Introduction to Topic Maps

Dr. H. Holger Rath
Director Research & Development
empolis GmbH
holger.rath@empolis.com – http://www.empolis.com
Why do we need Topic Maps?

Some arguments ...
Manage the Info Glut
Sounds challenging ... 
But what are Topic Maps? 

An overview ...
Topic Maps are ...

  - enabling standard to describe knowledge structures, electronic indices, classification schemes, ...

- Web enabled:
  - XML Topic Maps (XTM) are ready to use

- Designed to:
  - manage the info glut
  - build valuable information networks above any kind of resources / data objects
  - enable the structuring of unstructured information
OK, Topic Maps seemed to be needed ... But how do they work?

A first example ...
<table>
<thead>
<tr>
<th>Location</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gorda Sound</td>
<td>see North Sound 89</td>
</tr>
<tr>
<td>Little Dix Bay</td>
<td>89</td>
</tr>
<tr>
<td>North Sound</td>
<td>90</td>
</tr>
<tr>
<td>Road Harbour</td>
<td>see also Road Town 73</td>
</tr>
<tr>
<td>Road Town</td>
<td>69, 71</td>
</tr>
<tr>
<td>Spanish Town</td>
<td>81, 82</td>
</tr>
<tr>
<td>Tortola</td>
<td>67</td>
</tr>
<tr>
<td>Virgin Gorda</td>
<td>77</td>
</tr>
</tbody>
</table>
Back-of-the-Book Index “British Virgin Islands”

Gorda Sound see North Sound
Little Dix Bay .................................. 89
North Sound .................................... 90
Road Harbour see also Road Town ... 73
Road Town ...................................... 69,71
Spanish Town .................................. 81,82
Tortola ........................................... 67
Virgin Gorda .................................... 77
Back-of-the-Book Index “British Virgin Islands”

Gorda Sound see North Sound
Little Dix Bay ......................... 89
North Sound .......................... 90
Road Harbour see also Road Town ... 73
Road Town ............................ 69, 71
Spanish Town .......................... 81, 82
Tortola .................................. 67
Virgin Gorda ............................ 77
Back-of-the-Book Index “British Virgin Islands”

Gorda Sound see North Sound
Little Dix Bay .......................... 89
North Sound .............................. 90
Road Harbour see also Road Town ... 73
Road Town ............................... 69, 71
Spanish Town .............................. 81, 82
Virgín Gorda .............................. 77

Different topic classes
Back-of-the-Book Index “British Virgin Islands”

- Gorda Sound see North Sound
- Little Dix Bay .......................... 89
- North Sound .............................. 90
- Road Harbour see also Road Town .......................... 73
- Road Town ..................................... 69, 71
- Spanish Town .......................... 81, 82
- Tortola ........................................... 67
- Virgin Gorda ........................................... 77

Different occurrences classes
Gorda Sound see North Sound
Little Dix Bay ......................... 89
North Sound .......................... 90
Road Harbour see also Road Town ... 73
Road Town ............................. 69, 71
Spanish Town .......................... 81, 82
Tortola .................................. 67
Virgin Gorda ........................... 77

Multiple topic names
<table>
<thead>
<tr>
<th>Location</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gorda Sound</td>
<td>89</td>
</tr>
<tr>
<td>Little Dix Bay</td>
<td>89</td>
</tr>
<tr>
<td>North Sound</td>
<td>90</td>
</tr>
<tr>
<td>Road Harbour</td>
<td>73</td>
</tr>
<tr>
<td>Road Town</td>
<td>69, 71</td>
</tr>
<tr>
<td>Spanish Town</td>
<td>81, 82</td>
</tr>
<tr>
<td>Tortola</td>
<td>67</td>
</tr>
<tr>
<td>Virgin Gorda</td>
<td>77</td>
</tr>
</tbody>
</table>
The Topic Map building blocks

Topic, occurrences, associations, classes
Occurrences
Occurrences

Topics

Occurrence classes

Resources

BVI Welcome

SurfBVI

CaribNet
Associations

Part-Whole

Vicinity

Geo Containment

Part-Whole

Vicinity

Geo Containment

Part-Whole

Geo Containment

Vicinity

Geo Containment

Little Dix Bay

North Sound

Virgin Gorda

Tortola

Road Town

Spanish Town

Road Harbour

Associations

Topics
Associations

Association classes

Part-Whole
Vicinity
Geo Containment

Little Dix Bay
North Sound
Virgin Gorda
Tortola
Road Town
Spanish Town
Road Harbour

Topics
Associations cont’d

- Ferry Connection
- Ferry Line
  - Speedy’s
  - Spanish Town
- Harbour stop
- Road Harbour

Association
class
Association
instance
Assoc. roles
Role playing
topics
Class Hierarchies (Taxonomy)
Class Hierarchies (Taxonomy)

Super-classes

Sub-classes

Topics
Information in context

Scopes
Scopes

- Geo Containment
- Political Dependency
- Caribbean
- Brit. Virgin Islands
- Great Britain

BVI Welcome
SurfBVI
CaribNet
Scopes

Caribbean
Karibik

Brit. Virgin Islands
Brit. Jungferninseln

Great Britain
Großbritannien

Geo Containment
Geo Umschließung

Political Dependency
Politische Abhängigkeit

Names:

English
Deutsch
Scopes

Occurrences:
- Public
- Confidential

Associations:
- Geography
- Politics

Names:
- English
- Deutsch

Scopes
Scope Examples: **English, Public, Politics**

Scopes

- Geography
- Politics

Names:
- English
- Deutsch

Occurrences:
- Public
- Confidential

Associations:
- Geography
- Politics

BVI Welcome
- Article
- Map
- Image

SurfBVI

CaribNet

empolis GmbH
Knowing where we talk about

Addressable and Non-Addressable Subjects
Addressable & Non-Addressable Subjects

- **Addressable subject**
  - if subject itself is a resource in the computer – e.g., it has an URI
  - example: the web site “http://bviwelcome.com”

- **Non-addressable subject**
  - if subject exists outside the bounds of the computer
  - example: the British Virgin Islands
Subject Indicator

- Subject indicator is an addressable resource representing a non-addressable subject
  - e.g., URI http://www.topicmaps.org/xtm/1.0/country.xtm#VG representing the country “Virgin Islands (British)” (as defined by ISO 3166)
- Subject indicator is kind of proxy for the subject

- If a subject indicator is published and maintained at an advertised location (web site) it is called a Published Subject Indicator (PSI)
  - PSIs are necessary to standardize and re-use non-addressable subjects in various topic maps
  - Several OASIS TCs define PSIs for vertical application domains
Bringing things together

Identity, Topic Naming Constraint, Merging, Reification
Identity of a Topic

http://www.topicmaps.org/xtm/1.0/country.xtm#VG

Brit. Virgin Islands

BVI

Topic map 1

Topic map 2

Identity established by Resource Reference or Subject Indicator
Topic Naming Constraint (TNC)

Topics with the same name in the same scope are to be seen as the same topic by the processing software.

Controversial, because of danger of merging even if the topics are not about the same subject.

they just have accidentally the same name.
Merging

http://www.topicmaps.org/xtm/1.0/country.xtm#VG

Identity

Before Merging
Merging

http://www.topicmaps.org/xtm/1.0/country.xtm#VG

Identity

Merging

Topic map 1 merged with topic map 2
Merging

http://www.topicmaps.org/xtm/1.0/country.xtm#VG

Brit. Virgin Islands

BVI

Caribbean

Great Britain

Topic naming constraint

Topic map 1 merged with topic map 2
Merging

http://www.topicmaps.org/xtm/1.0/country.xtm#VG

Identity
Brit. Virgin Islands
BVI
Caribbean
Great Britain

Union of characteristics
Duplicates removed
Duplicates removed

Topic map 1 merged with topic map 2

After Merging
Reification

- Holiday background
- Associated associations
- Association names
- Association occurrences
- Occurrence names

- Nat. holiday "Territory Day"
- "BVI are UK’s overseas territory"
- Great Britain
- Brit. Virgin Islands
- "BVI Territory Contract"
- "UK’s involvement in the Caribbean"
- "Map of the 16th century"
Summary: Topic Map Concepts

- Topic, occurrence, association
- Classes for topics occurrences, and associations
- Class-instance, super-subclass associations
- Scope and scoping topic
- Addressable and non-addressable subjects
- Identity and subject indicator
- Merging
- Reification
The Family of Topic Map Standards

ISO/IEC 13250, Reference Model, Standard Application Model, TMQL, TMCL
Family of TM Standards

- Reference Model
- ISO/IEC 13250
- HyTM Syntax
- XTM Syntax
- TMQL
- TMCL

Mapping relationships:
- HyTM Syntax to Reference Model
- XTM Syntax to Reference Model
- TMQL to TMCL
- TMQL to Reference Model
- XTM Syntax to Reference Model
Family of TM Standards cont’d

- ISO/IEC 13250:2000
  - ISO standard defining general concepts and interchange syntax
    (HyTM = SGML/HyTime, XTM = XML/Xlink)

- Reference Model (RM)
  - ISO project
  - foundation of the TM paradigm
  - independent of any particular (storage/interchange) syntax

- Standard Application Model (SAM)
  - ISO project
  - infoset-based TM model for implementers on top of RM
  - mapping to HyTM and XTM interchange syntax
Family of TM Standards cont’d

- TMQL – TM Query Language
  - ISO project
  - ‘SQL’ for TMs
  - Standardized creation/modification of TMs stored in TM Management Systems

- TMCL – TM Constraint Language
  - ISO project
  - Framework for the definition of ontologies / schemas for vertical or domain specific applications
  - Support for semantic validation
Topic Maps and Related W3C Standards

XML, XLink, RDF
Topic Maps and XML

- XML is format to exchange TMs between software tools
- XML is not the internal data structure of TM tools but the ASCII serialization of a complex knowledge network
- Resources could be of any notation / format; they just have to be addressable
- If resources are XML occurrences can easily point into the resources
**Topic Maps and XLink**

- XLink is format to exchange the links in TMs
- TMs can point to resources which are addressable by XLink
- TMs organize link networks like XML structures data
Topic Maps and RDF

- Resource Description Framework
  - RDF Model and Syntax
  - RDF Schema
  - DAML, OIL, DAML+OIL

Attend Nocturne “RDF Schema for Topic Maps” at 07:00 pm today
Topic Maps and RDF cont’d

- TM / RDF – Similarities
  - Structured, complex metadata
  - Based on graphs
  - Knowledge representation, ontologies
  - Help power the Semantic Web idea
  - TMs on top of RDF ⇔ RDF on top of TMs
## Topic Maps and RDF cont’d

### TM / RDF – Differences

<table>
<thead>
<tr>
<th>TM</th>
<th>RDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>topic-centric</td>
<td>resource-centric</td>
</tr>
<tr>
<td>pre-defined semantics</td>
<td>simple data structure</td>
</tr>
<tr>
<td>$n$-ary associations with role players</td>
<td>directed binary relations</td>
</tr>
<tr>
<td>(instead of direction)</td>
<td></td>
</tr>
<tr>
<td>distinguishes between</td>
<td></td>
</tr>
<tr>
<td>addressable and non-addressable subjects</td>
<td></td>
</tr>
<tr>
<td>merging</td>
<td></td>
</tr>
</tbody>
</table>
Some TM application scenarios, please

Here they are ...
Typical Topical Applications

- Subject classification
  - Connecting resources with a classification scheme
- Knowledge representation
  - Model domain knowledge in a topic map explicitly
- Publishing knowledge networks
  - Selling added-value (e.g., commentaries, rich metadata)
- Search engines
  - Intelligent ‘find’ technologies
Advanced Topical Applications

- New database paradigm
  - Multi-dimensional graphs are more powerful than relations (RDB) or hierarchies (XML)

- Application integration, Web Services, Semantic Web
  - Bridging interface barriers
  - Ontology-based, syntax-less information/ knowledge interchange
Who is using TMs?

Some real-world projects
Who is Using TMs?

- Publishers
  - encyclopedia
  - legal
  - eLearning
  - media, news
- Web portal provider
  - site organization principle
- Industry
  - intelligent call center, knowledge gateway
  - corporate memory
  - next generation content management
Thank You!

Any Questions?