„We are drowning in information, but starving for knowledge. “

(John Naisbitt, *1929)
Towards Knowledge Discovery through Context Explication

Stefan Kremer¹ – Stefan Smolnik² – Lutz Kolbe¹

¹University of St. Gallen, Switzerland
²University of Paderborn, Germany

37th Annual Hawaii International Conference on System Sciences (HICSS-37)
Big Island, Hawaii, January, 05 – 08, 2004
Overview

- Motivation & challenge
- Delimitation of terms
- Knowledge discovery within portals
- The continuum of context explication
- Small cases
- Conclusions and future areas of research
Motivation & challenge

Dichotomy of deficits in the process of knowledge discovery:
- Lack of information
- Overload of information

Possible solutions:
- Control of semantic, i.e. meaning of terms
- Availability of explicated context
- Orderly classification of information objects

Comprehensive overview of several distinct solutions, their underlying principle, benefits and constraints

Selection criteria to support organizations to choose the appropriate technological solution
Delimitation of terms

- Knowledge: information and person-specific aspects like experiences, values, and insights
- The implicit and the explicit dimension of knowledge are complementary parts (Polanyi)
- Context: Any information that can be used to characterize the situation of an entity (Dey and Abowd)

- Focus: information objects and their contexts
- Goal: Providing approaches to discover, explicate, and use these contexts in different situations
Knowledge discovery: method for the identification or use of knowledge

Increase of transparency of knowledge in organizations

Support for users finding relevant information objects

Portals: web-based, personalized, and integrated access systems to internal and external applications and information repositories

Support of knowledge-oriented processes
Knowledge discovery within portals

Special challenges for knowledge discovery in portals
- Heterogeneous structures and formats
- Distributed and redundant data
- Protected content

Major approaches for knowledge discovery in portals
- Full-text searching
- Attribute-based searching
- Topic maps
Full-text searching

- Search query
- Indexing of each information object
- Representation of an information object by a set of weighted words

**Benefits**
- Speed
- Ease of use
- Pre-structuring unnecessary

**Constraints**
- ‘Wrong’ or too many results
- Spelling important

**Focus:** Content of an information object

**No explication (of context)**
Explication of an information objects’ context in metadata

Metadata are stored with the information object itself

Structured queries on context explicated in information objects’ metadata

Metadata standards, e.g. the Dublin Core Metadata Element Set
**Benefits**
- Reduced result set
- Controlled vocabulary
- Personalization capabilities

**Constraints**
- Maintenance of controlled vocabulary
- Metadata are stored with the information object itself
  - Need for re-classification

**Information objects contain content and explicated context**
- Maintained by the information object’s author at creation-time or during maintenance changes
Topic maps

- Model and architecture for the semantic structuring of link networks
- Associative network between subjects, which represent information objects
- Key concepts: topics, topic associations, and occurrences
Topic maps (contd.)

**Benefits**
- Creation of knowledge structures
- Creation of meta layers
- Discovery of new context
- Support of human thinking

**Constraints**
- Effort of topic map creation and maintenance
- New search paradigm

**Concepts and paradigms to discover and explicate contexts of information objects**

**Explicated context is not part of the information object**

**Explicated context can even be used without the information object**
Approaches focus on contextual information in different ways and varying intensity

Focus
- Data, information objects, and knowledge as basic subjects in portals
- Different embodied degree of context explication
- Approaches to find and use information objects and contextual information

Five approaches depending on a different degree of context and its ease of explication

Selection criteria and advice for choosing the appropriate solution, based on given pre-requisites
### Continuum of context explication

<table>
<thead>
<tr>
<th>Data approach</th>
<th>Information approach</th>
<th>Descriptor approach</th>
<th>Meta context approach</th>
<th>Knowledge approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>user</td>
<td>user</td>
<td>user</td>
<td>user</td>
<td>user</td>
</tr>
<tr>
<td>knowledge</td>
<td>knowledge</td>
<td>knowledge</td>
<td>knowledge</td>
<td>knowledge</td>
</tr>
<tr>
<td>worker</td>
<td>worker</td>
<td>worker</td>
<td>worker</td>
<td>worker</td>
</tr>
</tbody>
</table>

- **Large structured data sets**
- **Unstructured information objects**
- **Semi-structured information objects with descriptions**
- **Large sets of heterogeneous information resources**
- **Information objects in person-specific contexts**

---

**Data**

- **Context**
- **Content**

**Pattern identification**

- **Full-text searching**
- **Attribute-based searching**
- **Resource Description Framework**
- **Topic maps**
- **Action:**
  - Communication
  - Cognition
  - Construction

**Ease of context explication**

**Contexual Collaboration**
CoCE – Data approach

- Data are (meaningless) symbols
- Data are without context
- Transition of data to information objects or even to domain-specific knowledge
  - E.g. knowledge discovery in databases and data mining
    - Pattern identification

- Typical characteristics of a scenario
  - No or little interaction with users, authors, or knowledge workers
  - Large structured data sets
  - Possibly automatic data generation or collection
Simple information objects contain some kind of content

No explicit context like descriptors or other contextual information, but it inherently contains context

Context is interwoven with the content and difficult to conceptualize

Full-text search engines use indexed contents of information objects

Typical characteristics of a scenario

Many users, who are inexperienced with enhanced searching approaches or who are unwilling to use them

Authors are not trained in describing their information objects

Numerous unstructured information objects
Information objects are often enriched with metadata
- Content and explicit contextual information
- Dublin Core Metadata Element Set and the Resource Description Framework (RDF)
- Effort to enrich an information object with explicit contextual information
- Attribute-based searching

Typical characteristics of a scenario
- Authors are trained and skilled in describing their information objects
- Information objects contain descriptors
- Numerous semi-structured information objects
Subjects, concepts, or themes form contextual information in a meta layer above the information objects and are not necessarily explicitly stored within the information object.

- **Topic maps**
- **Discovery of new contexts** ➔ creation of new knowledge
- **Substantial effort has to be invested to define, create, and maintain a topic map**
Typical characteristics of a scenario

- Knowledge workers, who are familiar with topic map concepts and the domain of interest
- Manageable domains of interest
- Possibly existing taxonomies for the domains of interest
- Users are experienced in searching and navigating topic maps
- Large sets of possibly heterogeneous information repositories
CoCE – Knowledge approach

- Competencies, experiences, values, and insights, form a rich, person-specific context ➔ characteristic of the implicit dimension of knowledge and hardly to explicate

- Active involvement of users is a necessary precondition to transfer information into knowledge ➔ communication, construction, or more intrinsic cognition

Typical characteristics of a scenario

- Competencies, experiences, values, and insights
- Information objects in person-specific contexts
- Creation of knowledge by human actions, e.g. cognition of information objects
Small cases: IWI

- Combination of full-text and attribute-based searching
- Semi-structured information objects and manageable responsibilities for maintaining meta-data
- Most important descriptor: ‘topic’
- Taxonomy regarding ‘topic’
- 350 ‘topics’ classify about 11,000 information objects

Success factors

- Simplicity
- Mixture of centralization and decentralization
Small cases: IWI (contd.)

Database query

With the help of the following form, you can search our publication database:

Search for:
- knowledge
- this exact phrase

Search in:
- Complete text

Author:
- All authors

Year:
- From
- Until

Topic:
- Portal

Chair:
- All chairs

Competence centre:
- All competence centres

Submit search  Reset form

Fulltext query

Search results

Your search returned 7 results:

Riempp, G.: Wissensorientierte Portale

PDF 273 KB

Blessing, D., Riempp, G., Österle, H.: Entwicklungsstand und -perspektiven des Managements dokumentierten Wissens bei grossen Beratungsunternehmen
PDF 460 KB

Jansen, C.: Prozessunterstützung durch Wissensplattformen für Business Engineers
PDF 215 KB

PDF 144 KB

Österle, H., Bach, V., Schmid, R.: Mit Customer Relationship Management zum Prozessportal
PDF 1946 KB

PDF 292 KB
Small cases: GCC

- Applying topic maps to groupware-based organizational memories
- GCC K-Pool: a process-driven knowledge management system for contextual collaboration spanning intranet to internet
- Highly skilled researchers, who are familiar with the topic map concepts and who know all subjects of daily business
- Users and authors are supported by a lightly distinctive taxonomy
- Success factors
  - Understanding of work contexts
  - Low maintenance
Small cases: GCC (contd.)
Future areas of research

- Evaluation of the distinctness of situations for applying the continuum’s elements by adopting GCC’s topic map framework on IWI’s content
- Analysis if certain patterns for transition can be found, e.g. from topic maps to full-text searching or full-text searching to attribute-based searching
- Extension of the continuum with advices concerning the implicit dimension of knowledge by including the explication of skills and skill management
- Validation and expansion our findings with further external partners
Conclusions

- Three approaches of knowledge discovery in portals
- Continuum of context explication
- Selection criteria
- Two reference implementations

Knowledge discovery through context explication provides a comprehensive framework to support knowledge management processes productively.

HICSS-37 Best Paper Award Nomination 😊

Research colloquium May 2003:
Questions & discussion

Dr. Lutz Kolbe
University of St. Gallen
Institute of Information Management
Phone: +41--71--224-3796
Email: Lutz.Kolbe@unisg.ch
Web: http://www.iwi.unisg.ch
Web: http://ccckp.iwi.unisg.ch

Stefan Kremer
University of St. Gallen
Institute of Information Management
Phone: +41--71--224-3616
Email: Stefan.Kremer@unisg.ch
Web: http://www.iwi.unisg.ch
Web: http://ccckp.iwi.unisg.ch

Stefan Smolnik
University of Paderborn
Business Computing 2 – Information Management & Office Systems
Faculty of Business Administration, Business Computing & Economics
Phone: +49--5251--60-3375
Email: Stefan.Smolnik@notes.upb.de
Web: http://gcc.upb.de/StefanSmolnik