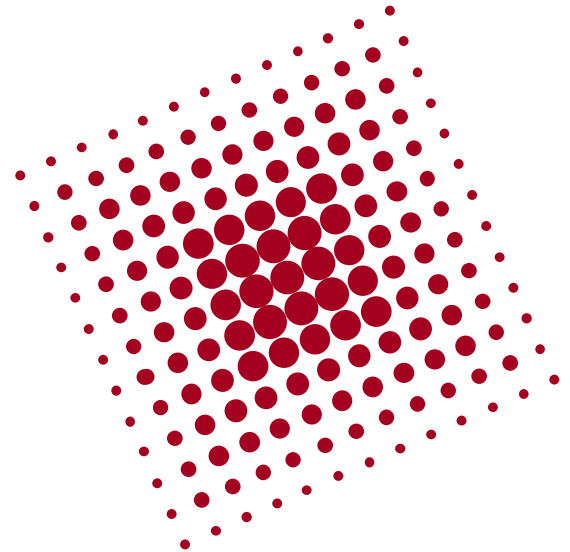# *CNRI Handle System*®

WIPO Int'l Conference
On Ecommerce and
Intellectual Property

The Handle System®

The Handle System® is a general purpose global name service enabling identifier resolution over the Internet.

Handles are managed identifiers for digital objects, with which multiple pieces of data are stored.

Identifiers

- ◆ Many kinds of identifiers have been in use in the "physical" world for a long time (ISBN, SICI, UPC).

- ◆ They work well. Organizations have made significant investments in them and want to keep using them.

- ◆ Both these legacy identifiers and new ones are now being put into use for digital objects.

- ◆ Identifiers for digital objects can be used to link to many kinds of resources and services (full texts, metadata, purchase options, etc.) as part of electronic "Identifier Systems".

Good Identifiers are...

◆ Unique - No possibility of duplicates.

◆ Persistent - Storage should facilitate accurate administration no matter how many times an object moves on a network.

◆ Actionable - Selecting an identifier should return one or more objects or services.

◆ Secure - Creating or modifying identifiers should require authentication.

◆ Simple to maintain - Accommodates existing schemes and internal work flows.

◆ Widely supported - A namespace based on Unicode 2.0 will enable international use.

Good Identifier <u>Systems</u>...

- ◆ Resolve identifiers - Locate associated objects.
- ◆ Accommodate existing schemes - Incorporate any string.
- ◆ Scale up - Add as many computers as needed.
- ◆ Are fast - Optimize speed for query results returned.
- ◆ Are secure - Tools for client and server authentication must be built in.

- ◆ Have strong support - System management must be committed to maintaining support.
- ◆ Store multiple types of data for any given identifier - The ability to return more than "one object or piece of data per identifier" is key to providing great customer service.

What kind of "Identifier System" is the Handle System?

A Handle *can be any string...*

4263537/4001

10.1061/(ASCE)1052-3928(1999)125:1(12)

1895.22/1011

100.1/dvlhome

loc.gov/loc.award/iencurt.cd07035

10.1045/may99-paskin

◆ Handles accommodate legacy numbering systems.

◆ They can be any Unicode v2.0 character.

◆ No restriction is imposed on length.

The Handle Syntax is...

$$1\ 8\ 9\ 5\ .\ 2\ 2\ /\ 1\ 0\ 1\ 1$$

$$\underline{\text{Naming Authority}} \Big/ \underline{\text{Local Name}}$$

- The "Naming Authority" identifies the Local Handle Service.
- The "Local Name" can be any unique string.
- Duplicates are prevented.

Handle Data...

- The Handle System stores both "public type" user data (URLs, email addresses) and "administrative type" system data in each handle record.
- There are defined data types for URLs, email, etc.; there can be multiples of any data type; handle administrators can create their own data types.
- Predefined "administrative type" data, such as HS_ADMIN or HS_SITE are used to store the identity of administrators, service location, and other information used by software clients.

Handle	Data Type	Handle Data
1895.22/1011	URL	http://acme.com/....
	URL	http://a-books.com/....
	EMAIL	Joe@acme.com
	HS_ADMIN	acme.admin/jsmith
	HS_SERV	1001110011110
	MY_HANDLE_TYPE	0010101011000

Persistence...

- Over time, the location (URL) of a digital object on the Web will most likely change (new server or even new owner) and known links to that object will break.

- If an object is identified on the Web by its handle instead of its URL, when the URL changes, the handle data can be updated with the new location, leaving the handle unchanged and known links unbroken.

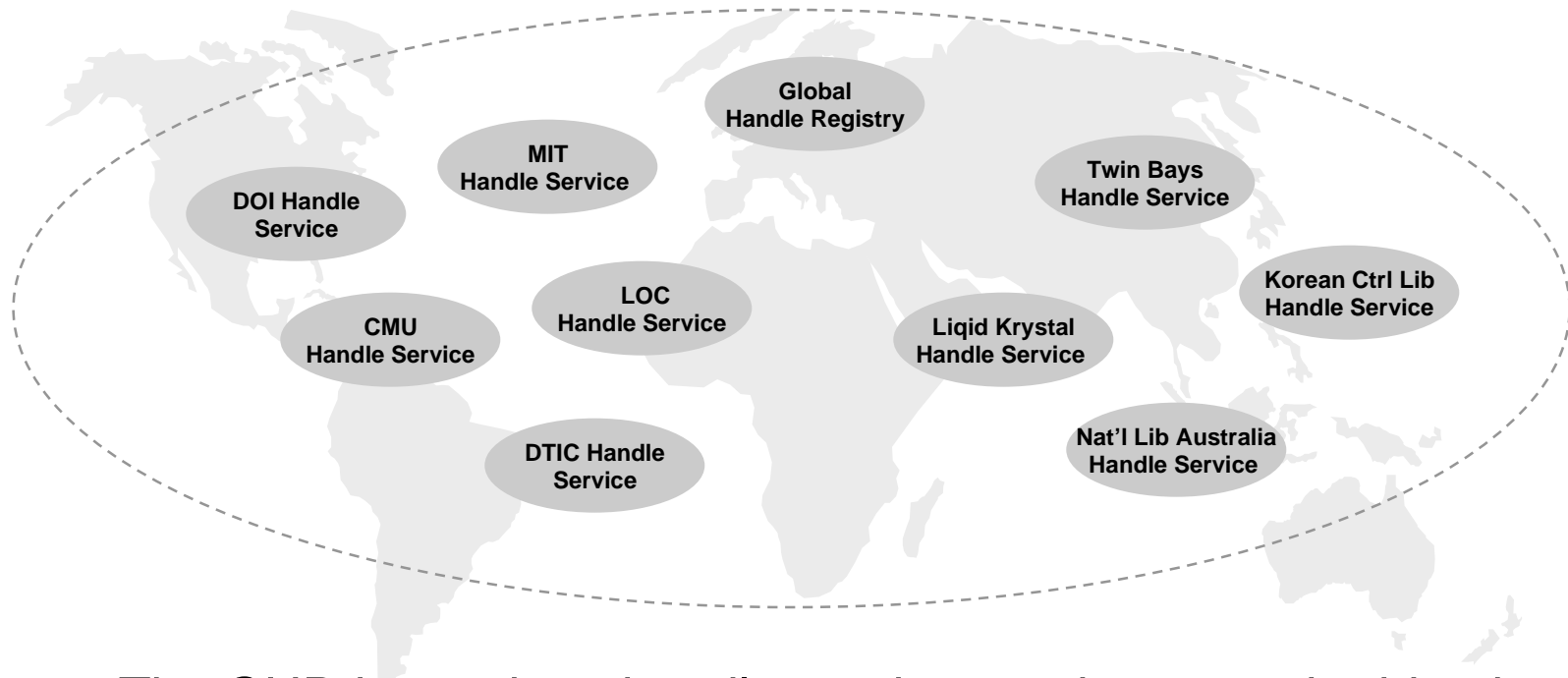| Handle	Data Type	Handle Data
1895.22/1011	URL	~~http://acme.com/....~~ http://newco.com/....
	URL	http://a-books.com/....
	EMAIL	Joe@acme.com
	HS_ADMIN	acme.admin/jsmith
	HS_SERV	1001110011110
	MY_HANDLE_TYPE	0010101011000

Granularity

- ◆ Physical objects typically had one identifier, i.e., each book had one ISBN. There was no way to separate out the components, and so no reason to identify them.

- ◆ Digital objects such as eBooks can contain other digital objects, e.g., chapters from other books, illustrations, audio segments, etc., and each of these can have its own handle.

- ◆ Tracking the individual pieces will help in digital rights management.
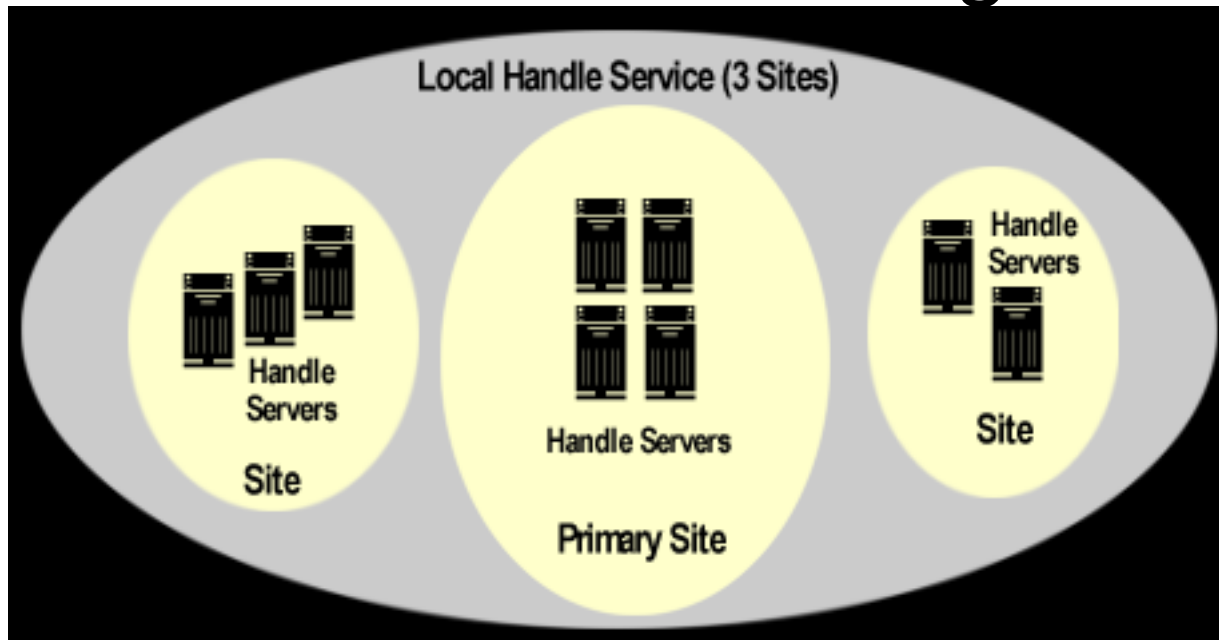
Setting up a *Local Handle Service...*

- ◆ Download the software from http://www.handle.net.

- ◆ Follow the instructions in the installation script.

- ◆ Send your "site bundle", containing the IP address of your server and your administrator information, to the Global Handle Registry (GHR) administrator.

The Global Handle Registry

Global
Handle Registry

MIT
Handle Service

Twin Bays
Handle Service

DOI Handle
Service

Korean Ctrl Lib
Handle Service

LOC
Handle Service

CMU
Handle Service

Liqid Krystal
Handle Service

DTIC Handle
Service

Nat'l Lib Australia
Handle Service

◆ The GHR is a unique handle service used to store the identity and location of all local handle services (LHS), and tells a handle client which service to query to resolve a handle.

◆ All handle clients (for resolution or administration) know how to contact and query the GHR.

Local handle service *configuration...*

Local Handle Service (3 Sites)

Handle Servers

Site

Handle Servers

Primary Site

Handle Servers

Site

- ◆ A local handle service is made up of one (or more sites), which contain one (or more) servers.

- ◆ Handle storage is distributed across the servers within a site.

- ◆ Multiple sites/servers provide replication, sites and servers can be distributed to any location, and any number of sites or servers can be added as needed.

Handle Clients

Administration

Use the Java™ Handle Client Tool provided in the distribution for creating or updating handles one-at-a-time or via a batch.

or

Develop your own administration client.
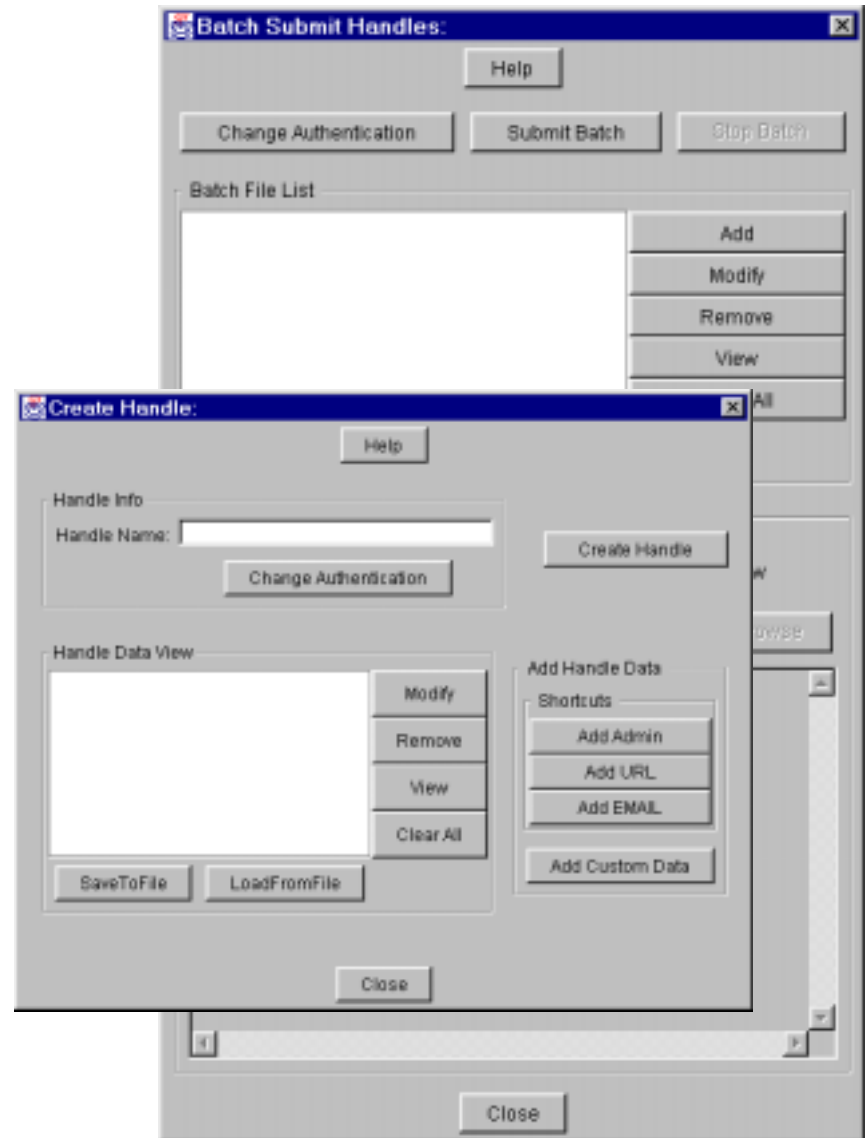
Handle Clients

Resolution

Download CNRI's web browser plug-in which enables browsers
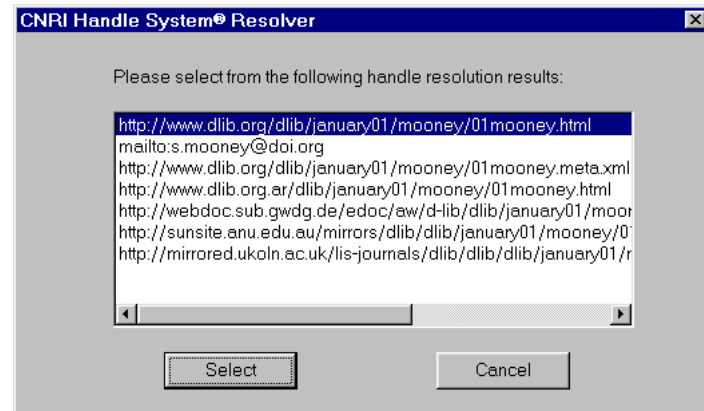to recognize the handle protocol.

        or

Append a handle to CNRI's
proxy server
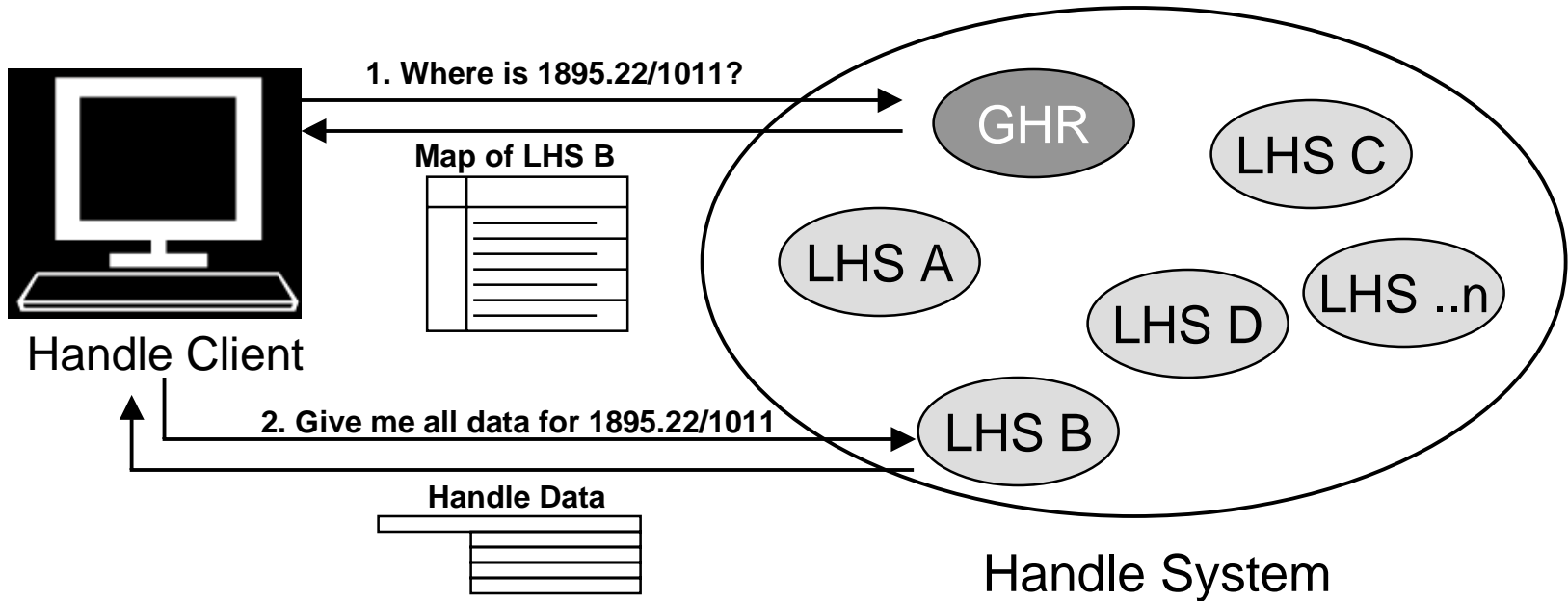(http://hdl.handle.net)
which understands both HTTP and HDL protocols.

        or

 Develop your own resolution client.

CNRI Handle System® Resolver

Please select from the following handle resolution results:

http://www.dlib.org/dlib/january01/mooney/01mooney.html
mailto:s.mooney@doi.org
http://www.dlib.org/dlib/january01/mooney/01mooney.meta.xml
http://www.dlib.org.ar/dlib/january01/mooney/01mooney.html
http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/january01/moon
http://sunsite.anu.edu.au/mirrors/dlib/dlib/january01/mooney/01
http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/january01/r

Select   Cancel

Handle Storage and System Scalabilty

- A "hashing mechanism" stores handles equally across the servers in a handle site. The "map" of the storage is called service information.

- When a new server is added to increase storage capacity, the handles are "remapped" across the servers.

- The GHR has the service information for every handle service. Admin and resolution clients know to query GHR for the service information, and read the information to get the address of the correct server.

- Handle resolution involves two steps no matter how large a service is -- to GHR to get service information, and a query directly to the responsible server.
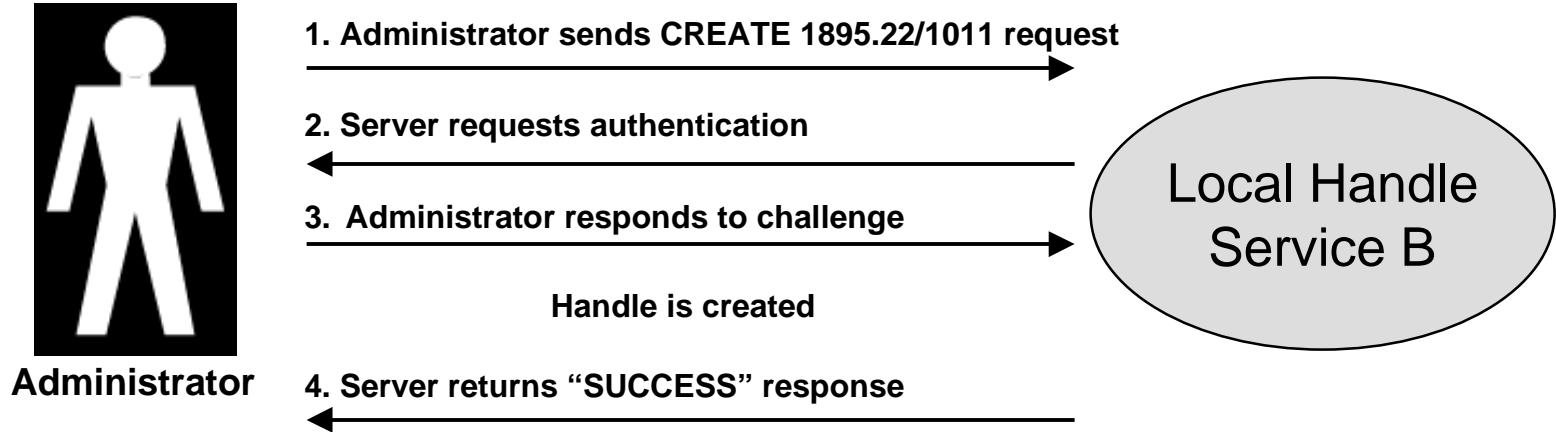
How to handles *resolve...*

1. Where is 1895.22/1011?

GHR

LHS C

Map of LHS B

LHS A

LHS D

LHS ..n

Handle Client

2. Give me all data for 1895.22/1011

LHS B

Handle Data

Handle System

Two steps to resolve a handle - -

- Client queries GHR: "Which Handle Service has 1895.22/1011?"
- GHR responds with a "map" showing the client which servers within the responsible LHS it can query for that handle .

Administering handles...

1. Administrator sends CREATE 1895.22/1011 request

2. Server requests authentication

3. Administrator responds to challenge

Handle is created

Administrator

4. Server returns "SUCCESS" response

Local Handle
Service B

- Administrators have "administrator handles" which identify them to the Handle System.

- Authentication is secret key or public/private key pair.

- Permissions determine who creates or updates a handle, but a handle can have multiple administrators.

Handle System *advantages...*

- **Security** -- The Handle System can provide authentication and data integrity services, depending on the client's request.

- **Speed** -- The current reference implementation of the Handle System is optimized for speed of resolution.

- **Scalability** -- The "hashing mechanism" by which handles are stored across sites enables unlimited scalability.

- **Services** -- The ability to associate multiple pieces of data with a handle -- location, copyright, metadata, price -- can help offer a variety of services to users.

Handle System Initiatives

- ◆ IDF (International DOI Foundation)
  - CrossRef (scholarly journal consortium)
  - Enpia (Korean content management technology firm)
  - Content Directions, Inc. (US content management technology firm)
- ◆ Library of Congress
- ◆ DTIC (Defense Technical Information Center)
- ◆ Other Interested Parties:
  - cIDF (Content ID Forum, based in Japan)
  - NTIS (National Technical Information Service)
  - NMPA (National Music Publishers' Association)
  - Numerous digital library research projects