TMRA'05
Application framework
based on Topic Maps

Oct 6, 2005

Knowledge Synergy Inc.

Motomu Naito

motom@green.ocn.ne.jp

http://www.knowledge-synergy.com

National Institute of Informatics

Frederic Andres

andres@nii.ac.jp

http://ww.nii.ac.jp

We are

Motomu Naito with Knowledge Synergy, and Frederic Andres with NII, the national Institute of Informatics.

The NII was founded as an inter-university research institute organized to conduct comprehensive research on informatics and to develop an advanced infrastructure for disseminating scientific information.

Today, we are going to talk about Application Framework based on Topic Maps which deals with Subjects

Table of Contents

1. **Introduction - Background**

✓ Information related to Tsunami
  - ✓ Explosion of the amount of information with evolution of IT
  - ✓ Various information existing on various servers and clients

✓ Why a Search engine can not solve the Information Tsunami?

✓ Limits and Shortcoming of information processing based on character string and keyword
  - ✓ Access any time to required information
  - ✓ Processing enabled to understand the meaning of word
  - ✓ Systematization and organization of information

Introduction - Background (2)

✓ However, the "view", the "timing", and "granularity" vary according to the position, a specific situation, a man and etc.

✓ Recently, many technical elements, tools and software components can be used to solve this problem

1 Introduction - Purpose

This presentation discusses about the Application Framework which can be built using those technical elements, tools and software components and its high-efficiency and usefulness.

2. The Technical Elements

The technical element which can be used for to build the
Application Framework.

Topic Maps, RDF

Information resources structured semantically

Published Subjects

A mechanism identify subjects globally

Ontology

Definition and organization of concepts and relationship
between concepts

Remote Access Protocol

Protocol for exchange, merge and filter of fragment of Topic
Maps and RDF on network

Query Language

Language for query and update Topic Maps and RDF

2.1   Topic Maps and RDF

✓Information resources structured semantically

- ✓Topic Maps made by ISO, RDF made by W3C
- ✓Both Topic Maps and RDF consist of standard family
- ✓Similarities between them are obvious, but differentia between them are also obvious
- ✓Possibility to use both technology mutually complementary

Topic Maps and RDF (2)

- ***RDF/Topic Maps Interoperability Task Force*** established in SWBPD (Semantic Web Best Practices and Deployment Working Group) for the purpose of interoperability of Topic Maps and RDF
(http://www.w3.org/2001/sw/BestPractices/RDFTM/)

  - First working draft "RDFTM: Survey of Interoperability Proposals" was published on 29 March 2005 (http://www.w3.org/2001/sw/BestPractices/RDFTM/survey)
  - Coordinator of the task force is Steve Pepper
  - There are five proposals in the survey
  - RDF to Topic Maps mapping is practical

2.2   Published Subjects

Defined as a mechanism which enable person and computer to identify subjects (topics). And it is permanently published on networks and is aimed at making easy share/exchange Topic Maps.

- ✓ It is included in Topic Maps standard
- ✓ It allocate URI to subject, and make possible to identify subjects
- ✓ On the occasion of merge of Topic Maps, it is used to identify topics (subjects)
    - ✓ The topic which have same subject is merged
- ✓ It provides a subject identifying mechanism to Computer and Human
- ✓ It is applicable both information resource and things in real world
- ✓ Any one can publish published subjects
- ✓ A good thing will survive

Example of PSI Subject Dolphin

http://www.knowledge-synergy.com/PSI/dolphin

This is a published subject indicator (PSI) conforming to the OASIS Published Subjects Standard

Subject: Dolphin

PSID:

http://www.knowledge-synergy.com/PSI/dolphin

Definition:

Dolphins are small cetaceans that have a long, beaklike snout, a falcate (sickle-shaped) dorsal fin, and conical teeth. They are Odontoceti (toothed whales).

Kingdom Animalia, phylum Chordata, subphylum Vertebrata, class Mammalia, order Cetacea, family Delphinidae

Existing PSI

ISO 639 Published Subjects for language code
(http://www.oasis-open.org/committees/download.php/1444/language.xtm)

ISO 3166 Published Subjects for country code
(http://www.oasis-open.org/committees/download.php/1442/country.xtm)

XTM (XML Topic Maps) Core Published Subjects
(http://www.topicmaps.org/xtm/1.0/core.xtm)

2.3 Ontology

✓ Definition and organization of concepts and relationships between concepts

✓ Several classifications of ontology

✓ For example:

   1. Upper ontology
- ✓ limited to general, abstract concepts
- ✓ some attempts have been made at creating standardized upper ontology

   2. Domain ontology
- ✓ targeted certain domain, such as plant, medical care, law, etc.

   3. Task ontology
- ✓ targeted special processes to solve problems, such as diagnosis, design, study support, etc.

Example of Ontology

- ✓ Many ontologies already exist
  Example of domain Ontology
- ✓ UNSPSC (Universal Standard Products and Services Classification)
  data format Topic Map (http://www.techquila.com/tm-samples.html)
- ✓ SWEET (Semantic Web for Earth and Environmental Terminology)
  data format OWL (http://sweet.jpl.nasa.gov/index.html)
- ✓ ISO 12207 SLCP (Software Life Cycle Process)
  data format Paper
- ✓ ISO 19115 Geographic information -- Metadata
  data format PDF

- ✓ Many ontologies can be downloaded from TAP project site
  http://tap.stanford.edu/
- ✓ Many ontologies can be found by using Swoogle
  http://swoogle.umbc.edu/

They are expected to be published as Published Subjects
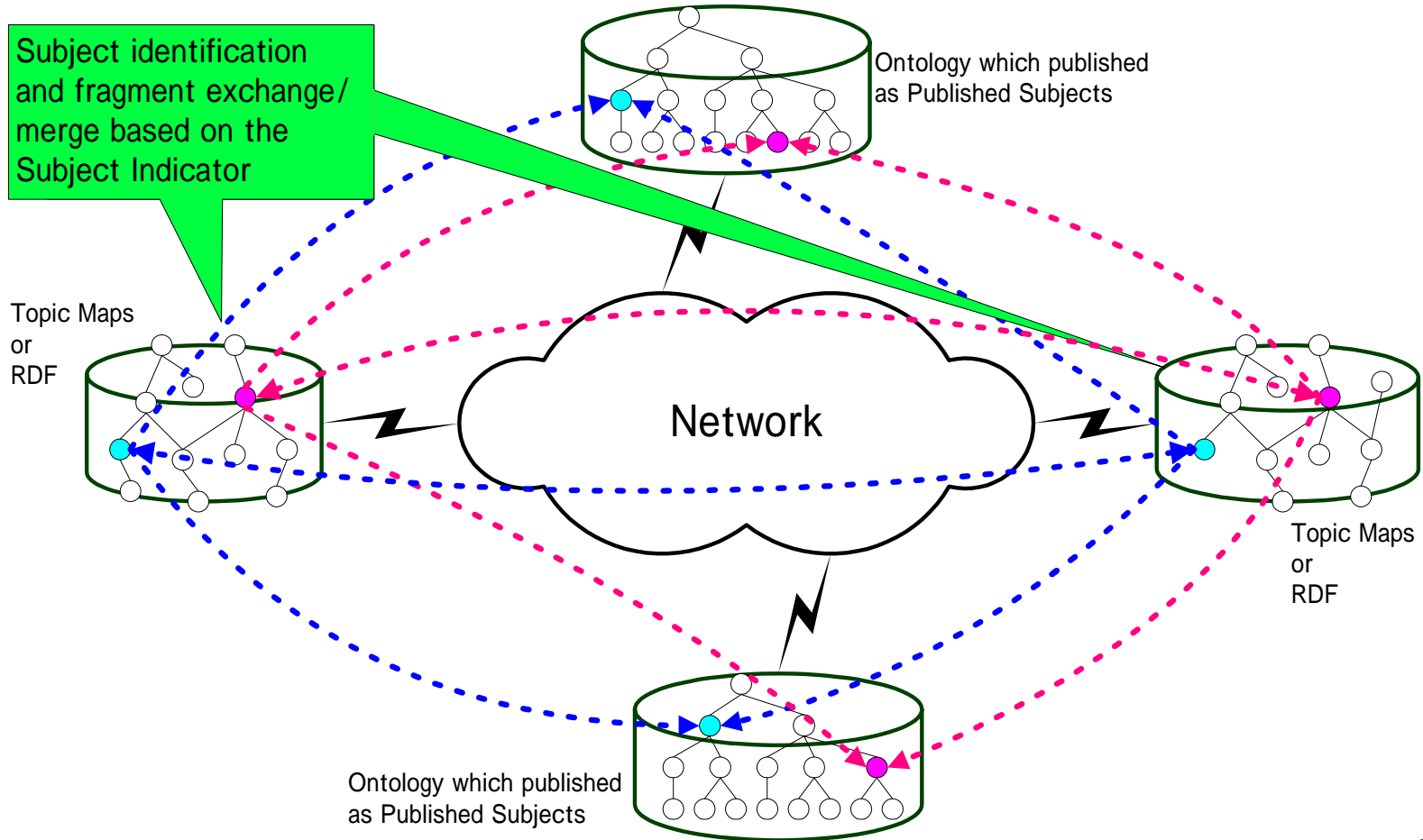
2.4   Remote Access Protocol

- ✓ Protocol for exchange, merge and filter of fragment of Topic Maps and RDF on network
- ✓ There are already many Topic Maps, RDF, PSI, Ontology, etc.
- ✓ The number of them is increasing rapidly
- ✓ They are located various place on network
- ✓ It is natural to feel the temptation to use those information resources
- ✓ There is no standard at this moment
- ✓ So standard is required
- ✓ There are some proposal
    - ✓ TMRAP (Ontopia As)
    - ✓ TopicMapster (techquila)
- ✓ Example of the Access Protocol
    - ✓ Select topic, select subject, select objects in scope, select whole Topic Maps
- ✓ Registration mechanism of PSI, Topic Maps, etc. such as UDDI registry seems to be needed

2.5 Query Language

✓ Querying and updating of Topic Maps and RDF
  ✓ TMQL by ISO and SPARQL by W3C will be available soon

✓ TMQL has three kinds of queries
  ✓ path expressions
  ✓ select queries
  ✓ FLWOR queries

✓ Get powerful way to find something and make diversified output
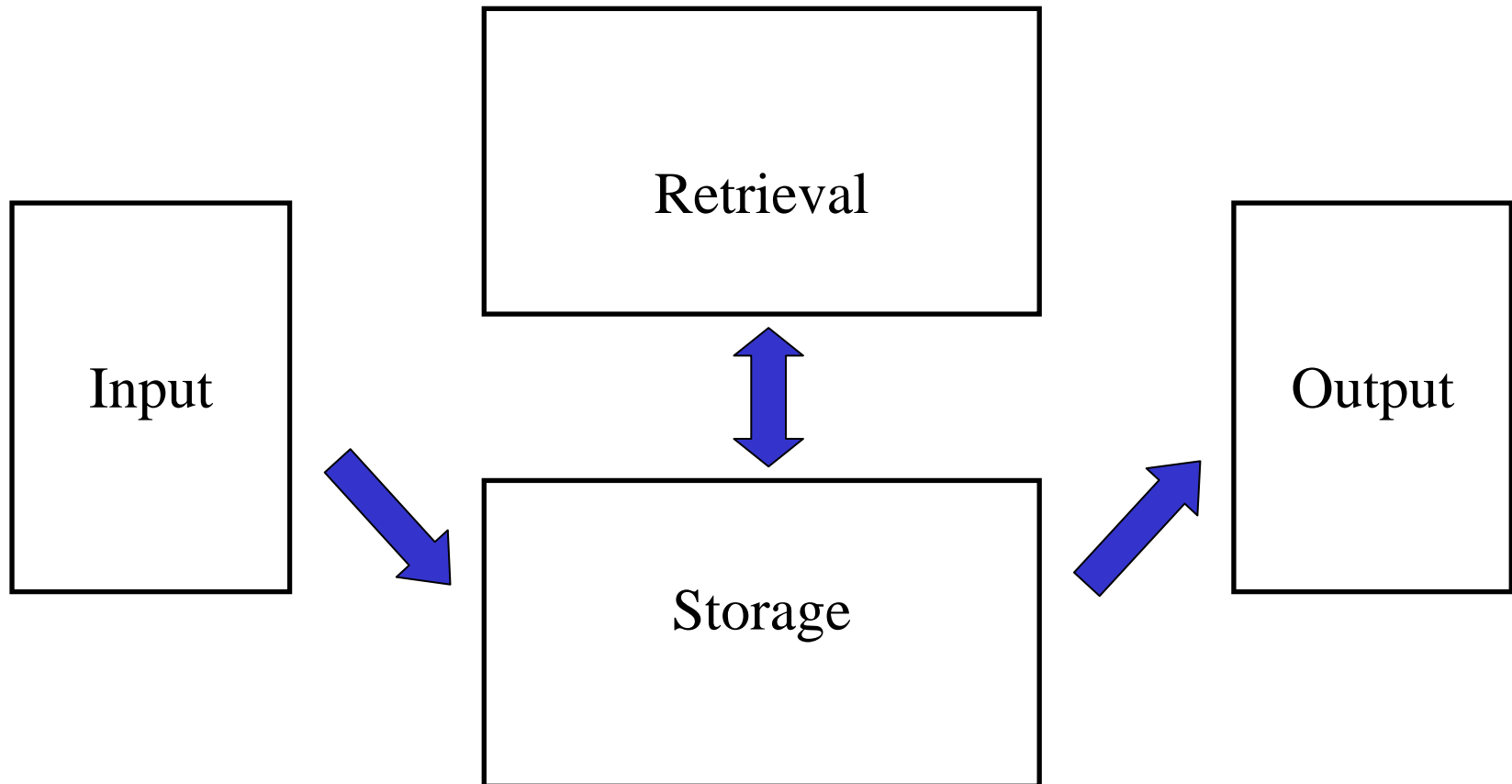
3. Application Framework

The application framework consists of the usable technology

(Information/Knowledge structure of Application Framework)

Subject identification and fragment exchange/ merge based on the Subject Indicator

Ontology which published as Published Subjects

Topic Maps or RDF

Network

Topic Maps or RDF

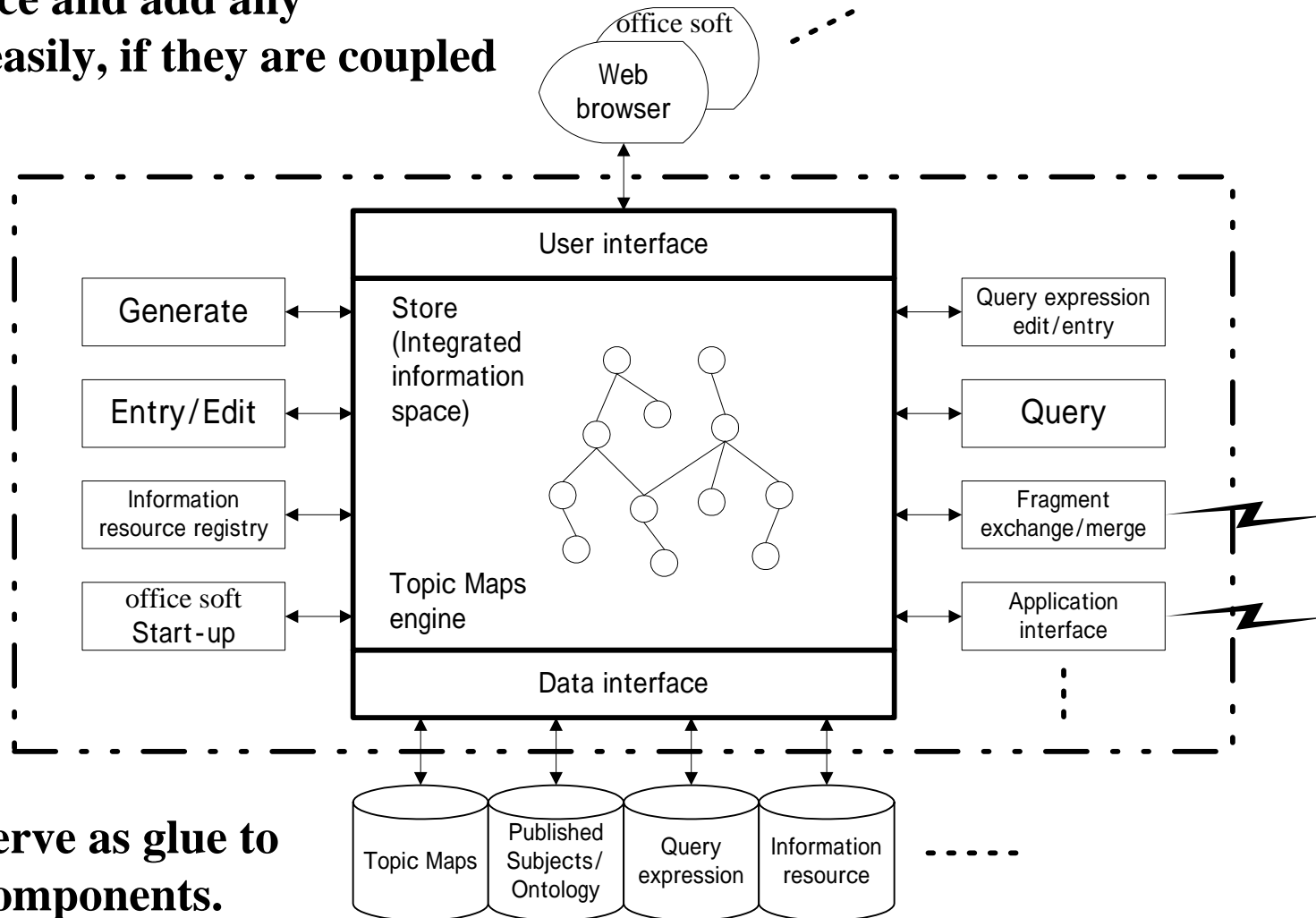Ontology which published as Published Subjects

Main function of Application Framework

Main functions of Application Framework are extremely simple

Structure of Application Framework

We can replace and add any components easily, if they are coupled loosely.

office soft

Web browser

User interface

Generate

Entry/Edit

Information resource registry

office soft

Start-up

Store (Integrated information space)

Topic Maps engine

Query expression edit/entry

Query

Fragment exchange/merge

Application interface

Data interface

Topic Maps	Published Subjects/ Ontology	Query expression	Information resource

he standards serve as glue to ombine those components.

Functional Components of the Framework

- ✓ Input function
  - ✓ Generation of Topic Maps
  - ✓ Entry/Edit of Topic Maps
  - ✓ Registry of information resources
- ✓ Store function
- ✓ Retrieve function
- ✓ Remote access (Fragment exchange) and merge function
- ✓ Output/display function
  - ✓ Translation from Topic Maps to HTML/XML
  - ✓ Translation from Topic Maps to PDF
  - ✓ Output of information resources
- ✓ More specialized and sophisticated components

Conditions of the Application Framework

✓ Compliance with standards (openness)
  ✓ De jure standards
    ✓ ISO/IEC 13250 Topic Maps
    ✓ (TMDM, XTM, Canonicalization, TMRM, CTM)
    ✓ ISO/IEC 18048 Topic Maps Query Language (TMQL)
    ✓ ISO/IEC 19756 Topic Maps Constraint Langauge (TMCL)
  ✓ De facto standards
    ✓ tolog
    ✓ TMAPI
  ✓ Need more standards

✓ Loose coupling structure
✓ Effective use of existing tools

Conditions of the Application Framework (2)

Application Framework needs good operational system, especially human

Both of buyable tools and open source tools are expected to have a large selection

4. Semantic Distance in Topic Maps

- The basic motivation
  - To add semantic scale inside topic maps structures
- Topic nodes semantically very close
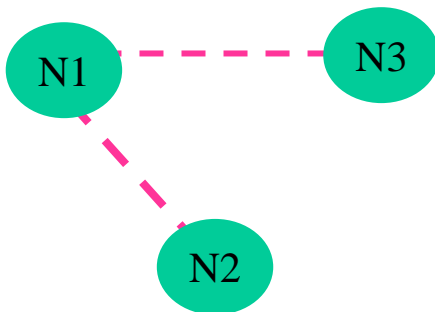
Example of Semantic Distance in Topic Maps
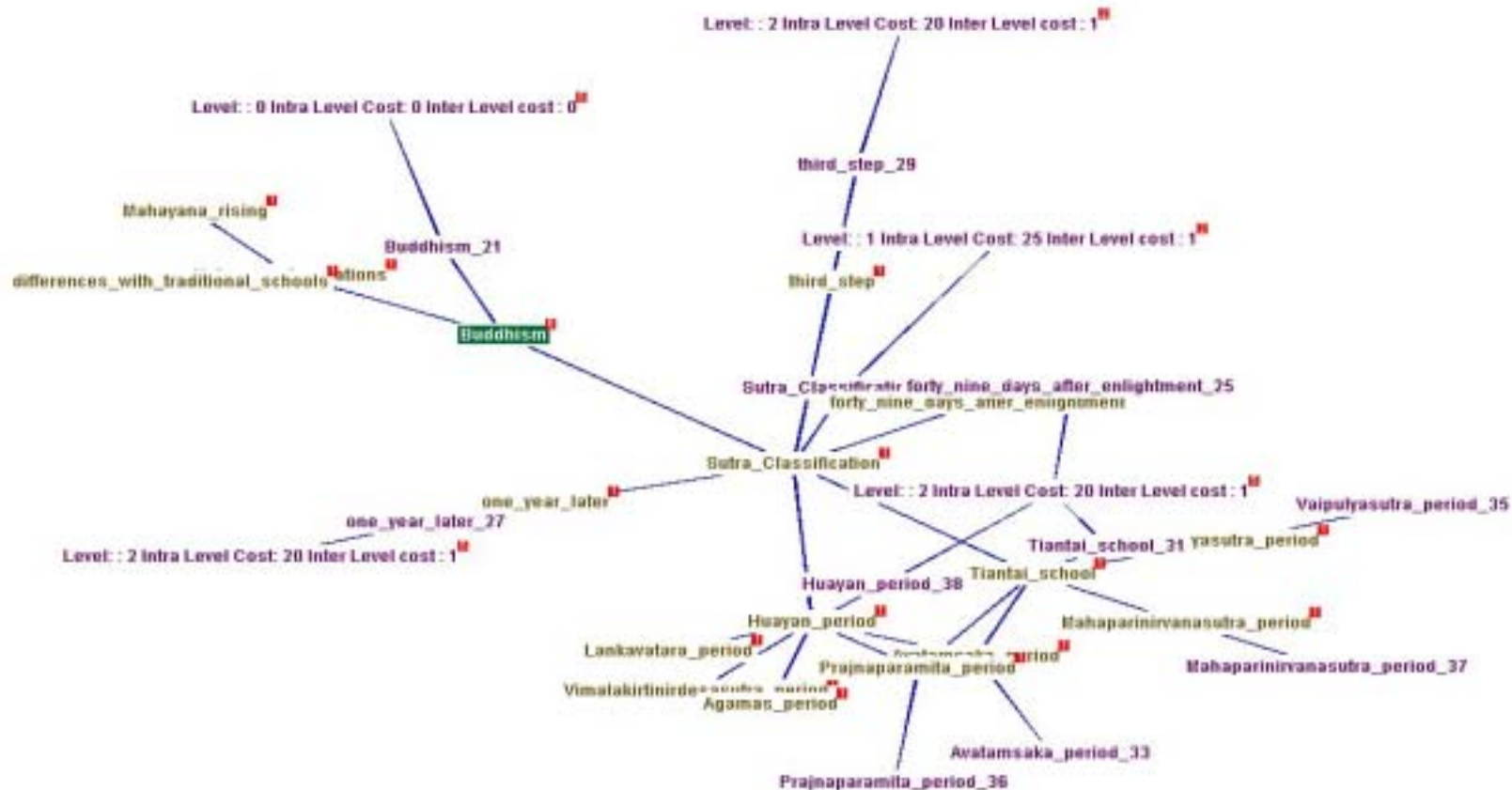
Semantic Distance in Topic Maps

- The semantic distance between two adjacent nodes $_{i,l}$ and $_{j,l-1}$ is defined by the following:

$S(\ _{i,l}\ ,\ \ _{j,l-1}) = D_{inter}\ * D_{intra}\ * W(\ _{i,l}\ ,\ \ _{j,l-1})$

Where $_{i,l}$ and $_{j,l-1}$ are the ith and the jth nodes located at levels l and l-1 of the target structure

Semantic Distance in Topic Maps

- D_{inter} = inter-level discrimination weight (Inter level cost) to maintain the semantic distance between two levels of a structure
  - D_{inter} = 1 in case of a graph

- D_{intra} = intra-level discrimination weight (Intra Level Cost) within a single level of a structure
- W = semantic path cost between two nodes $_{i,l}$ and $_{j,l-1}$

5. Challenge

Development of knowledge management environment for Software development & maintenance project

- ✓ Adoption of ISO/IEC 12207 Software Life Cycle Processes (SLCP) for main ontology (Topic Maps)
    - ✓ Work breakdown structure
    - ✓ Three level works (process, activity, task)
- ✓ We are creating various Topic Maps:
    - ✓ Person, Project, Document, Technology (skill), ….

Required Functional Components for the Environment

✓ Input/Edit or translate function for SLCP, person, project, etc

✓ Registry function for information resources

✓ Topic Maps engine

✓ Topic Maps storage

✓ Topic Maps query

✓ Display/output function for Topic Maps

✓ Display/output function for information resources

✓ More specialized and sophisticated functions

Expecting Use of the Environment

✓ Project management

✓ Process control (workflow management)

✓ Activities and tasks implementation support

✓ Content management

✓ Internal audit for ISO 9000

✓ Maturity evaluation
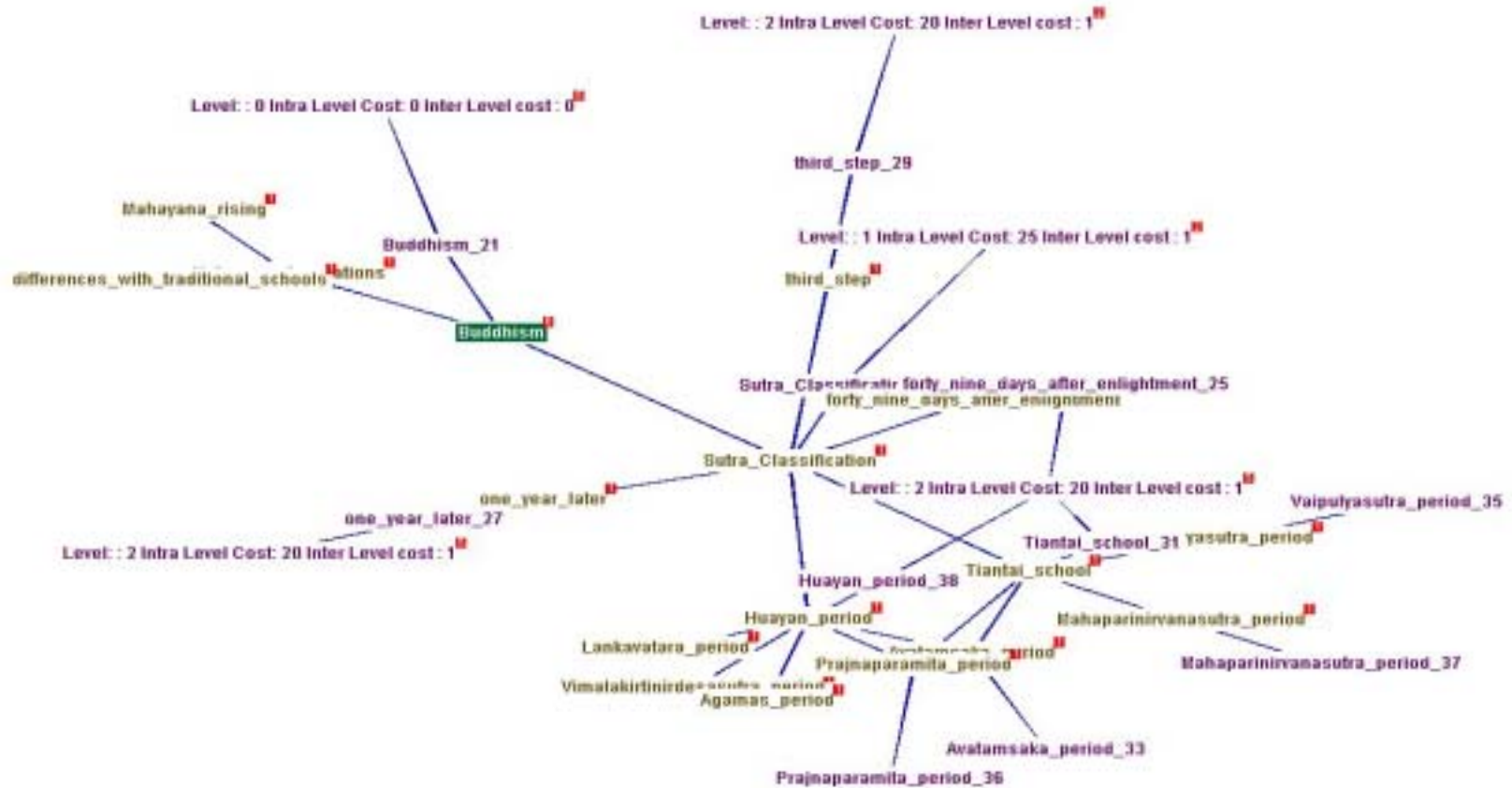
✓ Skill transfer

✓ Etc…

Demo

SLCP (Software Life Cycle Process) and Software document
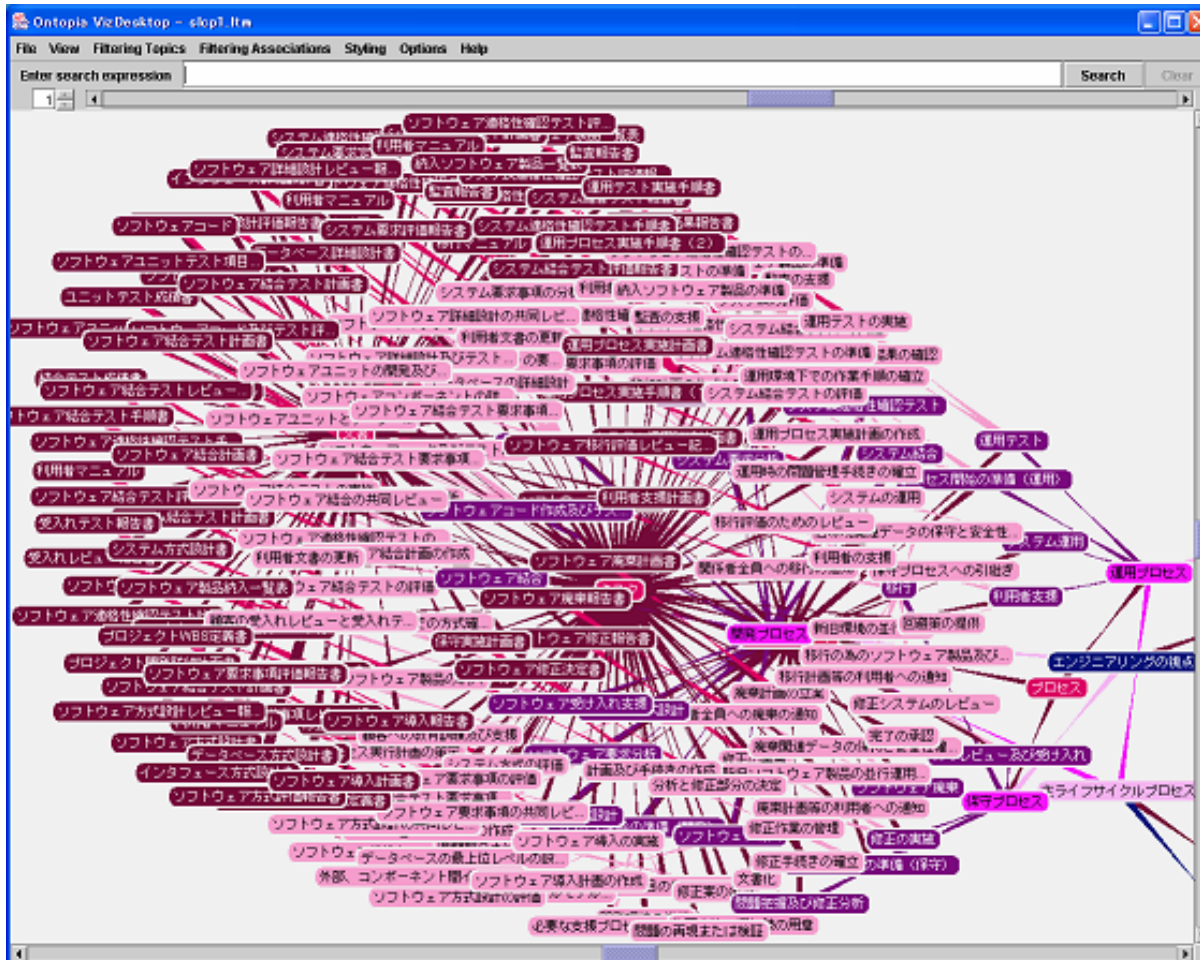
Graph representation of SLCP Topic Maps

28

Demo: Semantic Distance in Topic Maps

6. Conclusion & Future Work

✓ Improvement of findability and accessibility of information/ knowledge by making application based on Topic Maps

✓ Classification and organization of information/knowledge according to subjects and context

✓ High efficiency and cost performance to build application

   ✓ reuse and assemble of existing components

✓ Wide use for many application

✓ Further challenges:

   ✓ to develop more specialized and sophisticated components
     e.g. Semantic Distance Evaluation, Bayesian Network

Interesting Links

Event

INTAP Semantic Web conference

 2006.1.27 Keio University, Tokyo, Japan

 Web site

ISO SC34 WG3   www.isotopicmaps.org

topicmap.com   www.topicmap.com

Knowledge Synergy   www.knowledge-synergy.com

Ontopia AS   www.ontopia.net

TMAPI   www.tmapi.org

NetworkedPlanet   http://www.networkedplanet.com/

techquila   http://www.techquila.com/

Cerpus http://cerpus.com/

Jubik   http://www.jubik.com/en_index.html

 Free tools

Omnigator   www.ontopia.net

TM4J   tm4j.org

TM4L Editor, TM4L Browser

Thank you!