Crowdsourcing Terms for Thematic Exploration in the Catholic Portal

Ingrid Hsieh-Yee, Catholic University of America
Pat Lawton, University of Notre Dame

Catholic Research Resources Alliance

The Catholic Portal of the Catholic Research Resources Alliance (CRRA) facilitates access to resources related to studies of the Catholic experience and supports CRRA's mission to provide enduring global access to Catholic research resources. The Portal builds its collections on 12 themes, and CRRA has experimented with portal theme enhancements to support exploration and use of its collections. The crowdsourcing project is part of the subject access enhancement effort.

Project objectives

1. To analyze topical terms from users for patterns and subject coverage
2. To identify highly relevant expressions for inclusion in the Portal

Topical terms are compared with the Library of Congress subject headings and analyzed for possible linking to Portal subject headings. Findings will have implications for enhancing subject access to Catholic resources in local catalogs and the CRRA's Catholic Portal.
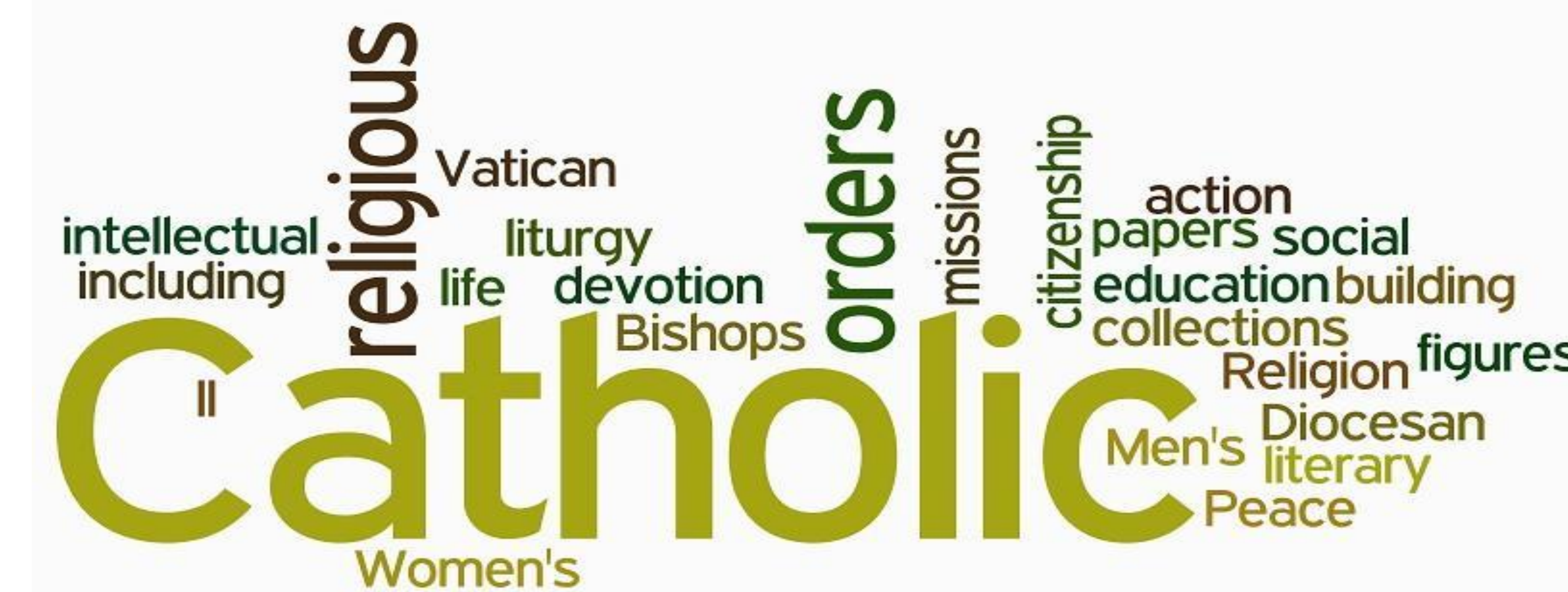
Methodology
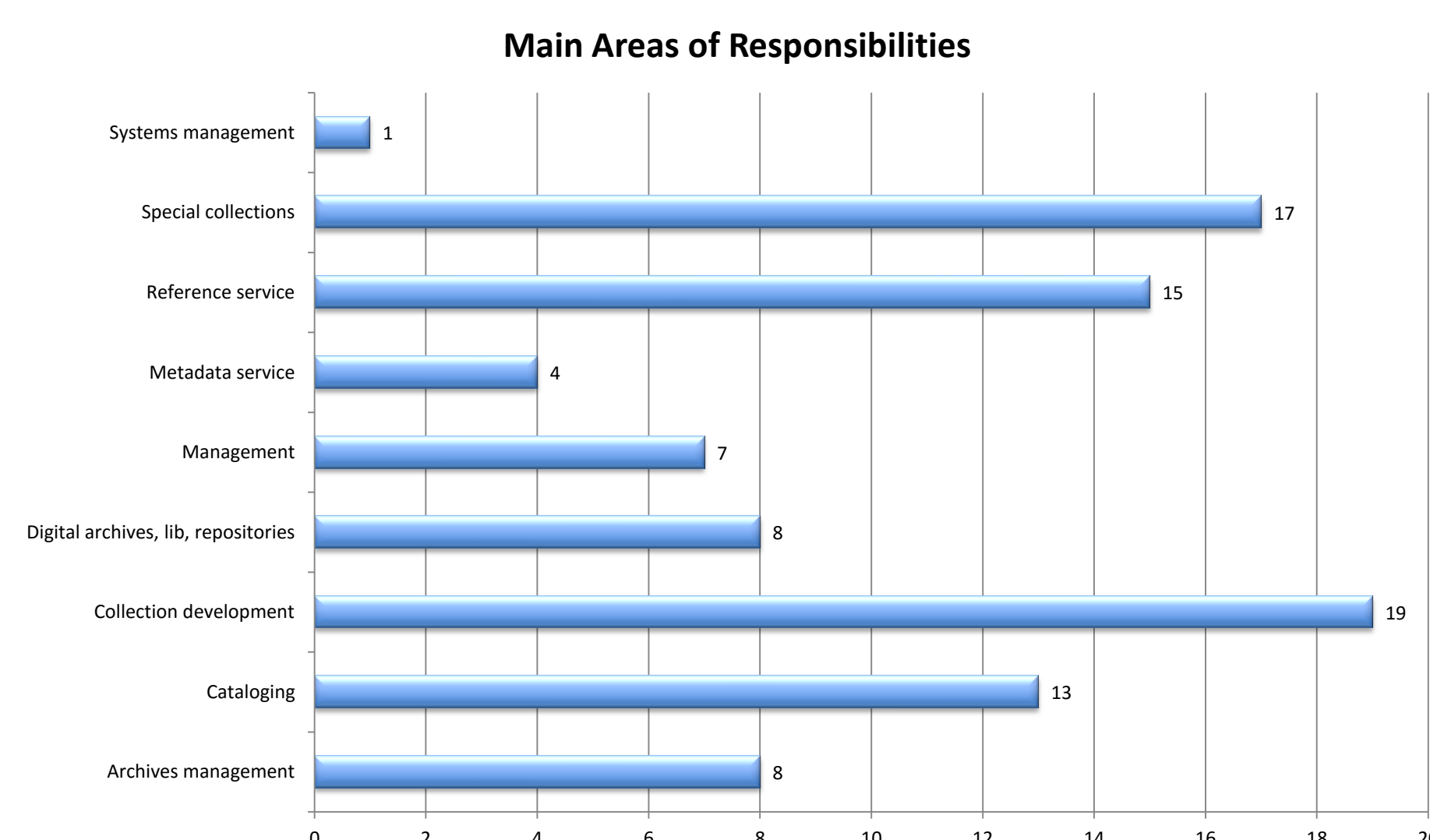
Crowdsourcing through an online survey

Survey Design

The study focuses on eight of the 12 CRRA Portal Themes. To make the tasks more manageable for participants, two online surveys were designed. The surveys were identical in all aspects, except one version covers Catholic Education, Liturgy, Peace-building, and Catholicism and Citizenship, while the other covers Diocesan Collections, Catholic Intellectual Life, Catholic Missions, and Catholic Social Action.

Data Collection

Fifty-eight librarians from members of the Catholic Research Resources Alliance were invited to participate in an online survey to provide topical terms for selected Portal themes. The survey has a 71% response rate, with 41 of the 58 invitees participating in the study. However, not all of them provided topical terms for the four portal themes presented in the survey.

Participant Background (n=41)
(up to 3 areas named)

Main Areas of Responsibilities

Area	Count
Systems management	1
Special collections	17
Reference service	15
Metadata service	4
Management	7
Digital archives, lib, repositories	8
Collection development	19
Cataloging	13
Archives management	8

Years of Professional Experience

Range	Count
0-4 years	6
5-10 years	6
11+ years	27

Terms Collected: 404 terms

Topic	Education	Liturgy	Peace-building	Catholicism & Citizenship
Number of terms collected	42	47	40	38

Topic	Diocesan	Intellectual life	Missions	Social actions
Number of terms collected	63	55	60	59

Project Status

Data analysis of **Catholic Education**, **Catholic Social Action, Liturgy,** and **Peace-building** has been completed, and preliminary findings are reported in this poster.

Next Steps

✧ Topical terms for the remaining four themes will be analyzed.
✧ Findings of this project will be compared with findings from a portal heading analysis project to assess the merits of each approach for portal enhancement.

Term Analysis

Step 1. Sort terms by level of relevance reported by each participant, resulting in groups of terms marked as HR (highly relevant), R (relevant), and N (neutral)
Step 2. Start with the HR terms and search against LCSH in the authority file at http://authorities.loc.gov to identify user terms that match LCSH or cross references.
Step 3. Check R terms and N terms against LCSH authority file to identify user terms that match LCSH or cross references.
Step 4. Results of Step 2 and Step 3 identify LC subject headings and cross references matched by user terms. Use the matched LC subject headings as the pool to code user terms that do not match LC subject headings. Data values for coding term relationships to identified LCSH include

Sample codes for term relationship
Va: Variant expressions
   "Catholic education" for "Catholic Church—Education" (an LCSH)
Na: New concept, associated with an identified subject heading
   "Seminary" to "Catholic Church—Education"
Nb: New concept, broader than an identified subject heading
Nn: New concept, narrower than an identified subject heading
   "Catholic teachers" to "Catholic Church—Education"
NC: New concept. No hierarchical relations with headings identified

EDUCATION

Relationship to LCSH Terms	Cases (n=42)	% of potential search success
LCSH MATCH	12	12/42=29% success
XREF to LCSH	1 if included	13/42=31% success
Va	8 if included	21/42=57% success
Na	5 if included	26/42=62% success
Nn	13 if included	37/42=88% success
Nb	0	
typo	0	
NC	0	
X to UNC	2	
N to UNC	1	

SOCIAL ACTION

Relationship to LCSH Terms	Cases (n=59)	% of potential search success
LCSH MATCH	26	26/59=44% success
XREF to LCSH	3 if included	29/59=49% success
Va	18 if included	47/59=80% success
Na	5 if included	52/59=88% success
Nn	4 if included	56/59=95% success
Nb	1	
typo	0	
NC	1	
X to UNC	1	
N to UNC	0	

PEACE-BUILDING

Relationship to LCSH Terms	Cases (n=40)	% of potential search success
LCSH MATCH	26	26/40=65% success
XREF to LCSH	4 if included	30/40=75% success
Va	2 if included	32/40=80% success
Na	0	
Nn	0	
Nb	0	
typo	0	
NC	8	
X to UNC	0	
N to UNC	0	

LITURGY

Relationship to LCSH Terms	Cases (n=47)	% of potential search success	
LCSH MATCH	27	27/47=57% success	
XREF to LCSH	2 if included	29/47=62% success	
Va*	8 if included	37/47=79% success	
Na	0		
Nn	2 if included	39/47=83% success	
Nb	0		
typo	2		
NC	4		
X to UNC	1		* 1word order, 7 va to UNC
N to UNC	1		

Pattern of Findings

Besides LCSH matches and cross references already in LCSH, these three types of term relationships, if accommodated, will increase success in retrieving items from the Catholic Portal:

✓ **Va: variant expression to a named subject heading**
✓ **Na: new concept, associated with a named subject heading**
✓ **Nn: new concept, narrower than a named subject heading**