Date: 01/06/2007



Cooperative Collection Development: Can it really be that easy?

Glenda Lammers

OCLC, Inc. Dublin, Ohio, USA

Meeting: 125 Acquisition and Collection Development

Simultaneous Interpretation: No

WORLD LIBRARY AND INFORMATION CONGRESS: 73RD IFLA GENERAL CONFERENCE AND COUNCIL
19-23 August 2007, Durban, South Africa
http://www.ifla.org/iv/ifla73/index.htm

Abstract

Using the library holdings in the OCLC WorldCat database for ten libraries in the GAELIC (Gauteng and Environs Library Consortium) and forming a group analysis of the selected libraries using the WorldCat Collection Analysis service, these collections were analyzed for unique items in each of the individual libraries in the group and overlap among group members by language, format, publication date and subject area of the materials. Additionally, the aggregated holdings of the Consortia were compared to over one billion holdings from member libraries to determine materials uniquely owned by only one library. From the results, libraries can make decisions about what to discard from their collections, what to acquire for their collections and what to preserve and digitize in their collections.

Introduction

Cooperative Collection Development is often the goal of libraries, especially libraries belonging to consortia. Before libraries can engage in successful cooperative collection development, each need to know what the other owns, where there are duplicate titles in the collections and what is unique in each collection. In the past, there have often been

technological barriers in doing cooperative collection development, with some of the difficulties arising from disparate collections, differing integrated library systems and geography (location of collections).

Using the library holdings in the OCLC WorldCat database of ten institutions across the three provinces of Gauteng, North-West and Limpopo and the University of Namibia (all members of the Gauteng and Environs Library Consortium (GAELIC)), and forming a group analysis of the libraries using the WorldCat Collection Analysis service, their collections are analyzed. The analysis will determine unique items in each of the individual libraries in the group and overlap among group members by language, format, publication date and subject area of the materials. Much of the data provided in the analysis will support the GAELIC Shared Collection Storage Focus Group Charter. The purpose of this focus group is to investigate the feasibility of establishing a shared storage facility of physical collections.

Additionally, from the results, libraries will make decisions about:

- what to discard from their collections
- what to acquire for their collections
- what to preserve/digitize in their collections
- what to move to their shared storage collection

Cooperative Collection Development is often the goal of libraries, especially libraries with very specific collections and policies. Before libraries can engage in successful cooperative collection development, each need to know what the other owns, where there are duplicate titles in collections and what is unique in each collection. In the past, there have often been technological barriers in doing cooperative collection development, with some of the difficulties arising from disparate collections, differing integrated library systems and geography (location of collections).

OCLC WorldCat Collection Analysis service

Taking advantage of the business intelligence OCLC has gathered about libraries,

OCLC makes it possible for library staff responsible for collection management to analyze

the age and subject content of their own collections, compare their collections with those of peer libraries, and compare, as a group, the level of overlap or uniqueness of their collections.

The OCLC WorldCat Collection Analysis service allows libraries to cost –effectively and routinely evaluate their collections. It enables library staff to communicate collection decisions to faculty, boards of trustees and administrators, as well as demonstrate financial needs and responsible stewardship of library acquisitions, budgets and collections. This is particularly important in times of tight budgets and funding cutbacks.

The OCLC WorldCat Collection Analysis service is available for all libraries and all collections, regardless of the types of materials to be analyzed, the type of library or the integrated library system used. A library that subscribes to the OCLC WorldCat Collection Analysis service may compare its collection to any library that maintains holdings in the WorldCat database.

For groups of libraries and consortia, the WorldCat Collection Analysis service provides a time- and labor- saving process to evaluate group collections and validate cooperative collection activities. The service allows libraries to review collection gaps, the degree of overlap of their collections with others, and their uniquely held titles.

The WorldCat Collection Analysis service is incomparable in the library community because it leverages the cooperative effort of thousands of librarians around the world who have built—and are continuing to build—WorldCat, the world's richest database of items held in libraries.

In the past, collection analysis projects could take months to complete. OCLC can set up a project for a library within a week of ordering, depending on the scope of the job. Once the project is set up, collection analysis can be completed online, instantaneously, through WorldCat.

As a membership cooperative organization, OCLC always works very closely with its member libraries as it develops new services, and the OCLC WorldCat Collection Analysis service was no exception. From the beginning, library staff and administrators were very enthusiastic about its potential.

Collecting and Reporting

This project analyzes the collections of ten of the sixteen library sites in the GAELIC Consortium in South Africa. The collections included in the analysis are North-West University, Potchefstroom Campus; Tshwane University of Technology, Pretoria Campus; University of South Africa, Muckleneuk Campus; University of Botswana; University of Johannesburg, APK Campus; University of Limpopo, Medunsa Campus; University of Namibia; University of Venda; University of Pretoria and Vaal University of Technology. Using the OCLC symbol associated with each site, OCLC identified the holdings of each library. Data is then mined from OCLC and compiled for each library's collection.

The GAELIC Collections

The library collections from the ten sites vary greatly in size, with the smallest collection containing a little over 16,000 holdings to the largest, with over 900,000 holdings. The total number of holdings in this analysis is 1.89 million. The breakout of the data is represented in Figure 1.

The information is in an easy-to-understand format, with the column values exposing the publication date of the materials and the rows indicating the library. The numbers in the

cells represent the number of titles in each collection by publication date range or single date.

Another view of the data shows the collections by the subject of the materials and the publication date ranges. There are 30 broad subject areas within the subject hierarchy used in the service. Each division is further divided into over 7,000 more descriptive subjects. Some of the subjects included in the display are agriculture, art and architecture, business and economics, engineering and technology, geography and earth sciences, law, medicine, philosophy and religion, sciences and several others. Figure 2 provides this view.

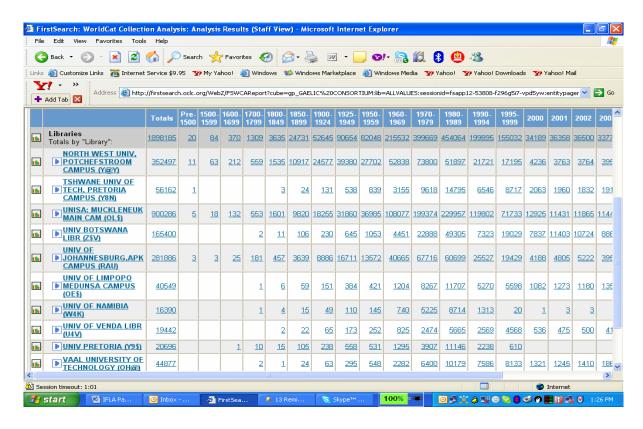


Figure 1: Each collection displayed by publication date.

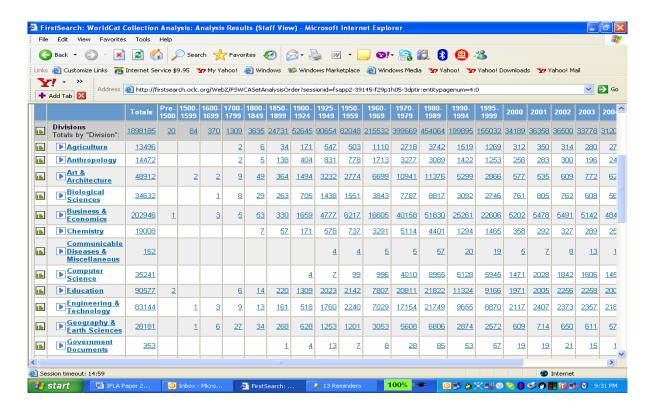


Figure 2: Collections arranged by publication date and subject.

A graphical view of the same data provides an illustration of the data that is easy to read, understand and present. Data in the graph shows actual number of titles by single year or ranges of years. The graph in figure 3 represents the collections of the ten libraries. It is easy to see that in the publication date range 1980-1989, there are about 450,000 titles in the collections representing nearly 24 percent of the materials. In the range 1970-1979, number of titles is nearly 400,000, representing 21 percent of the materials. In the range 1990-1994, there are about 200,000 titles representing 11 percent of the materials, and in the range 1995-1999, there are about 155,000 titles representing 8 percent of the materials.

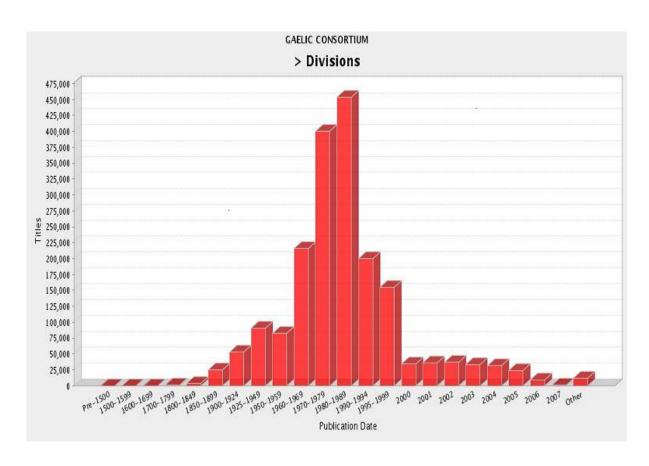


Figure 3: Graphical view of the collections by publication date and number of titles.

Looking at the age of the materials in the collection, 45 percent of the collection was published in the 1970's and 1980's, which on the surface may indicate an aging collection. Looking at the collection at the subject level, however, indicates much of these materials are in the subject areas of Anthropology, Art and Architecture, History and Auxiliary Sciences, Language, Literature and Linguistics, and Philosophy and Religion. Figure 4 provides a scoped view of the collections by subject and date range display.

Divisions	1960- 1969	1970- 1979	1980- 1989	1990- 1994	1995- 1999
Anthropology	11.84%	22.64%	21.34%	9.83%	8.66%
Art & Architecture	13.70%	22.37%	23.26%	10.83%	5.86%
History & Auxiliary Sciences Language, Linguistics &	15.18%	21.85%	20.56%	8.18%	5.13%
Literature	14.25%	22.21%	22.71%	9.27%	5.56%
Philosophy & Religion	13.95%	20.82%	21.18%	9.95%	7.03%

Figure 4: Scoped view of collections by subject and publication date.

Overlap and Uniqueness: The Overlap view of an analysis shows the number of titles one library holds in common with other libraries/institutions. Figure 5 shows overlap for each library in the group. The **bolded** numbers are the total holdings for the library indicated in each row. To interpret the amount of overlap from one library to another, simply locate the symbol in the column that corresponds with the comparisons library and move in from the row of the selected library. To illustrate, North-West University has 122,697 titles in common with UNISA (OL\$).

OVERLAP	Y@Y	Y8N	OL\$	Z\$V	RAU	OE\$	W4K	U4V	Y9\$	OH@
North-West Univ (Y@Y) Tshwane Univ of Tech	352497	12314	<mark>122697</mark>	23596	66711	5900	6330	5429	6710	8573
(Y8N)	12314	56162	19272	7861	11788	3371	2013	2685	2142	9261
UNISA (OL\$)	122697	19272	900286	69527	143471	11011	10582	9960	12701	14133
Univ of Botswana (Z\$V) Univ of Johannesburg	23596	7861	69527	165400	32130	4475	3922	4081	3429	5920
(RAU)	66711	11788	143471	32130	281886	6393	6236	5649	7446	8717
Univ of Limpopo (OE\$)	5900	3371	11011	4475	6393	40549	1248	1606	975	2444
Univ of Namibia(W4K)	6330	2013	10582	3922	6236	1248	16390	1266	1406	1400
Univ of Venda (U4V)	5429	2685	9960	4081	5649	1606	1266	19442	1027	1919
Univ of Pretoria (Y9\$)	6710	2142	12701	3429	7446	975	1406	1027	20696	1340
Vaal Univ of Tech (OH@)	8573	9261	14133	5920	8717	2444	1400	1919	1340	44877

Figure 5: Overlap among each collection in study.

Overlap can further be viewed by subject, language, format and publication date of materials in each library's collection. With further exploration of overlap, library's can determine where there are duplicate titles in multiple locations, which may assist in weeding projects, relocating materials throughout a consortium or moving to offsite storage and writing or updating collection development policies.

The Uniqueness view of an analysis shows how many of your library's titles are held by each library that no other library in the consortium own. Understanding what is uniquely held in the group provides important information about what may be preserved, digitized and moved to offsite storage for long-term housing. Figure 6 provides a high-level view of what is uniquely held by each library, and what titles are shared by several or all libraries in the

consortium. The one title held by each institution is <u>Organic synthesis</u>, the disconnection <u>approach</u> written by Stuart Warren. Based on holdings in all of WorldCat, 897 libraries own this title.

UNIQUENESS	Unique	Held by 2	Held by 3	Held by 4	Held by 5	Held by 6	Held by 7	Held by 8	Held by 9	Held by All
North-West Univ	197800	84642	46786	15979	5063	1660	435	110	21	1
Tshwane Univ of Tech	21640	15116	9363	5592	2804	1160	360	104	22	1
UNISA	620031	186213	66377	19416	5852	1807	458	109	22	1
Univ of Botswana	74535	49983	24697	11166	3536	1093	294	76	19	1
Univ of Johannesburg	106230	97718	53389	17074	5265	1652	427	108	22	1
Univ of Limpopo	21558	9249	4772	2654	1354	629	243	72	17	1
Univ of Namibia	2724	3839	3607	3174	1880	814	248	83	20	1
Univ of Venda	5262	4772	3818	2784	1657	773	281	78	16	1
Univ of Pretoria	4479	5497	4587	3451	1698	664	224	73	22	1
Vaal Univ of Tech	18187	12149	7187	4034	2031	896	292	83	17	1

Figure 6: Unique titles by library.

Another view of uniqueness (figure 7) shows in which subject the one shared title is found (Chemistry). The highest percentage of unique titles are found in the various medical subject areas, music and art & architecture. The subjects with the greatest overlap are in biological sciences, chemistry, math, and business and economics.

UNIQUENESS BY SUBJECT	Unique	Held by 2	Held by 3	Held by 4	Held by 5	Held by 6	Held by 7	Held by 8	Held by 9	Held by All
Agriculture	8303	3060	1302	600	145	72	14	0	0	
Anthropology	6930	3948	2148	1048	270	114	14	0	0	
Art & Architecture	32078	10618	3984	1508	545	150	21	8	0	
Biological Sciences	17133	8850	4587	2212	1210	426	140	56	18	
Business & Economics	94808	57910	30462	12856	4540	1620	539	184	27	
Chemistry	8130	5118	2991	1556	745	270	147	32	9	10
Communicable Diseases & Misc	97	32	15	8	10	0	0	0	0	
Computer Science	18679	9628	4476	1616	610	204	28	0	0	
Education	39223	25766	15006	6740	2665	876	245	56	0	
Engineering & Technology	46546	21228	9735	3868	1300	372	63	32	0	
Geography & Earth Sciences	13812	6902	4221	2004	850	294	56	24	18	
Government Documents	320	30	3	0	0	0	0	0	0	
Health Facilities, Nursing & History	340	88	84	52	35	30	28	8	0	
Health Professions & Public Health	1628	398	150	68	50	42	14	0	0	
History & Auxiliary Sciences	79175	32036	15927	5808	1680	660	161	24	0	
Language, Linguistics & Literature	183422	70078	33441	11144	4125	1308	469	80	45	

Law	43610	19314	8457	3132	1205	702	245	64	9	
Library Science, Generalities	31996	15820	8376	3268	1245	546	133	40	27	
Mathematics	21433	14046	7737	2916	875	258	77	24	9	
Medicine	55703	28676	13485	5640	2345	948	259	80	0	
Medicine By Body System	365	54	30	8	5	0	0	0	0	
Medicine By Discipline	940	236	111	40	10	12	0	8	0	
Music	33888	6004	1023	236	60	12	7	0	0	
Performing Arts	8739	2984	1119	344	135	42	28	8	0	
Philosophy & Religion	106950	45120	15762	3920	1125	246	28	0	0	
Physical Education & Recreation	6011	2332	1272	432	190	36	14	0	0	
Physical Sciences	19679	9718	4584	1744	670	264	77	32	0	
Political Science	26341	14170	7257	2524	825	270	98	32	0	
Preclinical Sciences	478	86	33	24	5	0	7	0	0	
Psychology	15603	10396	5712	2264	590	294	49	24	18	
Sociology	46811	27614	14631	6112	2335	750	217	64	9	
Unknown Classification	103275	16918	6462	1632	740	330	84	16	9	

Figure 7: Uniqueness and Overlap by Subject

Another important comparison for the GAELIC Consortium to view is called WorldCat Comparison. This comparison shows what is uniquely held by one GAELIC Consortium library that no other library is all of WorldCat own. Knowing this information is critical for all libraries to help determine what may be considered a "last copy". Libraries need to be cognizant of this so appropriate measures can be taken to preserve last known items. Figure 8 is a high-level overview of the Consortium's unique items as indicated from WorldCat.

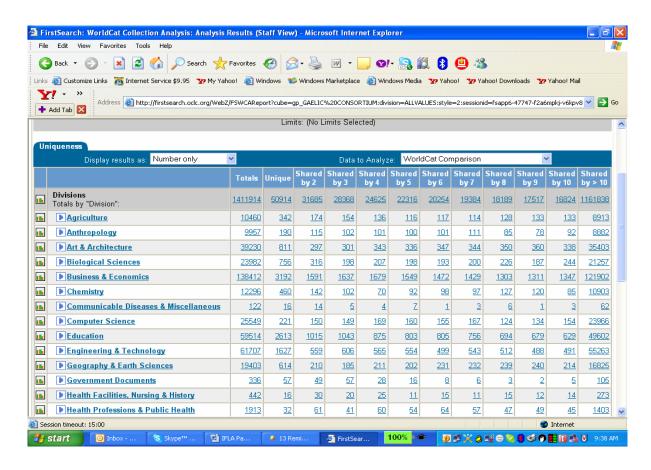


Figure 8: Uniquely held items as compared to all WorldCat libraries and holdings.

Nearly 51,000 holdings are held by only one library in the GAELIC Consortium AND by no other library as represented in the collections of libraries found in WorldCat. Additionally, over 31,000 titles are held by a GAELIC library AND one other library in WorldCat. The "Shared by" numbers indicate the relative rarity of library holdings.

Language Analysis. While the majority of the materials in the GAELIC Consortium libraries are published in English, there are significant other languages represented in the collections. North-West University has considerable titles in Afrikaans, Dutch and German; UNISA has considerable titles in Dutch, French, German, Italian; while University of Johannesburg has Dutch and German language materials. Figure 9 provides a summary of all languages represented in the Consortia collections, by library.

	Y@Y	Y8N	OL\$	Z\$V	RAU	OE\$	W4K	U4V	Y9\$	OH@
Afrikaans	31773	1989	10392	73	7334	260	1385	422	1402	749
Albanian	0	0	2	1	1	0	0	0	0	0
Arabic	21	3	1127	8	65	1	0	0	1	0
Armenian	0	0	7	0	1	0	0	0	0	0
Bengali	0	0	4	0	1	0	0	0	0	0
Bulgarian	2	1	23	0	1	0	0	0	0	0
Catalan	1	0	38	0	2	0	0	0	0	0
Chinese	7	1	687	5	5	3	1	1	0	1
Croatian	0	0	20	0	0	0	0	0	0	0
Czech	5	0	101	0	10	0	0	0	0	0
Danish	62	1	129	1	29	0	0	1	1	0
Dutch	32195	89	22985	34	11314	44	372	80	520	39
English	238186	53642	684450	163718	223542	40045	13519	18855	17496	43890
Finnish	0	0	217	1	6	0	0	0	0	1
French	6519	55	28488	1051	9518	42	54	16	242	40
German	34734	144	94278	141	25384	79	997	21	902	68
Greek, Modern [1453-]	158	1	255	0	225	0	0	0	1	0
Gujarati	1	0	1	0	0	0	0	0	0	0
Hebrew	98	0	3380	3	56	1	1	0	7	0
Hindi	2	0	9	0	2	0	0	0	0	0
Hungarian	13	0	103	1	5	1	0	0	0	0
Indonesian	8	0	9	0	3	0	0	0	0	0
Italian	406	25	12022	36	1628	6	4	1	10	13
Japanese	10	5	44	2	11	3	0	1	0	4
Korean	1	0	5	1	0	0	0	0	0	0
Latin	1627	7	4112	11	585	0	1	1	29	2
Macedonian	0	0	1	0	0	0	0	0	0	0
Malay	0	0	1	0	1	0	0	0	0	0
Marathi	0	0	2	0	0	0	0	0	0	0
Norwegian	23	1	89	0	6	0	0	0	0	4
Other	6248	162	16983	248	1773	48	52	41	77	50
Panjabi	0	1	4	0	0	0	0	0	0	1
Persian	0	0	120	0	1	0	0	0	0	0
Polish	8	0	254	5	12	2	0	0	2	0
Portuguese	137	14	5594	13	74	7	3	1	2	8
Romanian	3	0	83	0	6	0	0	0	0	0
Russian	40	4	5371	15	34	1	0	0	0	1
Sanskrit	2	0	37	0	2	0	0	0	0	0
Serbian	0	0	9	0	0	0	0	0	0	0
Slovak	0	0	3	0	1	0	0	0	0	0
Slovenian	0	0	6	1	0	0	0	0	0	0
Spanish	158	17	8508	23	219	5	0	1	3	3
Swedish	42	0	239	6	23	1	0	0	0	2
Tamil	0	0	3	0	1	0	0	0	0	0
Telugu	0	0	1	0	0	0	0	0	0	1
Thai	0	0	7	0	0	0	0	0	0	0
Tibetan	0	0	5	0	0	0	0	0	0	0
Turkish	3	0	22	2	1	0	1	0	0	0
Ukrainian	1	0	12	0	0	0	0	0	1	0
Urdu	0	0	7	0	0	0	0	0	0	0
Vietnamese	1	0	4	0	0	0	0	0	0	0
Yiddish	2	0	33	0	4	0	0	0	0	0

Figure 9: Language of materials by library.

Some languages identified in the "Other" language category include, but are not limited to: Akkadian, Aramaic, Middle and Old French, Middle High German, Middle and Old English, Ancient Greek and Swahili.

Format Analysis. While the majority of the materials in the GAELIC Consortium libraries are printed books, there are significant other formats represented in the collections. Microforms, serials (both print and electronic), sound recordings and visual materials are included in the collections. Figure 10 provides a summary of the formats represented in the collections, by library.

	Y@Y	Y8N	OL\$	Z\$V	RAU	OE\$	W4K	U4V	Y9\$	OH@
Formats	352497	56162	900286	165400	281886	40549	16390	19442	20696	44877
Articles	118	3	150	8	82	0	1	1	2	1
Books	330555	53953	850115	162318	275706	38010	16047	19004	20310	43719
Books, Microform	4459	25	1623	87	144	19	4	6	7	11
Computer Files	142	56	235	132	75	65	1	9	10	63
Maps	192	54	655	75	138	10	16	19	37	35
Musical Scores	4327	131	9976	6	25	2	5	26	5	2
Serial Publications Serial Publications,	10411	1270	22754	1080	5331	2231	307	323	263	815
Electronic Serial Publications,	289	18	298	24	35	18	2	1	0	9
Microform	62	5	136	7	30	9	2	4	0	3
Sound Recordings	1439	52	11902	185	24	2	1	5	8	11
Texts, Electronic	32	13	858	311	51	16	0	5	0	22
Visual Materials	462	581	1563	1165	243	167	3	39	54	184

Figure 10: Formats of materials by library

How to Use the Reported Data and Questions it Answers

Library decision-making. Using the data mined from WorldCat as business intelligence, the GAELIC Consortium can use the data to discover a variety of things about their collections. This allows library staff to make informed decisions about:

- Digitization and preservation based on
 - o publication date of materials
 - unique holdings

- low overlap
- Weeding based on
 - publication date
 - significant overlap
 - identified areas not supported by library
- Acquisitions based on
 - gaps determined in collection
 - items held by peer libraries
 - o publication date of materials available
- Academic program reviews by
 - benchmarking against peer collections
 - benchmarking against libraries strong in particular areas of study
 - benchmarking against bibliographies and authoritative lists
- Requests for funding by
 - o comparing with similar institutions to determine collection gaps
 - o looking at trends throughout the years by publication date of materials
 - identifying language and formats of materials not well-represented in collections

Summary

For Libraries that are or will be challenged to cooperatively develop collections because of limited budgets, changes in programs or changes in collection policies, must know what they hold, what their peers hold, what is unique to one collection in the group and what is duplicated throughout the collections. Once this information is known, steps can be taken to reduce cross-collection overlaps, freeing up library budgets for acquisitions of unique materials and to fill collection gaps. Libraries can make their informed collection decisions based on the intelligence provided.