KNOWLEDGE TECHNOLOGIES FOR NETWORK ORGANISATIONS

Mitja Jermol
Jozef Stefan INstitute
OUTLINE OF THE TALK

• Basic concepts
  – Networked organisations
  – Knowledge management
  – Knowledge technologies

• KT for NO: theoretical background

• Examples of using KT in NOI
  – Knowledge discovery
  – Knowledge mapping
  – Knowledge sharing: PKT
  – Knowledge formalisation: ontologies, Cyc
  – Knowledge use
ORGANISATION – BUSINESS PROCESS VIEW
NETWORKED ORGANISATIONS

Collaborative Networks (CN)

Collaborative Networked Organizations (CNO)

Virtual Laboratory (VL)

Virtual Organization (VO)

Virtual Enterprise (VE)

Professional Virtual Community (PVC)

Extended Enterprise

CNO Breeding Environment
WHAT WE UNDERSTAND BY CNO?

Network ...

Constituted by a variety of entities (e.g. organizations and people) that are:
- largely autonomous
- geographically distributed
- heterogeneous in terms of their: operating environment, culture, social capital and goals

Nevertheless these entities collaborate to better achieve common or compatible goals

The collaborative interactions are supported by a computer network.

Unlike other networks, in CNO collaboration is an intentional property that derives from the shared belief that together the network members can achieve goals that would not be possible or would have a higher cost if attempted by them individually.
SELF - ORGANISATION

Self-organised system
Self-directed system
Self-regulated system
Uncontroled active system

Organizer
Director
Regulator
Operator
Coordinator
Mediator
SELF-ORGANISATION – CHAORDIC SYSTEM

Complexity and Coherence

Limits to Growth

- E \(x\)
- NTE \(x\)
- FFE \(x\)

Complexity level \(x\)

- E \(x+1\)
- NTE \(x+1\)
- FC \(x\)

Complexity level \(x+1\)

Time
Knowledge technologies

• …are information technologies including data mining, machine learning, decision support, language technologies, knowledge sharing tools and other information technologies that support acquisition, retrieval, reuse, share, maintenance and modeling of knowledge.

• …adds a layer of intelligence to information technology, to filter appropriate information and deliver it when it is needed.

• …provide a means for computers to understand the knowledge that is being presented by knowledge integrated into documents and services on the Internet (Semantic Web).

• Among knowledge technologies are knowledge extraction tools, ontologies, knowledge representation formalisms, intelligent topic maps, blogs, groupware, document management, expertise locators, latent semantic analysis, semantic networks, social networking engines, wikis,…
• KM is an interdisciplinary business model dealing with all aspects of knowledge within the context of the firm, including knowledge creation, codification, sharing, and how these activities promote learning and innovation. In practice, KM encompasses both technological tools and organizational routines in overlapping parts.

http://sims.berkeley.edu
KM IN RELATION TO OTHER SYSTEMS

- Vision, strategy management
- Organisation, culture, models
- Tools, methods, production, IS

Strategic level
Organisational level
Operative level

ERP | CRM | SCM | FMS | TQM
HOLISTIC VIEW ON KM

Self-organisational

Society level

Corporate level

Resource level

Physical level

Predetermined

© Mitja Jermol

mitja.jermol@ijs.si
The Knowledge Process

Knowledge Generation
Build Knowledge Base through Research, Adaptation, Discovery, & Experience

Knowledge Need & Use
Implement & Adjust Knowledge to Meet Customers' Needs

Knowledge Transfer
Produce Documentation & People that Will Facilitate Knowledge Delivery

Knowledge Development
Transform Raw Knowledge into Codified Principles & Practices
Knowledge management in a real business environment

- Analysing, modeling, managing distributed networked organisations
  - Industry clusters, virtual organisations, virtual communities, living labs

- Knowledge discovery and knowledge storage
  - Competence directories, Knowledge mapping, K-bases, media and methods repositories

- Training and knowledge transfer
  - ICT supported training, virtual communities, distance learning, personalisation
SEVEN SCHOOLS ON KM

- Systems school – **codifying knowledge in databases**
- Cartographic school – **mapping knowledge**
- Engineering school – **process oriented**
- Commercial school – **intellectual assets, IPR**
- Organizational school – **knowledge sharing across communities**
- Spatial school – **work and relaxation spaces**
- Strategic school – **companies as a knowledge businesses**
ECOLEAD

KNOWLEDGE

Self-organisational

Tacit

Society level
Business level
Management level
Engineering level
Operator level
Physical level

Encultured knowledge (society)
Embrained knowledge (holographic)
Embodied knowledge (genetic)
Symbol type knowledge (bit)

Predefined

Explicit

Complexity
RESOURCE LEVEL

“soft”

Wisdom

Knowledge

Content & info

Tools, techniques

Commun. & media

Complexity

“hard”

Embodied

Actional

Situational

Relational

HtH

HtM

MtH

MtM

mitja.jermol@ijs.si
"Knowledge Management caters to the critical issues of organizational adaptation, survival and competence in face of increasingly discontinuous environmental change. Essentially, it embodies organizational processes that seek synergistic combination of data and information processing capacity of information technologies, and the creative and innovative capacity of human beings."

Y. Malhotra at http://www.brint.com/km/whatis.htm
Radical discontinuous Change (wicked environment)

Organisational need for new know. creation and know. renewal

Information-processing model of KM

Sense-making model of KM

Guiding framework for KM

Convergence & Compliance (Lockean/Leibnizian)

Tight Provides efficiencies

Loose Provides agility & flexibility

Optimisation-driven & Efficiency oriented

Knowledge creation & knowledge renewal

Divergent meanings (Hegelian/Kantian)

A few rules and a lot of freedom

Management of a VE
NEW PARADIGMS IN KM

- From K-resource to K-process
- From deterministic to complex – “A few rules and a lot of freedom”
- From cognitive factors to social factors – knowledge networks and social networks
- From know-how to creativity and innovation
  - Creativity: the crazy things that work
  - Innovation is the conversion of knowledge and ideas into a benefit.
KT FOR NO:
THEORETICAL BACKGROUND
KM IN NO

Organisation – rule based

Community – interest based

Individual

Knowledge of the structure
Business Knowledge
Organisational Culture
Operational Knowledge

Agent (Individual)
Agent (Individual)
Agent (Individual)
Agent (Individual)
Agent (Individual)

Group
Group
Group

Trust
Social network
Group dynamics

Values
Beliefs
Competences
Knowledge

Abilities – creativity and innovation
HIERARCHY OF BUSINESS PROCESSES

Value chain

- BP1
- BP2
- BP3

Business process level

- MP1
- MP2
- MP3

Microprocess process level

- AC1
- AC2
- AC3

Activity level

- Top-down modeling
- Bottom-up modeling

- experts
- data

© Mitja Jermol

ECOLEAD
BUSINESS PROCESS

Measuring

Disturbances (environment)

Data
Tools

Business process

Methods
Knowledge

Control

Assessment

Reference (set of goals)

Modeling

Out (set of results)

In
Business process

Analysing

Modeling

Simulations

Predictions

Trend detection
EXAMPLES OF USING KNOWLEDGE TECHNOLOGIES FOR NETWORKED ORGANISATIONS
TOOLS AND METHODS FOR K-PROCESSES

Knowledge storage and development (Content management)
tracking and workflow, reference materials, ontologies, semantics, BBS, yellow pages, content and media databases...

Knowledge transfer (Training)
(distance) learning, intelligent searching, portals, personalisation, machine learning, learning methods, didactics, pedagogy,…

Knowledge generation and acquisition
data analysis: x mining, statistic methods,…
business intelligence: scenario planning, game theory,…
content acquisition: digitalisation, classification, filtering, ranking,…
Modeling, decision support, expert systems,…

Knowledge use, customisation and resource networking
collaboration, community, teams, experts, virtual networks,…

mitja.jermol@ijs.si
(O'Leary 1998)

- **Converting individual to group knowledge**: knowledge sharing
- **Converting data to knowledge**: uncovering patterns in databases using data mining
- **Converting text to knowledge**: uncovering patterns in texts using text mining tools, Debating and negotiating meaning.
- **Connecting people to knowledge**: visualisation, knowledge mapping
- **Connecting knowledge to knowledge**: Agents, Ontologies and metadata, Boundary objects
- **Connecting people to people**: knowledge mapping, shared workspaces
- **Connecting knowledge to people**: agent systems, personalisation
• **Knowledge discovery**
  – Textgarden set of tools

• **Knowledge mapping**
  – Partner finding tool (IST-World)
  – Knowledge map (Document atlas)

• **Knowledge sharing**
  – Personalised Knowledge Transfer (videolectures.net)
  – Virtual Learning Communities (ECOLEAD VLC)

• **Knowledge formalisation (Ontologies)**
  – Ontogen
  – Standards
  – Cyc

• **Knowledge use**
  – Decision support systems (…intelligence, )
  – Reasoning mechanisms (CyC)
  – Knowledge in business process (ACTIVE, PROLIX, TENCOMPETENCE)
SOME CKM TOOLS
KNOWLEDGE DISCOVERY
WHAT IS TEXT-GARDEN?

• Text-Garden is a software library and collection of software tools for solving large scale tasks dealing with structured, semi-structured and unstructured data
  – …in particular, emphasis of functionality is on dealing with text

• It can be used in various ways covering research and applicative scenarios
  – used by several institutions such as CMU, BT, MSR, …
FUNCTIONALITY BLOCKS

Lexical text processing
(tokenization, stop-words, stemming, n-grams, Wordnet)

Unsupervised learning
(KMeans, Hierarchical-KMeans, OneClassSVM)

Semi-Supervised learning
(Uncertainty sampling)

Supervised learning
(SVM, Winnow, Perceptron, NBayes)

Dimensionality reduction
(LSI, PCA)

Visualization
(Graph based, Tiling, Density based, …)

Named Entity Extraction
(capitalization based)

Cross Correlation
(KCCA, matching text with other data)

Keyword Extraction
(contrast, centroid, taxonomy based)

Large Taxonomies
(dealing with DMOz, Medline)

Crawling Web and Search Eng.
(for large scale data acquisition)

Scalable Search (inverted index)
• Text Garden is almost entirely written in portable C++
  – …it compiles under Windows (Microsoft Visual C++, Borland C++) and Unix/Linux (GNU C)
  – …it runs under 32bit and 64bit platforms
  – …it consists of ~200,000 lines of code
Text-Garden has the following interfaces with the same API:

- C/C++ - through simplified DLL & native C++
- Java – through JNI
- .NET – e.g. accessible through C#, VB, …
- Matlab – through standard Matlab interface
- Python – through standard Python interface
- Mathematica, Prolog, R – in preparation

API has ~40 classes and ~250 functions
- …interfaces to the all above platforms are generated automatically from the master Text-Garden header file

Text-Garden is under LGPL licence
- It is available from www.textmining.net
KNOWLEDGE MAPPING
Knowledge Mapping (PROCESS) discovers:
- the constraints, assumptions, location, ownership, value and use of knowledge artifacts,
- agents (people, groups, objects) and their expertise,
- blocks to knowledge creation, and
- opportunities to leverage existing knowledge.

Knowledge Map (VISUALISATION TOOL) portrays:
- the sources, flows, constraints and sinks of explicit and tacit knowledge within an organization,
- relationships between knowledge stores and the dynamics.

Knowledge Repository (DATABASE):
- A model and a set of tools that covers formal and informal means of storing information of Knowledge Mapping

Knowledge Space (MODEL) describes:
- the dynamics of a knowledge evolution following the predescribed learning process
KNOWLEDGE MAPPING OVERVIEW (II)

- Knowledge mapping covers:
  - mind maps
  - concept maps
  - pattern languages
  - dialogue mapping
  - graphic facilitation
  - geographical information mapping (GIS)
  - quantitative charts and graphs
  - process and procedure flow charts
  - timelines
  - mapping public issues

- and all other forms of visual presentation of information:
  - semantic networks
  - relational diagrams
  - knowledge landscapes
  - dynamic knowledge fields
  - ...
IMPLEMENTATIONS

• IT tools to support knowledge management
  – helpdesk and customer service
  – data & text mining with data marts, OLAP and other backups
  – intranets
  – conversation servers
  – groupware and workflow
  – virtual communities, psychographics and e-commerce
  – search engines,
  – text & data visualization

• Business intelligence, push and scanning applications

• Training in systems thinking, creative thinking, on-line facilitation

• Specialized knowledge engineering services:
  – corporate memory,
  – heuristics,
  – expert systems,
  – ontology development,
  – concept extraction,
  – knowledge structuring,
  – patterns,
  – communities of practice.
AGENTS’ KNOWLEDGE

- **sociological data**: group and team effectiveness, sociological type, network intelligence NQ,…
- **psychological data**: personality type, creativity level, basic motivation, values and beliefs,…
- **partner’s cultural and organization knowledge**
- **group and team knowledge** (functional, group dynamics,…)
- **partner/group/individual background knowledge** (core, additional, specific), learning styles and communication characteristics
- **partner/group/individual connections** to other’s knowledge
- **process’s knowledge and object type knowledge** (tools, methods, techniques performances)
- **personal lessons learned**
KNOWLEDGE MAP (E-R MODEL)

Availability
- Motivation
- Learning style
- Personality
- Society
- Language
- Commun. Character.
- Learning Tool & media
- Socio-psychologic

Network
- Group
- Subject
- Object

Agent
- BGK
- Explicit Knowledge
- Tacit Knowledge
- Atom of Knowledge

Context
- KEV
- Background
- PEV

Past performance
- Project
- K-Process
- Cost
- IPR
- Contact

Knowledge
- Strategy
- Method
- Domain
- Term

S_A
S_A
M
S_L
M
M
S_L
M
S_A
S_A
S_P
S_P
S_S
S_S
S_S
S_S
KNOwLEDGE MAP (I)

The following diagram is a visualization of your query results. The competence map indicates clusters of competencies. The darker the background color, the deeper the competence. When moving the round circle, more specific competencies appear. Red dots indicate locations of individual results within the whole map. Decrease the range of discovery by zooming in/out, using the bottom tools.
Dynamic fluid fields
K- (DEVELOPMENT) MAP – KNOWLEDGE SPACE

Dimension 3: Complexity

Estimated knowledge space after 500 trials

Actual knowledge space used by simulator module
VISUALISATIONS
DOCUMENT ATLAS

• Given a collection of documents (time stamps, if available)
  – eg., news articles, publication record

• Visualize in 2D:
  – content of documents,
  – named entities,
  – relationships (social networks),
  – changes over time

• Installation is publicly available in binaries at http://docatlas.ijs.si/
APPROACH DESCRIPTION

Latent Semantic Indexing

>100

Multidimensional scaling

>10,000
KNOWLEDGE SHARING – PERSONALISED KNOWLEDGE TRANSFER
ICT BASED TRAINING

• KT (learning) is one of the main processes in KM.
• KM stimulates and fosters continuous learning.
• Introducing
  – new methods, concepts and ways of learning
  – new didactic concepts, methodology, pedagogy,…
  – new tools, content types, media,…
• There is no target population (personalization):
  – all subjects
  – all levels (Life long learning)
  – all types of education
ICT BASED TRAINING
Training is a challenge, because it is not something you give; but a process based on guidance.

Personalization of methods, content, tools, guidance, control, motivation and goal.
PERSONALISATION (MOD SCOPOS)

Knowledge source (Teacher)

- Knowledge
  - Formulation
    - Articulation
    - Encoding
      - Understanding the Knowledge source
        - Transformation (from source)
          - Perception (from source)
            - Decoding
  - Tools, Methods Medium

Interface (Translator)

- Knowledge
  - Interpretation
    - Perception (from source)
      - Articulation (to recipient)
        - Encoding
  - Tools, Methods Medium

Knowledge recipient (Learner)

- Knowledge
  - Formulation
    - Interpretation
      - Perception
        - Decoding
  - Tools, Methods Medium

© Mitja Jermol
PERSONALISATION ATTRIBUTES MATCHING

PERSONALISATION

BACKGROUND

Knowledge Map

Knowledge Space

Competence Directory

Codes

© Mitja Jermol
CONTENT META STANDARDS

- **RDF** Resource Description Framework (W3C)  
  http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/

- **TEI** Text Encoding Initiative (SGML/XML)  
  http://www.tei-c.org/

- **DOI** Document Object Identifier (Avtorske pravice)  
  http://www.doi.org/

- **GILS** Government Information Service  
  http://www.gils.net/

- **PICS** Platform for Internet Content Selection  
  http://www.w3.org/PICS/

- **IMS** Learning Resource Meta-data  
  http://www.imsproject.org/metadata/

- **LTSC-LOM** (Learning Technology Standards Committee  
  Learning Objects Metadata)  
  http://ltsc.ieee.org/wg12/

- **DC** Dublin Core  
  http://dublincore.org/

© Mitja Jermol
ECOLEAD TRAINING

Active Virtual Learning Community
386 members
50 self learning courses
172 video seminars from 16 events,
100 authors, 89 lectures
122 prepared documents
2 training DVDs
VIDREOLECTURES.NET

Status 02.9.2007
128 events,
1536 authors,
2022 lectures,
2659 videos

App 2000 visits/day
150,000 visits since 1.3.2007
500,000 pageviews

638 visits from CZ
KNOWLEDGE FORMALISATION
STANDARDS – CONCEPTUAL


- CycL - [http://www.cyc.com/cycdoc/ref/cycl-syntax.html](http://www.cyc.com/cycdoc/ref/cycl-syntax.html), CycL is a formal language whose syntax derives from first-order predicate calculus (the language of formal logic) and from Lisp.

- DAML+OIL - [http://www.daml.org/](http://www.daml.org/) - is a semantic markup language for Web resources. It builds on earlier W3C standards such as RDF and RDF Schema, and extends these languages with richer modelling primitives. The use of DAML+OIL is illustrated here via an annotated example

- DQL- DAML query language - [www.daml.org/dql/](http://www.daml.org/dql/)

© Mitja Jermol
STANDARDS – FORMATS

- **KIF** - [logic.stanford.edu/kif/kif.html](http://logic.stanford.edu/kif/kif.html) - Knowledge Interchange Format is a language designed for use in the interchange of knowledge among disparate computer systems

- **OKBC** - [http://www.ai.sri.com/~okbc/](http://www.ai.sri.com/~okbc/) - Open Knowledge Base Connectivity is an application programming interface for accessing knowledge bases stored in knowledge representation systems - DARPA

- **RuleML** - [http://www.ruleml.org/](http://www.ruleml.org/) - Rule Markup Language permitting both forward (bottom-up) and backward (top-down) rules in XML for deduction, rewriting, and further inferential-transformational tasks
STANDARDS - INITIATIVES

- SHOE, [http://www.cs.umd.edu/projects/plus/SHOE/](http://www.cs.umd.edu/projects/plus/SHOE/), is a small extension to HTML which allows web page authors to annotate their web documents with machine-readable knowledge – transition to OWL

- SUO (IEEE P1600.1), [http://suo.ieee.org/](http://suo.ieee.org/), Standard Upper Ontology, is limited to concepts that are meta, generic, abstract and philosophical, and therefore are general enough to address (at a high level) a broad range of domain areas

THE SEMANTIC WEB LAYER

16-bit character set that assigns unique character codes to characters

Uniform Resource Identifier
SEMANTIC WEB LAYERS

• XML layer
  – Syntactic basis

• RDF layer
  – RDF basic data model for facts
  – RDF Schema simple ontology language

• Ontology layer
  – More expressive languages than RDF Schema
  – Current Web standard: OWL
WHAT IS AN ONTOLOGY?

- In the context of knowledge sharing, the term ontology means a "specification of a conceptualization”
- We usually refer to an ontology as being a graph/network structure consisting from:
  - Set of concepts (vertices in a graph)
    - Types of data
  - Set of instances assigned to a particular concepts (data records assigned to vertices in a graph)
    - Actual data records
  - Set of relationships connecting concepts (directed edges in a graph),
    - Labeled links between classes denoting relationship
WHY DEVELOP AN ONTOLOGY?

• To make domain assumptions explicit
  – Easier to change domain assumptions
  – Easier to understand and update legacy data

• To separate domain knowledge from operational knowledge
  – Re-use domain and operational knowledge separately

• A community reference for applications

• To share a consistent understanding of what information means
TYPES OF ONTOLOGIES

Describe very general concepts like space, time, event, which are independent of a particular problem or domain. It seems reasonable to have unified top-level ontologies for large communities of users.

Top-level ontology

Domain ontology

Task ontology

Application ontology

Describe the vocabulary related to a generic domain by specializing the concepts introduced in the top-level ontology.

Describe the vocabulary related to a generic task or activity by specializing the top-level ontologies.

These are the most specific ontologies. Concepts in application ontologies often correspond to roles played by domain entities while performing a certain activity.
ONTOLOGIES AND THEIR RELATIVES

Front-End

Thesauri
Topic Maps
Navigation
Information Retrieval
Sharing of Knowledge
Information Retrieval
Query Expansion
Queries
Ontologies
Mediation
EAI
Extended ER-Models
Consistency Checking
Semantic Networks
Query Expansion
Reasoning
Predicate Logic
Back-End
ONTOLOGY DEVELOPMENT

- Scenario 1 – Top-down: Expert building the ontology - manually
- Scenario 2 – Bottom-up: Ontology build from documents, email logs, etc…
“Ontology evolution” appears in a situation where the modeling domain is not fixed or known in advance.

Initially we start with an empty or “upper-level” ontology.

Dynamics in “ontology evolution” is defined as a set of transformation operations:
- Add & delete a class
- Add & delete a relationship

Decision which operation to take is taken manually or (semi) automatically to optimize compliance of the current ontological conceptual structure and the data to be modeled and collected so far.
ONTOLOGY MANAGEMENT

- Idea: Shared vocabulary (concepts, relations, axioms) of the various actors in a KM information system

- Ontology Management is an important means to balance between local and global concerns in Distributed Organizational Memory scenarios
WHY TO USE ONTOLOGIES IN NO?

• NO development and user community
  – E4 “Constitution” - a reference formal description of common understanding
  – a common vocabulary, rules and formulations
  – a background for an integrated architecture
  – multilinguality and multimodality support
  – …

• Benefits to the user
  – Solving inconsistencies - Reference vocabulary
  – Sharing - Enabling the connectivity to NO community on the level of business processes, resources and BOM
  – Open and integrated - Connectivity to existing ontologies that are user defined and/or relevant for NO
  – …
OVERVIEW OF EXISTING ONTOLOGIES

- **Terminological ontologies**: WordNet, VerbNet, FrameNet, Sensus
- **Domain ontologies**: The Gene Ontology (GO), PSL (Process Specification Language Ontology), The CEO project
- **Upper ontologies**: SUMO (Suggested Upper Merged Ontology), Mikrokosmos, OpenCyc, Sowa’s top-level ontology
- **Ontologies with common-sense knowledge**: Cyc, ConceptNet
- **Business-oriented Ontologies**: The AIAI enterprise ontology, The TOronto Virtual Enterprise’s ontology (TOVE), The BPMO (Business Process Management Ontology)
ADVANCED SERVICES (E4 CASE)

Knowledge discovery and decision support

Intelligent search
• Semi-Automatic
  – Text-mining methods provide suggestions and insights into the domain
  – The user can interact with parameters of text-mining methods
  – All the final decisions are taken by the user

• Data-Driven
  – Most of the aid provided by the system is based on some underlying data provided by the system
  – Instances are described by features extracted from the data (e.g. bag-of-words vectors)

Installation package is publicly available in binaries at ontogen.ijs.si
ADCT

ADC Telecommunications, Inc. is a global supplier of broadband network equipment, software and systems integration services that enable communications service providers to deliver high-speed Internet, data, video and voice services to consumers and businesses worldwide. The Company offers its products and services through two segments, broadband infrastructure and access and integrated solutions. The broadband infrastructure and access business focuses on broadband connectivity products for a variety of communications network applications, digital subscriber line offerings for the telecommunications industry and Internet protocol (IP)-based offerings for the cable industry. The integrated solutions business focuses on systems integration services and operations support system software.
ONTOLOGY MERGING AND TEXT/SERVICES ANNOTATION

• Merging ontologies
  – Common vocabulary
  – Controlling consistencies

• Annotation methods and tools
  – Annotation of documents and services
  – Automating the annotation process as much as possible
  – Annotation software architecture
SOME TECHNIQUES

• Automating the annotation process as much as possible
  – Annotation by comparison of documents
  – Hypotheses checking
  – Google distance
ontoclassify

• System for scalable classification of text into large topic ontologies

• Available as Web service
  – for DMoz directory of Web pages
  – for Inspec ontology for annotating papers
    http://alchemist.ijs.si/ASP.NETv1.1/InspecClassify/OntoService.asmx
AI is used on toy problems.

Expert systems reason in narrow domains.

Lenat, Minsky, Feigenbaum, Kay, and others recognize need for a substrate of shared world knowledge; and realize it would take hundreds of person-years to "prime the pump".

Lenat leaves Stanford and pursue this project (Cyc) within MCC, the US answer to Japanese 5th Generation Computer Proj.

Cycorp is formed.

Extensive development of knowledge-intensive applications.

Cyc is the only developed Knowledge Base of its kind in the world.

CycEurope is formed.

2006

2000

1994

1984

1983

1977

1967

1965

© Mitja Jermol
WHAT IS CYC?

• the Cyc Knowledge Base™
  – is a formalized representation of a vast quantity of fundamental human knowledge: facts, rules of thumb, and heuristics for reasoning about the objects and events of everyday life. The medium of representation is the formal language CycL

• natural-language processing
  – The Cyc-NL system has three components: the lexicon, the syntactic parser, and the semantic interpreter.

• Semantic Integration Bus™
  – Computer-based information is stored in many forms, including data that is structured (databases), semi-structured (spreadsheets, web pages), and unstructured (text files and text fields). Cyc can turn this information into usable knowledge, and the remainder can be annotated for easier access by humans.

• developer toolsets
  – interface tools to browse, edit, and extend the Cyc KB, to pose queries to the inference engine, and to interact with the natural-language and database integration modules.
Concepts: 276,747
Buying, Sulfur, Authoring, TerroristOrganization, AbuSayyaf, Witbrock

Predicates: 14,844
owns, isa, beliefs, foundingDate
(foundingDate AbuSayyaf ?X)
(maritalStatus YassirArafat Married)

Assertions: 2,204,623
“Causes precede effects”
(Forall x  (Forall y
(implies
(causes x y)
(startsAfterStartOf y x))))
Reasoning via deduction, abduction, induction; with access to other reasoning (through heuristic modules)

Growing ability to support Knowledge Acquisition and use via Natural Language

General-purpose platform: Supports apps. involving search, question answering, knowledge management, scenario generation, situation reporting, (re-)planning, etc.
WHAT CAN BE DONE WITH CYC?

• **Capabilities**
  – Integration of Heterogeneous Databases
  – Intelligent Search
    • Knowledge-Enhanced Retrieval of Captioned Information
    • WWW Information Retrieval
  – Distributed AI

• **Potential Applications**
  – Sophisticated modeling
  – Online brokering of goods and services
  – “Smart” interfaces
  – Intelligent character simulation for games
  – Enhanced virtual reality
  – Improved machine translation and speech recognition
  – Semantic data mining
  – Advice Services: E-shopping Assistant
KNOWLEDGE USE
DECISION SUPPORT SYSTEMS

- Business intelligence tools
- Cargo intelligence (Euridice)
- Project intelligence (www.ist-world.org)
SUPPLY CHAIN PREDICTION

- Consumption prediction
- Consumption planning
- Long-term planning
- Mid-term planning
- Short-term planning

Understanding the customer behavior, generating patterns and aggregated model

Prediction of the consumption before the order is generated

Distribution planning
Selling planning
Order planning
Stock management

Poraba, Napoved, 3447tank1

Datum | Liter
--- | ---
12.10.2000 | 6000
19.10.2000 | 5000
26.10.2000 | 4000
2.11.2000 | 5000
9.11.2000 | 6000
16.11.2000 | 7000
23.11.2000 | 8000
30.11.2000 | 9000
7.12.2000 | 10000
14.12.2000 | 11000
21.12.2000 | 12000
28.12.2000 | 13000
4.1.2001 | 14000
11.1.2001 | 15000
18.1.2001 | 16000
25.1.2001 | 17000
1.2.2001 | 18000
8.2.2001 | 19000
15.2.2001 | 20000
22.2.2001 | 21000
1.3.2001 | 22000
8.3.2001 | 23000
15.3.2001 | 24000

© Mitja Jermol
KNO W L E D G E I N B U S I N E S S P R O C E S S

• ACTIVE FP7 IP unleashing tacit knowledge from available document sources
• KaaS in COIN FP7 IP
• PROLIX http://www.prolixproject.org,
• TENCOMPETENCE http://www.tencompetence.org) - lifelong competence development