OhioLINK—OCLC Collection and Circulation Analysis Project 2011

OhioLINK Collection Building Task Force

Julia Gammon
Head, Acquisitions Department
University of Akron

Edward T. O'Neill Senior Research Scientist OCLC Research





A publication of OCLC Research, in collaboration with OhioLINK

OhioLINK OCLC Collection and Circulation Analysis Project. 2011
OhioLINK Collection Building Task Force, Julia Gammon and Edward T. O'Neill

© 2011 The Ohio Library and Information Network (OhioLINK) and OCLC Online Computer Library Center, Inc.

Reuse of this document is permitted as long as it is consistent with the terms of the Creative Commons Attribution-3.0 Unported (USA) license (CC BY 3.0):

http://creativecommons.org/licenses/by/3.0/.

September 2011

Published by OCLC Research, in collaboration with OhioLINK Dublin, Ohio 43017 USA www.oclc.org

ISBN: 1-55653-432-9 (978-1-55653-432-4)

OCLC (WorldCat): 744669700

Please direct correspondence to:
Julia Gammon
Head, Acquisitions Department
University of Akron
jgammon@uakron.edu

or

Edward T. O'Neill Senior Research Scientist OCLC Research oneill@oclc.org

Suggested citation:

OhioLINK Collection Building Task Force, Julia Gammon and Edward T. O'Neill. 2011. *OhioLINK OCLC Collection and Circulation Analysis Project 2011*. Dublin, Ohio: OCLC Research. http://www.oclc.org/research/publications/library/2011/2011-06r.htm.

OhioLINK Collection Building Task Force Members

Current members

Joyce Baker Belmont Technical College jbaker@btc.edu

Danny Dotson
The Ohio State University
dotson.77@osu.edu

Phil Flynn Wright State University phil.flynn@wright.edu

Julia Gammon, Chair University of Akron jgammon@uakron.edu

Jessica Grim
Oberlin College
jessica.grim@oberlin.edu

Rob Kairis Kent State University, Stark Campus <u>rkairis@stark.kent.edu</u> Kevin R. Messner Miami University messnekr@muohio.edu

Andrew Whitis
Defiance College
awhitis@defiance.edu

Carol Zsulya Cleveland State University c.zsulya@csuohio.edu

Former members who participated

Margo Warner Curl The College of Wooster

Celeste Feather OhioLINK

Anne Gilliland OhioLINK

Institutions Participating in the Project

Antioch University Ashland University Athenaeum of Ohio Baldwin-Wallace College Belmont Technical College

Bluffton University

Bowling Green State University

Capital University

Case Western Reserve University

Cedarville University Central State University

Cincinnati Christian University

Cleveland Clinic

Cleveland State University College of Mount Saint Joseph

College of Wooster

Columbus College of Art and Design Columbus State Community College

Cuyahoga Community College

Defiance College Denison University

Edison State Community College

Franciscan University Franklin University Heidelberg University

Hiram College Hocking College

John Carroll University Kent State University

Kenyon College

Lakeland Community College Lorain Community College

Lourdes College Malone University Marietta College

Mercy College of Northwest Ohio

Miami University

Mount Carmel College of Nursing

Mount Union College

Mount Vernon Nazarene University

Muskingum University Notre Dame College Oberlin College

Ohio Dominican University Ohio Northern University Ohio State University

Ohio University

Ohio Wesleyan University

Otterbein College

Owens Community College Rio Grande Community College Shawnee State University Sinclair Community College

Southern State Community College

State Library of Ohio Terra Community College

Tiffin University
University of Akron
University of Cincinnati
University of Dayton
University of Findlay
University of Toledo
Urbana University
Ursuline College
Walsh University

Washington State Community College

Wilberforce University Wilmington College Wittenberg University Wright State University Xavier University

Youngstown State University

Contents

List of Illustrations 6
Introduction
Overview of the Study and Its Outputs
Purpose and Goal of the Study
Distinctive Aspects of the Study10
Gathering the Data11
Location and Form of the Analysis
Organization of Data
Analysis
Institutional Level Analysis
Data Overview: Large Institutions
Data Overview: Smaller Institutions
Collective Analysis of OhioLINK Libraries
Subject Analysis
Language Analysis
Uses for the Data
Conclusions
Glossary
Publications and Presentations
References
Appendix (external) http://www.oclc.org/research/publications/library/2011/2011-06a.htm

Illustrations

Tables

Table 1. Elements of the Circulation Record	. 12
Table 2. OhioLINK Resources.	. 15
Figures	
Figure 1. Group 1 entities and relationships in the FRBR hierarchy	. 14
Figure 2. Example of a three-level institutional hierarchy	. 17
Figure 3. The institutional and campus spreadsheets for Ohio State University	. 18
Figure 4. Example of institutional spreadsheet (Subjects analysis)	. 19
Figure 5. Example of an institutional spreadsheet (General tab)	. 20
Figure 6. Example of campus spreadsheet (OSU Marion)	. 20
Figure 8. Sample portion of the index for law	. 24
Figure 9. Portion of the Subjects Spreadsheet for International Law	. 24
Figure 10. Sort dialog for the default sort order (Comprehensiveness)	. 25
Figure 11. Portion of the Index for Languages.	. 26
Figure 12. Language Analysis for Danish	. 27
Figure 13. Danish-Language Books on History and Geography	. 28

Introduction

The goal of this study is to gain a better understanding of how the resources of OhioLINK libraries are being used and to identify how the limited resources of OhioLINK member libraries can be utilized more effectively. OhioLINK, the Ohio Library and Information Network, is a consortium of 88 Ohio college and university libraries, plus the State Library of Ohio, that work together to provide Ohio students, faculty and researchers with the information they need for teaching and research. OhioLINK serves more than 600,000 users at 89 institutions, including 16 public/research universities, 23 community/technical colleges, 49 private colleges and the State Library of Ohio (OhioLINK 2010).

To gain a better understanding of OhioLINK collections and their use, circulation data for virtually all books in Ohio academic libraries was collected and analyzed. Using this detailed data, the expectation is that the OhioLINK community can begin to create a set of collecting rubrics that will help reduce unnecessary duplication, allocate resources more effectively, and increase diversity of collections across the state. Primary examples of the use of this data would be the inference of how many copies of books in particular Library of Congress classification ranges are desirable in the system, or the ideal size of subject collections. It is also expected that engagement with and use of this data by selectors, individually and in groups, will contribute significantly to the evolution of selector thinking and behavior as we move toward more cohesive collecting practices across the state.

The OhioLINK Collection Building Task Force (CBTF) identified the data needed, and worked closely with OCLC Research to plan and test the study. OCLC Research is the investigative division of OCLC Online Computer Library Center, Inc., a nonprofit, membership, computer library service and research organization dedicated to the public purposes of furthering access to the world's information and reducing information costs. More than 72,000 libraries in 170 countries and territories around the world have used OCLC services. OCLC and its worldwide member libraries cooperatively produce and maintain WorldCat, the world's largest and richest database of bibliographic information. OCLC also publishes the Dewey Decimal Classification system, the most widely used library classification system in the world (OCLC 2011a). OCLC Research investigates trends in technology and library practice; identifies technological advances that enhance library services; and works with the

community to collaboratively identify problems and opportunities, prototype and test solutions, and share findings through publications, presentations and professional interactions.

During the spring of 2007 and again in the spring of 2008, library systems managers at each of the OhioLINK libraries were asked to generate files of circulation records for all their print books and manuscripts. Detailed instructions for the creation of these institutional files were written, rewritten and tested by member libraries to ensure that the results generated provided the necessary data for the analysis. The collection of circulation records on this scale is not straightforward and necessitated a new approach to the data collection and validation.

In the early 1990s, under the charge of the Cooperative Information Resources Management Committee (CIRM), the OhioLINK Approval Plan Task Force was formed to select a common book vendor for the state in an effort to better control unnecessary duplication of titles within the state. Soon after, Yankee Book Peddler (now YBP Library Services) was selected, coordinated firm ordering and approval plan profiling began, and the infrastructure for making informed collection decisions became available to consortial libraries via YBP's GOBI online product. The Approval Plan Task Force's charge was later expanded, and it was renamed the Collection Building Task Force (CBTF). CBTF's charge was (and remains) "...to expand the amount spent on cooperative purchases through identifying specific information resources of value to the OhioLINK community and encouraging the reduction of duplication in current purchases, thereby freeing funds to be reallocated to increase the depth and breadth of the collection; to investigate and foster specific methods to achieve the former; and to foster increased involvement in and coordination of local collection development activities with those of other OhioLINK member libraries." (OhioLINK 2004)

CBTF has initiated a variety of projects to encourage cooperative collection development activities throughout the state. Examples include fostering collaboration among the 28 subject groups of collection specialists, presenting "road shows" to market cooperative ideas and promote new vendor services, showcasing different library workflows for improved efficiency, and the "Not Bought in Ohio" feature developed in collaboration with YBP. To guide library selectors in making purchasing decisions, CBTF also provided guidelines for the number of copies per title in the OhioLINK collection (OhioLINK 2005). Some individual libraries have set their own "soft cap" purchasing limits based on these guidelines.

Despite the success of such efforts, however, a chief barrier to more effectively "selling" cooperative collection development to libraries and selectors was the lack of data demonstrating how many copies were actually needed within the consortium in order to meet OhioLINK's circulation needs.

CBTF knew from practical observation that there were too many copies of some books, not enough copies of others, and that some titles were probably being missed altogether. The CBTF approached the collection analysis by creating a list of questions (see external appendix). CBTF quickly learned that answers to the most important questions on usage of materials could not be supplied by the central Innopac circulation system. OhioLINK looked at several off-the-shelf collection-analysis products but at the time none were scalable to the scope and breadth required. In 2005, OhioLINK was approached by OCLC Research about working on a study of circulation of books within the consortium. OCLC Research had recently completed a circulation analysis at the University of Colorado (Knievel, Connaway, and Wicht 2006) and wanted to apply the methodology in a consortial setting. After some discussion, there was agreement that a joint project would be mutually beneficial and the project was initiated in 2006.

Overview of the Study and Its Outputs

Purpose and Goal of the Study

The goal of this study is to better understand the usage and collecting patterns within OhioLINK libraries. By using detailed data the expectation is that the OhioLINK community can begin to create a set of collecting rubrics that will help reduce unnecessary duplication, allocate resource dollars more effectively, and increase diversity of collections across the state. Primary examples of the use of this data would be to infer how many copies of books in particular LC classification ranges are desirable in the system, or to estimate the ideal size of subject collections. It is also expected that engagement with and use of this data by selectors, individually and in groups, will contribute significantly to the evolution of selector thinking and behavior as a general move takes place toward more cohesive collecting practices across the state.

This study is limited to books, including manuscripts, as these materials typically circulate and circulation is the most significant quantifiable element in evaluating book collections.

Distinctive Aspects of the Study

What makes this project unique? Parallels have been drawn between this project and the Pittsburgh Study of the 1970s (Kent 1979), one of the most comprehensive studies of collection usage patterns undertaken to date. The Pittsburgh Study, however, covered only a single institution's usage; not that of a large consortium.

OhioLINK libraries serve more than 600,000 students, faculty, and staff at 90 institutions. The group consists of 16 public/research universities including 5 ARLs, 23 community/technical colleges, 50 private colleges and the State Library of Ohio. The combined collection is comprised of 50 million books and other library materials. The size of the combined collection and the number of participating institutions make this by far the largest and most comprehensive study of academic library circulation ever undertaken.

While there have be a number of multi-institution studies, most have used holdings information from union catalogs, most notably OCLC's WorldCat. While those studies have been very valuable, they were based on manifestation-level holding data. This study used the item-level holding information obtained from the circulation system which provided detailed usage information at the item or individual copy level. This distinction becomes significant when institutions hold multiple copies of the same title. OCLC's work-set algorithm (Hickey and Toves 2005) was employed to identify works permitting FRBR (Functional Requirements for Bibliographic Records) analysis at the work, manifestation, and item levels (IFLA 2009, 17-24).

A two-phase data collection process was used to highlight both historical aspects, and recent circulation trends, of collections. The data collected included the cumulative circulation that occurred since a book was entered into the circulation system. For material acquired after the implementation of computerized circulation systems, the complete circulation history of the material is available. The data collected also included the annual circulation, permitting both an historical and a current view of usage patterns. The analysis utilized new metrics including OCLC's Audience Level and comprehensiveness, as well as more established metrics such as collection size and annual circulation rates.

Gathering the Data

The CBTF identified the data needed, and worked closely with OCLC Research to plan and test the study. During the spring of 2007 and again in the spring of 2008, the library-systems managers at each of the OhioLINK libraries generated a file of circulation records for all their print books. Detailed instructions for the creation of the file were written, rewritten and tested by member libraries to ensure that the results generated provided the necessary data for the analysis. See the appendix for a copy of the instructions.

The elements comprising the circulation records are shown in Table 1.

The collection of circulation records on this scale proved to be fairly complex; examples of issues that had to be contended with in the process of designing the study:

- Circulation records are at the physical item level.
- Item numbers are unique only within the source file.
- No distinction is made between local use and use by other OhioLINK users.

- Circulation data is cumulative (total circulation since acquired or cataloged or when a library joined OhioLINK), and the dates of cumulation vary from institution to institution.
- The standard numbers (OCLC No., LCCN, ISBN) are not present in all of the records.

Table 1. Elements of the Circulation Record.

Data Element	Definition
Item No.	An identifier that is unique for each item within a source file. The combination of source id and item number is unique within OhioLINK.
OCLC No.	The OCLC number for the corresponding WorldCat bibliographic record.
Title	The brief title. This frequently will be a shorter title than the title in the corresponding bibliographic record.
LCCN	The Library of Congress Control Number.
Location Code	The code for the specific location of the item within the source institution.
Circulation Status	A code indicating whether the item is circulating or non-circulating.
Circulations	The total number of times the item has circulated since being entered into the system.
Renewals	The total number of times the item has been renewed since being entered into the system.
Accession Date	The date the item was entered into the system.
Date of Last Use	The most recent date that the item was checked out.
ISBN	The International Standard Book Number
Source Id	The code for the source institution.

Once OCLC Research received the files of records, an extensive process to match them to the corresponding bibliographic record in OCLC's WorldCat was undertaken. When records contained an obsolete OCLC Number, it was replaced with the current OCLC Number. Records lacking an OCLC number were matched to WorldCat records using either the LCCN or the ISBN. All matches were validated by comparing the brief title in the circulation record with the title in the matching WorldCat record. Records lacking any standard number and those records where the titles were significantly different were excluded from the study. After the second phase of the data collection, the records for both years were combined into a single record. In addition to the information in the original records, the combined record includes the annual circulation for spring 2007 to spring 2008. Approximately 93% of the circulation records were successfully matched to a WorldCat bibliographic record. Only bibliographic records for books were considered for matching; many of the unmatched records were for non-book materials. Circulation records for items not present in both the 2007 and 2008 data were excluded. The resulting data set contained the circulation records for almost 30 million different items.

Location and Form of the Analysis

The analysis is presented in the form of Excel spreadsheets. The number and size of the Excel tables far exceeds what would be practical to print. In addition, the spreadsheets provide additional functionality since they can be sorted, filtered, and otherwise manipulated as necessary to provide customized views of the results. No attempt is made in this document to include the results of the analysis but rather to describe and provide examples of the resulting tables.

All of the spreadsheets are available from the OCLC website at: http://www.oclc.org/research/activities/ohiolink/collections.htm.

All of the spreadsheets generated for the institutional analysis as well as the spreadsheet generated for the collective analysis can be found at this site. A file of all of the circulation records and the file with the corresponding bibliographic records are also available (OCLC 2011b) so other researchers can use the data.

Organization of Data

This collection of circulation data, the largest known to date, can support extensive study and analysis. It is important to note that FRBR terminology is used to define and structure the study. The following FRBR terms are utilized:

Work: A distinct intellectual or artistic creation. For this study, the definition used by Hickey and Toves for the OCLC work set algorithm is used. Stieg Larsson's *The Girl with the Dragon Tattoo* is an example of a work.

Expression: The intellectual or artistic realization of a work. The English language translation by Reg Keeland of *The Girl with the Dragon Tattoo* is an example of a particular expression of the work. Although important in other contexts, expressions are not explicitly used in this study.

Manifestation: The physical embodiment of an expression of a work. In this study, each bibliographic record in WorldCat is considered to be a manifestation. The first American edition published by Alfred A. Knopf (2008) is an example of one the manifestations of the English language translation of *The Girl with the Dragon Tattoo*.

Item: A single exemplar of a manifestation. An item is a particular copy of a manifestation. In this study, each circulation record is assumed to represent a single item. The copy of the first American edition of *The Girl with the Dragon Tattoo* held by Ohio Wesleyan University is an example of an item.

Figure 1 illustrates the FRBR hierarchy used, which emphasizes the relationship between Group 1 entities.

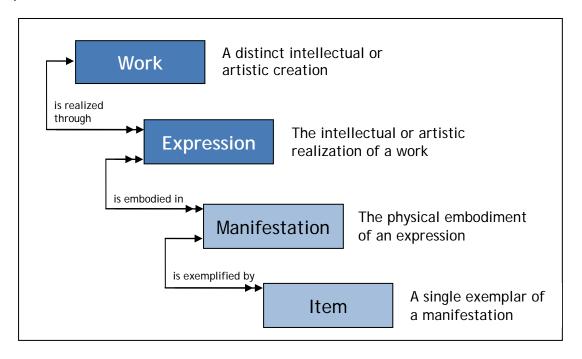


Figure 1. Group 1 entities and relationships in the FRBR hierarchy.

Applying FRBR to the OhioLINK circulating collection results in the collection summary figures shown in Table 2. In addition to viewing these composite totals, it is useful to view each institution's circulation activity in depth locally as well as in comparison to other institutions.

This required the use of location codes to identify non-circulating materials from the total, and to recognize the nature of specific holdings within institutions, their branches, affiliates, and units (all of which might be included in one catalog or overall collection). Labeling specific collections and locales allows for comparison. For example, circulation results of two branch campuses within a university can be compared as can two branches or two collections from two different schools.

Table 2. OhioLINK Resources.

FRBR Resource Types	Frequency
Items	29,570,205
Manifestations	6,955,505
Works	5,686,173

Analysis

The analysis is presented for individual institutions as well as for the collective collection of all OhioLINK member libraries.

Institutional Level Analysis

For larger institutions with multiple campuses and/or branch libraries, the analysis is done separately for the institution as a whole and also for each campus, and the results are tabulated in the associated Excel spreadsheets. When viewing these spreadsheets, the viewer will have the option of choosing a specific campus within an institution, or choosing the institution as a whole (i.e. all locations together). Note that different branch libraries on the same campus are included in the "campus" spreadsheet. For example, if a main campus has five branch locations, all of them will be included on the main campus spreadsheet.

To organize the results of the individual institutional analysis an institutional hierarchy was developed. Campuses, independent cultural institutions, and depositories are designated as top (first) level units, followed by independent administrative units (if present) within the campus, separate libraries (if present) within an administrative unit and, finally, distinct collections with unique location codes.

All of the statistics for the individual institutions are available from the project website at:

http://www.oclc.org/research/activities/ohiolink/institutions.htm

Figure 2 shows University of Akron's hierarchical arrangement as a representative example.

```
Akron Campus (Level 1)

Law Library (Level 2)

University Libraries (Level 2)

Bierce (Level 3)

Science (Level 3)

Archives (Level 3)

Local Storage (Level 3)

Wayne Campus (Level 1)

Wayne Library (Level 3)

NE Depository (Level 1)
```

Figure 2. Example of a three-level institutional hierarchy.

Circulation statistics for each of these units are available for comparative analysis across levels locally or with other institutions across the state. In addition, specific fields from the corresponding bibliographic record allowed for comparisons of specific subsets by language, subject, and publication date.

Data Overview: Large Institutions

As an example of how data is presented for large institutions, the statistics on library resources from all locations included under the collective entry "Ohio State University" is shown in figure 3. The statistics for just the Columbus campus are shown in the "main campus" entry. The statistics for each of the branch campuses and the depository are also shown separately. The institutional spreadsheet has information for all campuses together. Consult this sheet for statistics about a specific location within an institution. This spreadsheet has five tabs:

Holdings: Holdings at each location within a branch (and even sublocations within a location, if applicable - such as a reserves location code at a music library on a central campus)

General: Holdings and circulation information for each branch and its locations (does not go into details about sublocation codes)

Subjects: Information by subject area by location

Languages: Information by languages by location

Age: Information by age ranges by location

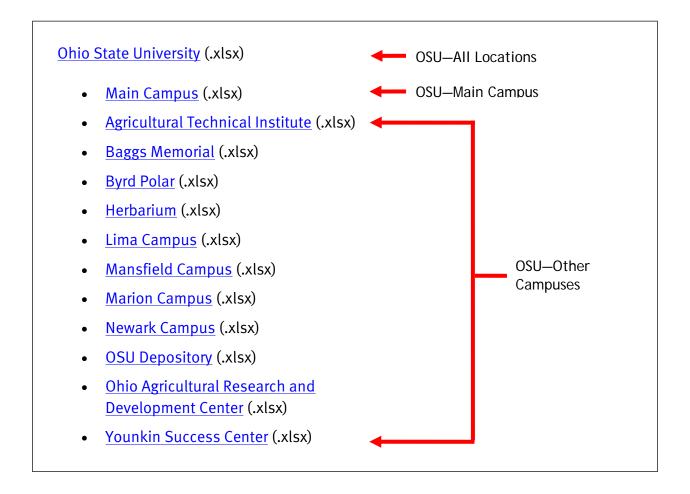


Figure 3. The institutional and campus spreadsheets for Ohio State University

In each tab, locations are listed vertically while the topic for the tab (Subjects, Languages, Age, etc.) is listed horizontally. Figure 4 presents an example, using Subjects as the topic.

		C D E	N2	OA	OB	OC.	00	30	OF	0G
	1			History & G	cography		History of the	Science 4	neral, Scholars	ship, Learning
	2			Sunnary				(A)	2)	10 To 25 10 10 10 10 10 10 10 10 10 10 10 10 10
	3	Library Unit	All Items	Circ. Items A	Annual Circ	Circ. Rate	All Items	Circ. Items	Annual Circ	Circ. Rate
	4			125,315	22,866	0.182	251	146	37	0.250
	5	University Libraries	167,266	122,304	22,580	0.184	247	142	37	0.26
_	6	Thompson Library	155,624	113,801	21,346	0.188	226	126	31	0.246
Branches	7	Architecture Library	981	830	167	0.201	3	2	5	2.500
	- 8	Archives	29	0	0		0	0	0	
	9	Biological Sciences/Pharmacy	57	46	10	0.217	1	1	0	0.000
	10	Business	747	740	51	0.069	4	4	0	0.000
	11	Cartoon Library and Museum	1,165	1,165	36	0.031	0	0	0	
	12	Center for Epigraphical Studies	1,020	11	0	0.000	0	0	0	
	13	Fine Arts	6,115	5,322	742	0.139	5	3	0	0.000
	14	Agriculture	379	362	59	0.163	3	3	0	0.000
A	15	Goology	674	232	62	0.267	0	0	0	
4	16	Music and Dance	243	217	16	0.074	3	3	1	0.333
_	17	Science and Engineering	203	175	31	0.520	2	0	0	
· 11	18	Stone Lab	19	0	0		0	0	0	
croll	19	Vet Med	4	3	0	0.000	0	0	0	
ertically	20	Health Sciences	62	14	1	0.071	0	0	0	
,	21	Nationwide Children's Hospital Libra	ry 4	2	0	0.000	0	0	0	
for more	22	Prior Health Sciences Library	58	12	1	0.083	0	0	0	
ocations	23	University Hospitals East Library	0	0	0		0	0	0	
ocations	24	Morite Law Library	3,561	2,397	285	0.119	4	4	0	0.000
	25	Agricultural Technical Institute	318	310	79	0.255	2	2	0	0.000
	26	Ohio State (Mansfield)	8,199	7,845	1,113	0.142	7	4	8	2.000
	27	Ohio State (Marion)	7,285	7,046	650	0.092	12	3	0	0.000
	28	Ohio State (Newark)	6,875	6,149	1,166	0.130	6	4	4	1.000

Figure 4. Example of institutional spreadsheet (Subjects analysis).

The institutional spreadsheets can be very large and thus a bit difficult to navigate. Only a small proportion of the spreadsheet can be viewed at one time. To view all of included statistics requires both horizontal and vertical scrolling. The Subjects statistics are presented in four columns: the total number of items, the number of circulating items, the annual circulation, and the circulation rate.

The General tab includes general statistical data. The General tab for The Ohio State University is shown in Figure 5.

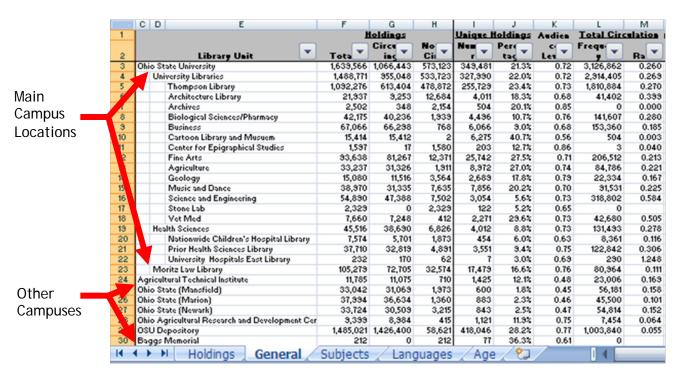


Figure 5. Example of an institutional spreadsheet (General tab).

Campus-Specific Spreadsheet

Each campus-specific spreadsheet presents information for an individual "campus." Since these sheets are simpler, information is displayed vertically. Figure 6 presents Ohio State's Marion campus as an example.

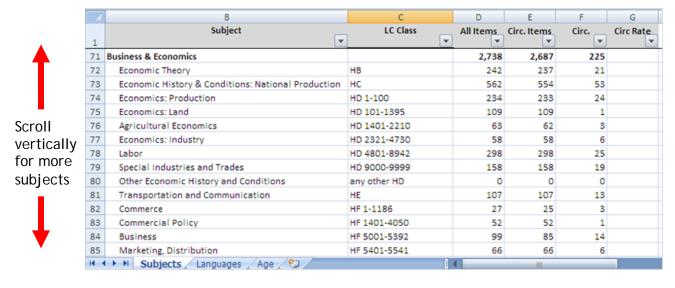


Figure 6. Example of campus spreadsheet (OSU Marion).

Data Overview: Smaller Institutions

For single-campus institutions without branch libraries, a single spreadsheet is used to present the same information as the two sheets used for multi-campus institutions since there is no distinction between the institution and the campus.

Holdings: For small institutions with only one main facility, various sub-locations are combined into one overall statistical presentation. The holdings tab will indicate which locations were used and whether they were designated as circulating or non-circulating. All of these locations are totaled in the top row for the composite institutional profile.

Sub-locations are listed below, vertically, in the *Unit ID* column and are coded Level 4. Statistical amounts pertaining to each are displayed horizontally. Exceptions to circulating sub-locations or collections are indicated in the *Excp.* column with the appropriate locally-assigned code.

Total Circulation refers to the cumulative number of circulation events recorded in the system as of 2008. (This usually is several years of data.)

General: Figures here are the total of all the locations. *Unique Holdings* means titles unique to the OhioLINK collection. Calculations used for frequency and rate are shown in the Glossary.

Languages: This is a summary of all locations. *Circ Rate* is derived by dividing *Circ.* by *Circ. Items*.

Age: These results are a summary of all locations. The *Average Age* is based on the imprint in the bibliographic records.

Subjects: Subject headings in this report are based on the North American Title Count project (NATC 1998). This is a summary of all locations.

Further information for interpreting the results and the definition of the terms can be found in the Glossary.

Collective Analysis of OhioLINK Libraries

Rather than emphasizing the separate collections, the collective analysis views the OhioLINK library collections as a whole and examines the overall characteristics of the collective collection and how those resources are distributed. Statistics for the collective analysis are available on the project website at:

http://www.oclc.org/research/activities/ohiolink/collective.htm

Subject Analysis

The subject analysis included a detailed analysis of over five hundred specific subjects derived from the North American Title Count identified by their Library of Congress Classification number. The analysis is presented by campus. Figure 7 presents a sample analysis for one of the subjects, Circuses and carnivals (GV 1800-1860). Branch campuses are listed separately from main campuses. All campuses holding at least 25 manifestations in the subject area are listed in descending order of comprehensiveness. Summary information for all of the system resources is shown on the bold subject row.

The Analysis of OhioLINK Subject Collections is grouped into nine broad subject groups:

- 0. Genre
- 1. Arts & Recreation
- 2. Business & Economics
- 3. History & Geography
- 4. Language & Literature
- 5. Science & Technology
- 6. Social Science
- 7. Medicine
- 8. Law

The first group, *Genre*, includes books that are categorized primarily by their type or genre rather than by their subject. This group includes encyclopedias, indexes, almanacs, bibliographies, and other similar materials.

The detailed analysis of individual subject areas is shown within the nine broad subject areas roughly in the order of their corresponding Library of Congress Classification Number. Because this spreadsheet contains nearly 32,000 rows with listings for 558 separate subjects, navigation can be difficult. By default, the subjects are grouped into the nine broad groups described earlier. To assist in navigation, an index to the detailed analysis is provided. A portion of the index for law is shown in figure 8. The entries in the index are linked to the portion of the analysis for that subject which be accessed by clicking on the name of the subject. The complete index is included in the subject spreadsheet which is available on the Collective Results Page (see link above).

Е	F	G	Н	1	K	L	M	N	0	Р	Q	R
벟	Institution	Number of	No. of	Number	Percent	Annual	Circulation	Percent	Comprehen-	Average	Percent	Audience
Subject		Works	Manifes-	of Items	Non-	Circulation	Rate	Coverage	siveness	Age	English	Level
S			tations		circulating							
] [•	~	•	•	~	•	~	~	~	•	•
Circu	ses, Carnivals [GV 1800-1860]	673	784	2,631	16.0%	521	0.236	100.0%	1.000	36.9	91%	0.37
	Bowling Green State University	306	326	357	57.7%	74	0.490	41.6%	0.743	25.7	98%	0.29
	Kent State University	151	157	177	10.2%	51	0.321	20.0%	0.478	40.3	98%	0.39
	Miami University	79	81	83	10.8%	45	0.608	10.3%	0.461	27.7	100%	0.32
	University of Cincinnati	89	89	90	4.4%	24	0.279	11.4%	0.399	23.1	99%	0.36
	Case Western Reserve University	67	71	77	6.5%	29	0.403	9.1%	0.367	33.1	97%	0.43
	University of Akron	98	100	107	7.5%	18	0.182	12.8%	0.363	29.6	100%	0.31
	Ohio State University	98	100	108	20.4%	15	0.174	12.8%	0.359	30.7	89%	0.46
	Oberlin College	54	56	56	3.6%	20	0.370	7.1%	0.344	38.7	98%	0.45
	Ohio University	99	101	104	5.8%	23	0.235	12.9%	0.338	36.2	97%	0.39
	Youngstown State University	60	61	61	3.3%	9	0.153	7.8%	0.261	34.4	100%	0.40
	University of Toledo	48	49	51	7.8%	20	0.426	6.3%	0.248	32.2	98%	0.38
	University of Dayton	41	41	43	9.3%	6	0.154	5.2%	0.246	33.0	98%	0.41
	Wright State University	27	27	27	3.7%	14	0.538	3.4%	0.232	16.7	100%	0.39
	Depositories	187	194	276	0.7%	15	0.055	24.7%	0.202	47.0	77%	0.49
	Cleveland State University	41	42	49	22.4%	2	0.053	5.4%	0.194	33.1	100%	0.34
	Kenyon College	30	30	30	0.0%	15	0.500	3.8%	0.186	30.0	97%	0.46
	Cuyahoga Community College	26	26	31	0.0%	2	0.065	3.3%	0.184	23.9	100%	0.36
	College of Wooster	31	32	32	0.0%	16	0.500	4.1%	0.167	30.1	97%	0.44
	Wittenberg University	35	36	37	0.0%	11	0.297	4.6%	0.165	41.9	97%	0.38
	Ohio Wesleyan University	26	26	26	7.7%	4	0.167	3.3%	0.161	23.7	100%	0.41
	Lakeland Community College	25	25	25	0.0%	8	0.320	3.2%	0.123	25.0	100%	0.32
	Hiram College	26	26	26	0.0%	4	0.154	3.3%	0.121	29.8	100%	0.36
	State Library of Ohio	65	65	76	30.3%	9	0.170	8.3%	0.104	42.0	100%	0.34
	Ohio Northern University	27	28	29	6.9%	3	0.111	3.6%	0.079	37.2	100%	0.36
	Heidelberg University	27	27	27	3.7%	3	0.115	3.4%	0.058	55.1	100%	0.42

Figure 7. Analysis of Circuses and Carnivals.

	Click on Subject for Detailed Analysis	LC Class
Law		
	International Law	JX
	Law (General)	K-KC
	Law: United Kingdom and Ireland	KD
	Law: Canada	KE
	Law: Canada - Provinces and Territories	KEA-KEY
	Law: U.S. (Federal)	KF
	Law: U.S. States and Territories	KFA-KFW
	Law: U.S. Cities	KFX
	Law: Latin America, West Indies	KG-KHW
	Law: Germany	KJ-KKZ
	Law: Other Jurisdictions	KL-KWX
	Law of Nations	KZ
	Copyright, Intellectual Property	Z 551-661
	<u>Unclassified Law</u>	

Figure 8. Sample portion of the index for law.

Within each subject area, the university or college campuses are listed by default in decreasing order of their comprehensiveness (see definition below) as seen in figure 9. However, the default order can be changed by sorting the entries. The sort dialog for default order (Comprehensiveness) is shown in figure 10.

International Law [JX]	17,303	20,826	128,154	9.8%	4,147
Ohio State University	5,735	6,042	8,007	28.0%	528
University of Cincinnati	4,910	5,094	6,839	5.4%	282
Depositories	9,975	10,920	24,081	4.3%	450
Ohio University	4,058	4,854	5,425	15.9%	228
Case Western Reserve University	5,418	6,386	8,189	26.6%	219
University of Akron	3,181	3,406	4,096	13.5%	139
University of Toledo	2,751	2,883	3,246	13.1%	125
Kent State University	4,225	4,413	5,082	2.5%	201
Oberlin College	3,399	3,863	4,558	2.3%	155
Miami University	2,904	2,961	3,247	40.1%	178
Cleveland State University	2,999	3,122	4,137	12.8%	116
Youngstown State University	2,140	2,589	2,839	1.1%	81
Ohio Wesleyan University	1,906	2,275	2,473	1.1%	93
College of Wooster	2,093	2,809	3,047	0.0%	87

Figure 9. Portion of the Subjects Spreadsheet for International Law

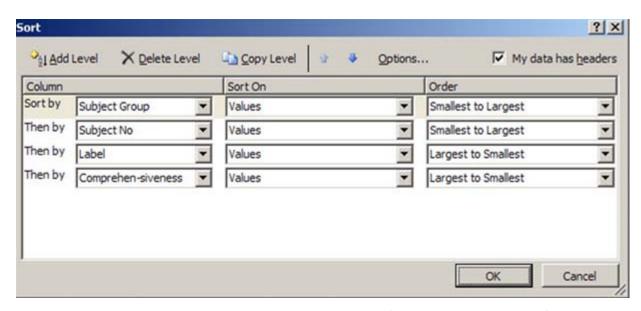


Figure 10. Sort dialog for the default sort order (Comprehensiveness).

To retain the order of the subjects, the first three levels should not be changed. For example, to order the table alphabetically by institution name, the fourth level would need to be changed from *Comprehensiveness* to *Institution*.

Language Analysis

The language analysis follows a pattern similar to the subject analysis. One hundred twenty-four different languages were analyzed. Since English is by far the dominant language, it was not included in the language analysis. In addition to the overall analysis, languages were also analyzed by broad subject areas. As was the case with subjects, the first sheet in the language spreadsheet is the index. A portion of the language index is shown in figure 11. The language index also includes the number of manifestations in the particular language/subject combination. An asterisk (*) indicates that no OhioLINK campus held ten or more manifestations in that language/subject combination. No results are provided for these combinations. The manifestation count for each language/subject combination is linked to the section of the spreadsheet (tab and cell) that presents the details underlying the manifestation count for that language and subject. Clicking on the manifestation count will take the user directly to the detail.

Index to Detailed Subject-Language Tables Click on Number of Manifestations for Details (*Detailed analysis unavailable; No libr Number of Manifestations by Subject Area											
Language / Subject Area	All Books	Arts & Recreatior	Business & Economic:	History & Geograph	Language &						
Achinese	<u>41</u>	*	*	*	<u>41</u>						
Afrikaans	<u>306</u>	*	*	<u>106</u>	<u>87</u>						
Akkadian	<u>32</u>	*	*	*	*						
Albanian	<u>117</u>	*	*	<u>18</u>	<u>77</u>						
Amharic	<u>99</u>	*	*	*	<u>19</u>						
Arabic	23,485	<u>526</u>	<u>709</u>	4,608	<u>9,126</u>						
Aramaic	<u>55</u>	*	*	*	*						
Armenian	<u>141</u>	<u>43</u>	*	<u>21</u>	<u>48</u>						
Austronesian (Other)	<u>264</u>	*	*	<u>15</u>	<u>129</u>						
Azerbaijani	<u>72</u>	*	*	*	<u>39</u>						

Figure 11. Portion of the Index for Languages.

As an example of the language analysis, the analysis of Danish is shown in figure 12. Only campuses holding at least ten manifestations are included in the tables. There were a total of 3,222 Danish language items held by OhioLINK libraries and half of the Danish items were in the depositories.

Figure 13 shows the statistics for Danish history and geography books. Only three campuses are shown since no other campuses held ten or more manifestations of Danish-language materials on history and geography. No analysis was done on language/subject combinations unless at least one campus held ten or more books.

В	С	D	Е	F	G	Н	I	J	K	L
	Language / Campus	Works	Number of Man.	Items	Circ. Items	Non-Circ. Items	Annual Circ.	Circ. Rate	Relative Circ. Rate	Average Age
Dani		1,797				12.2%	87	0.031	0.252	52.4
	Depositories	958	1,011			0.1%	8	0.005	0.041	57.1
	Ohio State University	270	306	466	368	21.0%	17	0.046	0.374	44.9
	University of Cincinnati	222	233	291	239	17.9%	11	0.046	0.374	63.2
	Oberlin College	171	181	216	195	9.7%	1	0.005	0.041	44.5
	Kent State University	96	101	106	84	20.8%	1	0.012	0.098	53.4
	Case Western Reserve University	66	66	89	65	27.0%	20	0.308	2.504	76.4
	Bowling Green State University	38	64	64	4	93.8%	4	1.000	8.130	40.8
	Ohio University	39	44	47	41	12.8%	6	0.146	1.187	28.2
	Miami University	27	28	32	20	37.5%	2	0.100	0.813	38.1
	University of Dayton	20	22	27	21	22.2%	0	0.000	0.000	60.2
	Columbus College of Art and Design	19	19	22	20	9.1%	3	0.150	1.220	14.0
	Wittenberg University	17	19	21	16	23.8%	0	0.000	0.000	68.3
	Wright State University	12	13	21	20	4.8%	0	0.000	0.000	18.0
	Cleveland Institute of Art	17	17	17	17	0.0%	5	0.294	2.390	28.7
	University of Akron	13	14	15	9	40.0%	1	0.111	0.902	35.3
	Ohio Wesleyan University	11	12	15	15	0.0%	0	0.000	0.000	60.9
	Dayton Art Institute	11	11	11	11	0.0%	0	0.000	0.000	28.3
	Antioch University	10	10	10	8	20.0%	0	0.000	0.000	64.2

Figure 12. Language Analysis for Danish.

В	С	D	E	F	G	Н	1	J	K	L
	Language / Campus	Number of	Number of	Number of	Circ. Items	Non-Circ.	Annual	Circ. Rate	Relative	Average
_		Works	Man.	ltems		ltems	Circ.		Circ. Rate	Age
~	_	_	_	_	_	~		_	_	_
Danis	Danish		222	390	371	4.9%	9	0.024	0.205	59.2
	Depositories	114	120	194	194	0.0%	3	0.015	0.128	64.5
	University of Cincinnati	65	65	99	90	9.1%	4	0.044	0.376	60.8
	Ohio State University	16	17	53	51	3.8%	2	0.039	0.333	40.5

Figure 13. Danish-Language Books on History and Geography.

Uses for the Data

The hope is that many potential uses for this data will be discovered, explored, and expanded upon by institutions across the state. Some of the uses of the data include:

- An overview of the quantity, breadth, and use of the OhioLINK state-wide collection is possible for the first time.
- Overview comparisons of institution-to-institution collection quantities, circulation rates, values, and subject class holdings are possible in relation to the statewide collection or to each other.
- Institutions of like size, academic level, organizational type, or subject nature might compare their data.
- The strengths and deficiencies of the statewide and local collections can be viewed by quantity, or circulation rate by LC subject class.
- A consideration of obsolescence trends in each discipline might guide purchasing versus rental decisions for e-book collections. For example: renting might be advisable for quickly obsolete computer science, medicine, and education books but purchases would be more advisable for history, language, and religion books with longer usevalue trends.
- Use the circulation rate by subject class data to identify high- and low-use collections for special collection management considerations.
- Areas with high duplication and low circulation rates might be targeted for deselection or moved to off-site storage.
- Use quantity by subject class across institutions to identify cooperative collection partners. Cooperative purchasing plans based upon similar collection interests by topic or publisher would diversify the collection through the intentional selection of more unique titles rather than unnecessary duplicates.

- The institutional comparative data will assist members interested in joint approval plan profiling.
- The circulation rates by subject area will inform subject-based decisions on state-wide purchase caps.
- Provide information on potential e-book purchases.

The analysis completed to date represents only a fraction of what can be done. Although the current project team will continue to study the data, it is recognized that to exploit its full potential, other expertise is needed.

Therefore the circulation records and the corresponding bibliographic records are being made available to download for study and research (OCLC 2011b). The data is available under an Open Data Commons Attribution license which allows the data to be used for study and research with attribution of works produced from the database. Uses other than study and research may be subject to additional limitations. OCLC and OhioLINK encourage other scholars and researchers to study and further analyze this data.

Conclusions

This study revealed that circulation rates vary widely by subject, language, institution, and age of the material. Limits on unnecessary duplication must logically include a differentiation by subject discipline. Rules regarding de-selection choices and depository pull criteria are inaccurate if applied uniformly across every discipline.

The academic richness and histories of the OhioLINK member institutions are reflected in the uniqueness of their library collections. Unique items are not limited to a few large institutions but are widely distributed across many different types of member institutions. The membership should avoid collection practices that homogenize the state-wide collection through unnecessary duplication.

Individual institution members commented with surprise on the low use of their non-English language collections. Further study is needed to discover potential causes and trends of these collections' usage patterns.

The most fascinating result of the study was a test of the "80/20" rule. Librarians have long espoused the belief that 80% of a library's circulation is driven by approximately 20% of the collection. The analysis of a year's statewide circulation statistics would indicate that 80% of the circulation is driven by just 6% of the collection.

The project data may ultimately change selectors' state-wide purchasing behavior. A careful consideration of the data should lead selectors at every institution to recognize that cooperative collecting behavior is not a luxury. Cooperative collection habits serve each institution's fiscal and curricular interests better than current selection habits because they result in increases in both depth and breadth of collections overall.

The data presents the OhioLINK Consortium members with critical information for beginning new cooperative ventures with confidence.

Glossary

Active Materials—Number: Number of books in the collection that are in non-circulating collections or have circulated since May 1, 2003.

Active Materials—Percentage: The proportion of the books in the collection that are active.

Accession Date: Either the date the item was acquired or the date the record for the item was entered into OhioLINK.

Annual Circulation: The total circulation generated by the collection for the year observed.

Annual Circulation—Rate: The average number of time books in the collection circulated during the year studied (2007-2008).

Audience Level: The type of audience for which the collection is most suitable (O'Neill, Connaway, and Dickey 2008). Most books in WorldCat have had their audience level estimated based on the type of libraries that have acquired the book. The audience level for an individual book ranges from maximum of one for very scholarly material to minimum of zero for picture books for preschoolers. The audience level for a collection is the average of the audience levels of the books in the collection. In general, collections with high audience levels will be more scholarly or research-focused than collections with lower audience levels.

Average Age: The average age of items in the collections as of May 2008 based on the book's publication date.

CBTF: Collection Building Task Force. A subcommittee of the OhioLINK Cooperative Information Resources Management Committee (CIRM) charged with identifying methods for eliminating collection duplication, encouraging cooperative purchases, and suggesting statewide collection-analysis procedures; all aimed at increasing the breadth and depth of the statewide collection.

Circulation: Included both check-outs and all renewals.

Circulation Rate: The average annual circulation for the circulating items in the collection or subset of collection being analyzed.

CIRM: Cooperative Information Resources Management Committee. An OhioLINK standing committee charged with cooperative collection development and resource management; reports to LAC.

Comprehensiveness: The proportion of all OhioLINK circulations that could have been met by the manifestations in a specified collection. For the University of Akron's *Ethics* collection:

- Collectively, all Ethics manifestations in OhioLINK libraries circulated 8,817 times
- Ethics manifestations held at the University of Akron accounted for 3,060 circulations, totaled across all libraries participating in the study¹

Therefore the comprehensiveness of Akron's ethics collection = 3,060 / 8,817 = 0.347. Note that the 3,060 circulations represent the total circulation from all OhioLINK libraries for the manifestations held by the University of Akron. Another way of viewing this is that comprehensiveness is the proportion of the year's circulation that would have been satisfied in Akron's collection.

Coverage: The percent of all OhioLINK manifestations held in a specified collection. Example: The University of Akron's *Ethics* collection:

- Collectively, OhioLINK libraries hold 13,308 different manifestations classified as Ethics
- 1,563 of these manifestations are held by the University of Akron

Therefore Akron's coverage of ethics = 1,563 / 13,308 = 11.7%.

Date of Last Use: The last date that the item circulated.

FRBR: Functional Requirements for Bibliographic Records. The format recommended by IFLA (International Federation of Library Associations and Institutions) to structure databases to reflect the relational aspects of resources and relate user needs with the bibliographic

¹ The total circulation from all OhioLINK libraries for the manifestations held by the University of Akron is not shown in the Collected Subjects table (http://fast.oclc.org/OhioLINK/Subjects.xlsx). It is calculated by summing the OhioLINK circulation for each manifestation held by the University of Akron.

description. Works, manifestations, and items were used for this study; expressions could not be identified from the information in the records and are not used in this study.

FRBR Work-Set Algorithm: Converts MARC21 databases to the FRBR model.

Holding: Item owned by specific institution or an item within a particular collection.

Item: A single example or copy of a manifestation. Each circulation record represents one single item.

Item No.: The number assigned to each item. Item numbers are unique within institutions but not across institutions.

ISBN: The International Standard Book Number.

LAC: Library Advisory Council. The oversight committee for OhioLINK standing committees. Approve and recommend policy and strategic initiatives; composed of library directors.

LC Classification: Library of Congress alphanumeric system for classifying materials.

Broad subject area: Eight broad subject classes used in this study are based on a variation of the LC Classification system and include: 1. Arts & Recreation, 2. Business & Economics, 3. History & Geography, 4. Language & Literature, 5. Science & Technology, 6. Social Science, 7. Medicine, 8. Law.

LCCN: Library of Congress Control Number.

Level: The level in the administrative hierarchy. The campus is at the top level (level 1). The next unit (such as a branch library) or a distinct administrative unit (such as a law or medical library) are level 2 units. Further subdivision would creates level 3 units, which usually are separate branch libraries. Level 4 units are distinct collections or locations within a library such as oversized, rare, reserve, etc.

Location Code: The code used in the OhioLINK circulation records identifying the location of the item.

Local Holdings: The items held by an individual library or branch library.

Manifestation: Represented by a unique OCLC number/bibliographic record. Each manifestation could have more than one item record, representing multiple copies.

Material Types: The study was limited to books and manuscripts. These were determined by the MARC21 material type codes of "a" or "t" and bib level of "m" in the bibliographic records.

Non-circulating Item: An item that is not allowed to circulate. As used in this study, an item was considered non-circulating if it couldn't be checked out by most members of the academic community (including undergraduate students).

Percent English: The percentage of the collection that consists of English language books.

OCLC: OCLC Online Computer Library Center, Inc.

OCLC No.: The number assigned to the manifestation in OCLC's WorldCat.

OCLC Research: OCLC's research division, which investigates trends in technology and library practice; identifies technological advances that enhance library services; and works with the community to collaboratively identify problems and opportunities, prototype and test solutions, and share findings through publications, presentations and professional interactions. Source: The institution supplying the circulation records.

Title: The title of the book or manuscript.

Unique Holdings: The number of the books in a collection that are not held by any other OhioLINK libraries.

WorldCat: A global network of library-management and user-facing services built upon databases of bibliographic and institutional metadata cooperatively produced and maintained by OCLC and its member libraries. WorldCat is the world's largest and richest database of bibliographic information.

WorldCAT Linking: Bibliographic records were matched by the OCLC number or, if that was not present, by the LCCN (LC control number) or ISBN.

Publications and Presentations

OCLC/OhioLINK Research Project Workshop Presentations

- Gilliland, Anne T. "The OhioLINK/OCLC Collection Analysis Project: A Preliminary Report." *Collection Management*, 33 (2008):161-172.
- Gammon, Julia, Anne T. Gilliland, and Edward T. O'Neill. "Preliminary Analysis." Presentation given 7 February 2008 at the OhioLINK Collection Building Workshop. http://platinum.ohiolink.edu/cbtf/oclcres/feb08.ppt.
- Gammon, Julia A., and Edward T. O'Neill. "Preliminary Analysis." Presentation given 13 April 2009 at the 2nd OhioLINK-OCLC Collection Analysis Project Workshop. http://platinum.ohiolink.edu/cbtf/oclcres/OCLCWorkshop09.ppt.
- O'Neill, Edward T., and Julia Gammon. "Building Collections Cooperatively: Analysis of Collection Use in the OhioLINK Library Consortium." In *Pushing the edge:* explore, engage, extend, : proceedings of the Fourteenth National Conference of the Association of College and Research Libraries, March 12-15, 2009, Seattle, Washington, ed. Dawn M. Mueller, 36-45. Chicago: Association of College and Research Libraries, 2009.
- Gammon, Julie, and Ed O'Neill. "Report on the OCLC/OhioLINK Circulation Study." OCLC Webinar. November 18, 2010. Streaming recording (59:40). Available in three formats:

M4V: http://www5.oclc.org/downloads/research/webinars/20101118ol.m4v
WMV: http://www5.oclc.org/downloads/research/webinars/20101118ol.wmv

OhioLINK—OCLC Research Project Workshop Breakout Session Notes

Notes are available from the February 7, 2008 and April, 13, 2009 workshops, in the external Appendix to this report. http://www.oclc.org/research/publications/library/2011/2011-06a.pdf.

References

Hickey, Thomas B., and Jenny Toves. 2005. FRBR Work Set Algorithm. Dublin, OH: OCLC Research. Accessed 16 December 2008,

http://www.oclc.org/research/projects/frbr/algorithm.htm.

IFLA (IFLA Study Group on the Functional Requirements for Bibliographic Records). 2009 Functional Requirements for Bibliographic Records: Final Report, September 1997, as amended and corrected through February 2009. Accessed 4 May 2011, http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records, p17-24.

Kent, Allen. 1979. *Use of Library Materials: The University of Pittsburgh Study*. New York: Marcel Dekker.

Knievel, Jennifer, Lynn Silipigni Connaway, and Heather Wicht. 2006. "Use of Circulation Statistics and Interlibrary Loan Data in Collection Management." *College and Research Libraries*, 67,1 (January): 35-49. http://crl.acrl.org/content/67/1/35.full.pdf+html.

NATC (North American title count: titles classified by Library of Congress and National Library of Medicine classifications, 1997). 1998. [Chicago, IL]: Association for Library Collections & Technical Services. OCLC: 248434674.

O'Neill, Edward T., Lynn Silipigni Connaway, and Timothy J. Dickey. 2008. "Estimating the Audience Level for Library Resources." *Journal of the American Society for Information Science and Technology* 59,13: 2042-2050.

OCLC (Online Computer Library Center, Inc.). 2011a. OCLC Boilerplate. http://www.oclc.org/about/styleguide/boilerplate.htm.

——. 2011b. OhioLINK Collection and Circulation Analysis—Circulation Data. Dublin, OH: OCLC Research. http://www.oclc.org/research/activities/ohiolink/circulation.htm.

OhioLINK. 2004. *Mission Statement. Collection Building Task Force*. (July). Accessed 17 June 2011, http://platinum.ohiolink.edu/cbtf/missionstatement.pdf.

——. 2005. *Guidelines for the Number of Copies per Title in the OhioLINK Collection*. Accessed 17 June 2011, http://platinum.ohiolink.edu/cbtf/numberofcopies.pdf.

——. 2010. OhioLINK: The Ohio Library and Information Network. http://www.ohiolink.edu/about/what-is-ol.html.