CURATED LIST OF LIBRARY RESOURCES (CULLR)

Metadata Working Group Presentation on 9/16/2011

CuLLR Charge

- Design and implement a process whereby print and electronic resources in the library catalog (in the form of metadata)…are extracted in accordance with the subject areas for a specific library.

- These resources may then be annotated to identify the subject areas for which the resource is useful and other attributes as determined by the project team.

- The process must include the ability to update data extracted from the catalog and other sources without destroying the existing annotations.

- The CuLLR project will be coordinated with the Discovery and Access Task Force, particularly with Group 3, to insure that our efforts are in line with the long-term plans of Cornell University Library (CUL).

CuLLR Team Members

Adam Chandler (for technical services skills)*
John Cline (for technical skills)
Jeremy Cusker (for engineering)*
Dianne Dietrich (for technical & metadata skills)*
Michael Engle (for humanities)*
Holly Mistlebauer (for project management)
Jim Reidy (for technical skills)
Marty Schlabach (for life sciences)

Rick Silterra (for technical skills)
Leah Solla (for physical sciences)

Member of CUL Discovery and Access Group.

Used PSL Website as Example

PSL Website does what CuLLR is doing, but CuLLR is automated…

- Faceted Searching
  - Search by Discipline and Format then filter results by Sub-Discipline, Format, & Content (like PSL website)
  - Search using keyword(s) then filter results by Discipline, Format, & Content (like Engineering website)
- Faceted Browsing (browse by Discipline then filter results by Sub-Discipline, Format, & Content)

http://physicalsciences.library.cornell.edu/

Additional Attributes

CuLLR contains data from Voyager, plus the following fields from other automated sources…

- Table of contents (for journals use Journal TOCs API)
- Open Access Journal (SHERPA/RoMEO API)
- Cover Image (LibraryThing API & OpenAccess API)
- Scientific Name (uBio API)—coming soon!
- Common Name (uBio API)—coming soon!
- Chemical/Physical Property (PSL provided database)—coming soon!
- Material/Compound Type (PSL provided database)—coming soon!

Additional Attributes (continued)

CuLLR also contains data entered manually…

☐ Discipline (via library-provided spreadsheet and curator interface)

☐ Sub-Discipline (via library-provided spreadsheet and curator interface)

☐ Weight (via library-provided spreadsheet and curator interface)

☐ Sub-Resource Type—now Content (via library-provided spreadsheet and curator interface)

☐ Annotation (via curator interface)

☐ Annotation Author (via curator interface)

☐ Suppress Record Flag (via curator interface)

☐ Free Text Keywords (via curator interface)

☐ Internal Note (via curator interface)

☐ Internal Note Author (via curator interface)

☐ Internal Note Date (via curator interface)

☐ Latest Update (via curator interface)

Key Tasks

- Which data elements do we want to extract from the catalog?

- How will we identify which titles need to be extracted from the catalog?
  - Perfect world: subject headings
  - Our world: package id (899), call number range, and bibid

Key Tasks (continued)

- How will we identify items in the catalog as books, ebooks, journals, ejournals, and databases?
  - Database has 948|f="fd" (have to add it if not there)
  - eBook has LDR 7&8 = "am" and 245|h = "electronic resource"
  - Book has LDR 7&8 = "am" and 245|h not = "electronic resource"
  - eJournal has 948|f = "¡" and 245|h = "electronic resource"
  - Journal has LDR 7&8 = "as"
  - Anything else is reported as "other" (to be reviewed)
  - We realize that serials will show up as journals

Key Tasks (continued)

- How will we identify items in the catalog as handbooks or standards?
  - Original idea was to use 650|v value
  - If not available, would ask CTS to populate 653|a (custom field used by CUL)
  - Doesn't work because each library could have it's own values
  - Getting info from input spreadsheet instead

Five Fields on Input Spreadsheet

- ☐ Identifier (bibid, call number range, or package id)
- ☐ Discipline (high-level field of study, such as Chemistry)
- ☐ Sub-Discipline (sub-field of study, such as Organic Chemistry)
- ☐ Weight (a number from 1 to 100, with 100 being the highest weighting)
- ☐ Sub-Resource Type--now Content (A further designation of the type of resource, such as Handbooks, Standards, Dictionaries, Thesis, etc.)

Spreadsheet Examples

	A	B	C	D	E
1	BIBID	DISCIPLINE	SUB-DISCIPLINE	WEIGHT	SUB-FORMAT
2	314182	Civil Engineering		80	Standard
3	4090658	Computer Science		1	
4	6189208	Computer Science		1	Handbook
5	6189208	Civil Engineering		1	Handbook
6	6887306	Earth Sciences		1	
7	7055187	Mechanical Engineering		50	
8	7055187	Biomedical Engineering		75	

	A	B	C	D	E
1	CALL NUMBER RANGE	DISCIPLINE	SUB-DISCIPLINE	WEIGHT	SUB-FORMAT
2	GB 3 | GB 5030	Civil Engineering		1	
3	Q 327	Computer Science		1	
4	QC | QE	Earth Sciences		1	
5	R 856 | R 857	Biomedical Engineering		75	
6	R 856 | R 857	Electrical Engineering		50	

	A	B	C	D	E
1	PACKAGE	DISCIPLINE	SUB-DISCIPLINE	WEIGHT	SUB-FORMAT
2	kmoddlbk	Mechanical Engineering		1	
3	knovelebksadhes	Biomedical Engineering		50	
4	knovelebksadhes	Materials Science		1	
5	knovelebksaero	Mechanical Engineering		1	Handbook
6	knovelebksaero	Electrical Engineering		25	Handbook
7	knovelebkselectric	Electrical Engineering		1	Handbook

CuLLR in Action

Demo new Engineering Library Website at
http://engineering.test1.library.cornell.edu/

CuLLR Phase II

- Improve curator interface, in particular the process of weighting resources.
- Create web service to allow the use of another library's annotations.
- Create consistent input spreadsheet format.
- Eventually replace SVN input data?
- Add feature to PLOP whereby curators search and select items for their unit's virtual library. There should be an option to exclude anything already selected.
- How will Scientific Name and Common Name be derived from uBio?
- How will Chemical/Physical Property and Material/Compound Type be derived from PSL tables?
- How should be handle more than one Open Access Journal color code?
- What will the process be for adding new attributes?
- Add additional field(s) to PLOP: Editor.

Technical Details

And now for the CuLLR technical details…