French Public Information Library (BPI)
FEDERATED SEARCH IMPLEMENTATION WITH LIBRARYFIND

Terry Reese
Gray Family Chair for Innovative Library Services
Oregon State University

Roger Essoh
Head of Business Development
Head of Innovation
Atos Origin

San Francisco, SUN PASIG, October 2009
Agenda

1. Introduction to Atos Origin and French Public Information Library - BPI
2. Challenges to implement a Federated Search system
3. LibraryFind : technical architecture and features (by Terry Reese)
4. Implement Federated Search using SCRUM
5. What is coming next...?
6. Q/R
Overview of Atos Origin

» Atos Origin is a leading IT services company providing consulting, systems integration and outsourcing solutions – delivering business outcomes globally.

» Annual revenues of € 5.4 billion
» Over 50,000 employees
» In 40 countries worldwide
We deliver what we promised, no matter how critical the environment

<table>
<thead>
<tr>
<th>Retail &amp; Distribution</th>
<th>Healthcare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every day, Atos Origin processes the purchases of 3 million customers, e.g. 34 per second</td>
<td>Every year we carry out 1 million medical assessments and look after the occupational health of 500,000 employees</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Olympic Games</th>
<th>Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>We’re in charge of the IT of the Olympic Games. In Beijing 2008, more than 4,000 IT specialists managed the IT systems of the largest sports event in the world.</td>
<td>Atos Origin is in charge of the command and control systems of all French and many international nuclear power plants – a challenging environment.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Payment solutions</th>
<th>Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atos Origin manages more than 21 million credit cards, 36 million loyalty cards. We also process more than 100 million internet payments per year.</td>
<td>Atos Origin manages the compensation systems of the British Rail – or the German air traffic control systems.</td>
</tr>
</tbody>
</table>
The Olympic Games – an example of an intimate client partnership

Atos Origin is the official “Worldwide Information Technology Partner” of the International Olympic Committee (IOC):
- Salt Lake City 2002
- Athens 2004
- Torino 2006
- Beijing 2008
- Vancouver 2010
- London 2012
- Sochi 2014
- RIO 2016

The Olympic Games project is a showcase of a giant design-build-run project (> 4,000 collaborators), demonstrating Atos Origin’s ability to deliver on time and under any circumstances.

“Atos Origin continues to demonstrate its ability to seamlessly coordinate the efforts of a complex array of best-of-breed technology partners and suppliers. This unique systems integration capability enables consistent performance across multiple Olympic Games, while simultaneously facilitating the management of costs and complexity.”

Jean-Benoît Gauthier, Technology Director, IOC
Our work in Preservation and Archive

Preservation and Archive is more than a business opportunity for us! It’s a real commitment to provide expertise, support, resources to the community.

Provide resources & support

- We can provide team to work in your open source project

Federated Search

- We believe that Federated Search is a mature trend today with many interesting IT challenges (semantic, de-dupe, ...)

Talking with Fedora to work on the OSGi Fedora project

Preservation system

- BPI, BHL-Europe
- We believe in a modular preservation system to guarantee a sustainable solution
- We are looking to contribute in any open source modular system

Current projects: BPI, BHL-Europe and we’re in partnership with SUN Belgium for a Public Private Partnership (20 years) for a Belgian Gov project

Storage

- Why invest in a storage system when you can just use Cloud?
- What is the Cloud provider SLA for long term archive?

Current projects: partnership with London NHM (ICT project, Datacenter), BHL with Duracloud
French Public Information Library mission

PROVIDE PUBLIC INFORMATION
(Books, Web sites, newspapers,...)

ACTU
Le Mur de Berlin
À l'occasion du 20e anniversaire de la chute du Mur de Berlin, la Bpi vous propose d'explorer l'histoire de ce mur... Consultez notre dossier sur le Mur.
Agenda

1. Introduction to Atos Origin and French Public Information Library - BPI
2. Challenges to implement a Federated Search system
   3. LibraryFind: technical architecture and features (by Terry Reese)
   4. Implement Federated Search using SCRUM
   5. What is coming next...?
   6. Q/R
Federated Search Oriented services solution

- Content Management System
- Learning Management System
- Portal
- Partner application
- ...
Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Search Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Features</td>
<td>Connectors</td>
</tr>
<tr>
<td>Alerts Management</td>
<td>System Administration</td>
</tr>
<tr>
<td>Architecture</td>
<td>Integration</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Service Level Agreement - SLA</td>
</tr>
</tbody>
</table>

---

10 Custom presentation- or document control name
# Challenges with a Federated Search solution

## Usability

<table>
<thead>
<tr>
<th>Search Features</th>
<th>Search Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alerts Management</td>
<td>Connectors</td>
</tr>
<tr>
<td>Architecture</td>
<td>System Administration</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Integration</td>
</tr>
<tr>
<td>Service Level Agreement - SLA</td>
<td></td>
</tr>
</tbody>
</table>

## Search Features

- Friendly user interface
- Easy to use
- Easy basic search
- Multilingual
- Clustering search results

## Search Results

- Faceted search
- Taxonomies Ontology support
- Intuitive advanced search
- Informative result pages

## Alerts Management

- Online help system
- Community features
- Peers Search and review
- Auto complete search
## Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Search Features</th>
<th>Search Results</th>
<th>Alerts Management</th>
<th>Connectors</th>
<th>System Administration</th>
<th>Infrastructure</th>
<th>Integration</th>
<th>Service Level Agreement - SLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search</td>
<td>Features</td>
<td></td>
<td>Management</td>
<td>Connectors</td>
<td>Administration</td>
<td>Infrastructure</td>
<td>Integration</td>
<td>SLA</td>
</tr>
<tr>
<td>Search in specific sources</td>
<td>Advanced search capabilities</td>
<td>Full text search</td>
<td>Search</td>
<td>Save users search</td>
<td>Share search criteria</td>
<td>Search Intelligent agent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search in groups of sources</td>
<td>Search with Phrases</td>
<td>Search pages view by user profile</td>
<td>Save users search</td>
<td>Keep users search history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full text search</td>
<td></td>
<td></td>
<td>Keep users search history</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search pages view by user</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>View by user profile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced search</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>capabilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with Phrases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Save users search</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keep users search history</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share search criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search Intelligent agent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- **Search in specific sources:** Searches through specific sources.
- **Search in groups of sources:** Searches through groups of sources.
- **Full text search:** Full-text search capability.
- **Search pages view by user profile:** View search results by user profile.
- **Advanced search capabilities:** Advanced searching features.
- **Search with Phrases:** Search with phrases.
- **Save users search:** Save users’ search history.
- **Keep users search history:** Keep users’ search history.
- **Customize search interface:** Customize search interface.
- **Spelling Suggestions:** Spelling suggestions.
- **Share search criteria:** Share search criteria.
- **Search Intelligent agent:** Intelligent search agent.
Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Search Features</th>
<th>Search Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alerts Management</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Architecture</td>
<td>Connectors</td>
</tr>
<tr>
<td></td>
<td>Infrastructure</td>
<td>System Administration</td>
</tr>
<tr>
<td></td>
<td>Service Level Agreement - SLA</td>
<td>Integration</td>
</tr>
</tbody>
</table>

- Ranking result
- De-dupe results
- Single result page with different resources
- Incremental result
- Narrowing results
- Download, Print, save a result
- Email results list
- Export results in RSS
- Export citations
- Organize results by content type
Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Search Features</th>
<th>Search Results</th>
<th>Connectors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alerts</strong></td>
<td><strong>Management</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td>System Administration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Level Agreement - SLA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- User alert on new search results
- User alert personalisation
- New search results by email, RSS
Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Search Features</th>
<th>Search Results</th>
<th>Alerts Management</th>
<th>Connectors</th>
<th>Basic Connectors (Z39.50, SRU, SRW, OAI, …)</th>
<th>Sessions &amp; Cookies Management</th>
<th>Integration with ILS Catalogs</th>
<th>Connectors Monitoring Remotely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Features</td>
<td></td>
<td></td>
<td></td>
<td>Connectors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alerts Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Level Agreement - SLA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Multiple sources management</th>
<th>Simultaneous users queries management</th>
<th>One search access to multiple sources</th>
<th>Multi protocols access management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Features</td>
<td>Search Results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alerts Management</td>
<td>Connectors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td>System Administration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Integration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Level Agreement - SLA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Users management</th>
<th>Users queries management</th>
<th>Sources management</th>
<th>Search terms management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Features</td>
<td>Search Results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alerts Management</td>
<td>Connectors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Integration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Level Agreement - SLA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

System Administration
Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Cloud Approach</th>
<th>Insourcing Platform</th>
<th>Outsourcing Platform</th>
<th>Platform SLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Features</td>
<td>Search Results</td>
<td>Connectors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alerts Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td>System Administration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Integration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Level Agreement - SLA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>API</th>
<th>WS SOAP</th>
<th>WS JSON</th>
<th>OpenURL resolver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Features</td>
<td>Integration with RIA application</td>
<td>Integration with Portal</td>
<td>Integration with CMS</td>
<td>...</td>
</tr>
<tr>
<td>Search Results</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alerts Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connectors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Administration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Integration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Level Agreement - SLA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Challenges with a Federated Search solution

<table>
<thead>
<tr>
<th>Usability</th>
<th>Bug reporting</th>
<th>Bug resolution</th>
<th>Support call for users</th>
<th>Connectors support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Features</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search Results</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alerts Management</td>
<td>Training</td>
<td>Document-</td>
<td>Support call for users</td>
<td>Connectors support</td>
</tr>
<tr>
<td>Connectors</td>
<td></td>
<td>tation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td></td>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>System Administration</td>
<td></td>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Infrastructure</td>
<td></td>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Integration</td>
<td></td>
<td></td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Service Level Agreement - SLA

[Image of a table and diagram related to the challenges of a Federated Search solution]
Agenda

1. Introduction to Atos Origin and French Public Information Library - BPI
2. Challenges to implement a Federated Search system
3. LibraryFind : technical architecture and roadmap (by Terry Reese)
4. Implement Federated Search using SCRUM
5. What is coming next…?
6. Q/R
(some under-the hood)
LibraryFind Features

» “Look ahead” OpenURL linking (for conditionally linked materials)
» Dynamic, on-the-fly facet creation
  » Integrated facets utilizing local (solr) facets & dynamically generated facets
» Extensively user-tested user interface
» Event caching
» Search caching
» Ability to locally index harvestable collections
» Metadata-based connector repositories
LibraryFind Architecture

- Group / Collection Knowledge Base
- OpenURL Resolver
- Holdings Knowledgebase

Local Data Store

Harvestable Dbs

Direct-query Dbs

WS = Web Services
RoR = Ruby on Rails
» LibraryFind design built around the concept of platform
  » User interface provides metasearch/federated search for the OSU Libraries
  » API/Services layer provides a sandbox allowing the library and its data to be utilized by researchers on our campus
  » Connector framework with allows for varied us of the application
    - As a digital collections portal
    - As a service registry federated system
    - As a mediator between local/remote resources and other library services
LibraryFind – Behind the scenes

» LibraryFind is a unique metasearch tool for a couple of reasons:

- Integrated OpenURL resolution
- Both a harvester/indexer and federated search tool
- Metadata driven knowledge-base
- Implements a 2-tier caching system
LibraryFind – Behind the scenes

» Federated Search tool
  » Utilizes a metadata-based knowledge-base system
    - Allows for the use of abstract connection classes rather than hardcoded connectors
    - Makes adding new resources and sharing the knowledge-base with other institutions possible
LibraryFind – Behind the scenes

» Caching System Infrastructure

- Daily Cache (3 day cache)
- Research Cache (Represents top 15% of queries)
Technologies/Protocols

» Ruby/Rails
» Solr/Lucene
» Z39.50 / SRU / SRW
» OAI-PMH
» OpenSearch
» SOAP / WSDL/JSON
» COinS
Agenda

1. Introduction to Atos Origin and French Public Information Library - BPI
2. Challenges to implement a Federated Search system
3. LibraryFind : technical architecture and roadmap (by Terry Reese)
4. Implement Federated Search using SCRUM
5. What is coming next...?
6. Q/R
Implement Federated Search with SCRUM

Using SCRUM we completed a 10 months project in 7 months !!!
Implement Federated Search with SCRUM

Example of a SCRUM Board

Daily Stand up
Agenda

1. Introduction to Atos Origin and French Public Information Library - BPI
2. Challenges to implement a Federated Search system
3. LibraryFind : technical architecture and roadmap (by Terry Reese)
4. Implement Federated Search using SCRUM
5. What is coming next...?
6. Q/R
What is coming next…

We plan to launch a first version of Federated Search engine for cloud platforms in 2010.

Federated Search Engine

Communication and services layer

Connectors layer

Content Management System

Learning Management System

Portal

application

…

Cloud Platform

Internal sources

External sources
Thank You

Questions

Terry Reese
Gray Family Chair for Innovative Library Services
Oregon State University
Terry.Reese@oregonstate.edu

Roger Essoh
Head of Business Development
Head of Innovation
Atos Origin
Roger.Essoh@atosorigin.com
» Boost performance
» Reduce cost
» Increase agility
» Enhance CRM
» Shorten time to market
» Drive innovation
» Improve efficiency
» Increase adaptivity
» Enable business transparency
» Ensure regulatory compliance

For more information please contact:
Roger Essoh

t  +31 (0)1 70 92 49 54
m  +31 (0)6 27 15 81 06
roger.essoh@atosorigin.com

Atos Origin (France and Integration)
5-6 place des Iris
92926, Paris la Défense
www.atosorigin.com