

**ELAG 2010**

# Web Services Bootcamp

**Karen A. Coombs**  
[coombsk@oclc.org](mailto:coombsk@oclc.org)

**~ 10 Years library web stuff**

**Self taught coder who loves to play**

**Work with developers regularly**

**Product Manager for Developer Network**

**Name**

**Institution**

**Techie skills**

- **HTML, Javascript, DOM, XML, Server side scripting (PHP, ASP, Rails, Coldfusion...)**

**What you want to learn this week**

# What are Web Services?

**Machine readable data or procedures  
access via HTTP**

**On a remote/external system (usually)**

**Use structured data for data exchange  
(often XML)**

**Come in 3 flavors : RPC, SOAP, REST**

# Why Should You Care?



**Web Services are the backbone of mashups**

**Web Services enable you to access and recombine data**

**Move and share data across systems in real time**

**Easiest web service style**

**Readable by Humans**

**Most popular form of web services today**

**Create, Read, Update Delete**

**POST, GET, PUT, DELETE**

## Javascript Object Notation

Text file used to pass information from one system to another

Javascript treats it as an object

- which makes it easily navigated and specific pieces cherry-picked



## XPath

Little bit like the DOM but more complex and smart

Allows you to cherry pick info in XML documents

Learn more about XPath -

<http://www.w3schools.com/xpath>

**Used to navigate XML and select information**

**Context is key - where are you in the tree?**

**Many mashup errors are XPath errors**

- **Many of these are namespace issues**

**Beware //somenode**

## **XML is a tree**

- **children**
- **descendants**
- **ancestors**
- **siblings**

**Namespace matters**

**Method to avoid name conflicts between elements**

**Used often when creating XML which uses different schemas**

**Namespace declaration(s)**

- **typically goes into root node**
- **always at least one if there is a schema being used**
- **doesn't have to be in root. can be defined in element where it is used**

**Way of adding prefix to elements to make it easier to distinguish which namespace they are part of.**

**Used when there are multiple namespaces declared in the root**

- Can only have one default namespace in root**

**Alternatively, can declare default namespaces in sub-elements to avoid using prefix**

**nodename** - selects the children of the named node

**/** - root node

**//nodename** - any of the named nodes within the document regardless of where they are in the structure

**.** - select current node

**..** - selects parent node

**@** - selects an attribute

**|** between statements is an OR

Allow you to select nodes that match certain conditions

Condition you want to match goes in brackets

- Value of an element or attribute =, <, >
- Nodes in particular positions
- Pattern matches
  - contains or starts-with
  - regular expressions
- position()

**Get all the possible forms of the name - Identities**

- `//match/establishedForm`

**Get all the entry links where the rel=alternate - Atom feed**

- `//entry/link[@rel='alternate']`

**Get all the 245 |a - WorldCat Search API**

- `//marc:record/marc:datafield[@tag='245']/marc:subfield[@code='a']`

**WorldCat Identities**

**Atom Feed - WorldCat Search API**

**Atom Feed - YouTube**

**Flickr XML**

**MARCXML**

# **Navigating XML With the DOM**



**Doesn't work with XML with namespaces**

**Simpler to understand than XPath**

**Works well for less complicated XML structures**

# Example Using the DOM



**\$xml->rsp->oclcnumber['url']**

**\$xml->Identity->authorityInfo->wikiLink**

**\$xml->Identity->authorityInfo->wikiLink**

**\$xml->feed->title**

**AJAX doesn't allow you to access “scripts” across servers**

- **XML, PHP or other scripting languages**

**Can access JSON across servers**

- **Reason why JSON popular**

**Workaround**

**Proxy the script via the webserver**

- **Apache mod\_proxy**

**Use an intermediary to transform content into JSON**

- **Use PHP to convert XML to JSON**

**Javascript library**

**Eliminates cross-browser compatibility issues**

**Capable of retrieving and manipulating JSON**

**Lots of nice effects**

# Basic JSON Manipulation With JQuery



**Get the JSON**

**Multiple records returned?**

- **Loop through them**

**Access specific fields you want to use**

# JQuery JSON Example

# Possible Types Of Data



**Bibliographic Data**

**Book Covers**

**Digital Library metadata and objects**

**Reviews**

**Ratings**

**Library Information**

**Map Data**

**Feeds from library catalog, Amazon or other sources**

**Covers, Ratings and Reviews from LibraryThing**

**Preview or fulltext from GoogleBooks**

**Info from dbpedia**

**Data from Open Library**

**Serial Solutions -**

**<http://journal.code4lib.org/articles/108>**

**SFX**

**Gold Rush Link Resolver -**

**[http://grx.coalliance.org/grxtest/dsp/grx\\_documentation.pdf](http://grx.coalliance.org/grxtest/dsp/grx_documentation.pdf)**

**Lots of library database platforms**

**Several library catalogs**

**Web Service, sort of**

**Meant to perform asynchronous interactions for  
harvesting data**

**Outputs XML**

**Can be queried**

**WorldCat Search API**

**WorldCat Identities**

**xID services**

- **xISBN**
- **xISSN**
- **xOCLCNum**

**Registry**

**QuestionPoint KB API**

**Search WorldCat in a programmatic fashion**

- **OpenSearch, SRU**

**Get data back as XML**

- **Atom, RSS, MARCXML, Dublin Core**

## Service for identifiers

- OCLC number, ISSN, ISBN

## Send an identifier

## Get related identifiers back

## Get information about item

- Peer Reviewed or not
- RSS for Table of Contents

## Data about libraries throughout the world

- Name
- Address
- Catalog url
- Type (public, academic, special ...)

**OpenURL based**

**Send an OpenURL**

**Get back holdings information**

**Uses**

- **Crosslisting ebooks**
- **Direct Links to Fulltext in LibraryFind**

# Serial Solutions API Response



**Citation**

**Holdings Info**

- **Provider - Proquest**
- **Database Name - Safari Tech Books**
- **Start & End Dates**

**URL(s)**

## Different types of URLs

Depends on OpenURL sent

- format
- granularity

Books - source, book

Journal - source, article, journal, volume, issue

# Crosslisting Content In the Catalog



## Serial Solutions OpenURL Resolver

- API

## WorldCat Search API

## ISBN/ISSN

Adds also available as link to ebook print book

# Crosslisting Ebooks



UNIVERSITY of HOUSTON

LIBRARY CATALOG

[New Search](#) [My Account](#) [Ask A Librarian](#) [Help](#)

[M.D. Anderson](#) [Architecture & Art](#) [Music](#) [Optometry](#) [Pharmacy](#)

[Start Over](#) [Request](#) [Save Records](#) [MARC Display](#) [Return to List](#) [Limit/Sort Search](#) [Another Search](#) [WebBridge](#) (Search History)

ISN: 0321461061 UH Main & Other Resources  
System Sorted Search  
 Available in library

Result page: [Previous](#) [Next](#)

**Author** [Bardzell, Jeffrey.](#)  
**Title** **Adobe Dreamweaver CS3 with ASP, ColdFusion, and PHP / Jeffrey Bardzell, Bob Flynn.**  
**Publisher** Berkeley, Calif. : Peachpit ; London : Pearson Education [distributor], 2007.  
**How to Cite** [APA, Chicago, Harvard, MLA, & Turabian](#)

[Persistent link to this title](#)

### Electronic Access:

[Table of contents](#)

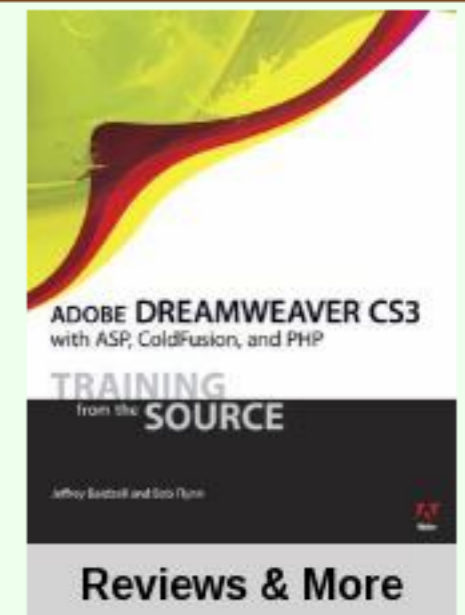
[Brief Record](#) [Details](#) [Find Similar Items](#) [Full Record](#)

[Print This Page](#)



LOCATION	CALL #	STATUS
<a href="#">ANDERSON/GENERAL COLL</a>	<a href="#">TK5105.8885.D74 B358 2007</a>	IN LIBRARY
<a href="#">UHCL/GENERAL STACKS</a>	<a href="#">TK5105.8885.D74 B369 2007</a>	DUE 02-20-10

### Additional Info:



Reviews & More

Item Rating ★★★★★

[Also Available as an Ebook](#)

# Direct Links to Fulltext in LibraryFind



[Article] [Characteristics of Educational and Learning Environments in Organization and](#)

The challenges of the emerging knowledge economy, globalization and competitiveness call for new approaches at the levels of organizations and society as a whole. The traditional conceptual approach to innovation is being challenged.

JAN, 2009 *Social Sciences (1392-0758)*

[Full Abstract](#) [Citation](#) [Save This Item](#) [more from 'SocINDEX with Full Text'](#)

**Blip.tv API - [http://wiki.blip.tv/index.php/Blip.tv\\_API](http://wiki.blip.tv/index.php/Blip.tv_API)**

**Flickr API - <http://www.flickr.com/services/api/>**

**YouTube API -**

**<http://code.google.com/apis/youtube/overview.html>**

**Google Maps -**

<http://code.google.com/apis/maps/>

**Yahoo Maps API -**

<http://developer.yahoo.com/maps/>

**New York Times -**

**<http://developer.nytimes.com/docs>**

**Guardian - <http://www.guardian.co.uk/open-platform>**

Last.fm - <http://www.last.fm/api>

dbpedia - <http://wiki.dbpedia.org/OnlineAccess>

Google Book Search -

<http://code.google.com/apis/books/>

Billboard API -

<http://developer.billboard.com/docs>

**Google Calendar API -**

**[http://code.google.com/apis/calendar/data/2.0/developers\\_guide.html](http://code.google.com/apis/calendar/data/2.0/developers_guide.html)**

**Delicious API - <http://delicious.com/help/api>**

**and <http://delicious.com/help/feeds>**

**Expression of the Semantic Web**

**Way to expose, share and connect data via  
dereferenceable URIs**

**Able to be explored by both people and  
machines**

**Share data in a non-library centric format**

**Share semantically rich data in a web-friendly way**

**Create a scenario where data is meaningful in new contexts**

**Look at data across different silos and make connections**

**URIs**

**Real world objects**

**Content Negotiation**

**RDF**

**Act like control numbers**

**Have several other advantages**

- **Domain name tells you who created the URI**
- **URI is typically a URL, which means that there is often something to go to**

**Type of URI that represents a “Concept”  
not Web page**

**303 redirect is a returned and tells you  
that this is a “concept” not just a page**

**Deals with the idea that different application will want to interact with different representations of a real world object**

- **human readable - HTML**
- **machine processable - RDF**

**Sent via the Accept Header**

- **\*.\*, text/xml, application/rdf+xml**

**Framework for describing things**

**Different Serializations: RDF/XML, Turtle, RDFa,  
N3**

**RDF statements have three parts:**

- **subject, predicate, object**

**Subject and predicate must be URIs**

**Object can be a value or URI**

```
<contact:Person  
  rdf:about="http://www.w3.org/People/EM/contact#  
me"> <contact:fullName>Eric  
  Miller</contact:fullName>  
</contact:Person>
```

Subject - page that is the URI

Predicate - node contact:fullName

Object - value of node contact:fullName Eric Miller

**Dublin Core**

**FOAF**

**SKOS**

**Library Linked Data Incubator Group**

- **working on what library data should look like in RDF**

**Search within Linked Data**

**Query language for RDF**

**Search webs of RDF data**

**Potentially query disparate data sources**

**Relatively few SPARQL “end-points”**

**dbpedia - Wikipedia as Linked Data**

**VIAF - authority files from several national libraries**

**Dewey.info - top three levels of Dewey classification in several languages**

**id.loc.gov - Library of Congress subject headings, etc**

## **Library of Congress Subject Headings**

- **Genre/Form headings**
- **Children's headings**

**MARC Relator Codes**

**Thesaurus of Graphic Materials**

## Feeds

- The poor man's web services
- XML
- Can have schemas added to them

## Feed2JS

- embed a feed in a web page

Drupal and Wordpress can have feeds  
embedded into pages

## Databases for the MBA Program

*You must be logged into ANGEL to access these resources.*

For more information on searching the databases, visit [Library Research in Business and Management: An Introduction to the Key Databases](#).

Decision-making in a business environment is best supported by facts, and analytical, thought-through arguments. There are several specialized websites and business databases that provide access to this kind of information. They are gathered below under the following headings:

- [Finding Factual Information about Companies](#)
- [Finding Articles on Business Topics](#)
- [Finding Factual Information about Countries](#)

### Finding Factual Information about Companies

Databases of business facts are typically used for investment analysis; financial, estate and insurance planning; contract bidding, and

### Library News

For distance students and faculty at Norwich University

#### Library Closed July 3-5

The Kreitzberg Library will be closed from Friday July 3rd through Sunday July 5th for the Independence Day holiday. We will not be available for instant message chat via Meebo on Friday July 3rd, and all reference emails and interlibrary loan requests made over the holiday weekend will not be reviewed until Monday July 6th.

#### Library Survey closes 3/31

Please help the Library by participating in a very important electronic survey designed to measure your

**Sometimes easier to use than the “full” API**

- **Flickr**
- **Delicious**

**Ubiquitous format**

- **Google APIs**

**Lots of systems provide them**

**Drag and drop UI for creating mashups**

**Free**

**Lots of examples**

**Some library examples**

- **RSS of Books + Covers**
- **Finding Alternative Copies of a Book**

**Programming for non-programmers**

**Good user community**

**Clone!!**

**End up with a variety of formats**

- **Widgets**
- **RSS**
- **JSON**
- **PHP**

## RSS + Cover Data + Yahoo Pipes

### ISBN essential

### How does it work?

- RSS feed from WorldCat Search API
- Extracts ISBN
- Gets cover data
- Builds new feed with covers

# **WorldCat Registry + Yahoo Maps**



**Send registry request for data for a particular location (zip code, city)**

**Retrieve results**

**Send address information to Yahoo Maps**

**Get library locations plotted on map**

# **WorldCat Holdings + Yahoo Maps**



**Send an OCLC number and location to  
WorldCat Search API**

**Returns a list of libraries nearest that  
location with holdings (includes address)**

**Send libraries' address to Yahoo Maps**

**See which libraries in that location have that  
book visualized**

**Google Calendar Embedding**

**Creating a Flickr Gallery**

**Embedding delicious links**

**Embedding YouTube videos**

**WorldCat Search + Book Covers + Ratings +  
Preview at Google Books**

**WorldCat Search Holdings + Google Maps**

**Journal List Enhancements**

**New York Times Reviews**

**Similar Artists using Last.fm**

**Lots of different ways in which people can access information**

**Different devices**

- **Mobile devices, computers, special purpose devices**

**Machine access to data**

**WorldCat Search API**

**Open Library/LibraryThing Covers**

**Summary - ISBNdb.com**

**Full text/Preview**

- **Hathi Trust**
- **Google Books**

**Reviews from New York Times**

**Screenscrape Item data**