Building a Global View of Your Data

Dick Weisinger
dweisinger@formtek.com
Twitter: @DickAtFormtek

May 20, 2010
Discussion Points

- Introduction
- Importance of Information to Business
- Challenges of Managing Content
- Content as a Service (CaaS)
- CMIS: a new standard to access content
  - History
  - Use Cases
Who am I?

- Dick Weisinger
- Vice President and Chief Technologist, Formtek, Inc
- 20+ years of experience in Content, Document, Image, and Records Management
- Regular blogger at
  http://www.formtek.com/blog
  http://keytocontent.blogspot.com
Formtek?

- An ECM software and services company
- Experts in general ECM and CM space
- Expertise in engineering data management
- Formtek Orion ECM Software
- Alfresco Gold Integration Partner
  - Largest privately-owned Open Source Company
  - Sharepoint-like Collaboration Features
  - First Implementation of CMIS 1.0 Standard
High Quality Information

- Can help Businesses:
  - Make Better Decisions
  - Serve Customers Better
  - Gain Competitive Advantage

Business Intelligence is the #1 priority of 1500 CEOs
Gartner, 2009

From www.foresightwm.co.uk
85% of organizations say Information is a strategic Business Asset.

95% of organizations say Information is essential to Business Success.

Forbes survey: Managing Information in the Enterprise, April 2010
Ackoff’s DIKW Hierarchy

1989

Wisdom

Knowledge

Information

Data

Content

Content grounds the Business Intelligence Pyramid

CKW Pyramid

formtek
Content is King

Source: Wikimedia Commons
If HP knew what HP knows, we would be three times as profitable.

-- Lew Platt
Former CEO of Hewlett-Packard
The Problem

- Content is a Vital Asset to Businesses

But...

- Problem: Content is too hard
  - Too Hard to Find
  - Too Hard to Access
  - And.. Much much too much of it
Barriers to Unifying Content

- Findability
  - 62% say findability is “imperative or significant” to their business success – AIIM
  - Half say they can’t find the data they need -- Forrester

- Accessibility
  - Data is often siloed across multiple systems, repositories, and spreadsheets

- Content Growth
  - Enterprise data volume doubles every 18 months – Gartner Research / IDC
The Network is the Computer
-- Sun Microsystems

The Network is the Content
The Content is Out There...

It just needs to be found...

From http://omega.enstb.org/yannis/cours/s31/i-want-to-believe.jpg
Content Resides in Repositories

Capture → Storage → Security → Services

- Storage
- Full-Text Indexes
- Metadata
- Volume

Metadata Services
Multiple Repositories

>6

DM    WCM
Application Repositories

- Often the application is inseparable from the data repository
Social Media Repositories

Wiki
Blog
Collaboration

Social Networks
Microblogging
The Cloud & Hybrid Systems
Siloed Content

- Structured Data (Database)
- Unstructured Data
- Communication Streams
- Social Media Streams
- Enterprise Applications
- Cloud Applications
Data Fragmentation

- Multiple Repositories
- Inconsistencies
- Duplication

Photo by horiavarlan (Flickr)
Master Data Management

- Data Dictionary
  - Common Vocabulary
  - Single Version of Truth
- Data Centralized in a Hub
  - MDM Registry Repository
- Needs to be On-going
  - Once and Done == Failure

Average G5000 company spend on an MDM project is $5.2 million over 12-24 months
-- The MDM Institute
Data Virtualization

- **Content as a Service** (CaaS)
- Map of Trusted Data Sources
  - Like a GPS Locator for the Enterprise
- Data is Presented as a Service
  - Single Access Point
  - Manage/View all Data regardless of data source or physical location
- Success depends on Governance

Source: flickr: aburt
Content Management Interoperability Services

“SQL for Document Management”
What is CMIS?

- Content Management Interoperability Services
  - Defines a lowest-denominator CM capability set
    - Objects Supported
      - Documents
      - Folders
      - Relationships [Source -> Target Object]
      - Policies [e.g., ACL or retention management policies]
  - A single application works identically with content from any CMIS vendor
  - Uses either SOAP or AtomPub REST web services
CMIS Timeline

- 2005: iECM sees need for new standard
- 2006: EMC, IBM, Microsoft begin development
- 2007: CMIS announced and released to OASIS
- 2008: CMIS 1.0 ratified. Work on 2.0 begins
- 2009: CMIS 1.0 released for public comment
- 2010:
Use Cases:
Mash-Up / Aggregation

- Data Virtualization
- Mashups / Portals
- Federated Search
- eDiscovery
- Business Workflows
- Gradual Migration from old to new repository
http://www.aiim-iecm.org/

Thomas Pole and Laurence Hart
Use Cases: Inter-Repository

- Federated Records Management
- Publish to Web
- MDM
- Migration Tool
Use Cases:
Write Once, Run Anywhere

- A single software app / extension
  - Portable to work with any CMIS repository
  - No need to learn proprietary CMS API

- Examples
  - Content-enabled Enterprise Applications
    - ERP, BPM, CRM
  - Vertical Applications
    - Case Management
    - Collaboration
    - Electronic Health Records
    - Plugin to Office Suite
CMIS Integrations Appearing

- Microsoft SharePoint Connector (June 2010)
- Alfresco Share Connector (Available Now)
  - Alfresco + (Lotus/Quickr | Drupal | Joomla)
- Nuxeo CMIS-enabled Digital Asset Mgmt
- EMC/Documentum CMIS (December 2010)
- Also Exo, Day, Jahia, WeWebU, ...
Post CMIS 1.0: Uncharted

- Business Process Management, WF
- Web Content Management
- Digital Asset Management
  - Streaming
- Records Management, Compliance
- Security: ACLs
- Metadata: Aspects
- Browser Binding
  - JSON
- ...

...
CMIS Resources

- OASIS CMIS documentation
  http://www.oasis-open.org/committees/cmis/

- Gilbane Group Beacons Report
  - CMIS – Addressing Contemporary Requirements for Content Integration
    http://gilbane.com/beacons.html

- CMIS Chemistry Project
Summary

- The Network is the Content
- Business Information is Siloed
- Develop a Common Vocabulary
- Expose Content as Services (CaaS)
- CMIS is now a standard!
  - It will help make it easier to unite separated Content repositories
Thank You!