

# The TAO of Topic Maps

Finding the Way in the Age of Information

**Steve Pepper, CEO, Ontopia**  
Convenor ISO/IEC JTC 1/SC 34/WG 3  
Editor XML Topic Maps (XTM)  
<pepper@ontopia.net>

# What is Topic Maps?

---

- **An international standard for the distributed management of information and knowledge**
  - ISO/IEC 13250:2003 – Topic Maps
- ***Other Perspectives on Topic Maps:***
- **The Information Management View**
  - A new paradigm for organizing, retrieving, and navigating information resources
- **The Knowledge Management View**
  - A knowledge representation formalism optimized for use in information management
- **The Library Science View**
  - A way to collocate all knowledge about a subject – in particular its relationship to other subjects and to information resources

# Some Uses of Topic Maps

---

## 1. Web Portals

- Topic Map-Driven Web Sites

## 2. Corporate Intranets

- Knowledge Management

## 3. Connecting Information Systems

- Enterprise Information Integration

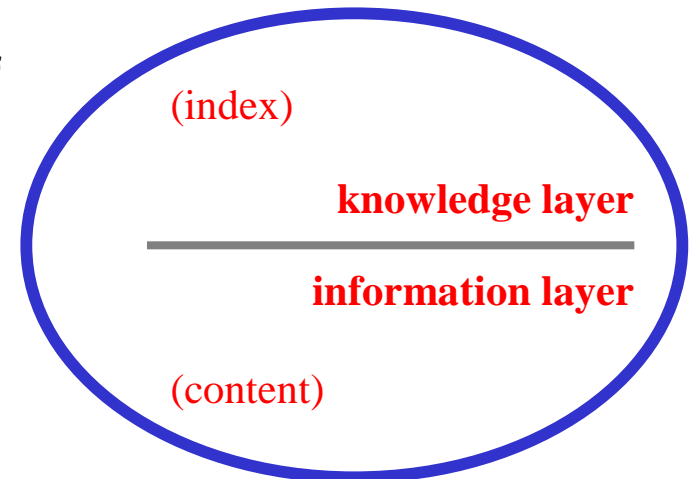
- **etc. etc. etc.**

- eLearning
- Business Process Modelling
- Application Integration
- ...

# Understanding the Topic Map Model

---

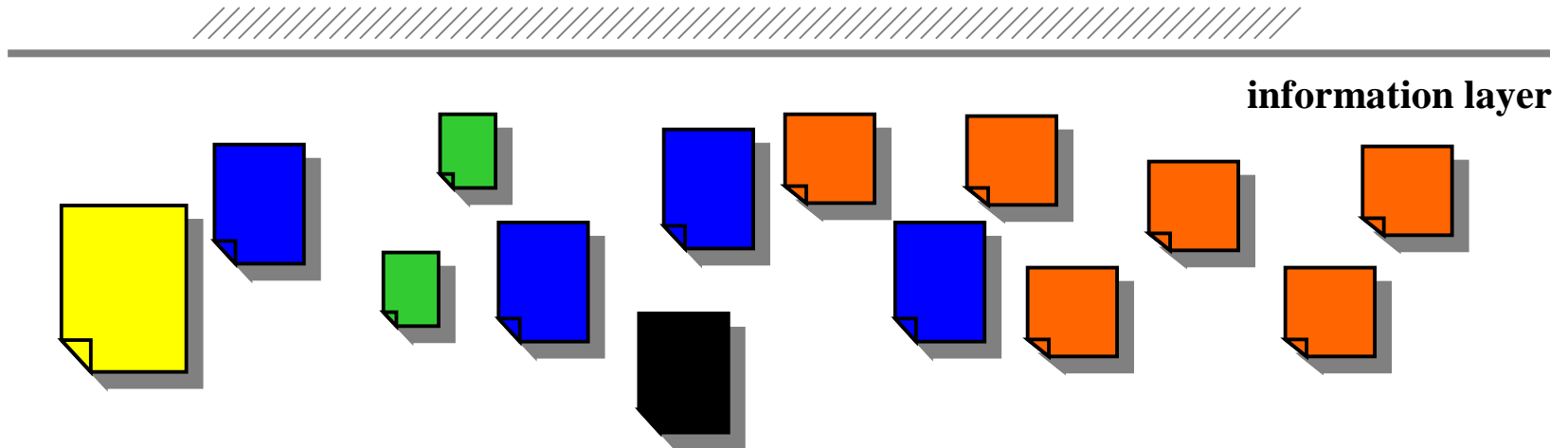
- **The core concepts of Topic Maps are based on those of the back-of-book index**
  - (The purpose of a back-of-book index is to make it easier to find information...)
- **The same basic concepts have been extended and generalized for use with digital information**
- **Envisage a 2-layer data model 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

---

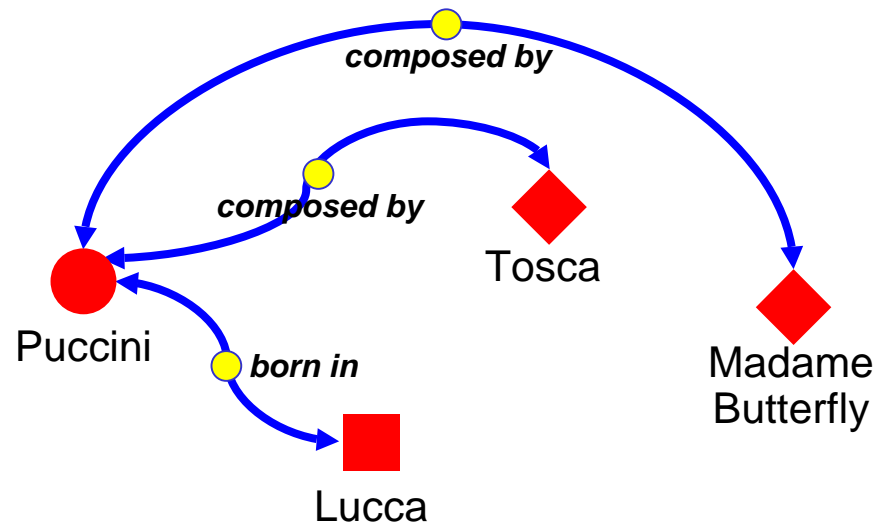
- **The lower layer contains the content**
  - usually digital, but need not be
  - can be in any format or notation
  - can be text, graphics, video, audio, etc.
- **This is like the content of the book to which the back-of-book index belongs**



## (2) The Knowledge Layer

---

- The upper layer consists of (typed) 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 the “see also” relationships in a back-of-book index

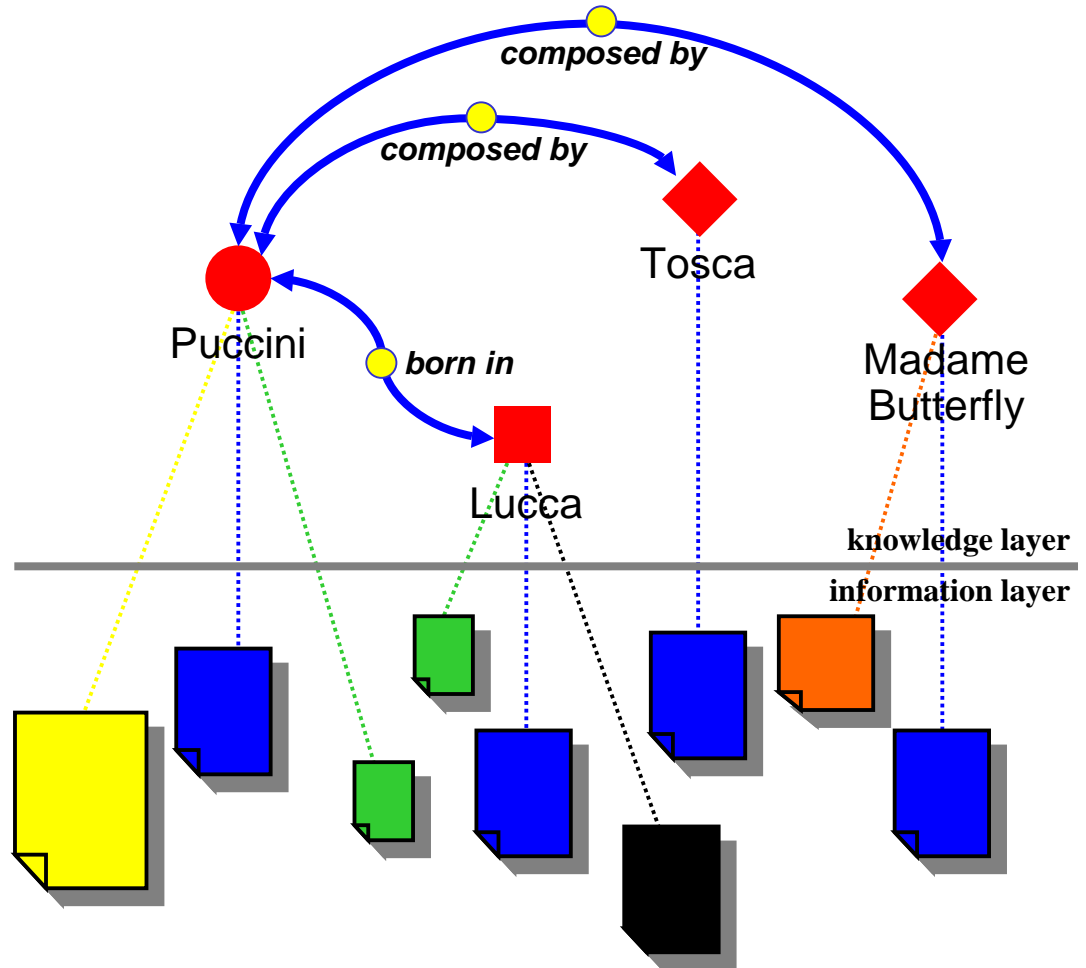


knowledge layer

---

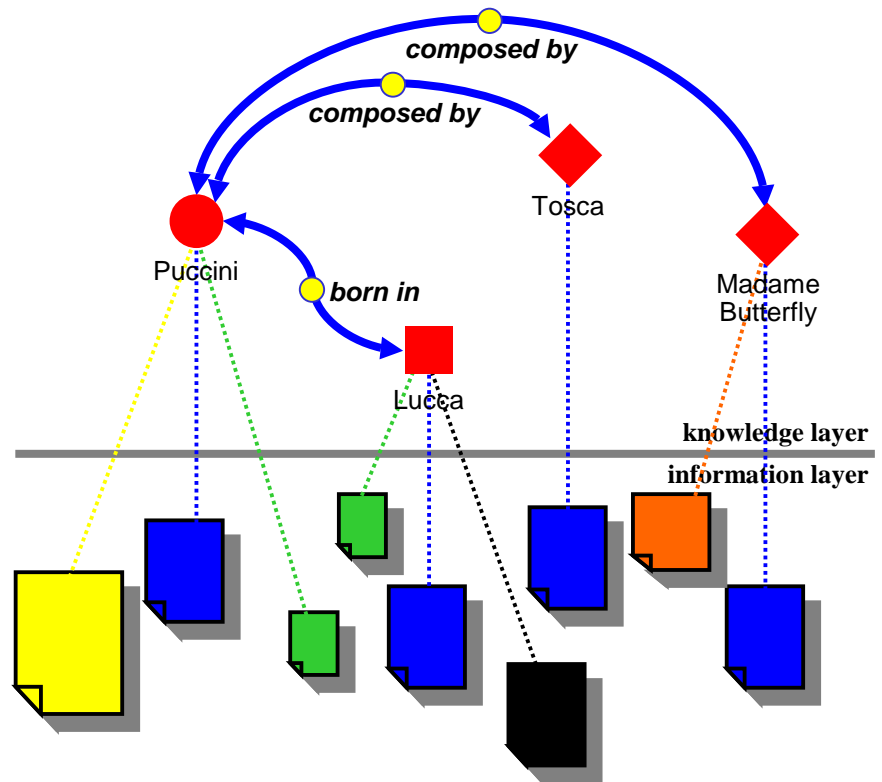
# Linking the Layers Through Occurrences

- The two layers are linked together
  - Occurrences are (typed) relationships between topics and information resources that are pertinent to them
  - The links (or locators) are like page numbers in a back-of-book index



# Summary of Core Topic Maps Concepts

- **Some pool of information or data**
  - any type, format, or location
- **A knowledge layer, consisting of:**
- **Topics**
  - a set of topics representing the key subjects of the domain in question
- **Associations**
  - expressing relationships between subjects
- **Occurrences**
  - links to information that is somehow relevant to a given subject
- **= The TAO of Topic Maps**





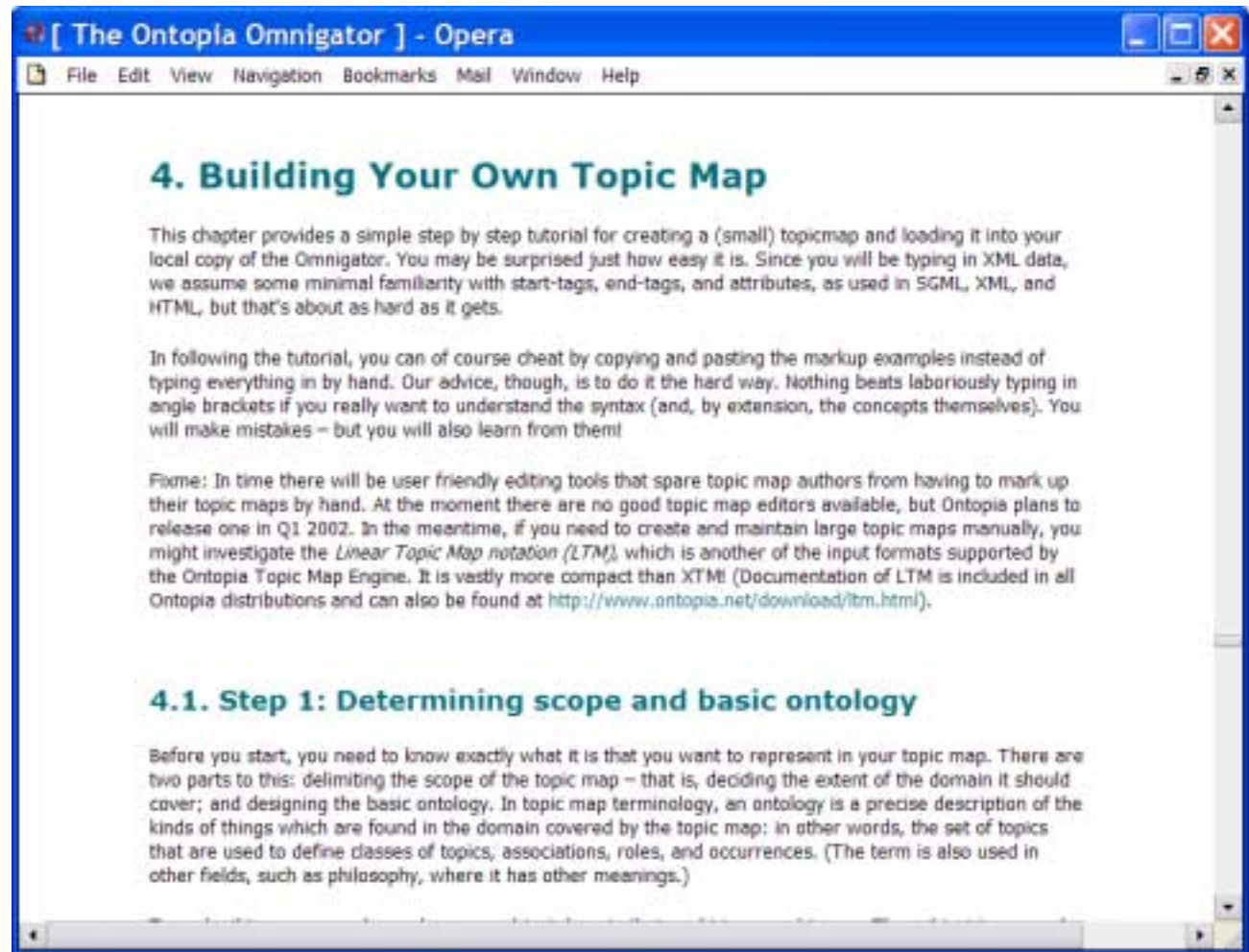
# The Omnigator – A Free Topic Maps Browser

The world's most popular topic map tool!

Download it for free from Ontopia's web site!

Learn about Topic Maps – and create your own!

See the demo during the breaks!



## With this Simple but Flexible Model You Can

---

- **Make knowledge explicit – by**
  - Identifying the *subjects* that your information is about
  - Expressing the *relationships* between those subjects
- **Bridge the domains of knowledge and information – by**
  - Describing where to find *information* about the subjects
  - Linking information about a common subject across multiple repositories
- **Transcend simple categories, hierarchies, and taxonomies – by**
  - Applying rich associative structures that capture the full complexity of the real world
- **And encompass virtually any kind of data – including**
  - Relational, hierarchical, and associative structures
  - Metadata tuples and triples

## Summary Thus Far

---

- **Some things are new and some things are old**
- **Established, well-tried, and functioning concepts have been taken from various fields:**
  - Indexing, glossaries, thesauri
  - Bibliographic principles (e.g. collocation)
  - Data modelling, formal logic, artificial intelligence
- **These have been welded into a Complete Model**
- **That model has been Formalized**
- **It has been given an Interchange Syntax**
- **And it has been made into an International Standard**
  - Standard software
  - Interchange between systems
  - Interoperability
- **But there is more, of course...**

# Supporting Context through Scope

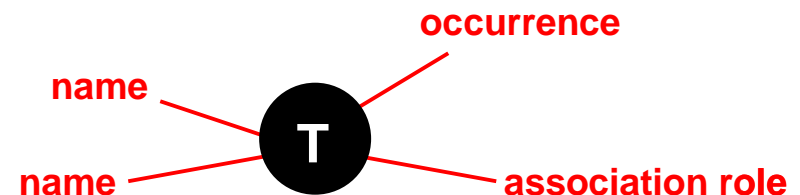
---

- **Topic Maps are about representing knowledge**
- **Knowledge is not absolute; it has a contextual aspect**
- **Context sensitivity is handled through the concept of **scope****
- **Scope makes it possible to**
  - Cater for the subjectivity of knowledge
  - Express multiple viewpoints in one knowledge base
  - Provide personalized views for different groups of users

# How Scope Works

---

- **Topics have “characteristics”**
  - Its names and occurrences, and the roles it plays in associations with other topics
  
- **Every characteristic is valid within some context (scope), e.g.**
  - the name “Allemagne” for the topic Germany in the scope “French”
  - a certain information occurrence in the scope “technician”
  - a given association is true in the scope (according to) “Authority X”



Filtering by scope

# How Topic Maps Improves Access to Information

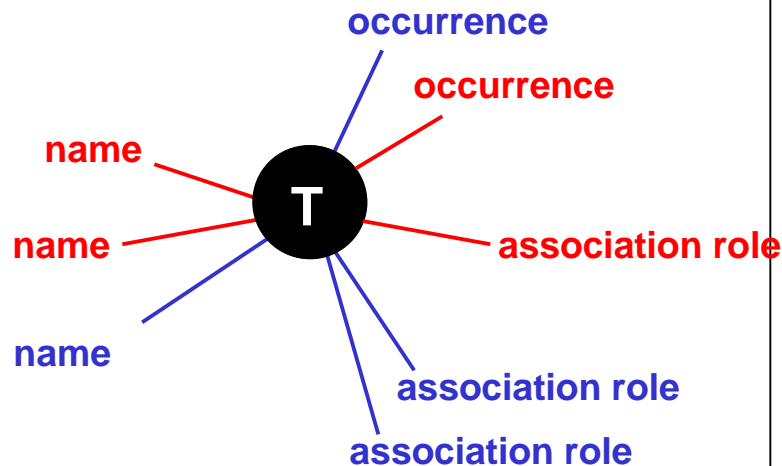
---

- **Intuitive navigational interfaces for humans**
  - The topic/association layer mirrors the way people think
  - See the demo of the Omnigator
- **Powerful semantic queries for applications**
  - A formal underlying data structure
  - *“Show me all composers who wrote operas that were based on plays that were written by Shakespeare”*
- **Customized views based on individual requirements**
  - Personalization based on scope
  - Set language preferences, skill levels, special interests, etc.
- **Information aggregation across systems and organizations**
  - Two topic maps can be merged automatically...
  - Perhaps the most powerful feature of Topic Maps...

# Principles of Merging in Topic Maps

---

1. In Topic Maps, every topic represents a subject
2. To achieve the *collocation objective*, there should be exactly one topic per subject
  - When **two topic maps** are merged, topics that represent the same subject should be merged to a single topic
  - When **two topics** are merged, the resulting topic has the union of the characteristics of the two original topics



Merge the two topics together...

...and the resulting topic has the union of the original characteristics

# Applications of Merging

---

- **Information integration**
  - Information that spans multiple repositories can be merged to provide a unified view of the whole
- **Knowledge sharing across the organization**
  - Knowledge captured in one part of an organization can be made available to the whole organization
- **Distributed knowledge management**
  - There is no need to centralize knowledge management in order to make it sharable
- **Knowledge sharing between organizations**
  - Information and knowledge can be shared without enforcing a common vocabulary



# What Makes Merging Possible?

---

- **NOT the use of names, which are notoriously unreliable**
  - Names are not unambiguous (the homonym problem)
  - Many topics have multiple names (the synonym problem)
- **Reliable knowledge aggregation is only possible through the use of unique global identifiers**
- **The issue of *identification of subjects* is crucial**
  - If subjects have unique identifiers, people can be free to use whatever names they like – and machines can still aggregate information
- **Topic Maps has solved this problem through**
  - Subject Indicators
  - Published Subjects

# What Makes Topic Maps Special?

---

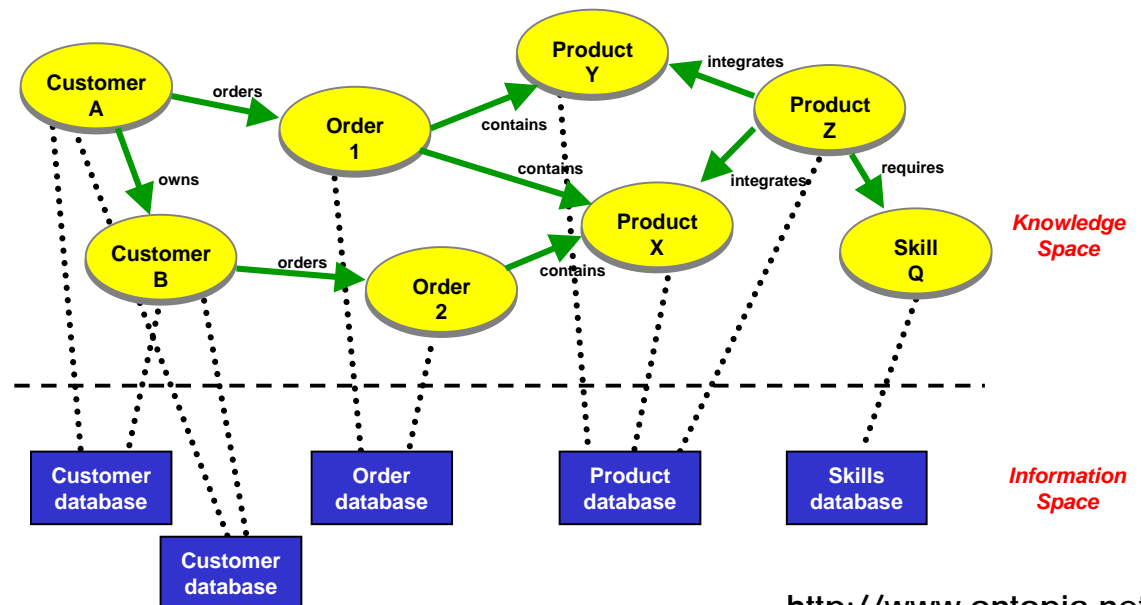
## *A unique combination of features:*

1. **A simple but flexible associative model that encompasses almost any kind of data**
  2. **“Just enough” semantics to make the model easy to use for humans but still tractable for computers**
  3. **The ability to merge arbitrary topic maps**
  4. **A robust identity model in support of the collocation objective**
  5. **Status as an ISO standard**
- These features make Topic Maps especially suitable for Enterprise Information Integration (EII)**

# Integrating Information Across Systems

- **Topic Maps are designed for ease of merging!**
  - Multiple Topic Maps can be created from many different repositories of information ... and then merged to provide a unified view of the whole
- **Typical Applications:**
  - Integration of hitherto disconnected information silos within an enterprise
  - Integration of information (and knowledge) across organizations

- **Advantages:**
  - Consolidated access to all information
  - Does not require existing content or applications to be changed
  - Results in a unified model of the organization's business



# Borland: Topic Maps for Application Integration

---

- **StarTeam products support software development teams**



- **Product suite consists of**
  - source code repository with version control
  - bug and feature tracking database
  - requirements definition and tracking database
  - test management software
  - CRM system
- **Products acquired from other companies; no integration**

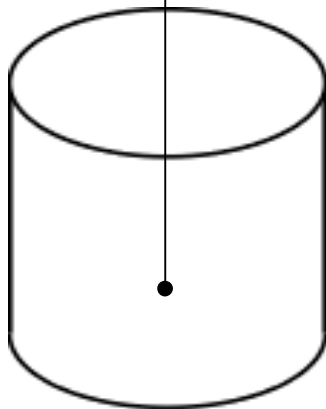
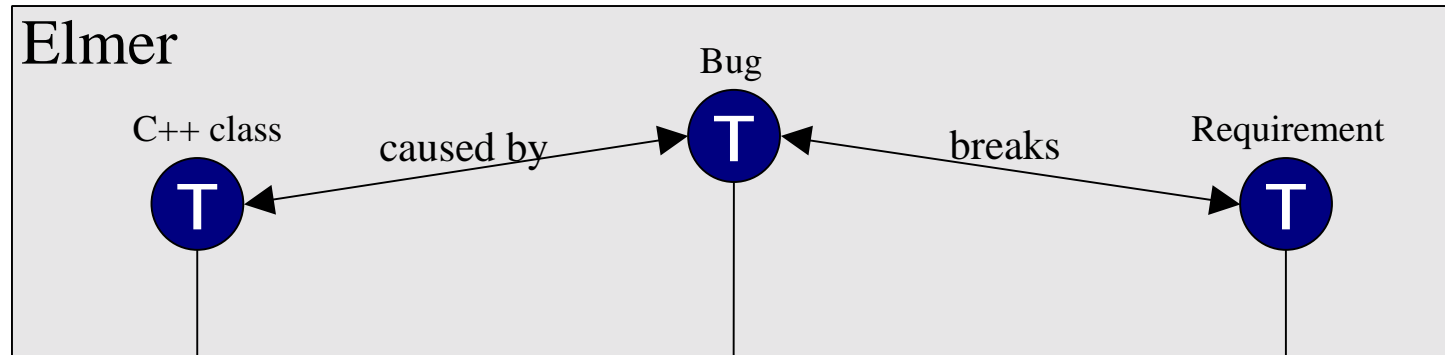
## Elmer Solution: Topic Maps

---

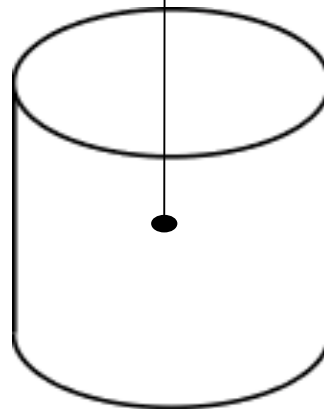
- **Elmer is an internal Borland project to develop the next generation of their suite**
- **Rather than rewrite the products, a new integration product is being added**
- **Elmer is a Topic Maps server which integrates information from different products**
- **Based on the Ontopia Knowledge Suite**

# Elmer: Data Example

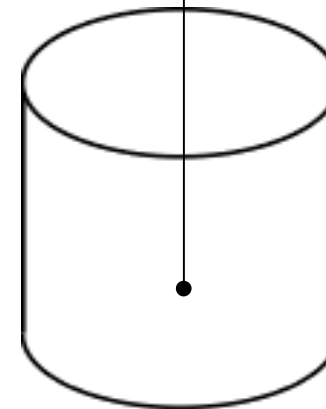
---



Source repository

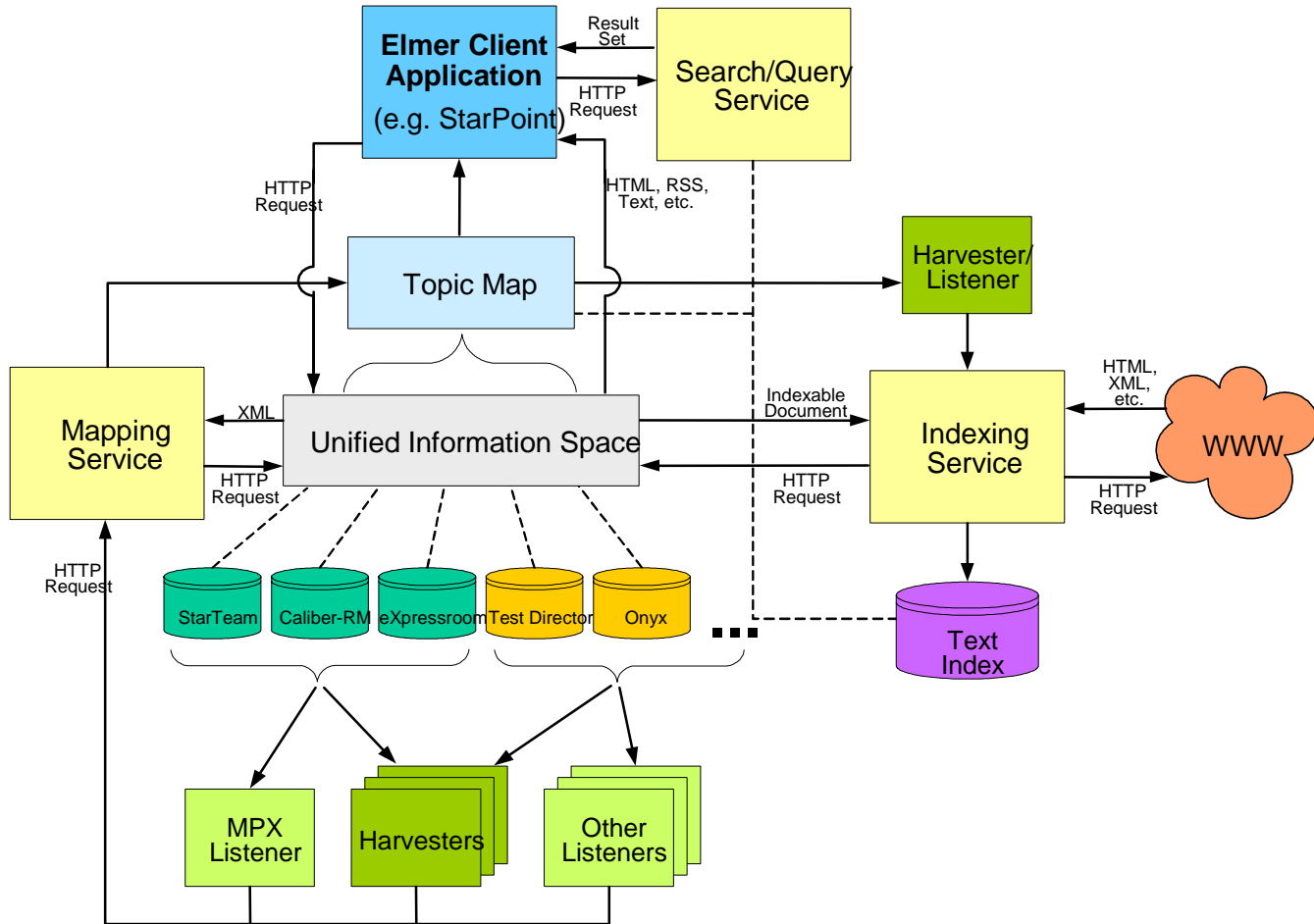


Bug database



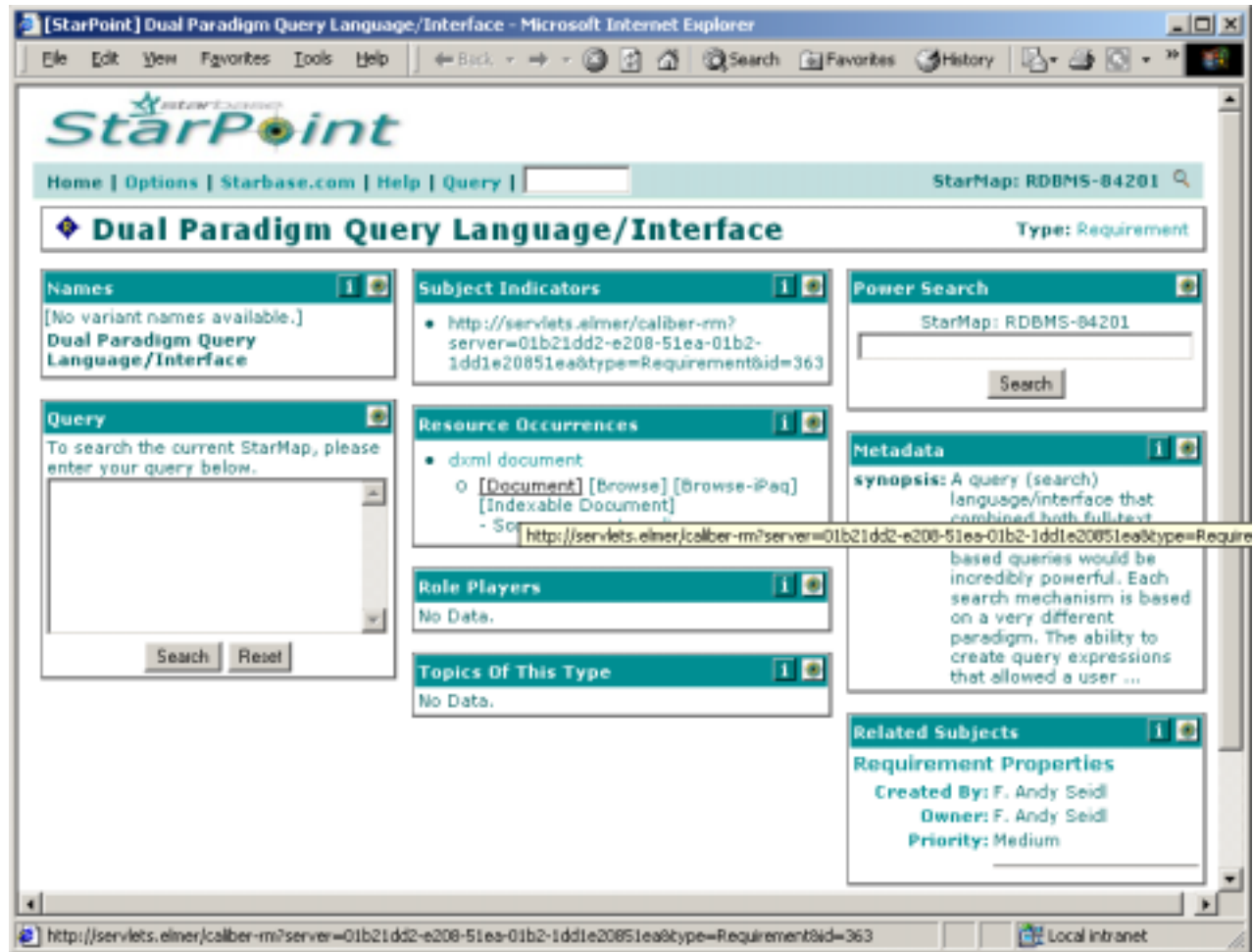
Requirements DB

# Elmer Architecture



# Uses of Information in Elmer (1)

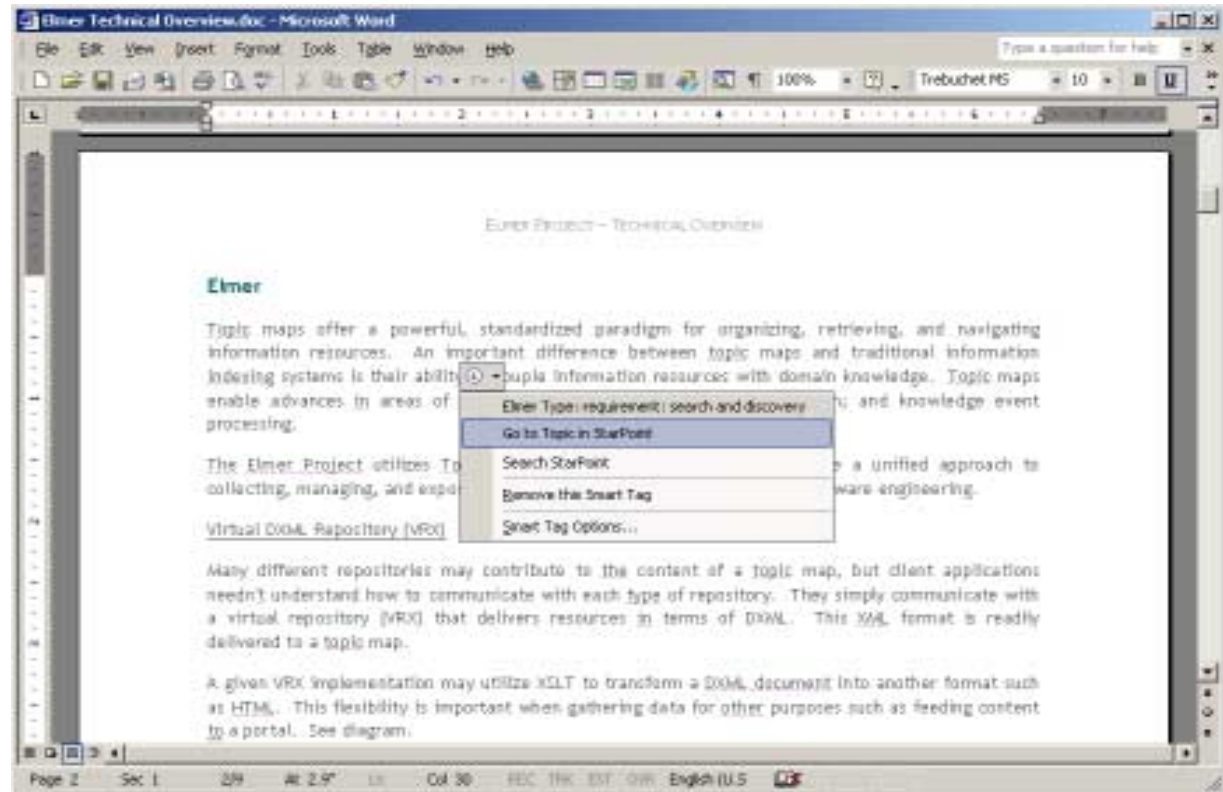
- All information provided through a portal
  - Straightforward navigation interface
  - Querying, both full-text and structured
  - Now possible to query across multiple databases





## Uses of Information in Elmer (2)

- **Drives Smart Tags integration with Microsoft Office**
  - Terms known from Elmer are highlighted
  - (Names of topics are used as a controlled vocabulary)
  - These appear as links back into the portal



## Key Benefits from Elmer

---

- **Application integration without changing the existing applications**
- **Highly flexible architecture**
  - Combination of Web Services and Topic Maps = Semantic Web Services
- **Ability to search across multiple repositories**
- **Much more consistent data**
- **Juxtaposition of data reveals new knowledge**
- **Collecting data in one place leads to new and unexpected uses of the content**

## For More Information

---

- **Useful papers:**
  - *“The TAO of Topic Maps: Finding the Way in the Age of Information”*  
<http://www.ontopia.net/topicmaps/materials/tao.html>
  - *“The XML Papers: Lessons on Applying Topic Maps”*  
<http://www.ontopia.net/topicmaps/materials/xmlconf.html>
  - *Pointers to other papers:*  
[http://www.ontopia.net/topicmaps/learn\\_more.html](http://www.ontopia.net/topicmaps/learn_more.html)
- **The Omnigator:**
  - *Online demo:* <http://www.ontopia.net/omnigator>
  - *Free download:* <http://www.ontopia.net/download/freedownload.html>
- **Standards activities:**
  - *ISO:* <http://www.isotopicmaps.org/>
  - *OASIS:* <http://www.oasis-open.org/committees/tm-pubsubj>
- **Community sites (with pointers to mailing lists, books, etc.):**
  - <http://easytopicmaps.com>
  - <http://www.topicmap.com>