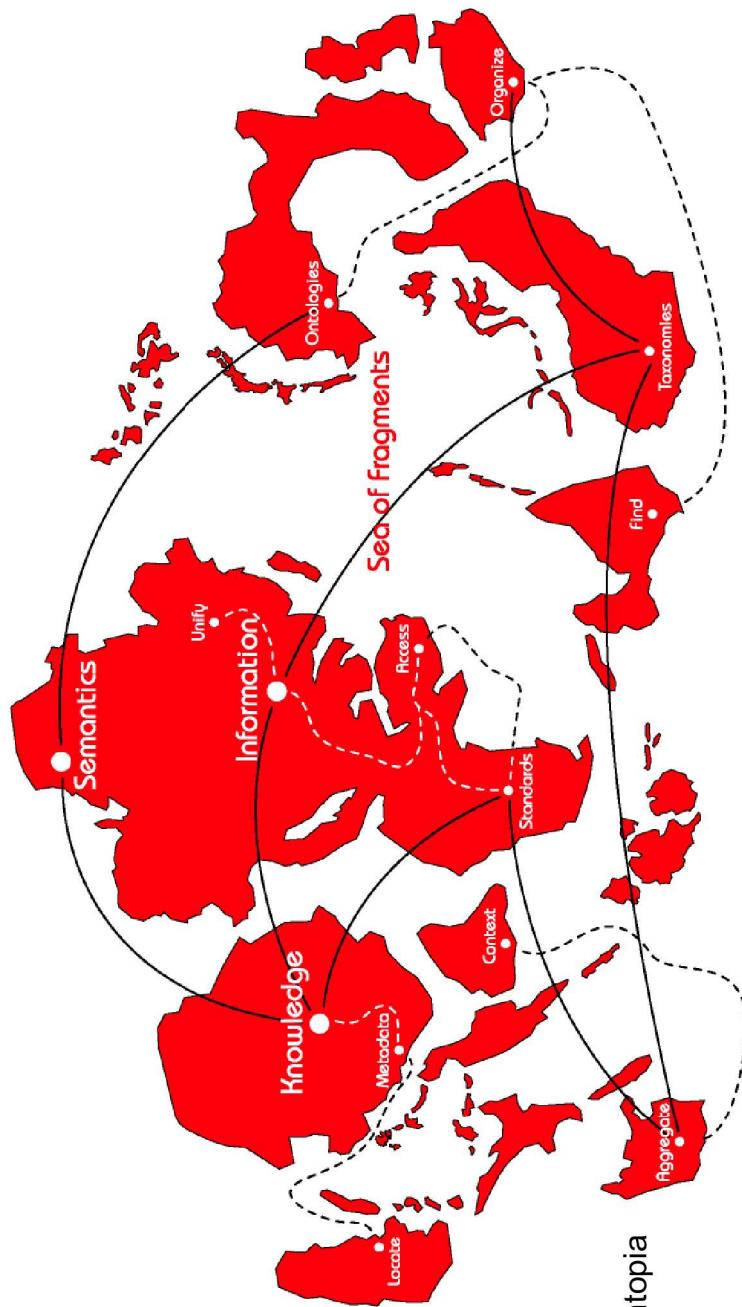


Topic Maps and ID

Introduction and Thesis Topic Suggestions



Lars Marius Garshol
Development Manager, Ontopia
<larsga@ontopia.net>

2005-01-05

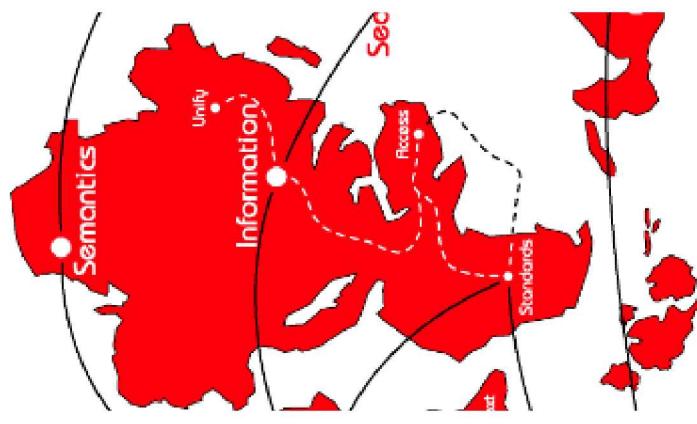
Who is speaking?

- **Lars Marius Garshol**
 - Development Manager at Ontopia
 - Former IFI student
 - Editor of ISO 13250-2 (Topic Maps), 13250-3 (XTM), and 18048 (TMQL)
- **Ontopia**
 - Topic map software vendor
 - Main product: Ontopia Knowledge Suite (OKS)
 - Active in ISO, W3C, and OASIS standardization related to topic maps

Overview

- **Introduction to topic maps**
- **Description of suggested thesis topics**
- **Where to find more information**

Introduction to topic maps



What they are

The TAO model

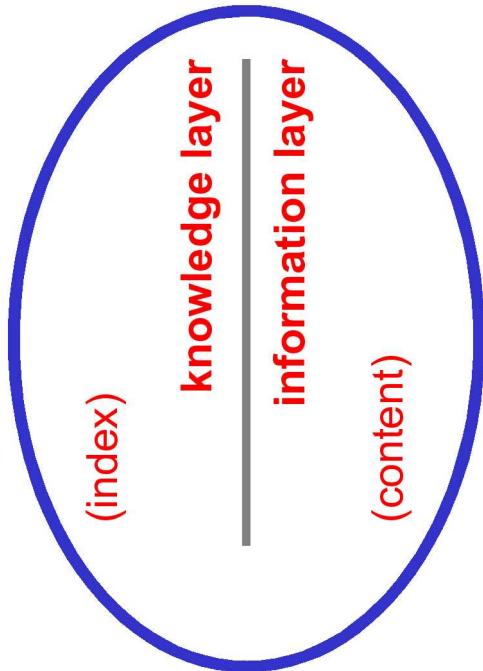
Uses of topic maps

What are topic maps?

- **A model for organizing information**
 - kind of like a back-of-book index on steroids
- **A subject-based classification of information**
 - that is, information resources are classified based on what they are *about*
- **A knowledge technology**
 - a topic map is a semantic network, and can be used for querying and logical inferencing
- **A data model for applications**
 - applications can be written using topic maps to represent their data
- **An international standard**
 - topic maps are an ISO standard with an XML-based interchange syntax

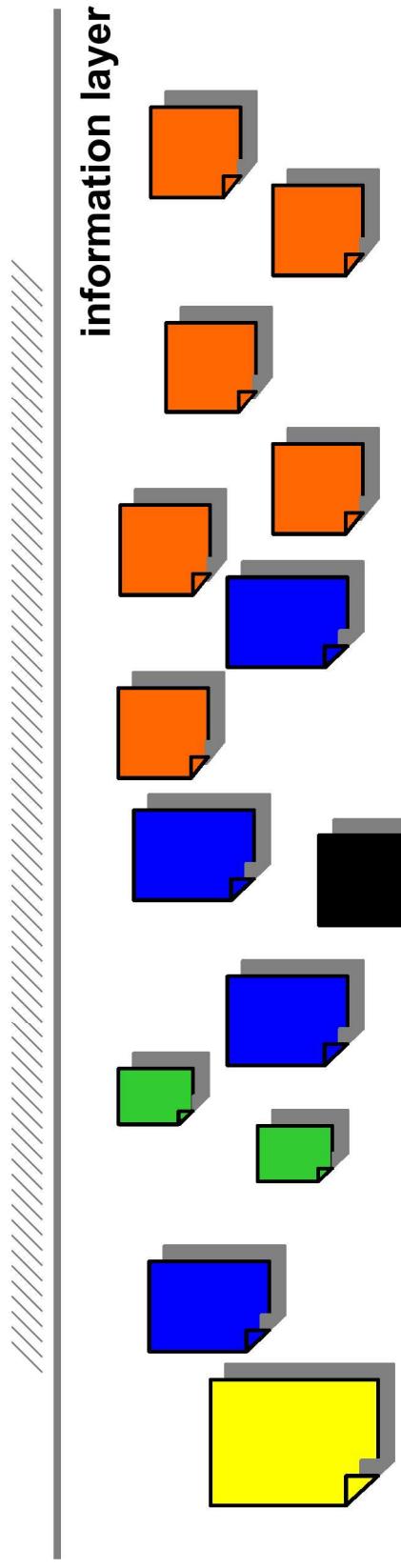
The 2-Layer Topic Map Model

- The core concepts of Topic Maps are based on those of the back-of-book index
- The same basic concepts have been extended and generalized for use with digital information
- Envisage a 2-layer data model consisting of
 - a set of information resources (below), and
 - a “knowledge map” (above)
- This is like the division of a book into content and index



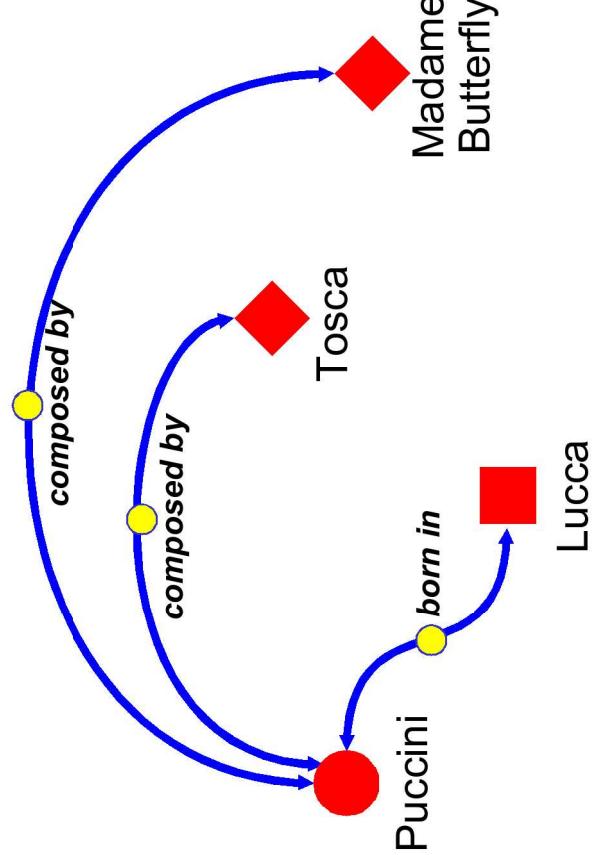
(1) The Information Layer

- **The lower layer contains the content**
 - usually digital, but need not be
 - can be in any format or notation
 - can be text, graphics, video, audio, etc.
- **This is like the content of the book to which the back-of-book index belongs**



(2) The Knowledge Layer

- The upper layer consists of **topics and associations**
 - **Topics** represent the subjects that the information is about
 - Like the list of topics that forms a back-of-book index
 - **Associations** represent relationships between those subjects
 - Like “see also” relationships in a back-of-book index

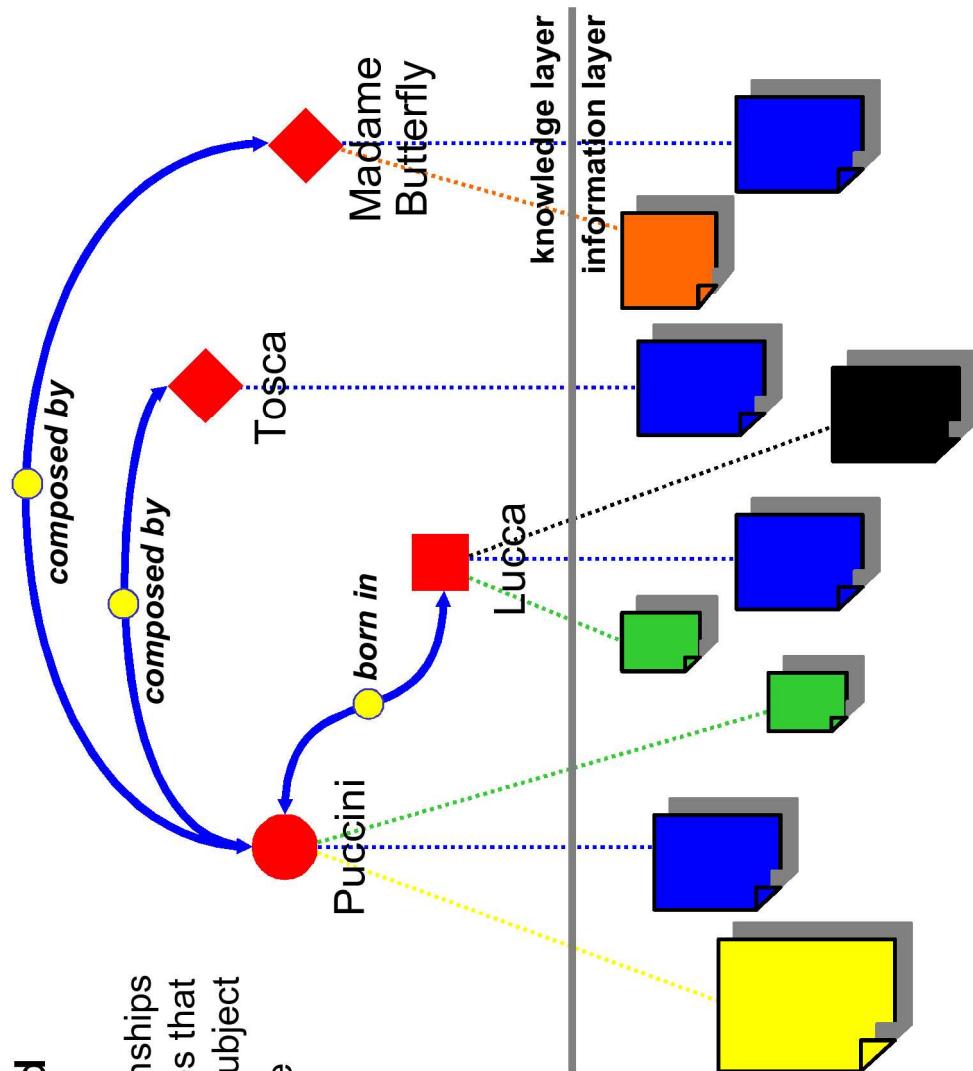


knowledge layer

Linking the Layers Through Occurrences

- The two layers are linked together

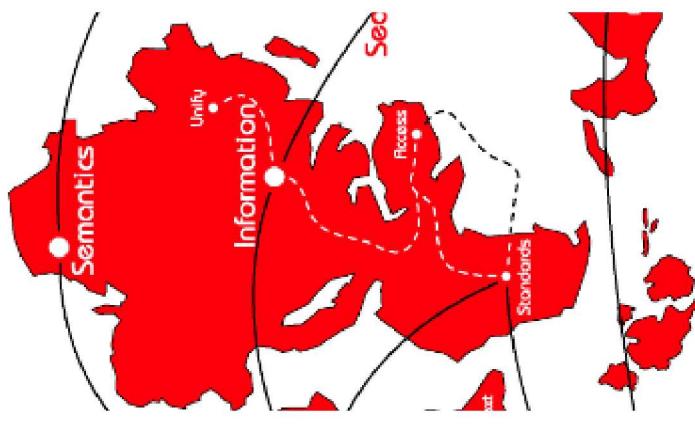
- **Occurrences** are relationships with information resources that are pertinent to a given subject
- The links (or locators) are like page numbers in a back-of-book index



Application areas for topic maps

- **Portals**
 - used to organize the portal content for better findability
- **E-learning**
 - used by students to organize concepts; also as courseware metadata
- **Business process modelling**
 - used to describe and relate business processes
- **Product configuration**
 - used to describe reusable components for easy creation of new products from existing components
- **Enterprise application integration**
 - integration of information from diverse applications using a topic map server
- **Content management**
 - organization of content and metadata vocabularies for content
- **Asset management**
 - description and mapping of organization assets for KM purposes

Suggested thesis topics



Natural language querying

Upconversion

Modelling guidelines

Natural language querying of topic maps

- **Topic maps are highly structured information graphs**
 - structured querying of these graphs a la SQL is possible
 - query languages for such querying already exist
- **However, most users don't speak a formal query language**
 - they would like to query using natural language
 - since all topics in topic maps have names natural language terms should be mappable to topics
- **Imagine the following**
 - “please show me all cities” → instance-of(\$CITY, city)?
 - “all cities in Norway, please” → instance-of(\$CITY, city), located-in(\$CITY, norway)?
 - “museums in Østfold about ships” → ...

Upconversion from X to topic maps

- **To use topic maps to find information in a repository you need a topic map of the repository**
 - creating this by hand is often economically infeasible
 - in many cases, however, it can be done automatically, or semi-automatically
- **Is it possible to create tools to simplify this conversion?**
 - can XML documents, databases, Excel spreadsheets, PDF files, HTML documents, ... be automatically converted to topic maps?
 - at the press of a button? with some configuration? with simple programming?
- **Can it be done in a more limited way?**
 - automated classification of documents uploaded to a CMS
 - automated classification of blog entries

Modelling guidelines for information retrieval

- **There are many approaches to classifying content with topic maps**
 - organizing subjects in a hierarchy/taxonomy (a la LivsIT)
 - using multiple hierarchies (a la forskning.no)
 - using faceted classification (a la kulturnett.no)
 - using a strong domain ontology (a la Italian Opera)
 - using traditional DC-like metadata (a la ODIN)
 - hybrid approaches (a la apollon.uio.no)
- **More fine-grained choices also exist**
 - model resources as occurrences or topics?
 - how to name topics?
 - the use of topic types and class hierarchies
- **Very little is known about**
 - the economic cost of each approach
 - the impact on precision/recall for each approach
 - the relation to stickiness of content/serendipitous discovery
 - consequences for searching, whether full-text or structured

Other possibilities

- visualizations (choice of metaphor, clustering, ...)
- graphical topic map editors
- graphical modelling languages for topic maps
- published subjects: repositories & discovery of
 - TM storage models, native TM databases, ...
- TMFS (topic map file system)
- personal information management with topic maps
 - TM-driven email client?
- TM-driven brainstorming tool (a la MindManager etc)
- automated classification of content with topic maps
- formal models for topic maps
- topic map-based web services (lookup, query, modify)
- logical inferencing in topic maps (with rules or description logic)
- annotating software with TMs (automate tasks, reasoning, ...)

For more information

- **Introductions to topic maps**
 - <http://www.ontopia.net/topicmaps/materials/tao.html>
 - <http://www.xml.com/pub/a/2002/09/11/topicmaps.html>
- **Sites about topic maps**
 - <http://www.topicmap.com>
 - <http://www.ontopia.net>
 - <http://www.isotopicmaps.org/>
- **These slides**
 - <http://www.ontopia.net/topicmaps/materials/uio-id-2005-01-05.pdf>
- **Topic maps in Norway**
 - <http://www.emnekart.no>
- **Topic maps in e-learning**
 - <http://topicmaps.estandard.no>
- **Topic map news**
 - <http://topicmaps.it.bond.edu.au/weblogs.mc>
- **Mailing list**
 - <http://www.infoloom.com/mailman/listinfo/topicmapmail>