Distributed Enterprise Knowledge Management: Balancing Individual and Organizational Needs

Ludger van Elst
German Research Center for Artificial Intelligence (DFKI)
ドイツ人工知能研究センター(DFKI)
- Knowledge Management Department -
elst@dfki.uni-kl.de

2004-11-05, Osaka Prefecture University

大阪府立大学

Outline

Starting Point	Supporting Knowledge Management with Organizational Memory Information Systems
Drawbacks and Solution Approach	From Centralized to Distributed Organizational Memories
Agent-based Realization	The FRODO Framework for Distributed Organizational Memories
Summary and Outlook	Towards Agent-Mediated Knowledge Management

Outline

Starting Point	Supporting Knowledge Management with Organizational Memory Information Systems
Drawbacks and Solution Approach	From Centralized to Distributed Organizational Memories
Agent-based Realization	The FRODO Framework for Distributed Organizational Memories
Summary and Outlook	Towards Agent-Mediated Knowledge Management

Knowledge Management (KM) Research is Strongly Driven by Real World Needs of Today's Enterprises

- Nonaka/Takeuchi attributed Japan's success over the US economy (in the eighties) to improved knowledge creation
- Many companies define themselves as becoming "Knowledge Organizations"
- Many companies had KM projects (often assessed as flops ☹)
- Many companies had Information/Document Management projects (often labeled as KM projects and rated as flops ☹)
- Many companies still have (Info/Document/___) Management projects that root in bad KM
- There are still public discussions about the transition of many countries into "Knowledge/Information Societies"

Of course, the buzzword lifecycle might also apply to KM.

Knowledge management intends a holistic approach

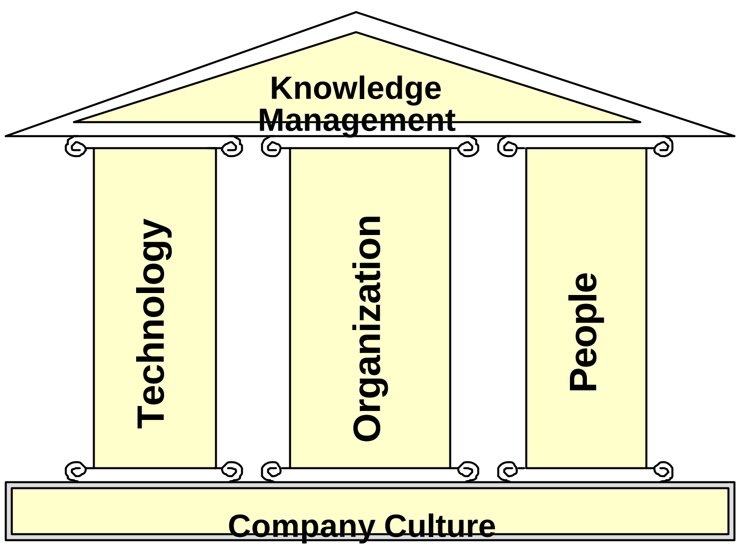
Knowledge Management is a

- structured, holistic approach
- to improve the handling of knowledge

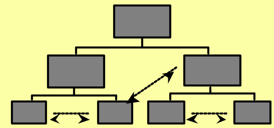
  (know-how, experience, skills, active documentation)
- on all levels (individual, group, organizational)
- in order to save costs, improve quality, support innovation

see: www.netacademy.org

Credo: Successful KM Needs a Holistic View

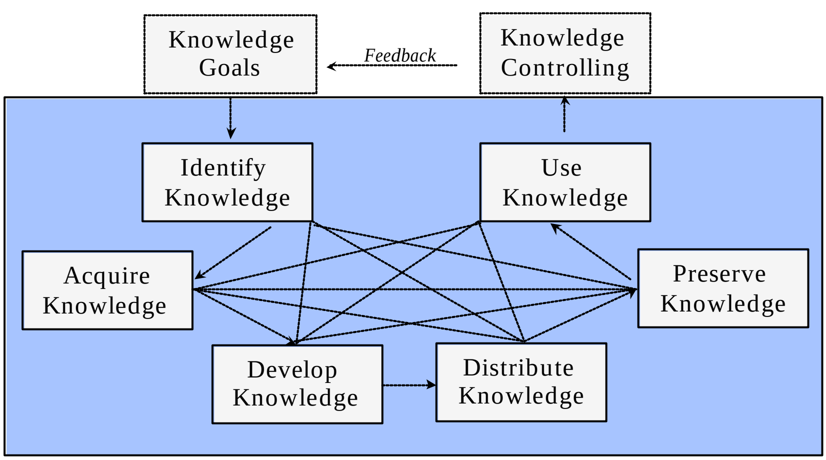


Knowledge Management Takes Place at Various Levels

Individual Level 	*Group Level* 	 *Organizational Level*
Intuition Competencies Knowledge Expectations ...	Routines Role allocation Shared language Complementary com- petencies ...	Core competencies Myths Secret rules Contracts Electronic knowledge base ...
KM as an individual competency	KM as a team tool	KM as an organizational method

Adapted from: M. Eppler/St. Gallen

Information Technology is often seen as an enabling factor for facilitating organizational Knowledge Management



Adapted from: Probst/Raub/Romhardt

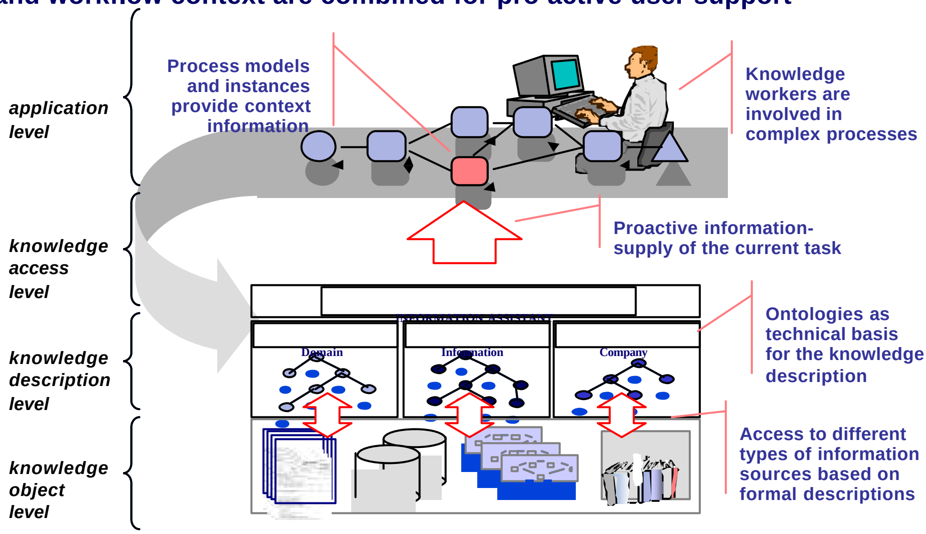
Consequences From History of the DFKI-KM Group (before 1996)

- *Assistant Systems* Instead of Expert Systems
- System as *Knowledge & Communication Medium*
- *Knowledge Evolution* as Task
- Integration of *Different Formality Levels* of Knowledge
- Integration with *Legacy Systems* and *Standard Applications*
- Links between *Heterogeneous Information Items*

This led to a

- working definition: *Knowledge = Information Made Actionable*
- basic research project: *Organizational Memories*

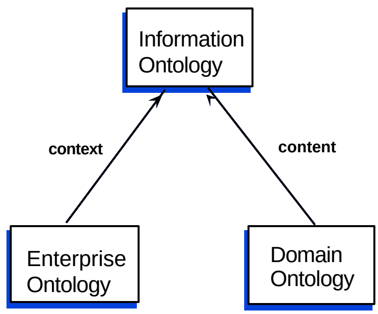
KnowMore Architecture (1998): Ontology-based information description and workflow context are combined for pro-active user support



Ontologies organize information models and background knowledge

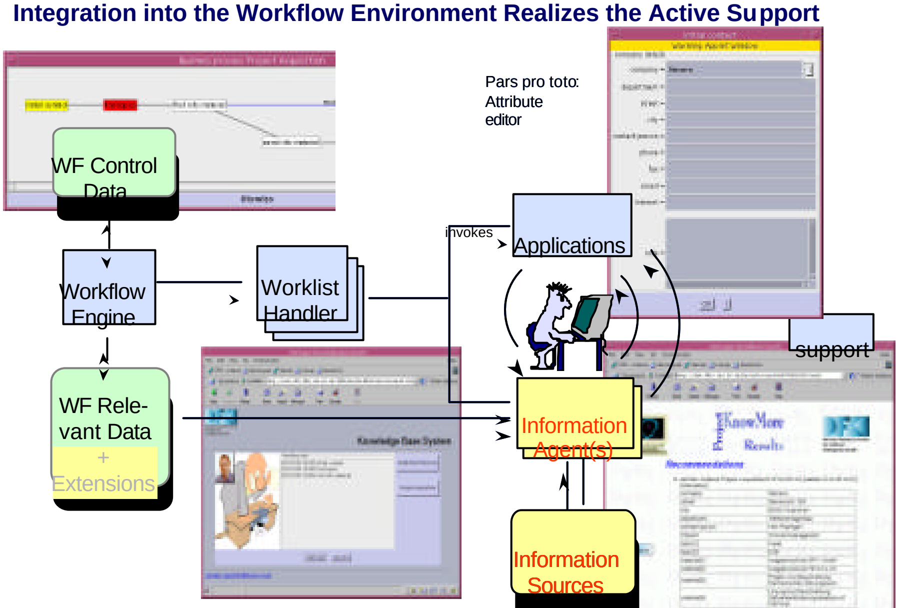
Ontology: *"An ontology is an explicit specification of a conceptualization."* (Gruber, 1993) – typically taxonomies, classes, relationships

DIMENSIONS OF INFORMATION MODELING



- **information ontology:**
  - kinds of information sources
  - logical structure -> relevance propagation
  - **meta properties** (reliability, message type, availability, creation context, intended usage context)
  - link to information content
- **enterprise ontology:**
  - provides usage and creation context
  - basis for BPMs
  - enterprise organizational structure
- **domain ontology:**
  - description of information content
  - usually incomplete

Integration into the Workflow Environment Realizes the Active Support



Outline

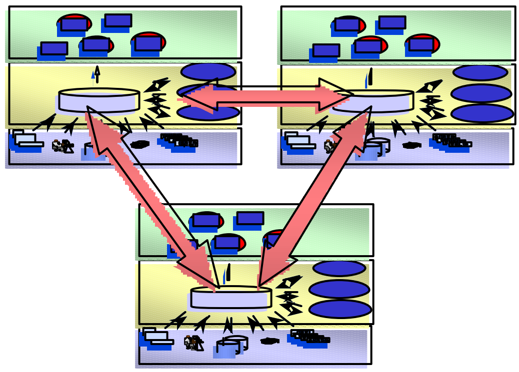
Starting Point	Supporting Knowledge Management with Organizational Memory Information Systems
Drawbacks and Solution Approach	From Centralized to Distributed Organizational Memories
Agent-based Realization	The FRODO Framework for Distributed Organizational Memories
Summary and Outlook	Towards Agent-Mediated Knowledge Management

Observation 1: A monolithic central OM is seldom feasible

- Various stakeholders in an organization have different requirements
  - individual knowledge sources
  - domain-specific knowledge structures
- Each stakeholder closely guards knowledge in its possession
  - Responsibility, competition, rivalry
- Information sources are structured according to the particular needs of the respective stakeholder
  - Explicit ontologies illustrate the respective organization principles
- Evolutionary grow-up of knowledge management solutions has advantages
  - high motivation by 'quick wins'
  - success stories in pilot areas convinces the top management
- ... but results in competing, dispersed results
  - individual solutions resist global standardizations

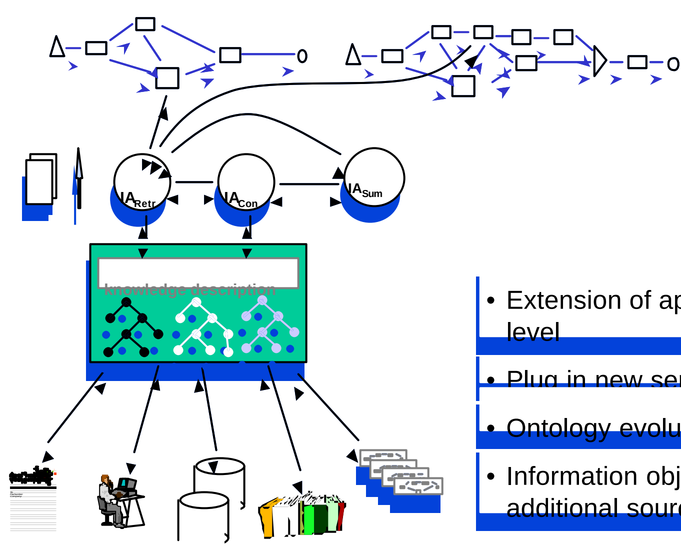
A flexible Framework for Distributed Organizational Memories (FRODO) facilitates the evolution of OMs by integrating different local solutions

FRODO extends the OM paradigm towards a less rigid, distributed scenario



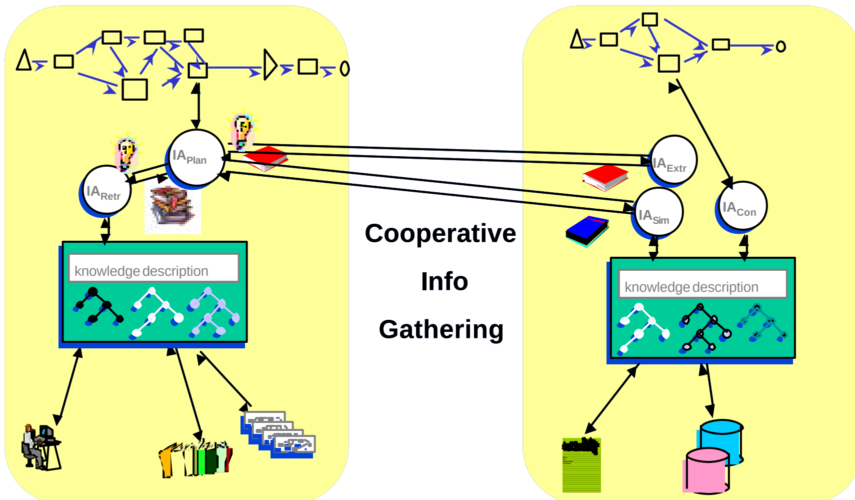
- OM introduction starts with ‚quick wins‘ and small pilots
- Several (group/department-wide) OM can be established
- To realize a comprehensive OM the islands must inter-operate
- Scalability is the key question

Vertical Scalability Allows to Extend One OM in All Relevant Dimensions



- Extension of application level
- Plug in new services
- Ontology evolution
- Information objects from additional sources

Horizontal Scalability Addresses Interoperability of Several OMs



Co-operation of OMs requires complex co-ordination mechanisms.

Observation 2: Different types of work require different support!

- Static process models / workflows provide reliable triggers and valuable context information
  - if the work in question is repetitive in nature
  - if the work in question can be modeled a priori
  - if information needs are determined once and for all
- Knowledge-intensive work typically can not be modeled by static process models
  - details of work are not repetitive
  - task sequences are not known a priori
  - information needs vary greatly

Process-oriented support for knowledge-intensive work require the notion of dynamically configured, agile knowledge workflows.

Outline

Starting Point	Supporting Knowledge Management with Organizational Memory Information Systems
Drawbacks and Solution Approach	From Centralized to Distributed Organizational Memories
Agent-based Realization	The FRODO Framework for Distributed Organizational Memories
Summary and Outlook	Towards Agent-Mediated Knowledge Management

The agent paradigm is appropriate to model distributed OM scenarios

- The characteristics of distributed organizational memories in realistic enterprise scenarios are described by the notion of agent societies
  - components have to be considered as autonomous units
    - individual business units with specific information sources and structures
    - individual goals result in different commitments
    - individual procedures cope with local particularities
  - cooperation relies on agreements between partners
    - societies of agents with agreed-upon roles
    - interactions are governed by rights and obligations
- Using the agent paradigm to model OM designs results in clear roles, responsibilities, and communication structures

Agent societies are characterized by underlying role models

- Role models reflect social competence of agents
  - modelled by rights and obligations
  - influence agent behaviour
  - resulting in typical speech acts and protocols for society build-up
- Role models allows to ensure some global system characteristics while also preserving individual flexibility
  - Explicit rights and obligations allow to commit to specific roles
  - roles guarantee global behaviour
  - role descriptions are represented by formal models
- The notion of socially-enabled agents is relevant for all FRODO framework components

A social model is defined by rules

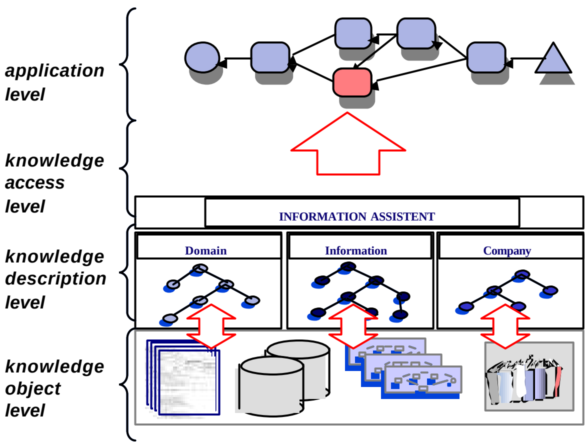
- SpeechAct ::= (FRODO_SA, Protocol^{0,1}).
- Competency ::= (ReceiverRole, SpeechAct)| Action.
- Right ::= **perform** Competency **if** Condition.
- Obligation ::=
  when SpeechAct **from** ReceiverRole **andif** Condition
  perform Competency |
  if Condition **perform** Competency.
- Role ::= rolename(Right*, Obligation*).
- Rolemodel ::= rolemodelname{Role*}.

Rights are modeled as *filter rules* on intentions

Obligations are modeled as *reactive* or *proactive rules*

A social layer in the platform ensures fair processing of rights & obligations

In FRODO, *all levels of a single OM* as well as *mediation services* between OM are designed and implemented as agents

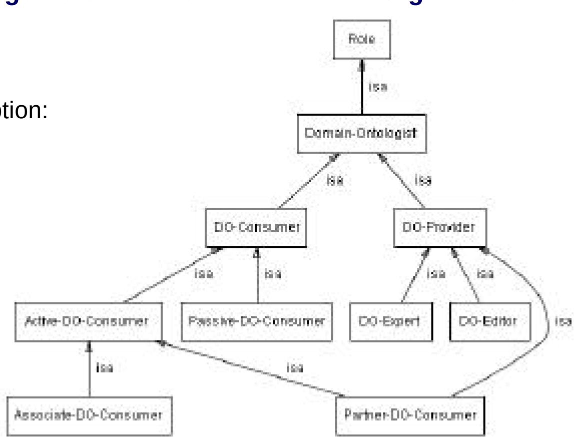


- An agent-based weakly-structured workflow system for specifying information needs and their context
- Information agents for satisfying specific information needs
- Ontology agents for the maintenance of domain vocabulary
- Wrappers for info sources

Different ontology related agents can be identified in the organization

Knowledge Level Description:

- Goals
- Knowledge
- Competencies
- Rights
- Obligations



The ontology society is formed by determining rights and obligations of specific agents

Concrete speech acts are derived from knowledge-level descriptions

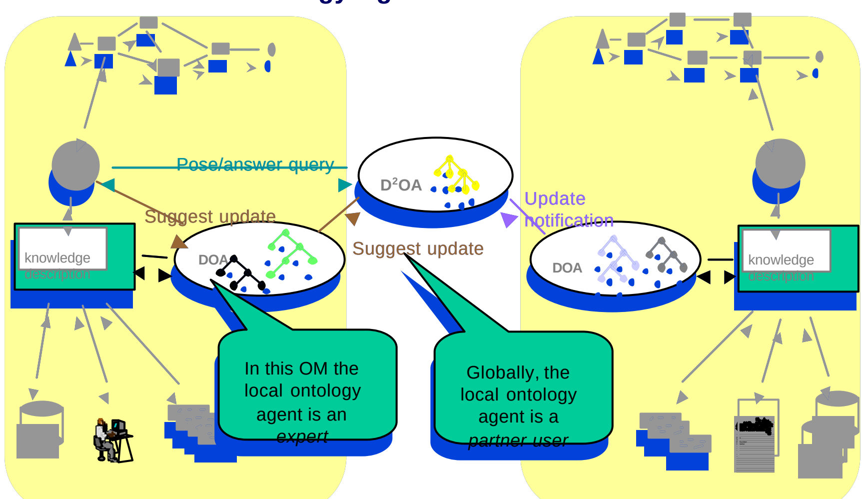
	Non User	Passive User	Associate User	Partner User	Expert	Editor
Query		R	R	R	R	R
Answer Queries					R/O	R
Receive Update			R	R	R	R
Suggest Update		R	R	R/O	R	R/O
Edit						R
Send Upd. Notif.						R/O
ApplyForRole	R	R	R	R		
Grant Guarantees						R
Guarantee Quality						O

- Ontology Utilization
- Ontology Evolution
- Ontology Socialization

R: has-the-right-to

O: is-obliged-to

Distributed Domain Ontology Agents Mediate Between Different OMs



Each agent can play different roles with respect to the various ontologies

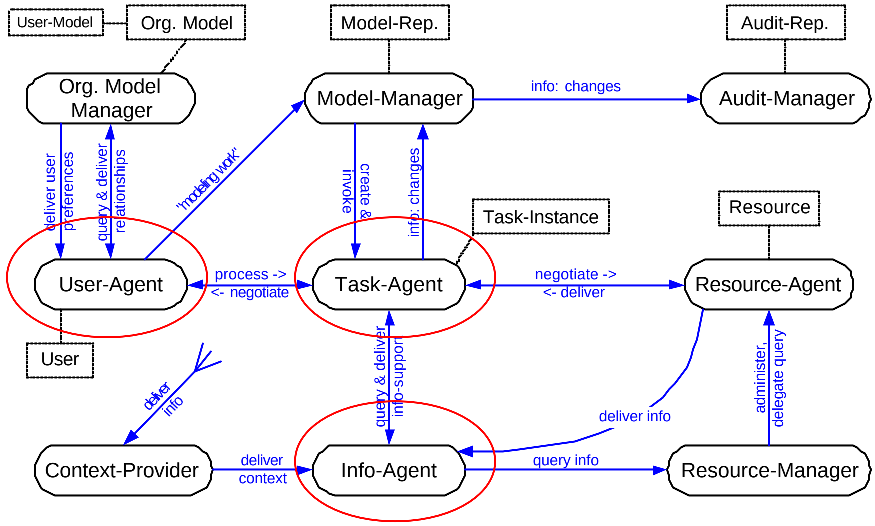
The application level of an OM asks for additional agent structures

- Workflow Agents enact agile knowledge processes
  - intertwining of workflow modeling and execution
    - new workflow blueprints and new tasks are created during execution
  - the individual task instance is represented as an agent
    - responsible for completion of the task
    - communicates e.g. with other task instances in the workflow
  - other relevant entities are realized as agents with specific roles
    - models, resources, manager
- Personal User Agents mediate between user and system
  - represent and observe individual goals and preferences
  - may act pro-actively, observing local and global context
  - interact and negotiate with other PUAs to realize collaborative aspects

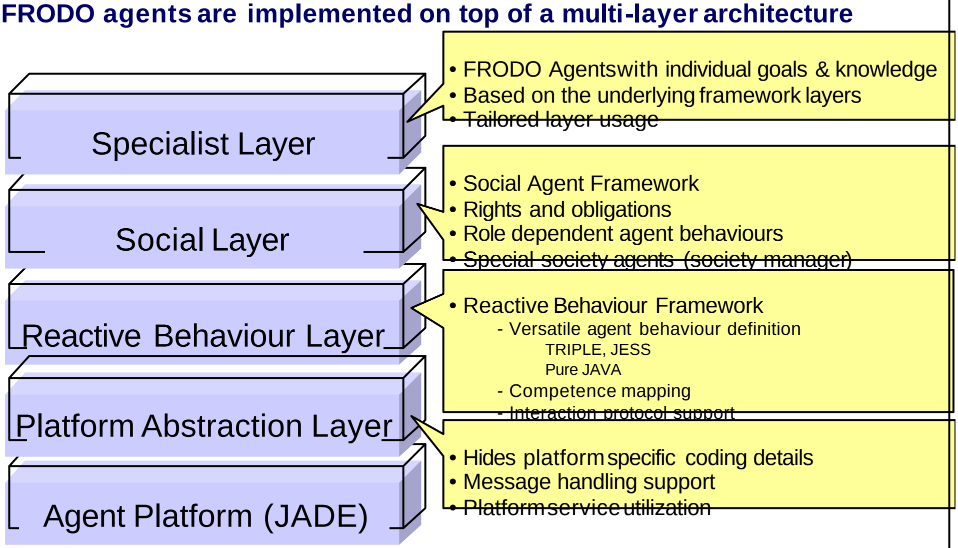
Agents support the knowledge object & access level of distributed OMs

- Info Agents access the information sources
  - multiple agent types, roles, and functionalities realize distributed information management
  - Info Agents interact with Context Agents and Personal User Agents
- Distributed information processing profits from agent structures
  - document analysis and classification
  - need-driven metadata extraction
  - wrappers
  - collaborative filtering

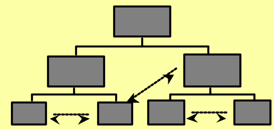
The agents interact in a complex society structure in order to realize information support based on weak workflows



FRODO agents are implemented on top of a multi-layer architecture



Looking back: Knowledge Management Takes Place at Various Levels

Individual Level 	***Group Level*** 	 ***Organizational Level***
?	Distributed Organizational Memories with Flexible Ontology Societies and Agile Knowledge Workflows	Monolithic Organizational Memory with Global Ontologies and (Standard) Workflows
EPOS (since 2003)	**FRODO (2000-2002)**	**KnowMore (1997-1999)**

Knowledge Management has to cope with contradictions between personal and organizational goals

- Organizations introduce organizational memories (OMs) to improve access to and use of critical knowledge
- Individuals do not and do not want to realize any benefit
- The introduction in almost all cases requires new duties
  - document activities
  - describe skills
  - categorize and structure information
  - answer additional questions
  - learn and accepts pre-given access modalities
  - formulate requests

Knowledge workers often do not accept knowledge management technology in order to keep their subjective productivity

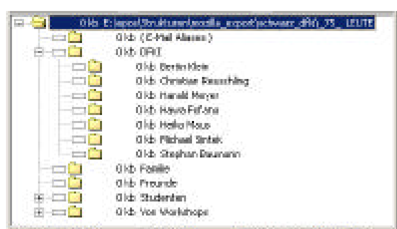
Effort in structuring individual information spaces provides valuable input for Knowledge Management

- Knowledge workers use various tools for conceptualizing their domains
  - Generic operating system structures (file folders)
  - Dedicated information management applications (address books, mail tools, outliners, mind managers,…)
- Advantages of these *native structures* concern knowledge utilization and acquisition
  - They reflect – at least temporarily – the worker's individual view and can therefore easily be exploited by the consultant
  - They are regularly extended and maintained
- Problems arise from the lack of clear semantics
  - With easily extendable structures, often redundant and contradictory models are created which are
    - difficult to utilize by automatic services
    - hard to share with other knowledge workers
  - The (ascribed) ad hoc semantics is typically not stable.
    Therefore, the usefulness over time even for one knowledge worker is not given.

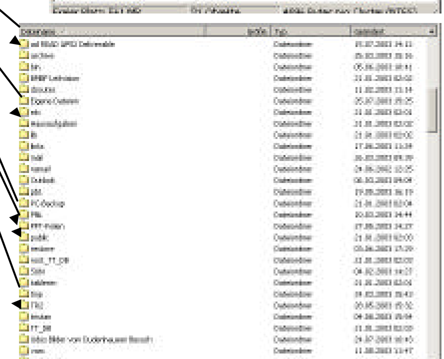
In EPOS, a formally grounded personal information model, fed by the native structures, bridges individual and organizational Knowledge Management.

From a modeling point of view native structures are often flawed

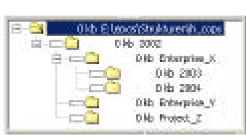
- No distinction between/mixture of `subclass-of` and `part-of` relations:
  - DFKI `subclass-of` LEUTE ???
  - Bertin Klein `part-of` DFKI !!!
  - Studenten `subclass-of` Leute !!!
- No separation of distinct domains



Project & document type
Tasks/To do
Document type
Access right
Persons



- Document centered view often justifies models
- But also:
  Real violations of semantics

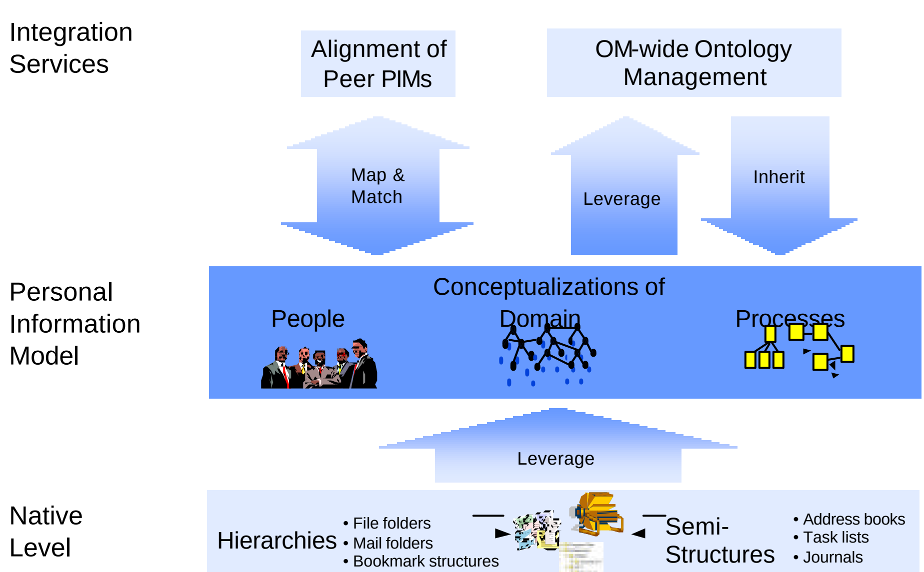


The EPOS vision: A Personal Information Model (PIM) as *semantic middleware* for knowledge services

- The PIM is a formally grounded model
- More global ontologies as well as native structures provide input
- A maintenance assistant will help with stepwise formalization of native structures
- The PIM can be utilized by various knowledge services (retrieval, personal information agent, visualization, …)

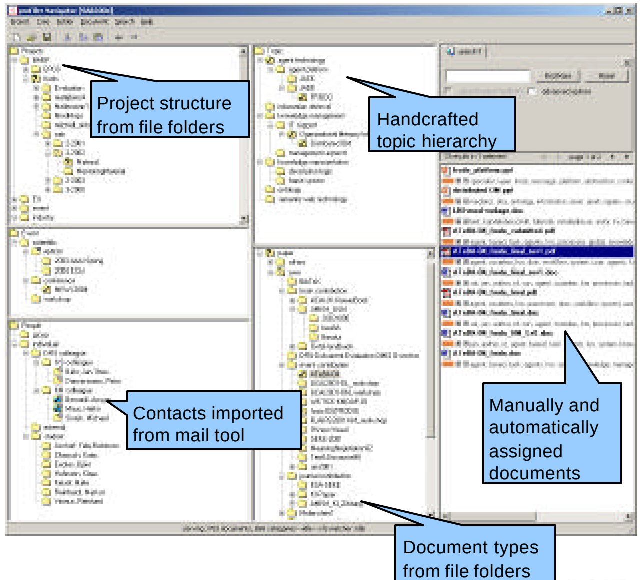
- Technical aspects:
  - Semantic Web technology allows for seamless integration into broader environments (group, company, Internet)
  - The JENA 2 framework (by HP) will allow for persistent handling of PIM base on RDF/S and OWL
  - The PIM implementation can be seen as a Semantic Web ontology Service

- Challenges:
  - Integration of existing ontologies
  - Leveraging native structures
  - Mappings between PIM

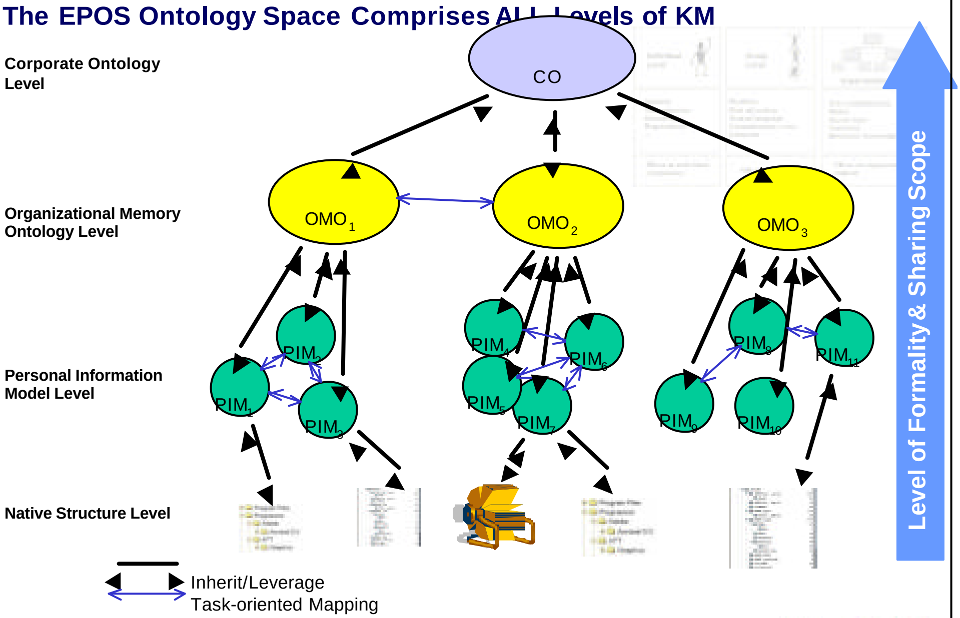


The brainFiler integrates various native structures in one Personal Information Model

- Application with explorer plus search engine-like look&feel
  - Multiple personal views on the native file system
  - Integration of bookmark and mail folders
  - Active synchronization with native structures
- Knowledge Model: taxonomies + attribute-value pairs
  - RDF/S im-/export
- For attached documents, the classifier realizes the is-a semantics
  - Set inclusion
  - Automatic, content-based classification suggestions
- Developed together with brainbot AG
  - brainbot is a DFKI spin-off company



The EPOS Ontology Space Comprises ALL Levels of KM

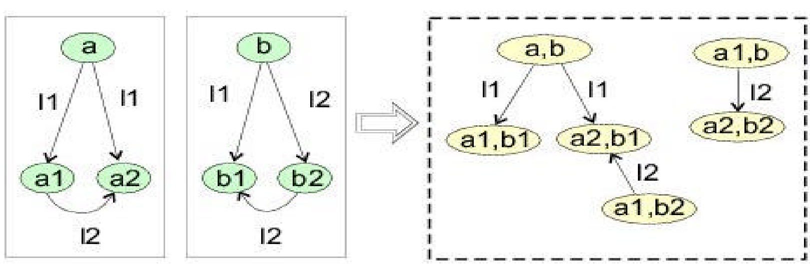


Mapping Personal Information Models is a core functionality in EPOS

- Semantic unification is a *complex and highly knowledge-intensive task* (Bachmann, 1997; Wache, 2003) and in general not solvable.
- In EPOS, we use three sources of evidence for concept mappings
  - *Term-based*: Similarity of concept names
    In many tools the only source of evidence (e.g., Noy's PROMPT tab in Protégé-2000, McGuiness' Chimaera)
  - *Topology-based*: Similar structures of concept graphs
    Promising results in DB schema matching (e.g., Rahm, 2001)
  - *Instance-based*: Similarities of instances give evidence for relations between the respective classes
    Rarely found in literature, but promising in document-centered applications
- A first prototype realizes these three types of evidence
  - Basis for the prototype: The Protégé-2000 PROMPT tab provides the basic environment for semi-automated ontology mapping
  - Term-based similarities are determined by evaluating overlap of 3-grams and 4-grams
  - Similarity of the concept graphs is determined by similarity flooding
  - The brainFiler is used to assess similarities of documents that are annotated with concepts

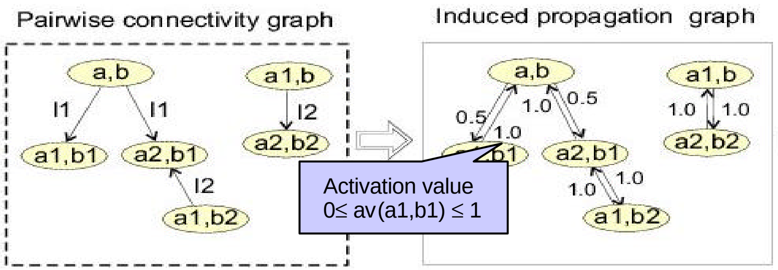
Similarity flooding for assessing overlap in the structure of two PIM (1) (adapted from Melnik & Rahm, 2001)

- **Input** : Two concept graphs
  Output : Set of map pairs with corresponding similarity values
- In the **pairwise connectivity graph PCG**, the nodes are map pair candidates and the edges are relation types of the original concept graphs.



Similarity flooding for assessing overlap in the structure of two PIM (2)

- From the PCG a **propagation graph** is constructed, where the directed connectivity edges are substituted by directed, weighted edges and nodes are supplemented with activation values.



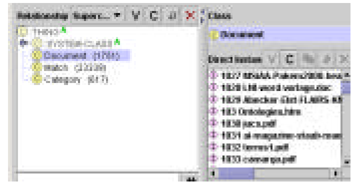
- The set of edges can be represented as 2-dim matrix EDGE with nodes as indices and edge weights as element value. The activation values can be seen as vector AV.
- **Fixpoint computation**: $AV^{i+1} = \text{normalize}(EDGE^* AV^i)$ until $|AV^{i+1} – AV^i| < \varepsilon$.
  Converges if not all AV[]=0 (Rodenhausen, 1992)

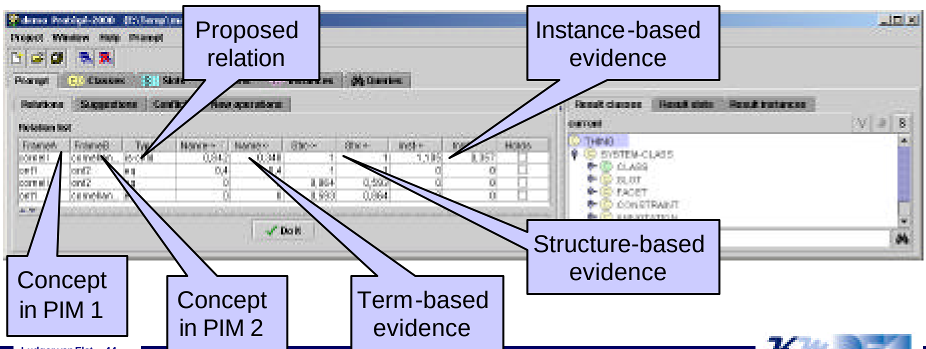
Instance-based evidence generation is based on the brainFiler functionality for assessing document similarities

- **Input**
  Documents $D_1,\ldots,D_m$ annotated with concepts $C_1,\ldots,C_k$ from PIM1
  Documents $D'_1,\ldots,D'_n$ annotated with concepts $C'_1,\ldots,C'_l$ from PIM2
- **Output**
  mapping values $mv(C_i,C'_j)$ indicating evidence C_i `subclass-of` C'_j
  mapping values $mv(C'_i,C_j)$ indicating evidence C'_i `subclass-of` C_j
- **Procedure**
  - Learn Classifiers (by proFiler engine)
    Classifier CF : $D \cup D' \times C \rightarrow [0,1]$ with (D_i annotated with C_j) $\Rightarrow CF(D_i,C_j) = 1$
    Classifier CF': $D \cup D' \times C' \rightarrow [0,1]$ with (D'_i annotated with C'_j) $\Rightarrow CF'(D'_i,C'_j) = 1$
  - Cross-classify documents
    $mv(C_i,C'_j) = mean_{Di}((CF(D_i,C_j)==1) * CF'(D_i,C'_j))$
    $mv(C'_i,C_j) = mean_{D'i}((CF'(D'_i,C'_j)==1) * CF(D'_i,C_j))$
- **Subclass-of semantics**
  IF all instances of C_i are instances of C'_j
  THEN C_i `subclass-of` C'_j
  Instance-evidence heuristic
  IF many documents that are annotated with C_i are classified as C'_j documents
  THEN propose(C_i `subclass-of` C'_j)

All three evidence generation mechanism have been integrated in the PROMPT ontology mapping environment for Protégé-2000

- The brainFiler was integrated to generate instance-based evidence
- A widget was developed for a spreadsheet-like “generate evidence – explore – decide – test” cycle
- Work in progress & next steps
  - Integration of different evidences
  - Elaboration of systematic evaluation scenario





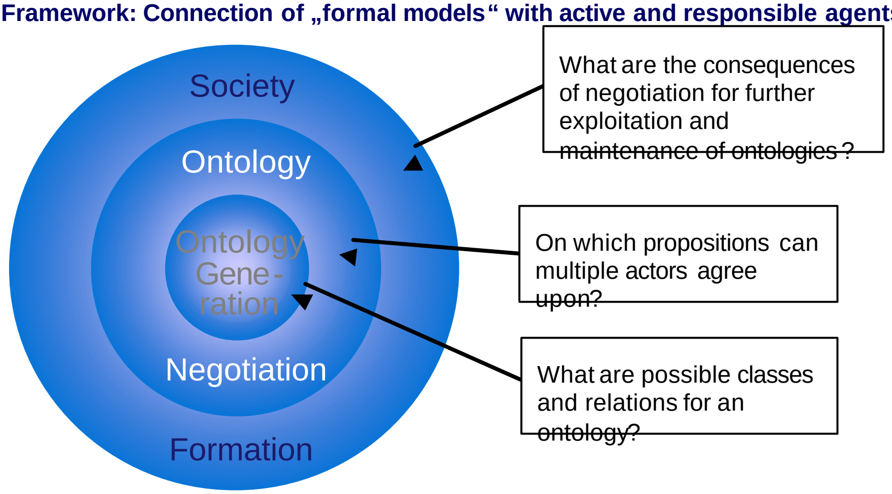
Outline

Starting Point	Supporting Knowledge Management with Organizational Memory Information Systems
Drawbacks and Solution Approach	From Centralized to Distributed Organizational Memories
Agent-based Realization	The FRODO Framework for Distributed Organizational Memories
Summary and Outlook	Towards Agent-Mediated Knowledge Management

Summary

- Organizational Memories as concept for supporting organizational KM
- Characterization of KM landscapes show drawbacks of centralized approaches
  - Distributed nature of knowledge
  - Distribution of (legacy) information systems
  - Flexibility of knowledge-intensive processes
- The FRODO framework for distributed OMs applies agent technology on all levels:
  - Socially-enabled agents reflect the social aspects of knowledge
    - Rights and obligations
    - Sets of rights and obligations form role models
    - Agents can commit to roles. This leads to societies.
  - The platform allows easy creation of KM specialists
    - flexible creation of and cooperation between agents
    - individual agent behaviour enhances the systems's adaptiveness
    - configuration by specification of formal models
- EPOS aims at a better coupling with the knowledge worker's needs and exploits the individual knowledge work for a better maintenance of organizational knowledge structures.

Integration of different elements into a comprehensive Ontology Society Framework: Connection of „formal models" with active and responsible agents



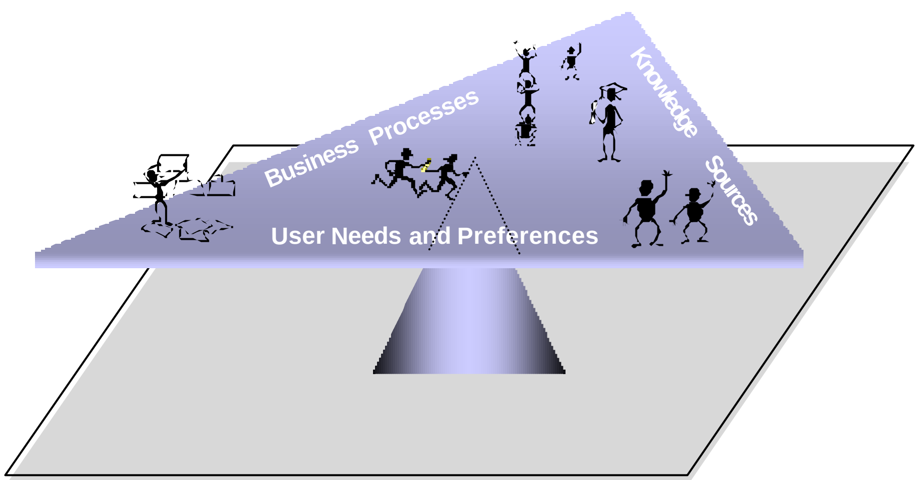
Overcoming the neglection of the social dimension of the concept of ontologies ("*shared*", "*commitment*") in today's IT praxis and research.

The vision of Agent-mediated Knowledge Management addresses all development levels of KM support systems

Organizational Analysis	Relevant actors/human agents, groups, tasks, competencies, etc. and their relations
System Architecture	Artificial agents, agencies, etc. and their relations, e.g., AUML
System Implementation	AgentClasses, Behaviors, e.g., JADE, JACK, ...

Methodological Support: CommonKADS, Tropos, Gaia

Agent-Mediated Knowledge Management



Make Societies of Agents Balance the “KM Seesaw”!

Thank you for your attention!

http://www.dfki.de/frodo/

Any Questions?

http://www.dfki.de/epos/

Contribute? More material?

Workshop Series on AMKM
http://www.dfki.uni-kl.de/~elst/AMKM
http://www.dfki.uni-kl.de/~elst/AMKM2004/

Springer LNAI 2926