

# Topic Maps

## A Practical Introduction With Case Studies

Kal Ahmed

*Techquila*

kal@techquila.com

Emnekart Norge 2002

Oslo

18th October 2002

# Introduction

# Goals Of This Presentation

- Discuss current application areas for topic maps
- Present some existing examples
- Discuss why topic maps were chosen
- Discuss what kinds of problems were faced by implementers

# Topic Map Applications

- Embedded vs. Engine
  - Not all applications say "topic maps" on the box.
  - Of those that do, commercial offerings are mainly toolkits or "engines"
  - This presentation deals with the uses of topic map engines, not the engines themselves.
- Engine Features Include:
  - Programming interface (usually Java)
  - Persistent storage
  - Indexing
  - Import and export of interchange syntax topic map files
  - Web application integration

# Current Topic Map Uses

- Current topic map application areas include
  - Published Indexes
  - Web Applications
  - Application Development
  - Application Integration

# Published Indexes

# Introduction

A stated goal of XTM is to:

*improve the findability and manageability of information*

*XTM 1.0 Specification*

The "first wave" of topic map applications were:

- Structured indexes
- Static content navigation aids
- Topic-based or concept-centric content organisation

# Case 1 - IRS

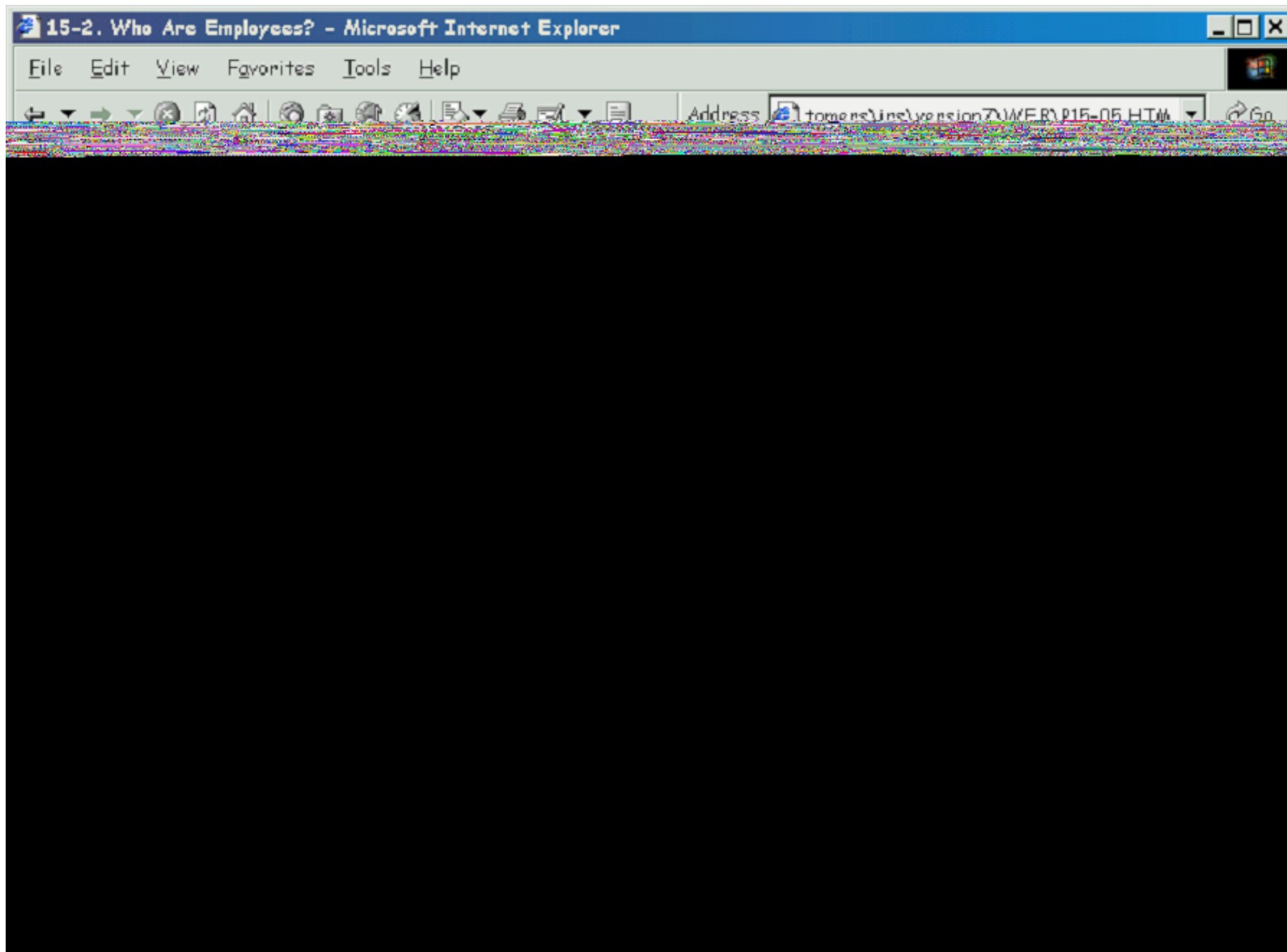
- The Customer:
  - The US Internal Revenue Service
- The Developer:
  - Michel Biezunski
- The Application:
  - Unified index of publications on tax issues
    - ▼ Published as HTML on a CD-ROM of the publications.



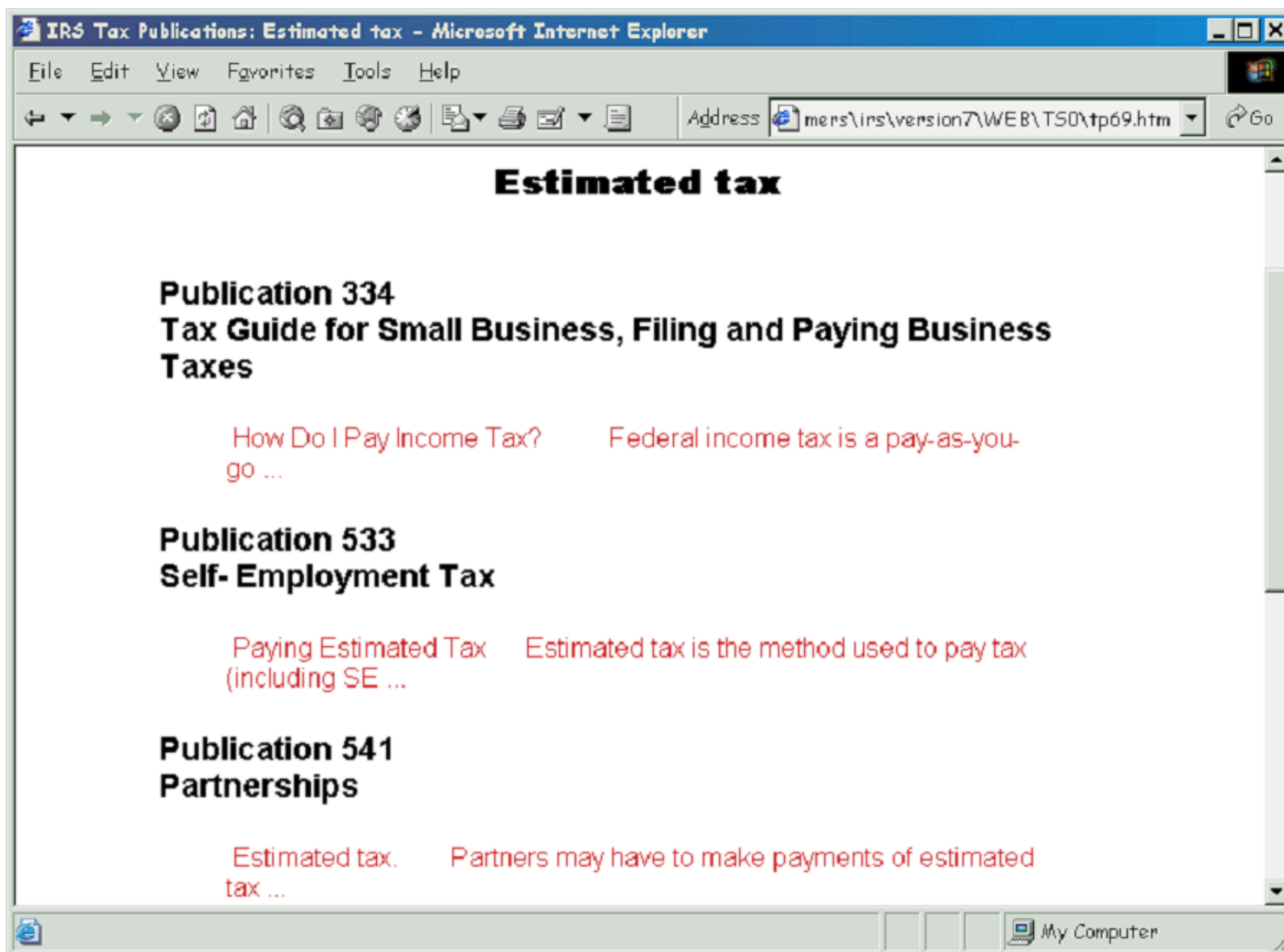
# IRS: Design

- Structured Sources
  - IRS already produced documentation from structured markup
  - Markup already contained harvestable index information
- Tightly Integrated Navigation
  - Links are directly inserted into the rendered content
    - ▼ Alphabetical index of topics
    - ▼ Navigate from topics in the index to other related topics or to the relevant information.
    - ▼ Topic occurrences are highlighted in the document content and provide a link back to the index.

# IRS: Navigation From Resources



# IRS: Navigation From Topics



# IRS: Benefits

- For the end user
  - Information is easily accessible
  - Topic-based navigation makes it simple to quickly locate all *relevant* information for a particular area of tax law.
- For the IRS
  - Promotion of the "index"
  - Quickly able to integrate indexes from a large number of publications
  - Combined indexes easily generated for different document sets
  - Revealed inconsistencies in the current index.

# Web Applications

# Introduction

- Web publishing challenges:
  - Dynamically generated content
  - Content tailored to viewer
  - Non-content data - e.g. users, groups
  - Containers and other organising principles: folder, project, subject area
- Not a suitable environment for a fixed schema
- Topic Maps are a data-driven structure
  - non-content data and content organisation may be determined at runtime

# Case 1 ITU: Introduction

- The Customer
  - ITU: The Network for IT Research and Competence in Education
  - Coordinates projects aimed at increasing use of IT into education
- The Developer
  - A collaborative development between Creuna and Ontopia
- The Application
  - A web site to:
    - ▼ Present information about organisations, people and projects
    - ▼ Provide an interactive forum for discussion and information sharing
    - ▼ Be populated by non-specialist users

# Case 1 ITU: Design

- Topic-base navigation structure
  - Topics represent concepts such as
    - ▼ Organisation
    - ▼ Project
    - ▼ Person
    - ▼ Report
  - Each topic has its own page
    - ▼ Content generated from topic occurrences
    - ▼ Links to other pages generated from associations to other topics
    - ▼ An "abstract" of the most closely related topics is presented in an engaging graphic display.



# Case 1 ITU: A Sample Page

hjem

**Status:** Pågående  
**Geografisk omfang:** Nasjonalt



## PROGRAM FOR LÆRERUTDANNING, TEKNOLOGI OG OMSTILLING (PLUTO)

← Klikk her for å åpne og lukke hovedmenyen

Program for LærerUtdanning, Teknologi og Omstilling (PLUTO) er et tiltak i ITUs spiss-satsing knyttet til innovativ omstilling av lærerutdanning. Programmet er forankret i KUFs handlingsplan "IKT i norsk utdanning - Plan for 2000-2003", der "IKT i lærerutdanningen" er ett av seks innsatsområder, med tilhørende årsplaner.

### Bakgrunn:

PLUTO programmet ble opprettet av ITU. Etter avviklingen av Arbeidsgruppen for Digitale Læremidler ble prosjekter under LUT overført og underlagt PLUTO. PLUTO består per oktober 2001 av i alt 10 prosjekter fordelt på 8 institusjoner, der pedagogisk, teknologisk eller organisasjonsmessig utvikling og omstilling av lærerutdanning ved bruk av IKT står i sentrum.

### Mål

PLUTO programmet har som overordnet mål: Å bidra til innovativ omstilling av lærerutdanning gjennom bruk av IKT.

Prosjektvirksomheten i PLUTO har som mål: Å utvikle pedagogiske og organisatoriske modeller for tilrettelegging og gjennomføring av studie- og læringsvirksomhet innen lærerutdanningen, der IKT utgjør en substansiell del.

WISE/VITEN ITU  
Vibeke Kløvstad Sten R. Ludvigsen  
DOCTA  
Program for Lære...  
Morten Søby  
Fra kaos til kun... IKT i flerkultur...  
skriv inn søketeksten

**Medvirkende:**  
Morten Søby →  
Vibeke Kløvstad →

**Består av:**  
Innovasjon med IKT i språklærerutdanningen →  
Heilskapleg integrering av IKT →  
IKT som endringsfaktor i lærerutdanningen →  
IKT i lærerutdanningen →  
DIA Danning, Informasjonsvurdering og  
Argumentering i naturvitenskap →  
FPPU - Fleksibel Praktisk-Pedagogisk Utdanning →  
IKT og nye læreprosesser →  
Det åpne rommet for læring og samspill →

**Rådgivere:**  
Vibeke Kløvstad →

**Har kontaktperson:**  
Sten R. Ludvigsen →

# Case 1 ITU: A Sample Page

hjem

**Tittel:** Prosjektleder  
**Telefon fast:** 22 84 05 91  
**E-Post:** morten.soby@itu.uio.no  
**Hjemmeside:** http://folk.uio.no/mortenso

Kharon 4a  
ITU  
Sten R. Ludvigsen

Ola Erstad  
Kirke-, utdanning...  
**Morten Søby**  
Program for Lærere  
Vilje, Sonja M. Mork  
Ola Erstad


ITU-magasinet

skriv inn søketeksten

søk

**MORTEN SØBY**

← Klikk her for å åpne og lukke hovedmenyen



**Bakgrunn:**

Cand.Polit.

Var som redaktør i PROFIL opptatt av den postmoderne tilstand. Har arbeidet med e-læring i NKS Fjernundervisning (87-91), og levd deltid som cyborg på nettet. Leder for 6th International Conference on Cyberspace (97). Har publisert en rekke artikler, rapporter om IKT og læring og om kulturens digitale felt. Kan alltid høres på Oslo Lufthavn Gardermoen gjennom en lyd-dusj - et hypertekstuel manifest om interaksjon mellom menneske og maskin ("Cyber Researcher" utviklet av Anna Karin Rynander). Redaktør av Futuramas arena: utdanning.

**Forfatter av:**

**RAPPORT**  
Workshop: Interaktivitet, teknologi og læring →

**ARTIKKEL**  
Bedre matematikkresultater med IKT →

**Medvirker i:**

**WEBTJENESTE**  
Futurama →

**PROSJEKT**  
IKT i flerkulturelle skoler i Oslo →  
Program for LærerUtdanning, Teknologi og Omstilling (PLUTO) →  
EduAction →  
DOCTA →  
IKT som mediator for kunnskapsproduksjon →

# Case 1 ITU: Implementation

- Reusable Framework
  - Base on the Zope database
  - Plans to integrate more tightly with Zope Content Management Framework
  - Planned for release as open source

# Case 1 ITU: Benefits

- Benefits for ITU
  - Intuitive navigation system
  - Topic map is integrated with the content management system
  - Easy to incorporate new content and new types of content into the site
  - System was delivered on time and within budget!

# Case 1 ITU: Benefits

- Benefits for Creuna
  - Topic maps provide the basic principles for site organisation
  - Reusable components
    - ▼ Reduces time taken to produce further topic map-driven web sites
  - Standards-based
    - ▼ Make use of existing and future tools for topic map creation and maintenance
  - Robust information interchange
    - ▼ Possibility of enabling related sites to more easily exchange information and links

# Case 1 ITU: Benefits

*Being our first Zope project and our first topic map projects, we still delivered on time and on budget. The customer can maintain and rearrange content, information structure and template layout through the web. The customer is very satisfied...We couldn't come up with a better concept in the couple of weeks we could spend on analysis and design."*

*Stian Danenbarger, Creuna*

# Case 1 ITU: Hurdles

- Hurdles

- Problems of adopting a new technology:

- ▼ Relatively immature tools

- ▼ Lack of "best practices" on topic map design

- Overcoming the hurdles

- Creuna believes that the development of more projects like ZTM in the open source community will help to address both of these problems

# Case 2 Patrimoine: Introduction

- The Customer
  - Patrimoine, publisher of professional and legal documentation
- The End User
  - Major financial institution
- The Developer
  - Mondeca
- The Application
  - Intranet for sales consultants, combining
    - ▼ Professional and legal documentation (from Patrimoine)
    - ▼ Product information (own and competitors)
    - ▼ Business process documentation
    - ▼ Financial simulators (Java applets)



# Case 2 Patrimoine: Design

- Integrated solution
  - Documentation held in 3rd party CMS (Documentum)
  - Topic map structures provided by Mondeca's Intelligent Topic Manager (ITM)
  - Navigation delivered via a web portal
- Multiple Views
  - Thesaurus
  - Hierarchical classification schema
  - Index of business subjects
- Editable topic map
  - Users can add new information to a topic or create new links between topics via the portal interface.

# Case 2 Patrimoine: Benefits

- For the end user
  - Wide span of documentation available from a single source
  - Tight integration of 3rd party information with business information
  - Improved productivity in sales and consulting
- For the information provider (Patrimoine)
  - Robust platform for delivery of highly tailored information sets
  - Integration with customers own documentation is a great selling point

# Application Development

# Introduction

- Topic maps provide:
  - Ability to model relationships between data objects
  - Data-driven schema for the object model
  - A single API for processing the data
- Allows developers to:
  - Easily modify and refine schema as application development progresses
  - Modify the schema of deployed applications
  - Enable users to configure or design their own schemas

# Case 1 Bravo: Introduction

- Knowledge management tool
  - Developed by GlobalWisdom Inc.
  - Combines explicit and automated categorisation with user feedback and user tracking to determine the most relevant documents to a given query
  - System identifies "experts" in subject areas and uses their profile to influence the categorisation of the documents they use

# Case 1 Bravo: Design

- Topic map stores
  - Document information
  - Information about the concepts found in one or more documents
  - User information
  - Concept relationships
- Uses the K42 topic map engine from Empolis

# Case 1 Bravo: Benefits

- Benefits for GlobalWisdom
  - Scalability and flexibility
  - Rich basic architecture
  - Use of 3rd party engine enabled more rapid development

*We used a topic map engine to provide a sophisticated and scalable information architecture. This serves as a differentiator for us and helps us to address more upscale markets.*

*Bryan Thompson, CEO GlobalWisdom*

# Case 1 Bravo: Benefits

- Benefits for customers
  - Intuitive, concept-centric organisation
  - Faster, more accurate access to relevant information
  - Reduced learning curve - users are never aware of creating or maintaining a topic map



# Application Integration

# Introduction

- Integration of meta data
  - Diverse meta data sources for a single entity
  - Diverse syntaxes for meta data values
  - Multiple APIs for accessing meta data
  - Links between applications become exponentially more complex
- Potential solutions
  - Map meta data properties
    - ▼ A headache even with just two systems to integrate
  - Define a common, neutral meta data set

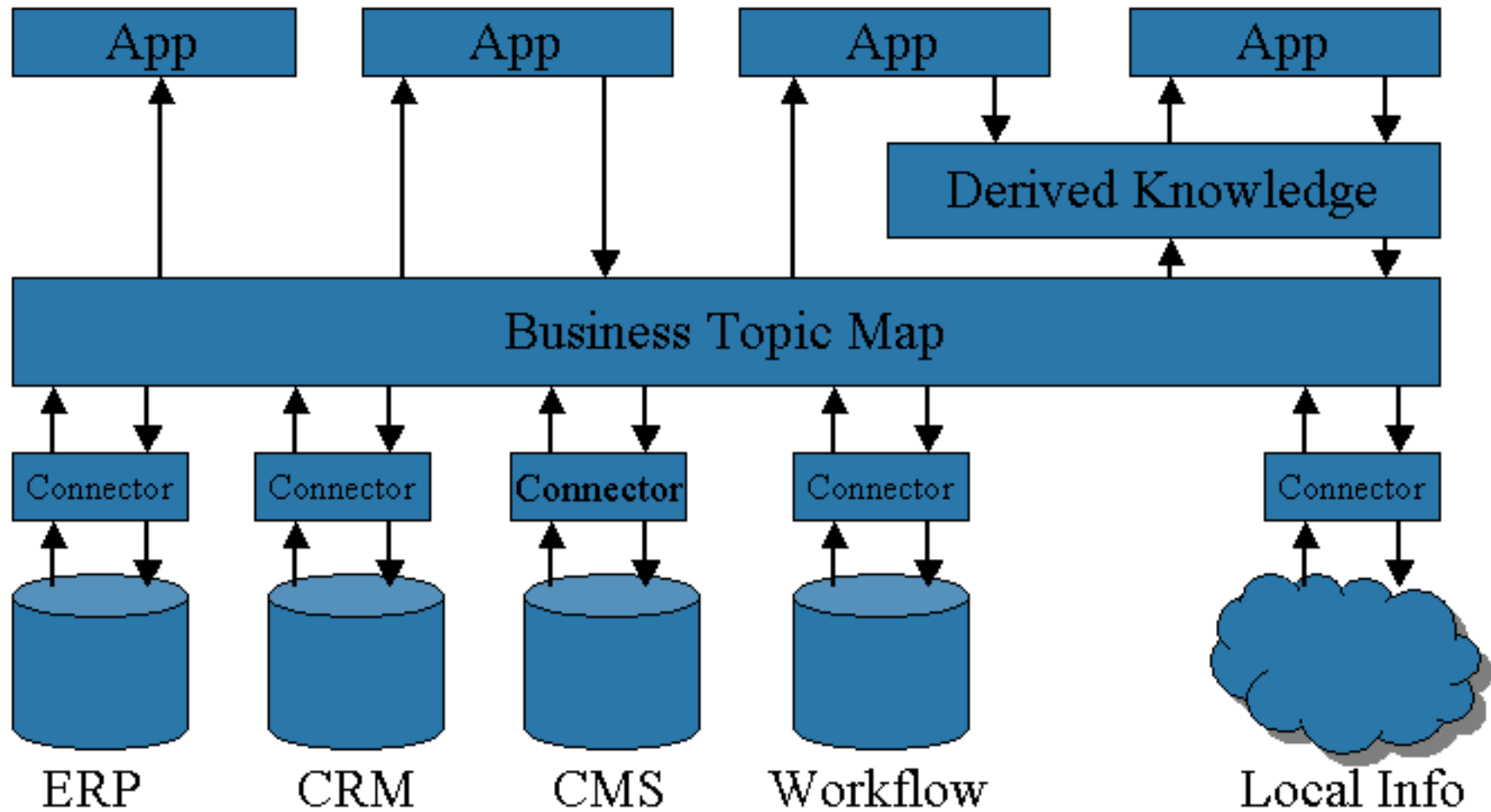
# Topic Maps In Application Integration

- A flexible data structure
  - Entities (topics)
  - Entity meta data (occurrences)
  - Relationships between entities (associations)
  - Sources of entities and entity meta data (subject indicators)
- Enables use of 3rd party tools
- Have a standardised interchange syntax

# Topic Map Approaches

- Topic map-based integration
  - Create a common topic map "schema"
  - Write a connector from each meta data source to the topic map schema
  - Developers access all topic-mapped information through a single API
- Connectors may be
  - *Static* - populate topic map from sources
  - *Dynamic* - interpret requests for topic map data into queries against the underlying resources, presenting the results as fragments of a topic map

# A Sample Architecture



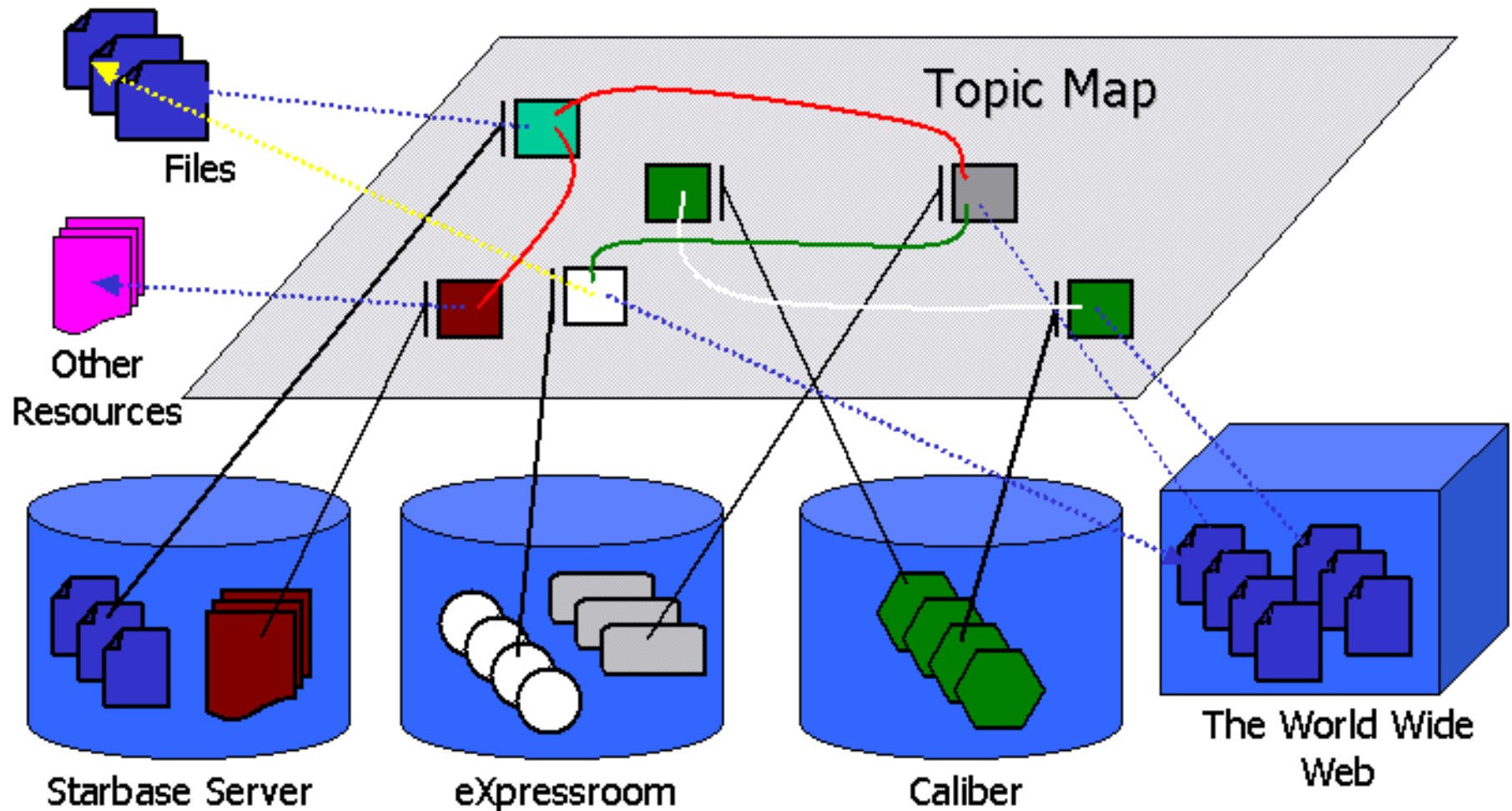
# Case 1 Starbase: Introduction

- Starbase provide collaborative solutions for business application development
- Flagship products
  - StarTeam - version control, incident tracking, workflow
  - Caliber RM - requirements planning
  - eXpressroom, Codewright - code and content development
- Each product has a separate repository
  - Repositories hold huge amounts of information which may be related in complex ways.
- End users have other sources of information
  - e.g Email applications

# Case 1 Starbase: Design

- Unified Information Space
  - Integrates information from a wide variety of different sources
    - ▼ Includes Starbase applications and third-party applications
  - Topic Maps provide the infrastructure for:
    - ▼ Relating information so that it can be found based on semantic queries
    - ▼ Guiding traditional text search engines
    - ▼ Navigating information spaces
    - ▼ Providing a framework for notifying users of "knowledge events"
  - Created using the Ontopia Knowledge Suite (OKS)

# Case 1 Starbase: Architecture





# Case 1 Starbase: Benefits

- Long-lived technology
  - XML-based
  - Defined by a standard
  - Proven interoperability
- Simple
  - Easy to get started with implementation
- Flexible
  - New data sources are easily integrated
- Extensible
  - Topic map structures open up applications beyond the original intention of the integration

# Case 1 Starbase: Hurdles

- New technology
  - Steeper learning curve
  - Scarcity of topic map applications and development methods

# Open Source Projects

# A Brief Word About Open Source

- A number of open source and free software packages:
  - Engines: TM4J, GooseWorks Topic Map Toolkit, tmproc
  - Editors: TMTab
  - Applications: Nexist, SemanText

# Engines

- **tmproc**
  - The very first topic map processor!
  - Written in Python
- **TM4J**
  - Java topic map engine with in-memory or persistent storage models.
  - Implementation includes a simple query language (tolog) and a number of utilities for processing topic map data.
- **GooseWorks Topic Map Toolkit**
  - Written in C with C and Python interfaces
  - In-memory and persistent storage models
  - Includes tools for manipulating and querying topic maps.

# Editors

- TMTab
  - Structured, forms-based editing environment
  - Based on Protege-2000
  - Export to XTM only (currently no round-trip)

# Applications

- **Nexist**
  - An experimental hyperdocument system
  - Intended as an educational tool using the constructivist approach
- **SemanText**
  - Semantic network application
  - Uses topic maps to hold the knowledge-base
  - Uses inference rules to derive new knowledge
  - Includes tools for extracting topics and associations from XML and RDF files.

# Conclusions



# State of the art

- Although new, topic maps are already finding a range of practical applications.
- Benefits identified by users/developers:
  - Expressive data structure
  - Intuitive model
  - Data-driven, flexible schema
  - Simple serialized format
  - Explicit rules for combination - enables knowledge-sharing
  - Implementations available