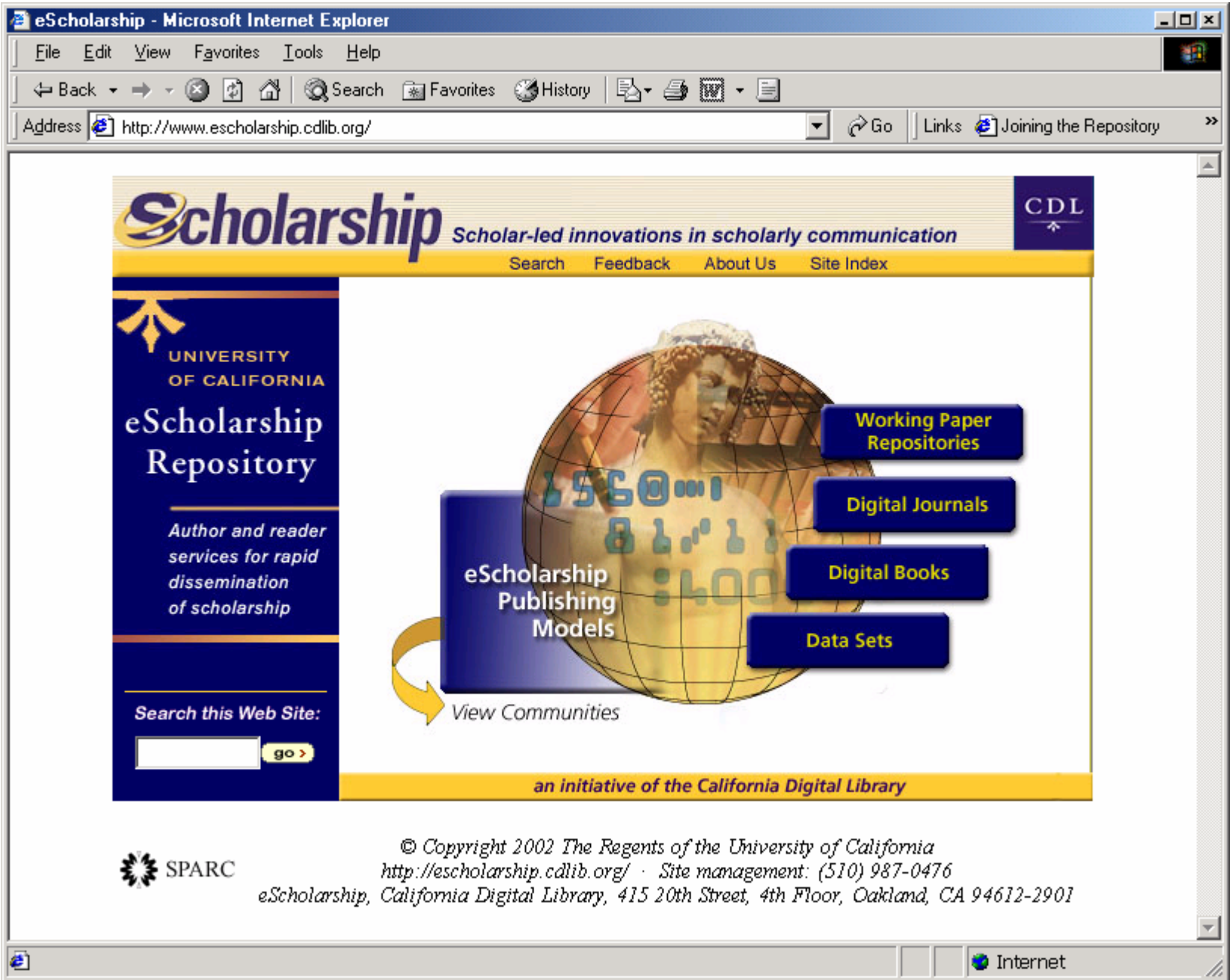


Acquiring and Ingesting Digital Content: Methods of the CDL's eScholarship Program



Suzanne Samuel and
David Loy
California Digital Library



eScholarship

Scholar-led innovations in scholarly communication

CDL

Search Feedback About Us Site Index

UNIVERSITY OF CALIFORNIA

eScholarship Repository

Author and reader services for rapid dissemination of scholarship

Search this Web Site:

go >

eScholarship Publishing Models

View Communities

- Working Paper Repositories
- Digital Journals
- Digital Books
- Data Sets

an initiative of the California Digital Library



© Copyright 2002 The Regents of the University of California
http://escholarship.cdlib.org/ · Site management: (510) 987-0476
eScholarship, California Digital Library, 415 20th Street, 4th Floor, Oakland, CA 94612-2901

Program Examples





- ⌘ The eScholarship Repository and UC International and Area Studies Digital Collection
- ⌘ Handbook of Visual Optics
- ⌘ University of California Press eScholarship Editions

eScholarship Repository - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites History Print Copy Paste

Address <http://repositories.cdlib.org/escholarship/> Go Links Joining the Repository >>





eScholarship

- home
- search
- help
- my account

The eScholarship Repository offers faculty a central location for depositing any research or scholarly output deemed appropriate by their participating University of California research unit, center, or department, including working papers and pre-publication scholarship. This is a service of the [eScholarship initiative](#) of the [California Digital Library](#). For more information, see [About the Repository](#) and [Joining the Repository](#).

Search all papers:

- [Berkeley Olin Program in Law & Economics, Working Paper Series](#) (Berkeley)
- [Center for Comparative Immigration Studies](#) (San Diego)
- [Center for Global, International and Regional Studies](#) (Santa Cruz)
- [Center for Research on Education, Diversity & Excellence](#) (Santa Cruz)
- [Center for U.S.-Mexican Studies](#) (San Diego)
- [Institute of Business and Economic Research](#) (Berkeley)
 - [Berkeley Program on Housing and Urban Policy](#)
 - [Robert D. Burch Center for Tax Policy and Public Finance](#)



Done Internet



UNIVERSITY OF CALIFORNIA
eScholarship Repository



eScholarship > SIO

[home](#)[search](#)[help](#)[my account](#)SCRIPPS
INSTITUTION OF
OCEANOGRAPHY**Scripps Institution of
Oceanography**

University of California, San Diego

[SIO Papers](#) • [SIO Website](#) • [Policies](#) • [Search SIO](#) • [Submit a Paper](#)[search
melvyl](#)

Scripps Institution of Oceanography, founded in 1903, is the oldest oceanographic institution in the Western Hemisphere. Its mission is "To seek, teach, and communicate scientific understanding of the oceans, atmosphere, Earth, and other planets for the benefit of society and the environment."

Collections for *Scripps Institution of Oceanography*:

- [Scripps Institution of Oceanography Technical Report](#)
- [SIO Reference](#)
- [Scripps Institution of Oceanography Archives](#)

Collections for Affiliated Groups:

- [Center for Marine Biodiversity and Conservation](#)



OPEN ARCHIVES

eScholarship > ISEEEES > BPS > Paper 2002_02-wert

home

search

help

my account

search
melvyl



eScholarship



OPEN ARCHIVES



UCIAS

Institute of Slavic, East European, and Eurasian Studies

Berkeley Program in Soviet and Post-Soviet Studies

University of California, Berkeley

BPS Papers • BPS Website • Policies • Search • Submit a Paper

Language "Purity" and the De-Russification of Tatar

Suzanne Wertheim, *University of California, Berkeley*

[Download the Paper](#) (PDF format) - May 1, 2002

[Tell a colleague](#) about it.

Printing Tips: Select 'print as image' in the Acrobat print dialog if you have trouble printing.

ABSTRACT:

Tatar is a Turkic language spoken in Tatarstan, Russia, by one quarter of its four million residents. Tatar ethnic identity, as constructed through linguistic performance, is inextricably linked with orientation towards or away from Russian language and culture, such that the integrity and cultural "purity" of post-Soviet Tatars -- thought by many to be necessary for the survival of the Tatar language, culture, and nation -- is equated with de-Russification. "Purification" practices usually involve the cleansing of Russian influence, while the influences of other cultures (e.g., Persian, Arabic) are allowed to remain. This de-Russification, as expressed through a variety of purification movements, takes place in a post-Soviet sociolinguistic setting where, although the climate is more conducive to the use of Tatar than in years past, language shift to Russian continues apace.

SUGGESTED CITATION:

Wertheim, Suzanne, "Language "Purity" and the De-Russification of Tatar" (May 1, 2002). *Berkeley Program in Soviet and Post-Soviet Studies*. Working Paper 2002_02-wert. http://repositories.cdlib.org/iseees/bps/2002_02-wert

University of California Digital Repositories - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites History Print Copy Paste

Address <http://repositories.cdlib.org/> Go Links Joining t >>

 **University of California Digital Repositories**

[eScholarship Repository](#) [UCIAS Digital Collection](#)

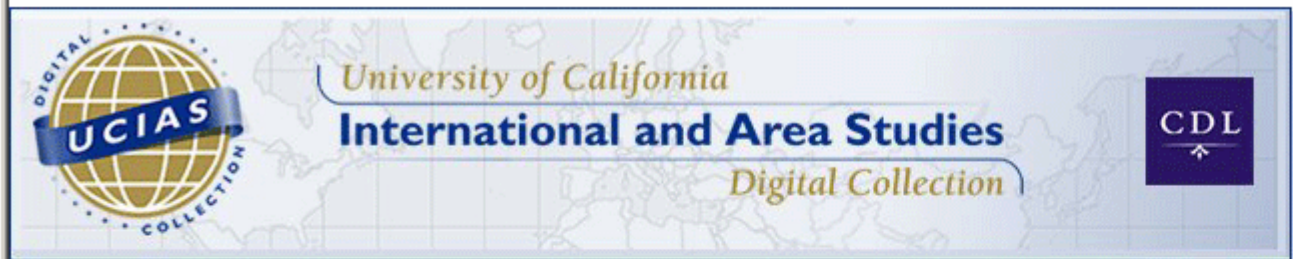
The University of California Digital Repositories offer faculty a central location for depositing scholarship. This is a service of the [eScholarship](#) initiative of the [California Digital Library](#).

Search all Repositories:

Repositories

- [eScholarship Repository](#) -- working papers, research and technical reports, etc.
- [University of California International and Area Studies Digital Collection](#) -- peer-reviewed papers, edited volumes, and monographs.

Internet



HOME

- About UCIAS
- Editorial Board
- Policies
- UCIAS Working Papers
- Search
- Help
- My Account

Welcome to the new online home of **The University of California International and Area Studies (UCIAS) Digital Collection** — a peer-reviewed electronic publications program.

UCIAS is a partnership of the University of California Press, the California Digital Library (CDL), and internationally oriented research units on eight UC campuses. The Digital Collection publishes articles, monographs, and edited volumes that are peer-reviewed according to standards set by an interdisciplinary UCIAS Editorial Board and approved by the University of California Press. UCIAS makes digital versions of these works available free of charge to a global network of scholars and encourages international intellectual exchange and research collaboration. Selected volumes may also be published in hard copy by the University of California Press.

For more information on the project, see our [press release](#).

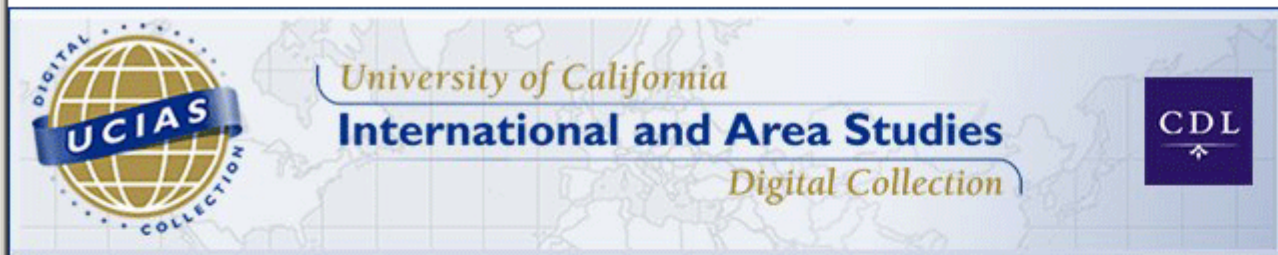
Search the Collection:

[UCIAS Articles](#)

[UCIAS Edited Volumes](#)

[UCIAS Monographs](#)

[Research Series \(previously published by IAS\)](#)



HOME > EDITED VOLUMES > Volume 3

About UCIAS

Editorial Board

Policies

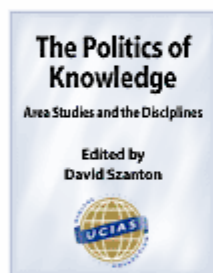
UCIAS Working Papers

Search

Help

My Account

UCIAS Edited Volumes



The Politics of Knowledge: Area Studies and the Disciplines

Edited by David L. Szanton

2002 — [View Table of Contents](#)

Published in association with the University of California Press

SUGGESTED CITATION: David L. Szanton, ed. *The Politics of Knowledge: Area Studies and the Disciplines*. University of California Press/University of California International and Area Studies Digital Collection, Edited Volume #3, 2002.
<http://repositories.cdlib.org/uciaspubs/editedvolumes/3>

Contents

7. [Eastern Europe or Central Europe? Exploring a Distinct Regional Identity](#), *Ellen Comisso and Brad Gutierrez*

8. [The Transformation of Contemporary China Studies, 1977-2002](#), *Andrew G. Walder*

University of California Press eScholarship Editions - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites History Print Copy Paste

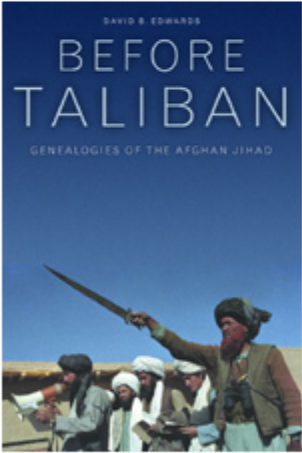
Address <http://escholarship.cdlib.org/rtenant/test/ucpress/> Go Links >>

California

The University of California Press Scholarship Editions

Over 500 online editions are available here through a partnership with the [University of California Press](#) and the California Digital Library's [eScholarship](#) program. Nearly 300 of the titles are available to the general public; the rest are for University of California faculty, staff, and students only. [[about us](#) | [help](#)]

Featured Title:



» **Search** [[more options](#)]

» **Browse**

- [By Subject](#)
- [By Author](#)
- [By Title](#)

Comments? Questions?

University of California Press eScholarship Editions are published by [eScholarship](#), the [California Digital Library](#)
© 2002 The Regents of the University of California

UC Press Description Search - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites History Print Mail

Links >>

California
The University of California Press eScholarship Editions

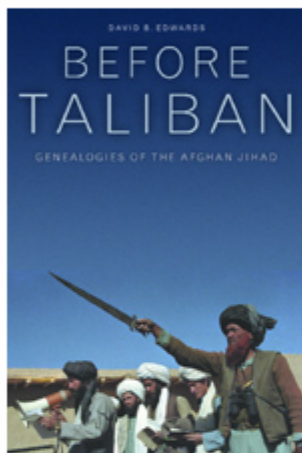
Search "Anthropology" ucp in the descriptions or full text [[Advanced search](#)]

Your search for "'Anthropology'" found 91 books, sorted by relevance. [Change to the brief display.](#)

Public = Available to Everyone; All others are UC Only

Results on this page: 1-20 • Jump to: 1... 2... 3... 4... 5...

1. **[From Savage to Negro: Anthropology and the Construction of Race, 1896-1954](#)** — Lee Baker
Lee D. Baker explores what racial categories mean to the American public and how these meanings are reinforced by anthropology, popular culture, and the law. Focusing on the period between two landmark Supreme Court decisions - *Plessy v. Ferguson* (the so-called "separate but equal" doctrine established in 1896) and *Brown v. Board of Education* (the public school desegregation decision of 1954) - Baker shows how racial categories change over time. Baker paints a vivid picture of the relationships between specific African American and white scholars, who orchestrated a paradigm shift within the social sciences from ideas based on Social Darwinism to those based on cultural relativism. He demonstrates that the greatest impact on the way the law codifies racial differences has been made by organizations such as the NAACP, which skillfully appropriated the new social science to exploit the politics of the Cold War.
Subjects: [Anthropology](#) | [Cultural Anthropology](#) | [Ethnic Studies](#) | [African American Studies](#) | [United States History](#) | [Sociology](#) | [Racism in anthropology -- United States -- History](#) | [Anthropology -- United States -- History](#) | [Racism in popular culture -- United States -- History](#) | [Public opinion -- United States](#) | [Afro-Americans -- Public opinion](#) | [United States -- Race relations](#) |
2. **[Expectations of Modernity: Myths and Meanings of Urban Life on the Zambian Copperbelt](#)** — James Ferguson
Once lauded as the wave of the African future, Zambia's economic boom in the 1960s and early 1970s was fueled by the export of copper and other primary materials. Since the mid-1970s, however, the urban economy has rapidly deteriorated, leaving workers scrambling to get by. *Expectations of Modernity* explores the social and cultural responses to this prolonged period of sharp economic decline. Focusing on the experiences of mineworkers in the Copperbelt region, James Ferguson



Before Taliban *Genealogies of the Afghan Jihad*

David B. Edwards

Sample citation: Edwards, David B. *Before Taliban: Genealogies of the Afghan Jihad*. Berkeley : University of California Press, 2002.
<http://ark.cdlib.org/ark:/13030/ft3p30056w/>

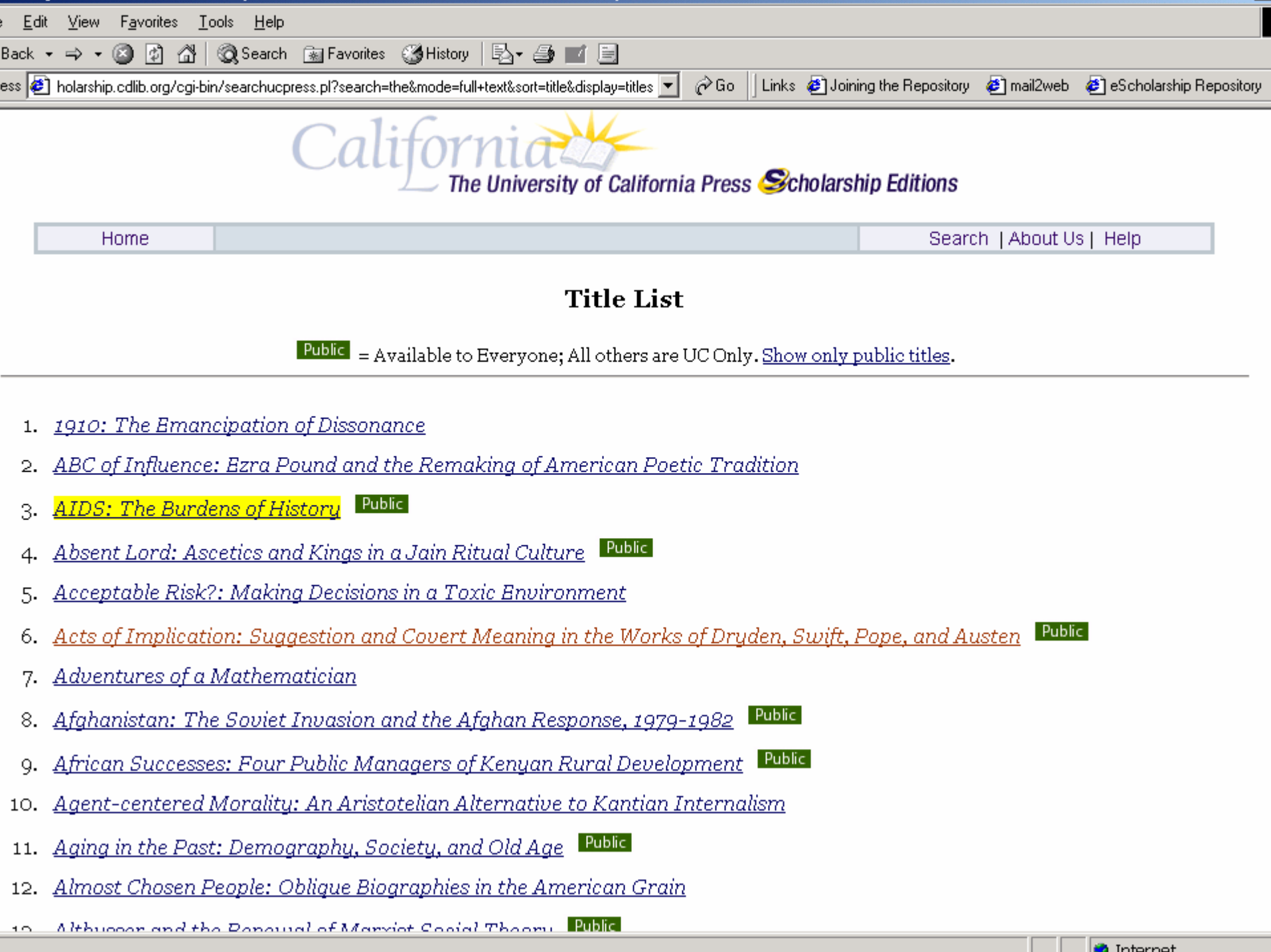
Contents

Search this Book:

- [Significant Persons and Parties](#)
- [Author's Note](#)
- [Preface](#)

- [Introduction](#)
- [Part I The Saur Revolution](#)
 - [1. Lives of the Party](#)
 - [2. The Armature of Khalqi Power](#)
 - [3. Coda The Death of a President](#)
- [Part II The Pech Uprising](#)
 - [4. A Son of Safi](#)
 - [5. Anatomy of a Tribal Uprising](#)
 - [6. Coda The Death of a Safi Daughter](#)
- [Part III The Islamic Jihad](#)



Title List

Public = Available to Everyone; All others are UC Only. [Show only public titles.](#)

1. [1910: The Emancipation of Dissonance](#)
2. [ABC of Influence: Ezra Pound and the Remaking of American Poetic Tradition](#)
3. [AIDS: The Burdens of History](#) **Public**
4. [Absent Lord: Ascetics and Kings in a Jain Ritual Culture](#) **Public**
5. [Acceptable Risk?: Making Decisions in a Toxic Environment](#)
6. [Acts of Implication: Suggestion and Covert Meaning in the Works of Dryden, Swift, Pope, and Austen](#) **Public**
7. [Adventures of a Mathematician](#)
8. [Afghanistan: The Soviet Invasion and the Afghan Response, 1979-1982](#) **Public**
9. [African Successes: Four Public Managers of Kenyan Rural Development](#) **Public**
10. [Agent-centered Morality: An Aristotelian Alternative to Kantian Internalism](#)
11. [Aging in the Past: Demography, Society, and Old Age](#) **Public**
12. [Almost Chosen People: Oblique Biographies in the American Grain](#)
13. [Althusser and the Renewal of Marxist Social Theory](#) **Public**

Building a Digital Books System



User Functionality: Collection Search



⌘ Search books

- ☑ Search book text

- ☑ Search book metadata

⌘ Melvyl - UC online Catalog

- ☑ Search Marc records for books

User Functionality: Collection Browsing

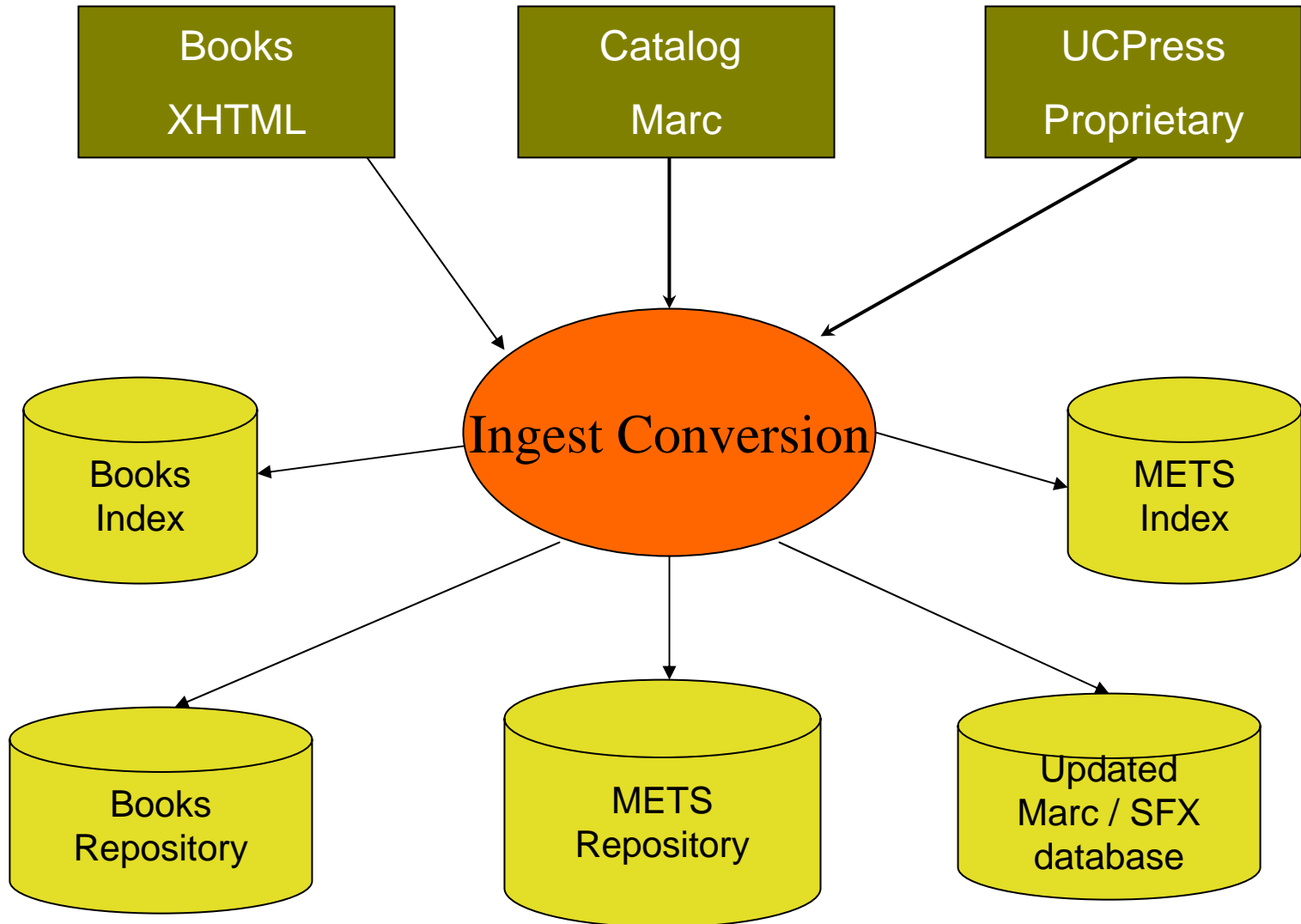


⌘ Title

⌘ Author

⌘ Subject

Ingest Conversion



The Inputs



- ⌘ The digital books - XHTML format
- ⌘ UCPress database - metadata and administrative information
- ⌘ LC Marc

The Outputs



- ⌘ TEI representation of the digital books in XML
- ⌘ METS metadata for each book built from MODS and UC Press proprietary data
- ⌘ SFX and MARC data to locate the electronic form of book

Technical Goals



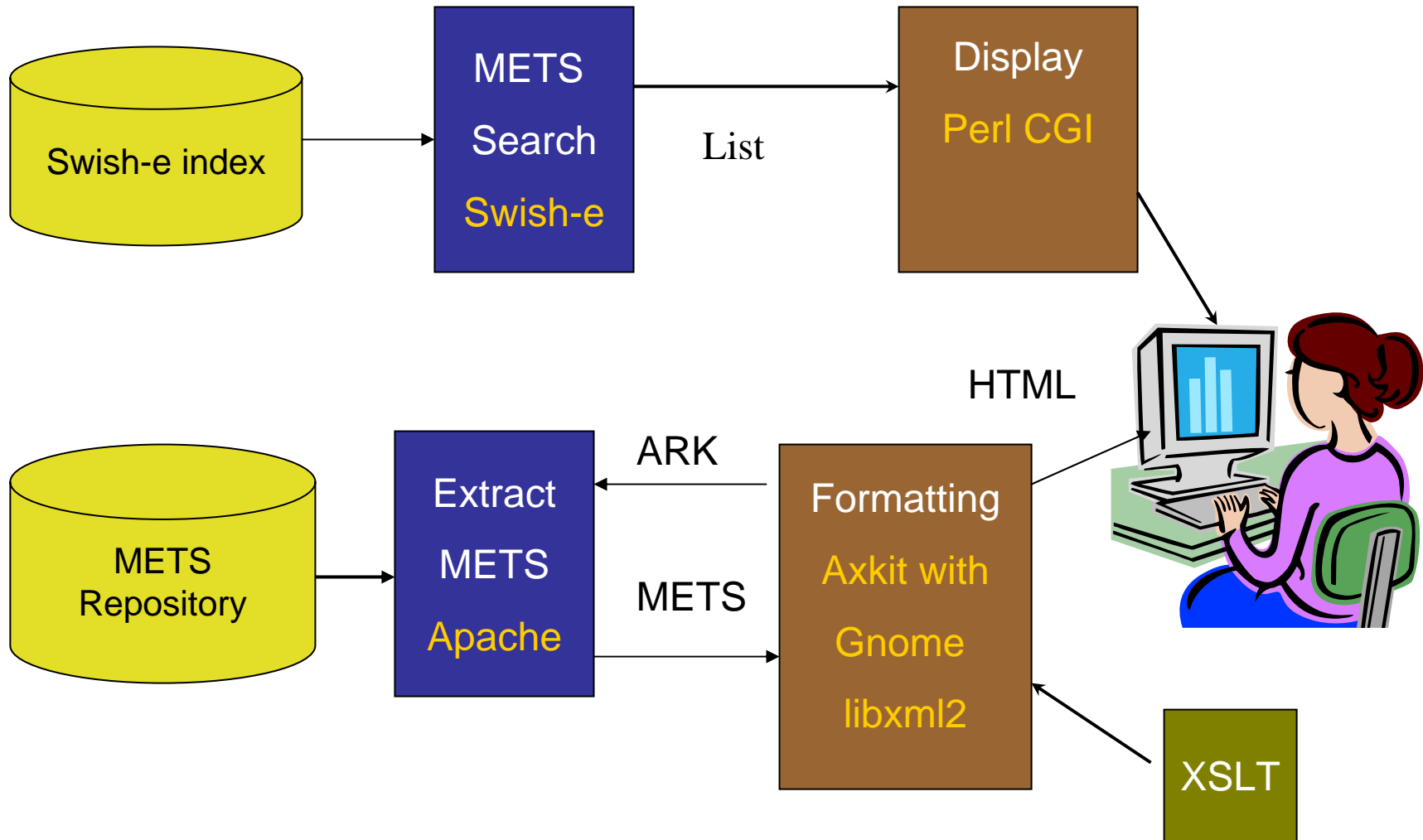
- ⌘ Use national standard formats for representing digital objects
- ⌘ When possible use Open Source technologies
- ⌘ Provide OAI compliant services to allow harvesting data

Standards for digital objects

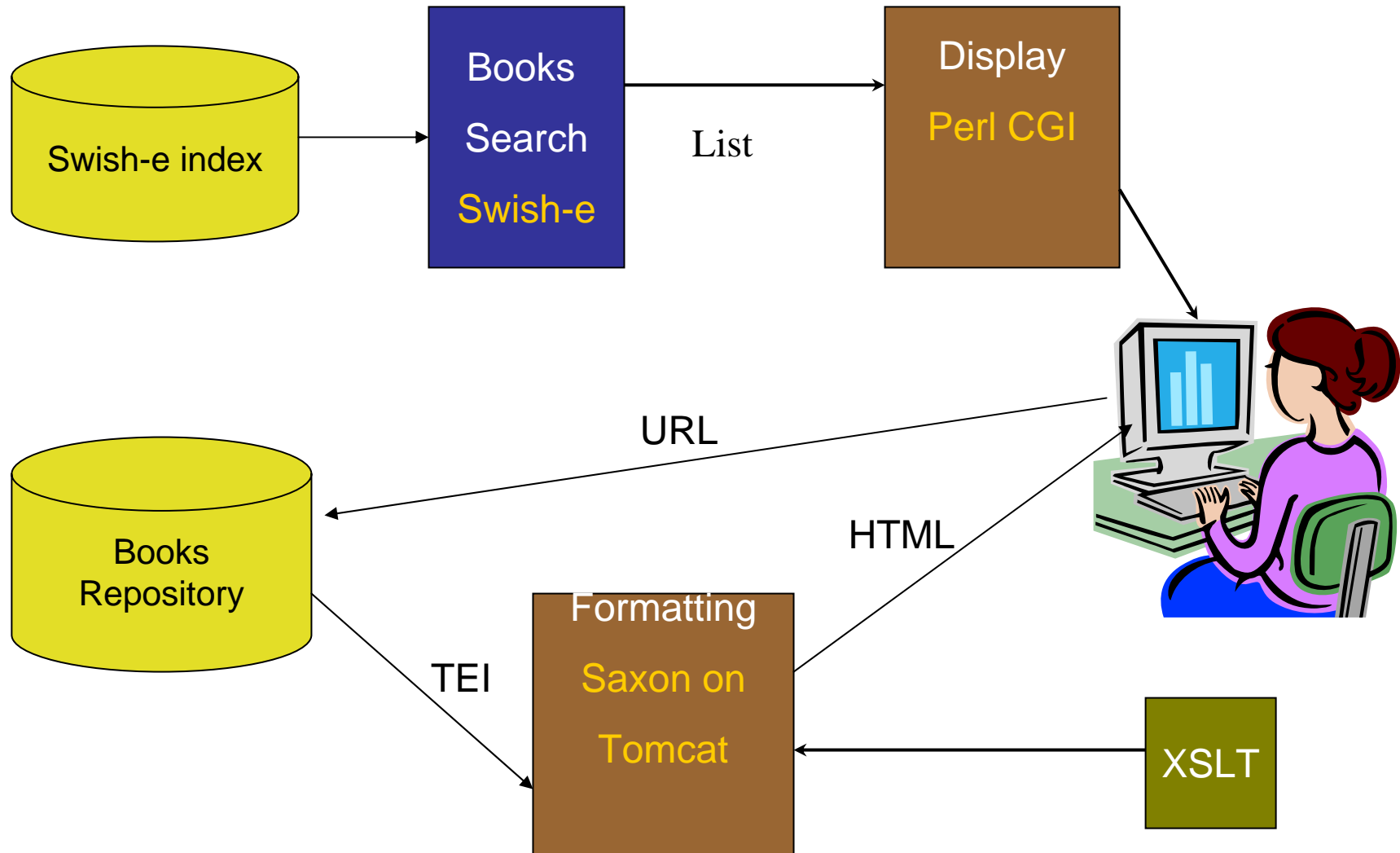


- ⌘ **TEI - Text Encoding Initiative**
the digital book encoding
- ⌘ **METS - Metadata Encoding & Transmission Standard**
represents each book's metadata, structure, identifier, and link
- ⌘ **MODS - Metadata Object Description Schema**
XML record derived from Marc metadata
- ⌘ **ARK- Archival Resource Key**
Persistent object identifier for METS->TEI record

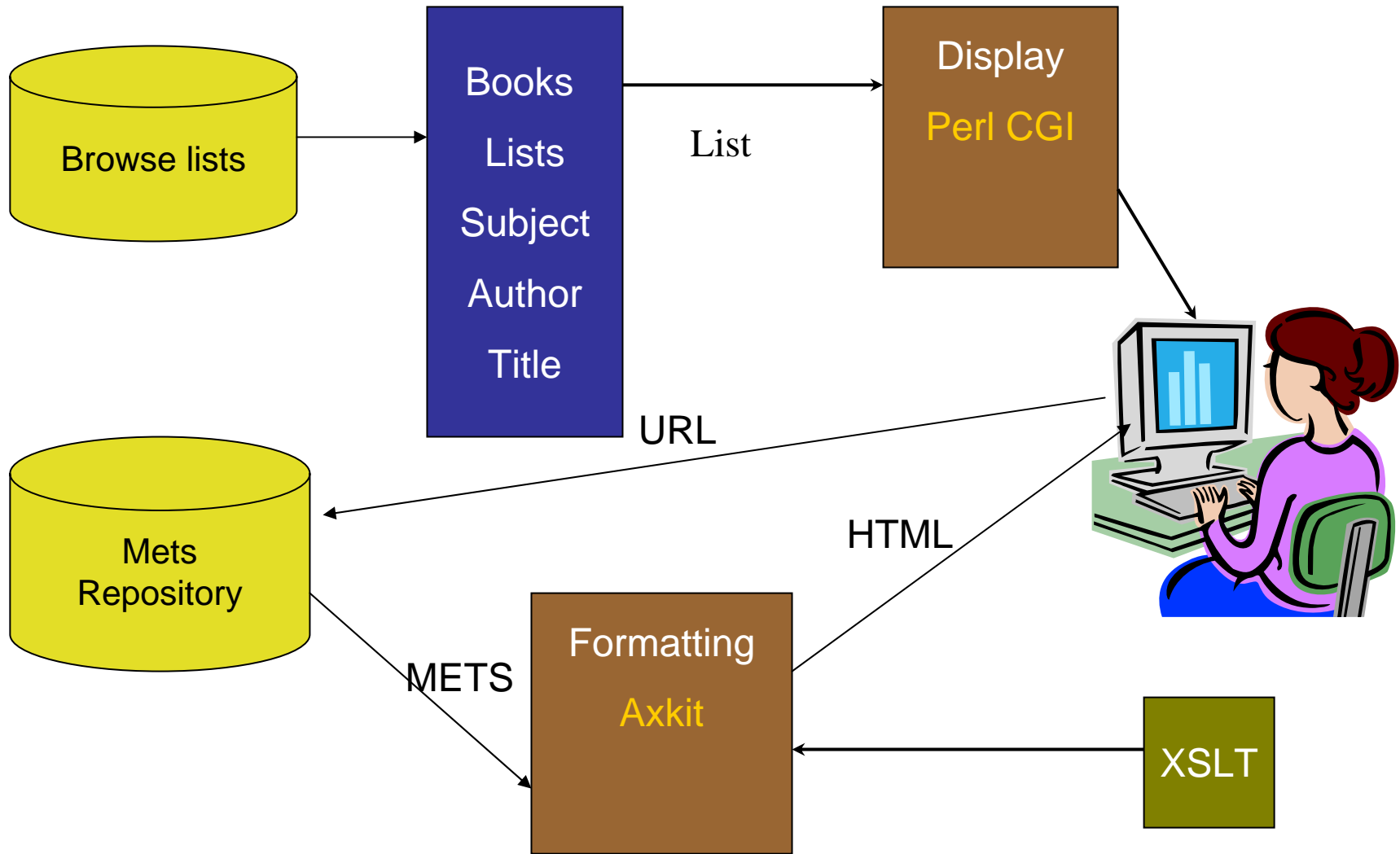
Searching METS



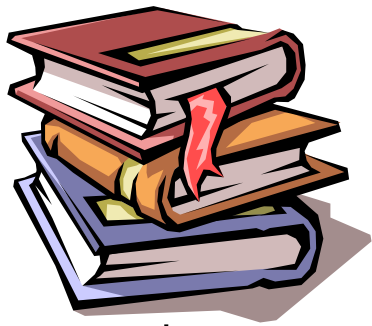
Searching Books



Browsing Books

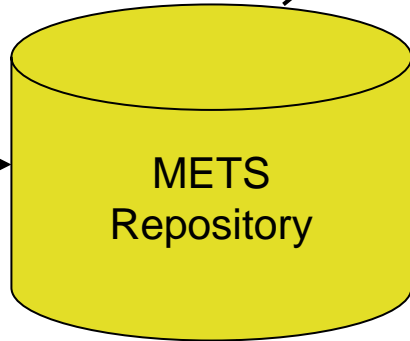


METS creation



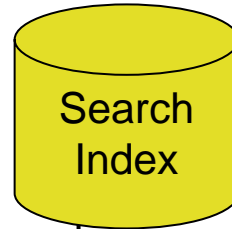
Structure

Record
Creation
Program



Field
Extraction
Program

Indexing
Software

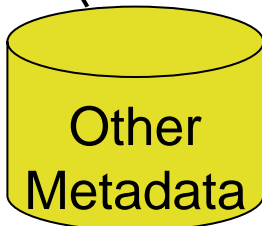
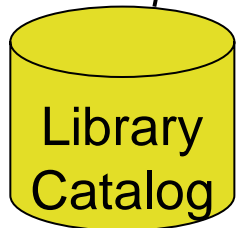


Results
in XML

XSLT

MODS record

DC record

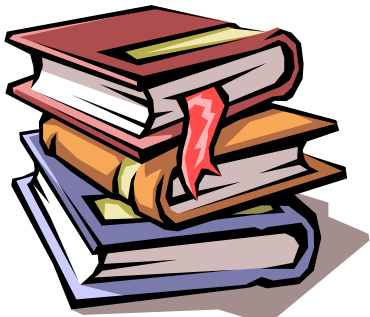


Object in XML

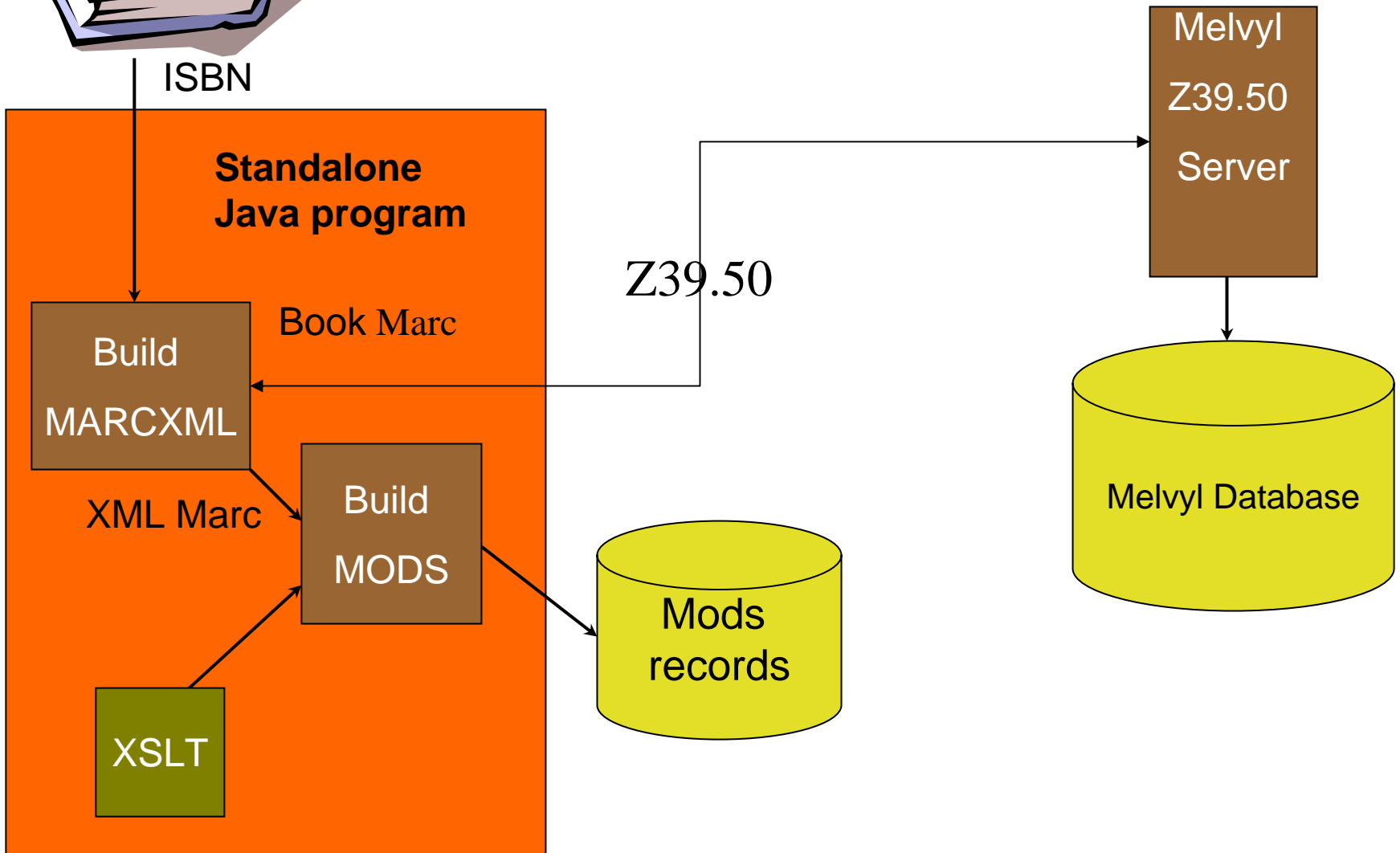
XSLT

Project
Profile

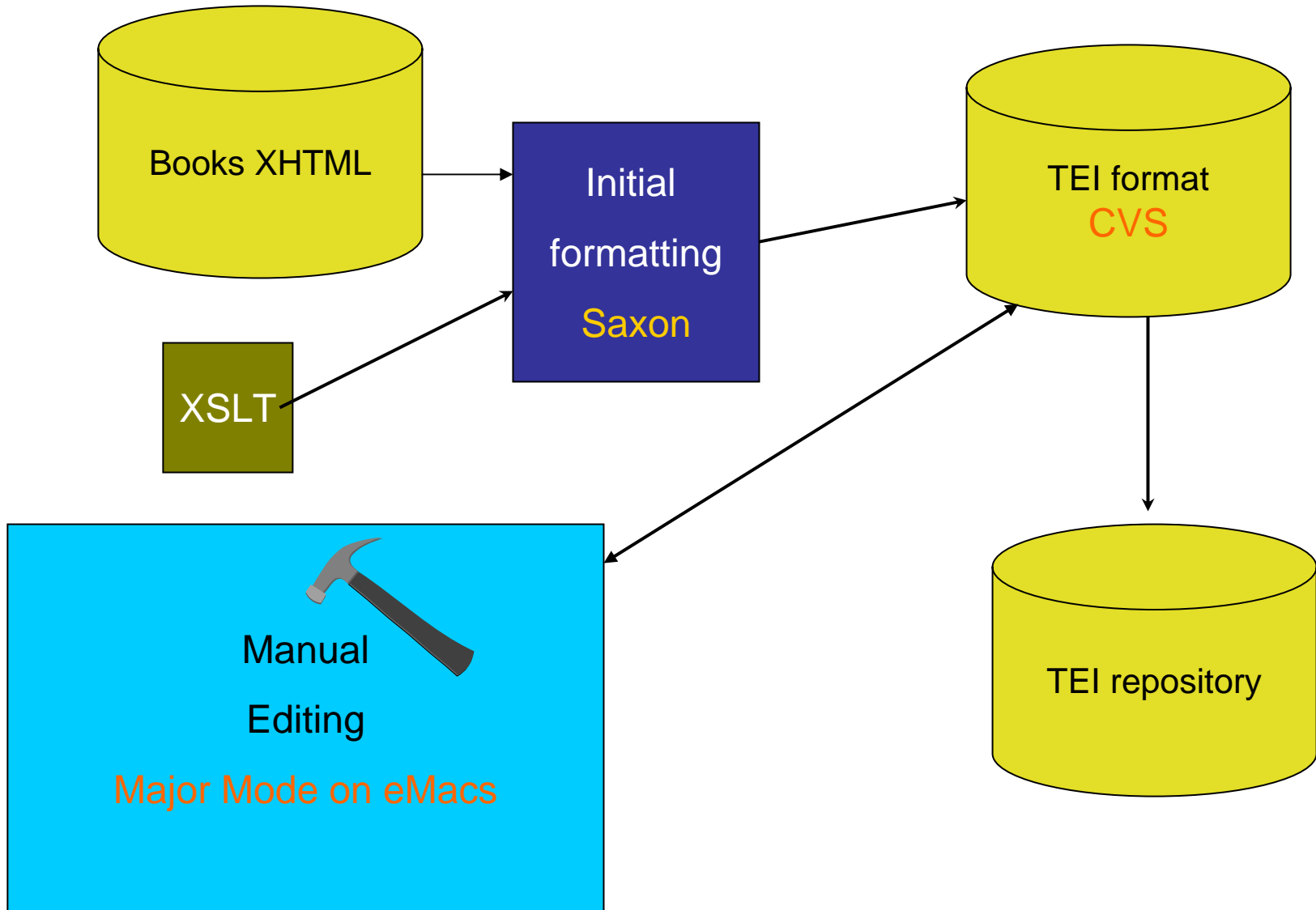




Building MODS



Books From XHTML to XML



Summary of Software by Functionality



⌘ XML Searching

- ⌘ Swish-e

⌘ XML editing

- ⌘ Major Mode plugin for eMac

⌘ Unix Programming

- ⌘ Perl

- ⌘ Java

Summary of Software by Functionality



⌘ Serving dynamic web pages:

- ☑ Tomcat

- ☑ apache


⌘ XML to HTML conversion using XSLT

- ☑ A Toolkit with Gnome (modperl and C)

- ☑ Saxon (java based)

- ☑ Xalan (java based)

Summary of Technologies



- ⌘ Axkit - run from Apache using Gnome's libxmlt and libxml2 - formatting xml
- ⌘ tomcat 4.0 running Saxon - formatting xml
- ⌘ swish-e 2.1- searching extracted xml
- ⌘ Apache 1.3 - web server
- ⌘ Major Mode plugin for eMacs