



Linking to LCSH and LCC: Controlled Subject Headings and Classification Systems through the Web

Barbara Tillett

Libby Dechman and

Loche McLean
Library of Congress
Washington DC, USA

Meeting:

**149 —Bridging domains, communities and systems—
Classification and Indexing Section**

Abstract:

The Library of Congress is making its controlled vocabularies and classification system freely accessible on the Web. This presentation describes our services related to two of those controlled vocabularies and classification schemes: the Library of Congress Subject Headings and Library of Congress Classification. It also describes the work to link various language versions of LCSH together.

*The initial structure for this service uses SKOS, Simple Knowledge Organization System, which “Provides a model for expressing the basic structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, folksonomies, and other similar types of controlled vocabulary. ” —SKOS Primer
Additionally we are exploring RDF XML as another structure to use for presenting this data.*

The service can be found at <http://id.loc.gov> and the LCSH/SKOS is at <http://id.loc.gov/authorities>

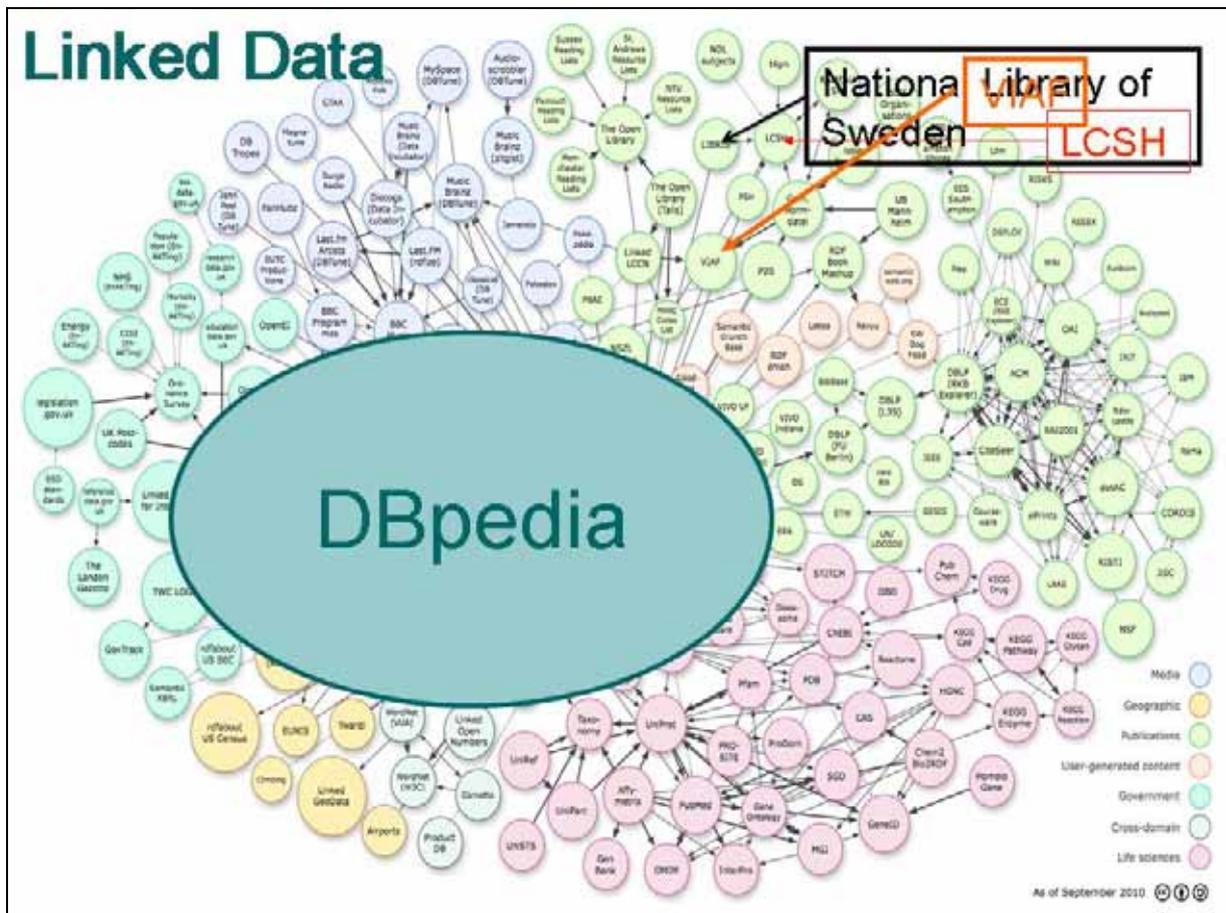


Linking to LCSH and LCC

Controlled Subject Headings and Classification Systems through the Web

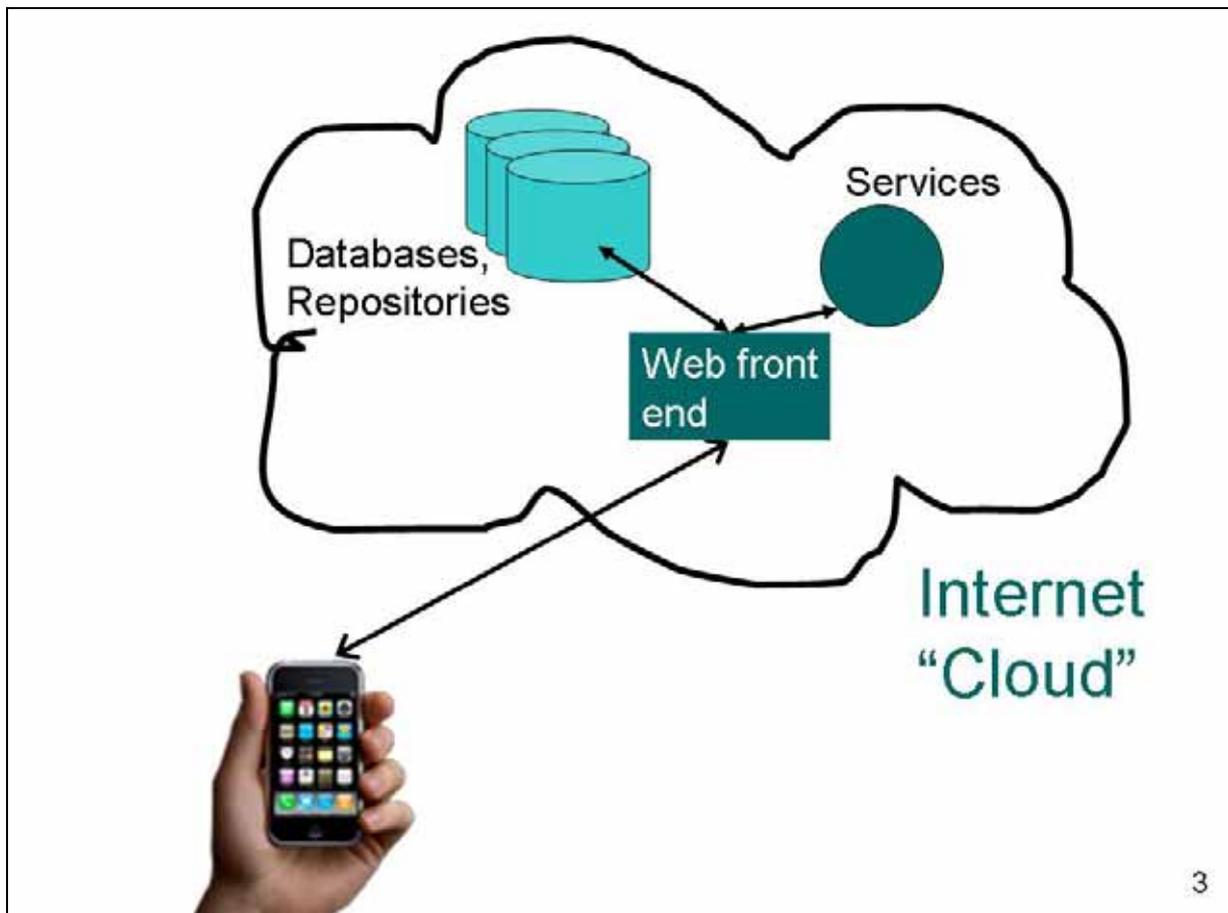
Presentation by
Barbara Tillett, Libby Dechman, and Loche McLean
Library of Congress Policy & Standards Division
For IFLA 2011
Puerto Rico

This brief communication presentation shares news of our work at the Library of Congress to make our *Library of Congress Subject Headings* and other controlled vocabularies freely available on the Web. This is part of our experimentation into the world of Linked Data and the Semantic Web. We offer our controlled vocabularies as an essential building block for the Semantic Web. You may find more about the value and benefits of controlled vocabularies for the Semantic Web from Aida Slavic in the United Kingdom, such as her 2008 presentation to CILIP (Chartered Institute of Library and Information Professionals).

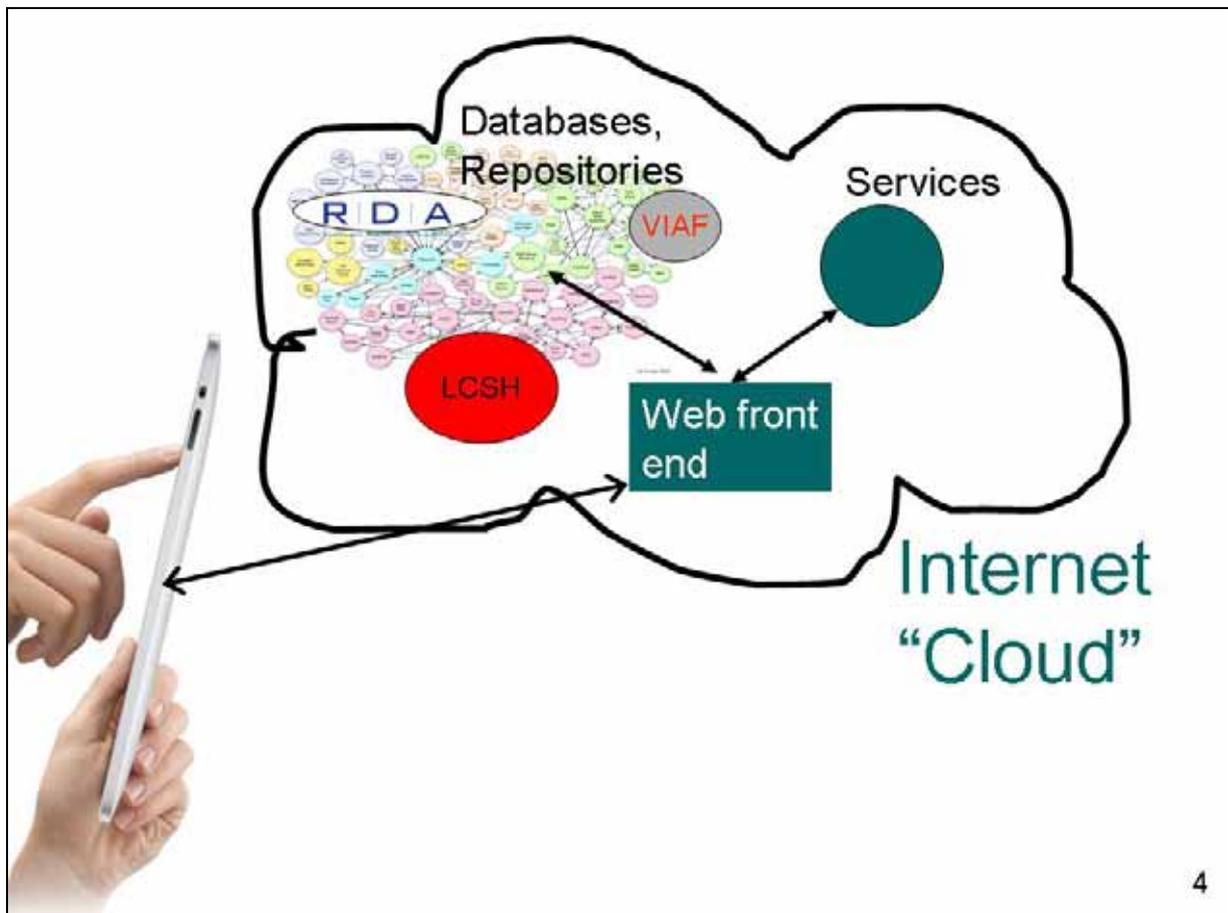


“Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. <http://lod-cloud.net/>” Also see: <http://richard.cyganiak.de/2007/10/lod/>
(Modified notes from Karen Smith-Yoshimura)

DBpedia is a community effort to extract structured information from Wikipedia and to make this information available on the Web. DBpedia allows you to ask sophisticated queries against Wikipedia, and to link other data sets on the Web to Wikipedia data. The idea is that this will make it easier for the amazing amount of information in Wikipedia to be used in new and interesting ways, and that it might inspire new mechanisms for navigating, linking and improving the encyclopedia itself as well as the linked data, including connections to library data – for example, the database at the National Library of Sweden is linked – also the VIAF And LCSH are part of this linked data environment. The DBpedia knowledge base describes more than 3 million things, in at least 91 different languages.



Information systems and content, like digital resources, images on Flickr, and our bibliographic and authority data are now part of the Internet cloud computing environment. The elements that describe our resources are available to libraries and users everywhere in the world – not just on an institution’s computer, but shared and available to everyone through the Internet. The data comes from publishers, from the creators of the resources, from trusted libraries and other institutions, and can be augmented by further descriptive information from anyone who wants to help. All of the data about information resources in our bibliographic universe of things we have in our library collections and in archives will be accessible by any user anywhere at anytime.



Bibliographic data and digital resources are on the Web now and we've started adding the controlled vocabularies to help identify resources – such as the controlled values for naming the types of content (like sound, text, still images, and so on), and other elements in RDA: Resource Description and Access that have controlled lists of values. Some of the RDA vocabularies are already being registered on the Web and can be used to present displays and show pathways to related resources.

So today, I'd like to show you one of these services or building blocks for this linked Web environment that the Library of Congress is involved in with other libraries: namely the LCSH – Library of Congress Subject Headings.

The other two we are involved with are VIAF (Virtual International Authority File), and the controlled vocabularies from the new cataloging code, RDA: Resource Description and Access.



SKOS

- Simple Knowledge Organization System
 - “Provides a model for expressing the basic structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, folksonomies, and other similar types of controlled vocabulary”—*SKOS Primer*

5

The Library of Congress has posted our subject headings – the Library of Congress Subject Headings on the Web in SKOS format – Simple Knowledge Organization System. We are also in the process of posting our Library of Congress Classification system on the Web as well.

SKOS is “a model for expressing the basic structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, folksonomies, and other similar types of controlled vocabulary” (from the SKOS Primer).

SKOS

- Based on the Resource Description Framework (RDF)
 - Resources can be exchanged between software applications and published on the Web
 - Interconnects data on the Web, helping create the Semantic Web

SKOS is based on the Resource Description Framework (RDF) from the World Wide Web Consortium (W3C). The framework provides a structure for exchanging resources between software applications and publishing resources on the Web and is one of the key components in the development of the Semantic Web of linked data.



id.loc.gov/authorities

- “Authorities & Vocabularies” from the Library of Congress
- Intent: To provide human and programmatic access to commonly found standards and vocabularies developed by the Library of Congress

So the Library of Congress used SKOS for several vocabularies we want to share with the global information community. We established the website: id.loc.gov/authorities for the purpose of providing human and computer access to the standards and vocabularies developed at the Library of Congress

“Authorities & Vocabularies”

- *LCSH* is the first offering
 - Subject headings
 - Genre/form headings
 - Children’s subject headings
 - Subdivision records
 - Validation records
- Provides links from *LCSH* headings to RAMEAU headings
 - Exploring Répertoire de vedettes-matière (RVM) and others

8

The Library of Congress Subject Headings were the first of our vocabularies to be posted. It includes nearly 500,000 subject headings, genre/form headings, the children’s subject headings, all of our subdivision records and what we call our “validation records” which are LCSH heading strings that formerly were only in our bibliographic records but now have authority records created for them.

As part of our initial test of LCSH in SKOS we collaborated with colleagues in Europe and included a snapshot of the RAMEAU headings from the Bibliothèque nationale de France. We are also exploring adding the terms from the Répertoire de vedettes-matière (RVM) from the University of Laval in Canada (for French-Canadian terms), as well as other subject heading lists in other languages based on LCSH, such as the Spanish version from the Biblioteca nacional de España, and the Latin American Spanish version from the National Library of Chile, as well as an Arabic version from the Bibliotheca Alexandrina in Egypt.

“Authorities & Vocabularies”

○ Also includes:

- *Thesaurus for Graphic Materials (TGM)*
- MARC geographic area codes
- MARC language codes
- MARC relator codes
- Preservation Events ... etc.

9

Some of the other vocabularies on the site include the Thesaurus for Graphic Materials, MARC geographic area codes, and other MARC code lists.

“Authorities & Vocabularies”

○ Benefits

- Servers can download entire controlled vocabularies and the values within them, in multiple formats
- Available for free on the Web

10

We see some major benefits in making our controlled vocabularies freely available on the web. For example, computer systems and servers around the world can take advantage of these large vocabularies and can load them and use them in multiple formats. And again – these vocabularies are freely available on the Web.

“Authorities & Vocabularies”

- Human end-users can
 - **Search** and **view** individual headings and data elements
 - Details of the record
 - Visualization
 - **Suggest** additions, changes

11

In addition the LCSH/SKOS file is viewable by human beings. You can search and view individual headings and data elements and see the details of the authority data for the term or subject heading. There is a visualization tool based on AquaBrowser option.

The Web page also has a place for you to submit suggestions for additions or changes or to ask questions.

“Authorities & Vocabularies”

- URI for specific LCSH records/
concepts:
`id.loc.gov/authorities/[LCCN]`
`id.loc.gov/authorities/sh8508803`

The Website for LCSH includes the Universal Resource Identifier for each term in the LCSH file – the URIs are structured to begin with `id.loc.gov/authorities` and then the Library of Congress Control Number (LCCN), for example:
`id.loc.gov/authorities/sh8508803`.

Slide 13

The screenshot shows the Library of Congress website's "Authorities & Vocabularies" page. At the top, there is a navigation bar with the Library of Congress logo, "ASK A LIBRARIAN", "DIGITAL COLLECTIONS", "LIBRARY CATALOGS", a search box, and a "GO Options" button. Below this is a breadcrumb trail: "The Library of Congress > Authorities & Vocabularies". The main content area is titled "Authorities & Vocabularies" and features a decorative image of a classical architectural frieze. On the left, a sidebar menu lists: "Authorities & Vocabularies", "About", "Technical Center", "Search & Download", and "Contact Us". The main text area contains an "About" section with the following text: "The Library of Congress Authorities and Vocabularies service enables both humans and machines to programmatically access authority data at the Library of Congress via URIs." followed by a "Read more..." link. Below that is an "Available Authorities & Vocabularies" section with a link to "Library of Congress Subject Headings" and a paragraph: "LCSH has been actively maintained since 1898 to catalog materials held at the Library of Congress. By virtue of cooperative cataloging other libraries around the United States also use LCSH to provide subject access to their collections. In addition LCSH is used internationally, often in translation."

The "Authorities & Vocabularies" service is also known as id.loc.gov.

Slide 14

LIBRARY OF CONGRESS ASK A LIBRARIAN DIGITAL COLLECTIONS LIBRARY CATALOGS GO Options

The Library of Congress > Authorities & Vocabularies > LC Subject Headings

Library of Congress Subject Headings

Search Authorities & Vocabularies

Enter Keyword or Phrase

Hold CONTROL key for 2 or more

- All
- Library of Congress Subject Headings
- Thesaurus of Graphic Materials
- MARC Code List for Relators
- Preservation Events

Go

Details Suggest Terminology

Library of Congress Subject Headings

URI
<<http://id.loc.gov/authorities#conceptscheme>>

About
Library of Congress Subject Headings (LCSH) has been actively maintained since 1898 to catalog materials held at the Library of Congress. By virtue of cooperative cataloging other libraries around the United States also use LCSH to provide subject access to their collections. In addition LCSH is used internationally, often in translation. LCSH in this service includes all Library of Congress Subject Headings, free-floating subdivisions (topical and form), Genre/Form headings, Children's (AC) headings, and validation strings* for which authority records have been created. The content includes a few name headings

14

Each page representing a term includes a “Suggest Terminology” tab, which enables anyone to suggest new terminology or changes to existing terminology and to alert LC concerning any problems.

Terminology suggestions

- Suggestions for new headings or changes to existing headings: implemented June 2010
- 26 suggestions have been received to date
- 13 of the suggestions have resulted in changes to terminology or in new subject headings
 - **New subject headings:** Glaucous gull; Shapeshifters; Childfree choice; etc.
 - **Changes to existing headings:** Handbell duets; Barbershop quartets; Poetry therapy; Swindlers and swindling; etc.

That option to send us terminology suggestions is for new heading suggestions or reports about changes that people feel would be helpful. So far we have received 26 suggestions. Thirteen of them have resulted in changes to the LCSH terminology or new subject headings.

New headings, which were added as a result of terminology suggestions include **Glaucous gull**, **Shapeshifters**, and **Childfree choice**. Terminology suggestions have also resulted in changes to existing headings, such as, **Handbell duets**, **Barbershop quartets**, **Poetry therapy**, and **Swindlers and swindling**.

Slide 16

Search

Your search for "animated films" retrieved 17 results. Do you need [assistance with your search?](#)

Heading	Concept Type	LCCN
Animated films	Topical Term	sh85088032
Animated films	Genre/Form Term	sh2007025313
Animated films	Topical Term	sj96004798
Animated films and children	Topical Term	sh2005002226
Animated films--Bulgaria	Topical Term	sh85088034
Animated films--France	Topical Term	sh85088035
Animated films--History and criticism	Topical Term	sh2003003086
Animated films--Hungary	Topical Term	sh85088036
Animated films--Japan	Topical Term	sh99004613
Animated films--Kazakhstan	Topical Term	sh86006561
Animated films--Sound effects	Topical Term	sh2003010291
Animated films--Soviet Union	Topical Term	sh86005226
Animated films--Switzerland	Topical Term	sh85088037
Animated films--Ukraine	Topical Term	sh86004276
Animated films--United States--History and criticism	Topical Term	sh2007101146

A search for the heading **Animated films** in id.loc.gov, results in 3 possibilities –each with a different LC Control number – the prefix on the LCCNs indicate 3 different types of headings –

- sh (LCSH – topical heading)
- sj (Children’s Subject Headings)
- g/f (Genre/Form).

All three of these types of headings are included in this system.

Animated films

This heading is used as a topical heading for works about films that create the illusion of movement in drawings, clay, inanimate objects, or the like, through an animation technique. When used as a topical heading it is subdivided by the appropriate geographic, topical, and/or form subdivisions.

Works on the technical aspects of making animated films, television programs, or videos are entered under Animation (Cinematography).

URI: <<http://id.loc.gov/authorities/sh85088032#concept>>

Type: Topical Term

Alternate Labels: Animated cartoons (Motion pictures); Cartoons, Animated (Motion pictures); Motion picture cartoons; Moving-picture cartoons

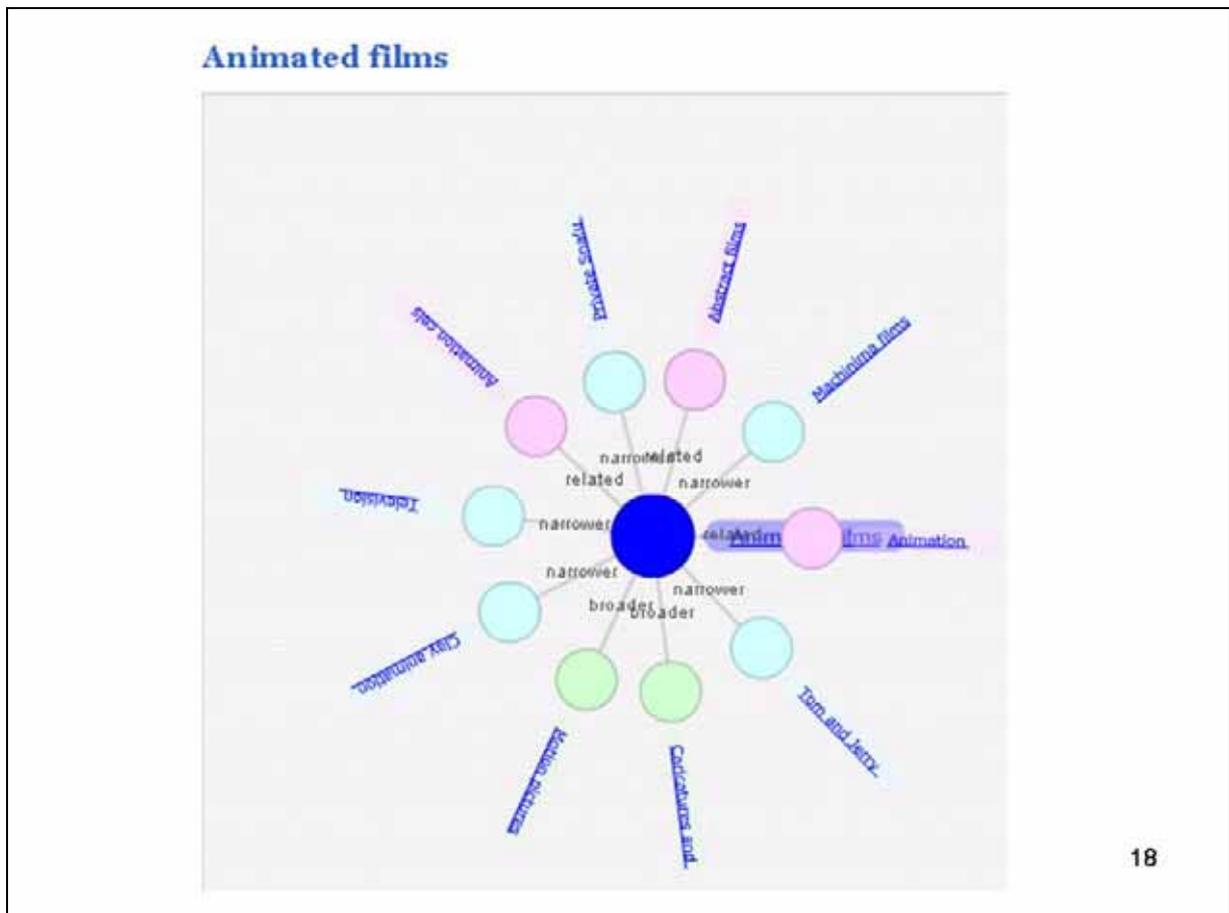
Broader Terms:

- [Caricatures and cartoons](#)
- [Motion pictures](#)

Narrower Terms:

- [Clay animation films](#)
- [Machinima films](#)
- [Private Snafu films](#)
- [Television broadcasting of animated films](#)
- [Tom and Jerry films](#)

Each element in the MARC subject heading record is displayed in a user-friendly form including scope notes, alternate labels, broader terms, narrower terms, and examples.



We also offer the ability to visualize the LCSH terms using AquaBrowser – the visual display indicates narrower terms in blue, broader terms in green, and related terms in pink. Each of these graphics is clickable, so the relationships can be explored.

Slide 19

Your Search
Search Term: antiques
Vocabulary:

of Results: 22 results in 0.23 seconds
 Do you need [assistance with your search?](#)

Search Authorities & Vocabularies

Enter Keyword or Phrase Hold CONTROL key for 2 or more

All
 Library of Congress Subject Hea
 Thesaurus of Graphic Materials
 MARC List for Relators
 MARC List for Countries

Go

Next >

Label	Vocabulary	Concept Type	Identifier
Antiques Antique collecting; Antiques--Collectors and collecting	LCSH	Topical Term	sh85005746
Selling--Antiques Antiques--Selling	LCSH	Topical Term	sh85119821
Antiques, Victorian Victorian antiques	LCSH	Topical Term	sh85005754
Antiques--Asia	LCSH	Topical Term	sh85005753

Searches can be limited to any one, more than one, or all of the vocabularies.

Slide 20

Antiques

Here are entered works on decorative or utilitarian objects, generally more than one hundred years old, that have aesthetic or historic importance and financial value. General works on any objects of interest to collectors, including mass produced items of little intrinsic value, are entered under Collectibles. Works on decorative art objects are entered under Art objects.

URI
<<http://id.loc.gov/authorities/sh85005746#concept>>

Type
Topical Term

Alternate Labels

- > [Antique collecting](#)
- > [Antiques--Collectors and collecting](#)

Broader Terms

- > [Antiquities](#)
- > [Art](#)
- > [Collectibles](#)
- > [Collectors and collecting](#)
- > [Decoration and ornament](#)
- > [Decorative arts](#)

Narrower Terms

- > [Antiquarian booksellers](#)

Related Terms

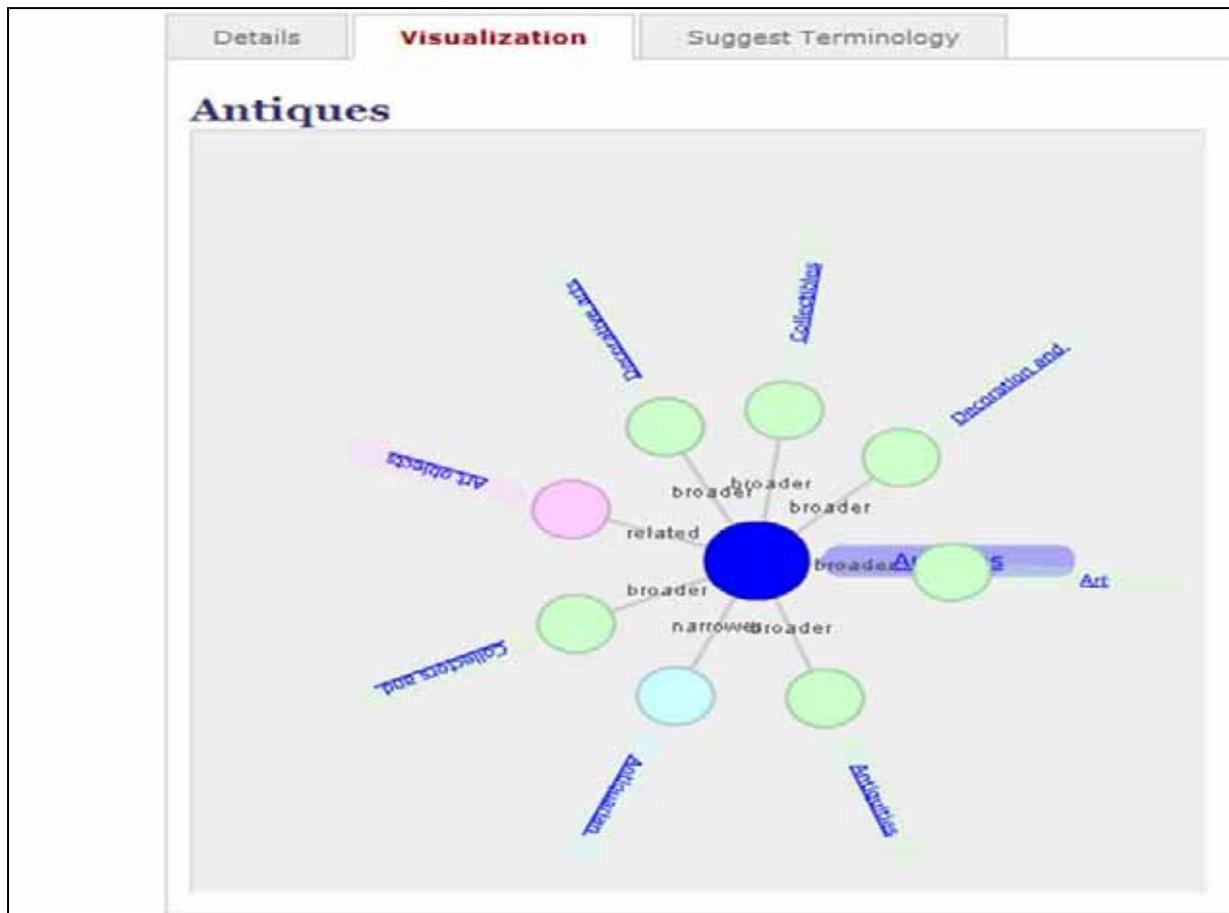
- > [Art objects](#)

Examples

- > Notes under Art objects; Collectibles

Searching for Antiques will display the scope note, the URI, the type of term, alternate labels, the broader terms, narrower terms, related terms, and examples.

Slide 21



As stated before any search, such as the one for Antiques, may be viewed as a visualization and the relationships explored by clicking on the various colors.

Similar concepts from other vocabularies

> <<http://stitch.cs.vu.nl/vocabularies/rameau/ark:/12148/cb11934773k>> 

Concept information

URI	http://stitch.cs.vu.nl/vocabularies/rameau/ark:/12148/cb11934773k	
prefLabel	x-notation	FRBNF119347731
	fr	Antiquités (objets anciens)
altLabel	fr	Antiquités (objets anciens) -- Collectionneurs et collections
	fr	Antiquités (objets anciens) -- Collections
	fr	Objets anciens
scopeNote	fr	Sous cette vedette on trouve les ouvrages sur les objets décoratifs ou utilitaires anciens ayant une valeur esthétique, historique et monétaire. Les ouvrages sur les objets d'art décoratifs se trouvent sous la vedette Objets d'art
	fr	Voir aussi aux noms d'objets anciens surtout lorsqu'ils sont suivis des subdivisions Catalogues, Collectionneurs et collections, Expositions
inScheme	Rameau	
	Rameau - Noms Communs	
broader	Art	
narrower	Trésors d'églises	
	Papiers de collection	
	Objets de marine	

When there are links to your searched term used in parallel LCSH-based systems in other languages, for example, from RAMEAU or from the Bibliothèque nationale de France, the system alerts you by showing the statement “Similar concepts from other vocabularies” on the screen.



“Authorities & Vocabularies”

- Contact information
 - Content of site:
Libby Dechman, edec@loc.gov
 - Technical questions:
Larry Dixon, ldix@loc.gov

23

If you would like more information or have further questions about the content of this site – about LCSH or LCC, please contact Libby Dechman in our Policy & Standards Division at edec@loc.gov.

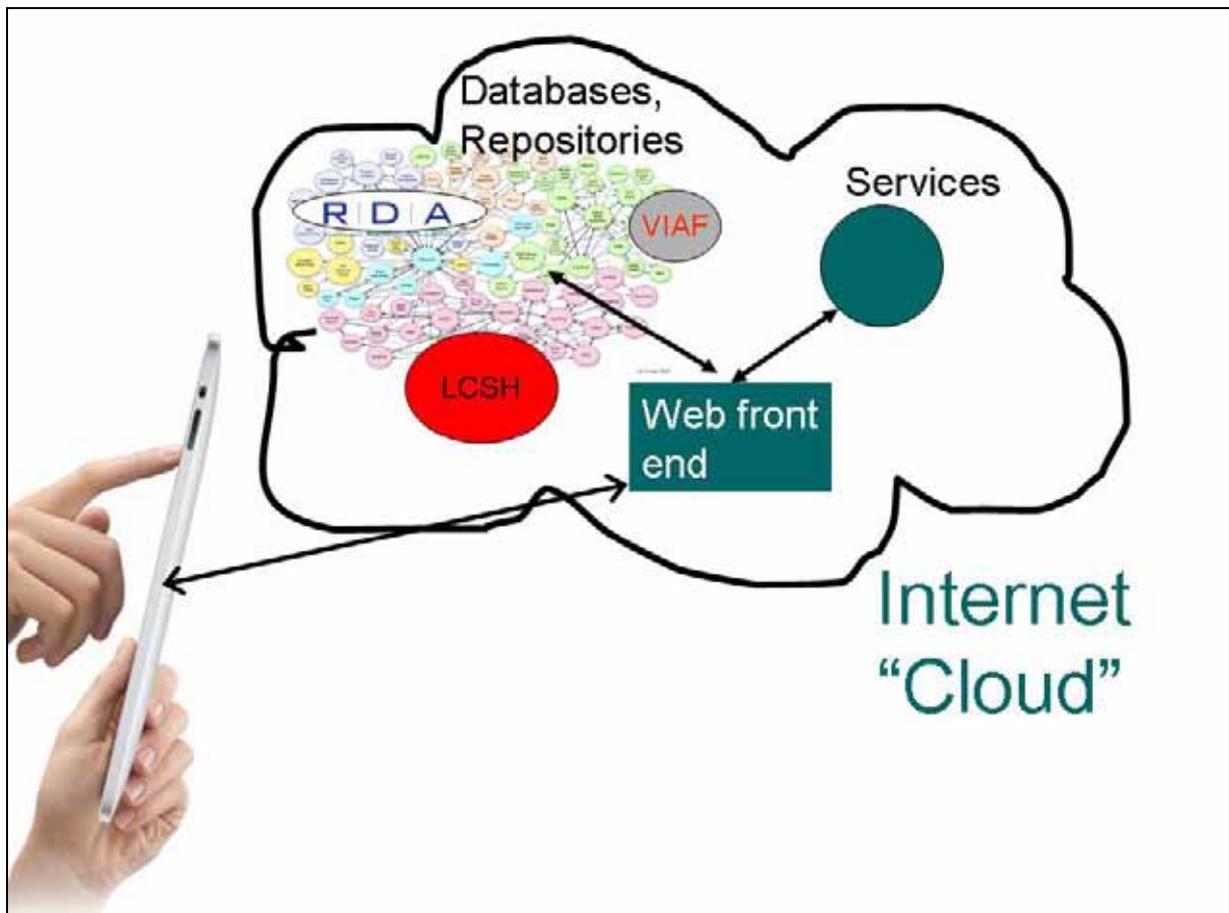
For any technical questions, please contact Larry Dixon in our Network Development and MARC Standards Office at ldix@loc.gov.

“Authorities & Vocabularies”

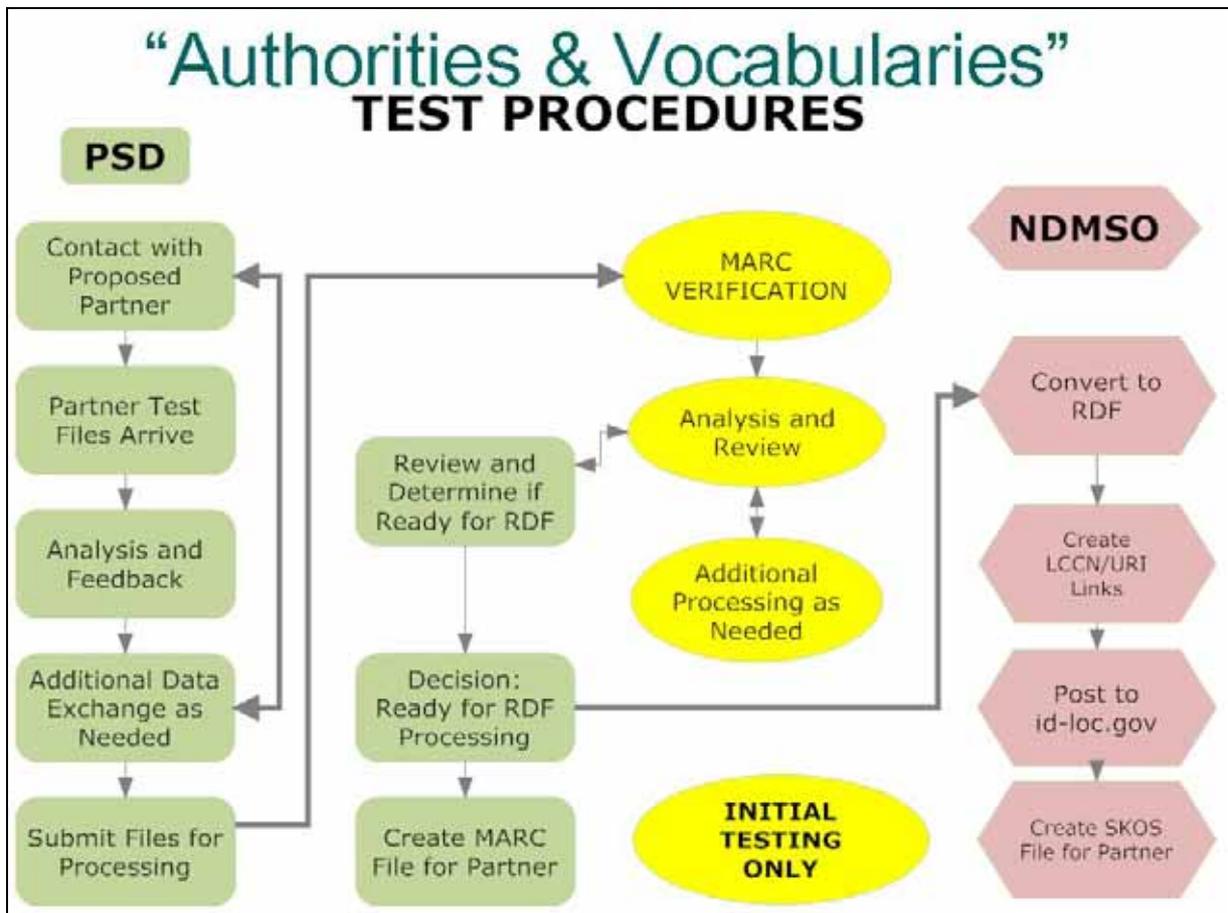
- A comment form and discussion list are available at

<http://id.loc.gov/authorities/contact.html>

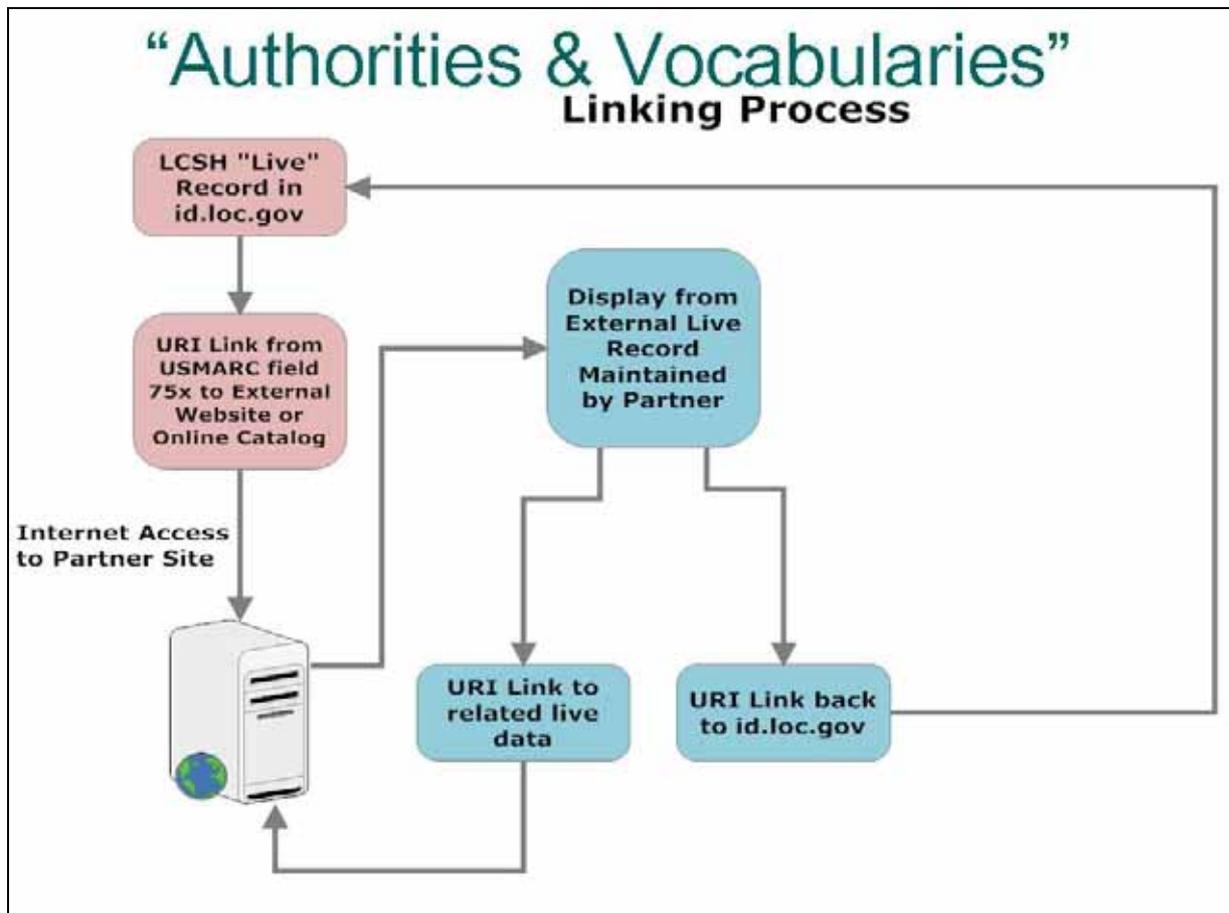
We have a comment form and a discussion list at this URL:
<http://id.loc.gov/authorities/contact.html>



Google and Yahoo work in a cloud computing environment with linked data and a global perspective, and now libraries are finally getting into the act. I hope to see really creative services emerge that take advantage of the building blocks libraries provide through our controlled vocabularies and bibliographic data, so users can connect to our collections.



We have tested creating data links using on LCSH/SKOS data as a model with a few libraries. Based on conversations with a proposed partner, next steps are determined. These include receiving test data, analyzing that data, converting to SKOS, and sending test MARC and SKOS files back to the partner. After the partner agrees to maintain links and contribute data, actual interagency agreements would be signed.



The ideal linking process would have each participant in an agreement maintaining links on their own sites. No single agency would need to have copies of the others data residing on local servers. This reduces the amount of work for each agency and allows the linking data to be locally controlled and available to all via the World Wide Web.

Brief biographical information about the authors/presenter:

Barbara B. Tillett: Chief, Policy and Standards Division, Library of Congress - *presenter*

Libby Dechman: Senior Subject Policy Specialist, Policy and Standards Division, Library of Congress

Loche McLean: Coordinator for Special Products and Services, Policy and Standards Division, Library of Congress