The Interactions of Emerging Gather/Create/Share End-User Tools with Digital Libraries

DLF Fall Forum October 26, 2004

Raymond Yee, Technology Architect

Goal of the Talk

- Problem:
    - the quality, quantity, and diversity of networked scholarly information is growing quickly
    - end-users tools to access and manage this bewildering array of information have been rapidly evolving
- This talk:
    - summarizes the range of current strategies and tools that enable users to effectively "gather, create, and share" digital information

Support References

- http://raymondyee.net/wiki/DlfOrg_2fFallForum2004_2fMyTalk
- Or
- **http://tinyurl.com/5h3hw**

UCB IU Mission

The Interactive University uses the Internet to democratize the content and community of the Berkeley campus. We seek to make the campus's and the nation's extraordinary digital resources far more usable for both higher education and the public, especially K-12 schools.

Vision: Our vision is that UC Berkeley be a national leader in the use of information technology for opening resources to the public.

Fundamental to the IU model is to find ways to open up the content and knowledge of the campus so that it adds value both to the campus community and to the public.
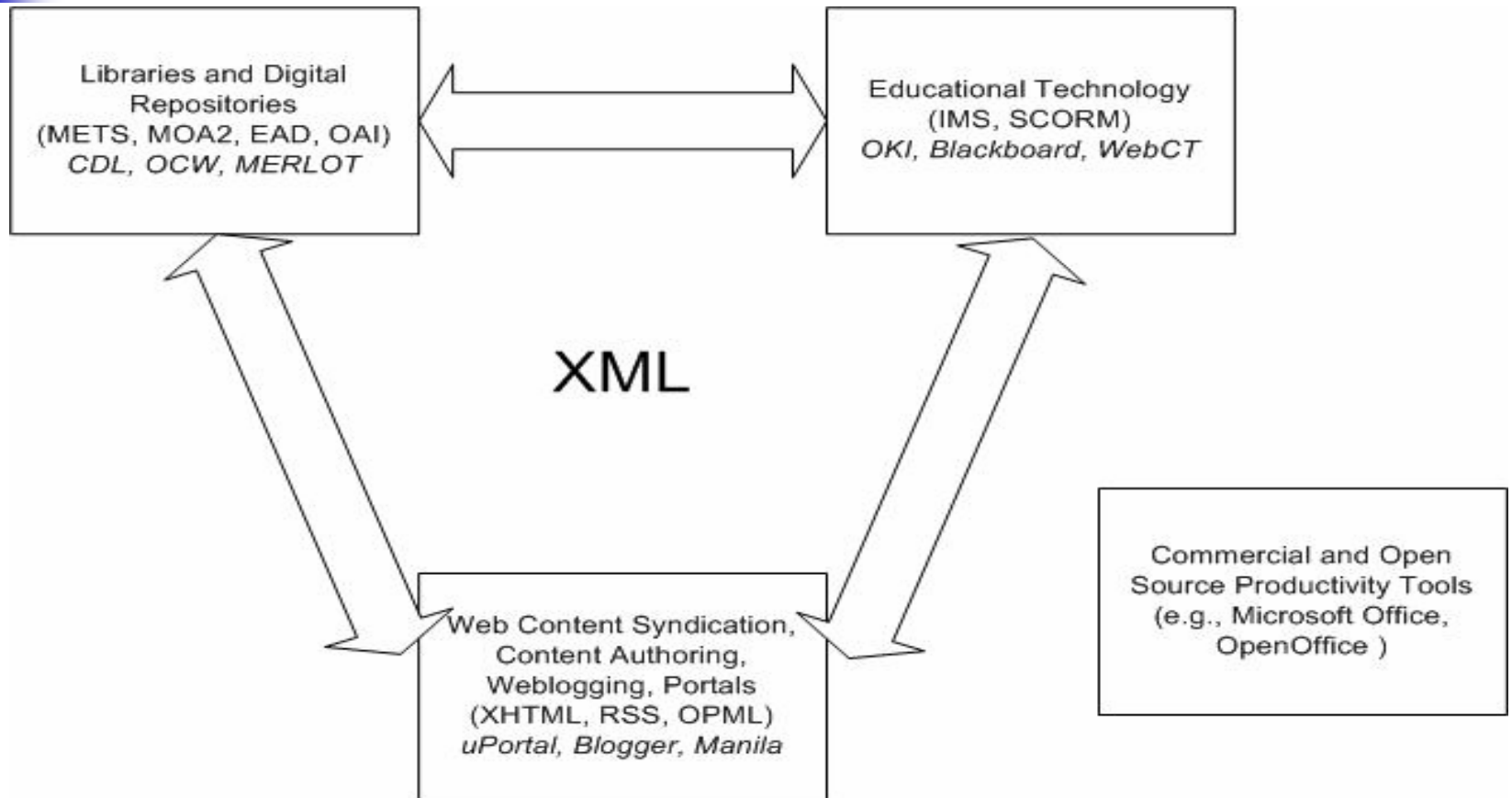
http://iu.berkeley.edu

Some "idealities" for us

- frictionless movement of content and knowledge and data

- ability to join content and services arbitrarily and easily

- tied into this is the notion of maximizing reuse, enabling bricolage authoring

- four domains of interaction

Four Intersecting Domains

Libraries and Digital Repositories (METS, MOA2, EAD, OAI) CDL, OCW, MERLOT

Educational Technology (IMS, SCORM) OKI, Blackboard, WebCT

XML

Web Content Syndication, Content Authoring, Weblogging, Portals (XHTML, RSS, OPML) uPortal, Blogger, Manila

Commercial and Open Source Productivity Tools (e.g., Microsoft Office, OpenOffice)

From music …

"*Rip. Mix. Burn.* Apple, of course, wants to sell computers. Yet its ad touches an ideal that runs very deep in our history. For the technology that they (and of course others) sell could enable this generation to do with our culture what generations have done from the very beginning of human society: to take what is our culture; to "rip" it - meaning to copy it; to "mix" it - meaning to reform it however the user wants; and finally, and most important, "burn" it - to publish it in a way that others can see and hear. Digital technology could enable an extraordinary range of ordinary people to become part of a creative process." *Lawrence Lessig, The Future of Ideas*

Gather. *Create.* *Share.*

Sources   **Scholar's Box Tool and Products**   **Destinations**

Digital Libraries

Museums

Learning Object
Repositories

Metadata
Harvesting/
Metasearch

WWW

Materials from
Personal
collections

→

Scholar's Box

Personal and
Themed Collections

PowerPoint

Web-based
Exhibition

Reading and
Resource
Lists

Flash X

Small Learning
Objects

Endnote, etc.

Word and OO Text:
Object-Embedded
Narratives

Weblogs / RSS

METs + IMS
Objects

Other Scholarly
Formats

→

IU Digital Library

Digital Libraries
and Museums
(e.g., CDL)

Learning Object
Repositories

Authoring Tools:
- Powerpoint
- Endnote
- Weblogs

LMS
Environments
- Sakai

Users Seem to Want this also

- See studies by D. Harley *et al*. <u>Digital Resource Study: Conclusions and Next Steps (first year report)</u>
- C. Borgman, <u>Creating individual spaces for innovation</u>
- W. Brockman *et al*. <u>Scholarly Work in the Humanities and the Evolving Information Environment</u>
- Notes on <u>my wiki</u>: **http://tinyurl.com/5ysv9**

Types of Tools to be Examined

- next generation web browser technology (e.g., Mozilla FireFox and its extensions);
- personal information managers such as Chandler,
- web-services enabled- and XML-aware office suites (such as Microsoft Office 2003 and OpenOffice.org);
- academic projects such as the **Scholar's Box**,
- high profile open source "Collaboration and Learning Environment (CLE) software" such as Sakai;
- evolving next generation operating systems, such Microsoft Longhorn
- Metasearch systems

The Scholar's Box approach

- We can **envision** possibilities for new information environments for scholars – with a focus on the scholar's point of view

- We **prototype** these possibilities and **test** ideas using Scholar's Box "in the small" as a **sandbox** to see how new flow of content and commentary can happen

- We **foster** functionality, **partnering** with others to **scale up** and **institutionalize**
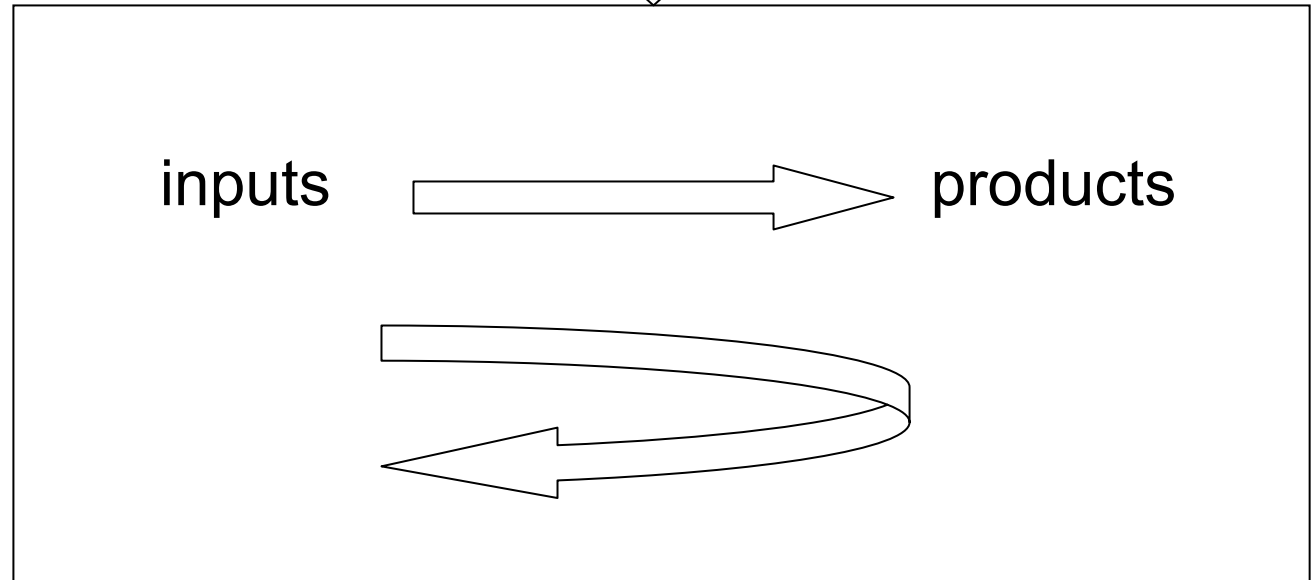
Scholar's Box Overview

Interactive University Project, UC Berkeley

A prototype tool for scholars to *gather*, *create*, *and share* digital content and documents. http://raymondyee.net/wiki/ScholarsBox

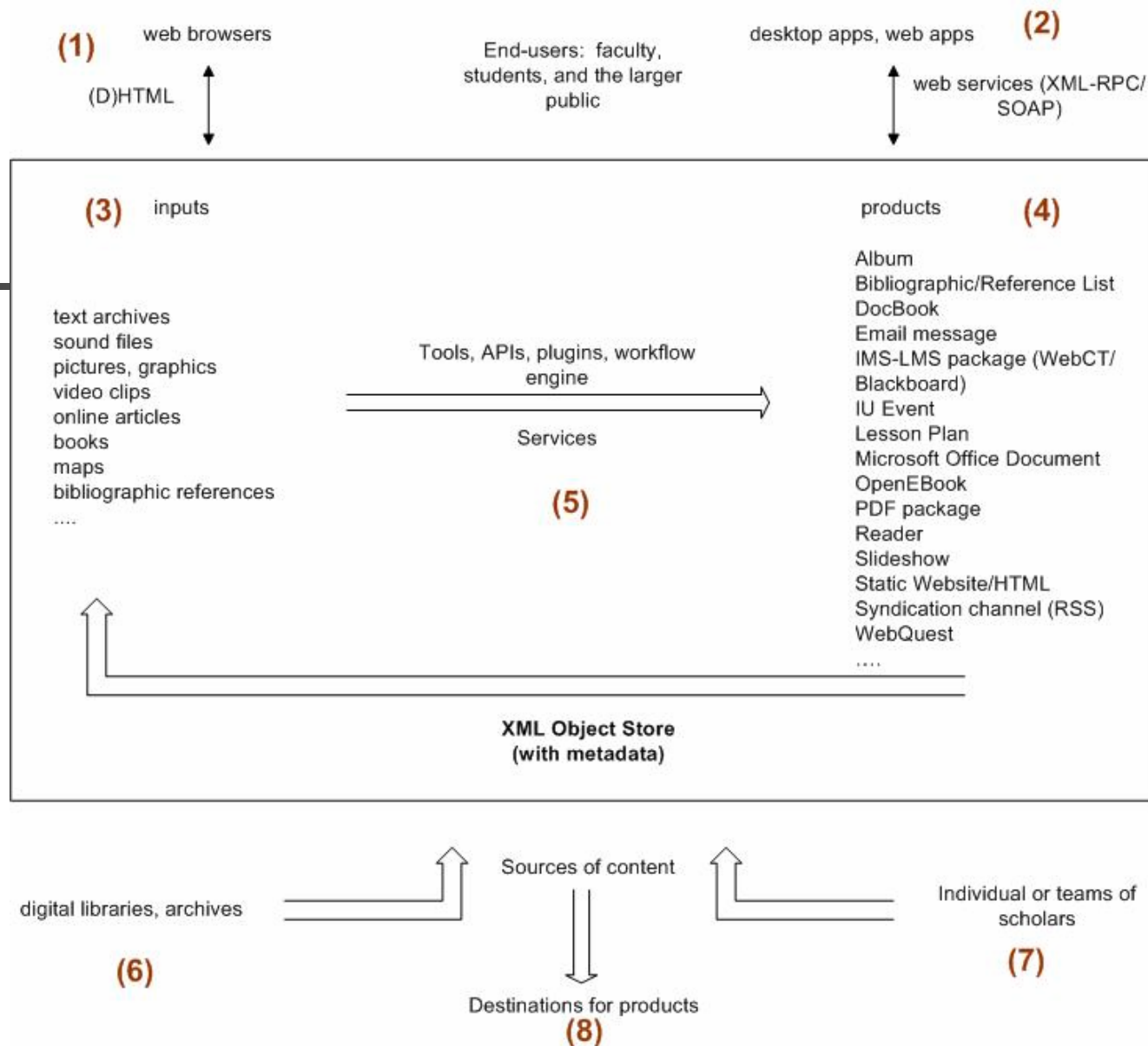|	
Gather	From California Digital Library, amazon.com, google.com, NSDL, RSS feeds, METS (digital library), WWW, and the local file system.
Create	Data and metadata gathered, annotated, and organized into personal collections via drag and drop
Share	IMS-CP, OpenOffice.org Presentation or Text document, PDF, HTML, a METS document, a set of Endnote references, Chandler Parcel, or sent to a weblog via blogger api

Technology Architecture

Interfaces (DHTML/web services)

inputs → products

Data sources

repositories   personal documents

(1) web browsers

(D)HTML

End-users: faculty, students, and the larger public

desktop apps, web apps **(2)**

web services (XML-RPC/ SOAP)

(3) inputs

text archives
sound files
pictures, graphics
video clips
online articles
books
maps
bibliographic references
....

Tools, APIs, plugins, workflow engine

Services

(5)

products **(4)**

Album
Bibliographic/Reference List
DocBook
Email message
IMS-LMS package (WebCT/ Blackboard)
IU Event
Lesson Plan
Microsoft Office Document
OpenEBook
PDF package
Reader
Slideshow
Static Website/HTML
Syndication channel (RSS)
WebQuest
....

**XML Object Store
(with metadata)**

digital libraries, archives

Sources of content

Individual or teams of scholars

(6)

(7)

Destinations for products
(8)

1 (D)HTML browser-based user-interface 2. Desktop applications, web applications via web services 3. Input materials
4. Products from the services 5. Technical infrastructure: XML object stores, metadata handling, tools, APIs, plugins,
extensible workflow engine 6. Interface to digital libraries and archives 7. Content producer tools 8. Destinations for
products

Next generation Web Browsers: Rich clients

- The web browser is a dominant human interface to digital libraries
- Renaissance of the web browser, possibly moving beyond Microsoft IE hegemony ("rich clients")
- XAML, Macromedia's Flex/MXML
- Opens the possibility of "augmented browsing"

Mozilla Firefox extensions

- What is so exciting here?
  - XML, web-services based frameworks (open specifications)
  - Really simple but powerful extension mechanisms (high extensibility, easy for folks to install, transparency of extensions)
  - Can implement gather/create/share functionality internally or interacting with other tools
- Demo Paracite-based extension

del.icio.us, unalog, furl.net, flickr: bibliographic and images

- Social bookmarking systems
- Unalog (Dan Chudnov) is academically oriented – and has groups
- Look at the tagging as a fascinating system of social metadata
- A way for the "community" to gather and make sense of stuff on the web
- But how to move collections around?

Chandler

- Open source PIM from OSAF (Mitch Kapor)
- CSG and Mellon funding to get higher education functionality
- Rich repository infrastructure – ultimately. Scholars will want to organize disparate types of info – hope to connect to SB
- Heterogeneous database might help; Microsoft Longhorn eventually
- Update

OO.o and Microsoft Office 2003

- Connecting writing environments directly with data sources
- XML-aware
- Research pane
- Bibliographic project
- Embedding provenance info?
- Universal canvas?

Bibliographic systems

- Bruce D'Arcus' effort to push MODS as a standard bibliographic representation

Blogs and Wikis: decentralization of publishing

- weblogs and wikis -- when we started in 2000 with weblogs, it wasn't really having much of an impact. but now, pretty mainstream
- http://www.educause.edu/apps/er/erm04/erm045.asp

Microsoft Longhorn

- Great ambition to create next generation OS with data, metadata, XML, service oriented architectures all baked in
- Will it work?
- What will it do to the competition?
- What should we do about it?
- What is happening on Linux and Mac OS X?

Metasearch

- How is the library community doing via metasearch?

Common observations about tools

- A lot of these different services but few (no?) common APIs and data representations
- Trying to make everyone conform to the same data framework is futile, I think

Discussion

- Let's talk about "So what?" and maybe…."What next?"
- Personal Collections BOF this afternoon!

Acknowledgements

- Special thanks to:
    - David Greenbaum (Director, IU)
    - Tom Schirmer (Software Developer, Interactive University Project, UC Berkeley) for transforming the Scholar's Box from skeletal prototype to functional software
    - Chris Ashley (Manager, New Program Development, Interactive University Project, UC Berkeley) for seminal ideas that ultimately gave rise to the Scholar's Box, input on user interaction, and helpful feedback on this talk

Talk to us!

- Raymond Yee (yee@berkeley.edu, http://iu.berkeley.edu/rdhyee, http://raymondyee.net/wiki/ScholarsBox)

- **http://tinyurl.com/5h3hw**

Acknowledgements

- Special thanks to:
    - David Greenbaum (Director of IU Project)
    - Tom Schirmer
    - Chris Ashley
    - California Digital Library
    - Thank you to funders: CDL, Mellon Foundation, NSF, Department of Commerce, Department of Education

Motivation for Scholar's Box

Scholarship: In creating and presenting ideas, scholars build custom collections from which they create their desired product. Gathering, manipulating, organizing, annotating, and sharing personal collections of cultural objects is a core activity.

Teaching: Every teacher possesses a box full of hand-collected teaching materials, "stuff" gathered over the years from various, often ephemeral sources. These primary source materials may include pictures, maps, news articles, short stories, speeches, graphs, and charts.

The Need for Better Tools

- Users can **look** but not easily manipulate digital content (texts, images, video) -- **data silos**

- We want malleable, reusable pieces that work together regardless of **data type** or **data origin**.

- How do we get content and ideas to **flow freely** among sources, tools, and destinations?

Current functionality in SB (some not demoed)

- Gathering via search
- Sources: CDL, melvyl, amazon.com, google.com, CalPhotos, NSDL, RSS, METS files, desktop (local filesystem)
- Data and metadata gathered
- Building personal collections via drag and drop
- Organizing subcollections
- Annotations of items and collections
- Creates OO.o, PDF, HTML, RSS, IMS-CP, METS, Endnote references, weblog posts