



# Taking Control of Library Metadata and Websites using the eXtensible Catalog

Jennifer Bowen

University of Rochester/eXtensible Catalog Organization

Code4lib 2010, Asheville, North Carolina

Feb.23, 2010



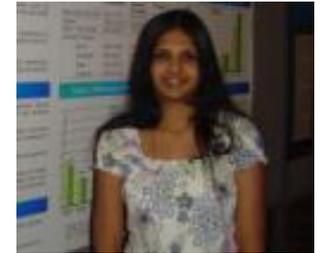
## **Péter Király (pkiraly)**

Drupal Toolkit, OAI Toolkit



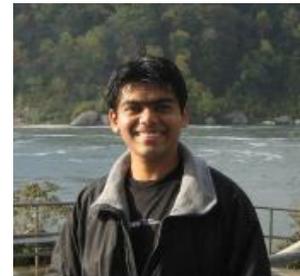
## **Sharmila Ranganathan (sranganathan)**

Metadata Services Toolkit



## **Shreyansh Vakil (svakil)**

OAI Toolkit, NCIP Toolkit



## **Randy Cook (rcook)**

Project Manager





- **eXtensible Catalog** is open source, user-centered, next generation software for libraries.
- Four software components
  - can be used independently to address a particular need
  - or
  - combined to provide an end-to-end discovery system to connect library users with resources.



- User Interface
  - Faceted, FRBRized, customizable search interface
  - Web application framework for libraries
- Metadata Tools
  - Automated processing of large batches of metadata
- Connectivity Tools
  - Between XC and an ILS
  - Harvest and synchronize metadata with OAI-PMH
  - Circulation and authentication with NCIP



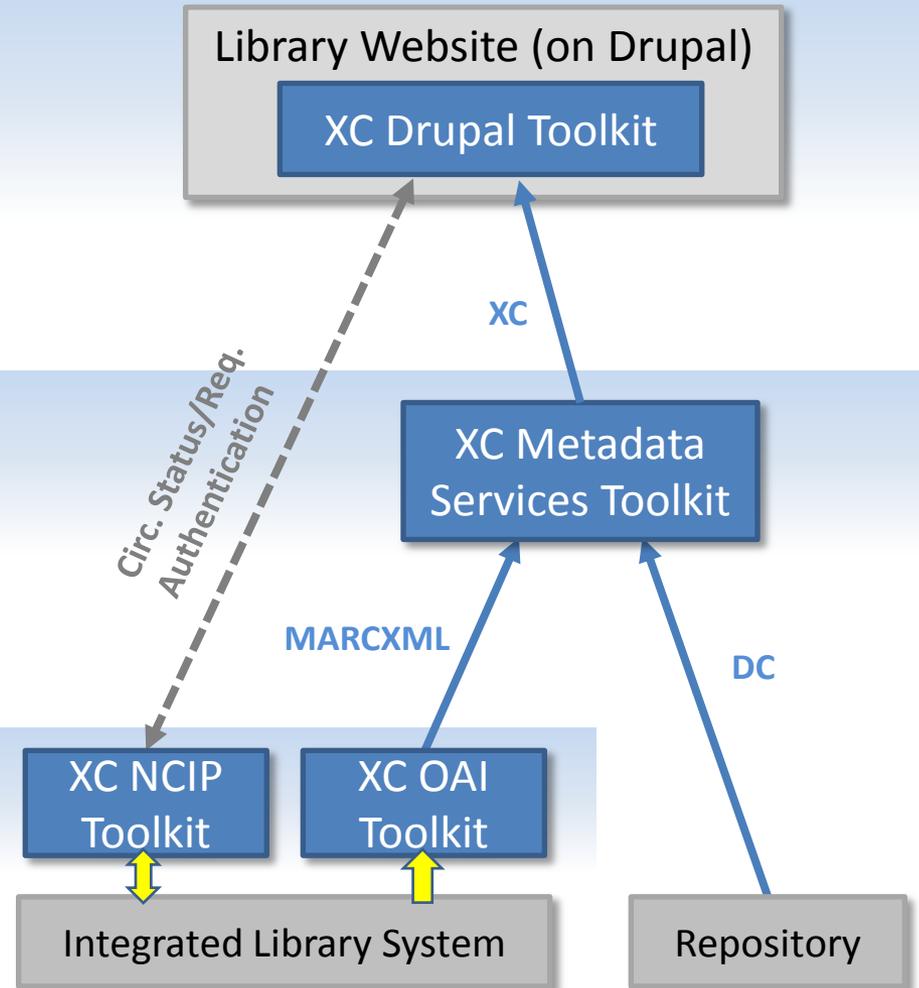
# XC Software Components

Take Control.

User Interface for searching and browsing

Tools for automated processing of large batches of metadata

Tools for connectivity between XC and an ILS





User Interface for searching and browsing

Library Website (on Drupal)

XC Drupal Toolkit

Tools for automated processing of large batches of metadata

XC Metadata Services Toolkit

Tools for connectivity between XC and an ILS

XC NCIP Toolkit

XC OAI Toolkit

Integrated Library System

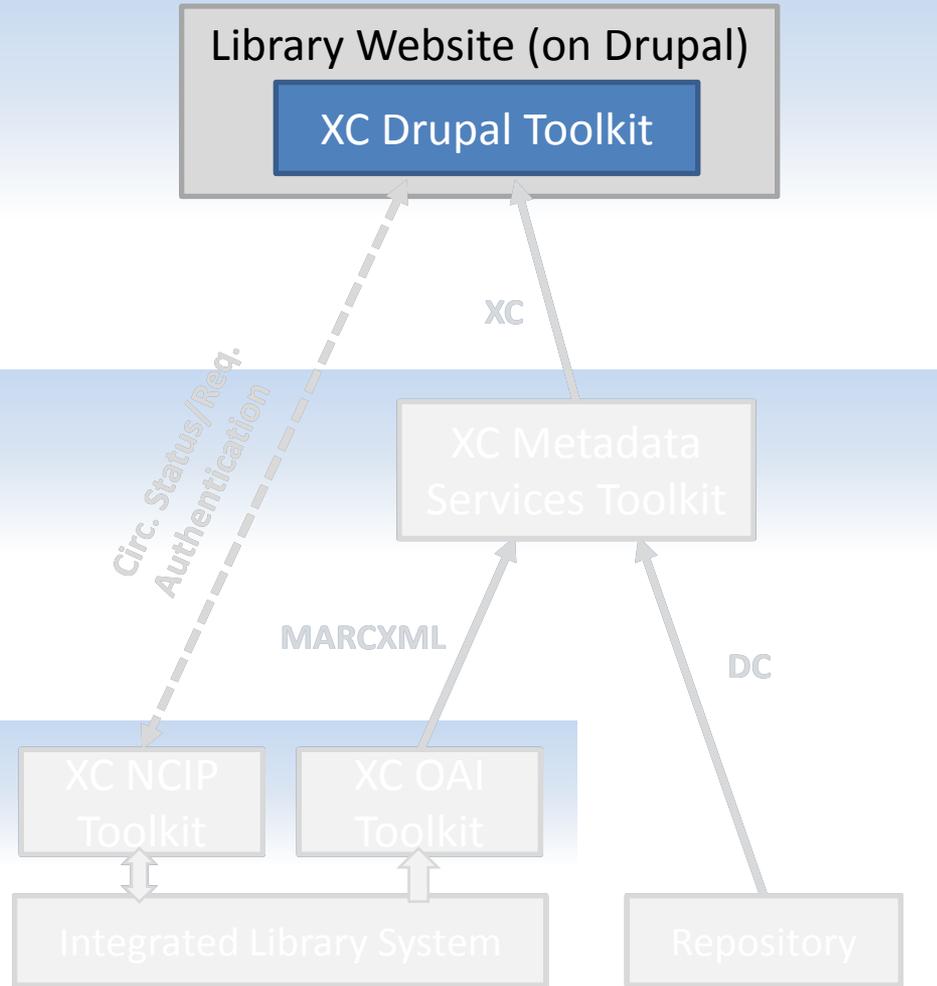
Repository

*Circ. Status/Req. Authentication*

MARCXML

DC

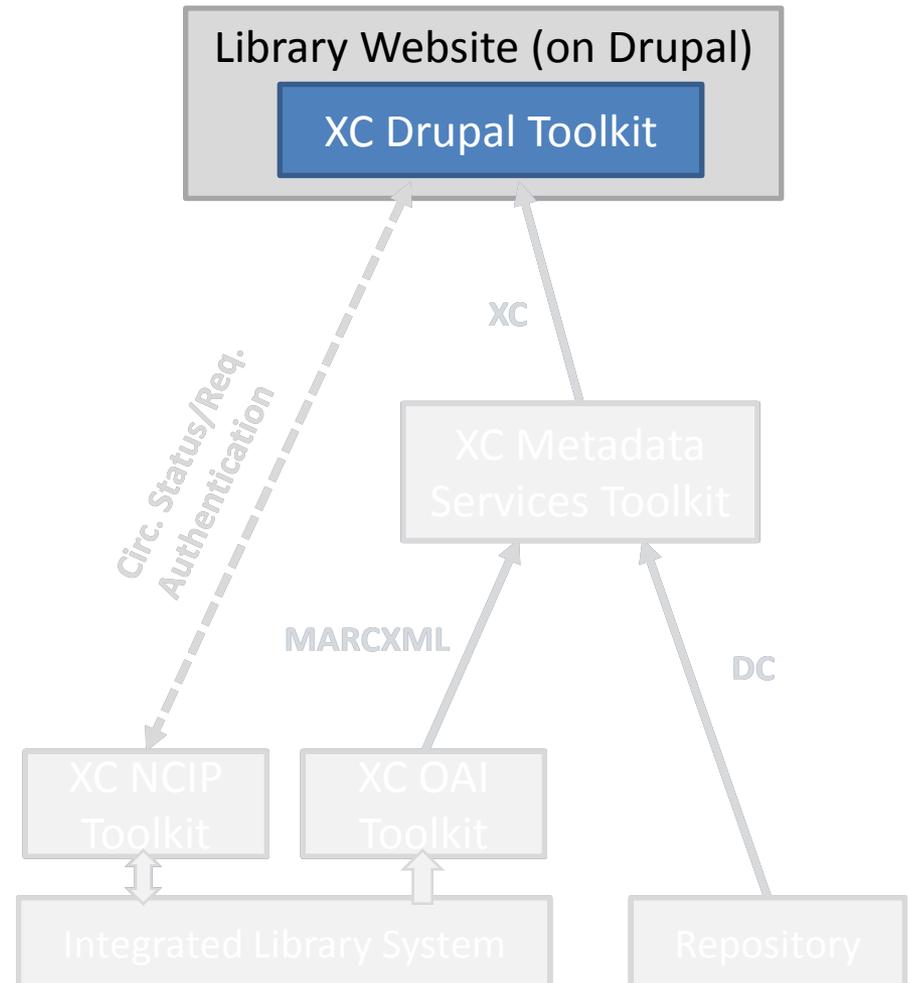
XC





## Features

- Adds support for library metadata to Drupal
- Single search interface across:
  - Library catalog
  - Digital repository
  - Website resources
- Faceted browsing of results
- FRBRized metadata
- Extensive customization
- Integration with ILS circulation system





# Drupal Toolkit: Search Interface

Take Control.

The screenshot shows the Drupal search interface on a localhost. The search bar contains 'xml' and the results show 42 records found. The search results are displayed in a list format with details for each item, including title, author, publication information, and availability. The interface includes a search bar, a search button, and a list of search results.

Search this site:  Search

Searched for: The query is: xml

Authors/Contributors:

- ACM Sigsoft. (1)
- Agouris, P. (Peggy), 1963- (1)
- Akoka, Jacky, 1959- (1)
- Ali, Moonis. (1)
- Althoff, Klaus-Dieter. (1)

other contributors:

- ACM Sigsoft. (1)
- Agouris, P. (Peggy), 1963- (1)
- Akoka, Jacky, 1959- (1)
- Ali, Moonis. (1)
- Althoff, Klaus-Dieter. (1)

Formats:

- Text (41)
- 24 cm. (27)
- 23 cm. (5)
- 28 cm. (4)
- 25 cm. (2)

Search results:

- XML and Java : developing Web applications / by Maruyama, Hiroshi, 1958-**  
published: Addison Wesley, c1999. ix, 386 p. ;  
format: Text  
Availability: Rush Rhees Library: *Not Available*
- Data on the web : from relations to semistructured data and XML / by Abiteboul, S. (Serge)**  
published: Morgan Kaufmann, c2000. xiii, 257 p. :  
format: Text  
Availability: Rush Rhees Library: *Possibly Available*
- http://webtext.library.yale.edu/xml/2html/music.ives-sinclair.nav.html Online version. ...**  
Availability: Rush Rhees Library: *Possibly Available*
- The LaTeX Web companion : integrating TeX, HTML, and XML / by Goossens, Michel.**  
published: Addison-Wesley, c1999. xxii, 522 p. :  
format: Text  
Availability: Rush Rhees Library: *Possibly Available*

**Features:**

- Simple keyword search
- Facets
- Customizable metadata display
- Circulation status
- Syndetic Solutions content
- Why am I getting these results for my query?



# Drupal Toolkit: Staff Interface

Take Control.

localhost

Solr delete delete all links log modules node schedule

Home > eXtensible Catalog (XC) > XC Search > Displaying > Display template element

Display template element settings. View Edit

**element name:**  
creator  
The name of the display type element.

**metadata schema:**  
XC  
The name of the metadata schema.

**entity type:**  
manifestation  
The entity type (record type).

**element template:**

```
if(isset({rdarole:artist})) {  
    return {rdarole:artist};  
} else if(isset({rdarole:author})) {  
    return {rdarole:author};  
} else if(isset({rdarole:compiler})) {  
    return {rdarole:compiler};  
} else if(isset({rdarole:composer})) {  
    return {rdarole:composer};  
} else if(isset({rdarole:speaker})) {  
    return {rdarole:speaker};  
} else if(isset({xc:creator})) {  
    return {xc:creator};  
} else if(isset({dcterms:creator})) {  
    return {dcterms:creator};  
}  
}
```

The display template.

Save Delete

Form to customize element display:

Determines how to display the creator element on the results screen.

The “element template” box allows libraries to write simple PHP code and offers more flexibility.

All options are pre-filled with “good” defaults.



- Create custom web applications to browse catalog content by filling out web forms
- Pre-set search limits, customized facets
- Possibilities:
  - Search/Browse eJournals
  - Search/Browse databases
  - Search/Browse collections
- Example: browsable list of videos and DVDs created without any programming



# Drupal Toolkit: Custom Search/Browse Pages

Take Control.

localhost

Solr delete delete all links log modules node schedule

Home

Find DVDs and Videos

Search  Browse by Language  Browse by Director  Browse by Genre

Keyword(s):: Albanian (2)

Note: in 24 document this field does not exist

type:

All  VHS ONLY  DVD ONLY

Submit

This example allows users to browse a subset of the catalog: the videos and DVDs collection

On this page, users can browse by language.

Other tabs support browse by director and genre.

Limits can be set for format: VHS or DVD



# Drupal Toolkit: Provided Modules

The Drupal Toolkit is actually a collection of Drupal add-on modules:

- OAI-PMH Harvester
- Metadata Import and export
- MySQL and Solr storage
- NCIP Integration
- Authentication
- EZProxy
- Syndetic Solutions
- Search

Enabled	Name	Version	Description
<input checked="" type="checkbox"/>	<b>EZProxy URL Rewrite</b>	6.x-1.1-alpha3	Rewrite URL links based on a user's IP address and other advanced options; allowing IP-protected content to be forwarded to an EZProxy or another proxy referral server
<input checked="" type="checkbox"/>	<b>ILS</b>	6.x-1.1-alpha3	ILS settings Depends on: XC NCIP Toolkit Integration ( <b>enabled</b> ), XC Core ( <b>enabled</b> ), NCIP Integration ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>Metadata Import and Export</b>	6.x-1.1-alpha3	Import and export metadata to and from the XC system Depends on: XC Metadata ( <b>enabled</b> ), XC Core ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>NCIP Integration</b>	6.x-1.1-alpha3	Client and library for NISO Circulation Interchange Protocol (NCIP/Z39.83) Required by: XC NCIP Toolkit Integration ( <b>enabled</b> ), ILS ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>OAI Harvester</b>	6.x-1.1-alpha3	OAI-PMH Harvester Depends on: XC Utility ( <b>enabled</b> ) Required by: XC OAI Harvester Bridge ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>OAI-DC Schema</b>	6.x-1.1-alpha3	Handles import and export of Dublin Core (oai_dc) schema records Depends on: XC Core ( <b>enabled</b> ), XC Metadata ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>Syndetics Solution</b>	6.x-1.1-alpha3	Syndetics Solution services Depends on: XC Core ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC Authentication</b>	6.x-1.1-alpha3	Delegates Drupal user account login and registration for XC to external authentication system
<input checked="" type="checkbox"/>	<b>XC Browse</b>	6.x-1.1-alpha3	Browse records
<input checked="" type="checkbox"/>	<b>XC Core</b>	6.x-1.1-alpha3	Integrates the eXtensible Catalog (XC) into a Drupal environment Required by: Syndetics Solution ( <b>enabled</b> ), XC eXternal services framework API ( <b>enabled</b> ), XC Metadata ( <b>enabled</b> ), XC NCIP Toolkit Integration ( <b>enabled</b> ), OAI-DC Schema ( <b>enabled</b> ), XC OAI Harvester Bridge ( <b>enabled</b> ), XC Schema ( <b>enabled</b> ), XC Search ( <b>enabled</b> ), XC Solr ( <b>enabled</b> ), XC SQL Database Storage ( <b>enabled</b> ), ILS ( <b>enabled</b> ), Metadata Import and Export ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC eXternal services framework API</b>	6.x-1.1-alpha3	mashupable external services framework API Depends on: XC Core ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC Metadata</b>	6.x-1.1-alpha3	Adds support for creating, managing, and manipulating metadata structures and their containers and facilitates the import, export, storage, retrieval, and relational organization of metadata Depends on: XC Core ( <b>enabled</b> ) Required by: Metadata Import and Export ( <b>enabled</b> ), OAI-DC Schema ( <b>enabled</b> ), XC OAI Harvester Bridge ( <b>enabled</b> ), XC Schema ( <b>enabled</b> ), XC Solr ( <b>enabled</b> ), XC SQL Database Storage ( <b>enabled</b> ), XC Search ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC NCIP Toolkit Integration</b>	6.x-1.1-alpha3	Connects to NCIP Toolkit instances Depends on: XC Core ( <b>enabled</b> ), NCIP Integration ( <b>enabled</b> ) Required by: ILS ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC OAI Harvester Bridge</b>	6.x-1.1-alpha3	Handles the import of metadata into XC via OAI-PMH. This is a bridge between OAI Harvester and XC modules Depends on: XC Core ( <b>enabled</b> ), XC Metadata ( <b>enabled</b> ), OAI Harvester ( <b>enabled</b> ), XC Utility ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC Schema</b>	6.x-1.1-alpha3	Handles import and export of eXtensible Catalog (xc) schema records Depends on: XC Core ( <b>enabled</b> ), XC Metadata ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC Search</b>	6.x-1.1-alpha3	Searching records in the XC system Depends on: XC Core ( <b>enabled</b> ), XC Solr ( <b>enabled</b> ), XC Metadata ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC Solr</b>	6.x-1.1-alpha3	Handles storage, retrieval, and search of metadata to, from, and on multiple Apache Solr servers Depends on: XC Core ( <b>enabled</b> ), XC Metadata ( <b>enabled</b> ) Required by: XC Search ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC SQL Database Storage</b>	6.x-1.1-alpha3	Handles storage and retrieval of metadata using sets of SQL database tables Depends on: XC Core ( <b>enabled</b> ), XC Metadata ( <b>enabled</b> )
<input checked="" type="checkbox"/>	<b>XC Utility</b>	6.x-1.1-alpha3	Utility functions for the rest of XC modules Required by: OAI Harvester ( <b>enabled</b> ), XC OAI Harvester Bridge ( <b>enabled</b> )



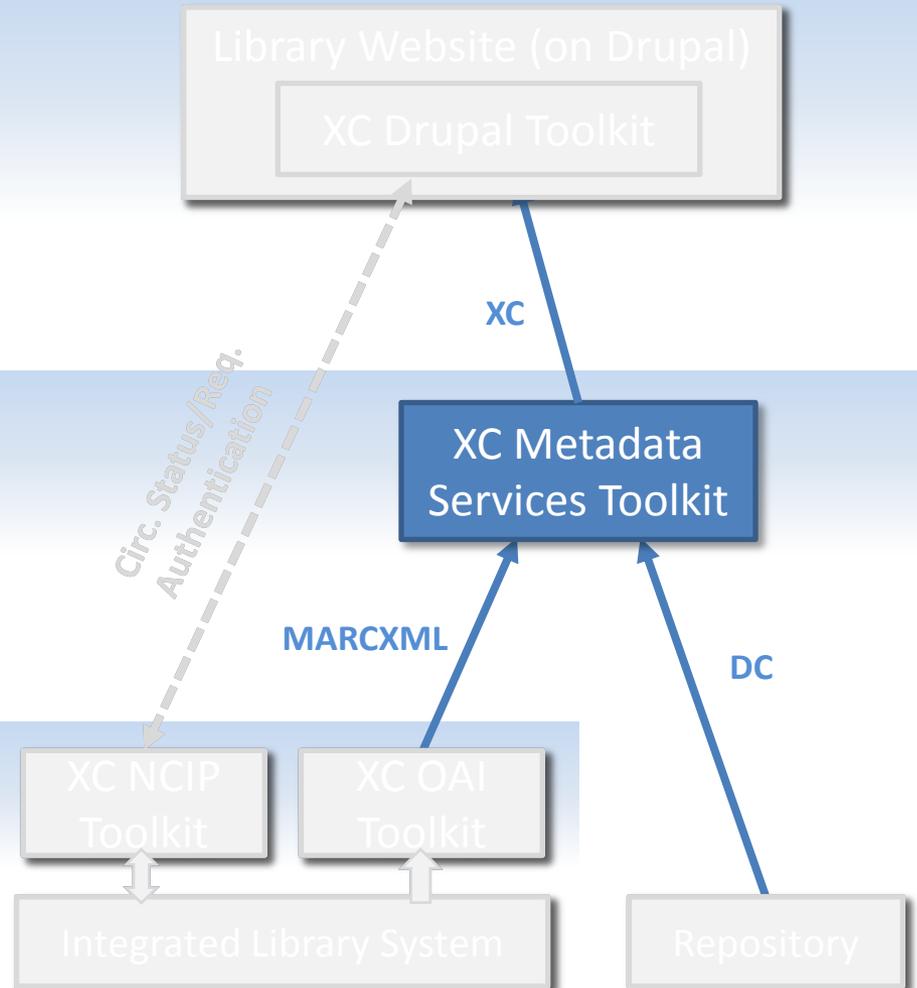
# Metadata Services Toolkit

Take Control.

User Interface for searching and browsing

Tools for automated processing of large batches of metadata

Tools for connectivity between XC and an ILS





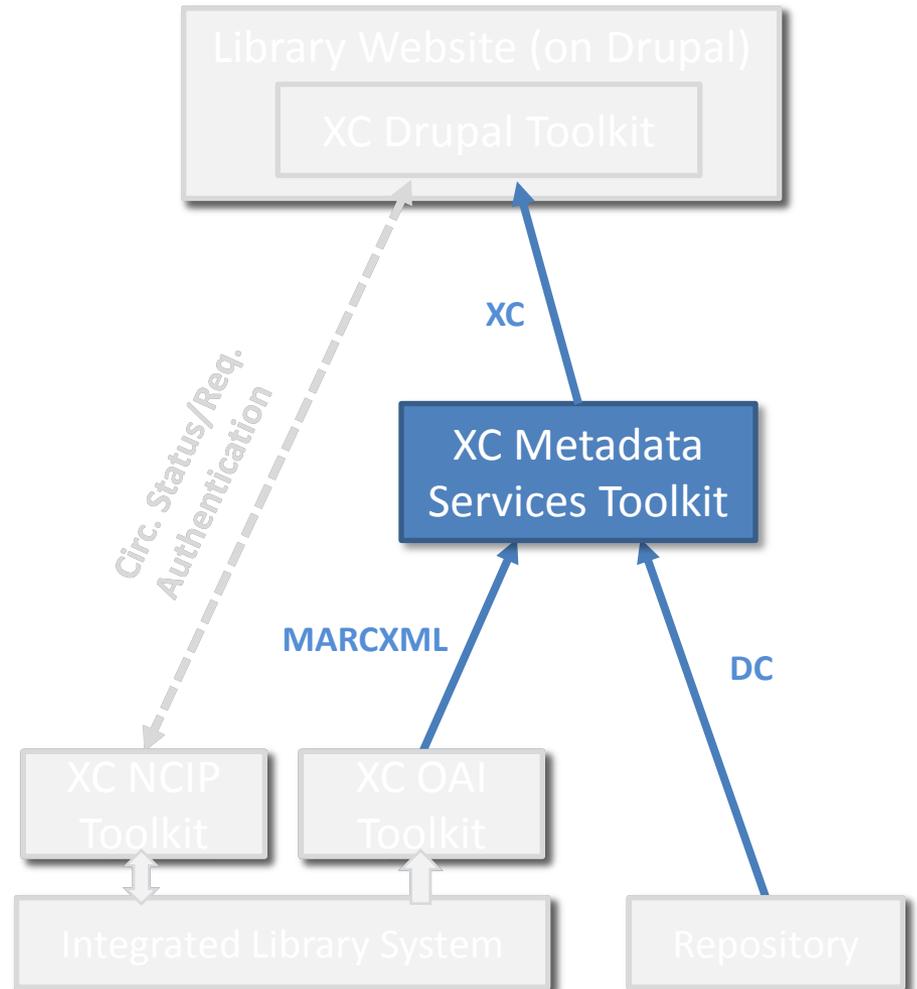
## Features:

New type of staff client for processing large batches of metadata through an orchestrated set of services.

Harvest from multiple sources (silos) to address format and quality issues.

Clean up, transform, and aggregate metadata .

Automatic synchronization propagates changes in source metadata through services and on to discovery interface.





## Functions:

- Add Repositories
- Schedule Harvests
- Add Metadata Services
- Add Processing Rules
- Browse Records
- View Logs
- Set up Users/Groups

**Metadata Services Toolkit**

Repository ▾ Harvest ▾ Services ▾ Processing Rules ▾ Browse Records Logs ▾ Users/Gr

Harvest | New Harvest: Step 1

1 Select Repository & Schedule 2 Select Sets & Formats

This will allow you to harvest records from the service and schedule incremental updates

Which **Repository** do you want to Harvest?  
Test repository ▾

How often do you want to harvest updated records?

**Schedule**

Hourly Run at 59 ▾ minutes past the hour

Daily Run at Select ▾ hour Select ▾ minutes each day

Weekly Run Every Select ▾ at Select ▾ each week

**Start Date :** 09/22/2009 Enter a future date for delayed start

**End Date :**  clear Leave blank for no end date



# MST Service Orchestration: Add Processing Rule

Take Control.

**Metadata Services Toolkit** Hi admin | My Account | Logout

Repository ▾ Harvest ▾ Services ▾ **Processing Rules ▾** Browse Records ▾ Logs ▾ Users/Groups ▾ Configuration ▾

Processing Rules | Add Processing Rule (Step 1)

**1** Select Source & Service **2** Select Sets & Formats & Services

Cancel Continue to Step 2 →

**Select Source for records to be Processed**

**External Repositories**

Test repository

**Input Records to Services**

Normalization Service

Transformation Service

**Output Records from Services**

Normalization Service

Transformation Service

Cancel Continue to Step 2 →

MetadataServicesToolkit v 0.2.6



# MST: Browse Records

Take Control.

**Metadata Services Toolkit** Hi admin | My Account | Logout

Repository ▾ Harvest ▾ Services ▾ Processing Rules ▾ **Browse Records** Logs ▾ Users/Groups ▾ Configuration ▾

Browse Records

**Repository** (173)

**Service** (173)

**Schema** (346)

**Set**

Test-repository (173)

MARCXML Bibliographic Records (114)

Test-repository:bib (114)

MARCXML Holding Records (59)

Test-repository:hold (59)

**Harvest**

Test repository 2009-09-22 08:59:26 (173)

**Error**

3-107: Invalid 035 Data Field (035s should not contain a \$b subfield) (51)

3-107: Invalid 035 Data Field (035s should not contain a \$9 subfield) (50)

3-106: Unrecognized language code: (4)

3-107: Invalid 035 Data Field (035s should not contain a \$z subfield) (2)

Search  Search Full XML content

“Browse Records” includes full-text search and faceted browse .

- Examples of Facets:
- Repository
  - Service
  - Schema
  - Errors



# MST: Browse Records

Take Control.

The screenshot shows the Metadata Services Toolkit (MST) interface. At the top, there is a navigation bar with the MST logo and the text 'Metadata Services Toolkit'. To the right of the logo, there is a user profile 'Hi admin' and links for 'My Account' and 'Logout'. Below the navigation bar, there is a search bar with the text 'Browse Records | Search Results'. The search criteria is 'Repository: Test repository'. There is a search button and a checkbox for 'Search Full XML content'. Below the search bar, there is a pagination bar showing 'Page 1 of 9 | 173 results' and a set of numbered links from 1 to 9, with 'Next' and 'Last' buttons. The main content area displays a list of search results. The first result is '1. Test repository 10000' with details: Schema: marcxml, Repository: Test repository, Harvest: Test repository Hourly. Below this, there is a red error message: 'Error: 3-107: Invalid 035 Data Field (035s should not contain a \$b subfield), 3-10 Field (035s should not contain a \$9 subfield)'. A green arrow points from the 'Error' section on the left sidebar to this error message. Below the error message, there is a link for '1 Successor'. The second result is '2. Test repository 10023' with details: Schema: marcxml, Repository: Test repository, Harvest: Test repository Hourly. Below this, there is a link for '1 Successor'. The third result is '3. Test repository 10022' with details: Schema: marcxml, Repository: Test repository, Harvest: Test repository Hourly. Below this, there is a red error message: 'Error: 3-107: Invalid 035 Data Field (035s should not contain a \$b subfield), 3-10 Field (035s should not contain a \$9 subfield)'. Below the error message, there is a link for '1 Successor'. On the left sidebar, there are sections for 'Repository', 'Service', 'Schema', 'Set', 'Harvest', and 'Error'. The 'Error' section lists three error messages: '3-107: Invalid 035 Data Field (035s should not contain a \$b subfield) (51)', '3-107: Invalid 035 Data Field (035s should not contain a \$9 subfield) (50)', and '3-106: Unrecognized language code.'.

Library staff use “Browse Records” to verify that services are functioning properly and to debug any issues.

The MST preserves the original record and new records produced by each metadata service. Navigation links take you to predecessor and successor records.



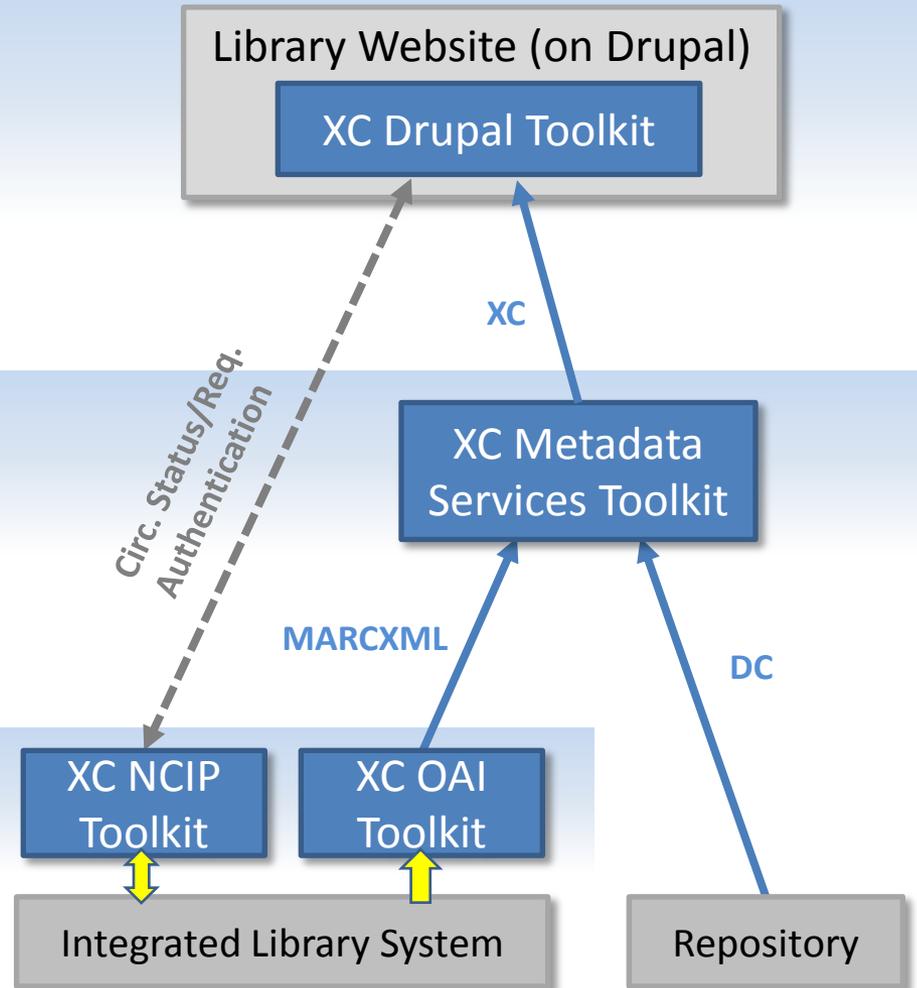
# XC Software Components

Take Control.

User Interface for searching and browsing

Tools for automated processing of large batches of metadata

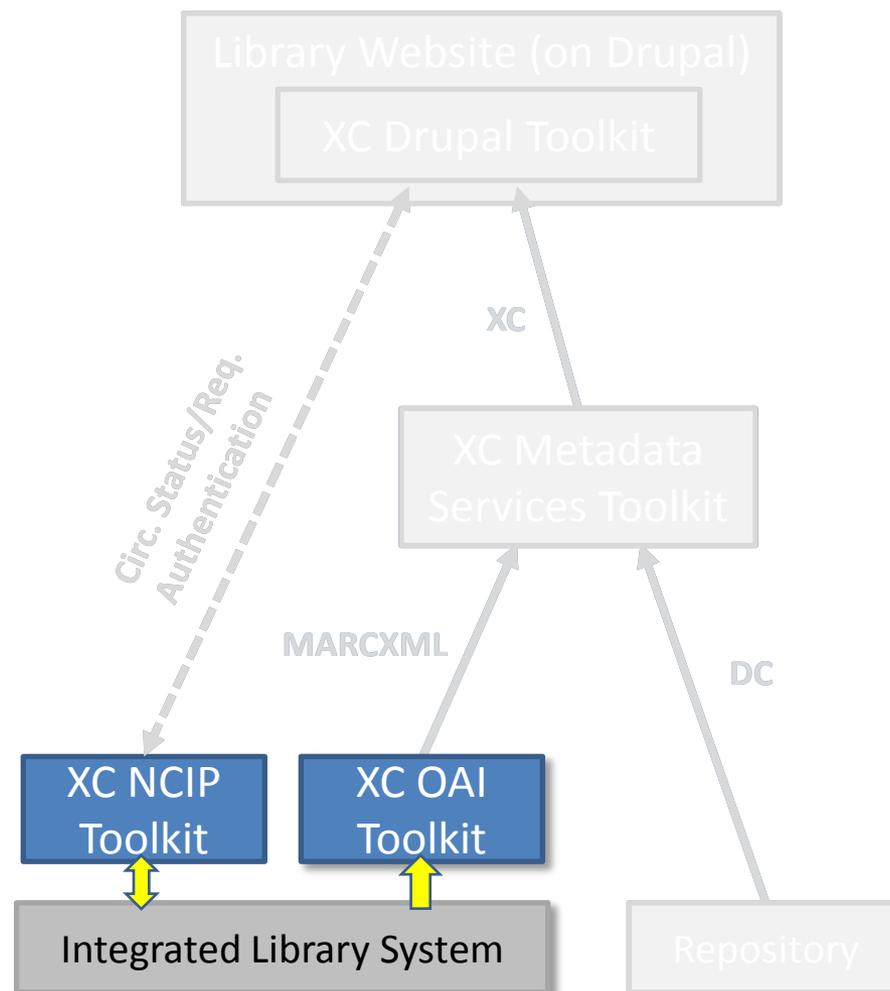
Tools for connectivity between XC and an ILS



## XC OAI Toolkit + XC NCIP Toolkit =

Connectivity between a range of ILSs and XC software

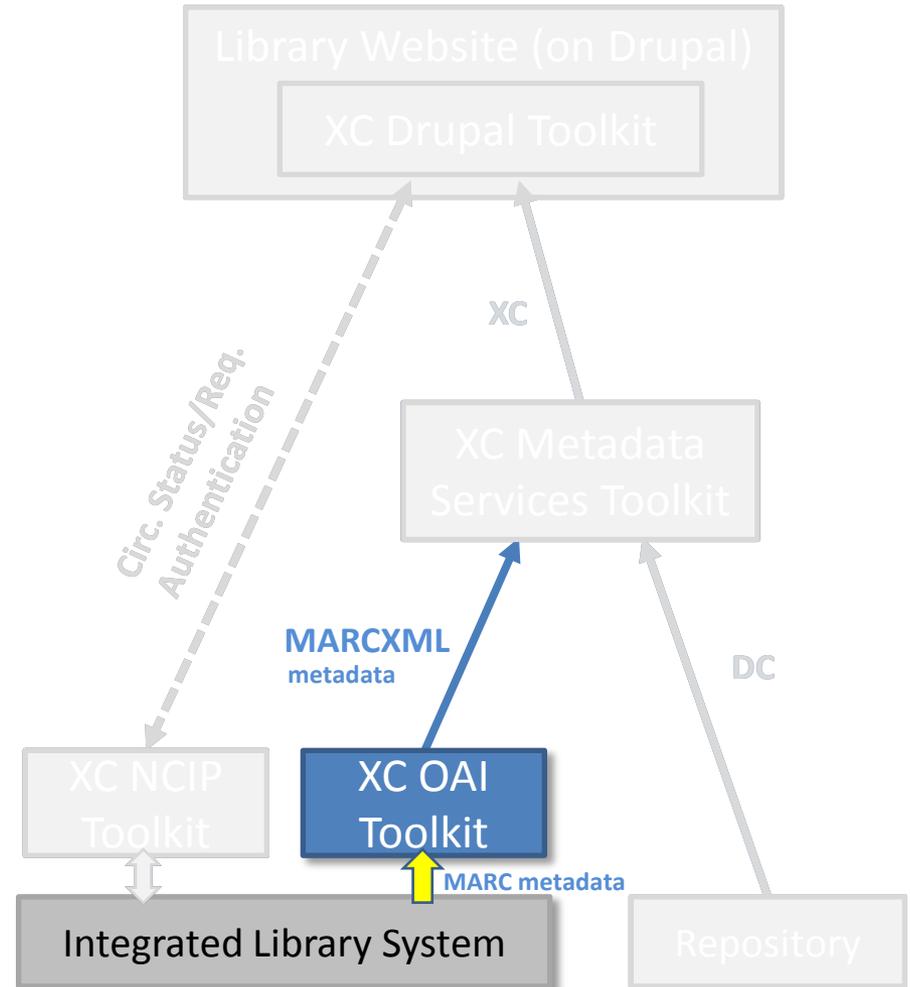
Standards-based to enable other next-generation catalog products to work with many legacy ILSs



## Overview:

Expose ILS metadata to XC's next generation catalog interface and metadata tools.

Synchronize ongoing changes in ILS records with XC software automatically.



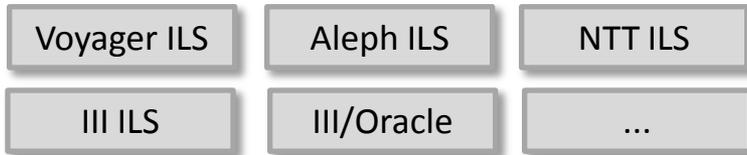


# OAI Toolkit: Compatibility

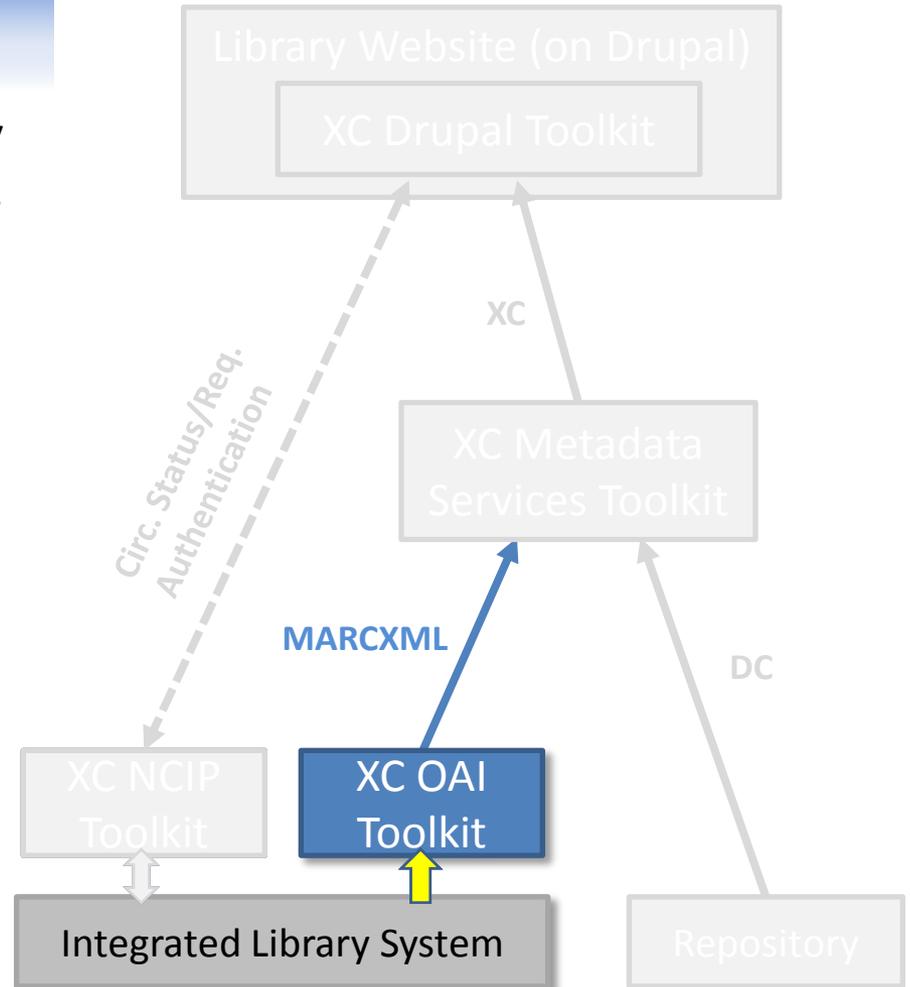
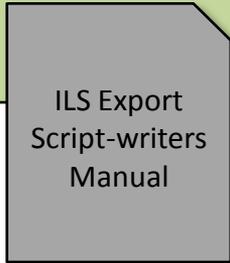
Take Control.

## Compatibility:

The XC OAI Toolkit works with virtually any ILS that includes MARC export capabilities.



ILS system administrators can program the OAI Toolkit with a familiar scripting language. No Java programming is required to make the OAI Toolkit work with a particular ILS.



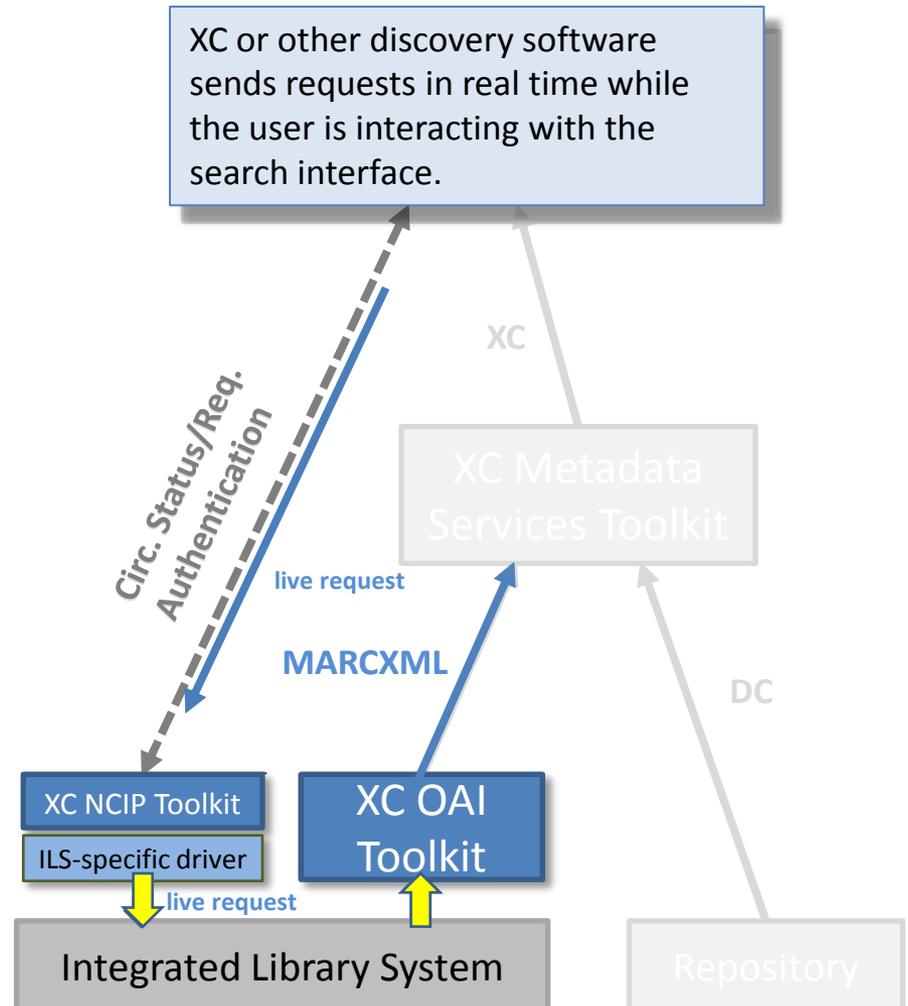


## Overview

Enable XC's next generation catalog interface access to ILS authentication, circulation and patron services.

## Example Requests

- Circulation status lookup
- Renew an item
- Lookup patron information



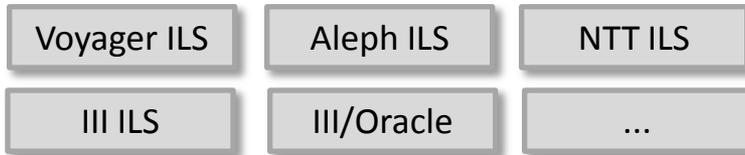


# NCIP Toolkit: Compatibility

Take Control.

## Compatibility:

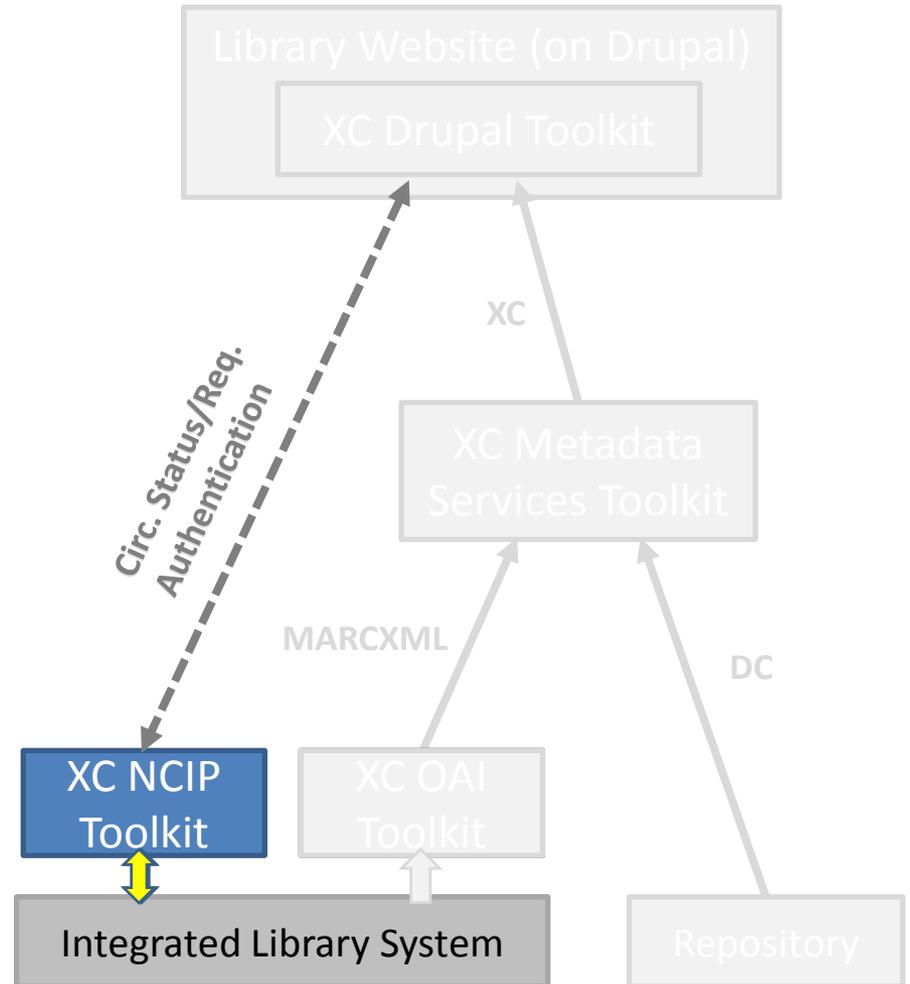
The XC NCIP Toolkit works with many Integrated Library Systems.



Build your own Java driver to allow the NCIP Toolkit to support an additional ILS.

The NCIP Toolkit allows developers to focus on interfacing with their ILS rather than learning the details of the NCIP protocol.

NCIP Developer Manual





- Download XC software at [www.eXtensibleCatalog.org](http://www.eXtensibleCatalog.org)
- Post bugs, questions, enhancement requests to XC Google Code or Drupal software sites
- New XC website: sign up for a mailing list
  - Announcements
  - General Discussion
  - Developer Discussion
- Become an eXtensible Catalog Organization (XCO) Sponsor
- Apply for invitation to XCO Meeting, May 6, 2010 Charlotte, NC



Take Control.

Download XC software at  
[eXtensibleCatalog.org](http://eXtensibleCatalog.org)