Library Catalog as Versatile Discovery Platform

Tito Sierra, Emily Lynema, and Markus Wust North Carolina State University Libraries DLF Fall Forum 2007 November 6, 2007

Outline

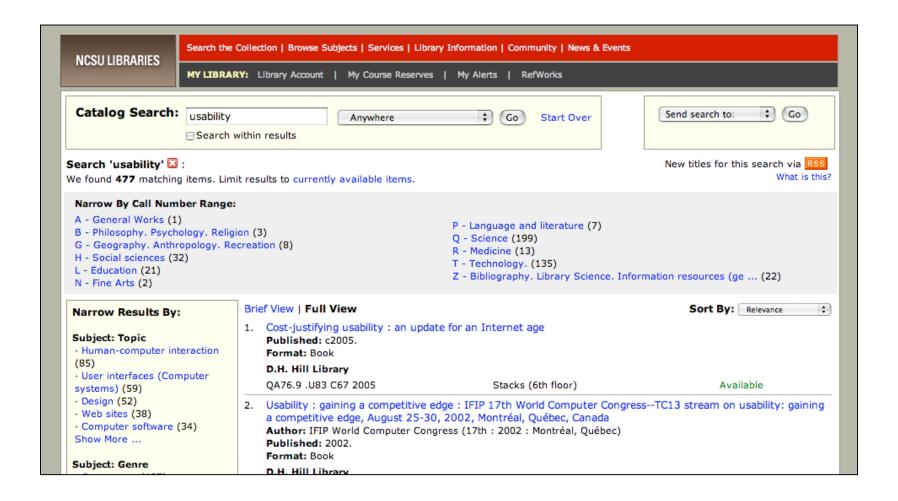
- Next Generation Catalogs
- Catalog as Discovery Platform
- CatalogWS API
- Current CatalogWS Applications
- Closing Thoughts

Next Generation Catalogs

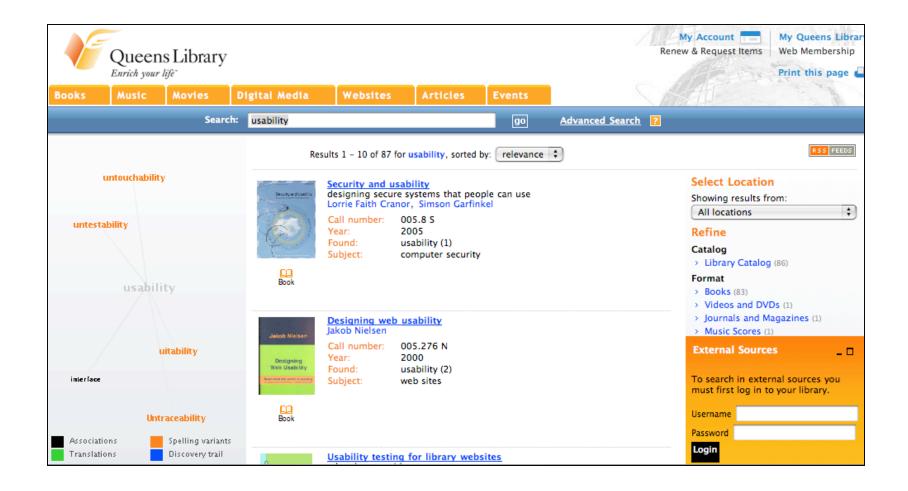
Background



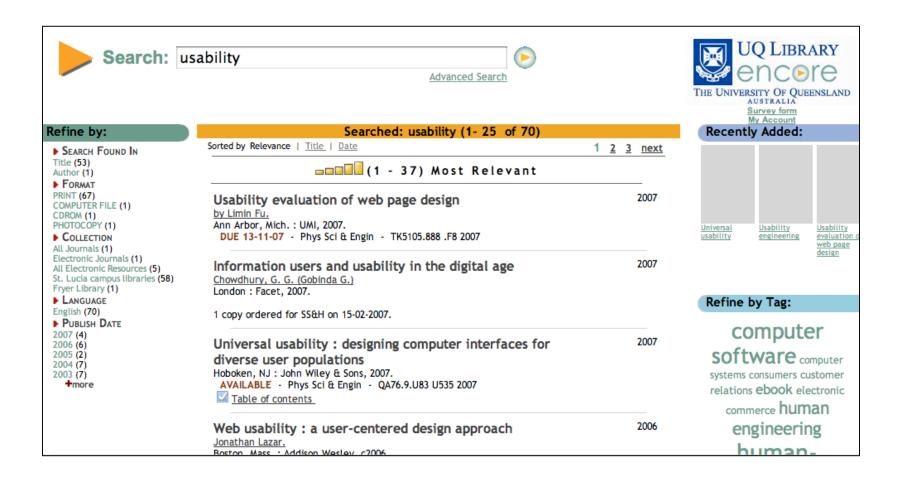
Examples - Endeca



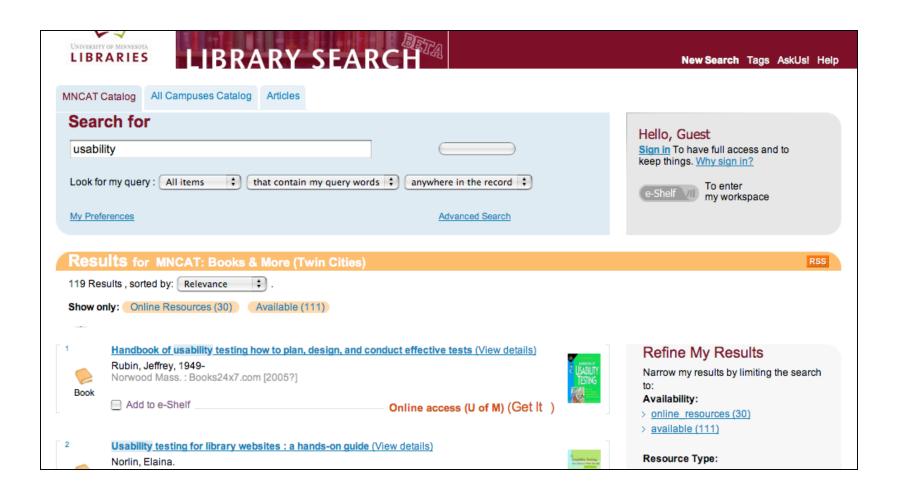
Examples - AquaBrowser



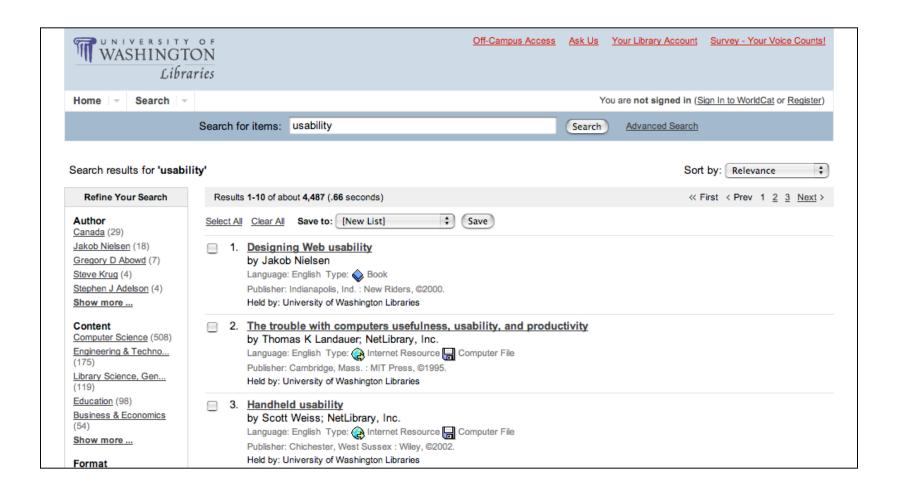
Examples - Encore



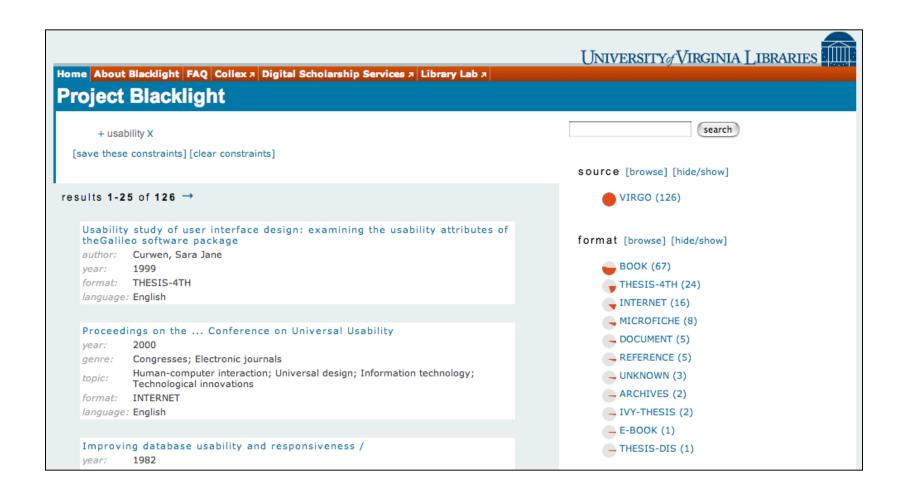
Examples - Primo



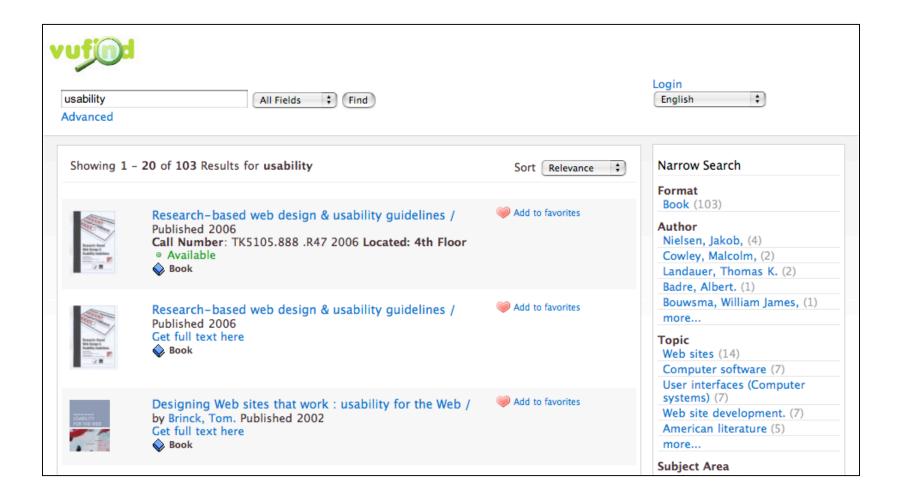
Examples - WorldCat Local



Examples - Solr-powered



Examples - Solr-powered



Next Generation Catalogs

- Modern search options
 - Relevance ranking
 - Faceted search
 - Tag/word clouds
- New content and functionality
 - User contributed content, social features
 - Enriched content (book covers, reviews)
 - Current awareness (RSS feeds)

Next Generation Catalogs

Current "next generation catalog" systems are largely focused on optimizing a single discovery context... the OPAC

Question...

Why should the discovery of cataloged library collections be limited to user interaction with a single catalog application?

Catalog as Discovery Platform

Platform?

- ... a "platform" is a system that can be reprogrammed and therefore customized by outside developers -- users -- and in that way, adapted to countless needs and niches that the platform's original developers could not have possibly contemplated, much less had time to accommodate.
- In contrast, an "application" is a system that cannot be reprogrammed by outside developers. It is a closed environment that does whatever its original developers intended it to do, and nothing more.
- Marc Andreessen, "Analyzing the Facebook Platform, three weeks in" (http://blog.pmarca.com/2007/06/analyzing_the_f.html)

Discovery Happens Elsewhere

No single website is the sole focus of a user's attention. Increasingly people discover websites, or encounter content from them, in a variety of places. These may be network level services (Google, ...), or personal services (my RSS aggregator or 'webtop'), or services which allow me to traverse from personal to network (Delicious, LibraryThing, ...).

— Lorcan Dempsey, "Discovery happens elsewhere" (http://orweblog.oclc.org/archives/001430.html)

Platform Motivation

- Move beyond the "one-size-fits-all" approach to catalog discovery
- Make it easier to reuse and repurpose catalog data outside the ILS/OPAC
- Build catalog interfaces optimized for different use contexts

CatalogWS API

CatalogWS Goals

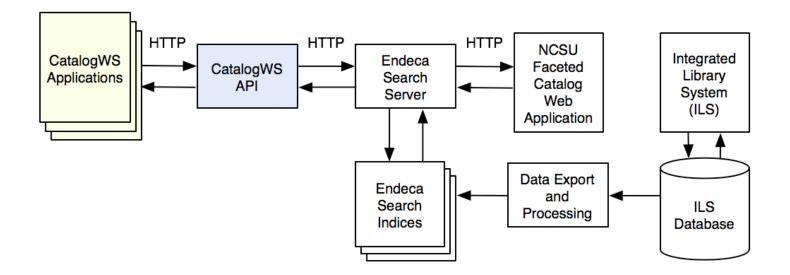
Initial impetus:

- Can we have RSS feeds for the catalog?
- Can we integrate catalog results into library website Quick Search?

Final result:

 Rich API for searching NCSU Libraries Catalog

NCSU Catalog Architecture



Search Indices as Data Source

Benefits:

- Performance
- Search features
- Consolidated and normalized data

Limitations:

- Subset of catalog data
- Read-only
- [Not real time]

Technical Design

- RESTful architecture
 - Define query using HTTP GET requests with URL parameters
- Separate web application handles API requests
 - Java, Tomcat, XOM, Saxon 8.8, org.json

API Functionality

Technical documentation:

http://www.lib.ncsu.edu/catalog/ws/

- Discovery-oriented
- Catalog availability
 - service=availability
 - Focus on known-item lookup (isbn)
- Catalog search
 - service=search
 - Support known-item and exploratory searching
 - Mimic functionality available on catalog search results page

Example Request

http://www.lib.ncsu.edu/catalogws/?service=search&query=deforestation

Required Parameters

- service
 - availability | search
- query
 - Any term(s)

Optional Parameters

- output
 - [xml] | rss | opensearch | json
- count
 - default: 30, max: 300
- offset
 - default: 0
- sort
 - [relevance] | date_desc | date_asc | call_number | most_popular | date_added
- style
 - URL of XSL to transform to custom output

XML Response

- Defined with Relax NG Schema
- Discovery-centric
 - Search metadata
 - Results
 - Basic bibliographic metadata (not MARC)
 - Holdings / Item metadata
 - Facet metadata
 - Links

```
<item id="1683670">
   <catalogLink>
       http://catalog.lib.ncsu.edu/web2/tramp2.exe/do_ccl_search/guest?se
   </catalogLink>
    <title>Causes of deforestation of the Brazilian Amazon</title>
    <author>Margulis, Sérgio.</author>
    <pubDate>2004.
    <format>Book</format>
   <isbn>0821356917</isbn>
   <holdings institution="ncsu">
        library name="D.H. Hill Library">
            <holdingsItem>
                <callNumber>SD418.3 .B6 M27 2003</callNumber>
                <location>Stacks (8th floor)</location>
                <status>Available</status>
            </holdingsItem>
        </library>
        library name="Natural Resources Library">
            <holdingsItem>
                <callNumber>SD418.3 .B6 M27 2003</callNumber>
                <location>Stacks</location>
                <status>Available</status>
            </holdingsItem>
       </library>
   </holdings>
</item>
```

```
<facet id="topic">
    <title>Subject: Topic</title>
   <catalogLink>
        http://www2.lib.ncsu.edu/catalog/?Ntt=deforestation&Ntk=Keyword&
   </catalogLink>
   (value)
        <requestUri>
            http://www2.lib.ncsu.edu:9921/catalogws/?service=search&quer
        </requestUri>
        <catalogLink>
            http://www2.lib.ncsu.edu/catalog/?Ntt=deforestation&Ntk=Keyw
        </catalogLink>
        <title>Deforestation</title>
        <count>205</count>
   </value>
   (value)
        <requestUri>
            http://www2.lib.ncsu.edu:9921/catalogws/?service=search&quer
        </requestUri>
        <catalogLink>
            http://www2.lib.ncsu.edu/catalog/?Ntt=deforestation&Ntk=Keyw
        </catalogLink>
        <title>Environmental aspects</title>
        <count>73</count>
   </value>
(/facet)
```

Current CatalogWS Applications

Current CatalogWS Applications

Three categories:

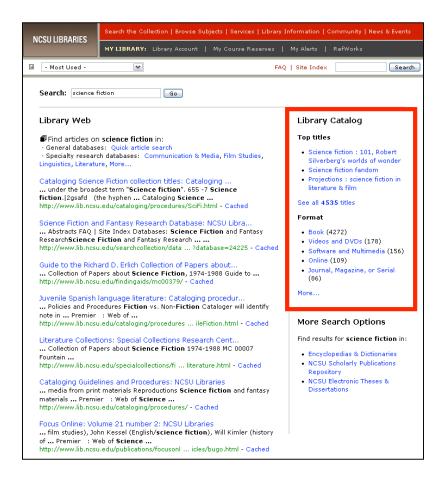
- 1. Integration with external applications
- 2. Alternative catalog interfaces
- 3. Collection promotion

Integration with External Applications

- Integrate catalog data in external library applications: Quick Search
- Integrate catalog data in external nonlibrary applications: iGoogle Widget

Quick Search

- Website Search Tool
- Combines:
 - Catalog search
 - Journal finder
 - Databases
 - Website search
 - Library FAQs



Quick Search

- Top matching titles
- Format facet results
- Links to full catalog search results

Library Catalog

Top titles

- Science fiction: 101, Robert Silverberg's worlds of wonder
- · Science fiction fandom
- Projections : science fiction in literature & film

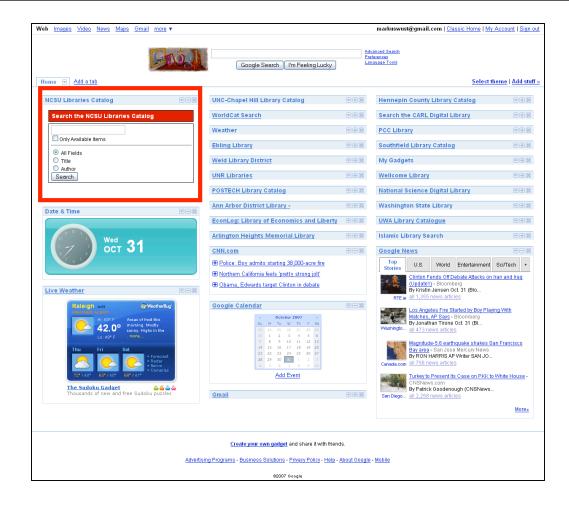
See all 4535 titles

Format

- Book (4272)
- Videos and DVDs (178)
- · Software and Multimedia (156)
- Online (109)
- Journal, Magazine, or Serial (86)

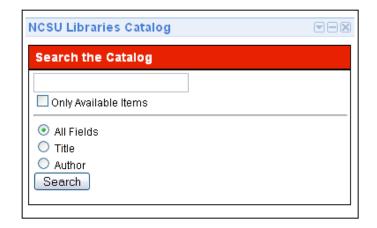
More...

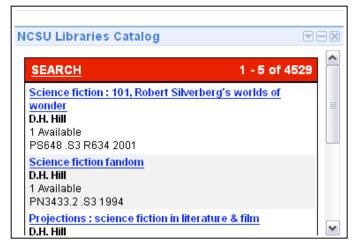
Google Widget



Google Widget

 Search and display catalog results within iGoogle





Alternative Catalog Applications

- MobiLIB Catalog: basic catalog interface for mobile devices
- FacetBrowser: experimental facetbased browsing interface

MobiLIB Overview

Current services:

- Catalog search
- Computer availability
- Library opening hours
- Campus directory
- Contact information for selected library departments and services
- Links to external sites
- Location of buses in campus transit system

- Usage scenario: Known-item search on mobile device (smart phone, PDA, ...)
- Bare-bones search interface
- Only one facet: Availability

Search fields:

- Keyword
- Title
- Author
- ISBN



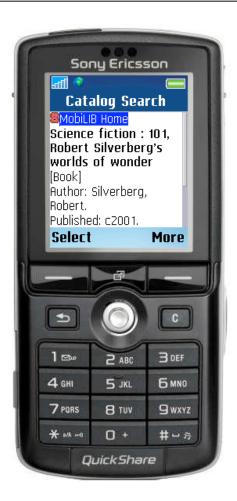
Results display:

- Short title
- Library
- Availability
- Call number if only one copy



Item display:

- Full title
- Format
- Author
- Date
- Library
- Availability
- Call number
- Location



FacetBrowser

- Exploratory search
- Focus on browsing with facets
- Same facets as OPAC
- Basic keyword search

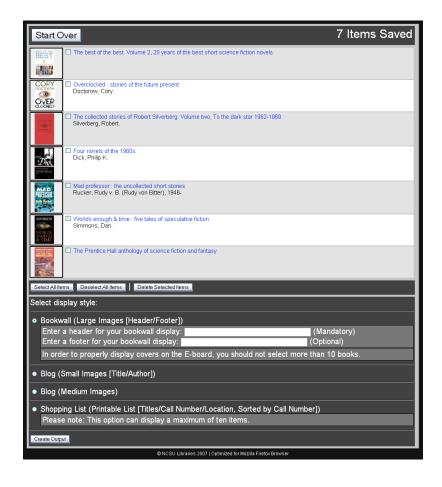


Collection Promotion

- FacetBrowser as a collection promotion authoring tool
- Automatically generated "bookwalls"
- RSS feeds for new titles

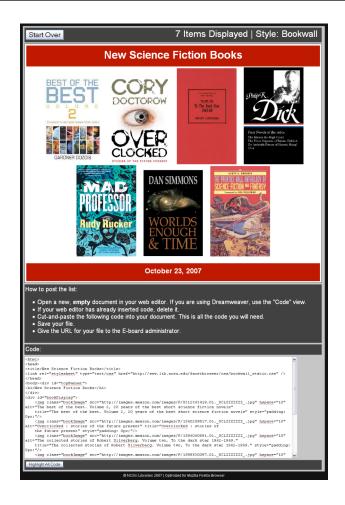
FacetBrowser

- Includes shopping cart for saving selected items
- Currently offers four different output styles



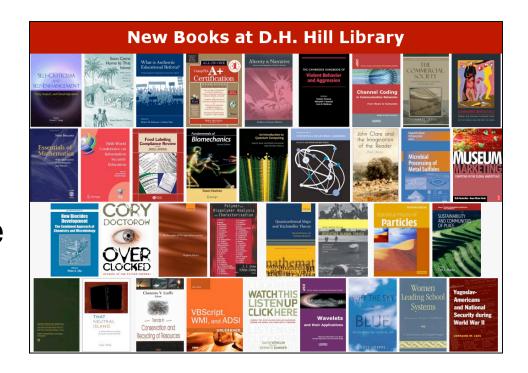
FacetBrowser

- Displays output
- Generates HTML code for bookwall displays and blogs



New Books Bookwall

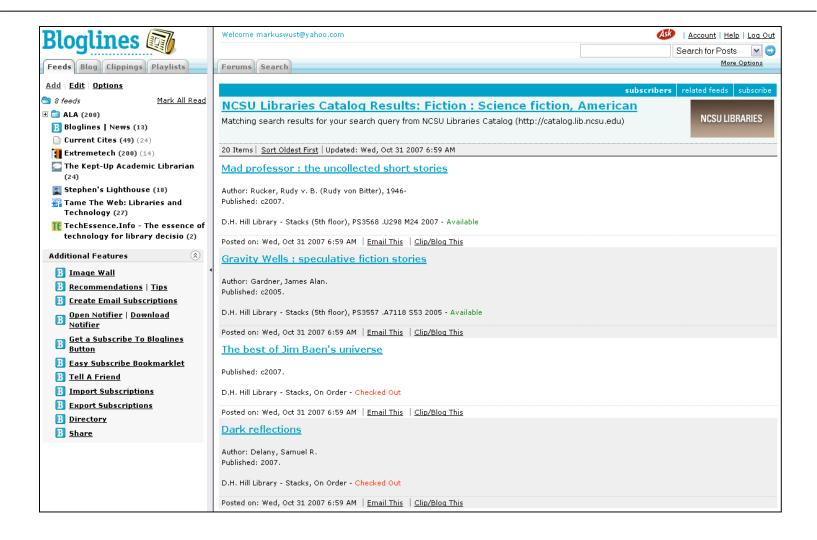
- Newest books in catalog
- Automatically generated
- Maintenance free



RSS Feeds

- Users can subscribe to RSS feeds based on their individual search preferences
- Active promotion
 - Information about newest items is pushed to users

RSS Feeds



Closing Thoughts

Benefits

Reduced development costs

The CatalogWS API has lowered the technical barriers for staff to develop new web applications that use catalog data.

Benefits

Versatile access to catalog data

The resource-oriented nature of the CatalogWS API makes it easy to integrate catalog data in external applications.

Platforms, Again

- ... a "platform" is a system that can be reprogrammed and therefore customized by outside developers -- users -- and in that way, adapted to countless needs and niches that the platform's original developers could not have possibly contemplated, much less had time to accommodate.
- Marc Andreessen, "Analyzing the Facebook Platform, three weeks in" (http://blog.pmarca.com/2007/06/analyzing_the_f.html)

Thanks!

Tito Sierra

tito_sierra@ncsu.edu

Emily Lynema

emily_lynema@ncsu.edu

Markus Wust

markus_wust@ncsu.edu

More Information

CatalogWS API:

http://www.lib.ncsu.edu/catalog/ws/

CatalogWS Applications:

http://www.lib.ncsu.edu/dli/projects/catalogwsapps/