The International Foundation for Information Technology (IF4IT)

Open Global Best Practices

Semantic Web Series

Using Ontologies to Facilitate Better Enterprise Knowledge Management (EKM)

This Presentation Covers:

- A breakdown of tasks for good/comprehensive KM
- An understanding of why KM breaks down for most enterprises
- A framing of the Enterprise Knowledge Management problem
- An understanding of Knowledge Fabrics
- Components/constructs of an Ontology-driven Knowledge Fabric
- Best practices for implementing Ontology-driven Knowledge Fabrics
- Supporting Evidence

Authored and Presented By

IF4IT Best Practice Series

Company: The International Foundation for Information Technology (IF4IT)
Presenter: Frank Guerino
Position: Chairman & Founder
Note: Feel free to connect via LinkedIn

International Coverage Across All Verticals

Alliance for Permanent Access to Records of Science Network (APARSEN)	Conseil Europeen pour la Recherche Nucleaire (CERN)	Morgan Stanley Smith Barney
Associazione Nazionale Archivistica Italiana (ANAI)	Consorzio Interuniversitario Nazionale Per L'Informatica (CINI)	Museiimpresa (l'associazione italiana dei musei e degli archivi d'impresa)
Austrian Parliament, AT/AP	Credit Suisse	Science & Technology Facilities Council (STFC) (of the United Kingdom)
Bank of America Merrill Lynch	Food and Agriculture Organization (FAO) of the United Nations (UN)	U. S. Army Aviation and Missile Research Dev. and Engineering Center (AMRDEC)
British Library (BL), The	United States Defence Information Technology Center (US DTIC)	UBS
CINAV-PT Navy Research Center, Lisboa, Portugal	Joint Learning Network (JLN) for Universal Health Coverage	University of Deusto
Computer Sciences Corporation (CSC)	Lufthansa Systems AG	Wells Fargo

Reflection: Some KM Tools & Technologies

IF4IT Best Practice Series

Traditional (Beginner KM)

- Intranets
- General Content Management Systems (CMSs)
  - Drupal, Wordpress, etc.
- Document Management Systems (collaborative)
- Wikis
- Blogs
- Forums
- Instant Messaging/Chat
- Phones
- Online/Web Conferencing & Desktops
- Electronic Documents (Documents, Spreadsheets, Drawings, etc.)
- MS Sharepoint

Non-Traditional (Advanced KM)

- Business Domain Systems
  - Marketing, Product, Sales, Fulfillment, Support, etc.
- Electronic Libraries
- Enterprise Directories
- Asset Management Systems
- Architecture Modeling Tools (AMTs)
- Enterprise Service Catalogs
- Configuration Management Databases (CMDBs)
- Autodiscovery Tools/Software
- Customer Relationship Mgmt. (CRM) Systems
- Data Warehouses and Datamarts
- Big Data Solutions
- Semantic Data Stores/Repositories

Common Patterns

- People "incompletely" deploy these solutions
- They're partially populate with content (i.e. data & information and then abandoned for the "next" new shiny tool/technology/technology

Common Patterns

- Usually pertain to Domain-specific knowledge
- Requires high level of technical expertise and staff dedication (e.g. Engineers & Librarians)
- Better "localized" knowledge but poor sharing

Good KM is Not Simple or Free

IF4IT Best Practice Series

Sample of Enterprise Librarian's Work

1. Design, build, and deploy the Electronic Library Infrastructure
2. Design, build, deploy, and maintain changes to Enterprise Taxonomies
3. Design, build, deploy, and maintain changes to the Enterprise Master Catalog
4. Design, build, deploy, and maintain changes to Domain Catalogs
5. Design, build, deploy, and maintain changes to Domain or Classification specific Indexes
6. Design, build, deploy, and maintain Domain/Categorical Data and Information Views and View Infrastructure
7. Design, build, deploy, and maintain Domain/Categorical Reports and Report Infrastructure
8. Design, build, deploy, and maintain Domain/Categorical Dashboards and Dashboard Infrastructure
9. Design, build, deploy, and maintain Domain/Categorical Visualizations and Visualization Infrastructure
10. Design, build, deploy, and maintain Semantic Relationship infrastructure

NOTE: Most people throw up a Wiki, Content Management System, or a Sharepoint sit and ignore most of the above!

Sample of Author/Publisher's Work

1. Identify Item to Be Documented
2. Document details about Item
3. Register Item with all appropriate Taxonomy Tags (for search)
4. Register Item with each and every appropriate Catalog (Master, Domain, etc.)
5. Register Item with each and every appropriate Index
6. Create all "Semantic Relationships" or "Links" to/from every other Item that is related
7. Create alternate Data and Information Views
8. Create Item Related Reports
9. Update impacted Item Category Related Reports
10. Create Item Related Dashboards
11. Update impacted Item Category Dashboards
12. Create Item Related Data Visualizations
13. Updated impacted Item Category Data Visualizations
14. Repeat for Hundreds of Thousands (or even Millions) of Enterprise Items

REPEAT all of the above hundreds of thousands (or even millions) of times, in order to identify, document, categorize, curate, and publish all relevant Knowledge Artifacts (Articles/Publications).

Reality: Manual Effort Does Not Scale

IF4IT Best Practice Series

- In an information age, where almost everything is digital, un-harvested Knowledge (in the form of relevant data and information) that has yet to be made available for consumption by others, is constantly outpacing humans' ability to work with it.

- Enterprises do not have the funds or resources to document hundreds of thousands, millions, or even tens of millions of Knowledge Artifacts for the purpose of effective Knowledge Sharing

- Enterprises do not have the funds or resources to "stitch" data together, throughout the enterprise, in order to create a solid Knowledge Fabric

The Common End Result: People deploy and quickly abandon many different tools and technologies that have no strategic plans or support for long term evolution. They also abandon content development (*witness most Intranets, Wikis, and Content Management Systems*) when they realize how much work is required to evolve and maintain content.

Summarizing the Enterprise KM Problem

IF4IT Best Practice Series

The Wikipedia Problem

How does an enterprise build an enterprise-wide Knowledge Sharing solution, very much like an Enterprise Library, that is more comprehensive, has better features, and which has a higher level of quality than Wikipedia… doing so with far less human resources, far less time, and far less funding?

Reminder: Wikipedia took millions of people, more than a decade, to build nothing more than a loosely structured content repository, that doesn't handle data well, and that does not leverage advanced data analytics or visualizations for advanced knowledge determination.

Comparison Highlights the Challenges

IF4IT Best Practice Series

Wikipedia

A. Supports very little data (mostly just informational content)
B. Content is mostly "static"
C. Information is all in one place (articles are Sources of Truth)
D. Millions of people helped develop over more than a decade
E. No library Catalogs
F. No library Indexes
G. Does not leverage Taxonomies
H. No requirement for alternate views of content
I. No requirement for context oriented glossaries
J. No requirement for data oriented reports
K. No requirement for data visualizations
L. No Semantic Relationships

Any Good Enterprise Library

A. Must support content that includes, both, data as well as information
B. Must deal with highly dynamic data
C. Data and information comes from other Sources of Truth)
D. Very limited pool of resources (thousands to tens of thousands) with other priorities
E. Strong need for Catalogs/Cataloging
F. Strong need for Indexes
G. Must leverage Taxonomies
H. Strong requirement for alternate views of content
I. Strong need for context oriented glossaries
J. Strong need for data oriented reports
K. Strong need for data visualizations
L. Strong need for Semantic Relationships

What is Enterprise Knowledge Management?

IF4IT Best Practice Series

- **Enterprise Knowledge Management (EKM)** is the very broad but scientific attempt to facilitate knowledge collection and sharing, at an enterprise level, for any stakeholder, pertaining to any interest that facilitates getting answers to anyone's questions, regardless of their position or function, within the enterprise.

- EKM makes it easier for people to find **Retrospective Knowledge** (i.e. things we've learned and documented, in the past)

- EKM makes it easier for people to develop and discover new **Prospective Knowledge**

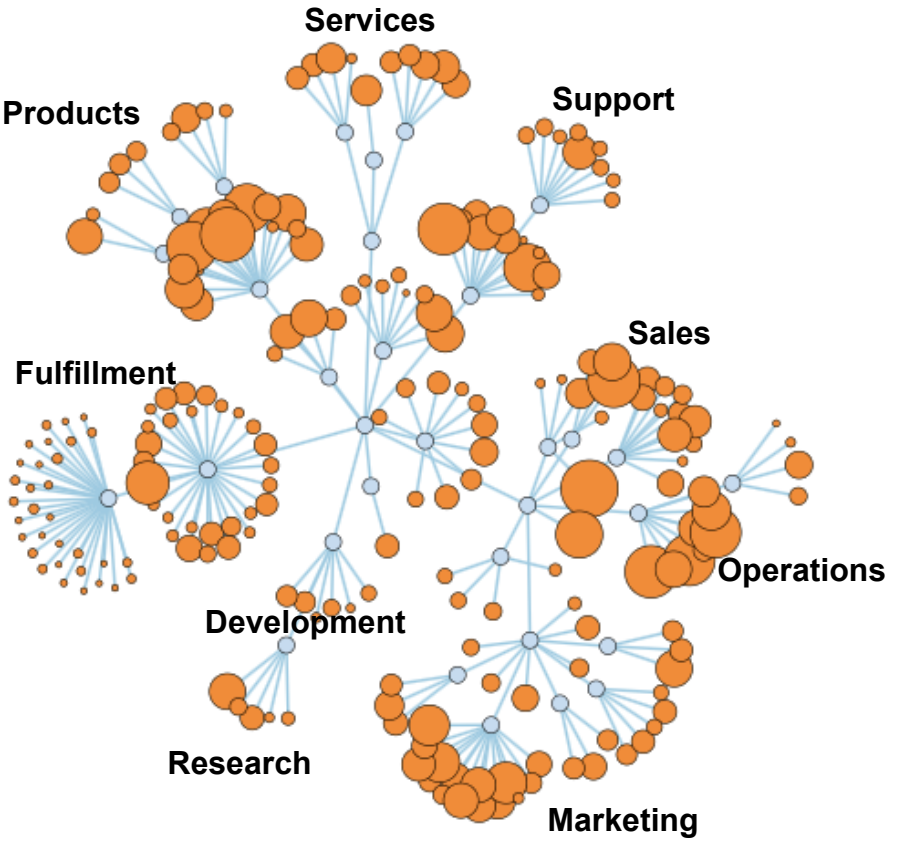
- Both, Retrospective and Prospective Knowledge are critical for any enterprise's success

What is a Knowledge Fabric?

IF4IT Best Practice Series

- **A Knowledge Fabric** is the end-to-end connectivity of knowledge, such that data and information flow from anyone or anywhere in the enterprise to any other person in the enterprise, regardless of their location, function, or purpose

- Knowledge Fabrics enable true **Enterprise Transparency**, by fluently tying together many areas of the enterprise, therefore allowing stakeholders to work more effectively and efficiently



Why Taxonomies Just Aren't Enough

IF4IT Best Practice Series

- Taxonomies are over-communicated and under-utilized
- Taxonomies are "only" a classification mechanism
  - While they help with context, they often don't help with true meaning or detailed structure of an entity
- In an Enterprise, classification is often context specific and may change based on role or function
- Enterprises often do a poor job of maintaining Taxonomy synchronization, across many systems and organizations
- **Most important:** Taxonomies represent only a small subset of all things that are important to an Enterprise

Ontologies are More Comprehensive

IF4IT Best Practice Series

- While Taxonomies only cover those things that can classified, *Ontologies cover all things that are meaningful*
- Therefore, an **Enterprise Ontology** covers all things that are meaningful to an Enterprise
- Done correctly, an Enterprise Ontology has many reusable constructs that are core to any Enterprise, regardless of its vertical industry or purpose
- Those things that fall outside such reusability are used to enrich an Enterprise's specific Ontology to include those things that are important to its vertical industry (i.e. its competitive space)

Components of an Enterprise Ontology

IF4IT Best Practice Series

(In no particular order and certainly never complete.)

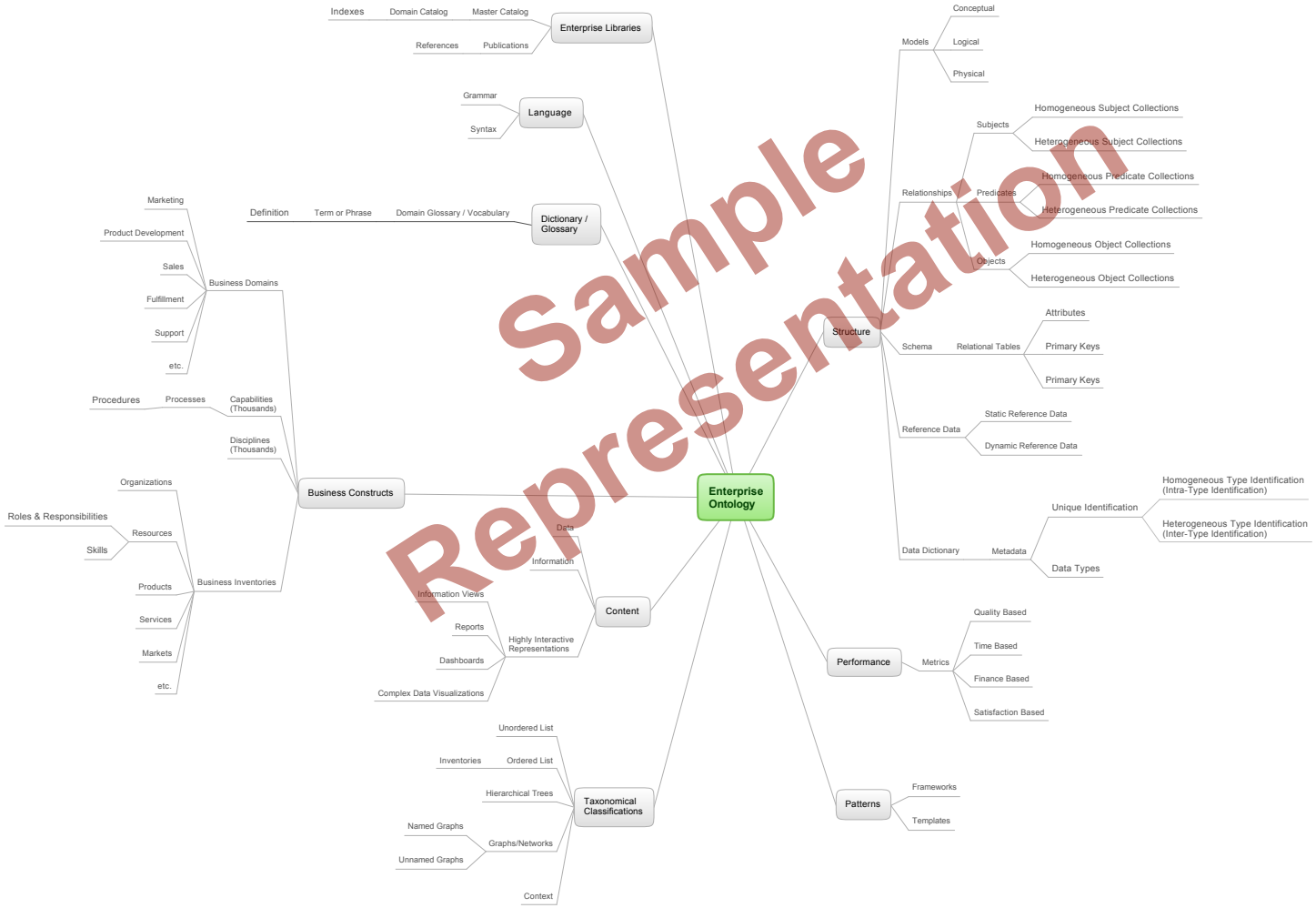
- Structure
  - Models
    - Conceptual
    - Logical
    - Physical
  - Relationships
    - Subjects
    - Predicates
    - Objects
  - Schema
    - Relational Tables
      - Attributes
      - Primary Keys
      - Primary Keys
  - Reference Data
    - Static Reference Data
    - Dynamic Reference Data
  - Data Dictionary
    - Metadata
      - Unique Identification
      - Data Types
- Performance
  - Metrics
    - Quality Based
    - Time Based
    - Finance Based
    - Satisfaction Based

- Patterns
  - Frameworks
  - Templates
- Taxonomical Classifications
  - Unordered List
  - Ordered List
    - Inventories
    - Controlled Vocabularies
  - Hierarchical Trees
  - Graphs/Networks
    - Named Graphs
    - Unnamed Graphs
  - Context
- Content
  - Data
  - Information
  - Highly Interactive Representations
    - Information Views
    - Reports
    - Dashboards
    - Complex Data Visualizations
- Enterprise Libraries
  - Master Catalog
    - Domain Catalog
      - Indexes
  - Publications
    - References

- Business Constructs
  - Business Domains
    - Marketing
    - Product Development
    - Sales
    - Fulfillment
    - Support
    - etc.
  - Capabilities (Thousands)
    - Processes
      - Procedures
  - Disciplines (Thousands)
  - Business Inventories
    - Organizations
    - Resources
      - Roles & Responsibilities
      - Skills
    - Products
    - Services
    - Markets
    - etc.
- Dictionary / Glossary
  - Domain Glossary / Vocabulary
    - Term or Phrase
      - Definition
- Language
  - Grammar
  - Syntax

Enterprise Ontologies Are Complex

IF4IT Best Practice Series



Critical Observations About Ontologies

IF4IT Best Practice Series

Challenges…

- They're complicated - Finding the skills that understand how to design, implement, and control such comprehensive Ontologies
- They're not clear - Convincing leadership to fund and allow such efforts
- Leveraging is difficult - There aren't too many truly reusable Enterprise Ontologies or related constructs

Benefits…

- Scale - They allow synchronization of massive quantities of data and information, across an enterprise
- The Next Level of KM - They allow quicker and easier achievement of higher order Knowledge Constructs

Example of Reusable Enterprise Ontology
The IF4IT Open Repository of Common and Reusable Resources
http://www.if4it.com/resources.html

Best Practices for Implementing

IF4IT Best Practice Series

1. Start with a well-maintained and well-organized Enterprise Electronic Library
   - Recommended to be your Intranet. Why introduce another tool when you already have one that reaches most of the enterprise?
2. Identify your Ontological structure(s) and design physical solutions to accommodate implementations
3. Use a Semantic Data Engine, or other Semantic binding structures, in order to ensure meaningful connectivity between Data Elements
   - This facilitates "stitching" for your Knowledge Fabric
4. Use Enterprise Inventories (sourced from throughout your enterprise) to seed and feed the Library
   - Inventories have "hidden" connectivity/relationships that allow you to bind things in one domain to things in other domains (e.g. Product XYZ, will have "traits" and some of those traits are Subjects or Objects in Semantic Relationships
5. Use software and rules to Harvest Semantic Relationships directly from elements within these Inventories to stitch elements of your Knowledge Fabric together
6. Minimize manual labor by embracing automation, wherever possible

Statistical Evidence Via Case Study

IF4IT Best Practice Series

An Electronic Library For A Very Well Known Insurance Company

Their Way

Quantities

Total Unique Topics:	11K
Total Web Pages:	11K
Links to "Documents":	Thousands (Manually created)

Time & Effort

Total Resources:	7 People Full Time
Total Duration:	8 Years
Total Effort Hours:	12,880 Person Hours (230 Hrs/Yr x 7 Pers x 8 Yrs) (Does not include time to set up and maintain technology)
Modification Recompilation Time:	13 hrs/week

Costs

Labor:	$1.29M (12,880hrs x $100/hr)
Technology:	$750K

The IF4IT Way

Quantities

Total Unique Topics:	33K (3X More)
Total Web Pages:	170K (15.5X More)
Links to "Documents":	Hundreds of Thousands

Time & Effort

Total Resources:	1 Person Part Time
Total Duration:	1 Week
Total Effort Hours:	8 Person Hours
Modification Recompilation Time:	13.5 minutes/week

Costs

Labor:	$1,600.00 (8hrs x $200/hr)
Technology:	$25K

End Result

Comprehensive Enterprise Ontologies, when applied correctly, resulted in a fraction of the cost, time, and energy.

Case Study Delivered Features

IF4IT Best Practice Series

What End Users Had Access To

Their Way

Library Features

Content Article Pages: Yes

Master/Enterprise Catalog: Partial (Manual)

Domain Specific Catalogs: Partial (Manual)

Indexes: Partial (Manual & Static Only)

Taxonomies: Partial (Manual)

Global Glossary: No

Controlled Context Vocabularies: Partial (Manual)

Semantic Relationships: No

Alternate Data & Info Views: No

Dynamic Sortable Tables: No (Static Only & non-"Drillable")

Interactive Reports: No

Interactive Dashboards (w/ Metrics): No

Interactive Data Visualizations: No

Abbreviations/Acronyms Catalog: Partial (Manual)

The IF4IT Way

Library Features

Content Article Pages: Yes (Manual + Automated)

Master/Enterprise Catalog: Yes (Automated)

Domain Specific Catalogs: Yes (Automated)

Indexes: Yes (Interactive & "Drillable, Automated)

Taxonomies: Yes (Automated)

Global Glossary: Yes (Automated)

Controlled Context Vocabularies: Yes (Automated)

Semantic Relationships: Yes (Automated)

Alternate Data & Info Views: Yes (Automated)

Dynamic Sortable Tables: Yes (Automated & "Drillable")

Interactive Reports: Yes (Automated)

Interactive Dashboards (w/ Metrics): Yes (Automated)

Interactive Data Visualizations: Yes (Automated)

Abbreviations/Acronyms Catalog: Yes (Automated)

End Result

Comprehensive Enterprise Ontologies, when applied correctly, resulted in many more feature rich Knowledge Constructs for End Users.

Closing…

IF4IT
Best Practice Series

Enterprise Ontologies for Better Enterprise Knowledge Management

- Yield Lower Costs
- Yield Quicker Turn-Around/Delivery Times
- Yield Higher Quality
- Yield More Features and Functionality for End Users
- Yield Higher Volumes of Knowledge Collection, Publishing, & Sharing
- Yield Higher End User Satisfaction

Questions/Comments?

(Note: Please feel free to contact and connect, directly, with the presenter for more detailed discussions on this topic.)