Conceptual Framework To User-Oriented Content Management

EJC’06, Trojanovice, May 30, 2006

May 30, 2006

Bernhard Thalheim
Technologie der Informationssysteme
Institut für Informatik, Christian-Albrechts-Universität zu Kiel, BRD

Kolmogorow-Professor e.h. der Lomonossow-Universität Moskau
Overview

- **Content and content management**
  The next hype, the next IT lie or the final solution?

- **The Kiel approach to content management**
  Information = Content + concepts + topics

- **Co-Design of structuring, functionality, interactivity, and distribution**
  Structuring: the classical and the non-classical case
  Functionality: the hidden programmer’s cave
  Interactivity: playout of scenarios, actors and interfacing
  Distribution: exchange of services based on media objects

- **The user dimension for content management**
  the most important part

- **User-oriented CMS**
  handling complexity

Maximal exploitation of database theory and technology for development of intelligent user-oriented content management systems
Content Management - All in one!

CMS is supposed to subsume:

- Portal management, Enterprize content management
- Content delivery, Agency solutions, Content Provider, Customer relationship management, E-Commerce solutions, E-Marketing, Online Payment
- Document management, archiving, and search, Support for documentation business processes
- Intelligent, user-specific generation of content
- ASP solutions, Media asset management
- Groupware solutions, Intranet solutions
- Redactional system, playout system, refreshment system
- Scalable solutions, Agent technologies, Performance monitoring, Security technology, High availability
- Open source solutions, Community solutions
CMS Support a Variety of Tasks

Managing web assets: variety of sources
   content from XML and databases
   Meta data creation and storage for services, e.g., search

Workflow: task support, complex business processes
   variety of protocols

Templates for entering content, for presentation

Source control and versioning: source code management capabilities,
   e.g.,
   versioning, merging changes, conflict resolution

Deployment and delivery services: deployment solutions, automated
   archival, expiration services, runtime delivery services

Management of distribution and adaptation: extraction, integration
deliverable to a large variety of customers.
Content $\approx$ complex and ready-to-use information

Content is structured variety (different user, environments, tasks)

Complex workflows and tasks

Distributed services

Adaptation and delivery

- integrate extraction, storage and delivery of complex structured objects,
- support workflows and tasks,
- based on service systems,
- deliver to users on demand and profile, at the right moment, and within the right format and size.

CMS $\approx$ information systems + extraction + intelligent storage + delivery
## Already Achievable Architecture for CMS

<table>
<thead>
<tr>
<th>Content Management System</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Story Space</strong></td>
</tr>
<tr>
<td>Stories</td>
</tr>
<tr>
<td>Scenarios</td>
</tr>
<tr>
<td><strong>Content types</strong></td>
</tr>
<tr>
<td>Structure</td>
</tr>
<tr>
<td>Service</td>
</tr>
<tr>
<td>Functionality</td>
</tr>
<tr>
<td>Container</td>
</tr>
<tr>
<td><strong>Structuring</strong></td>
</tr>
<tr>
<td>Structure</td>
</tr>
<tr>
<td>Static IC</td>
</tr>
<tr>
<td>(Pragmatics)</td>
</tr>
<tr>
<td><strong>Functionality</strong></td>
</tr>
<tr>
<td>Processes</td>
</tr>
<tr>
<td>Dynamic IC</td>
</tr>
<tr>
<td>(Pragmatics)</td>
</tr>
</tbody>
</table>
The Kiel Approach to Content Management

What do we need? INFORMATION!

Information as processed by humans, is data perceived or noticed, selected and organized by its receiver, because of his subjective human interests, originating from his instincts, feelings, experience, intuition, common sense, values, beliefs, personal knowledge, or wisdom simultaneously processed by his cognitive and mental processes, and seamlessly integrated in his recallable knowledge.

Knowledge is justified true belief and understanding

Skills as ability to do something well

T. S. Eliot (1888-1965), The rock, 1934:
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?

β nowadays:
Where is the information we have lost in news?
Where is the information we have lost in data?
The Kiel Approach to Content Management

Never solve all problems at the same time
Use differences for sophisticated management
Provide a theory that handles all aspects
Check, extend, integrate, purge existing technology
Provide a methodology that supports development of CMS
Challenge and proof the concepts in applications
Separation of Concerns Based on the Semiotic Triangle of Content, Concepts and Topics

- Computation theory
- Computation
- Syntax
- Semantical unit
- Semantics
- Infon
- Asset
- Pragmatics
- Content delivery
- Presentation
- Explanation
- Annotation
- Validation
- Model theory
- Interpretation
- Foundation
- Content
- Concept
- Topic
- Information

Content
Our Approach
Integrated Design
User Dimension
UO-CMS
Concluding
The Mappings of the Syntax, Semantics, and Pragmatics Dimensions

Content
Our Approach
Integrated Design
User Dimension
UO-CMS
Concluding

Computing and ETL of content

Derivation of concepts

Enrich, integrate topics

found by units
interpret through units
represent by infons
explain by infons
annotate through assets
deliver assets
Actor Profiles and Topics

Actors as groups of users with similar tasks, similar culture, similar behavior and similar interests

Actor specification frame

Actor profile: ⟨actor profile name⟩
Grouping criteria: ⟨characteristics of grouping of users⟩
Information demand: ⟨general description⟩
Utilisation pattern: ⟨general description⟩
Specific utilisation: ⟨general description⟩
Actor context: ⟨general description⟩

Topics are used by actors for annotating content and representing concepts
## Integrated Design: State Of The Art So Far

<table>
<thead>
<tr>
<th>Concept</th>
<th>Used in practice</th>
<th>Theoretical background</th>
<th>Earliest layer of specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structures</td>
<td>well done</td>
<td>well developed</td>
<td>strategic</td>
</tr>
<tr>
<td>Static semantics</td>
<td>partially used</td>
<td>well developed</td>
<td>conceptual</td>
</tr>
<tr>
<td>Processes</td>
<td>somehow done</td>
<td>parts and pieces</td>
<td>requirements</td>
</tr>
<tr>
<td>Dynamic semantics</td>
<td>some parts</td>
<td>parts and glimpses</td>
<td>implementation</td>
</tr>
<tr>
<td>Services</td>
<td>implementations</td>
<td>ad-hoc</td>
<td>implementation</td>
</tr>
<tr>
<td>Exchange frames</td>
<td>intentionally done</td>
<td>nothing</td>
<td>implementation</td>
</tr>
<tr>
<td>Interfaces</td>
<td>intuitive</td>
<td>nothing</td>
<td>implementation</td>
</tr>
<tr>
<td>Stories</td>
<td>intuitive</td>
<td>nothing</td>
<td>implementation</td>
</tr>
</tbody>
</table>

Late Specification, Inflexibility, and Unmaintainability

Extension, change management and integration become a nightmare
Languages: Extended ER + DistrLang + SiteLang

Structuring on the basis of an extended ER model
that is based on hierarchical predicate logic

Functionality on the basis of HERM/LC
with HERM-algebra, HERM/QBE, query forms and transactions
with some kinds of dynamic integrity constraints, behavior
GCS integrity enforcement instead of rule triggering pitfalls

Interactivity in integrated form based on SiteLang
description of dialogue scenes, stories, story space
(actors, scenario, dialogue steps)

Distribution through service specification and exchange frames

Translation and transformation methods for compilation of design into other models (logical, physical) or XML

Development and engineering methods for pragmatism (see my homepage)
Constructs of the Co-Design Languages

Structuring as pair

Structuring := (Structure, Static Constraints)

Structure as (marked) expression on constructors

Functionality as pair (Operations, Dynamic constraints)

Functionality := ((StateChange ∪ Retrieval)Operations, DynamicConstraints)

Operations on the basis of the HERM algebra (for modification and retrieval)
providing a language for generalized views (media types)

Distributivity as pair (Services, Exchange Frames)

Distribution := (Service (Informational process, Service Manager, Competence),
ExchangeFrame (Architecture Collaboration Style, CollaborationPattern))

Services on the basis of generalized views (media types)

Interactivity as 4-tuple

Interaction := (StorySpace, Actors, MediaObjects, Presentation)

StorySpace as graph of scenes and activities
Actors are abstractions of user groups
MediaObjects are used by actors and are based on generalized views (media types)
Scene in a Storyboard

Course Data Entry Scene Extended With Internal Negotiations

Chairs Lecture Proposal Scene

- Login by chairs responsible
- Generate new course proposal
- Collect seminar proposals
- Settle data for proposal
- Settle data for seminar
- Accept course demand

- Entry of necessary data
- Auxiliary & historic data
- Formulation of side conditions
- Negotiation of assignments by members
- Assignment of courses to members
- Submission chair data

Content
Our Approach
Integrated Design
User Dimension
UO-CMS
Concluding
Making Co-Design Working: Abstraction Design Layers

- Content
- Our Approach
- Integrated Design
- User Dimension
- UO-CMS
- Concluding
Practicality of the Co-Design Approach

Our Experience

- Internet and cable net information-intensive services
  - Cottbusnet and 35 other website projects
    Codesign of structuring, functionality, and interactivity
    Content-on-interest + Content-on-profile-and-portfolio + Delivery-on-context !!! + Adaptation-on-context !!!
  - Intelligent personalized internet in set-top-box-based TV
    Codesign of structuring, functionality, distribution and interactivity
    TV/Radio-on-interest + Cinema-on-demand + Internet-on-profile + Internal-communication-among-users

- Conceptual modeling for story spaces and story boards for services

Experience Gained by Partners

- Co-Design of structuring and functionality by Turkey System Inc.
  more than 1.500 applications

- Re-Design of SAP R/3
  Co-Design of structuring and functionality
Content Management as a Steel Bullet

**e - Learning** on the basis of the Da(ta)Mi(ning)T(utoring work-bench) project

**Theory of open learning units**

**Content** structuring, functionality, distributivity and interactivity

**Logistics of content playout** through adaptivity to context and user

http://damit.dfki.de

---

Open Problems and an Invitation to Join Research

The triad and the integrated usage of concepts, topics and content

**Content service architectures**

Other real applications

not like WWW10,..., ICWE'04
The User Dimension

We all are different and do not want to be unified

There are as many opinions as there are people; each has his own correct way.

Terence (190–165 BC)

- history of education, history of knowledge, different understanding
- language understanding, language use
- understanding, education, culture, preferences, abilities, skills
- working context, environment, success
- 3 facets of intelligence: problem-solving, emotional, self-reflection
- collaboration of people based on communication, coordination, and cooperation

Memory chunks - memes
Specification of User Profiles

User profile: 〈user profile name〉

Education profile: 〈general description〉
   Education: 〈list of degrees〉
   Capabilities: 〈list of skills〉
   Knowledge: 〈description of knowledge in the application〉

Work profile: 〈general description〉
   Task expertise: 〈description of knowledge〉
   Task experience: 〈positive and negative experience〉
   System experience: 〈experience with infrastructure planned〉

Information profile: 〈information need〉

Interaction profile: 〈interaction properties〉

Personality profile: 〈general description〉
   General properties: 〈list of user properties〉
   Preferences: 〈list of input/output/dialogue preferences〉
   Polarity profile: 〈list of polarity properties〉

Derivable profiles: 〈profile description and enforcement〉
   Security profile: 〈access control and privacy〉
   Safety profile: 〈safety requirements〉

Based On: 〈user goals〉
Based On: 〈user types〉
Specification of Portfolio

Explicit scoping of information

Party portfolio: ⟨party portfolio name⟩
Task: ⟨general description⟩
  Characterisation: ⟨general description⟩
    Initial state: ⟨characterisation of the initial state⟩
    Target state: ⟨characterisation of the target state⟩
    Profile: ⟨profile presupposed for solution⟩
    Instruments: ⟨list of instruments for solution⟩
    Collaboration: ⟨specification of collaboration required⟩
    Auxiliary: ⟨list of auxiliary conditions⟩
  Execution: ⟨list of activities, control, data⟩
  Result: ⟨final state, satisfied target conditions⟩
Party involvement: ⟨general description⟩
  Role: ⟨description of role⟩
  Part: ⟨behavioural categories and stereotypes⟩
Collaboration: ⟨general description⟩
Restrictions: ⟨general description⟩
  Party restrictions: ⟨general description⟩
Environment: ⟨general description⟩
Based On: ⟨life cases⟩
Memes of the Referent or User Dimension for CMS

- names,
- a number of properties,
- a variety of associations with different adhesions to other memes,
- and a variety of groupings for different purposes.
Mappings of the CMS

Concept dimension

Content dimension

Topics dimension

User Dimension

Our Approach

Integrated Design

Content

UO-CMS

Concluding
Mappings from and to the User Dimension

User2Topic-Mapping: associate/express by utterance

User2Content-Mapping: query/provide view data

view data := \( \pi_{\text{content}} (\text{utterance} \triangleleft \text{asset}) \)

User2Concept-Mapping: understand/describe meaning by chunks

chunk := \( \pi_{\text{concept}} (\text{utterance} \triangleleft \text{infon}) \)

Consistency by requirement

\( \pi_{\text{content}} (\text{utterance} \triangleleft \text{asset}) = \pi_{\text{content}} (\text{utterance} \triangleleft \text{infon} \triangleleft \text{semantical} \text{unit}) \)

or \( \subseteq, \subset, \supset, \text{ or } \supseteq \)
Result of user-oriented CMS

Not trapped in the SQL trap

Tina Musterfrau, causal user

Content
Our Approach
Integrated Design
User Dimension
UO-CMS
Concluding

Content
Information
Concept
Topic

Search request

topic
welt
concepts

parametric HERM expressions

relational database schema

SQL query

query interface

DBS

DBS query representation
Mappings from and to the User Dimension

User Dimension

Content

Our Approach

Integrated Design

User Understanding

UO-CMS

Concluding
Data, Functions for the User Dimension

Content
Our Approach
Integrated Design
User Dimension
UO-CMS
Concluding

Content
Information
Concept
Topic

EJC’06
May 30, 2006
B. Thalheim
Faithful, consistent and well-founded user-oriented CMS

Towards provable quality of CMS

Well-foundedness:

\[
\text{interpretation(explanation(t))} \subseteq \text{delivery(t)} \quad \text{and} \quad \text{presentation(foundation(cs))} \subseteq \text{annotation(cs)}
\]

are valid for any topic \(t\) and any content suite \(cs\).

Faithfulness:

\[
\text{interpretation(explanation(associate(m)))} \subseteq \text{delivery(associate(m))}
\]

for any meme \(m\).

Saturatedness:

\[
\text{interpretation(explanation(t))} \supseteq \text{delivery(t)} \quad \text{and} \quad \text{presentation(foundation(cs))} \supseteq \text{annotation(cs)}
\]

are valid for any topic \(t\) and any content suite \(cs\).

Consistency:

\[
\text{interpretation(explanation(associate(m)))} \supseteq \text{delivery(associate(m))}
\]

for any meme \(m\).
Challenges for user-oriented CMS

Foundation problem: no topic exists that may be associated with a concept or a concept set which are associated to content data which are not annotated by the given topic

Saturation problem: all topics that are associated to content data that are founded for this topic

Faithfulness problem: all memes of users are represented by faithful topics

Consistency problem: detect those memes that are not associated to saturated topics and repair

Profile genericity problem
Profile initialization problem
Profile extension problem
Portfolio genericity problem
Portfolio initialization problem
Portfolio extension problem
The data layers of well-founded and saturated CMS

- Layer 4: Memes of the users
- Layer 3-4: Privacy protection layer
- Layer 3: Topics of topic landscapes for annotation/representation
- Layer 2: Concepts of concept bases for foundation/explanation
- Layer 1: Content of content bases as macro-data or aggregations
- Layer 0: Data and documents of underlying databases as micro-data

- The user understands chunks of concepts.
- The user expresses data needs through utterances based on association to topics.
- The user queries for content or data through views.
Facilities of the top-layer of user-oriented CMS

*Utterance interpreter and analyzer:* analysis of utterances
generation of the appropriate topic landscape

*Portfolio manager:* derive, manage, change, retrieve and associate portfolio
specific task glossary that supports utterances

*Profile manager:* storage, retrieval, change, and introduction
specific slang-like vocabularies

*Meme manager:* storage, manipulation, and retrieval

*Privacy preserver:* preserving privacy by protecting ownership against possessors
Proposal for an architecture of user-oriented content management systems

Towards this century CMS

Web Playout System
- Story Space
- Stories
- Actors
- Scenarios
- Context

User Management System
- Profile manager
- Portfolio manager
- Association generator / Natural language engine

Privacy Protection System

Content Management System
- Content types
  - Structure
  - Functionality
  - Service
  - Container
- Structuring
- Functionality
- Process
- Dynamic IC
- (Pragmatics)

Topic Management System
- Topic manager
- Community manager
- Asset manager / Infon representer

Concept Management System
- Concept manager
- Derivation engine
- Unit manager / Infon representer

Content base
- Database
- Topic landscape
- Concept base

Content
Information
Concept
Topic
Concluding

Feasibility, realism, research agenda

separating concerns in dimensions for data, logical foundations and representational (topic) worlds

handling each of the dimensions separately by providing sophisticated functionality for the dimension

adding the user worlds through explicit representation of their understandings

mapping facilities between the syntax, semantics, pragmatics and referent dimension

uniform and feasible framework for user-oriented content management