Expose Your Content with XTF

Martin Haye (martin.haye@ucop.edu)

Kirk Hastings (kirk.hastings@ucop.edu)

Lisa Schiff (lisa.schiff@ucop.edu)

California Digital Library

University of California

XTF

eXtensible Text Framework

Session Outline

Part 1: *Customizing an XTF instance*
- Install XTF
- Swap in new content
- Implement bookreader features
- Add a new language to the globalized interface
- Display and customize a facet

Part 2: *How do I...?*
- Participant questions—discussions & demonstrations

What is XTF?

- A framework for building robust digital content applications

- Java Servlets + Lucene + Saxon + XSLT 2.0 = XTF

- Used by institutions worldwide

- Open source ➔ http://xtf.cdlib.org/download

- Maintained and supported by CDL developers

XTF 3.0

- New features
  - Globalization – users choose interface language
  - OpenLibrary Book Reader
  - Updated tutorial package
  - Bug fixes

- Release dates
  - Alpha—October 29, 2010
  - Beta—November 2010
  - Production—early December 2010

Part 1: Customizations

- Install XTF from scratch

- Then we'll customize

Follow Along: XTF in 5 Easy Steps

1. Download XTF: *Windows* or *Mac/Linux*
2. Extract content: *Double click the package*
3. Index sample documents:
   - Double click ../xtf3Workshop/cmdPrompt.[bat]
   - Type: textIndexer -index default
4. Start the web servlet:
   - *Double click ../xtf3Workshop/tomcat.[bat]*
5. Use the application!
   - Browse * Search * View EADs * View PDFs

Swap in New Content (1)

- Adding content to the XTF index is easy
  - Formats: PDF, WORD, ASCII, HTML, XML, TEI, NLM, EAD, and now Book Reader
- Metadata
  - 1 metadata file alongside each document, or stored within XML documents
  - Can use Dublin Core for basic metadata
  - Can work with METS (extra config. required)
  - Extend to use another or your own data format

Swap in New Content (2)

- Download
  - Sample open access books from HathiTrust:
    - copy from USB stick, or download http://xtf.cdlib.org/wp-content/uploads/2010/10/bookreader.zip
  - Unpack here: ../xtf3Workshop/data/ (creates new 'bookreader' folder)
- Index
  - Double click../xtf3Workshop/cmdPrompt.bat
  - Incremental index
    - **textIndexer -index default**
- View
  - Look at your content here: http://localhost:8080/xtf/search?browse-all=yes

Book Reader Customization

Switch search results to left side

Make font match other XTF pages

Bibliographic info in object view

Globalization - Adding a Langage (1)

- Users can now select a language for the interface.
  - Covers all UI text strings (e.g. labels, form buttons, drop-down options, etc.)
  - Content and metadata are NOT translated.

- XTF 3.0 default interface translations:
  - English, French, Spanish, German, Italian

- Add/substitute any languages you want

Globalization - Adding a Language (2)

- Create a mapping template

- Translate the English terms into the <to> element

- Expand with additional words and phrases

- Add a new radio button to the language form

Add a Facet

- Why would I use facets?
- Add type facet to the indexer (textIndexer)
  - 1. Name and attributes
  - 2. Flat or hierarchical?
- Add facet-type query support (queryParser)
  - Sorting options
  - Select statement
- Update results display (resultFormatter)
  - Change display label
  - Remove +/- expansion icon

Part 2: *How do I ???*

- Discussion and demonstration

    ▫ What do you want to know how to do?

    ▫ What kind of content do you have?

    ▫ Who are your users and what will they want to do on your site?