Artificial Intelligence & Knowledge Management

Nick Bassiliades, Ioannis Vlahavas, Fotis Kokkoras
Aristotle University of Thessaloniki
Department of Informatics
Programming Languages and Software Engineering (PLaSE) Laboratory

What is Knowledge Management?

- Knowledge Management (**KM**) is the formal management of knowledge for facilitating creation, access, and reuse of knowledge, typically using advanced technology.
- KM systems contain numerous knowledge bases, made up of numeric and qualitative data (searchable Web pages, for example).

What is Knowledge Management?

- In addition, KM systems often allow discussion groups that focus on a single set of issues or a specific activity, such as particular software or a single consulting engagement.
- KM is a process of:
  - **converting** knowledge from the sources accessible to an organization and
  - **connecting** people with that knowledge.

Typical KM tools

- World Wide Web
- Internet και intranets
- Microsoft Backoffice
- Lotus Notes
- AI developments also play an important role in KM:
  - Intelligent Agents,
  - Knowledge Bases,
  - Knowledge Discovery in Databases (KDD),
  - Ontologies.

Why Knowledge Management?

- Organizations' environmental pressures
  - Increasingly competitive global marketplace.
  - Downsizing. Employees who remain are unaware of critical information resources.
  - Globalization and geographical dispersion.
- Technological advancements
  - Technology made available from Internet developments facilitate KM.
  - Advanced tools for independent Internet use have been developed (e.g. intelligent agents).
- Creating valuable information
  - KM systems help converting individually available knowledge into group or organizationally available knowledge.

Converting and Connecting

- Classic KM thinking assumes that a firm gathers all its important knowledge in a single place.
  - Employees use it to make good decisions that will benefit the organization.
- This classic approach ignores knowledge generated from data and text-based information.
- It also stresses connecting knowledge and people, but it doesn't link knowledge to other knowledge or push knowledge out to employees.

KM: Converting

Knowledge Source	Converting Activity	Example Knowledge
Individuals and groups	Knowledge harvesting and sharing	Operations, products, and sales
Data	Knowledge discovery (KDD)	Fraud or customer-service analysis
Text	Knowledge discovery (KDD)	News concerning management changes

Converting Data to Knowledge

- Knowledge discovery (KDD) is a new and rapidly evolving discipline
  - KDD uses many different techniques and tools from both AI and statistics to tease knowledge out of data warehouses and other sources
- KDD is the nontrivial extraction of implicit, previously unknown, and potentially useful information from data.
- KDD tools focus on turning data (e.g. financial, medical) into knowledge.

Converting Text to Knowledge

- Other important sources of knowledge include text-based information
  - e.g. news articles.
- A number of systems help users generate knowledge from text.
  - NewBot is an intelligent agent that help users to monitor over 100 different news-oriented sites (Business Week, USA Today).
  - ODIE (On Demand Information Extractor) is a system that reads over 1,000 news stories nightly to generate knowledge about management changes.

KM: Connecting

Connect	People	Knowledge
People to...	Emphasize personal networks and replace old networks.	Use "pull" technology. Search engines and Intelligent Agents.
Knowledge to...	Use "push" technology to send knowledge directly to people.	Focus on content instead of document information and link knowledge with multiple paths.

Connecting People to People

- The greatest source of knowledge is other people.
- Human-resource capabilities are being captured on homepages of organizational Intranets.
  - e.g. knowledge base "Knowledge On-Line 2.0" (Booz Allen)
- Intelligent agents can also be used to connect people.
  - ContactFinder is a proactive intelligent agent that reads and responds to electronic bulletin-board messages.
  - Rather than trying to solve users' problems, ContactFinder offers assistance by referring people to others who might be able to solve their problems
  - When it does have a referral, it posts its knowledge as if it were another user

Connecting People to Knowledge (1/2)

- KM converting can result in very large sets of knowledge to navigate.
  - Ford Motor Company is reported to have over 30,000 pages available to its users.
- In addition to search engines, some firms are generating unique tools to help connect users and information.
  - Coopers & Lybrand is building "Michelin Guides" into its knowledge bases to help people understand what is available and where it is located.
  - These guides summarize available knowledge and suggest other places with the same or similar information.

Connecting People to Knowledge (2/2)

- Intelligent agents offer another approach.
  - InfoFinder is an intelligent agent that learns about a user's information interests in a document repository (Lotus Notes).
  - InfoFinder generates a user profile based on sample documents that the user investigates while browsing.
  - InfoFinder learns profiles from documents and guides the user to other new and existing documents in the repository.

Connecting Knowledge to People

- Classic (passive) KM systems assume that people ***pull*** knowledge from the system.
  - Result: unfound or unused knowledge and unsolved problems.
- Alternative solution: ***push*** the knowledge to the user.
  - PointCast brings news, weather, and sports directly to a user's desktop.
- The ***push*** approach can be used with a ***proactive*** KM.
  - Advantages: There is a larger chance that knowledge will be found and used with less effort.
  - Disadvantages: Higher administrative costs and less security.

Connecting Knowledge to Knowledge

- KM systems have a document-oriented focus.
  - Documents do not establish the content of a topic in the KB.
  - Content comes from facilitating links between different documents using (hypertext).
- Users don't all have the same views and thus don't connect knowledge in the same way.
  - KM systems must accommodate these multiple views.
  - Coopers & Lybrand system offers several routes (business geography maps) to each destination because its 70,000 employees don't all look for the same information in the same way.

Knowledge Bases

- KM Systems (**KMS**) employ a wide range of Knowledge Bases (**KB**)
  - Proposals - engagements
  - Best-practices. These KBs capture information and knowledge about the best way to do things (business processes).
- In order to use Knowledge Bases effectively, the consulting firms must be able to generate ontologies that allow users to pinpoint what resources they need and want.

Ontologies for Knowledge Management

- Ontologies are explicit specifications of conceptualizations
  - They describe a taxonomy of the tasks that define the knowledge.
  - Within the context of knowledge-management systems, ontologies are specifications of discourse in the form of a shared vocabulary.
- Ontologies are needed in knowledge management.
  - Ontologies give the ability to filter information based on content.
  - Users are facilitated in locating an appropriate discussion group for either raising or responding to an issue.
  - Searching for specific information in the company's Intranet web-pages is facilitated by providing filtering capabilities based on content and not just keyword matching.

AI & KM
Intelligent Assistant Systems

- An important AI goal has been to build knowledge-based systems that solve challenging problems on their own.
- A new direction for AI is ***intelligent assistants***, which cooperate with a human user in solving a problem.
- In order to solve important problems:
  - it is often better to let the computer work out what might be done,
  - but to let the human user decide,
  - thus distinguishing workload versus decision competence and responsibility.

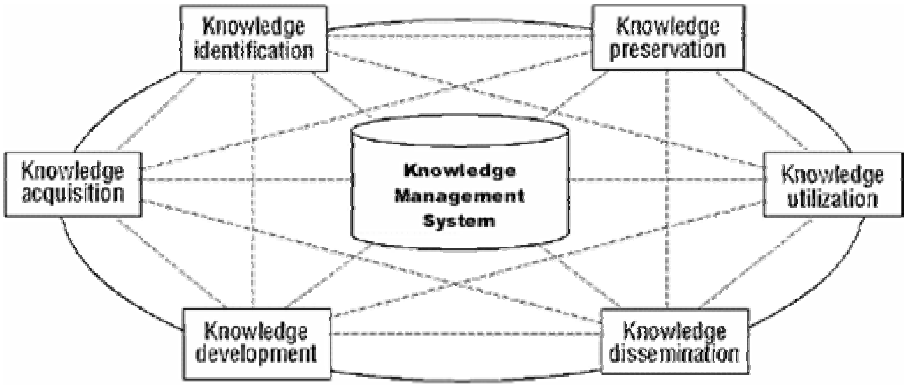
IA > AI

- IA (Intelligence Amplification)
  - Combination of a machine and a human mind
- AI (Artificial Intelligence)
  - An imitation of human intelligence by a machine
- The cooperation of a machine with an intelligent mind can beat an AI system that imitates human intelligence, working by itself.

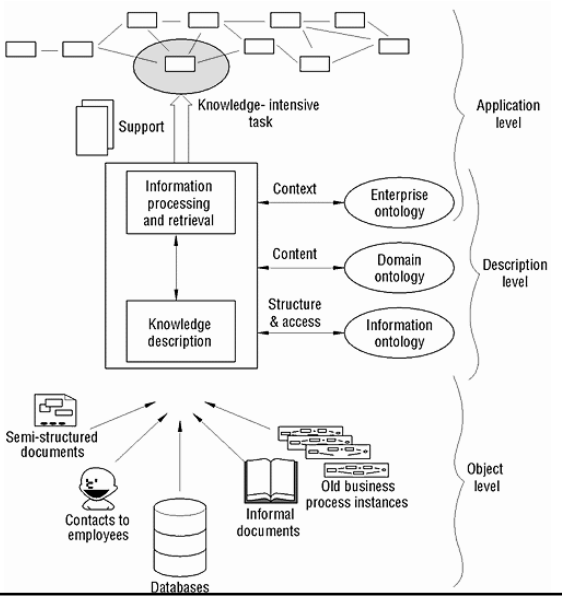
Knowledge Management Systems

- A KMS is not just a Management Information System (MIS)
  - It must help users and managers to convert information-knowledge to action (decisions)
- Activities:
  - Identification
  - Acquisition
  - Development
  - Dissemination
  - Use
  - Preservation
    - Explication of tacit knowledge (e.g. rule-based expert systems)

KM: Activities



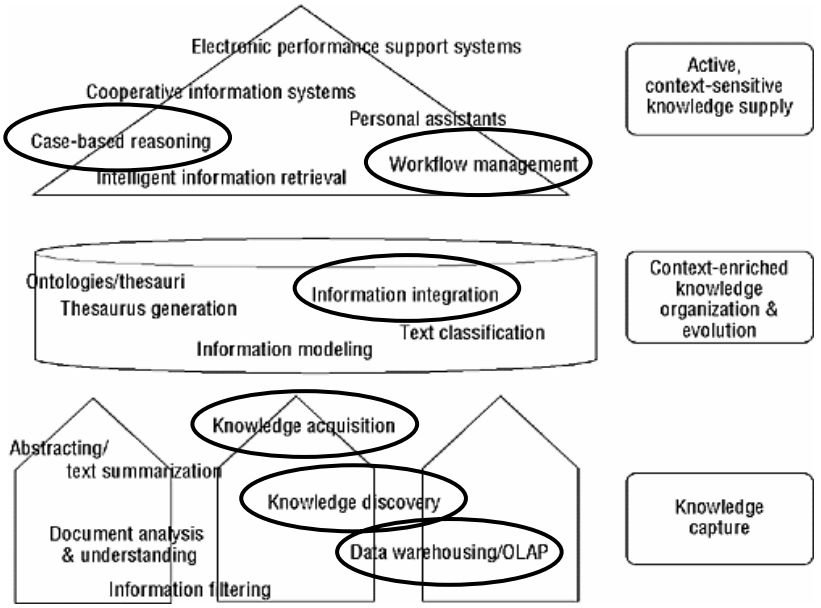
KMS Architecture



Business Intelligent Solutions

- BIS is a set of software products for:
  - Visualizing information
  - Complex report generation
  - Data mining
  - Query formulation
  - E-commerce applications
  - Integrating heterogeneous data (data warehouses, portals)
- All the above services are offered in a secured manner over the Internet, Intranets και Extranets.
- It has been used by various governmental organizations:
  - Financial, for supporting taxpayers
  - Traffic engineering, for reducing road traffic and accident risks
  - Housing, for consulting over loans
  - Ministry of Education, for monitoring and improving educational programmes

Research Areas



Knowledge Management Research at PLaSE Lab

- InterBase^{KB}
  - A data/knowledge warehouse system
- COMFRESH
  - A knowledge-based hypertext system
- SVT
  - A system for intelligent video data retrieval
- KOD
  - An adaptive, personalized educational environment

Technologies for Information and Knowledge Sharing

Information Sharing

- Information sharing is only possible with the help of 3 different approaches:
  - syntactical, structural and semantical
- The Internet has a most challenging problem, it's size.
  - One way to cope with this problem is the use of support technology for browsing, searching and filtering of information.
  - In order to support access to this information, "freedom" has to be reduced by providing machine-readable and/or machine understandable information about the content of a web page.

HTML
Visualizing Information

- Creating a web page on the Internet using the hypertext markup language (HTML) is the most frequently used technique for sharing information.
- HTML does not refer to the content of the information provided, but only covers the way it should be presented on the page.
  - HTML was created to make information processable by machines, but not understandable.

XML
Exchanging Information

- XML was proposed to overcome the purely visualization-oriented annotation provided by HTML
- XML is an extensible language allowing the user to define his own tags in order to indicate the type of content annotated by the tag.
- XML documents have associated type definitions which constrain the logical structure of the documents.
- An XML schema defines the structure of data and provides no information about the content or the potential use for others.

RDF
Representing Meta-data

- Meta information systems provide information about the location where information about a specific subject can be found
  - They can assist in linking Web pages based on their content (semantic translation)
- The RDF standard is a data model for representing meta-data about web pages and their content using XML syntax.
  - RDF actually borrows from frame systems well known from the area of knowledge representation in AI.

Semantic Modeling

- Neither XML nor RDF provide sufficient support for handling the structure and meaning of a document in order to integrate heterogeneous information.
- **Solution**: Middleware (Mediators) and Ontologies
- The middleware breaks the classical client-server architecture by introducing a layer between the server and the client.
  - A middleware remove functionality from the client and/or server in order to share the functionality or information to several other components.
  - The heterogeneities between the servers can be made transparent
- Example of middleware components are
  - Application servers
  - CORBA
  - Mediators (and wrappers)

Mediators

- Mediators provide read access to shared information stored in different heterogeneous information sources.
- One important advantage of mediators is that the knowledge how to access distributed information is removed from the clients.
  - The client "sees" a mediator as a “virtual information source”.
- The underlying information sources can be:
  - Database systems
  - Sources with unstructured or semi-structured information (i.e. Web pages)
- Mediators are able to solve syntactical and structural problems but not semantical problems.

Ontologies for Semantic Modeling

- Ontologies have set out to overcome the problem of implicit and hidden knowledge by making the conceptualization of a domain explicit.
  - Ontology is used to make assumptions about the meaning of a specific term.
  - It can also be seen as an explication of the context for which a term is normally used.
- Ontologies can be represented in:
  - Purely informal natural language description of a term, corresponding to a glossary
  - Strictly formal approaches (e.g. first order predicate logic)

Ontology Modeling Languages

- Ontology Interchange Layer (OIL)
  - Frame-based modeling features,
  - Reasoning facilities from description logic,
  - RDF-and XML-Schemes.
- DARPA Agents Markup Language (DAML), which describes to a browser the meaning of the information contained in a Web page
  - Extension of simple ontologies for local use,
  - Explicit representation of services, processes and business models.
  - Bottom-up design of meaning and sharing of higher-level concepts.
  - Semantic interoperability at the level we currently have syntactic interoperability in XML.