

Alfresco as a Metadata hub

Robin Bramley
Chief Scientific Officer, Ixxus
@rbramley



A Global Strategic-level Alfresco partner

Working with Alfresco since Alfresco v0.6

Over 140 Alfresco implementations

Alfresco & WCM Architecture Design Patterns

First worldwide implementation of the first Alfresco WCM product in 2007

Excellent Alfresco knowledge and highly trained and experienced staff

Alfresco Certified Engineers and Alfresco Certified Administrators

Authoring of leading Alfresco books

Alfresco Technical blogs <http://www.ixxus.com/blog/>

Multiple speaking engagements at the Alfresco Conferences (2011-2014),

Alfresco Days and BeeCon 2016

Alfresco Million \$ Club (May 2012) - Only 2 companies awarded

Alfresco Solution Partner of the Year 2013 and 2014

Alfresco Best Solution 2015 and Partner of the Year 2015 (Northern Europe)

Wholly owned subsidiary of Copyright Clearance Center

A Safe Pair of Hands

Recognised as one of the largest, most experienced Alfresco partners in the world

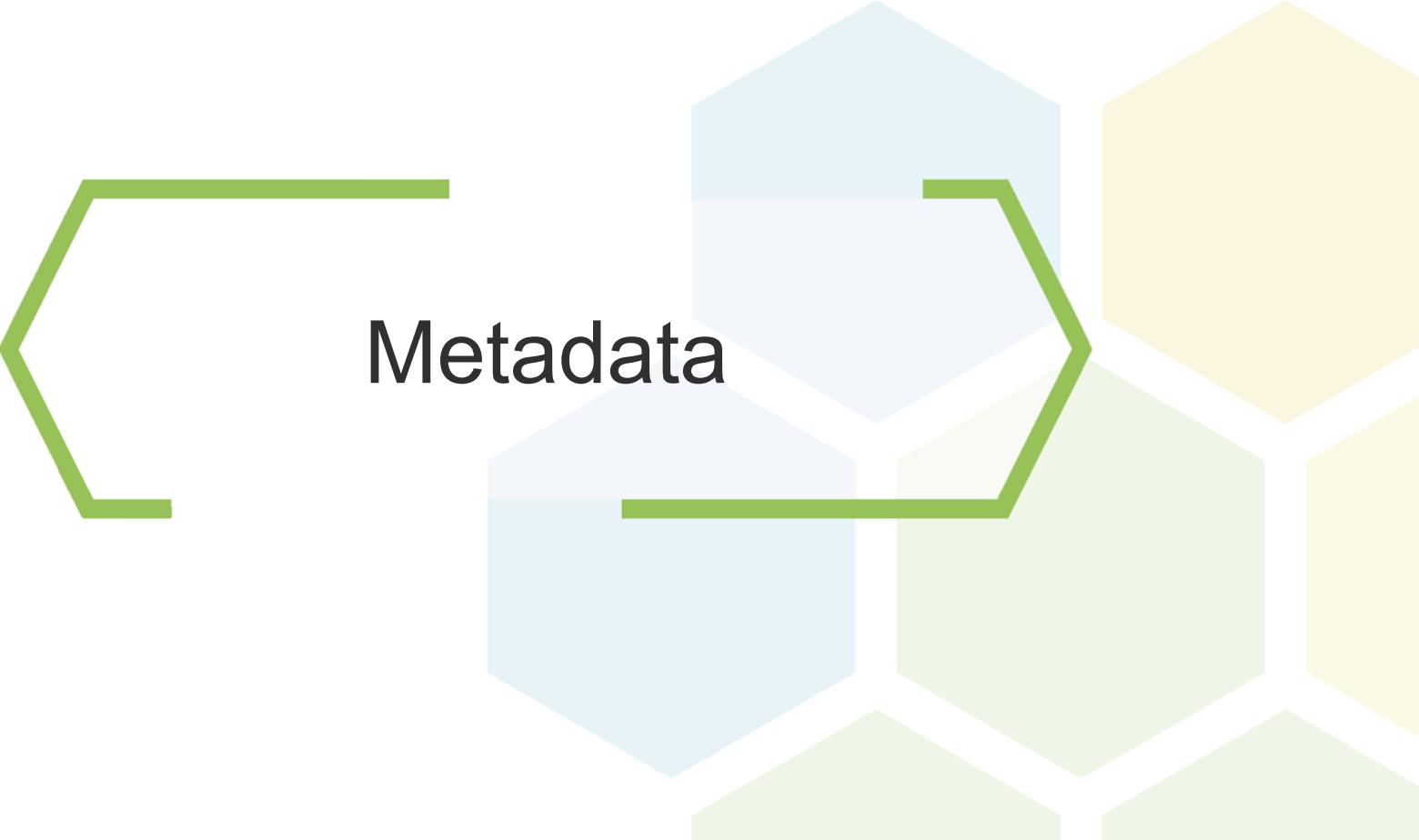


Presented at:



Published in:



A background pattern of overlapping hexagons in various shades of light blue, light green, and yellow, creating a sense of depth and connectivity.

Metadata

meta-

1. a prefix appearing in loanwords from Greek, with the meanings “after,” “along with,” “beyond,” “among,” “behind,” and productive in English on the Greek model: *metacarpus; metagenesis*.
2. a prefix added to the name of a subject and designating another subject that analyzes the original one but at a more abstract, higher level: *metaphilosophy; metalinguistics*.
3. a prefix added to the name of something that consciously references or comments upon its own subject or features:
a meta-painting of an artist painting a canvas.

```
@article {Dictionary.com2017,  
    title = {Random House Dictionary},  
    month = {Apr},  
    day = {10},  
    year = {2017},  
    url = {http://www.dictionary.com/browse/meta},  
}
```

**METADATA IS A
LOVE NOTE
TO THE FUTURE**



<https://www.flickr.com/photos/centralasian/8071729256>



metadata
“If music be the food of love, play on;”

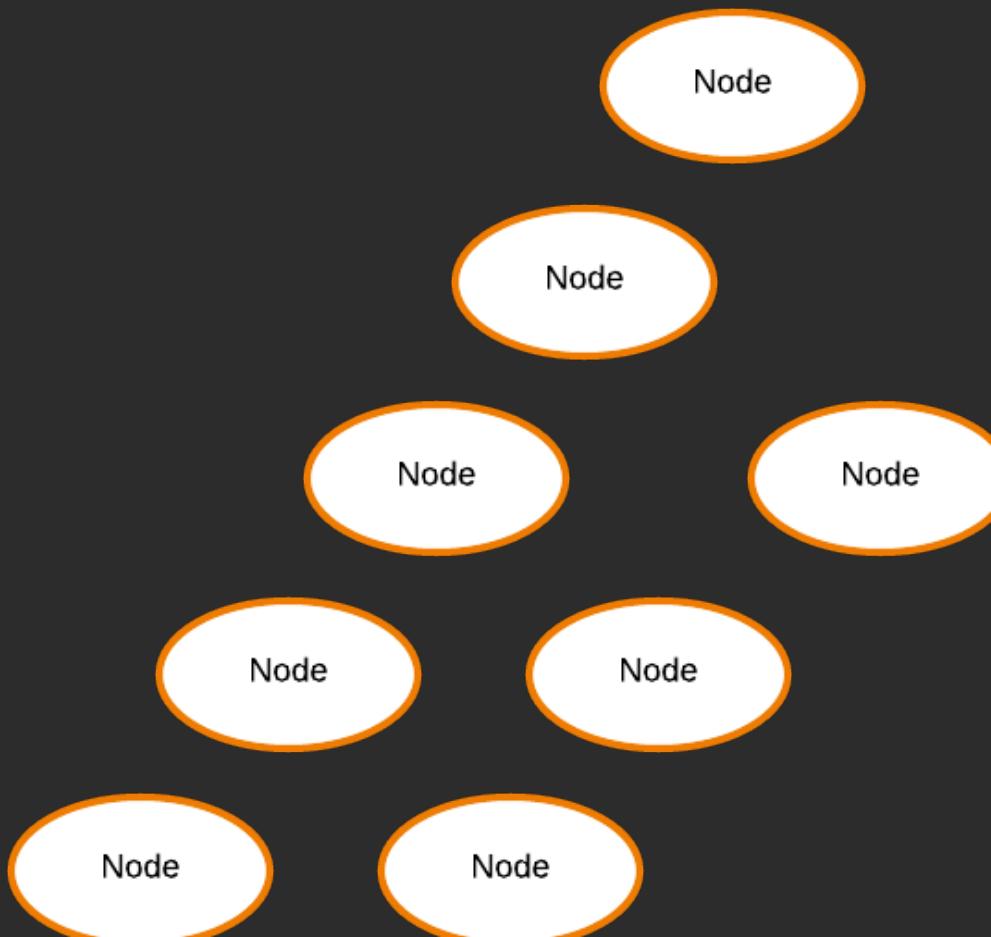
Orsino, Shakespeare’s Twelfth Night



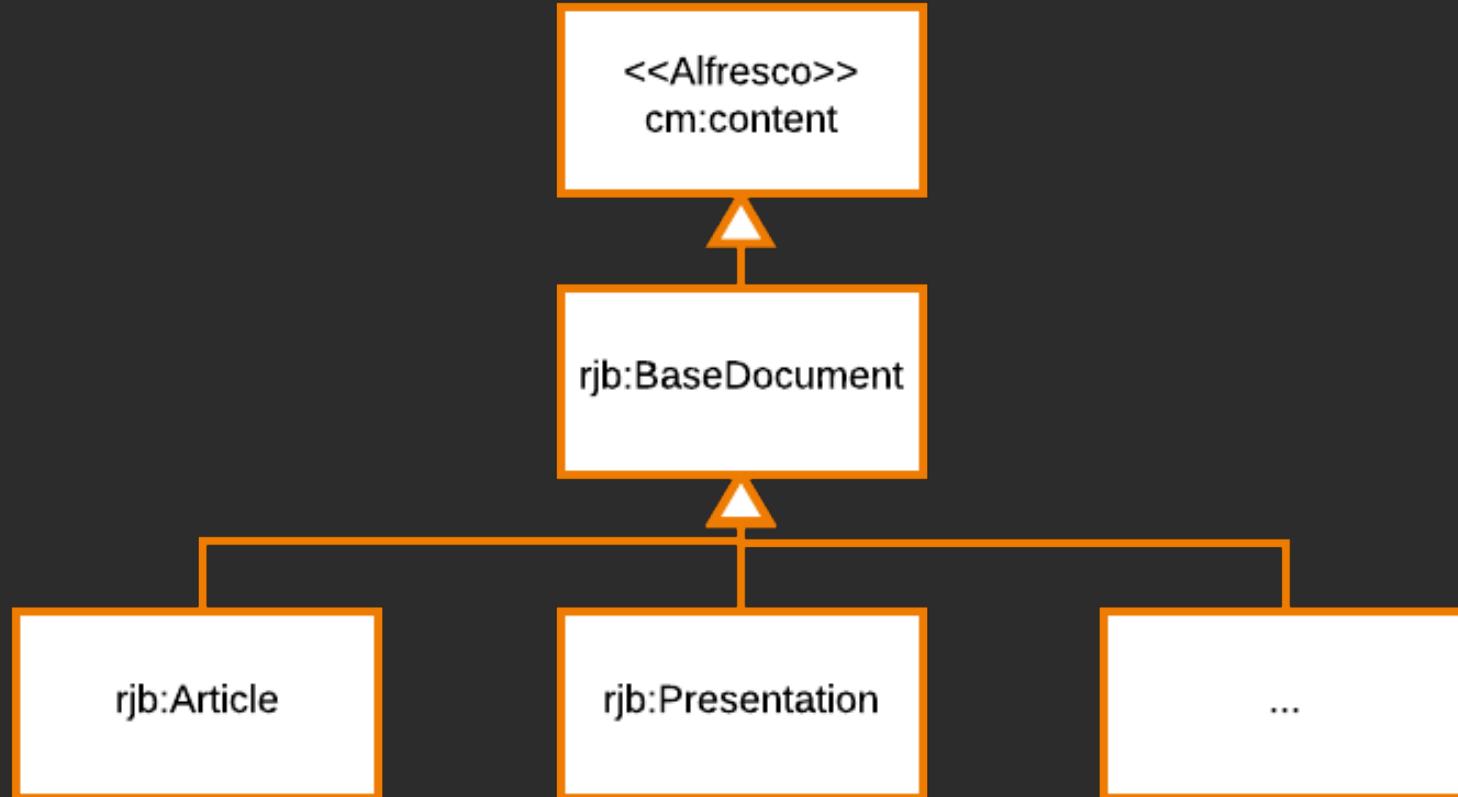
OOTB



Nodes



Type Model



Model admin



Model Manager

▼ Tools

- Application
- Category Manager
- Module Browser
- Node Browser
- Search Manager
- Tag Manager
- Model Manager
- Sites Manager
- Repository
- Replication Jobs
- Users and Groups

RJB-Knowledge

[<< Show Models](#) [Create Custom Type](#) [Create Aspect](#)

Custom Types

Name	Display Label	Parent
rjb:BaseDocument	Document	cm:content
rjb:Presentation	Presentation	rjb:BaseDocument
rjb:Article	Article	rjb:BaseDocument

Properties

- Properties allow us to capture metadata values
- Numbers, date / datetime, boolean, text & multi-lingual text
- *Mandatory or optional
- Multi-valued
- Constrained

Edit Properties: Bramley Metadata Hub.pptx

* Required Fields

Name:	*	Bramley Metadata Hub.pptx
Title:	Alfresco as a Metadata hub	
Description:	BeeCon 2017 session	

Constraints

- Regex
- min/max length
- min/max value
- pick from a list
- or use a custom Java constraint class

Edit Properties: Bramley Metadata Hub.pptx

Name: *
Bramley Metadata Hub.pptx

Title:
Alfresco as a Metadata hub

Description:
BeeCon 2017 session

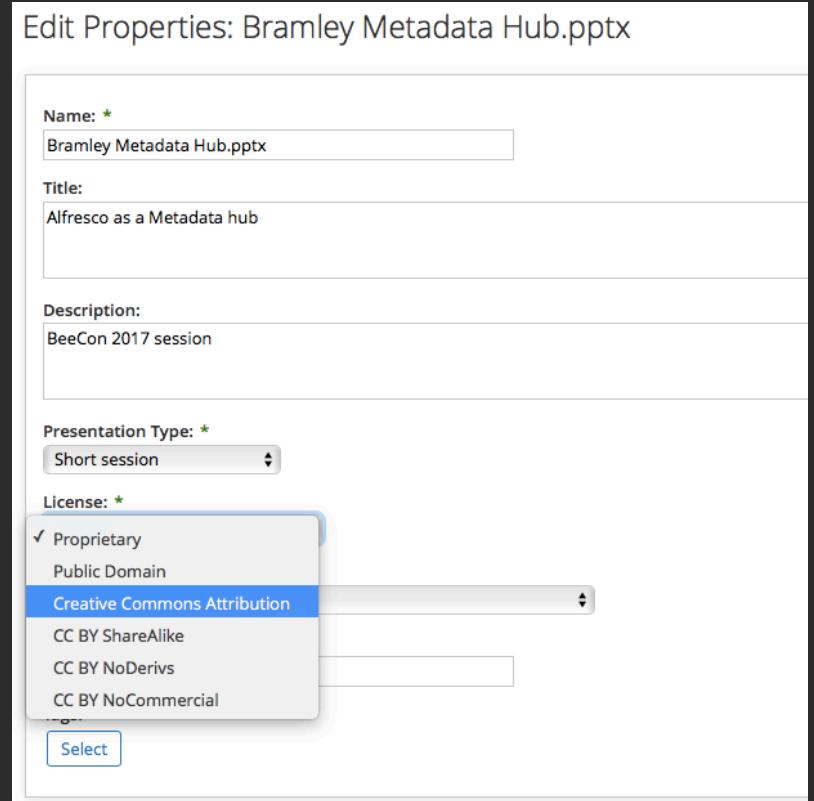
Presentation Type: *
Short session

License: *

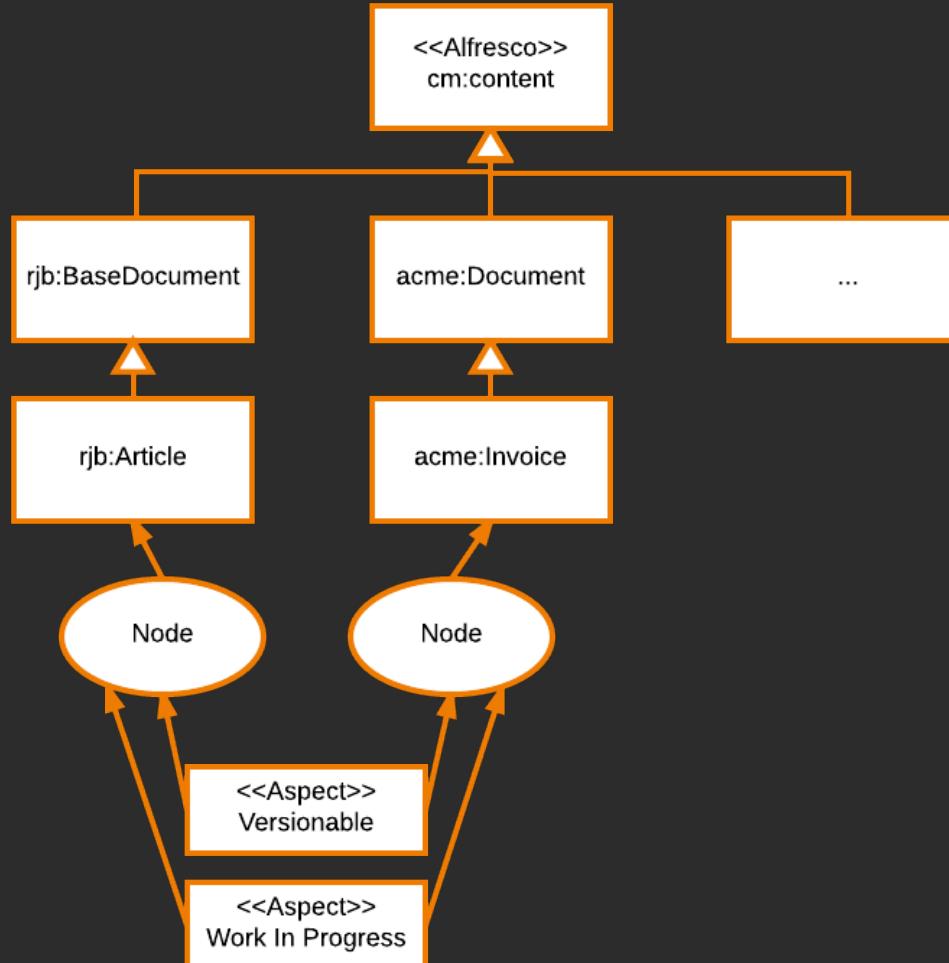
Creative Commons Attribution

✓ Proprietary
Public Domain
Creative Commons Attribution
CC BY ShareAlike
CC BY NoDerivs
CC BY NoCommercial

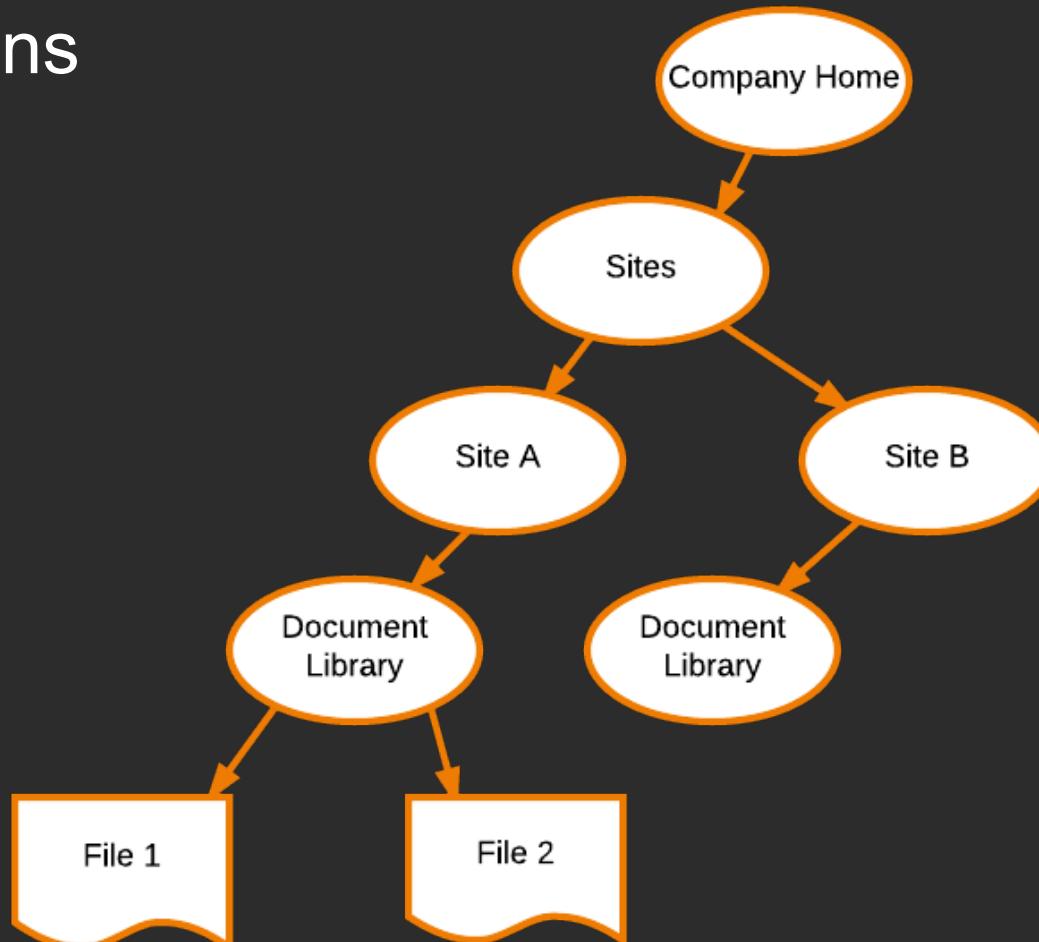
Select



Aspects



Associations



Browse

 Repository Browser

▶ Documents
▼ Repository
 └ Repository
 ├ Data Dictionary
 ├ Guest Home
 ├ Imap Attachments
 ├ IMAP Home
 ├ Shared
 └ Sites
 └ presentations
 └ documentLibrary

Select ▾ + Create... ▾ Upload Selected Items... ▾

Repository > Sites > presentations > documentLibrary

Conferences	Modified a minute ago by Robin Bramley
	Conference presentations
No Tags	
★ Favorite	Like 0
	Comment
Meetups	Modified just now by Robin Bramley
	No Description
meetup	
★ Favorite	Like 0
	Comment

1 - 2 of 2 << 1 >>

Search

Search Moderated

Site Dashboard Document Library Site Members Site News Site Settings

Search in: Presentations ▾

Jenkins Search

Filter by:

- Creator Robin Bramley 4
- File Type Microsoft PowerPoint 3
- Microsoft PowerPoint 2007 1
- Created
- Size 100KB to 1MB 1
- 1 to 16MB 3
- Modifier Robin Bramley 4
- Modified

Selected Items... 4 - results found

- Bramley_DevOps.pptx (Applying DevOps to Alfresco) ⓘ
Modified 10 hours ago by Robin Bramley | Size:4 MB | Site: Presentations
In folder: /Conferences/Alfresco Summit 2013/DevOps
- LondonGGUG_2010_10_Grails_Hudson.ppt (Grails & Hudson) ⓘ
Modified 8 hours ago by Robin Bramley | Size:393 KB | Site: Presentations
In folder: /Meetups/London GGUG
- LuceneEurocon2011_Bramley_Solr_at_VMG.ppt (Solr at VirginMoney Giving) ⓘ
Case study presented at Lucene EuroCon 2011 in Barcelona
Modified 10 hours ago by Robin Bramley | Size:1 MB | Site: Presentations
In folder: /Conferences/Lucene EuroCon 2011
- Bramley - Grails ECM DevCon Slides.ppt (Using Enterprise Content in Grails) ⓘ
Modified 10 hours ago by Robin Bramley | Size:1 MB | Site: Presentations
In folder: /Conferences/Alfresco DevCon 2011/Grails

Relevance Relevance

Name

Title

Description

Author

Modifier

Modified date

Creator

Created date

Size

Mime type

Type

Advanced search

Look for: **Content** ▾

Keywords:

Name:

Title:

Description:

Mimetype:

Modified Date:
From: To:
DD/MM/YYYY DD/MM/YYYY

Modifier:

Configured Advanced Search

Advanced Search

Look for: ACME Contracts ▾

Keywords:

Name:

Title:

Description:

Mimetype:

Unknown

Modified Date:

From: To:

DD/MM/YYYY DD/MM/YYYY

Modifier:

ACME Document Information

ACME Document Id:

Security Classification:

Contract Information

Contract Name:

Contract Id:

Required Fields

Categories

Presentations Moderated

Site Dashboard Document Library

Selected Items... ▾

Category Root > Regions > EUROPE > Southern Europe > Spain

▶ Documents ▶ Library ▾ Categories

Category Root

- Languages
- Regions
 - AFRICA
 - ASIA
 - EUROPE
 - Eastern Europe
 - Northern Europe
 - Southern Europe
 - Albania
 - Andorra
 - Bosnia and Herzegovina
 - Croatia
 - Gibraltar
 - Greece
 - Holy See
 - Italy
 - Malta
 - Portugal
 - San Marino
 - Slovenia
 - Spain

Bramley_2013_Scalability_with_notes.pptx (Beating the benchmarks)
Modified 9 hours ago by Robin Bramley 11 MB
No Description
alfresco graph scalability
EUROPE, Spain, English
Favorite Like 0 Comment Share

Bramley_DevOps.pptx (Applying DevOps to Alfresco)
Modified 10 hours ago by Robin Bramley 4 MB
No Description
alfresco puppet vagrant jenkins
EUROPE, Spain, English
Favorite Like 0 Comment Share

Bramley_MemoryLane_LT.pptx (Memory Lane ignite talk)
Modified 9 hours ago by Robin Bramley 11 MB
No Description
alfresco
EUROPE, Spain, English
Favorite Like 0 Comment Share

LuceneEurocon2011_Bramley_Solr_at_VMG.ppt (Solr at VirginMoney Giving)
Modified 10 hours ago by Robin Bramley 1 MB
Case study presented at Lucene EuroCon 2011 in Barcelona
solr spring java case study
EUROPE, Spain, English

Friction



<https://www.flickr.com/photos/gtweb/11328349694>

All In One Edit

“Bring efficiency to content creation and manual enrichment tasks in an editorial workflow by being able to read / edit content and metadata in one single interface.”

B I U | E A A Ix

Hotel Baldwin

Middle Eastern themed hotel on the **waterfront** with a world class Mezze restaurant led by chef Benoit Haffreingue. All rooms have four poster beds and separate seating areas meaning you can easily relax after enjoying the bustle of the neighbouring fish market and cathedral quarter. Don't forget to look up when you pass through the souk themed lobby area - the painted ceiling is truly spectacular!



Price: \$\$\$
Restaurant: Yes
WiFi: Yes

Residence Coquelin

Stay in this charming family run guesthouse under the walls of the castle for a true Northern French experience. Proprietor Blaise Monluc ensures everyone has a warm welcome with a glass of wine on the shady terrace, before guests retire to traditional rooms with beamed ceilings and traditional French fin de siècle furnishings. While the Residence Coquelin does not have a restaurant it is very close to an array of great places to eat so guest will certainly not go hungry!

Price: \$\$
Restaurant: No
WiFi: Yes

Chateau Maurice

Name: *
hotels-in-boulogne.html

Authors:
 Grace Hopper

Article Type: *
Review Word Count: 251

Approval

Assigned To:
ian Synge (ian.synge)

Lifecycle Status:
SME Review Due Date:
07 Jun 2016 dd MMM yyyy HH:MM (24 Hour)

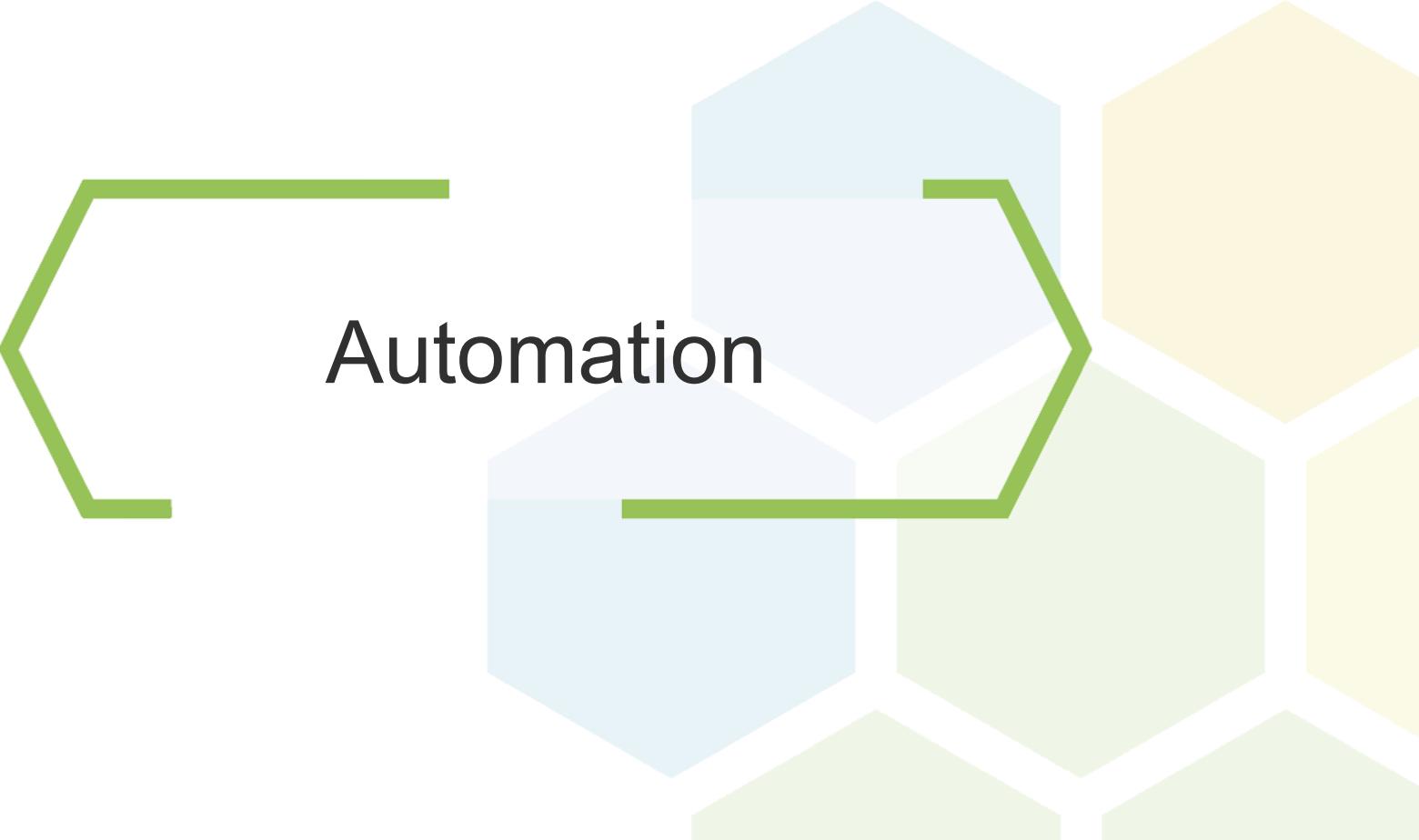
Access Level

Reading Level: *
One

Rights

“If music be the food of love, play on;
Give me excess of it; that surfeiting,
The appetite may sicken, and so die.”

Orsino, Shakespeare’s Twelfth Night

A background pattern of overlapping hexagons in various shades of light blue, light green, and yellow, creating a sense of depth and texture.

Automation

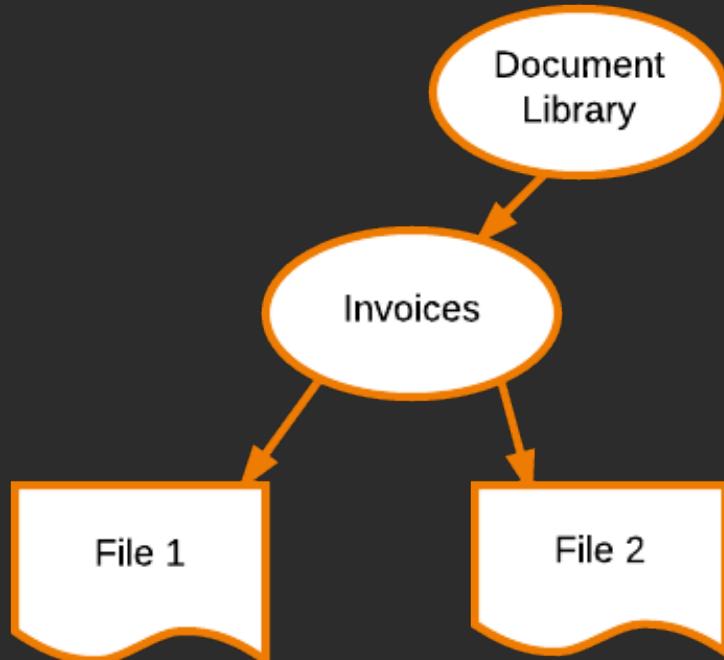
Metadata Extraction

1. Apache Tika FTW!
2. XPath extraction from XML files



Inheritance from the parent

E.g. Glean context from the folder

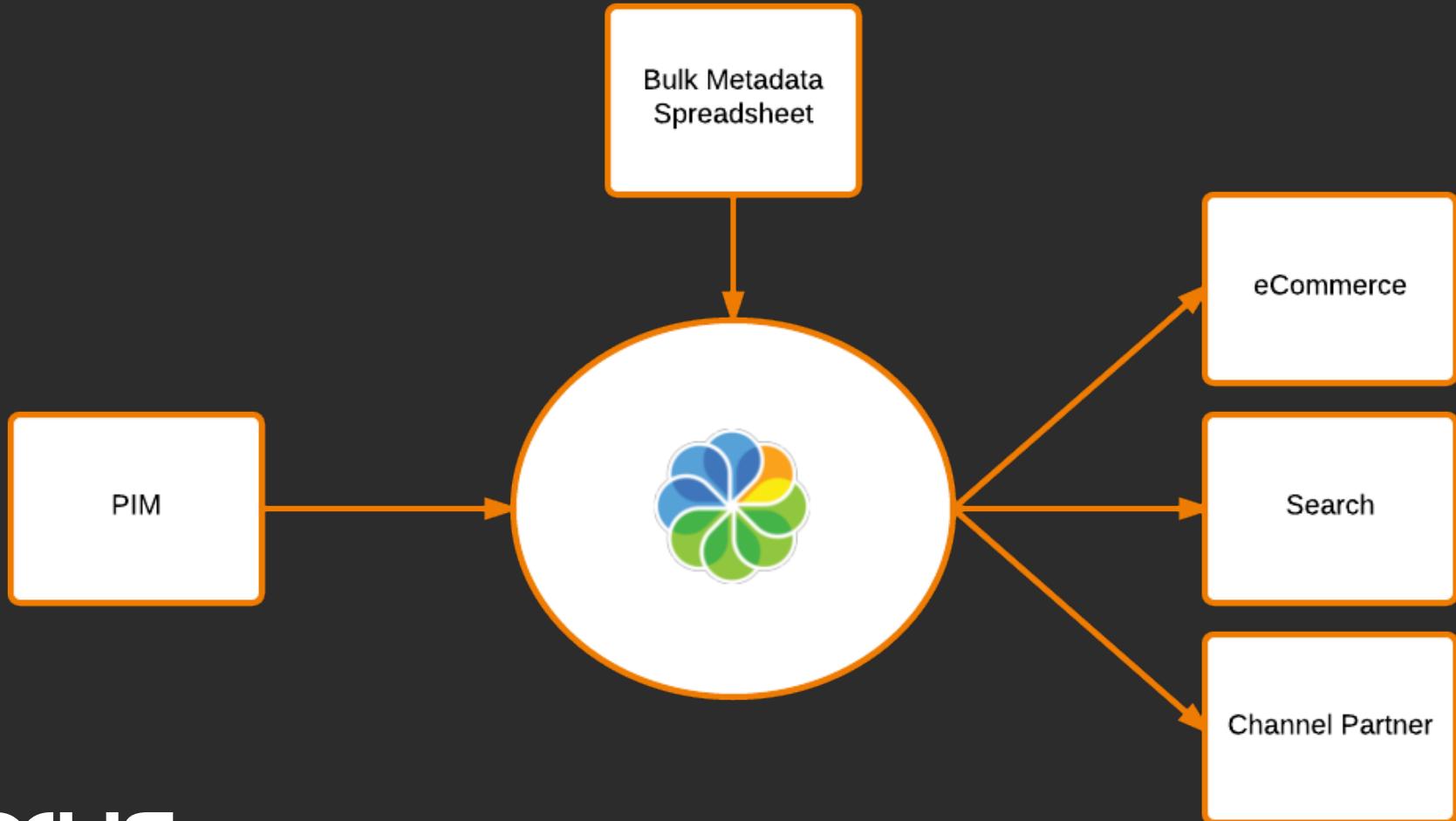


Derived metadata

E.g. File naming conventions

FSA0057-D103-20170421-ARB-minutes.docx





Complex metadata gathering approaches

1. Integration with a Product Information Management system
2. Bulk metadata import (not necessarily automated)
3. Machine Learning
4. Semantic extraction from text

A background pattern of overlapping hexagons in various shades of green and yellow, creating a sense of depth and texture.

Trends & Standards

HOW STANDARDS PROLIFERATE: (SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)

SITUATION:
THERE ARE
14 COMPETING
STANDARDS.

14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE'S
USE CASES.



YEAH!

SOON:

SITUATION:
THERE ARE
15 COMPETING
STANDARDS.



<https://imgs.xkcd.com/comics/standards.png>

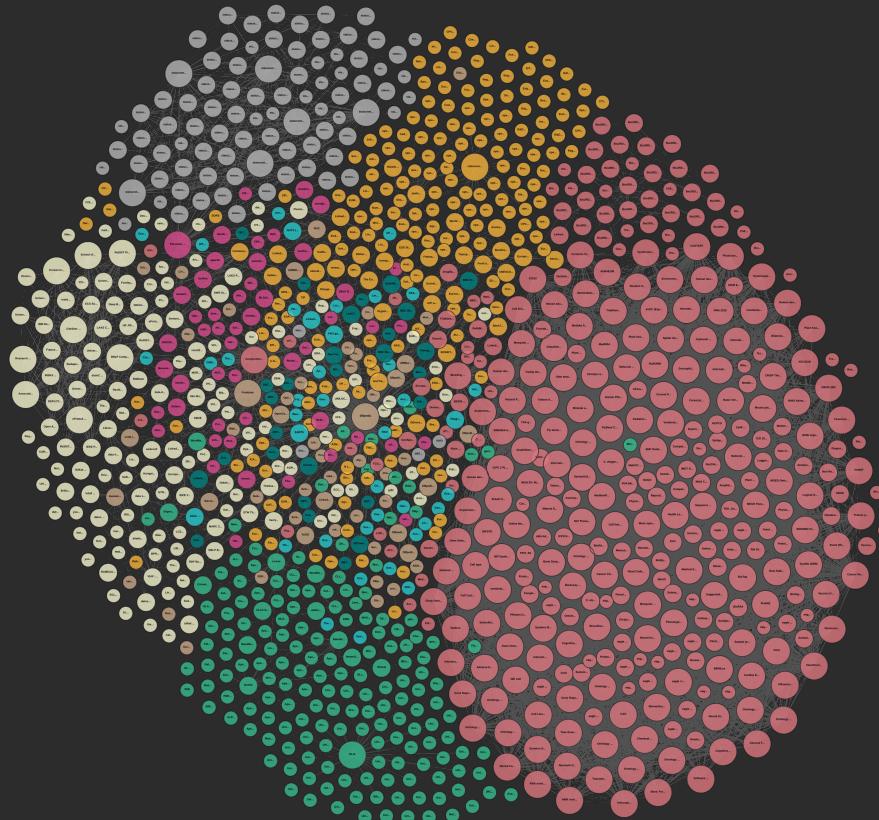
Resource Descriptor Framework

- RDF standard from the W3C
- Relies on *subject-predicate-object triples*
- John knows Jane
- DocumentA isA Contract
- Triples form a directed graph which can be navigated



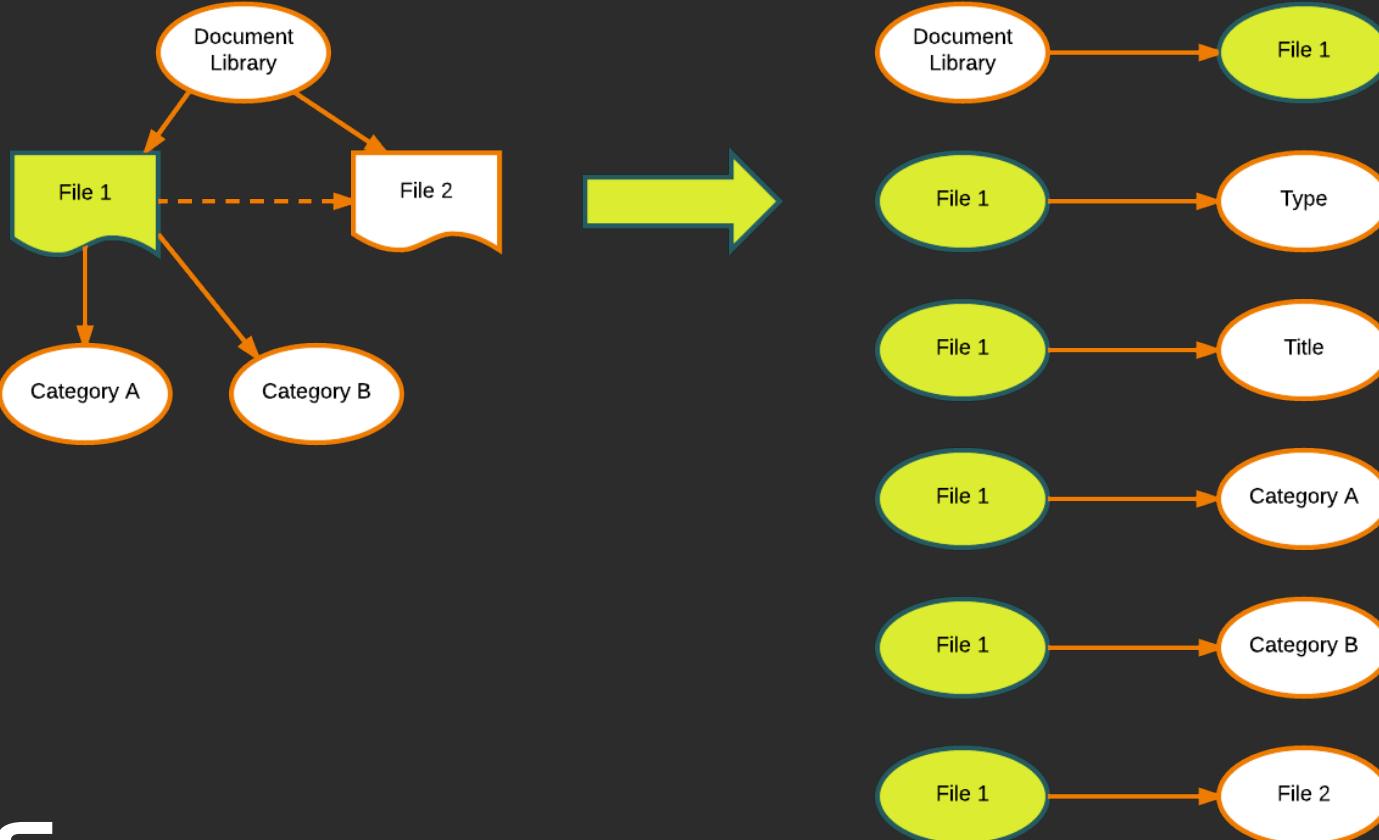
Linked Online Data

Legend
Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated
Incoming Links
Outgoing Links



Linking Open Data cloud diagram 2017, by Andrejs Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak. <http://lod-cloud.net/>

Outbound serialization



A background pattern of overlapping hexagons in various shades of green and yellow, creating a sense of depth and complexity.

Complex needs



Where is
the tipping



point for
Share?



Or the repo itself

Be aware of the limitations:

- Flexibility of the content model lagging behind current industry trends
- Internal model is proprietary
- Tricky to model some metadata standards
- Weak associative searches
- Many old file-based protocols don't support metadata capture

Don't forget Jeff Potts' Anti-patterns

See <http://ecmarchitect.com/archives/2014/06/26/3896>

- Anti-Pattern #3: Highly relational solutions
- Anti-Pattern #4: JSON/XML object store
- Anti-Pattern #5: Storing lots of content-less objects

A large, abstract graphic composed of several overlapping hexagons in various shades of light blue, light green, and yellow. A thick, dark green line forms a stylized bracket shape across the middle of the slide, enclosing the word "Summary".

Summary

Conclusion

Alfresco Content Services with Share provides a **great foundation** for capturing and storing metadata

- Content modelling capabilities, properties exposed through forms, workflow

Platform is **easy to integrate** for upstream/downstream systems

- Open principles (Open Source, open standards, open architecture)
- Java/Spring underpinnings
- Node crawler (from the Transfer Service)
- Transformation capabilities

Twitter: [@rbramley](https://twitter.com/rbramley)

Blog: <https://leanjavaengineering.wordpress.com>

GitHub: <https://github.com/rbramley>

Work: <https://www.ixxus.com>