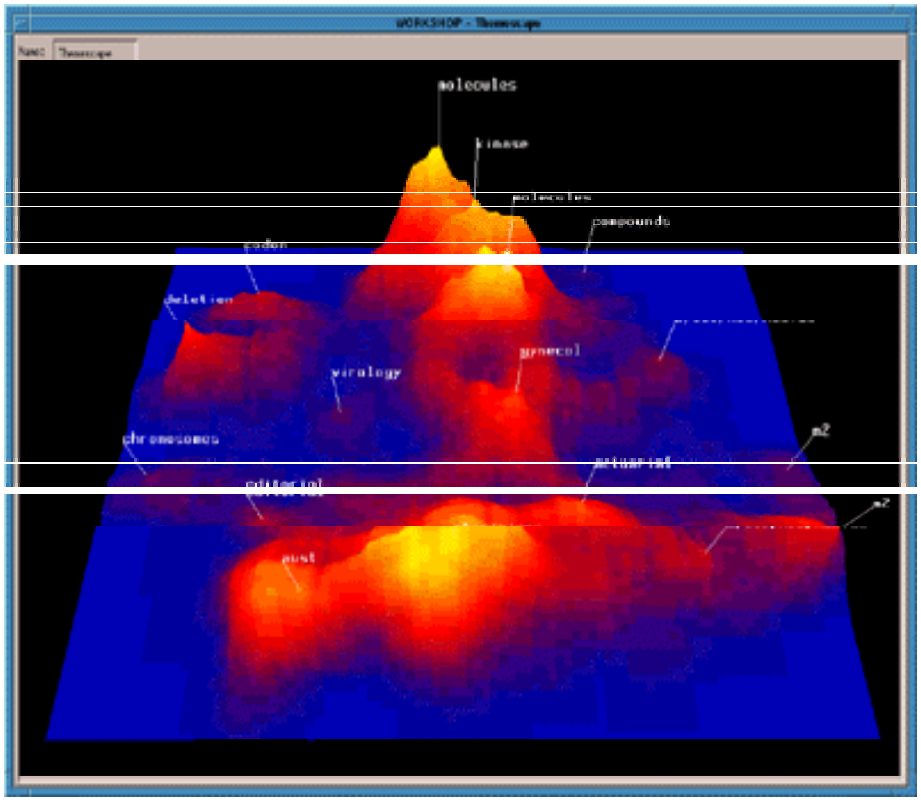


- Enables bigger graphs a 2D Viewer
- Suitable for Tree Structure

[T. Munzner 1998]

Document Visualization

- Visual Support for Navigation and Browsing in a Set of Document
- Topic distribution in a large document space:



•content abstraction and spatialization of the document

[J.A. Wise 1995]

AIFB

Document Visualization

- **Visual Support for Navigation and Browsing in a Set of Document**

- As close as possible to a real world-book

[S.K. Card et al. (1996)]

Knowledge Management Methodology:

CommonKADS

[Schreiber et al. 99]

AIFB

- Why Methodology?
  - Guidance for developing a KM System or Introducing a KM System in the Organization
  - Definition of Templates
  - Can be supported by Tools

4.0 The CommonKADS Methodology

- CommonKADS is a Knowledge Engineering Methodology (Methodology for developing Knowledge (-Based) Systems)

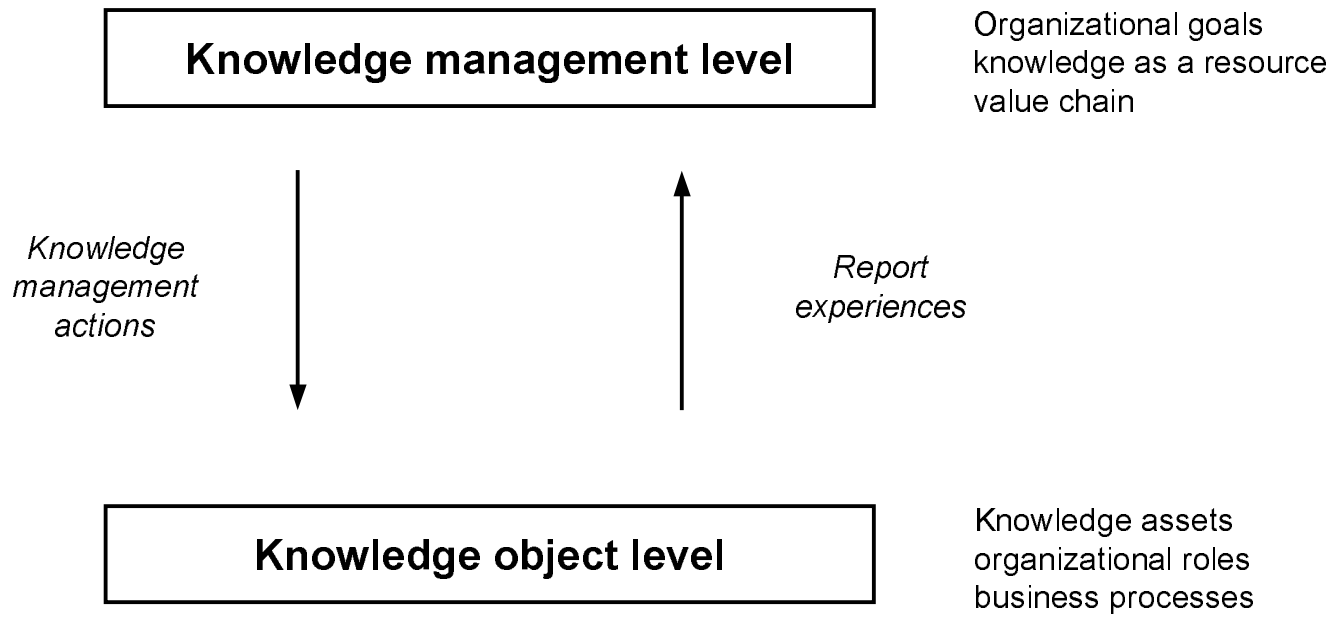
- Relationship between Knowledge Engineering and Knowledge Management

  – Knowledge Systems are Knowledge Management Methods and Tools of advanced Information

  – Embedding of Knowledge (-Based) Systems have Organizational and Human Resources context

- Presented parts of CommonKADS
  - The overall knowledge management framework as defined by CommonKADS
  - Those models of the CommonKADS model suite that are relevant for knowledge management
- Approach will be illustrated by a case study

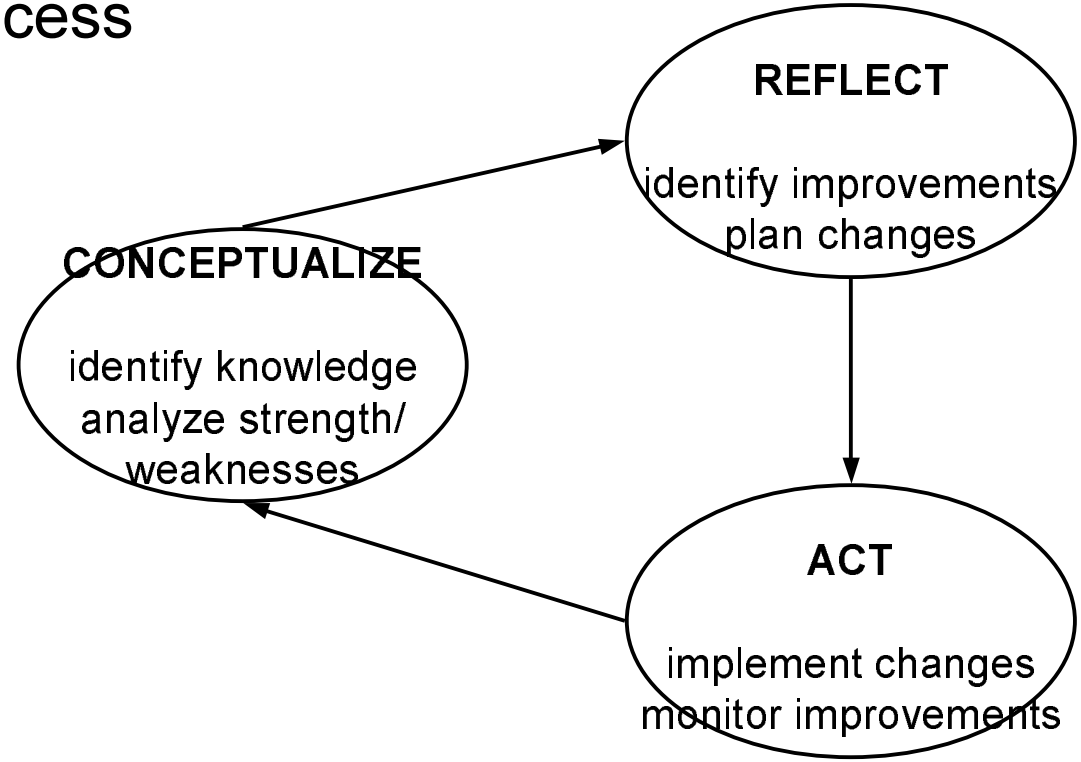
4.1 The Knowledge Managementg FrameWork



Knowledge management, like other management tasks, can be seen as a metalevel activity that acts on an object level.

4.1 The Knowledge Management Framework

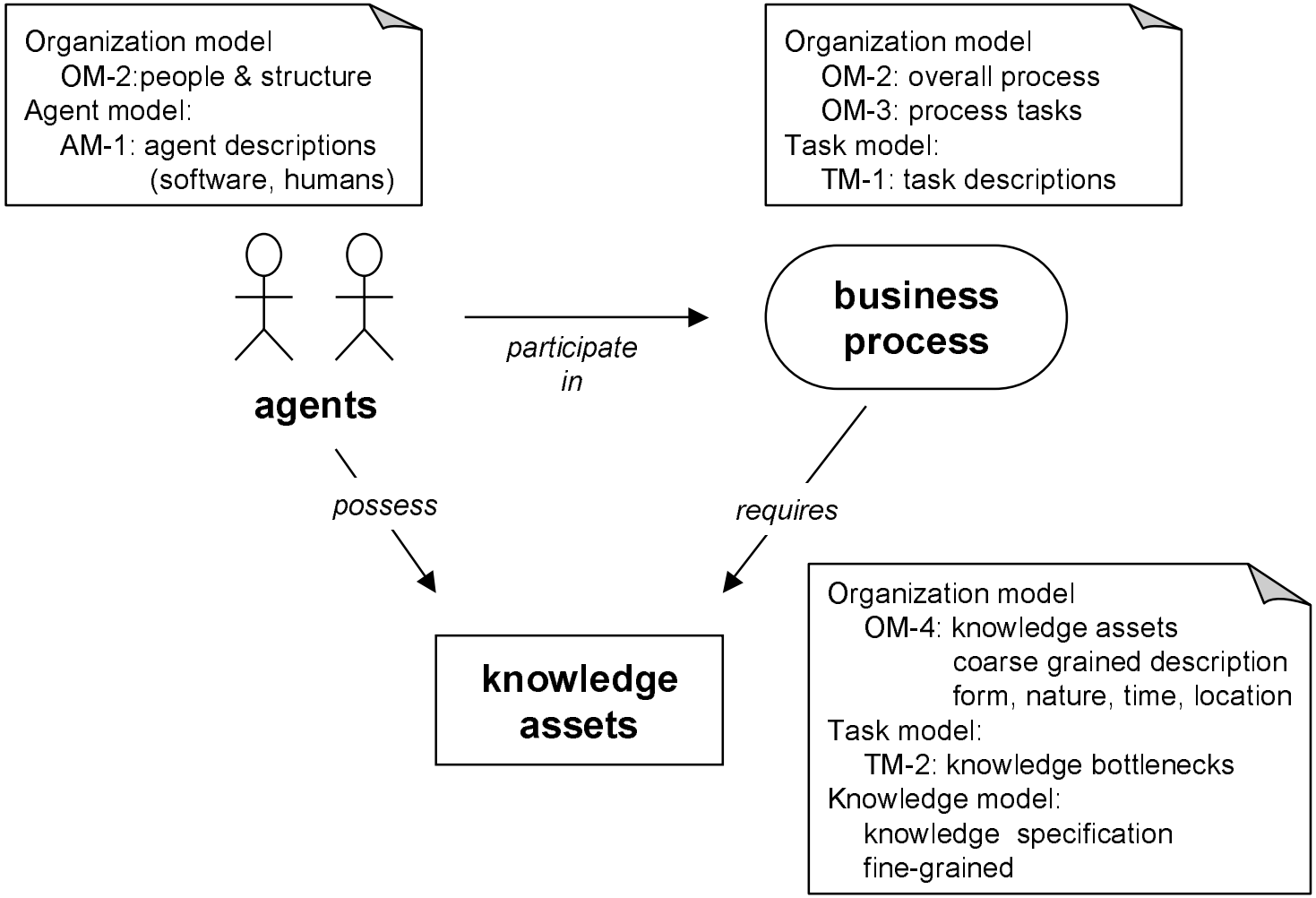
- Management level consists of three types of management activities, embedded in a cyclic process



Knowledge management consists of a cyclic execution of three main activities: conceptualize, reflect, and act.

- Knowledge object level
  - Knowledge object level is composed of
    - **Knowledge assets**
    - **Organizational roles**
    - **Business processes**
  - Some models of the CommonKADS model suite address the relevant aspects:
    - **Organization model**
    - **Agent model**
    - **Task model**

4.1 The Knowledge Management Framework



Knowledge-management actions are defined in terms of three objects: agents that possess knowledge assets and participate in the business process. The notes indicate which parts of the CommonKADS models describe these objects.

- The analysis of organizational and task aspects is divided in 2 phases:

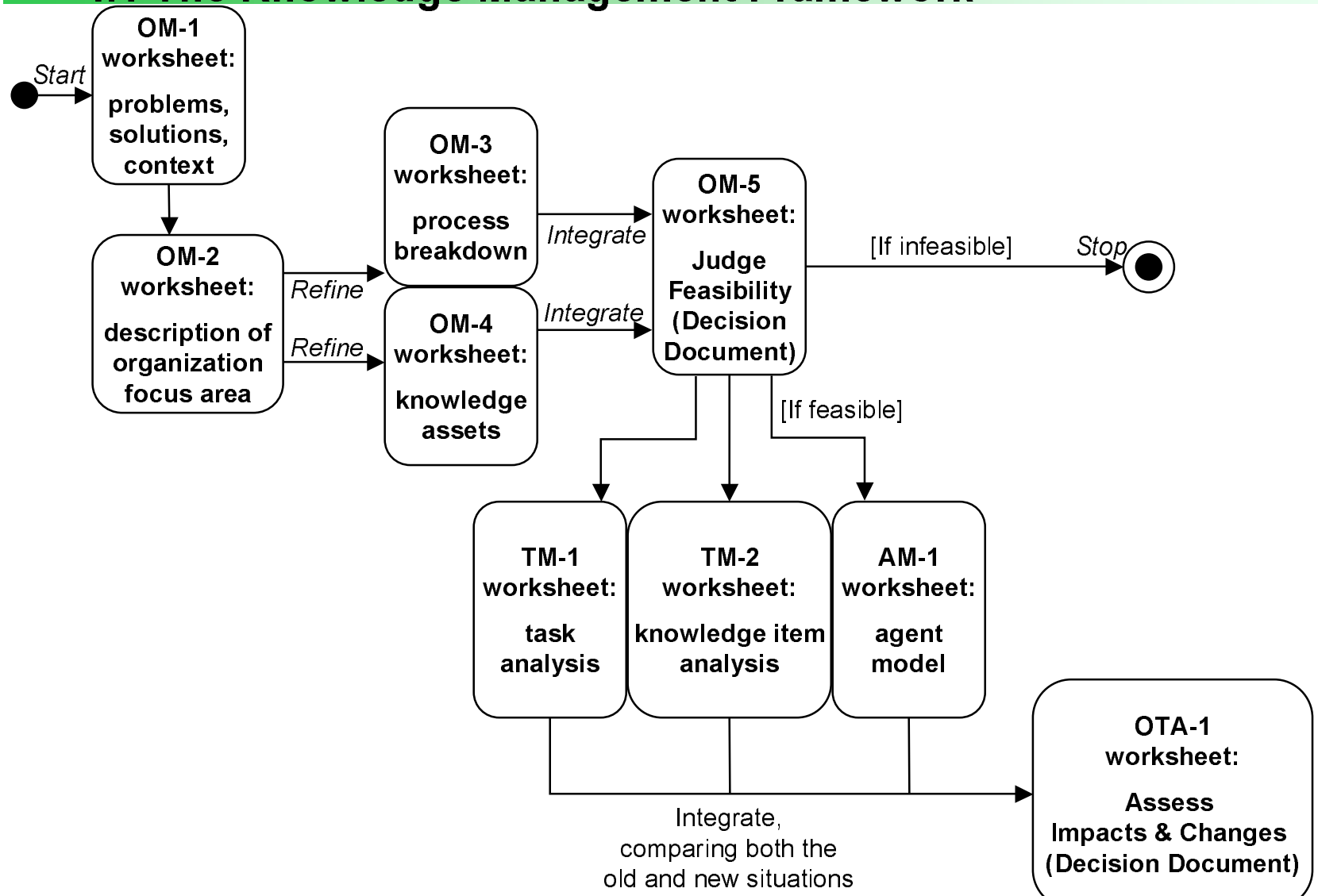
  Phase 1: Scoping and feasibility study

  - **Identify problem/opportunity areas and potential solutions, embedded into an organizational perspective**
    - Oriented towards modeling and analysis
  - **Decide about economic, organizational, technical feasibility in order to select the most promising focus area**
    - Oriented towards managerial decision making
  - **Organization model is used for this purpose**

Phase 2: Impact and improvement study

- **Gather insights into the interrelationship between task, agents involved and use of knowledge and potential improvements**
  - Oriented towards modeling and analysis
- **Identify required organizational measures and task changes in order to ensure organizational acceptance**
  - Oriented towards managerial decision making
- **Two models are offered**
  - Task model
  - Agent model

4.1 The Knowledge Management Framework

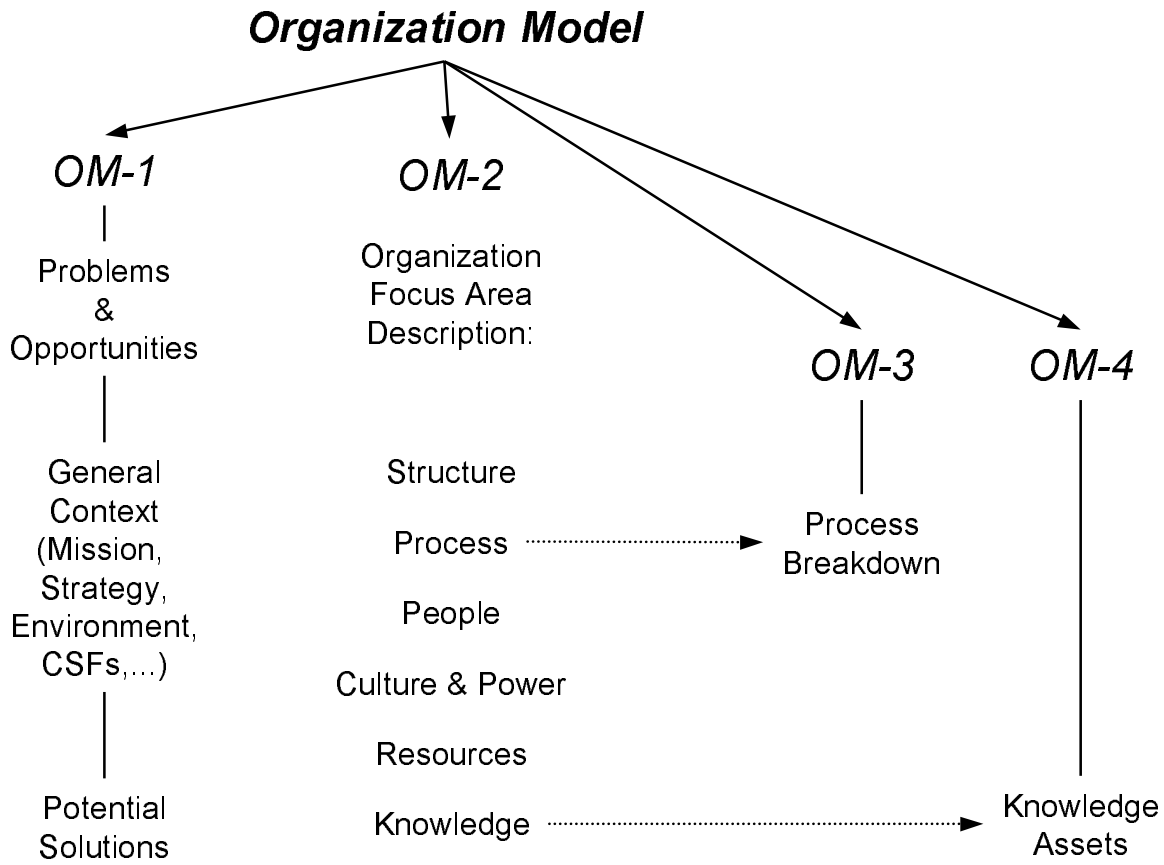


A road map for carrying out knowledge-oriented organization and task analysis.

4.2 Organizational Modeling

- Organization is analyzed from a KM point of view
- Integrate aspects from organization theory, business process analysis, information management
- Model is composed of different components addressing different aspects like
  - Organization structure
  - Processes
  - Staff
  - Resources

4.2 Organizational Modeling



Overview of the components of the CommonKADS organization model.

4.2 Organizational Modeling

Organization Model	Problems and Opportunities Worksheet OM-1
Problems and opportunities	Make a shortlist of perceived problems and opportunities, based on interviews, brainstorm and visioning meetings, discussions with managers, etc.
Organizational context	Indicate in a concise manner key features of the wider organizational context, so as to put the listed opportunities and problems into proper perspective. Important features to consider are: 1. Mission, vision, goals of the organization 2. Important external factors the organization has to deal with 3. Strategy of the organization 4. Its value chain and the major value drivers
Solutions	List possible solutions for the perceived problems and opportunities, as suggested by the interviews and discussions held, and the above features of the organizational context.

Worksheet OM-1: Identifying knowledge-oriented problems and opportunities in the organization

4.2 Organizational Modeling

Organizational Model	Variant Aspects Worksheet OM-2
STRUCTURE	Give an organization chart of the considered (part of the) organization in terms of its departments, groups, units, sections,…
PROCESS	Sketch the layout (e.g., with the help of a UML activity diagram) of the business process at hand. A process is the relevant part of the value chain that is focused upon. A process is decomposed into tasks, which are detailed in worksheet OM-3.
PEOPLE	Indicate which staff members are involved, as actors or stakeholders, including decision makers, providers, users or beneficiaries ("customers") of knowledge. These people do not need to be actual people, but can be functional roles played by people in the organization (e.g., director, consultant)

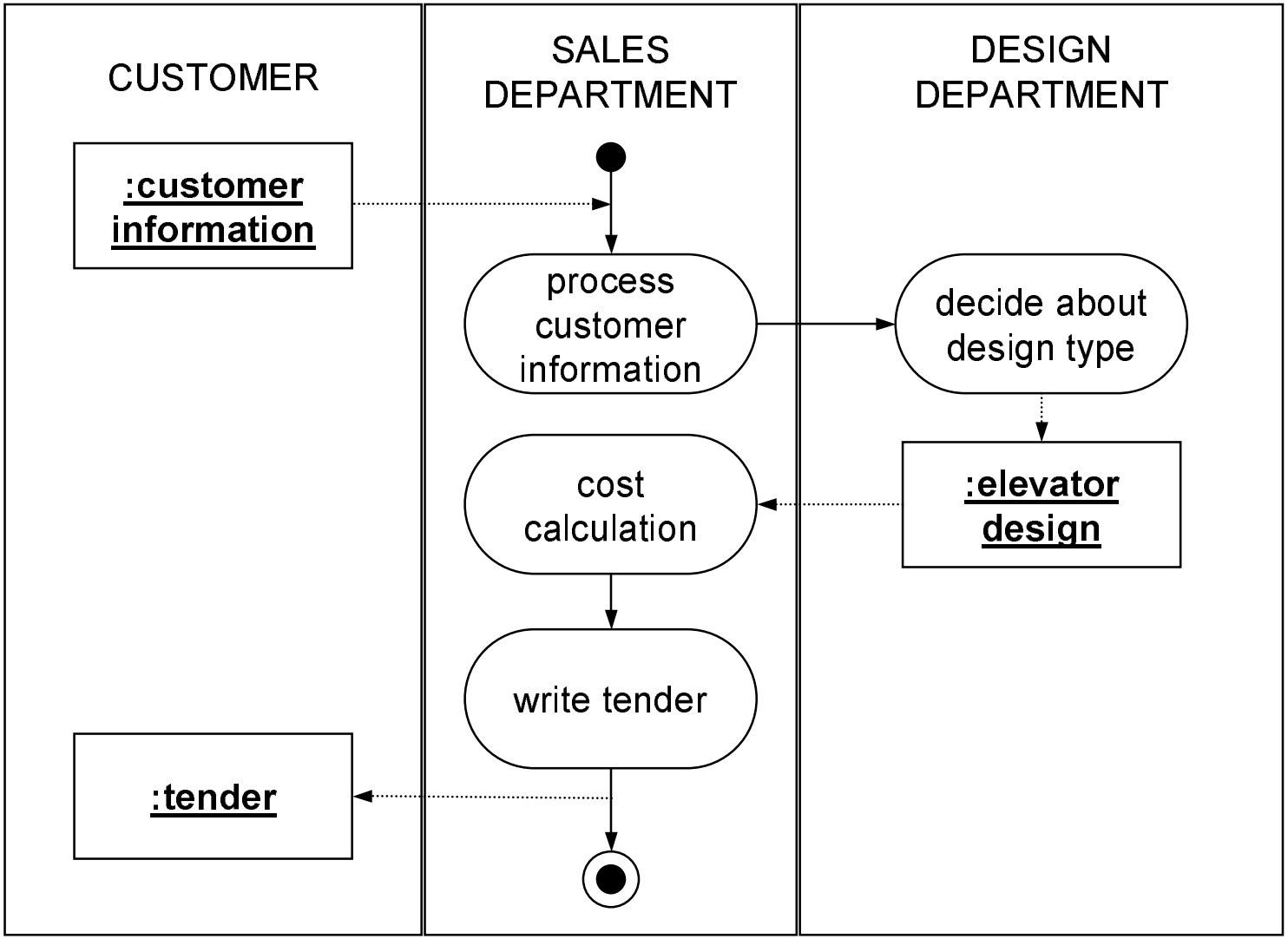
Worksheet OM-2: Description of organizational aspects that have an impact on and/or are affected by chosen knowledge solutions. (Part I)

4.2 Organizational Modeling

Organizational Model	Variant Aspects Worksheet OM-2 (continued)
RESOURCES	Describe the resources that are utilized for the business process. These may cover different types, such as: 1. Information systems and other computing resources 2. Equipment and materials 3. Technology, patents, rights
KNOWLEDGE	Knowledge represents a special resource exploited in a business process. Because of its key importance in the present context, it is set apart here. The description of this component of the organization model is given separately, in worksheet OM-4 on knowledge assets.
CULTURE & POWER	Pay attention to the unwritten rules of the game, including styles of working and communicating ("the way we do things around here"), related social and interpersonal (nonknowledge) skills, and formal as well as informal relationships and networks.

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

4.2 Organizational Modeling



Business process of a company designing and selling elevators, specified through a UML activity diagram

4.2 Organizational Modeling

Organization Model		Process Breakdown Worksheet OM-3				
NO.	TASK	PER-FORMED BY	WHERE?	KNOWL-EDGE ASSET	INTEN-SIVE?	SIGNIFI-CANCE
Task identifier	Task name (some part of the process in OM-2)	A certain agent, either a human (see „People“ in OM-2) or a software system (see „Resource“ in OM-2)	Some location in the organization structure (see OM-2)	List of knowledge resources used by this task	Boolean indicating whether the task is considered knowledge-intensive?	Indication of how significant the task is considered to be (e.g., on a five point scale in terms of frequency, costs, resources or mission criticality

Worksheet OM-3: Description of the process in terms of the task of which it is composed.

AIFB

First overview about important knowledge assets

Organization Model		Knowledge Assets Worksheet OM-4				
KNOWL-EDGE ASSET	POS-SESSED BY	USED IN	RIGHT FORM?	RIGHT PLACE?	RIGHT TIME?	RIGHT QUALITY?
Name (cf. worksheet OM-3)	Agent (cf. worksheet OM-3)	Task (cf. worksheet OM-3)	(Yes or no; comments)	(Yes or no; comments)	(Yes or no; comments)	(Yes or no; comments)

Worksheet OM-4: Description of the knowledge component of the organization model.

4.2 Organizational Modeling

Organizational Model	Checklist for Feasibility Decision Document: Worksheet OM-5
BUSINESS FEASIBILITY	For a given problem/opportunity area and a suggested solution, the following question have to be answered: 1. What are the expected benefits for the organization form the considered solution? Both tangible economic and intangible business benefits should be identified here. 2. How large is this expected added value? 3. What are the expected costs for the considered solution? 4. How does this compare to possible alternative solutions? 5. Are organizational changes required? 6. To what extent are economic and business risks and uncertainties involved regarding the considered solution direction?
TECHNICAL FEASIBILITY	For a given problem/opportunity area and a suggested solution, the following questions have to be answered: 1. How complex, in terms of knowledge stored and reasoning processes to carried out, is the task to be performed by the considered knowledge-system solution? Are state-of-the-art methods and techniques available and adequate? 2. Are there critical aspects involved, relating to time, quality, needed resources, or otherwise? If so, how to go about them? 3. Is it clear what the success measures are and how to test for validity, quality, and satisfactory performance? 4. How complex is the required interaction with end users (user interfaces)? Are stat-of-the-art methods and techniques available an adequate? 5. How complex is the interaction with other information systems and possible other resources (interoperability, systems integration)? Are stat-of-the-art methods and techniques available an adequate? 6. Are there further technical risks and uncertainties?

Worksheet OM-5: Checklist for the feasibility decision document (Part I).

4.2 Organizational Modeling

Organizational Model	Checklist for Feasibility Decision Document: Worksheet OM-5 (continued)
PROJECT FEASIBILITY	For a given problem/opportunity area and a suggested solution, the following question have to be answered: 1. Is there adequate *commitment* from the actors and stakeholders (managers, experts, users, customers, project team members) for further project steps? 2. Can the needed *resources* in terms of time, budget, equipment, staffing be made available? 3. Are the required *knowledge* and other *competences* available? 4. Are the *expectations* regarding the project and its results realistic? 5. Are the *project organization* and its internal as well as external *communication* adequate? 6. Are there further project risks and uncertainties?
PROPOSED ACTIONS	This is the part of the feasibility decision document that is directly subject to managerial commitment and decision making. It weights and integrates the previous analysis results into recommended concrete steps for action: *1.* *Focus:* What is the recommended focus in the identified problem.opportunity areas? *2.* *Target solution:* What is the recommended solution direction for this focus area? 3. What are the expected *results*, *costs*, and *benefits*? 4. What *project actions* are required to get there? *5.* *Risks:* If circumstances inside or outside the organization change, under what *conditions* is it wise to reconsider the proposed decisions?

Worksheet OM-5: Checklist for the feasibility decision document (Part II).

Remarks

- CommonKADS approach is biased towards initiating a development project for a knowledge (-based) system

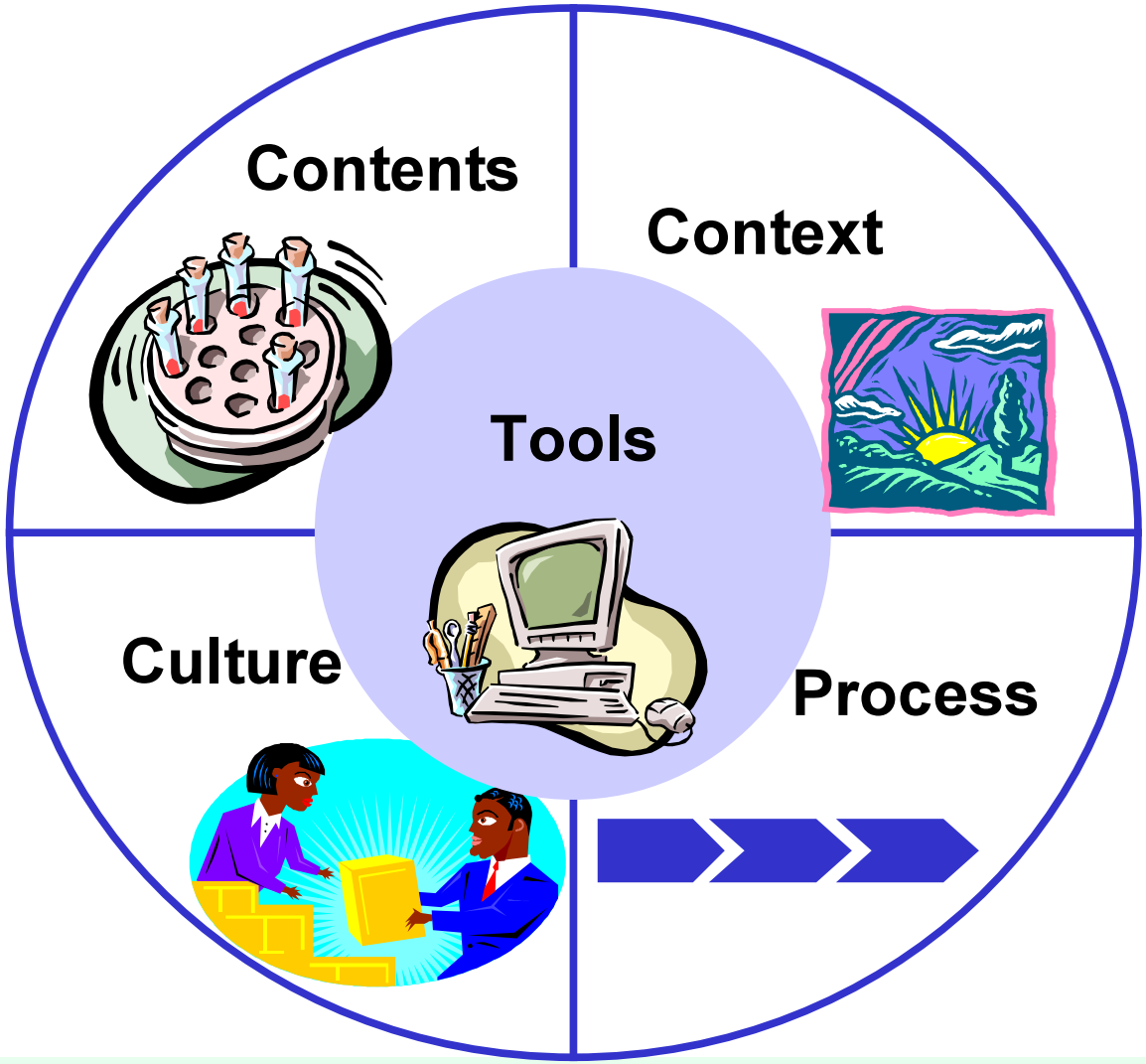
- Feasibility has to include aspects like
  - Are the required organizational changes feasible?
  - Are the required changes for human resource management feasible?

The Chemical Information Network - CIN

Scenario

- Global consulting company
- High fluctuation of experts
- Many projects
- Many clients

Foundations for CIN



AIFB

Foundations for CIN

- Culture - Open Door (in spite of high workload)
- Content
  - new, innovative knowledge
  - lessons learned
  - focused on one particular domain area (Chemicals Practice)
- Process
  - Dedicated Knowledge Manager
  - Specific points for debriefings of expert knowledge (touchdowns of projects etc.)
  - Evaluation Process for knowledge pieces

Foundations for CIN

- Context
  - Thesauri (content-wise, regional thesauri, etc.)
  - Relevant views (which person knows about chemical practice X in South America?; etc.)

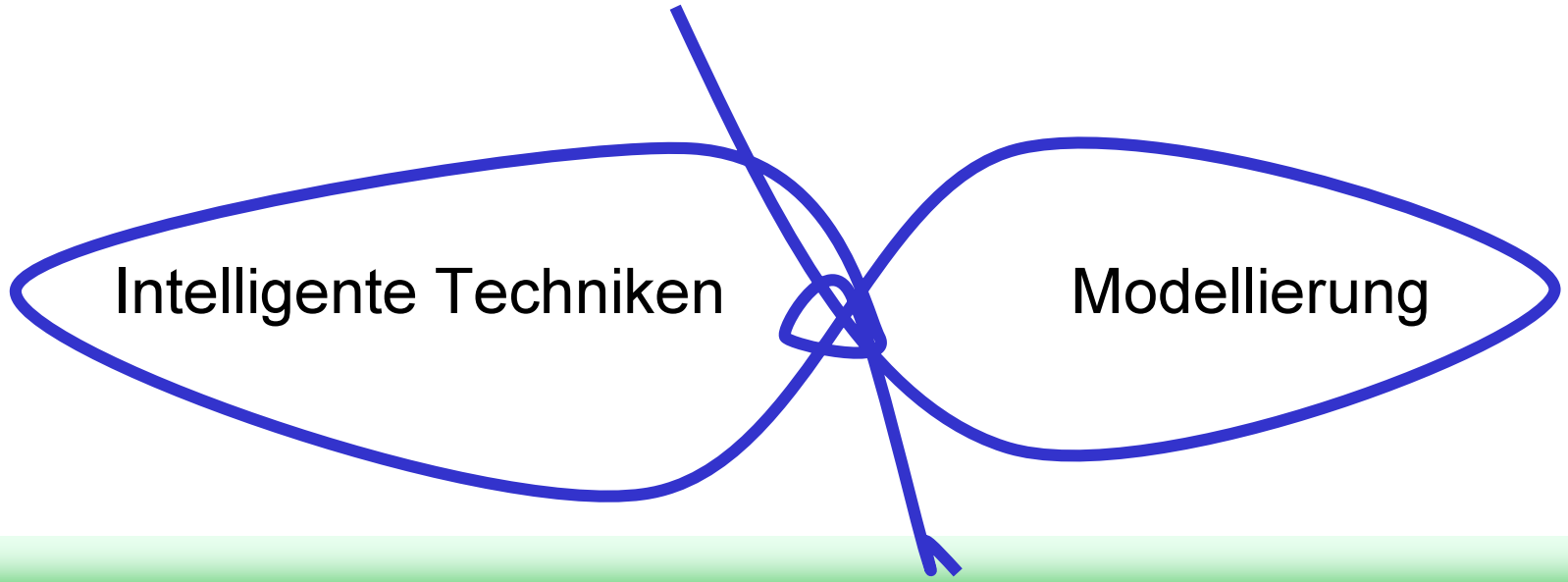
- Tools
  - Web-based document management
  - metadata about documents
    (or “empty documents”, i.e. just facts structured according to thesaurus)
  - thesaurus-based information retrieval

Outcome

- Everyday practice!!!!
- Significant performance improvements of consulting business
- Knowledge base represents an explication of formerly implicit knowledge, measures allow evaluation of strategy!!

Outlook

- Semantic Web
- E-Learning
- Virtual Enterprises



AIFB

Acknowledgements for their support

To

Stefan Decker, Stanford University, CA, USA
All colleagues at AIFB, University of Karlsruhe, Germany
Jürgen Angele & Hans-Peter Schnurr, Ontoprise, Karlsruhe, Germany
Klaus-Dieter Althoff, Fraunhofer IESE, Kaiserslautern, Germany
Robin Burke, University of California, Irvine, CA, USA
Knut Hinkelmann, Insiders, Kaiserslautern Germany
Ulrich Reimer, Swiss Life, Zürich, Switzerland

Literature

- A. Abecker, S. Decker, N. Matta, F. Maurer, U. Reimer (eds.). *Building, Maintaining, and Using Organizational Memories (OM-98).* At ECAI-98. Brighton, August 25th, 1998. http://www.aifb.uni-karlsruhe.de/WBS/ECAI98OM/
- M. Ackerman and E. Mandel. Memory in the Small: Combining Collective Memory and Task Support for a Scientific Community. *Journal of Organizational Computing and Electronic Commerce*, 1999.
- D. Aha & H. Munoz-Avila. Exploring Synergies of Knowledge Management and Case-based Reasoning: Papers from the AAAI 1999 Workshop. Technical Report. AAAI 1999.
- D. Aha, *The AAAI-99 KM/CBR Workshop: Summary of Contributions,* Navy Center for Applied Research in AI Naval Research Laboratory, Washington, DC 20375. http://www.aic.nrl.navy.mil/papers/1999/AIC-99-009.ps
- K.-D. Althoff, F. Bomarius, C. Tautz. Using Case-based Reasoning Technology to Build Learning Software Organizations. In: (Abecker et al. 1998).
- V. Basili, G. Caldiera, F. McGarry, R. Pajerski, G. page, and S. Waligora. The Software Engineering Laboratory --- An Operational Software Experience Factory. In *Proceedings of the IEEE International Conference on Software Engineering*, pages 370--381, 1992.
- D. Billsus & M. J. Pazzani. Learning collaborative information filters. In *Proceedings of the Fifteenth International Conference on Machine Learning*, pages 46-54, Madison, WI, 1998. Morgan Kaufman.
- U. Borghoff & R. Pareschi (Eds). *Information Technology for Knowledge Management.* Springer, Berlin, 1998.
- D. Brickley, R. Guha (eds.): Resource Description Framework (RDF) Schema Specification, W3C Candidate Recommendation 27 March 2000, http://www.w3.org/TR/2000/CR-rdf-schema-20000327
- P. Buitelaar & K. Hasida (eds.), COLING 2000 Workshop Semantic Annotation and Intelligent Content Centre Universitaire, Luxembourg, 5/6 August, 2000.
- Burke, R.D., Hammond, K.J., Kulyukin, V., Lytinen, S.L., Tomuro, N., & Schoenberg, S. (1997). Question answering from frequently asked question files: Experiences with the FAQ Finder system. *AI Magazine*, 18, 57-66.
- P. Chapman, R. Kerber, J. Clinton, T. Khabaza, T. Reinartz, R. Wirth. *The CRISP-DM Process Model*. Discussion Paper, March 1999. http://www.crisp-dm.org/
- S.K Card, G.G. Robertson, and W. York. 1996. The WebBook and the WebForager: an information workspace for the World Wide Web, CHI 96, ACM Conference on Human Factors in Software, ACM Press, New York. 111-117.
- John Davies, Richard Weeks & Mike Revett. Jasper: Communicating Information Agents for WWW, *4th International WWW Conference*, Boston, USA, December 1995. http://www.w3j.com/1/davies.180/paper/180.html
- S. Decker, M. Erdmann, D. Fensel, & R. Studer: Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information. In R. Meersman et al. (eds.), *Semantic Issues in Multimedia Systems*, Kluwer Academic Publisher, Boston, 1999.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.

Literature

- M. Erdmann, A. Mädche, H.-P. Schnurr, & S. Staab. From Manual to Semi-Automatic Semantic Annotation. TechReport. AIFB, Univ. Karlsruhe, 2000.
- Usama Fayyad & Evangelos Simoudis. *Data Mining*. Tutorial at the Fourteenth National Conference on Artificial Intelligence. July 227, 1997. AAAI.
- D.J. Foskett. Thesaurus. *Encyclopaedia of Library and Information Science*, 30, 416-462. Reprinted in: Sparck-Jones & Willett 1997.111-134.
- Y. Freund, R. Iyer, R.E. Schapire, & Y. Singer. An efficient boosting algorithm for combining preferences. In Machine Learning: Proceedings of the Fifteenth International Conference, 1998.
- D. Goldberg, D. Nichols, B.M. Oki, & D. Terry. Using Collaborative Filtering to Weave an Information Tapestry. *Communications of the ACM*, 35(12) , 61-70, 1992.
- G. v. Heijst, R. v. der Spek, & E. Kruizinga. The Lessons Learned Cycle. In: Borghoff & Pareschi (eds.) 1998.
- J. Justeson & S. Katz. Technical Terminology: Some Linguistic Properties and an Algorithm for Identification in Texts. *Natural Language Engineering* 1(1), 9-27, 1995.
- Jörg-Uwe Kietz, Katharina Morik: A Polynomial Approach to the Constructive Induction of Structural Knowledge. *Machine Learning* 14(1): 193-217 (1994).
- J. Kolodner. Case-Based Reasoning, Morgan Kaufmann Publishers, San Mateo, CA, 1993.
- J. Lamping, Ramana Rao: The Hyperbolic Browser: A Focus + Context Technique for Visualizing Large Hierarchies. Journal of Visual Languages and Computing 7(1): 33-55 (1996)
- O. Lassila, Ralph Swick (eds).: Resource Description Framework (RDF) Model and Syntax Specification, W3C Recommendation 22 February 1999, http://www.w3.org/TR/REC-rdf-syntax/
- M. Lenz, B. Bartsch-Spörl, H.-D. Burkhard, & S. Wess (Eds). Case-based Reasoning Technology. LNAI 1400. Springer, Heidelberg, 1998.
- M. Lenz. Managing the Knowledge Contained in Technical Documents. In: (Reimer 1998).
- A. Maedche & S. Staab. Semi-automatic Engineering of Ontologies from Text. In: *Proceedings of the Twelfth International Conference on Software Engineering and Knowledge Engineering (SEKE'2000).*
- R. Magaldi. A Case-based Approach to the Management an Deployment of Knowledge Assets. In: *(Aha & Munoz-Avila 1999).*
- T. Munzner: Exploring Large Graphs in 3D Hyperbolic Space IEEE Computer Graphics and Applications, Vol. 18, No. 4, pp 18-23, July/August 1998.
- M. Pazzani & D. Billsus. Learning and Revising User Profiles: The identification of interesting web sites. *Machine Learning*, 27(3), 313-331, Kluwer Academic Publishers, 1997.
- Pazzani, M., Muramatsu J., & Billsus, D. (1996). Syskill & Webert: Identifying interesting web sites. Proceedings of the National Conference on Artificial Intelligence, Portland, OR.
- L. Rau, Conceptual information extraction and retrieval from natural language input. In: *RIAO 88*, pp. 424-437. Paris, 1988. Reprinted in: Sparck-Jones & Willett 1997.527-533.

Literature

- U. Reimer (ed.). *PAKM-98: Proceedings of the Second International Conference on Practical Aspects of Knowledge Management*. Basel, CH, October 29-30, 1998.
- Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, John Riedl. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, Chapel Hill, NC: pp. 175-186.
- M. Staudt, J.-U. Kietz, U. Reimer, A Data Mining Support Environment and its Application on Insurance Data, In: Proceedings of the 4th Int. Conf. On Knowledge Discovery and Data Mining (KDD-98), AAAI Press, New York, USA, August 1998, pp.105-111.
- K. Sparck-Jones & P. Willett (eds.), *Readings in Information Retrieval*. Morgan Kaufmann, San Francisco, CA, 1997.
- S. Staab. *Grading Knowledge: Extracting Degree Information from Texts*. LNAI 1744. Springer, Heidelberg, 1999.
- S. Staab, A. Mädche, C. Nedellec, P. Wiemer-Hastings (eds.). *Ontology Learning. Proceedings of the ECAI-2000 Workshop*. Berlin, August 25, 2000.
- S. Staab, H.-P. Schnurr, R. Studer, Y. Sure: Knowledge Processes and Ontologies. *IEEE Intelligent Systems*, 16(1), January/February 2001 (Special issue on knowledge management).
- A. Mädche, S. Staab: Ontology Learning for the Semantic Web. *IEEE Intelligent Systems*, 16(2), March/April 2001 (Special issue on semantic web).
- S. Staab, A. Mädche: Knowledge Portals - Ontologies at Work. To appear in: *AI Magazine*, 21(2), Summer 2001.
- S. Staab, J. Angele, S. Decker, M. Erdmann, A. Hotho, A. Mädche, H.-P. Schnurr, R. Studer, Y. Sure. Semantic Community Web Portals. In: *WWW9 / Computer Networks (Special Issue: WWW9 - Proceedings of the 9th International World Wide Web Conference, Amsterdam, The Netherlands, May, 15-19, 2000)*, 33(1-6): 473-491. Elsevier, 2000.
- S. Staab, H.-P. Schnurr. Smart Task Support through Proactive Access to Organizational Memory. *Knowledge-based Systems*, 3(5): 251-260. Elsevier, 2000.
- E. Tsui, B. Garner, S. Staab. The role of Artificial Intelligence in Knowledge Management. *Knowledge-based Systems*, 13(5): 235-239. Elsevier, 2000.
- Topic Navigation Maps: ISO-document: http://www.ornl.gov/sgml/sc34/document/0058.htm/ Demonstration: http://www.stepuk.star.co.uk:82/ See also: http://www.topic-maps.com/, http://www.infoloom.com/
- World Development Report 98/99, Worldbank, http://www.worldbank.org/wdr/
- E. Wiener, J.O. Pedersen, & A.S. Weigend. A Neural Network Approach to Topic Spotting. In Proceedings of the 4th Symposium on Document Analysis and Information Retrieval (SDAIR 95), pages 317-332, Las Vegas, NV, USA, April 24-26 1995.
- J.A. Wise, J.A., J.J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow. 1995. Visualizing the non-visual: spatial analysis and interaction with information from text documents. Proceedings of Information Visualization, October 20-21, 1995. IEEE Computer Society Press, Los Alamitos, CA. 51-58.

AIFB

Thank You!