GLOBAL COLLECTIVE RESOURCES:
A Study of Monographic Bibliographic Records in WorldCat

Report of a Study conducted under the auspices of an OCLC/ ALISE 2001 Research Grant

by

Anna H. Perrault
Associate Professor
School of Library and Information Science
University of South Florida

July 2002

GLOBAL COLLECTIVE RESOURCES

Abstract

In 2001, WorldCat, the primary international bibliographic utility, contained 45 million records with over 750 million library location listings. These records span over 4,000 years of recorded knowledge in 377 languages.[1] Under the auspices of an OCLC/ALISE research grant, a bibliometric study was conducted of WorldCat. A 10% systematic random sample of the database was analyzed utilizing the OCLC iCAS product to profile the monographic bibliographic records in WorldCat by type of library, subject, language, and publication date parameters. The profile details the Ainformation commons@ of global publication made accessible through the OCLC international network.

There were 3,378,272 usable records from the 10% systematic random sample of which 2,199,165 records had call numbers and could be analyzed by subject. Five types of library groupings were established for the study: research, academic, public, special, and school. The research libraries grouping has the largest number or records in the sample with call numbers at 1,745,034. The missions of the different types of libraries can be discerned in the subject profiles for each library grouping.

Among the findings of the study are that the profile of WorldCat by time period and by subject divisions is mirrored in the profile of the grouping of research libraries. Of all of the records in the 10% sample, approximately 65% are English language materials with 35% for foreign language materials. The analysis by number of unique records and title overlap demonstrate that the universe of materials under bibliographic control in WorldCat shows a high level of diversity of resources with 53% of records having only one library location symbol. The number of records in the analysis show a sharp decline by most measures from 1992 to the last imprint year in the study.

An analysis was performed of the records in the sample with ISBN numbers, finding that only 21% of the 3 million plus records in the study had ISBN numbers. This can be due to the amount of retrospective titles published before the numbering system came into use and also the number of publications that are not from mainstream publishers. But for publications since 1970, 57% of all records with call numbers have ISBN numbers, leaving an intriguing 43% of records with call numbers that do not have ISBN numbers.

The findings establish that WorldCat is a rich resource for cataloging records, verification of the existence of titles, and identifying prospective materials for resources sharing. As OCLC continues to implement its Global Strategy, AExtending the Cooperative,@ the number of international members and thus foreign language records and unique titles may continue to increase.

---

[1] *OCLC Newsletter* Jan/Feb. 2001 No. 249, p.6-7.

GLOBAL COLLECTIVE RESOURCES

Table of Contents*

Abstract
Table of Contents
List of Tables and Figures
Acknowledgments
About the Author

Chapter One: *Global Collective Resources*
- **Introduction**
- **The Problem**
- **Review of Related Research**
- **Methods**

Chapter Two: *WorldCat: The Profile*
- **The Sample**
- **Subject Analysis**
- **Summary**

Chapter Three: *Library Groupings*
- **Research Libraries**
- **Academic Libraries**
- **Special Libraries**
- **Public Libraries and School Libraries: Bibliographic Records by Audience Level**
- **Public Libraries Subject Analysis**
- **School Libraries Subject Analysis**

Chapter Four: *Diversity of Resources*
- **Unique Titles**
  - **WorldCat and Research Libraries**
  - **Library Groupings**
- **Title Overlap**

Chapter Five: *Language Analysis*
- **English, Non-English**
- **Foreign Language Groupings**
- **Subject Analysis for the Seven Language Groupings**

Chapter Six: *ISBN Analysis*

Chapter Seven: *Summary and Conclusions*

- **Findings**
  - **Decline in the Number of Records for Current Years**
- **Implications**
  - **Further Research**
  - **Conclusion**

Site Links

- **OCLC list of libraries http://www.oclc.org/contacts/libraries/**
- **OCLC iCAS http://www.oclc.org/western/products/aca/icas.htm**
- **OCLC Annual Reports http://www.oclc.org/about/annualreport/**

***Organization of the Report**

This report has been designed as a web document with hyperlinks for easy navigation. Each chapter is self-contained with attached tables and figures. Chapters can be reached through the hyperlinks in the Table of Contents and also through the links at the end of each chapter to the next chapter. Each table or figure is hyperlinked from the title within the text to the attached table or figure. In chapter one, the references are hyperlinked from the superscript number to the notes at the end. Chapters with a small number of references have the references in footnotes on that page.

GLOBAL COLLECTIVE RESOURCES

List of Tables and Figures

Chapter Two (http://www.oclc.org/research/grants/reports/perrault/chapter2final.pdf)

Table 2-1 WorldCat Monographic Bibliographic Records: All Titles Held by Date (All Subject Divisions)

Table 2-2 WorldCat Monographic bibliographic Records: Percentage Increase/Decrease by Time Period

Table 2-3 WorldCat Records: Subject Divisions, 50 Year Range

Table 2-4 WorldCat Records: Subject Divisions, 10-Year Range

Chapter Three (http://www.oclc.org/research/grants/reports/perrault/chapter3final2.pdf)

Table 3-1 Library Groupings: Total Number of Bibliographic Records from Subject Analysis

Table 3-2 Research Libraries: Percentage Increase/Decrease by Years

Table 3-3 Research Libraries: Percentage Increase/Decrease by Years

Table 3-4 Academic Libraries: Percentage Increase/Decrease by Years

Table 3-5 Academic Libraries: Percentage of Total by Subject Division

Table 3-6 Special Libraries: Percentage Increase/Decrease by Years

Table 3-7 Special Libraries: Percentage of Total by Subject Division

Table 3-8 Public Libraries: Percentage of Adult/Juvenile

Table 3-9 Public and School Libraries by Years

Table 3-10 School Libraries: Percentage of Adult/Juvenile

Table 3-11 Each Library Group by Subject Division

Chapter Four (http://www.oclc.org/research/grants/reports/perrault/chapter4final.pdf)

Table 4-1 WorldCat Unique Records: All Titles Held by Date (All Divisions)

Table 4-2 Unique Records Research Libraries: All Titles Held by Date (All Divisions)

Table 4-3 Total Number of Unique Bibliographic Records: WorldCat and Research Libraries

Table 4-4 Total Number of Unique Records by Decade: WorldCat and Research Libraries

Table 4-5 Unique Records by Subject Division: WorldCat and Research Libraries

Table 4-6 Unique Bibliographic Records From Subject Analysis by Library Grouping

Table 4-7 Unique Records by Subject by Library Groupings

Table 4-8 All Libraries: Title Overlap Between Groups

Table 4-9 Total Number of Shared Records by Library Grouping

Figure 4-1 WorldCat and Unique Research Record Comparison

Chapter Five (http://www.oclc.org/research/grants/reports/perrault/chapter5final.pdf)

Table 5-1 WorldCat: English and All Non-English as a Percentage of Total Subject Records

Table 5-2 Research Libraries: English and All Non-English as a Percentage of Total Subject Records

Table 5-3 Academic Libraries: English and All Non-English as a Percentage of Total Subject Records

Table 5-4 Increase/Decrease by Decades for English/non-English Records: Academic and Research Libraries

Table 5-5 Increase/Decrease by Five-Year Periods, 1980-1999 for English/Non-English Language Groupings: Academic and Research Libraries

Table 5-6 Foreign Language Groupings as a Percentage of Total Records

Table 5-7 Foreign Language Groupings: Number of titles by Time Period

Table 5-8 Language Records 1985-1999: Academic and Research Libraries

Table 5-9 WorldCat: Language Titles by Subject Division

Chapter Six (http://www.oclc.org/research/grants/reports/perrault/chapter6final.pdf)

Table 6-1 WorldCat ISBN Analysis: All titles Held by Date (All Divisions)

Table 6-2 WorldCat English ISBN

Table 6-3 WorldCat Foreign ISBN

Table 6-4 WorldCat, English, and Foreign ISBN Analysis by Time Period

Table 6-5 Comparison of Total WorldCat Records with ISBN Totals

Table 6-6 WorldCat ISBN Tables: Records With Call Number Present by Subject

Chapter Seven (http://www.oclc.org/research/grants/reports/perrault/chapter7final.pdf)

Table 7-1 and Figure 7-1 Annual Increase/Decrease in Number of WorldCat Records 1990-2000

Table 7-2 and Figure 7-2 Annual Increase/Decrease in Number of Records 1990-2000 Academic and Research Libraries

Table 7-3 and Figure 7-3 WorldCat and Research Libraries Unique Records 1990-2000

Table 7-4a and Figure 7-4a WorldCat English and Non-English Language Records 1990-2000

Table 7-4b and Figure 7-4b Research Libraries English and Non-English Language Records 1990-2000

Table 7-4c and Figure 7-4c Academic Libraries English and Non-English Language Records 1990-2000

Table 7-5 and Figure 7-5 WorldCat Records with ISBN Numbers

Table 7-6 and Figure 7-6 WorldCat, English, and Foreign Language Records with ISBN Numbers

Acknowledgments

A number of people contributed to the success of this project. Sally Loken of WLN and Ed O=Neill of the OCLC Office of Research both lent their support to the idea for the project when I first broached it to them. The project was endorsed and supported by the administration of the Lacey Product Center and the staff who produce the iCAS products, including Scott Barringer, Paul Brogger, Eric Kraig, Will Ryan, Ann Marie Wehrer and Glenda Lins.

The University of South Florida granted me a full semester of sabbatical leave in Fall 2001. The analysis of data began during that time. The support and cooperation I received from the Director of the School of Library and Information Science at the University of South Florida, Vicki L. Gregory, and my colleagues are greatly appreciated. Graduate assistants Jennifer Boucher and Monica Jenkins took an interest in the project and helped with the data analysis. Rich Austin, also of USF, readied the manuscript for the web. I am thankful to all of these people.

About the Author

Anna H. Perrault is an Associate Professor in the School of Library and Information Science at the University of South Florida in Tampa. Her research in collection analysis and assessment has been frequently cited and has won several awards including the ALCTS/Blackwell=s Scholarship Award, the LAPT Research Award, and an OCLC/ALISE Research Grant. Dr. Perrault has conducted collection analysis and assessment projects in Louisiana and with the College Center for Library Automation (CCLA), the network for community colleges in Florida. She has conducted collection assessment workshops throughout the Southeast for SOLINET. Perrault is a member of the Center for Research Libraries/Big 12 Plus Working Group to develop measures for quantifying, evaluating and maximizing the economic benefits of coordinated cooperative collection development projects.

A complete vita and publication bibliography including an Impact statement can be accessed at the web site below.

Contact information:
Anna H. Perrault perrault@chuma1.cas.usf.edu
(813) 974 6844 FAX (813) 974 6840
http://www.cas.usf.edu/lis/faculty/perra.html

Chapter One

GLOBAL COLLECTIVE RESOURCES

Introduction

In 2001, OCLC staff and member libraries, and many librarians and information specialists worldwide, celebrated the 30th anniversary of WorldCAT. OCLC was founded in 1967 as the Ohio College Library Consortium. The Union Catalog, later renamed WorldCAT, was founded in 1971. WorldCAT is now the largest and most comprehensive database of bibliographic information in the world. Through the merged catalogs of libraries around the world, resources are made available to libraries and their users that no single library could provide alone.

For the first 20 years of its existence WorldCAT was primarily a cataloging database also heavily utilized for interlibrary loan. Beginning in the latter 1980s, reference and other database services were added. By the year 2001, there were 37,297 libraries in 76 countries participating in the OCLC Library Cooperative. As of June 30, 2000, there were 42,476,614 total records in WorldCAT, with 767 million location listings for those materials. Of these bibiographic records, 35,570,633 or 83.7% were for books with the remaining 16.3% for other formats such as serials, maps, scores, sound recordings, computer files, and mixed materials. The language distribution at that time was 26,869,669 records for English language materials with the remainder divided among 376 languages.[1] These records are for materials spanning four millennia of recorded knowledge, from 2,000 BC to the present.

In 1999, WLN became a division of OCLC and the collection assessment and analysis programs of WLN became OCLC products. This study profiles the monographic

bibliographic contents of WorldCAT utilizing the OCLC/WLN iCAS software. Analysis is by variables of subject, imprint year, and language, for separate groupings of research, academic, public, school and special/other libraries. Public and school libraries are also analyzed with juvenile and adult audience levels. The results of the study will contribute to a greater appreciation and understanding of WorldCAT as an international resource including bibliographic records on the universe of publication available for research and resource sharing worldwide.

The Problem

Users of WorldCat view bibliographic records one at a time. The user sees the Atrees@ individually but has little idea of what the Aforest@ contains. Subject and key word searching do allow the construction of bibliographies so that subsets of the database can be viewed. But for the majority of users, there is no concept of WorldCat as a coherent whole or a single aggregated resources base that reflects the global universe of publication. The scope of the resources under bibliographic control in WorldCat has not been analyzed in a multi-dimensional mode, that is single analyses have been performed of imprints by year or number of records by language, but not analyses with more than one variable combined together.

The problem of defining a universe of publication for subject areas has long plagued collection development. It is difficult to judge the strengths or weaknesses of library collections when there are no benchmarks to use for relative measurement. Viewing WorldCat as reflecting an international universe of publication offers at least a partial solution to the problem of defining a universe of publication for subject analysis of library collections.

Another aspect of collection analysis is that of identifying publishing and collecting patterns over time. While statistics have been published annually for U.S. book production and a few Western European countries, statistics for international book production are still unreliable for comparative analysis. Can WorldCat, as the largest international bibliographic database be used as a surrogate for studying publishing trends and library collecting patterns?

The next section reviews the existing body of research on collecting patterns using WorldCat and OCLC collection analysis products.

Review of Related Research

Prior to the development of the OCLC Online Union Catalog, there were a number of automated collections analyses, primarily overlap studies, conducted by groups of libraries to determine the feasibility of establishing regional network consortia. The research literature of overlap studies has been reviewed by Potter[2] and Medina.[3] Both Perrault[4] and McNeely[5] have reviewed studies on the methodology and feasibility of using database records to analyze library holdings and collection analysis studies using automated data analysis. Nisonger has provided a useful annotated bibliography of *Collection Evaluation in Academic Libraries*[6] which has been supplemented by bibliographies appearing in *Collection Building*.[7] The bibliographies provide a broad overview of the research in collection evaluation.

This review only covers research that has utilized WorldCat or the OCLC/AMIGOS CACD to study aggregated collections at the national level. The majority of these studies have been on academic library collections.

Four studies by Schwartz have been reported which utilized both the OCLC Union

Catalog database and the OCLC/AMIGOS CACD for data collection.[8] Schwartz used the two systems in combination to measure literature loss in anthropology, Judaic studies, international relations and psychology. Book production for a specified time period was measured by subject searching the OCLC Union catalog. Those records retrieved were assumed to represent the universe of publication in the disciplines under study. Book production was compared to the aggregate holdings of the ARL libraries for the same time period derived from the OCLC/AMIGOS CACD database. The methodology is posited as a model for determining the growing gap between total book publication output and the aggregate holdings of a peer group of libraries.[9] Schwartz found that in international relations overall, "ARL holdings dropped from 75% of book output in the late 1970s to 65% of book output in the late 1980s." In anthropology for the same time period, holdings dropped from 66% of book production in the late 1970s to 53% in the late 1980s. For anthropology, 40% of the total book output was not held by any of the 70 ARL libraries and 10% of the titles were held by only one or a few of the libraries. The remaining 30% Schwartz characterizes as "mainstream literature" of which the average was 10% for any one of the 70 libraries. He concluded that literature loss defined as the "growing gap between book publication output and the aggregate holdings of 70 institutions in the Association of Research Libraries ...grows rather steadily".[10]

Research by Perrault was designed to determine the effects of the decline in book acquisitions on the collective resources base of 72 ARL libraries.[11] The study utilized the 1991 OCLC/AMIGOS CACD to compare the two imprint years 1985 and 1989. This was the first reported research to study the ARL library collections as an aggregated resources base by both subject and language grouping parameters. The findings corroborated the trends in monographs acquisitions that had been tracked in the *ARL Statistics* series for

those years. The 1990/91 *ARL Statistics* showed a decrease of 14% in monograph purchases for the previous five years, with 1988 having the lowest point at a 20% decrease.[12] Among the findings of Perrault=s research were that, as of the time of data extraction in 1991, overall there was a 27.76% decline in the total number of books acquired by the libraries in 1989 as compared to 1985. English language titles had declined 12.34%, while foreign language books declined more steeply by 43.33%. A shift in collecting patterns occurred between 1985 and 1989 within the three broad knowledge groupings with the humanities declining in proportional share of total by 2.35 percentage points and the sciences increasing share by 2.99 percentage points. The social sciences remained constant in percentage share of total as well as percentage of unique titles to total and in the mean number of holding libraries.[13]

In an article on the methodology used in the ARL study, Perrault provided examples of collection analysis utilizing the OCLC/AMIGOS CACD product, updating the findings of the ARL study with data extracted from the 1996 OCLC/AMIGOS CACD.[14] This article reported on three peer groups of academic libraries, with all records in the CACD database composing a fourth peer group. The size of the collections for the four peer groups were compared for the time frame 1987-1995. The analysis concentrated on the trends in collecting for six major subject areas over the nine years. The conclusions of the study were that the aggregated resources of academic libraries in the United States are becoming increasingly composed of English language publications; the budget crisis of the 1980s is still visible in reduced numbers of titles for those years across all subject areas in 1996; smaller academic libraries have more stable collecting patterns indicating that the larger research institutions were harder hit by the serials crisis in the 1980s; and the number of titles for all categories rises in the early 1990s indicating somewhat of a

recovery from the decline in the 1980s.[15]

A broader study of the decline in foreign acquisitions in academic libraries on the 1980s was conducted for the AAU/ARL Foreign Acquisitions Project. The project was Adirected toward developing a clear understanding of the forces influencing North American research libraries= ability to build and maintain collections of publications produced outside the United States and Canada.@[16] The study utilized database scans for foreign publications in OCLC WorldCat with imprint dates of 1988-1993. The data were extracted for publications by country of origin but not by language parameters so that foreign English language publications are included in the counts. These data are not comparable to data from the CACD database in the research by Perrault. No subject analyses were conducted for the AAU/ARL Foreign Acquisitions Project. The emphasis of the investigation was on materials published in eight different geographic areas outside the United States. The results of the four-year study were analyzed in *Scholarship, Research Libraries, and Global Publishing*, authored by Reed-Scott.[17]

The trend analysis for foreign publishing included in *Scholarship, Research Libraries, and Global Publishing*, was partially replicated and updated in an ARL AGlobal Resources Report@ in 1999.[18] The update showed that publishing output outside of the U.S. had continued to increase in the latter 1990s. The data also show that the number of foreign titles with imprint dates from 1988-1994 had considerably increased in the four years after the initial data extraction. The increases in the number of imprints on the older end (1988) which were added are smaller than the increases beginning with 1991 imprints. The largest increase was for 1994 titles, the last imprint year of the earlier study, an increase of 67% being added in the 1995-1999 time frame. From the updated data, it appears that the time lag in adding foreign imprints to WorldCat is in the three to five year

range, with the number added beginning to slow after five years. Nonetheless, the updated data substantiated the earlier study in 1995, that North American Aresearch libraries collection coverage of global publishing output is declining.@[19]

Both Holleman[20] and McNeely[21] have conducted studies updating the Perrault studies. Holleman performed a small scale study with only the world history LC subject divisions using the 1995 edition of the OCLC/AMIGOS CACD. He found that by 1995 the number of 1989 imprints showed a .56% increase over the number of 1985 imprints as found by Perrault in the 1991 edition of the CACD. In the four years between the two editions of the CACD, a decline of 27.76% in the number of titles from 1985-1989 had become a modest increase as libraries continued to add titles from the latter 1980s. Holleman=s findings called attention to a problem with the reliability of the numbers of records in the last few imprint years in each edition of the CACD due to a time lag in library acquisitions and cataloging.

The McNeely study utilizes the Perrault methodology to analyze the changes in book collecting patterns of three peer groups: (1) members of the Association of Research Libraries (ARL); (2) libraries included in the OCLC/AMIGOS CACD, 1998 edition; and (3) TexShare Libraries. Data were analyzed for the period 1985-1995 by 104 Library of Congress classification ranges and for seven language groupings formulated for the Perrault studies. Unique titles by the three peer groups and subject categories were aggregated into humanities, social sciences, and sciences. Yearly changes, rate of change, and yearly average rate of change were calculated for LC subject categories, language groups, broad subject groups, and unique title counts. Findings were that the ARL and OCLC peer groups had an average yearly rate of growth between 2-3% for the ten year period. TexShare libraries had between 1-2%. OCLC had the largest growth in

the humanities; ARL had the largest in the social sciences, and TexShare had the largest growth in the sciences. The number of unique titles(diversity) in the collections increased as measured by yearly average for all peer groups: OCLC 3.76%, ARL 3.l8%, and TexShare, 1.64%. There was a 3% average yearly decline in percentage of non-English titles for this period. McNeely also found that the annual rate of change went into a negative mode for the last two imprint years in the study, as had the earlier studies by Perrault, further substantiating Holleman=s findings of the number of titles by subject divisions by imprint years continuing to increase over time. Overall, McNeely=s longitudinal study shows that Athe >national collection= continued to grow at an average yearly rate of 2-3 percent for the period studied (1985-1995).@[22]

Hardesty and Mak conducted a study of college library collections using both the CACD and data extracted from WorldCat. The collection size range was 100,000 to 299,000 volumes. The purpose of the study was to ascertain if there was an identifiable core of materials being collected by the group of college libraries. The findings were that there was not an identifiable core, but that the holdings were scattered.[23]

Senkevitch and Sweetland studied adult fiction collections in over 4,000 OCLC member public libraries under the auspices of an OCLC research grant.[24] The study examined the extent of change over a one-year period in the most widely held adult fiction titles and sought to answer the question AIs there a consensus core of adult fiction?@ Findings were that the most widely held adult fiction titles were Arecent, popular works that form a stable core from one year to the next. Results also suggest that such a listing of widely held titles might be suitable as an evaluation tool in smaller public libraries.@[25]

Two studies were reported which utilized data from OCLC or the CACD to compare library holdings with *Choice* AOutstanding Academic Books@ lists. Budd and Craven[26]

investigated the holdings of four peer groups of academic libraries for the intervals 1984/85, 1990, and 1995. Data were extracted from two editions of the OCLC/AMIGOS CACD. The researchers took a purposive sample of titles from the *Choice* AOutstanding Academic Books@ lists for those years. Titles were chosen for major academic disciplines representing the humanities, social sciences, and sciences. Findings demonstrated that Aacross all subjects and for broad subject areas, there have been significant declines in holdings over time and for all groupings of libraries.@ The researchers maintain that the findings of other studies with regard to the Ashrinking of library resources is corroborated in this study.@[27]

Although not a study of academic library collections in the aggregate, a survey by Sweetland and Christensen of selection practices for language and literature collections in Wisconsin academic libraries used OCLC data to ascertain the number of holding libraries for *Choice* titles.[28] The researchers used *Choice* 1993 AOutstanding@ titles in languages and literature. The list was published in January 1993. The data were extracted from the OCLC Online Catalog in March, 1995, two years later. Findings were that, on the average, about 10 libraries held each title with the range being from 4 to 23 libraries per title. The authors observe, A...considering that these titles are supposedly the most outstanding of all those recommended by *Choice* for four-year college libraries, the small number of holdings is curious, to say the least.@ Part of the problem is that from the study data it appeared that Aa substantial number of libraries did not add books to their collection until sometime in 1994 (or very late in 1993).@[29] The researchers give as a possible explanation for the delay in adding material the almost total reliance upon faculty recommendations with faculty using scholarly reviews to identify new titles. The researchers conclude that Afaculty still control selection in most [college] libraries;@ Athe libraries in the study still consider current,

immediate needs, and rarely consider other library holdings in selecting material to purchase;@ Aeven though libraries feel they are buying current material, their definition of ›current= means within the last couple of years;@ and Arespondents are currently buying very little outside the English language.@[30]

The majority of the research utilizing data from WorldCat or the OCLC/AMIGOS CACD has been on national collecting patterns in North American libraries. It has been heavily focused on the ARL research library group. There have been fewer reported national studies of smaller academic or public library collections. This review of related research shows that there is still much to be learned about the collective resources of the libraries contributing to WorldCat and the state of those collections in the latter years of the 20th century. While broad profiles of the database have been produced by the OCLC research office, the monographic bibliographic records in WorldCat have not been formally studied for collecting patterns by subject, language, and type of library and level of audience parameters. This study is the first reported research on WorldCat as reflective of global collective resources and the international universe of publication. The methodology for the study is outlined in the next section.

Methods

The study is a bibliometric content analysis of OCLC WorldCat, a merged, electronic catalog of the bibliographic and holdings information of OCLC member libraries, founded in 1971. The more than 45 million records in WorldCat as of 2001 represent more than 750 million items held in libraries around the world. The study was designed to take advantage of the data analysis capabilities of the OCLC/WLN iCAS software product which is a CD-ROM tool for collection analysis. (For information on iCAS go to

http://www.oclc.org/western/products/aca/icas.htm)

The iCAS software performs basic age and content analyses. Data are shown in a grid with the customer library=s choice of publication date ranges across the top and WLN conspectus division, categories, and subjects down the side. Each cell contains the number of titles held for that subject area and publication date/date range, as well as its percentage of the division, category or subject. Graphing, printing and export options are included.

Data Collection and Analysis

A 10% systematic random sample of the monographic bibliographic records in WorldCat was extracted in the OCLC Office of Research and forwarded to the Lacey Product Center of OCLC in February 2001. The number of MARC bibliographic records in the sample was 3,555,907. Each MARC record in the sample carried an identifier for a type of library grouping as research, academic, public, school, or special\other. The ICAS software was utilized to profile the bibliographic records in the WorldCat sample dataset according to the following parameters:

$ imprint year analyses -- Dates and date ranges are: by century until 1800; 1800 - 1849,1850-1899; by decade until 1980; 1980-1984; then annual through 2000.

$ type of library categories -- Research, academic, public libraries, special/other, and school

$ subjects WorldCat, research grouping and the academic group holdings are analyzed by WLN Library of Congress Conspectus categories; the public and school library groups holdings are analyzed as WLN Dewey Conspectus categories

$ juvenile and adult subtotals B for the public and school library groupings, analyzed as Dewey Conspectus categories and publication year

$ language groupings BEnglish, Chinese, French, German, Japanese, Russian and

Spanish for WorldCat, the research and the academic library groupings analyzed by publication year and WLN LC Conspectus categories

The iCAS data analysis for WorldCat was completed in June, 2001. A number of different analyses were conducted by staff at WLN on the monographic bibliographic records in the 10% systematic sample of the WorldCat database. The total number of records in the sample is 3,555,907. Of those records a small number, 4% of the records, had no library holdings symbol or were not owned by any of the libraries in the sample. These records were not used at all. Of the remaining records, 3,376,272 records were used in the language analysis. The language analysis is performed by sorting on the language codes in the fixed field of the record. The same dataset was used to run an analysis of the records to determine the number of records with ISBN numbers.

With regard to publication dates, an Aother@ category in the analysis by imprint years contains the number of records that had no publication date. Approximately 1% of the records did not have publication dates, too small a number to have any effect on the findings.

In order to perform subject analyses the bibliographic record has to have a Dewey, Library of Congress, or NLM call number. Of the 3,376,272 usable records, 1,179,107 did not have call numbers. Hence 35% of the sample records could not be used for subject analysis. This set of records formed a second dataset that appears as ANo call number present@ in many of the tables in the study. Throughout this report it is indicated which of the datasets have been used for the particular analysis under discussion.

After preliminary analysis of the data and conferring with Lacey Product Center staff, the researchers requested that a number of additional analyses be run. These included an analysis of WorldCat, the research libraries, and the academic libraries with the

ANo Call Number@ records included as one line so that a total number of records from the sample could be accounted for in each of the three largest data groupings. Separate analyses were run for unique records for these three groupings and an analysis of records by defined ISBN number groupings for WorldCat. The second round of iCAS analyses were delivered in October 2001.

It is important to bear in mind that the data in this study form a Asnapshot@ of the contents of WorldCat at one point in time. Bibliographic records are continually added by OCLC member libraries for imprints across the entire time span defined for the study. The analysis in this study concentrates on proportions, percentages of total, and trends. These proportions should represent the profile of the population of monographic bibliographic records in WorldCat from which the sample was drawn. Since the absolute numbers are from the sample, these numbers cannot be utilized for benchmarks, only for the findings of proportions, percentages, and trends. Were another sample to be extracted in several years, the absolute numbers would be different, but the proportions or trends should be nearly the same. While discernible shifts over time can occur in the profile of the database, such shifts would necessarily involve a large number of bibliographic records in a database of over 45 million records. There are a small number of studies in the literature review in this chapter with which the findings of the WorldCat study can be compared. But because the data for these previous studies were extracted from the WorldCat database at different times, again, the findings can only be compared in proportions or trends and not in exact numbers of records.

The proportions found in the data analysis in this study can be influenced by the collecting patterns of the library membership of OCLC. In the separate analysis for each of the type of library groupings, the composition of member libraries for each group is

described in general terms. A listing of library membership in OCLC with the type of library code can be accessed at http://www.oclc.org/contacts/libraries/.

The report of findings begins with Chapter two which profiles WorldCat as reflecting the international universe of publication over time. Subsequent chapters analyze the profiles of the aggregated collections of the type of library groupings, the language groupings, unique titles, and an ISBN analysis.

Click here for Chapter Two.

References

1. OCLC Annual Report, 1999/2000: 46-47.

2. William G. Potter, "Studies of Collection Overlap: a Literature Review," *Library Research* 4 (1982): 2-21.

3. Sue O. Medina, "Duplication and Overlap Among Library Collections: A Chronological Review of the Literature," *Advances in Collection Development and Resource Management* v.1, ed. Thomas W. Leonhardt (Greenwich CT: JAI Press, 1995): 1-60.

4. Anna H. Perrault. AThe Changing Print Resource Base of Academic Libraries in the United States: a Comparison of Collection Patterns in Seventy-two ARL Academic Libraries of Non-serial Imprints for the Years 1985 and 1989.@ Ph.D. dissertation, The Florida State University, 1994.

5. Bonnie Michaels McNeely. ALongitudinal Study of the Collection Development Patterns of Libraries in the Association of Research Libraries and the Online Computer Library Center, and Texas Academic Libraries Using the OCLC/AMIGOS Collection Analysis Compact Disc to Assess Strengths, Weaknesses, and Overlap for 1985-1995.@ Ph.D. dissertation, Texas Woman=s University, 2001.

6. Thomas E. Nisonger. *Collection Evaluation in Academic Libraries: A Literature Guide and Annotated Bibliography*. (Englewood, CO: Libraries Unlimited, 1992)

7. Thomas E. Nisonger. AThe Collection Development Literature of 1996: a Bibliographic Essay.@ *Collection Building* 17, no. 1(1998): 29-39; AA Review of the 1997 Collection Development and Management Literature.@ *Collection Building* 18, no.2(1999): 67-80.

8. Charles A. Schwartz, "Literature Loss in Anthropology," *Current Anthropology* 33 (June 1992) 315-317; "Empirical Analysis of Literature Loss," *Library Resources and Technical Services*, 38, no.2 (April 1994), 133-138; "Literature Loss in International Relations," *PS: Political Science & Policy* 25 (December 1992), 720-723; "Gap Analysis of Book Publication Output and Aggregate Research Collections in Psychology," *American Psychologist* 48 (1993): 1151-52.

9. Schwartz, AEmpirical Analysis of literature loss.@

10. Ibid.

11. Anna H. Perrault, "The Shrinking National Collection: a Study of the Effects of the Diversion of Funds from Monographs to Serials on the Monograph Collections of Research Libraries," *Library Acquisitions Practice & Theory*, 18, no.1 (1994): 3-22.; "The Changing Print Resource Base of Academic Research Libraries in the United States," *JELIS* (*Journal of Education for Library and Information Science*) 36, no.4 (Fall 1995): 1-16.

12. *ARL Statistics*, 1990/91.

13. Perrault, AThe Shrinking National collection...@ p.

14. Anna H. Perrault. ANational Collecting Trends: Collection Analysis Methods and Findings.@ *Library & Information Science Research* 21, no.1(1999): 47-67.

15. Ibid., p.65.

16. Jutta Reed-Scott, *Scholarship, Research Libraries, and Global Publishing*, (Washington, DC: Association of Research Libraries, 1996), xvii.

17. Ibid.

18. Deborah Jakubs. AThe AAU/ARL Global Resources Program: Both Macrocosm and Microcosm.@ *ARL* No. 206 Special Issue on Global Resources (October 1999).

19. Ibid., p.8.

20. Curt Holleman. AThe Study of Subject Strengths, Overlap, and National Collecting Patterns: The Uses of the OCLC/AMIGOS collection Analysis CD and Alternatives to It.@ Collection management 22, no. 2: 57-69.

21. McNeely. ALongitudinal Study of the Collection Development Patterns....

22. Ibid. v-vi.

23. Larry Hardesty and Collette Mak. ASearching for the Holy Grail.@ *The Journal of Academic Librarianship* 19, no.6 (1994): pp. 362-371.

24. Judith J. Senkevitch and James H. Sweetland. APublic Libraries and Adult Fiction: Another Look at a Core List of Classics.@ *Library Resources & Technical Services* 42, no.2 (1998): p.102-112.

25. Ibid., p.102.

26. John M. Budd and Catherine K. Craven. AAcademic Library monographic Acquisitions: Selection of *Choice*=s Outstanding Academic Books.@ *Library Collections, Acquisitions, & Technical Services* 23, no.1(1999): 15-26.

27. Ibid., p.122.

28. James H. Sweetland and Peter G. Christensen. ADeveloping language and Literature Collections in Academic Libraries: a Survey. *Journal of Academic Librarianship* v (March 1997):119-125.

29. Ibid., p.122.

30. Ibid.,p.123.

Chapter Two

WORLDCAT: The PROFILE

The Sample

The data analysis and interpretation in this report construct a profile of the monographic bibliographic records in WorldCat by a number of parameters. All 3,378,272 usable records from the sample are analyzed according to the imprint year time periods and subject divisions defined for the study in the main Table 2-1, *WorldCat Records, All Titles Held by Date (all subject divisions)*, attached to this chapter. It must be borne in mind that all absolute numbers in this report are from the 10% systematic sample of monographic bibliographic records in WorldCat which should reflect the population of records in WorldCat. Because the numbers are from a sample, they will not directly correspond to numbers in actual collections or the analyses of WorldCat in *OCLC Annual Reports*.

Table 2-1 is the ABig Picture,@ showing the distribution of monographic bibliographic records in WorldCat by imprint year and subject division across the 500 plus years of the printed book. Table 2-1 has a Subject Analysis subtotal line which contains the total for all records with call numbers used in the subject analysis (2,199,1650). After the Subject Analysis subtotal line, the number of records with ANo call number present@ are shown by time period (1,179,107). There is also an Aother@ column at the end of the dated columns which shows the total by subject for all those records which did not have publication dates. The AGrand Total@ line adds the subtotal, Subject Analysis, and the ANo call number present@ lines to account for all usable records in the10% sample (3,378,272).

Table 2-2, *WorldCat Monographic Bibliographic Records: Percentage Increase/Decrease by Time Period*, is a summary table of records by time period. The table contains the total number of records in the WorldCat sample divided into those records with usable call numbers and those without call numbers. The increase/decrease in the number of records in the sample over the 500 year time span is shown.

As defined for the study, the early centuries beginning with 1500 are shown in 100

year blocks. The number of records for the earliest period, pre-1500, may appear to be high, but these no doubt include records for manuscript works as well as incunabula, those works printed before 1500 or facsimiles of these early works. The first full century of printing, 1500-1599, naturally shows a very large increase in titles over the first time interval of pre-1500.

The publication explosion of the nineteenth century can be seen in the huge difference between the number of records for the first half of the century, which is more than doubled in the second half (133%). From 1900 forward, the time periods are by decade until 1985, in which the analysis becomes annual. The number of records increases steadily throughout the 20th century with the exception of the decades of the two world wars which each decrease from the previous decade. Even in the decade of the 1930s, the period of the Great Depression, the number of records increases 18% over the more prosperous decade of the 1920s. The decrease is slight in the 1940s, only 2 percent. The explosion in publishing which followed World War II is obvious in the huge increase in records for the 1950s (49%), with an even larger increase of 82% in the 1960s. The number keeps going up as the 1970s saw another increase of 61 percent. Adding together the number of records for the 1980s gives a total of 670,660 records, an increase of 27% over the previous decade.

For the fifteen years 1985-1999, the number of records increases annually until 1993, in which the number begins to decrease. From an annual high of 75,460 records in 1992, there is a 39 percentage point difference in the number of records between 1992 and 1999. Since the data were extracted in February 2001, it could be expected that many libraries had not cataloged or set holdings for all year 2000 imprints. The differences, however, between 1999 and years immediately preceding would seem to show that there is a time lag during which bibliographic records for a considerable number of published works are not available in WorldCat. Previous research as reviewed in Chapter One shows that the annual number of records rises as time goes on and the numbers tend to level out over time.

In Table 2-2, the records for ANo call number present@ do not follow closely the patterns of increase/decrease for those with call numbers. Prior to the mid 19th century, the number of records without call numbers exceeds the number of records with call numbers. It is not until the 1850B1899 time period that the number of records with call numbers exceeds the number without call numbers. The large number of records without call numbers pre-1850 may reflect records from special collections and non-US libraries which do not use Dewey or Library of Congress classification schemes.

In 1988 both groupings of records have a similar increase over the previous year of plus 3 percent. For those records without call numbers, the number of records declines each year beginning in 1991, while those with call numbers begin to decline in 1993. The declines are larger in the number of records with no call number than in those with call numbers. In 1996 the ratio of records with call number to those without call number is nearly 3/1. The number of no call number records peaks slightly earlier, 1990, than those with call numbers (1992) and the decrease is more pronounced, but the same pattern of decline during the 1990s is present. The number of records with call numbers declines 72 percentage points between 1992 and 2000. The number of records without call numbers declines 88 percentage points from 1990 to 2000. Further analysis of the decreasing numbers of records in WorldCat for the most current years is contained in the foreign language analysis, the unique title analysis, and the ISBN analysis.

Subject Analysis

Subject analysis can only be conducted on those records with call numbers as the call numbers are used to perform the subject sort. If we look at the vertical ATotal@ column in Table 2-1, the main table, we can see the total number of records with call numbers for each of the 24 subject divisions which correspond to the Library of Congress A-Z broad subject divisions. It is obvious that the Language, Linguistics, and Literature (LLL) subject division has the largest number of titles in the sample out of the 24 subject divisions, with 464,138 records. The LLL division represents the majority of the titles in the LC AP@ classifications and comprises 21% of all of the records with call numbers. The history, biography and travel classifications, LC AC-F,@ have 292,037 or 13.3% of all records with

call numbers. The business and economics classes are the only other grouping to have over 200,000 titles and comprise 9.6% of the total number of records with call numbers. Only two other subject divisions have more than 100,000 records: philosophy/religion has 180,309 (8.2%) and engineering/technology with 150,042 (6.8%).

Of those subject divisions with less than 100,000 records in the sample, education has 97,094 (4.4%); art and architecture with 93,013 for (4%); law with 85,621 (3.9%); sociology with 84,886 (3.9%); political science with 74,097 or (3.4%). Library science and bibliography and medicine have almost the same number of records for 3.1% of total each. The remaining subject divisions have very small percentages of the total number of records in the sample. These proportions to total for the records in the sample should closely reflect the proportions of those subject divisions in the WorldCat database.

Those subject divisions which are known to be monographically oriented disciplines are the subject divisions with the highest number of records. The LLL and history subject divisions have much larger proportions of records than any other subject divisions. These two subject areas comprise 34.3%, or over 1/3 of the records with call numbers in the sample. Adding in the only other subject division with over 200,000 records, business/economics at 9.6%, yields 43.9% of total comprised of the three largest subject divisions. The next two largest subject divisions, those with over 100,000 titles comprise 15% of total. The largest five subject divisions account for nearly 60% of the records in the sample, the remaining subject areas thus comprising the other 40 percent.

Computer science predictably shows few records until the 1960s and is one of the few subject divisions with increases in annual numbers in the 1990s, with 1999 having the highest number of records of any time period. Chemistry, a classic science field, more closely follows the pattern of increase/decrease for the database. It fluctuates after 1984 with the highest annual number reached in 1989 and declining numbers after that. The LLL division has small but rather steady increases each year beginning in 1985 with the highest annual total reached in 1994 and a small decline in numbers each year until a more pronounced decrease of 15% in 1999. The LLL division has by far the highest

number of records in 2000 evincing a strong pattern of acquisitions support for those subject areas. The history areas follow an almost identical pattern to the LLL division, again with the highest number in 1994 but a larger 23% decline in 1999.

If we look at the imprint year which has the highest number of records after 1984 for each of the 24 subject divisions, the earliest year in which any division peaks is 1990. Education and engineering/technology are the two subject divisions which have their highest number of annual records in 1990 with the number steadily declining after that. ANo call number present@ also peaks in 1990. Five subject divisions peak in 1991: chemistry, computer science, music, and physical sciences. There are fours areas which peak in 1992: business/economics, medicine, philosophy/religion, and political science. 1993 has the largest number of subject division peaks; agriculture, biological sciences, mathematics, physical education/recreation, and sociology. In 1994 five subject areas peak: anthropology, history, LLL, law, and psychology. Only library science and bibliography have their highest number of records in 1995, perhaps reflecting the current nature of the materials, and one subject, art/architecture has the largest number of records in 1996. Thus is can be seen that the number of records in the sample reflect a pattern of the highest annual number of records peaking between 1990 and 1996, with the majority in the 1990-1994 time frame.

The patterns, both for absolute numbers and the annual increase/decline after 1984, are remarkably similar for all of the subject divisions. The differences in absolute numbers between the various subject divisions are due in part to the way in which the subject divisions are constituted in the WLN Conspectus and the iCAS products (This organization is also the North American Title Count subject division). The LLL division is the largest because it includes all language, literature, and linguistics worldwide. The history division is likewise inclusive. These two areas are the major disciplines that are studied internationally and will figure prominently in the language analysis. They are also the most monographically oriented of all disciplines.

The absolute numbers for all of the subject divisions are also a reflection of the universe of publication by subject fields. The relative size of each subject to the others reflects the proportion of the universe of publication which each field represents. The relative lack of fluctuation over the last 15 years of the 20th century can either be a reflection of the universe of publication, or a reflection of the numbers of titles by subject fields the libraries in WorldCat were acquiring and cataloging. It is likely that the acquisitions patterns reflect publication patterns, with the exception of the last 5-6 years of data.

The data from Table 2-1 have been aggregated into several summary tables for subject analysis. Table 2-3, *WorldCat Records: Subject Divisions, 50 Year Range*, shows the growth in numbers of titles by 50 year ranges beginning in 1800.

It is easy to see the explosion in publishing which accelerated from the first half of the 19th century through the last half of the 20th century in Table 2-3. The total of records for publication years since 1800 make up 95% of the records in the study. Just the fifty years from 1950-1999 contain 70% of all records in the 10% sample from WorldCat. While there is undoubtedly much global publication which is not reflected in WorldCat, we can be certain that these patterns are at least reflective of Western European and North American publication rates.

Table 2-4, *WorldCat RecordsBSubject Divisions, 10 Year Range*, shows the decades of the 20th century. It has been already been shown from Table 2-1 that the number of records progressively declines throughout most of the 1990s, yet it can be seen in Table 2-4 that the decade of the 1990s has the highest number of records for anthropology, art and architecture, biological sciences, business and economics, computer science, history, LLL, law, mathematics and medicine, performing arts, philosophy/religion, physical education and recreation, political science, psychology, and sociology. (Highlighted in Table 2-4)

When the magnitude of the increase in publication numbers from the 1950s forward is seen, there can be no wonder that the rate of increase for monographs (not to mention

average price increases) made it a near impossibility for libraries to continue to acquire a substantial proportion of the universe of publication. Those subject areas which have a decrease in numbers in the last decade are, agriculture, chemistry, education, engineering and technology, geography and earth science, library science, general reference and bibliography, music, physical sciences. These are mainly fields which are not monographically oriented disciplines. The exception would be education and the library science, reference and bibliography areas. The decrease for the AZ@ classifications may be partially explained by a change in classification by the Library of Congress in which bibliographies began to be assigned to the subject classification of the work instead of AZ.@ It is also suspected that with the prevalence of online bibliographic databases beginning in the early 1990s, printed bibliographies became less useful as they became out of date very quickly and it had become easy to construct up-to-date bibliographies from databases.

The increase for the monographically oriented humanities and social sciences fields in the 1990s over the 1980s may be an indication that libraries were endeavoring to support acquisitions for those areas because of the large increases in serial subscription prices in the scientific, technical, medical fields.

It is obvious that the rate of monographic publication did not abate in the 1990s as even with the steep declines in the number of records annually for most of the decade, the total number of records for the 1990s is only 5,000 less than the total for the decade of the 1980s.

The rate of decrease in the number of records by imprint year in the latter half of the 1990s points to a pattern of Aslow accretion@ in the way in which libraries add new titles. The analysis provided for the WorldCat profile cannot solely identify the causes of the time lag in records appearing in WorldCat. Other studies have similar findings with regard to the time lag. It may be that libraries are slow in making selection and thus purchasing decisions. It may be that the majority of libraries wait for a bibliographic record to show up in WorldCat to avoid original cataloging. And it could be a combination of both of these

factors. This problem will again be explored in the subsequent chapters with the analysis by type of library grouping, the Unique Titles, and the Language Analysis.

Summary for WorldCat

The profile of WorldCat by subject divisions and time periods depicted in Chapter Two reveals the parameters of the aggregated resources base of the 9,000 library membership of OCLC. The collective holdings of these libraries comprise an Ainformation commons@ of titles available through resources sharing. The profile is a broad outline provided through the iCAS analysis of the 10% sample of monographic bibliographic records extracted from the WorldCat database. Subsequent chapters explore the variables in the study providing a more detailed profile of those resources.

In Chapter Three the five types of library groupings are each analyzed separately by the same parameters of distribution by imprint date and subject divisions.

Click here for Chapter 3.

Table 2-1a
WorldCat Monographic Bibliographic Records - All Titles Held by Date (all subject divisions)

SUBJECT DIVISION	Pre-1500	1500-1599	1600-1699	1700-1799	1800-1849	1850-1899	1900-1909	1910-1919	1920-1929
AGRICULTURE	0	26	75	244	441	1,987	1,285	1,536	1,642
ANTHROPOLOGY	1	20	30	52	160	650	354	296	419
ART AND ARCHITECTURE	3	92	151	401	942	3,101	2,159	1,641	2,728
BIOLOGICAL SCIENCES	2	34	76	424	1,020	3,043	1,399	1,009	1,263
BUSINESS AND ECONOMICS	2	59	209	2,371	4,342	5,128	3,316	3,253	4,610
CHEMISTRY	0	9	44	76	118	358	230	177	327
COMPUTER SCIENCE	0	0	0	1	0	4	10	3	2
EDUCATION	1	15	56	144	565	1,992	1,148	1,143	1,693
ENGINEERING AND TECHNOLOGY	5	77	177	428	969	3,966	2,575	2,273	2,976
GEOGRAPHY AND EARTH SCIENCES	1	39	99	356	819	2,676	1,196	973	1,197
HISTORY AND AUXILIARY SCIENCES	8	401	2,100	4,141	7,677	23,375	11,360	9,465	9,745
LANGUAGE, LINGUISTICS, AND LITERATURE	51	962	2,332	6,530	12,932	34,256	18,443	12,916	18,031
LAW	15	237	417	688	1,602	3,571	1,168	950	1,053
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	6	33	324	738	1,567	3,262	1,407	1,007	1,611
MATHEMATICS	1	50	150	263	410	1,047	502	382	449
MEDICINE	5	128	222	466	957	2,301	1,077	1,003	1,257
MUSIC	3	49	83	402	567	1,594	973	686	943
PERFORMING ARTS	0	13	31	69	109	371	257	195	362
PHILOSOPHY AND RELIGION	75	1,492	3,638	5,029	7,405	18,945	7,683	5,405	6,394
PHYSICAL EDUCATION AND RECREATION	0	6	19	47	127	547	352	282	443
PHYSICAL SCIENCES	16	169	308	612	820	1,926	902	760	924
POLITICAL SCIENCE	0	61	245	640	1,062	3,234	1,625	1,932	1,785
PSYCHOLOGY	1	24	99	82	153	621	447	384	631
SOCIOLOGY	1	12	64	242	866	2,315	1,145	1,074	1,317
Subject Analysis	**197**	**4,008**	**10,949**	**24,446**	**45,630**	**120,270**	**61,013**	**48,745**	**61,802**
No Call Number Present	1,189	7,369	16,124	32,265	48,052	98,366	42,986	37,472	42,853
Grand Total	**1,386**	**11,377**	**27,073**	**56,711**	**93,682**	**218,636**	**103,999**	**86,217**	**104,655**

Table 2-1b
WorldCat Monographic Bibliographic Records - All Titles Held by Date (all subject divisions)

SUBJECT DIVISION	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1984	1985	1986	1987
AGRICULTURE	1,993	1,986	3,418	5,046	8,595	5,361	1,080	1,013	993
ANTHROPOLOGY	484	405	732	1,231	2,396	1,418	308	303	305
ART AND ARCHITECTURE	2,742	2,795	5,094	10,317	16,617	10,330	2,083	2,013	2,068
BIOLOGICAL SCIENCES	1,556	1,340	2,241	4,332	7,738	4,489	920	923	937
BUSINESS AND ECONOMICS	6,159	6,525	9,513	20,515	39,372	24,406	4,952	5,153	5,127
CHEMISTRY	368	372	759	1,527	1,894	955	197	176	214
COMPUTER SCIENCE	1	19	66	546	1,606	1,785	499	519	564
EDUCATION	2,848	2,310	4,385	10,821	21,417	11,854	2,444	2,451	2,356
ENGINEERING AND TECHNOLOGY	3,532	4,840	8,399	15,575	30,872	18,954	3,841	3,816	3,705
GEOGRAPHY AND EARTH SCIENCES	1,412	1,370	3,518	6,237	9,633	5,753	1,122	1,137	1,143
HISTORY AND AUXILIARY SCIENCES	12,657	13,033	14,621	28,175	43,055	24,808	5,160	4,967	5,235
LANGUAGE, LINGUISTICS, AND LITERATURE	17,940	18,404	27,146	50,991	69,595	37,742	7,734	7,770	7,726
LAW	1,555	1,522	2,765	5,338	13,230	10,233	2,259	2,362	2,344
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	2,681	2,033	3,177	7,950	13,960	7,624	1,557	1,256	1,574
MATHEMATICS	446	537	977	2,681	3,749	2,045	445	420	429
MEDICINE	1,578	1,665	3,016	5,761	11,792	7,685	1,544	1,698	1,739
MUSIC	1,077	1,141	1,656	2,814	4,510	2,552	561	473	529
PERFORMING ARTS	522	502	880	1,495	2,780	1,793	401	373	367
PHILOSOPHY AND RELIGION	6,828	6,799	10,441	17,433	22,776	13,635	2,914	2,822	2,699
PHYSICAL EDUCATION AND RECREATION	606	545	1,176	2,138	4,559	2,793	504	509	549
PHYSICAL SCIENCES	1,122	1,114	2,150	4,513	5,124	2,865	590	624	659
POLITICAL SCIENCE	2,412	3,047	4,366	7,737	12,072	6,889	1,426	1,458	1,557
PSYCHOLOGY	612	524	1,021	2,161	4,520	2,358	500	560	514
SOCIOLOGY	1,710	1,570	2,722	7,259	17,726	10,069	2,032	2,121	2,156
Subject Analysis	**72,841**	**74,398**	**114,239**	**222,593**	**369,588**	**218,396**	**45,073**	**44,917**	**45,489**
No Call Number Present	50,466	46,145	65,353	104,778	158,019	108,674	21,891	21,830	22,413
Grand Total	**123,307**	**120,543**	**179,592**	**327,371**	**527,607**	**327,070**	**66,964**	**66,747**	**67,902**

Table 2-1c
WorldCat Monographic Bibliographic Records - All Titles Held by Date (all subject divisions)

SUBJECT DIVISION	1988	1989	1990	1991	1992	1993	1994	1995
AGRICULTURE	1,076	1,067	1,133	1,147	1,156	1,156	1,100	943
ANTHROPOLOGY	329	338	370	405	411	401	431	378
ART AND ARCHITECTURE	2,181	2,228	2,342	2,342	2,355	2,225	2,215	2,231
BIOLOGICAL SCIENCES	994	952	1,086	1,126	1,122	1,134	1,080	932
BUSINESS AND ECONOMICS	5,322	5,487	5,702	5,792	6,037	5,783	5,736	5,366
CHEMISTRY	197	232	225	228	223	201	178	154
COMPUTER SCIENCE	553	672	735	795	774	776	724	700
EDUCATION	2,539	2,511	2,737	2,571	2,749	2,689	2,469	2,328
ENGINEERING AND TECHNOLOGY	3,791	3,780	4,085	4,044	4,088	3,700	3,544	3,191
GEOGRAPHY AND EARTH SCIENCES	1,144	1,272	1,232	1,395	1,331	1,294	1,181	1,118
HISTORY AND AUXILIARY SCIENCES	5,541	5,634	6,088	5,912	6,130	5,980	6,241	6,136
LANGUAGE, LINGUISTICS, AND LITERATURE	7,908	8,440	8,813	9,206	9,262	9,561	9,600	9,251
LAW	2,465	2,618	2,902	3,076	3,028	3,030	3,204	2,901
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	1,410	1,287	1,437	1,324	1,370	1,227	1,170	1,278
MATHEMATICS	456	467	507	540	550	551	486	468
MEDICINE	1,918	1,861	1,976	2,141	2,246	2,111	2,089	1,866
MUSIC	530	535	521	548	508	524	513	471
PERFORMING ARTS	381	375	431	404	446	473	408	427
PHILOSOPHY AND RELIGION	2,886	2,965	3,252	3,164	3,267	3,172	3,134	2,908
PHYSICAL EDUCATION AND RECREATION	610	567	612	695	663	699	653	659
PHYSICAL SCIENCES	673	646	659	754	713	683	596	654
POLITICAL SCIENCE	1,624	1,690	1,845	1,985	2,041	1,915	1,970	1,893
PSYCHOLOGY	539	573	550	566	609	601	640	523
SOCIOLOGY	2,188	2,341	2,459	2,631	2,695	2,729	2,579	2,588
Subject Analysis	**47,255**	**48,538**	**51,699**	**52,791**	**53,774**	**52,615**	**51,941**	**49,364**
No Call Number Present	23,130	23,054	23,291	22,208	21,686	21,255	20,287	18,341
Grand Total	**70,385**	**71,592**	**74,990**	**74,999**	**75,460**	**73,870**	**72,228**	**67,705**

Table 2-1d
WorldCat Monographic Bibliographic Records - All Titles Held by Date (all subject divisions)

SUBJECT DIVISION	1996	1997	1998	1999	2000	2001	Other	Total
AGRICULTURE	964	946	766	617	258	12	365	49,427
ANTHROPOLOGY	396	421	355	293	108	7	211	14,418
ART AND ARCHITECTURE	2,249	2,179	1,973	1,553	523	12	1,128	93,013
BIOLOGICAL SCIENCES	833	837	789	600	313	23	559	45,126
BUSINESS AND ECONOMICS	4,832	4,652	4,055	3,096	1,186	69	1,962	210,089
CHEMISTRY	148	114	102	107	66	5	74	10,055
COMPUTER SCIENCE	638	648	626	749	340	21	728	15,104
EDUCATION	1,888	1,649	1,569	1,134	383	47	2,188	97,094
ENGINEERING AND TECHNOLOGY	2,988	2,941	2,584	2,190	916	52	1,168	150,042
GEOGRAPHY AND EARTH SCIENCES	1,097	998	961	812	434	18	812	53,780
HISTORY AND AUXILIARY SCIENCES	5,888	5,577	5,168	3,995	1,350	84	2,330	292,037
LANGUAGE, LINGUISTICS, AND LITERATURE	8,977	8,639	8,155	6,942	3,222	156	4,505	464,138
LAW	2,910	2,807	2,513	1,911	529	16	402	85,621
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	950	1,553	689	519	210	5	2,742	68,938
MATHEMATICS	395	390	381	297	133	20	251	20,875
MEDICINE	1,834	1,797	1,595	1,510	761	52	932	68,583
MUSIC	392	407	375	371	144	13	494	26,959
PERFORMING ARTS	441	411	381	349	153	9	174	15,783
PHILOSOPHY AND RELIGION	2,972	2,831	2,493	2,062	865	51	1,874	180,309
PHYSICAL EDUCATION AND RECREATION	714	688	603	598	300	21	283	23,567
PHYSICAL SCIENCES	519	504	452	418	199	19	316	33,003
POLITICAL SCIENCE	1,822	1,663	1,430	1,146	536	13	976	74,097
PSYCHOLOGY	511	511	474	486	239	16	171	22,221
SOCIOLOGY	2,383	2,343	2,089	1,604	727	50	1,079	84,886
Subject Analysis	**46,741**	**45,506**	**40,578**	**33,359**	**13,895**	**791**	**25,724**	**2,199,165**
No Call Number Present	16,407	17,017	15,060	12,037	2,795	11	16,283	1,179,107
Grand Total	**63,148**	**62,523**	**55,638**	**45,396**	**16,690**	**802**	**42,007**	**3,378,272**

Table 2-2
WorldCat Monographic Bibliographic Records
Percentage Increase/Decrease by Time Period

	Call Number Present		No Call Number Present		Total Sample Records	
Years	**Number of Records**	**Percent Increase/Decrease**	**Number of Records**	**Percent Increase/Decrease**	**Number of Records**	**Percent Increase/Decrease**
Pre-1500	197		1,189		1,386	
1500-1599	4,008	1934.52%	7,369	519.76%	11,377	720.85%
1600-1699	10,949	173.18%	16,124	118.81%	27,073	137.96%
1700-1799	24,446	123.27%	32,265	100.11%	56,711	109.47%
1800-1849	45,630	86.66%	48,052	48.93%	93,682	65.19%
1850-1899	120,270	163.58%	98,366	104.71%	218,636	133.38%
1900-1909	61,013		42,986		103,999	
1910-1919	48,745	-20.11%	37,472	-12.83%	86,217	-17.10%
1920-1929	61,802	26.79%	42,853	14.36%	104,655	21.39%
1930-1939	72,841	17.86%	50,466	17.77%	123,307	17.82%
1940-1949	74,398	2.14%	46,145	-8.56%	120,543	-2.24%
1950-1959	114,239	53.55%	65,353	41.63%	179,592	48.99%
1960-1969	222,593	94.85%	104,778	60.33%	327,371	82.29%
1970-1979	369,588	66.04%	158,019	50.81%	527,607	61.16%
1980-1984	218,396		108,674		327,070	
1985	45,073		21,891		66,964	
1986	44,917	-0.35%	21,830	-0.28%	66,747	-0.32%
1987	45,489	1.27%	22,413	2.67%	67,902	1.73%
1988	47,255	3.88%	23,130	3.20%	70,385	3.66%
1989	48,538	2.72%	23,054	-0.33%	71,592	1.71%
1990	51,699	6.51%	23,291	1.03%	74,990	4.75%
1991	52,791	2.11%	22,208	-4.65%	74,999	0.01%
1992	53,774	1.86%	21,686	-2.35%	75,460	0.61%
1993	52,615	-2.16%	21,255	-1.99%	73,870	-2.11%
1994	51,941	-1.28%	20,287	-4.55%	72,228	-2.22%
1995	49,364	-4.96%	18,341	-9.59%	67,705	-6.26%
1996	46,741	-5.31%	16,407	-10.54%	63,148	-6.73%
1997	45,506	-2.64%	17,017	3.72%	62,523	-0.99%
1998	40,578	-10.83%	15,060	-11.50%	55,638	-11.01%
1999	33,359	-17.79%	12,037	-20.07%	45,396	-18.41%
2000	13,895	-58.35%	2,795	-76.78%	16,690	-63.23%
2001	791		11		802	
Other	25,724		16,283		42,007	
Totals	**2,199,165**		**1,179,107**		3,378,272	

Table 2-3
WorldCat Records - Subject Divisions, 50 Year Range

SUBJECT DIVISION	1800-1849	1850-1899	1900-1949	1950-1999	TOTAL 1800-1999
AGRICULTURE	441	1,987	8,442	37,577	48,447
ANTHROPOLOGY	160	650	1,958	11,221	13,989
ART AND ARCHITECTURE	942	3,101	12,065	74,595	90,703
BIOLOGICAL SCIENCES	1,020	3,043	6,567	33,065	43,695
BUSINESS AND ECONOMICS	4,342	5,128	23,863	170,898	204,231
CHEMISTRY	118	358	1,474	7,831	9,781
COMPUTER SCIENCE	0	4	35	13,975	14,014
EDUCATION	565	1,992	9,142	82,561	94,260
ENGINEERING AND TECHNOLOGY	969	3,966	16,196	126,088	147,219
GEOGRAPHY AND EARTH SCIENCES	819	2,676	6,148	42,378	52,021
HISTORY AND AUXILIARY SCIENCES	7,677	23,375	56,260	194,311	281,623
LANGUAGE, LINGUISTICS, AND LITERATURE	12,932	34,256	85,734	313,458	446,380
LAW	1,602	3,571	6,248	71,896	83,317
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	1,567	3,262	8,739	51,312	64,880
MATHEMATICS	410	1,047	2,316	16,234	20,007
MEDICINE	957	2,301	6,580	56,179	66,017
MUSIC	567	1,594	4,820	18,790	25,771
PERFORMING ARTS	109	371	1,838	13,016	15,334
PHILOSOPHY AND RELIGION	7,405	18,945	33,109	107,826	167,285
PHYSICAL EDUCATION AND RECREATION	127	547	2,228	19,989	22,891
PHYSICAL SCIENCES	820	1,926	4,822	23,796	31,364
POLITICAL SCIENCE	1,062	3,234	10,801	56,529	71,626
PSYCHOLOGY	153	621	2,598	18,217	21,589
SOCIOLOGY	866	2,315	6,816	72,714	82,711
Subject Analysis	**45,630**	**120,270**	**318,799**	**1,634,456**	**2,119,155**
No Call Number Present	48,052	98,366	219,922	736,731	1,103,071
Grand Total	**93,682**	**218,636**	**538,721**	**2,371,187**	**3,222,226**

Table 2-4
WorldCat Records - Subject Divisions, 10 Year Range

SUBJECT DIVISION	1900-1909	1910-1919	1920-1929	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1989	1990-1999
AGRICULTURE	1,285	1,536	1,642	1,993	1,986	3,418	5,046	8,595	10,590	9,928
ANTHROPOLOGY	354	296	419	484	405	732	1,231	2,396	3,001	**3,861**
ART AND ARCHITECTURE	2,159	1,641	2,728	2,742	2,795	5,094	10,317	16,617	20,903	**21,664**
BIOLOGICAL SCIENCES	1,399	1,009	1,263	1,556	1,340	2,241	4,332	7,738	9,215	**9,539**
BUSINESS AND ECONOMICS	3,316	3,253	4,610	6,159	6,525	9,513	20,515	39,372	50,447	**51,051**
CHEMISTRY	230	177	327	368	372	759	1,527	1,894	1,971	1,680
COMPUTER SCIENCE	10	3	2	1	19	66	546	1,606	4,592	**7,165**
EDUCATION	1,148	1,143	1,693	2,848	2,310	4,385	10,821	21,417	24,155	21,783
ENGINEERING AND TECHNOLOGY	2,575	2,273	2,976	3,532	4,840	8,399	15,575	30,872	37,887	33,355
GEOGRAPHY AND EARTH SCIENCES	1,196	973	1,197	1,412	1,370	3,518	6,237	9,633	11,571	11,419
HISTORY AND AUXILIARY SCIENCES	11,360	9,465	9,745	12,657	13,033	14,621	28,175	43,055	51,345	**57,115**
LANGUAGE, LINGUISTICS, AND LITERATURE	18,443	12,916	18,031	17,940	18,404	27,146	50,991	69,595	77,320	**88,406**
LAW	1,168	950	1,053	1,555	1,522	2,765	5,338	13,230	22,281	**28,282**
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	1,407	1,007	1,611	2,681	2,033	3,177	7,950	13,960	14,708	11,517
MATHEMATICS	502	382	449	446	537	977	2,681	3,749	4,262	**4,565**
MEDICINE	1,077	1,003	1,257	1,578	1,665	3,016	5,761	11,792	16,445	**19,165**
MUSIC	973	686	943	1,077	1,141	1,656	2,814	4,510	5,180	4,630
PERFORMING ARTS	257	195	362	522	502	880	1,495	2,780	3,690	**4,171**
PHILOSOPHY AND RELIGION	7,683	5,405	6,394	6,828	6,799	10,441	17,433	22,776	27,921	**29,255**
PHYSICAL EDUCATION AND RECREATION	352	282	443	606	545	1,176	2,138	4,559	5,532	**6,584**
PHYSICAL SCIENCES	902	760	924	1,122	1,114	2,150	4,513	5,124	6,057	5,952
POLITICAL SCIENCE	1,625	1,932	1,785	2,412	3,047	4,366	7,737	12,072	14,644	**17,710**
PSYCHOLOGY	447	384	631	612	524	1,021	2,161	4,520	5,044	**5,471**
SOCIOLOGY	1,145	1,074	1,317	1,710	1,570	2,722	7,259	17,726	20,907	**24,100**
Subject Analysis	**61,013**	**48,745**	**61,802**	**72,841**	**74,398**	**114,239**	**222,593**	**369,588**	**449,668**	**478,368**
No Call Number Present	42,986	37,472	42,853	50,466	46,145	65,353	104,778	158,019	220,992	187,589
Grand Total	**103,999**	**86,217**	**104,655**	**123,307**	**120,543**	**179,592**	**327,371**	**527,607**	**670,660**	**665,957**

CHAPTER THREE

Library Groupings

For the WorldCat study, libraries were sorted into five groupings for analysis by type of library: research, academic, public, school, and special/other. OCLC maintains a list of member libraries at http://www.oclc.org/contacts/libraries/. In this list libraries are coded into six types with a zero code for unknown/unclassified. For purposes of this study the seven codes were collapsed into five as follows.

Research The research libraries grouping in this study has all of the Association of Research Libraries (ARL) plus the addition of a number of national libraries and large historical or public libraries. The British Library, Cambridge University England, Trinity College, Dublin, The National Library of Australia, the National Library of Scotland, the National Library of Ireland, Hoover Institution, Boston Atheneaum, the Atheneaum of Philadelphia, and the Massachusetts Historical Society were among those libraries added into the research libraries grouping. All separate libraries affiliated with the institutions of higher education in the research grouping are also in this grouping, so that health sciences, museums, law libraries, and other libraries which might be regarded as special libraries are included in the research grouping if affiliated with one of those primary institutions.

Academic The academic libraries grouping contains all libraries affiliated with institutions of higher education not in the research libraries grouping. As with the research libraries grouping, all separate libraries affiliated with the institutions in the academic grouping are included. The number of libraries is very large, as all four year institutions, community or junior colleges, vocational or technical colleges, any academic degree granting institution in higher education would be in this grouping. A number of non-U.S. institutions of higher education in

France, Japan, Africa, Great Britain, Australia, and the Catalan region of Spain and others are included in the academic libraries grouping.

Public The public library grouping combines all public libraries, and library systems and consortia, including national libraries and state libraries not placed in the research libraries grouping. While there are a few non-U.S. libraries or library systems in this grouping, it is comprised almost entirely of U.S. public libraries of all sizes and jurisdictions.

Special The special/other grouping contains a mélange of different types of libraries including corporate, health sciences, legal, museum, and other libraries not affiliated with institutions of higher education. Government libraries are also included in this grouping. The U.S. Supreme Court and the large U.S. government agencies, the Census Bureau, the Department of Housing & Urban Development, the U.S. Department of Justice, and many military libraries are included.

School The smallest of the library groupings is the school group. There are a number of separate school libraries as well as school systems in this grouping. These are almost totally U.S. schools, both private and public.

All libraries in the study were placed in one of these five types of library groupings. The subject analysis dataset of bibliographic records with usable call numbers was used for the analyses in this chapter. The number of records with call numbers for each library grouping are shown in Table 3-1 below.

Table 3-1

Library Groupings: Total Number of Records from Subject Analysis

Research 1,745,034

Academic 1,323,165

Special 736,165

Public 727,643

School 130,309

In the *OCLC Annual Report 2000/2001,* the number and percentage share of libraries by type are reported. The type of library categories do not correspond exactly to the five types of library groupings defined for this study. As of June 30, 2001, the public library grouping comprised 17% of total libraries in OCLC. The School libraries comprised 7% of total. Adding together the different categories of special libraries and others results in 33% share for those libraries. Higher education institutions of all levels and types are another 42%. The last category of state/national libraries has 1%. These categories match fairly closely to the five types defined for the study, with the exception that the higher education category is split in the study into research and academic, with the 1% in state/national split between public and research. While 1% is not significant, in terms of the analysis, the few national libraries and other research libraries which were transferred to the research grouping hold large numbers of titles. Nonetheless, the proportions from the 2000/2001 annual report do give an idea of the contributions the type of libraries groupings each would make to the total international information commons.[1]

In this chapter, each of the library groupings is analyzed separately with a table of percentage increase/decrease in the number of bibliographic records by the time periods defined for the study. Each library grouping is also analyzed by subject divisions. The research libraries are analyzed first as the largest of the library groupings in terms of

[1]OCLC Annual Report, 2000/2001: 13.

number of titles held.

Research Libraries

The profile of the research libraries in WorldCat closely parallels the profile of the WorldCat database. The 1,745,034, the subject analysis dataset, (bibliographic records with usable call numbers), are analyzed in this section for the research libraries.

The pattern of increase/decrease in the number of records over time is shown in Table 3-2 *Research Libraries, Percentage Increase/Decrease by Years*. The same decreases for the decades of the two world wars are seen as in WorldCat. There is a 7% decrease in the 1910-1909 time frame. In the 1940-1949 decade, the number of records does not decrease but is virtually static. The publication explosion after World War II can be seen in the increases starting in the 1950s. The fiscal difficulties in research libraries, not only in the United States, but internationally, can be see in the pattern of increase/decrease beginning in the 1980s. There are small decreases in 1986 and 1987; then a number of years with small increases in numbers from 1988-1992. From a slight decline in 1994 and 1995, the decrease becomes larger in 1995-1996. After a smaller decrease in 1997, the decline is precipitous in 1998 and 1999. Since the 2000 imprint year was just ended when the data extraction for the study took place in February 2001, the last two imprint years in the study are not far enough removed at that point to be comparable to earlier years.

The pattern of a slowly declining number of records beginning in the early 1990s and becoming steeper in the latter years of the decade has been found in other similar studies on North American research libraries. (These studies are in the Review of Research in Chapter One.) While it was first thought that the decline in the late 1980s and 1990s was precipitated by the increase in both monograph and serial prices in that same time frame, more recent studies have shown that the number of titles for those years has increased in WorldCat as the imprint years become more distant from the present. The pattern which this research is finding is that for the past five years, the true acquisitions patterns for those years do not show in cataloging records data. Up until the last 10 years in the study, the number of titles per year increase annually in WorldCat. Within the last 10 years the number of records per year begins to taper off, until the last 3-5 years, in which

the decline is steep. Thus while, in current acquisitions the majority of titles being purchased may have been published in the last five years, the processes of cataloging or setting holdings in most libraries are more likely covering a longer time span. As we proceed through the analyses by library grouping we will see that the problem is most acute in the research libraries.

The percentage of total by Subject Division for the research libraries is shown in Table 3-3 *Research Libraries: Percentage of Total by Subject Division*. The language, linguistics, and literature division (LLL) has by far the largest number of titles and comprises 23.25% of all records with call numbers. Since this one grouping includes the global publication for those subject areas, it quite naturally makes up a large proportion of total records. All of the history classifications from Library of Congress AC@ through AF@ are in the one history division line which has the second highest total number of records and comprises nearly 14% of total bibliographic records in the study for the research libraries.

Another subject area with voluminous publication rates is business and economics which has the third largest total, not quite 10% percent share of total. Philosophy and religion comes in fourth with 7.46% of total and engineering and technology make up 6 percent. These five subject divisions together account for 60% of the total number of records in the sample for research library collections. The monographically oriented liberal arts disciplines in LLL, history, and philosophy/religion make up 44.6% of the total aggregated collections. The sciences, in which the book is not the primary publication format, have smaller numbers of records and thus small percentages of total. Disciplines in the social sciences, such as psychology, sociology, and education, are in the middle with regard to primary publication format, but also have small percentages of total.

The academic library group is analyzed next by the same parameters, rate of publication by time frame and subject division.

Academic Libraries

The academic library grouping is composed of libraries from institutions of higher education which are not included in the research libraries grouping. This includes all U.S. academic libraries from four-year, two-year and vocational institutions, and all non-U.S.

academic libraries that are members of OCLC. The number of academic libraries is much larger than the number of research libraries. The total number of records with call numbers for the academic library grouping is 1,323,165. Table 3-4 *Academic Libraries, Percentage Increase/Decrease by Years* shows the number by the time periods defined for the study.

We can begin to see that the increases/decreases in the number of titles by time period follow a similar pattern in all of the type of library groupings as well as WorldCat. The percentages of total for both the research libraries and the academic libraries for the three decades from 1950-1979 are very close. The research libraries have a 53% increase in the 1950s over the 1940s; the academic libraries have a 60% increase. In the 1960s the research library total increases 96%; the academic libraries increase 102%. In the 1970s both increase at approximately 59 percent.

Likewise, the same pattern of decline each year after 1992 is seen in the academic library collections. While the absolute number of titles is lower for the academic grouping, the pattern of increase and decrease is the same for both, with a difference between the research libraries and the academic libraries hovering in the range of 10,000 titles most years. While there are fewer libraries in the research grouping with much larger collections, the larger number of libraries in the academic grouping contribute almost as many titles to WorldCat as the research libraries. As we will see later, the major differences between the research libraries and the academic libraries lie in the records without call numbers.

The patterns of subject concentration are also similar in the research and academic library groupings as shown in Table 3-5 *Academic Libraries: Percentage of Total by Subject Division*. The LLL division has a 21.7% share of total in the academic libraries to a 23.25% share in the research libraries. History has 12.81% in academic libraries and 13.88% in research libraries. In the academic libraries, philosophy/religion is ranked third with 10% of total. This is relatively high in comparison to the research libraries in which that subject division is fourth with 7.5% of total. In the academic libraries, business and economics is fourth with 8.6%, whereas this subject division is third in the research libraries with 9.75%. The third and fourth ranked subject divisions by percentage of total flip flop in the two library groupings.

Engineering /technology is fifth in both groupings. The majority of the remaining

subject areas have very close to the same percentage of share in both groupings. Education has a slightly higher percentage of share in the academic library grouping at 4.54%, compared to 3.39% in the research libraries. Many of the subject areas have an uncannily close percentage of total: law with 3.89% for academic and 3.90% for research; library science, 2.94% for academic and 3.04% for research; mathematics with 1.03% for academic and .95% for research; medicine 3.54% for academic and 3.07% for research; the physical sciences, 1.48% for academic, 1.49% for research; sociology, 3.75% for academic and 3.77% for research. Thus, in terms of concentration of number of titles by subject division, the patterns are very similar for WorldCat, the research libraries, and the academic libraries.

We look next at the special/other grouping of libraries to see if these patterns are also evident in that grouping.

Special Libraries

The composition of the grouping of special/other in WorldCat includes a miscellany of subject specializations. If one scans through the list at http://www/oclc.org/contacts/libraries, corporate libraries, government agency libraries, law libraries, historical society libraries, art museum libraries and theological libraries can be identified. Some of these libraries, such as corporate libraries, most probably are concerned with current business and technical information. The historical and art museum libraries would have retrospective collections. With a few exceptions, these libraries would be smaller libraries by collection size in comparison to those in the academic library and public library groupings. The total number of titles with call numbers analyzed in this chapter is very similar for the special library and public library groupings. The first table for the special/other libraries grouping shows number of titles by imprint year.

In Table 3-6 *Special Libraries: Percentage Increase/Decrease by Years* for the special libraries we can see the patterns are very nearly the same as for the research and academic libraries with a decrease during the 1910-1919 time frame. Unlike the other two groupings there is an increase in number of records in the 1940s, but it is much smaller than the decades before and afterward. From 1985 through 1992 there are modest

increases in total each year, similar to the research and academic libraries. The decreases beginning in 1993 are not as severe as in the other two library groupings, but the same pattern nonetheless.

The totals by subject division for the special libraries might be expected to be somewhat different from the other four type of library groupings. Table 3-7 shows the subject distribution for the special libraries. The subject divisions with the highest number of titles are still LLL and history, but in the special libraries grouping these two have nearly identical percentages of total at 14 percent. Ranked third is business and economics at 10%, a not surprising finding for corporate libraries. Engineering and technology is a very close fourth rank at 9 percent. Philosophy/religion has a nearly 8% share of total, nearly the same as the research libraries. Philosophy/religion has the highest share of total in the academic libraries. Art and architecture ranks sixth, a higher ranking than in any other library grouping, most probably because of the museum libraries in the special libraries grouping. Medicine has a higher share of total than in the research and academic libraries, the highest share in the sciences other than engineering and technology. The biological sciences also have a higher share of total in the special libraries than the other two groupings. Law similarly has a higher percentage of total in the special libraries, almost one percentage point higher. Sociology has the same share of total in all three groupings with 3.5%-3.77% of total.

The subject concentrations in the special libraries are not as similar to the research and academic libraries as the percentages of total by subject in those two groupings. There are differences which can easily be attributed to the nature of the professional and discipline emphases in the special libraries. The large number of corporate, legal, medical, and technical libraries account for the higher percentages of total in business, legal, and scientific subject divisions.

Public Libraries and School Libraries: Bibliographic Records by Audience Level

The public library grouping is composed mainly of U.S. public libraries. State libraries and consortia are included in the public libraries grouping. Both the public library and the school library groupings were analyzed by adult and juvenile audience levels. This

analysis is a standard feature of the iCAS product. These two library groupings are also routinely analyzed using Dewey conspectus subject divisions. The public library monographic titles are first analyzed here for distribution of records by time period. Table 3-8 *Public Libraries: Percentage of Adult/Juvenile* shows the two audience levels of adult and juvenile with the percentage of total each comprises by time period.

As had been seen in both the research and academic library groupings, the number of titles by time period begins to increase substantially in the latter half of the 19th century, which has nearly three times the number of titles as the first half of the century in the public libraries grouping. The number of titles for public libraries in the 1st decade of the 20th century is 70% of the number for the preceding fifty years. As with the other library groupings and WorldCat, the number decreases during the decade of World War I, and begins to increase again in the1920s. The 1930s and 1940s have almost an identical number of records in public libraries. The 1950s have 40% more records than the 1940s. The number of records in the 1960s is double the number for the 1950s. The rate of increase slows somewhat in the 1970s with a 75% increase over the 1960s. Adding the last five years of the 1980s together gives a total of 77,152 records for public libraries, which added to the first half of the decade gives a total of 148,484 records for the 1980s. This is only a 17% increase over the 1970s. From 1990 through 1992, the number of records is virtually static. A slow decline annually in the number of records begins in 1993, but from 1993-1995, the numbers are again very similar with only a slight decline each year. In the last five imprint years, 1997-2001, closest to the present, the total number of records for the public libraries do show a marked decrease. The decrease in public library titles is much less, however, proportionately than the decrease in research and academic libraries.

The balance between adult and juvenile shows the records to be almost totally for adult titles in public libraries, over 95 percent. The data here probably do not truly represent the number of juvenile titles owned. AEasy books,@ and many children=s books are not classified at all by public libraries and many are not cataloged at all. If these titles are cataloged but not classified, they would be in the large dataset of ANo call number present@ records for this study. The public library grouping was not analyzed for this dataset; thus

only those records with a usable call number appear in the public library data analysis. The data here most probably do represent the distribution of titles in the adult collections, but are probably not representative of all juvenile materials.

If we compare the audience level numbers for the public library grouping with school libraries we see similar patterns. Table 3-9 *Public and School Libraries by Years* shows the two audience levels for both public and school libraries by time period. The differences in the balance of the collections can be seen in the ratios between the two audience levels within the two library groupings. The public libraries, with adult collections comprising over 90% of total, have a ratio of 21 to one, adult to juvenile. For school libraries the ratio of adult to juvenile is 5 to one.

Table 3-10 *School Libraries: Percentage of Adult/Juvenile* shows the percentage of adult/juvenile by time period for the school library group. From 1985 to present, the percentage of total records for juvenile titles steadily increases. As with the public library group, it is not known what proportion of all juvenile materials owned by school libraries is represented in this study by cataloged records in WorldCat. As with public libraries, the materials suitable for the pre-school and elementary grades may be uncataloged in many school libraries. Juvenile materials represent 15% of the total records for school libraries, with 85% for titles categorized as adult. When it is taken into account that the adult materials would be reference, instructional support for teachers, and most high school library titles and that much of the juvenile collections may not be classified, 15% seems to be a fairly substantial proportion. The bibliographic records for public libraries are analyzed next by subject.

Public Libraries Subject Analysis

One table has been constructed for both public and school library titles by subject: Table 3-11 *Public and School Libraries Records by Subject*. If we look at the total number of titles by subject divisions for public libraries it can be seen that the juvenile titles cluster in a few subject areas. The largest number of juvenile titles is in the LLL division with16,492 records. The history division, by comparison, only has 2,373 titles, which is only slightly more than the biological sciences (2,019). The remaining areas with over 1,000

titles are all sciences with the exception of philosophy/religion (1,025) and geography (1,038).

The records for the adult titles analyzed by subject division have a similar pattern to that found in all of the library groupings. The largest number of titles (146,479) occurs in the LLL division, which has the largest publication rate. The second largest total (98,331), although considerably less, is for the history division. Business and economics is third with 63,250 titles. Both engineering/technology (47,008) and philosophy/religion (45,076) have similar totals. Art/architecture (35,765) and sociology (33,827) are next in the hierarchy of most to least. Education, geography, law, library science/reference, medicine, all have totals in the low 20,000 range, with political science at 26,068. Biological sciences (12,613) and physical education (10,798) have over 10,000 titles. The remaining subject divisions have under 10,000 titles.

School Libraries Subject Analysis

Both adult and juvenile titles are also shown for school libraries in Table 3-11 *Public and School Libraries Records by Subject*. Out of a total of 33,060 juvenile titles, the highest number, 9,376, are in the LLL division, 28% of the total. The history division has a very small number in comparison (1,672), less than one percent. Biological sciences has almost as many as history at 1,491. The remaining subject divisions all have less than 1,000 titles. Anthropology, engineering, geography, physical education, and physical sciences have over 500 titles. From these numbers it appears that history is on a par with the sciences and the social/cultural areas of anthropology and geography. These proportions would seem to reflect elementary school curricula in which the concentration is on learning to read. History is considered a more advanced subject usually taught beginning in the middle school grades.

The subject concentrations by adult level do show less disparity between the number of titles in LLL (25,471) and in history (13,207). The third highest number of records in adult school library collections in the aggregate is in philosophy/religion with 7,498. The art and architecture division is next with 6,148 titles. School libraries are the only grouping in which art and architecture has such a high proportion of the collection.

Sociology has 5,910. Law has 5,480 titles followed closely by business and economics with 5,381, and then engineering and technology with 5,070 titles. Education is next with 4,686 titles, closely matched by political science with 4,587, and medicine 4,345. The remaining subject divisions have fewer titles. With the exception of the engineering/technology division, the sciences have the lowest numbers. It must be borne in mind that publication rates are lower in the sciences also so that these concentrations by subject may reflect the universe of publication by subject rather than curricular emphases. The number for education seems on the low side as the schools are educational institutions.

While these data for the school libraries grouping are interesting there is a such a small number of schools or school districts in the sample dataset that it cannot be claimed that the data in this study are representative of school libraries in general. The percentages by adult/juvenile audience levels may be typical, but again, cannot be generalized beyond this grouping of school libraries.

Table 3-12 *Library Groups by Subject Divisions* shows the number of records by the 24 subject division for each of the library groupings. The distribution of records by subject division in this table shows the distribution of titles which comprise the subject analysis set for WorldCat. In this master table, the comparative differences in the number of titles can be seen. The differences in the numbers are as much a function of size and number of libraries within each group as of collection emphases. As has become obvious from all preceding analyses, the LLL division has the largest number of records in WorldCat and all type of libraries groupings. History is also the clear second place division. Business and economics is third in three of the groupings-- special, public, and research. Philosophy is third in the other two library groupings, academic and school. The fourth ranked subject divisions vary with business and economics in the academic libraries, engineering and technology in special and public, and philosophy/religion in the research libraries. The fifth ranked subject division varies also with engineering and technology in the academic and research libraries, philosophy/religion in the special and public libraries, and sociology in the school libraries.

These distributions by subject, again, do reflect collection emphases by type of library, but they also reflect the proportions of publications by subject divisions within the

universe of publication. Subject divisions which have a high volume of publication, such as history and LLL, have the largest share of total no matter what the type of library. Subject divisions in which the volume of publication for monographs is very low, have small percentage shares, no matter what the type of library. Thus, it is the differences in concentration between the types of libraries that follow the nature of the missions of those institutions.

Now that the profiles by imprint year and subject divisions have been analyzed for both WorldCat and the five type of library groupings, specific subsets of records can be examined. Chapter four contains the analysis for unique records and title overlap between library groupings. Click here for Chapter 4.

Table 3-2
Research Libraries, Percentage Increase/Decrease by Years

Years	Total Number	Percentage Increase/Decrease
Pre-1500	149	
1500-1599	3,259	
1600-1699	9,032	177.14%
1700-1799	20,823	130.55%
1800-1849	38,319	
1850-1899	99,511	159.69%
1900-1909	44,181	
1910-1919	41,098	-6.98%
1920-1929	52,575	27.93%
1930-1939	60,885	15.81%
1940-1949	61,373	0.80%
1950-1959	93,933	53.05%
1960-1969	183,099	94.93%
1970-1979	292,594	59.80%
1980-1984	170,512	
1985	35,749	
1986	35,570	-0.50%
1987	35,442	-0.36%
1988	37,169	4.87%
1989	38,252	2.91%
1990	40,108	4.85%
1991	41,740	4.07%
1992	42,187	1.07%
1993	41,357	-1.97%
1994	40,816	-1.31%
1995	38,525	-5.61%
1996	36,495	-5.27%
1997	35,193	-3.57%
1998	30,920	-12.14%
1999	24,543	-20.62%
2000	11,178	-54.46%
2001	784	-92.99%
Other	7,663	
Totals	1,745,034	

Table 3-3
Research Libraries: Percentage of Total by Subject Division

SUBJECT DIVISION	Total by Subject	Percentage of Total
AGRICULTURE	39,269	2.25%
ANTHROPOLOGY	11,898	0.68%
ART AND ARCHITECTURE	77,830	4.46%
BIOLOGICAL SCIENCES	35,080	2.01%
BUSINESS AND ECONOMICS	170,168	9.75%
CHEMISTRY	7,827	0.45%
COMPUTER SCIENCE	10,598	0.61%
EDUCATION	59,122	3.39%
ENGINEERING AND TECHNOLOGY	106,844	6.12%
GEOGRAPHY AND EARTH SCIENCES	39,677	2.27%
HISTORY AND AUXILIARY SCIENCES	242,274	13.88%
LANGUAGE, LINGUISTICS, AND LITERATURE	405,678	23.25%
LAW	68,062	3.90%
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	53,020	3.04%
MATHEMATICS	16,603	0.95%
MEDICINE	53,623	3.07%
MUSIC	21,263	1.22%
PERFORMING ARTS	13,446	0.77%
PHILOSOPHY AND RELIGION	130,185	7.46%
PHYSICAL EDUCATION AND RECREATION	16,943	0.97%
PHYSICAL SCIENCES	26,087	1.49%
POLITICAL SCIENCE	57,386	3.29%
PSYCHOLOGY	16,281	0.93%
SOCIOLOGY	65,870	3.77%
TOTAL	1,745,034	

Table 3-4
Academic Libraries, Percentage Increase/Decrease by Years

Years	Total Number	Percentage Increase/Decrease
Pre-1500	89	
1500-1599	1,736	
1600-1699	6,378	267.40%
1700-1799	13,552	112.48%
1800-1849	28,327	109.02%
1850-1899	73,957	161.08%
1900-1909	33,376	
1910-1919	29,465	-11.72%
1920-1929	39,016	32.41%
1930-1939	44,334	13.63%
1940-1949	45,908	3.55%
1950-1959	73,380	59.84%
1960-1969	148,103	101.83%
1970-1979	236,103	59.42%
1980-1984	129,344	
1985	26,383	
1986	26,268	-0.44%
1987	26,786	1.97%
1988	27,916	4.22%
1989	28,517	2.15%
1990	30,111	5.59%
1991	30,508	1.32%
1992	30,751	0.80%
1993	29,624	-3.66%
1994	29,473	-0.51%
1995	28,135	-4.54%
1996	27,010	-4.00%
1997	25,599	-5.22%
1998	23,252	-9.17%
1999	18,247	-21.53%
2000	5,207	-71.46%
2001	41	-99.21%
Other	6,269	
Totals	1,323,165	

Table 3-5
Academic Libraries: Percentage of Total by Subject Division

SUBJECT DIVISION	Total by Subject	Percentage of Total
AGRICULTURE	25,007	1.89%
ANTHROPOLOGY	8,376	0.63%
ART AND ARCHITECTURE	57,128	4.32%
BIOLOGICAL SCIENCES	27,169	2.05%
BUSINESS AND ECONOMICS	113,385	8.57%
CHEMISTRY	6,210	0.47%
COMPUTER SCIENCE	10,188	0.77%
EDUCATION	60,092	4.54%
ENGINEERING AND TECHNOLOGY	78,209	5.91%
GEOGRAPHY AND EARTH SCIENCES	29,572	2.23%
HISTORY AND AUXILIARY SCIENCES	169,503	12.81%
LANGUAGE, LINGUISTICS, AND LITERATURE	287,065	21.70%
LAW	51,487	3.89%
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	38,930	2.94%
MATHEMATICS	13,676	1.03%
MEDICINE	46,808	3.54%
MUSIC	18,004	1.36%
PERFORMING ARTS	10,091	0.76%
PHILOSOPHY AND RELIGION	132,127	9.99%
PHYSICAL EDUCATION AND RECREATION	14,597	1.10%
PHYSICAL SCIENCES	19,555	1.48%
POLITICAL SCIENCE	41,840	3.16%
PSYCHOLOGY	14,544	1.10%
SOCIOLOGY	49,602	3.75%
TOTAL	1,323,165	

Table 3-6
Special Libraries
Percentage Increase/Decrease by Years

Years	Total Number	Percentage Increase/Decrease
Pre-1500	15	
1500-1599	1,083	
1600-1699	4,451	310.99%
1700-1799	7,646	71.78%
1800-1849	16,201	111.89%
1850-1899	40,549	150.29%
1900-1909	14,721	
1910-1919	13,913	-5.49%
1920-1929	17,213	23.72%
1930-1939	20,434	18.71%
1940-1949	22,399	9.62%
1950-1959	36,727	63.97%
1960-1969	72,211	96.62%
1970-1979	122,664	69.87%
1980-1984	75,254	
1985	15,331	
1986	15,406	0.49%
1987	15,689	1.84%
1988	16,270	3.70%
1989	16,871	3.69%
1990	17,578	4.19%
1991	17,814	1.34%
1992	18,180	2.05%
1993	17,890	-1.60%
1994	18,022	0.74%
1995	17,486	-2.97%
1996	17,208	-1.59%
1997	17,415	1.20%
1998	16,370	-6.00%
1999	13,999	-14.48%
2000	4,314	-69.18%
2001	15	-99.65%
Other	14,826	
Totals	736,165	

Table 3-7
Special Libraries: Percentage of Total By Subject Division

SUBJECT DIVISION	Total by Subject	Percentage of Total
AGRICULTURE	17,268	2.35%
ANTHROPOLOGY	4,919	0.67%
ART AND ARCHITECTURE	42,679	5.80%
BIOLOGICAL SCIENCES	23,772	3.23%
BUSINESS AND ECONOMICS	73,713	10.01%
CHEMISTRY	4,101	0.56%
COMPUTER SCIENCE	7,830	1.06%
EDUCATION	16,415	2.23%
ENGINEERING AND TECHNOLOGY	66,712	9.06%
GEOGRAPHY AND EARTH SCIENCES	24,910	3.38%
HISTORY AND AUXILIARY SCIENCES	103,356	14.04%
LANGUAGE, LINGUISTICS, AND LITERATURE	107,075	14.54%
LAW	35,667	4.84%
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	19,389	2.63%
MATHEMATICS	6,859	0.93%
MEDICINE	31,536	4.28%
MUSIC	6,285	0.85%
PERFORMING ARTS	4,839	0.66%
PHILOSOPHY AND RELIGION	56,055	7.61%
PHYSICAL EDUCATION AND RECREATION	8,508	1.16%
PHYSICAL SCIENCES	15,602	2.12%
POLITICAL SCIENCE	26,765	3.64%
PSYCHOLOGY	6,154	0.84%
SOCIOLOGY	25,756	3.50%
TOTAL	736,165	

Table 3-8
Public Libraries - Percentage of Adult/Juvenile

Public Libraries					
Years	**Total**	**Adult**	**Percentage Adult**	**Juvenile**	**Percentage Juvenile**
Pre-1500	24	23	95.83%	0	0.00%
1500-1599	483	482	99.79%	0	0.00%
1600-1699	3,058	3,057	99.97%	0	0.00%
1700-1799	4,828	4,744	98.26%	83	1.72%
1800-1849	10,646	10,364	97.35%	281	2.64%
1850-1899	35,386	34,975	98.84%	410	1.16%
1900-1909	24,824	24,513	98.75%	310	1.25%
1910-1919	17,113	16,870	98.58%	242	1.41%
1920-1929	22,564	22,215	98.45%	348	1.54%
1930-1939	25,248	24,783	98.16%	464	1.84%
1940-1949	25,839	25,381	98.23%	457	1.77%
1950-1959	36,224	35,352	97.59%	871	2.40%
1960-1969	72,245	69,569	96.30%	2,675	3.70%
1970-1979	126,866	121,705	95.93%	5,160	4.07%
1980-1984	71,332	67,843	95.11%	3,488	4.89%
1985	14,674	13,864	94.48%	809	5.51%
1986	14,854	13,946	93.89%	907	6.11%
1987	15,181	14,194	93.50%	986	6.49%
1988	15,818	14,844	93.84%	973	6.15%
1989	16,625	15,492	93.18%	1,132	6.81%
1990	18,042	16,818	93.22%	1,223	6.78%
1991	18,208	16,834	92.45%	1,373	7.54%
1992	18,429	16,972	92.09%	1,456	7.90%
1993	17,927	16,496	92.02%	1,430	7.98%
1994	17,799	16,439	92.36%	1,359	7.64%
1995	17,097	15,882	92.89%	1,214	7.10%
1996	16,385	15,177	92.63%	1,207	7.37%
1997	15,813	14,501	91.70%	1,311	8.29%
1998	14,527	13,343	91.85%	1,183	8.14%
1999	11,855	10,748	90.66%	1,106	9.33%
2000	4,093	3,585	87.59%	507	12.39%
2001	24	17	70.83%	6	25.00%
Other	3,644	3,555	97.56%	89	2.44%
Totals	**727,675**	**694,583**		**33,060**	

Table 3-9
Public and School Libraries by Years

	Public Libraries		School Libraries	
Years	**Adult**	**Juvenile**	**Adult**	**Juvenile**
Pre-1500	23	0	0	0
1500-1599	482	0	11	0
1600-1699	3,057	0	16	0
1700-1799	4,744	83	49	0
1800-1849	10,364	281	222	4
1850-1899	34,975	410	1,484	17
1900-1909	24,513	310	1,242	35
1910-1919	16,870	242	1,208	20
1920-1929	22,215	348	2,003	49
1930-1939	24,783	464	2,679	119
1940-1949	25,381	457	3,422	204
1950-1959	35,352	871	6,794	530
1960-1969	69,569	2,675	15,626	1,946
1970-1979	121,705	5,160	22,695	3,416
1980-1984	67,843	3,488	12,722	2,101
1985	13,864	809	2,717	503
1986	13,946	907	2,747	543
1987	14,194	986	2,782	606
1988	14,844	973	2,807	633
1989	15,492	1,132	2,962	698
1990	16,818	1,223	2,981	792
1991	16,834	1,373	3,030	847
1992	16,972	1,456	3,003	919
1993	16,496	1,430	2,840	902
1994	16,439	1,359	2,769	841
1995	15,882	1,214	2,697	759
1996	15,177	1,207	2,582	790
1997	14,501	1,311	2,274	838
1998	13,343	1,183	1,978	767
1999	10,748	1,106	1,517	722
2000	3,585	507	354	262
2001	17	6	1	0
Other	3,555	89	224	8
Totals	**694,583**	**33,060**	**110,438**	**19,871**

Table 3-10
School Libraries - Percentage of Adult/Juvenile

School Libraries					
Years	**Total**	**Adult**	**Percentage Adult**	**Juvenile**	**Percentage Juvenile**
1500-1599	11	11	100.00%	0	0.00%
1600-1699	16	16	100.00%	0	0.00%
1700-1799	49	49	100.00%	0	0.00%
1800-1849	226	222	98.23%	4	1.77%
1850-1899	1,501	1,484	98.87%	17	1.13%
1900-1909	1,277	1,242	97.26%	35	2.74%
1910-1919	1,228	1,208	98.37%	20	1.63%
1920-1929	2,052	2,003	97.61%	49	2.39%
1930-1939	2,798	2,679	95.75%	119	4.25%
1940-1949	3,626	3,422	94.37%	204	5.63%
1950-1959	7,324	6,794	92.76%	530	7.24%
1960-1969	17,572	15,626	88.93%	1,946	11.07%
1970-1979	26,111	22,695	86.92%	3,416	13.08%
1980-1984	14,823	12,722	85.83%	2,101	14.17%
1985	3,220	2,717	84.38%	503	15.62%
1986	3,290	2,747	83.50%	543	16.50%
1987	3,388	2,782	82.11%	606	17.89%
1988	3,440	2,807	81.60%	633	18.40%
1989	3,660	2,962	80.93%	698	19.07%
1990	3,773	2,981	79.01%	792	20.99%
1991	3,877	3,030	78.15%	847	21.85%
1992	3,922	3,003	76.57%	919	23.43%
1993	3,742	2,840	75.90%	902	24.10%
1994	3,610	2,769	76.70%	841	23.30%
1995	3,456	2,697	78.04%	759	21.96%
1996	3,372	2,582	76.57%	790	23.43%
1997	3,112	2,274	73.07%	838	26.93%
1998	2,745	1,978	72.06%	767	27.94%
1999	2,239	1,517	67.75%	722	32.25%
2000	616	354	57.47%	262	42.53%
2001	1	1	100.00%	0	0.00%
Other	232	224	96.55%	8	3.45%
Totals	**130,309**	**110,438**		**19,871**	

Table 3-11
Public and School Libraries Records by Subject

SUBJECT DIVISION		Public				School		
		Adult		Juvenile		Adult		Juvenile
AGRICULTURE		14,631		425		1,442		309
ANTHROPOLOGY		5,666	[4]	1,585		1,162	[4]	948
ART AND ARCHITECTURE		35,765		750		6,148		436
BIOLOGICAL SCIENCES		12,613	[3]	2,019		2,545	[3]	1,491
BUSINESS AND ECONOMICS	[3]	63,250		411	[5]	5,381		251
CHEMISTRY		2,628		37		345		24
COMPUTER SCIENCE		4,357		48		595		31
EDUCATION		20,686		450		4,686		231
ENGINEERING AND TECHNOLOGY	[4]	47,008		1,033		5,070		682
GEOGRAPHY AND EARTH SCIENCES		23,110	[5]	1,038		3,811		718
HISTORY AND AUXILIARY SCIENCES	[2]	98,331	[2]	2,373	[2]	13,207	[2]	1,672
LANGUAGE, LINGUISTICS, AND LITERATURE	[1]	146,479	[1]	16,492	[1]	25,471	[1]	9,376
LAW		22,261		76		5,480		64
LIBRARY SCIENCE, GENERALITIES AND REFERENCE		21,552		444		2,767		243
MATHEMATICS		5,749		309		1,114		136
MEDICINE		21,973		479		4,345		325
MUSIC		9,378		257		1,384		169
PERFORMING ARTS		6,236		206		1,386		148
PHILOSOPHY AND RELIGION	[5]	45,076		1,025	[3]	7,498		304
PHYSICAL EDUCATION AND RECREATION		10,798		932		2,685	[5]	722
PHYSICAL SCIENCES		8,364		745		1,515		505
POLITICAL SCIENCE		26,068		995		4,587		474
PSYCHOLOGY		8,777		185		1,904		136
SOCIOLOGY		33,827		746	[4]	5,910		476
TOTAL		**694,583**		**14,631**		**110,438**		**19,871**

Table 3-12
Each Library Group by Subject Division

		All Library Groups								
SUBJECT DIVISION		**Academic**		**Special**		**Public**		**School**		**Research**
AGRICULTURE		25,007		17,268		15,056		1,751		39,269
ANTHROPOLOGY		8,376		4,919		7,251		2,110		11,898
ART AND ARCHITECTURE		57,128		42,679		36,515	**[4]**	6,584		77,830
BIOLOGICAL SCIENCES		27,169		23,772		14,632		4,036		35,080
BUSINESS AND ECONOMICS	**[4]**	113,385	**[3]**	73,713	**[3]**	63,661		5,632	**[3]**	170,168
CHEMISTRY		6,210		4,101		2,665		369		7,827
COMPUTER SCIENCE		10,188		7,830		4,405		626		10,598
EDUCATION		60,092		16,415		21,136		4,917		59,122
ENGINEERING AND TECHNOLOGY	**[5]**	78,209	**[4]**	66,712	**[4]**	48,041		5,752	**[5]**	106,844
GEOGRAPHY AND EARTH SCIENCES		29,572		24,910		24,148		4,529		39,677
HISTORY AND AUXILIARY SCIENCES	**[2]**	169,503	**[2]**	103,356	**[2]**	100,704	**[2]**	14,879	**[2]**	242,274
LANGUAGE, LINGUISTICS, AND LITERATURE	**[1]**	287,065	**[1]**	107,075	**[1]**	162,971	**[1]**	34,847	**[1]**	405,678
LAW		51,487		35,667		22,337		5,544		68,062
LIBRARY SCIENCE, GENERALITIES AND REFERENCE		38,930		19,389		21,996		3,010		53,020
MATHEMATICS		13,676		6,859		6,058		1,250		16,603
MEDICINE		46,808		31,536		22,452		4,670		53,623
MUSIC		18,004		6,285		9,635		1,553		21,263
PERFORMING ARTS		10,091		4,839		6,442		1,534		13,446
PHILOSOPHY AND RELIGION	**[3]**	132,127	**[5]**	56,055	**[5]**	46,101	**[3]**	7,802	**[4]**	130,185
PHYSICAL EDUCATION AND RECREATION		14,597		8,508		11,730		3,407		16,943
PHYSICAL SCIENCES		19,555		15,602		9,109		2,020		26,087
POLITICAL SCIENCE		41,840		26,765		27,063		5,061		57,386
PSYCHOLOGY		14,544		6,154		8,962		2,040		16,281
SOCIOLOGY		49,602		25,756		34,573	**[5]**	6,386		65,870
TOTAL		**1,323,165**		**736,165**		**727,643**		**130,309**		**1,745,034**

*** Highlighted bracketed numbers indicated ranking.**

Chapter Four

Diversity of Resources

Unique Titles

The variable, number of unique records, is a measure of the variety or diversity of resources within an aggregated resources base. A record is considered to be unique if there is only one library holding symbol attached to the record. Thus, as far as the WorldCat database is concerned, there is only one known copy of that edition of the title. There can be variant records which do inflate the number of Aunique@ titles. Duplicate detection scans are run periodically at OCLC to eliminate records which are obvious duplicates. For purposes of data analysis, the assumption in this study is that the number of unique records do, by and large, reflect the number of different titles in the database that have only one holding library location.

The data analysis for unique records in WorldCat and the research libraries grouping was performed in the second phase of the iCAS analysis in October 2001. The totals from the October analysis are used for WorldCat and the research libraries, but the unique analysis for the other four library groupings is from the June 15, 2001 first phase data analysis. The first section in this chapter analyzes the unique records in WorldCat and the research libraries, after which the distribution of unique records for the other library groupings is analyzed.

WorldCat and Research Libraries

The tables from the second phase analysis for WorldCat and the research libraries show the distribution of all unique records in the study. Table 4-1 *WorldCat Unique RecordsBAll titles Held by Date (All Divisions)* shows the number of unique records in the sample according to the time periods established for the study and by the 24 subject divisions. These records are sub-totaled for the number of unique records with call number which can be analyzed by subject. The number of records without call numbers, ANo Call Number Present,@ appears in the table after the subject subtotal line. These records can

only be analyzed by imprint year. The grand total for each time period is the last line across the bottom of the table. Table 4-2 *Unique Research Records—All Titles Held by Date (All Divisions)* has the same format for unique records for the research libraries grouping. WorldCat and the research libraries grouping have the number of records with no call number present analyzed for unique records. The unique records for the other library groupings were only analyzed from the dataset of records with call numbers.

In considering the variety of resources represented by monographic bibliographic records in WorldCat, the ratio of unique records to total records demonstrates that there is, indeed, great variety. The total number of unique records, that is records which have only one library holding location in the sample, for WorldCat, is 1,794,913. Thus, of the sample records, 3,378,272, analyzed for WorldCat, 53% of those records have only one library holding location. The same data for the group of research libraries can be compared with the WorldCat data. Of 1,745,034, total records for the research libraries, 1,108,262 or 63.5% of the records are unique. This is an expected pattern, in that within the research libraries grouping, there would be a higher percentage of unique, or one of a kind, materials than in the other library groupings within WorldCat. As the research group contains many of the major research libraries in the world, including the Library of Congress, the British Library, the New York Public Research Libraries, and a number of national libraries around the world, it is not surprising that the percentage share of unique records is high.

Table 4-3 *Total Number of Unique Bibliographic Records: WorldCat and Research Libraries*, shows the total number of records for WorldCat and the research libraries, the number of unique titles for each, divided into those unique records with call number and those unique records without call number, and the percentages of total for each category.

For WorldCat, the unique records with call numbers comprise 27.82% of all the usable records in the sample. The unique records without call numbers comprise 25.31% of total records. Added together the total unique records for WorldCat comprise 53.13% of all records in the study.

The research libraries have a higher percentage of total in unique records comprising 63.51% of research library records. Those unique records with call numbers for the research libraries are 33.23% of the total research library records. The percentage of total records for those unique records without call numbers is 30.38%. For both WorldCat and the research libraries, the unique, no call number records comprise approximately 48% of the total number of unique records.

When the percentages are compared for unique records divided into those with, and those without call numbers, to all records in the sample, WorldCat and the research libraries have very similar percentages. In WorldCat the records with call numbers make up nearly 28% of total records while those without call numbers are right at 25% of total records. Thus, the 53.13% of total records consisting of unique records in WorldCat is almost evenly divided between unique records with and without call numbers. The situation is similar for the research libraries, with close to 30% in each category of unique records adding up to the 63.51% of unique records in the aggregated research libraries collections. Overall, the number of unique records for the research libraries is 62% of the number of unique records in WorldCat. While 62% is a large proportion, it is evident that there are unique titles in WorldCat being contributed by libraries in the other type of library groupings.

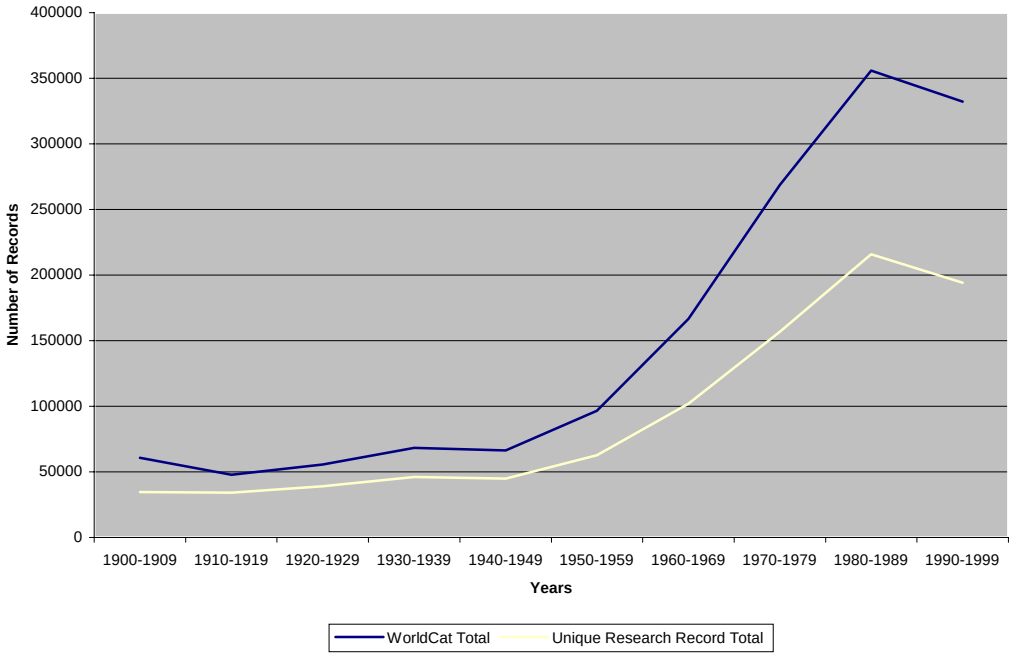
The number of unique records by publication year ranges for WorldCat and the research libraries are shown in Tables 4-1 and 4-2. The increase in publication from the first half of the nineteenth century to the latter half can be seen in the 65% increase in the research library collections and 66% increase in the number of WorldCat unique records between the two time periods.

Beginning with the 20th century, the number of unique records in WorldCat rises slowly but steadily until the 1940-1949 time period in which there is a decline, no doubt caused by the decline in publication rates during World War II. As would be expected there is a dramatic increase in unique records in both WorldCat and the research libraries beginning in 1950, another hefty increase in the 1960s, and increases again in the

decades of the 1970s and 1980s. The year by year analysis beginning in 1985 shows that the number of unique records only increases slightly until 1989 and then levels off during the 1989-1993 years. Beginning in 1994, the number declines and plummets precipitously between 1998 and 2000.

Table 4-4 *Total Number of Unique Records by Decade, WorldCat and Research Libraries* shows the number of unique records for the 20th century with the percentage of unique records in WorldCat contributed by the research libraries. Figure 4-1 *WorldCat and Unique Record Comparison* graphs the data in Table 4-4. A comparison of the unique records for WorldCat and the research libraries in Table 4-4 shows that the research libraries contribute a large percentage of the unique records in WorldCat.

Figure 4-1
WorldCat Unique Records By All Titles Held by Date (All Divisions)



The data in Table 4-4 contribute to the examination of the drop-off in numbers for the latter part of the 1990s. As can be seen, the percentage of unique research records of

the total unique records in WorldCat is very similar for the first and last decades of the 20th century. The percentage of the unique records in WorldCat attributable to the research libraries stays within the 60% range from the 1930s through the 1960s. The percentage for the 1970s and the 1990s is virtually identical. The highest percentages of total are from 1910-1929, at approximately 70 percent. The lowest is in the decade of the 1960s, but that is at 60 percent. While the number of unique records in WorldCat declines by 7% from the decade of the 1980s through the decade of the 1990s, the decline is 10% in the research library collections. The declines in the number of unique records are not as steep as the decline in overall total of records in the 1990s. This shows that there are many records in the 1990s to which other libraries have not added their holdings location. Since the research libraries contribute the majority of the unique records in WorldCat, the pattern for WorldCat is essentially set by the research libraries.

The unique records for WorldCat and the research libraries are analyzed by subject divisions in Table 4-5 *Unique Records by Subject Division: WorldCat and Research Libraries*. The subject analysis for unique records was performed upon the dataset of those records with call numbers for all of the library groupings and WorldCat. The rankings by subject are in brackets and highlighted in Table 4-5.

For those unique records with call numbers in WorldCat, nearly 10% (175,353) are in the languages/linguistics/literature subject division. The next highest concentration of unique records occurs in the history subject division with 114,394 unique records. The third highest concentration is in business and economics (97,507) followed by philosophy and religion (75,763). These rankings follow the pattern for total records in the database. These are the subject areas most oriented toward monograph publishing, and it is not unexpected to find the highest concentrations of unique records in these subject areas.

Among the social sciences, education has a relatively high number of unique records at 62,281. Law, library science, political science, and sociology all have unique records in the 30,000-39,000 range. Psychology has one of the lowest concentrations of unique titles for a field with a high publication rate, only 10,206 unique records in

psychology. This could point to a pattern of a core of titles which are purchased by many libraries.

As might be expected, the majority of the science subject divisions have lower numbers of unique records as contrasted with the high numbers for the humanities and the core social sciences. Engineering and technology as one subject division have the highest number of unique records in the sciences with 71,342 records, fifth in the number of unique records in WorldCat. Medicine and agriculture have similar numbers of unique records 26,174 and 23,882 respectively. Geography and earth sciences are in the same range with 22,834. The biological sciences and physical sciences have 17,486 and 13,961 unique records. Mathematics is a science field with more monograph publication, but it has a relatively low number of unique records comparatively at 9,075. As with psychology, it is possible that in mathematics there is a tight core of titles being acquired by a number of libraries.

It can be seen that there are large differences in the number of unique records by subject divisions and that these numbers reflect the universe of publication in those fields. Those with the highest publication rate for monographs, the humanities and the core socials sciences, have the largest numbers of unique records. The smaller fields in the social sciences and the core science fields have lower publication rates and also lower numbers of unique records.

When the number of unique records by subject division and publication year for WorldCat are compared with the same set of records for the group of research libraries, in Table 4-5, the similarities are very apparent. The subject division totals are also given for the academic libraries grouping. The research libraries have a total of 579,955 unique records with call numbers in the subject analysis. The ranking of subject divisions from highest to lowest number of unique records is virtually the same for both WorldCat and the research libraries. The first through third ranked subject divisions are the same: LLL, history, and business and economics. The fourth and fifth ranked subject divisions flip flop in ranking between the two: philosophy/religion and engineering/technology.

When the academic library grouping is compared with WorldCat and the research libraries, it can be seen that the ranking of unique titles by subject division is quite different. Philosophy/religion ranks first in the academic libraries, education second, and LLL third. History is fourth and engineering/technology fifth.

The proportion of the records in WorldCat that the research libraries hold is also shown in Table 4-5. The top ranked division in unique titles for the research libraries, LLL, represents 75% of the unique titles in that subject division in WorldCat. The majority of the subject divisions for the research libraries are in the 60% range for proportion of unique titles in WorldCat. Only four subject divisions represent less than a 50% share of the unique titles in WorldCat: computer science, education, philosophy/religion, and physical education and recreation. The percentage share is not calculated for the academic libraries which with a much smaller total of unique records does not contribute as large percentages of share to WorldCat as the research libraries.

The patterns of distribution and proportions of unique records by subject division are almost identical between WorldCat and the grouping of research libraries. As has been observed earlier, the research libraries, to a great extent, set the pattern of distribution for unique records in WorldCat.

The unique records with call numbers are analyzed in the next section by the type of library groupings.

Unique Records by Library Grouping

The number of unique records in WorldCat and all of the library groupings is shown in Table 4-6 *Unique Bibliographic Records from Subject Analysis by Library Grouping*. From Table 4-6 the percentage of total records the unique records represent for each grouping can be seen. It is obvious that the concentration of unique records in WorldCat and the grouping of research libraries is much higher than in the other library groupings. The smaller number of total titles and the missions of these other libraries make it less likely there will to be large numbers of unique materials that are not also held within the

grouping of research libraries. Even though the academic libraries have slightly over 15% in unique titles, that is far below the 33% rate for research libraries. The difference between the missions of academic libraries and research libraries shows clearly in the percentage of unique records for the two groups.

The differences in concentrations of unique records in the library groupings by subject division are shown in Table 4-7 *Unique Records by Subject by Library Groupings*. The top six subject divisions for each library grouping are shown by the rankings in brackets highlighted in the table. It is interesting that the subject division with the highest number of unique titles is different for each library grouping. The LLL division has the highest number in the research libraries. For academic libraries it is philosophy/religion. For special libraries the highest number of unique titles is in engineering and technology. The public libraries have the most unique titles in history. Although the school group has a very small number of unique titles, the highest number is in education. These differences seem to characterize the collections and define the subject emphases of the library groupings.

The academic group would have four-year institutions which grant graduate degrees in a small number of subject areas. Many of the academic libraries are in liberal arts institutions and also in institutions with religious affiliation. Both of these emphases would explain the contribution of unique titles in philosophy/religion to the database by the academic libraries. In absolute numbers the totals for philosophy/religion in the research libraries and the academic libraries are very close showing just how strong that subject division is in the academic library collections.

As many special libraries are corporate libraries, the ranking of engineering and technology as having the highest number of unique titles, with business/economics the second highest, points to the nature of those libraries.

The public libraries have the highest number of unique titles in history, with LLL a close second. Local history may be the reason that history comes out on top in the public library group. The fiction and literature collections in public libraries may overlap a great deal with those same subject areas in academic and research libraries, thus resulting in a

lower number of unique titles in LLL in the special libraries.

History is second in the research libraries in number of unique titles, but fourth in the academic library group which has education as second. In the research libraries business/economics comes in third, with engineering/technology fourth, and philosophy/religion fifth highest. For the academic group LLL is third, history is fourth, and business/economics fifth. Whereas education ranks second in the academic libraries, it ranks sixth in the research libraries.

In the special libraries, history ranks third with law fourth and philosophy/religion fifth. There are a number of historical societies in the special library grouping which can account for the contribution of a number of unique titles to the history and philosophy subject divisions.

In the public library grouping, political science ranks fifth and sociology sixth, the only group in which those titles are in the top six.

The analysis by unique title rankings in the five different types of library groupings does highlight the different composition of the collections in those groupings. The smaller or larger number of total titles for the groupings does not affect the analysis by grouping because the rankings show the different emphases in collections.

For resources sharing, the analysis of unique titles by type of library and subject groupings shows that each type and size of library are contributing unique records to the WorldCat database. Although the percentage of unique records in the research library grouping is much larger than the other groupings, those other four groupings contribute 359,959 unique records to the sample database. These records are 38% of the total unique records in the WorldCat sample.

Title Overlap

In a way, title overlap is an opposite measure from that of unique records. Whereas the number of unique records is an indication of variety in resources, title overlap provides a gauge of the amount of duplication or overlap of the same titles among groups of

libraries. In Table 4-8 *All LibrariesBTitle Overlap,* the percentage overlap between each of the five library groupings is shown. For each grouping, the number of titles in common between the primary library and another grouping is shown in the cell in which the two library groupings intersect. The table is read horizontally from left to right. For example, of the total number of titles in academic libraries, 43.5% of those titles are also held in the public library grouping.

The overlap between academic and research libraries is the highest overlap for the academic libraries at 78.4%. It would be expected that these two groupings would have a large number of titles in common.

The special libraries grouping has a high degree of overlap with both the academic and research library groupings: 77.3% with academic libraries and 81.4% with research libraries. Public libraries have 78.4% of titles in common with academic libraries and 80.7% in common with research libraries. For public libraries these numbers of records represent titles with call numbers present and these would be the titles most likely to overlap with other library groups. It is not known to what extent, fiction, children=s materials and other non-cataloged materials which might add considerably to the total number of titles in public library collections are unrepresented in the data in this WorldCat study.

It has been shown in the data on unique titles that WorldCat and the research libraries have a large proportion of unique titles to total titles. While the public libraries own 80.7% of the same titles as research libraries and 78.4% the same as academic libraries, research libraries, by virtue of much larger size, have only a 34% overlap with public libraries. In other words, only 34% of the titles that are owned by the libraries in the research grouping are also owned by libraries in the public library grouping.

Size is destiny in that the larger a collection, the greater the possibility that there will be low overlap with libraries of smaller size. The research libraries have 1,036,807 titles which overlap with the holdings of the other library groups and that figure represents 59.4% of the total number of titles owned by the research libraries. That 1,036,807 titles is a 60% percent overlap with the titles owned by academic libraries. Of the titles owned by

academic libraries 78.4% of them are also owned within the research libraries grouping. Research libraries own 33.9% of the same titles as public libraries and 34.3% of the same titles as special libraries. If we look back at the data on unique titles, we can see the correspondence between the low overlap of titles owned by the research libraries and the high proportion of unique titles within the group of research libraries.

Table 4-8 shows that the libraries in the special libraries grouping have high overlap with academic and research and lower overlap with the collections in public libraries. The overlap between special and research is at 81.4%, the highest rate of overlap between any of the library groupings with the exception of the school grouping.

The school grouping has a very small number of titles in comparison to the other four library groupings. The overlap of the school titles with the other four groups is highest with the academic grouping at 96.3% with public next at 92.3% and research only slightly less at 91.7%. The degree of overlap is probably reflective of the academic and reference titles acquired for instructional support and staff development and also overlap with the children=s and young adult materials of public libraries.

Another way of viewing title overlap is to look at the number of titles shared by more than one library grouping. In the iCAS analysis, unique records and shared records by library grouping are analyzed in one set of tables. In the analysis in this chapter, uniqueness and overlap are analyzed separately. The iCAS tables show the number of unique records and the shared 2-5 by subject for each library Table 4-9 *Total Number of Shared Records by Library Grouping* shows the number of unique records by library grouping, but also the number of those records with two, three or four library holdings symbols.

Table 4-9 shows the number of titles each library group shares with the other library groups. The research libraries have 579,955 unique titles that are not shared with any of the other four library groupings. The research libraries have the highest number of unique titles as has been seen in the analysis above. There are 425,731 titles in the research library grouping which are shared with at least one other library grouping, making each of

those titles have two library locations within the sample dataset. Research libraries share 342,731 titles with two other library groupings, almost the same absolute number as academic libraries at 341,453. The percentages of total differ in that research libraries share nearly one fifth of titles with two other groups, but for academic libraries the percentage is higher at 28% shared with two other groups. For special (27%) and public (26.3%), a similar number of titles shared with two other library groupings, also results in similar percentages of total.

For the research and academic libraries the number and percentage of total for titles shared decreases as the number of library groups sharing increase. But the opposite obtains for the other three library groupings: as the number of libraries increases, so does the number of titles shared by each library group. Special libraries have 13% of titles shared with one other library group, but 27% shared with two, and 36% shared with three and 14% shared with all four other groups. Public libraries have close to the same numbers and percentages as the special libraries: 13% shared with one, 26% shared with two, 37% shared with three, and 14% shared with all four other groups. Less than one percent of titles are shared by the school grouping with one other group. School shared with two is still less than 1% at .5%. School shared with three other groups is 15% of all school titles. A high 77% of titles are shared by the school libraries with all of the other library groups. This stands to reason, in that a much smaller number of titles than owned by the other type of library groupings would repeat many of the same titles.

The 101,378 titles which are shared by all five library groupings are slightly less than 5% of the total number of titles in the subject analysis dataset for WorldCat. Seen in this way, the extent of overlap between the library groupings is not great. It is determined more by the smaller library groupings, in that they have fewer titles, and the overlap is constrained by the overall number of titles in those groupings. The academic libraries have a low rate of unique titles at 15.4% and a large number of titles shared with at least one other library groupings, 403,731, right at 33% of all academic library group titles. The research libraries share only 24% of total subject analysis titles with another library

grouping. Both public libraries and special libraries share 13% of titles with another library grouping. The school group has over one fifth of titles (22.4%) shared with another library group.

The data on title overlap among the five type of library groupings lead to an assumption that if, the school library grouping with a very small number of titles in comparison to the other library groupings were removed from the analysis, the number of titles in common among the other four groups might be much larger. The number of titles shared by four of the five groups is very similar for all four of the groups other than the school grouping. The range for all four is 261,903 to 272, 235, just over 10,000 titles difference in the absolute numbers of titles shared by 4. The academic libraries own the most titles in common with the other three, followed by public libraries, research libraries, and then special libraries.

This pattern does not obtain for the four library groupings with titles shared by three and by two library groups. The numbers for the research libraries and the academic libraries are much larger than those for special, public, and school, leading to the conclusions that research and academic are sharing many of the same titles between the two groups. As the number of library groups sharing increases, the research and academic groups share fewer titles with the other three groups. This is probably mostly a function of size again. The two groups with the largest collections have more titles unique within the study database than the other three with smaller collections which contribute fewer titles to the analysis. What is somewhat surprising is that the special library group and the public library group have such similar absolute numbers. At least for the study sample from WorldCat, there appear to be many similarities between the collection patterns of these two groups of libraries.

The small number of titles the school group has in common with the other libraries under the 101,378 it shares with the other four type of library groupings and the very small percentage of unique titles to total the school libraries contribute to the total titles in the subject analysis indicate that the school grouping only owns titles also owned by all the

other libraries. In other words, there are 101,378 titles that are owned by at least one library in each of the five library groupings. If we add together the total of unique titles with the number of shared titles we have the total number of titles in the WorldCat dataset of records with call numbers, the dataset which is used for the subject analysis.

Table 4-10 has been constructed from the iCAS All LibrariesBUniqueness (Subject Divisions). In Table 4-10 *Overlap Analysis by Library Grouping,* both commonalities and differences between the library groupings can be seen by subject division. This analysis is most useful when used in a consortia because it can show the overlap by subject within the consortia.

Table 4-10 shows the number of shared titles for each of the library groupings by subject division, one table for each library grouping. The last column, AShared by 5" has the same number in each of the library groupings, as these are the titles all five type of library groupings have in common. In other words, there are 101,378 titles that are owned by at least one library in each of the five library groupings. And the distribution of those 101,378 titles by subject division are shown in Table 4-10.

While it was observed above that the special and public library groupings have very similar absolute numbers in Table 4-9, the differences between the collections can be seen in 4-10. The differences are by and large, discipline based. The special libraries have higher numbers of shared titles in the sciences while public libraries have much higher numbers in history and the languages, literature, linguistics classifications. Both have about the same numbers in business. Likewise, the numbers are almost the same for library science, generalities and reference. Public libraries have fewer titles in art and architecture shared with the other groups until the AShared by 4" column, indicating that public libraries own more unique titles than the special library grouping. The AShared by 4" column for both public and special has similar numbers by subject with the research and academic, but this makes sense in that those are the four groups which would be sharing the same titles. As observed above, the number of AShared by 4" titles for those four groups are very close to the same values.

The AShared by Four@ category is probably a reflection of the overlap between the four type of library groupings without the school titles included. The AShared by 4" total number of titles is over 150% more than the AShared by 5" number of titles in all four library groupings. If we look at each of the type of library groupings we can see in AShared by 4" that the different library grouping have different rankings by subject divisions. And that as we back down to the lower numbers of shared titles, the character of the collections by library grouping becomes more apparent until the unique titles by subject division which shows the difference in the collections most clearly.

The subject divisions which have the largest number of shared titles are naturally those divisions with the largest total number of titles. The universe of publication is at work in that the overall rate of publication by subject field provides a large number of titles from which to select. Thus subject fields with high rates of publication have more unique titles, the high rate providing variety. The fields with the largest number of shared titles, also are those same fields: LLL, history, philosophy/religion, engineering/technology. The exception to the usual pattern of concentration we have seen by subject areas is that business and economics which is a highly collected subject division does not rank near the top in the AShared by 5" category. Instead, art and architecture which does not rank near the top in total numbers or in the rankings of unique titles, does rank fifth in the AShared by 5 Acategory. The influence of the school library grouping can be seen here, in that business and economics is not heavily collected by school libraries, and, as has been observed earlier, a title has to be held by at least one school library to be in the AShared by 5" category.

Summary

Both WorldCat and the research libraries have close to the same absolute numbers in the datasets for unique records, both those with call numbers and without call numbers. The other four types of library groupings have low percentages of unique titles and higher overlap among the four. The academic library records by subject naturally have a high

overlap with the research libraries. The research libraries grouping contributes 62% of the unique records in the sample database, while the other four library groupings contribute the remaining 38 percent. The core set of records (101,378) shared by all of the library groupings represents only 3% of the total records in the sample. It represents 5% of the records used in the subject analysis of which the 101,378 records is a subset.

The data analyses on number of unique records and title overlap demonstrate that the universe of materials under bibliographic control in WorldCat show a high level of diversity of resources with 53% of records in the analysis having only one library location symbol. This aspect of WorldCat makes it a rich resource for cataloging records, verification of the existence of titles, and identifying prospective materials for resources sharing.

The analysis by language groupings is in Chapter Five. Click here for Chapter 5.

Table 4-1a
WorldCat Unique Records - All Titles Held by Date
(All Divisions)

SUBJECT DIVISION	Pre-1500	1500-1599	1600-1699	1700-1799	1800-1849	1850-1899	1900-1909	1910-1919	1920-1929
AGRICULTURE	0	14	23	115	196	1,037	732	758	895
ANTHROPOLOGY	0	12	15	28	77	241	175	118	144
ART AND ARCHITECTURE	2	62	72	172	361	1,037	861	548	767
BIOLOGICAL SCIENCES	0	12	17	128	304	1,011	592	337	428
BUSINESS AND ECONOMICS	1	28	76	337	827	2,606	2,026	1,583	2,109
CHEMISTRY	0	4	20	39	46	128	98	57	125
COMPUTER SCIENCE	0	0	0	1	0	4	6	1	1
EDUCATION	1	10	26	67	299	853	656	507	832
ENGINEERING AND TECHNOLOGY	4	44	91	212	467	1,939	1,401	952	1,249
GEOGRAPHY AND EARTH SCIENCES	1	16	43	130	274	1,053	641	375	448
HISTORY AND AUXILIARY SCIENCES	8	219	906	1,991	2,847	8,022	5,089	3,675	3,522
LANGUAGE, LINGUISTICS, AND LITERATURE	32	559	1,010	3,047	5,956	13,523	7,800	4,693	5,979
LAW	15	174	236	394	474	1,004	550	431	465
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	4	12	90	399	752	1,625	762	457	693
MATHEMATICS	1	29	87	149	168	468	275	142	165
MEDICINE	4	66	81	200	387	1,061	560	430	479
MUSIC	2	23	48	210	327	704	459	272	319
PERFORMING ARTS	0	13	28	38	45	107	142	80	132
PHILOSOPHY AND RELIGION	56	746	1,584	2,667	3,308	7,541	3,419	2,060	2,533
PHYSICAL EDUCATION AND RECREATION	0	3	5	23	60	200	160	124	168
PHYSICAL SCIENCES	6	63	100	193	210	753	422	310	338
POLITICAL SCIENCE	0	41	97	310	509	1,450	959	791	753
PSYCHOLOGY	1	13	48	33	59	210	219	133	188
SOCIOLOGY	0	5	23	111	288	865	647	433	541
Subject Analysis	**138**	**2,168**	**4,726**	**10,994**	**18,241**	**47,442**	**28,651**	**19,267**	**23,273**
No Call Number Present	867	5,531	9,945	23,167	29,012	71,754	32,320	28,423	32,444
Grand Total	**1,005**	**7,699**	**14,671**	**34,161**	**47,253**	**119,196**	**60,971**	**47,690**	**55,717**

Table 4-1b
WorldCat Unique Records - All Titles Held by Date
(All Divisions)

SUBJECT DIVISION	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1984	1985	1986	1987
AGRICULTURE	1,139	1,064	1,691	2,367	4,016	2,588	488	475	483
ANTHROPOLOGY	176	155	240	393	814	568	137	133	132
ART AND ARCHITECTURE	868	857	1,580	3,188	5,563	3,875	784	786	802
BIOLOGICAL SCIENCES	590	470	727	1,483	2,982	1,945	370	406	401
BUSINESS AND ECONOMICS	2,808	2,978	4,055	8,624	17,451	11,571	2,360	2,482	2,580
CHEMISTRY	177	166	384	736	856	507	93	89	115
COMPUTER SCIENCE	0	16	27	243	658	696	205	212	246
EDUCATION	1,645	1,360	2,697	6,444	13,477	8,140	1,736	1,756	1,616
ENGINEERING AND TECHNOLOGY	1,578	2,167	4,438	7,352	13,724	9,226	1,869	1,876	1,890
GEOGRAPHY AND EARTH SCIENCES	566	527	1,561	2,449	3,788	2,353	497	515	497
HISTORY AND AUXILIARY SCIENCES	5,007	5,100	5,198	9,390	15,639	10,075	2,110	2,000	2,105
LANGUAGE, LINGUISTICS, AND LITERATURE	6,182	6,914	9,524	17,968	25,253	14,154	2,795	2,812	2,768
LAW	722	738	1,315	2,016	4,852	3,944	967	958	958
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	1,335	852	1,245	3,415	6,591	3,866	832	619	902
MATHEMATICS	193	182	371	1,080	1,659	893	178	193	183
MEDICINE	632	612	1,037	1,825	4,109	2,871	547	647	665
MUSIC	403	394	606	918	1,724	1,043	220	195	198
PERFORMING ARTS	225	199	336	536	998	653	166	131	131
PHILOSOPHY AND RELIGION	2,874	2,810	4,047	6,368	9,074	5,979	1,308	1,264	1,184
PHYSICAL EDUCATION AND RECREATION	218	201	448	753	1,647	1,170	187	207	207
PHYSICAL SCIENCES	410	409	861	1,956	2,165	1,374	261	339	312
POLITICAL SCIENCE	1,048	1,364	1,946	3,283	5,270	3,027	619	657	667
PSYCHOLOGY	210	183	397	948	2,143	1,204	276	297	286
SOCIOLOGY	750	692	1,122	3,044	7,959	4,772	923	975	1,026
Subject Analysis	**29,756**	**30,410**	**45,853**	**86,779**	**152,412**	**96,494**	**19,928**	**20,024**	**20,354**
No Call Number Present	38,658	35,857	50,611	79,790	116,467	77,406	15,836	15,640	15,732
Grand Total	**68,414**	**66,267**	**96,464**	**166,569**	**268,879**	**173,900**	**35,764**	**35,664**	**36,086**

Table 4-1c
WorldCat Unique Records - All Titles Held by Date
(All Divisions)

SUBJECT DIVISION	1988	1989	1990	1991	1992	1993	1994	1995	1996
AGRICULTURE	507	485	514	511	542	510	503	384	390
ANTHROPOLOGY	125	128	171	146	173	165	174	163	172
ART AND ARCHITECTURE	796	875	865	918	904	849	832	841	854
BIOLOGICAL SCIENCES	419	395	471	470	477	489	458	363	299
BUSINESS AND ECONOMICS	2,642	2,798	2,852	2,915	3,042	2,953	2,883	2,653	2,331
CHEMISTRY	100	128	113	118	122	111	94	55	54
COMPUTER SCIENCE	215	275	309	344	312	307	258	233	199
EDUCATION	1,762	1,665	1,858	1,741	1,873	1,889	1,709	1,572	1,163
ENGINEERING AND TECHNOLOGY	1,901	1,811	2,046	1,995	2,089	1,812	1,692	1,407	1,293
GEOGRAPHY AND EARTH SCIENCES	484	534	548	607	629	566	511	476	496
HISTORY AND AUXILIARY SCIENCES	2,216	2,321	2,548	2,347	2,536	2,626	2,774	2,709	2,539
LANGUAGE, LINGUISTICS, AND LITERATURE	2,741	2,843	3,111	3,219	3,234	3,524	3,629	3,501	3,563
LAW	1,000	1,090	1,304	1,403	1,411	1,400	1,431	1,385	1,351
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	723	623	737	716	702	700	628	799	499
MATHEMATICS	194	205	210	266	273	240	223	176	161
MEDICINE	701	717	789	870	949	913	852	699	636
MUSIC	229	203	204	208	189	203	181	151	148
PERFORMING ARTS	138	141	151	132	179	177	150	147	166
PHILOSOPHY AND RELIGION	1,326	1,272	1,440	1,378	1,414	1,342	1,348	1,193	1,274
PHYSICAL EDUCATION AND RECREATION	241	217	241	270	274	287	237	257	294
PHYSICAL SCIENCES	317	297	274	323	338	327	261	263	185
POLITICAL SCIENCE	737	778	871	835	853	855	874	901	810
PSYCHOLOGY	275	271	260	284	308	300	304	250	215
SOCIOLOGY	1,035	1,126	1,175	1,244	1,290	1,345	1,266	1,227	1,120
Subject Analysis	**20,824**	**21,198**	**23,062**	**23,260**	**24,113**	**23,890**	**23,272**	**21,805**	**20,212**
No Call Number Present	16,274	16,245	16,243	15,449	15,468	14,869	14,335	12,519	10,478
Grand Total	**37,098**	**37,443**	**39,305**	**38,709**	**39,581**	**38,759**	**37,607**	**34,324**	**30,690**

Table 4-1d
WorldCat Unique Records - All Titles Held by Date
(All Divisions)

SUBJECT DIVISION	1997	1998	1999	2000	2001	Other	Total
AGRICULTURE	419	331	246	158	11	290	23,882
ANTHROPOLOGY	199	154	137	68	6	172	5,711
ART AND ARCHITECTURE	794	681	631	323	12	832	33,192
BIOLOGICAL SCIENCES	282	276	226	165	22	471	17,486
BUSINESS AND ECONOMICS	2,215	1,819	1,402	663	62	1,775	97,507
CHEMISTRY	39	44	35	32	4	63	4,752
COMPUTER SCIENCE	221	188	225	130	17	715	6,264
EDUCATION	993	943	670	183	41	2,097	62,281
ENGINEERING AND TECHNOLOGY	1,263	1,113	910	502	46	983	71,342
GEOGRAPHY AND EARTH SCIENCES	487	414	371	286	18	673	22,834
HISTORY AND AUXILIARY SCIENCES	2,342	2,110	1,802	782	81	1,758	114,394
LANGUAGE, LINGUISTICS, AND LITERATURE	3,486	3,420	3,079	1,998	151	2,885	175,353
LAW	1,248	1,153	878	267	16	350	35,900
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	1,150	335	273	121	5	2,596	35,058
MATHEMATICS	149	137	108	74	19	224	9,075
MEDICINE	638	520	499	279	47	852	26,174
MUSIC	120	123	124	76	13	389	10,626
PERFORMING ARTS	132	112	131	95	9	159	5,979
PHILOSOPHY AND RELIGION	1,203	1,010	835	470	49	1,387	75,763
PHYSICAL EDUCATION AND RECREATION	249	251	262	164	20	243	9,288
PHYSICAL SCIENCES	191	149	162	109	16	267	13,961
POLITICAL SCIENCE	811	681	541	342	12	878	33,570
PSYCHOLOGY	193	189	215	140	14	140	10,206
SOCIOLOGY	1,128	907	634	397	47	996	39,113
Subject Analysis	**19,952**	**17,060**	**14,396**	**7,824**	**738**	**21,195**	**939,711**
No Call Number Present	10,861	9,600	8,133	1,903	10	13,355	855,202
Grand Total	**30,813**	**26,660**	**22,529**	**9,727**	**748**	**34,550**	**1,794,913**

Table 4-2a
Unique Records Research Libraries - All Titles Held by Date (All Divisions)

SUBJECT DIVISION	Pre-1500	1500-1599	1600-1699	1700-1799	1800-1849	1850-1899	1900-1909	1910-1919	1920-1929
AGRICULTURE	0	11	14	94	135	768	546	626	762
ANTHROPOLOGY	0	10	11	17	56	168	82	87	101
ART AND ARCHITECTURE	2	22	53	127	259	705	468	390	495
BIOLOGICAL SCIENCES	0	4	9	70	174	660	315	246	282
BUSINESS AND ECONOMICS	0	14	43	208	498	1,694	984	1,182	1,594
CHEMISTRY	0	4	13	31	37	93	56	50	106
COMPUTER SCIENCE	0	0	0	0	0	4	5	1	1
EDUCATION	0	9	15	40	183	502	289	309	493
ENGINEERING AND TECHNOLOGY	1	37	77	137	334	1,138	636	526	673
GEOGRAPHY AND EARTH SCIENCES	1	10	25	81	155	624	276	238	269
HISTORY AND AUXILIARY SCIENCES	6	174	689	1,572	2,008	5,416	2,317	2,624	2,495
LANGUAGE, LINGUISTICS, AND LITERATURE	16	454	748	2,485	4,224	9,483	4,221	3,426	4,364
LAW	15	155	177	313	340	646	295	248	300
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	4	9	61	300	507	1,237	451	359	506
MATHEMATICS	1	26	74	129	110	354	154	111	123
MEDICINE	4	55	59	152	310	833	291	296	351
MUSIC	0	15	27	158	220	376	185	118	144
PERFORMING ARTS	0	12	27	31	39	64	63	59	95
PHILOSOPHY AND RELIGION	37	350	733	1,461	1,576	3,011	1,026	830	977
PHYSICAL EDUCATION AND RECREATION	0	2	4	17	36	98	62	71	68
PHYSICAL SCIENCES	4	51	60	138	145	445	186	195	239
POLITICAL SCIENCE	0	32	78	208	336	918	471	484	507
PSYCHOLOGY	1	9	34	26	44	130	99	78	112
SOCIOLOGY	0	4	18	87	189	514	278	272	357
Subject Analysis	**92**	**1,469**	**3,049**	**7,882**	**11,915**	**29,881**	**13,756**	**12,826**	**15,414**
No Call Number Present	652	3,740	7,193	17,724	21,281	54,376	20,964	21,251	23,646
Grand Total	**744**	**5,209**	**10,242**	**25,606**	**33,196**	**84,257**	**34,720**	**34,077**	**39,060**

Table 4-2b
Unique Records Research Libraries - All Titles Held by Date (All Divisions)

SUBJECT DIVISION	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1984	1985	1986	1987
AGRICULTURE	909	826	1,267	1,619	2,587	1,590	314	328	291
ANTHROPOLOGY	124	115	174	302	586	405	103	92	93
ART AND ARCHITECTURE	563	551	1,095	2,067	3,654	2,576	550	537	520
BIOLOGICAL SCIENCES	336	298	451	786	1,495	1,013	214	241	240
BUSINESS AND ECONOMICS	2,121	2,124	2,794	6,001	11,469	7,674	1,638	1,740	1,808
CHEMISTRY	136	129	252	415	450	298	59	56	76
COMPUTER SCIENCE	0	13	19	158	353	350	88	128	150
EDUCATION	1,007	710	1,189	2,752	5,950	3,888	832	856	817
ENGINEERING AND TECHNOLOGY	825	1,204	2,723	3,905	6,815	4,821	1,096	1,095	975
GEOGRAPHY AND EARTH SCIENCES	345	300	991	1,243	1,820	1,196	275	268	251
HISTORY AND AUXILIARY SCIENCES	3,492	3,397	3,270	6,184	10,530	6,984	1,488	1,426	1,493
LANGUAGE, LINGUISTICS, AND LITERATURE	4,334	5,055	7,572	14,233	20,001	11,292	2,273	2,251	2,215
LAW	414	451	688	1,341	2,998	2,459	640	637	607
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	1,031	583	839	2,339	3,872	2,501	586	393	387
MATHEMATICS	129	145	283	710	1,043	566	115	125	120
MEDICINE	471	426	673	996	2,149	1,433	292	347	347
MUSIC	184	167	291	482	863	607	153	113	124
PERFORMING ARTS	153	126	234	379	676	463	114	100	89
PHILOSOPHY AND RELIGION	1,152	1,023	1,440	2,903	4,297	3,009	673	639	648
PHYSICAL EDUCATION AND RECREATION	114	111	231	343	767	567	91	98	87
PHYSICAL SCIENCES	216	214	554	1,109	1,254	838	179	228	214
POLITICAL SCIENCE	744	794	985	1,885	2,953	1,862	372	430	435
PSYCHOLOGY	110	81	241	545	982	625	149	180	167
SOCIOLOGY	509	452	674	1,880	4,686	2,908	584	660	664
Subject Analysis	**19,419**	**19,295**	**28,930**	**54,577**	**92,250**	**59,925**	**12,878**	**12,968**	**12,818**
No Call Number Present	26,561	25,636	33,634	47,359	64,690	43,952	9,213	9,007	9,048
Grand Total	**45,980**	**44,931**	**62,564**	**101,936**	**156,940**	**103,877**	**22,091**	**21,975**	**21,866**

Table 4-2c
Unique Records Research Libraries - All Titles Held by Date (All Divisions)

SUBJECT DIVISION	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
AGRICULTURE	297	270	316	320	341	341	312	228	252	285
ANTHROPOLOGY	83	96	113	109	124	123	121	117	118	112
ART AND ARCHITECTURE	548	593	579	653	619	594	575	571	578	513
BIOLOGICAL SCIENCES	253	241	292	294	287	262	258	179	173	134
BUSINESS AND ECONOMICS	1,886	2,009	2,013	2,012	2,152	2,118	2,003	1,814	1,586	1,513
CHEMISTRY	67	85	71	86	80	63	57	21	34	18
COMPUTER SCIENCE	107	135	162	184	160	139	118	99	79	85
EDUCATION	884	793	847	856	894	932	838	730	475	368
ENGINEERING AND TECHNOLOGY	1,017	990	1,125	1,089	1,089	989	921	690	635	685
GEOGRAPHY AND EARTH SCIENCES	261	313	307	373	375	328	297	271	287	291
HISTORY AND AUXILIARY SCIENCES	1,554	1,660	1,703	1,685	1,826	1,930	2,068	1,909	1,813	1,597
LANGUAGE, LINGUISTICS, AND LITERATURE	2,135	2,185	2,401	2,532	2,526	2,761	2,836	2,695	2,536	2,416
LAW	597	670	804	942	925	979	998	931	910	833
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	460	388	439	441	396	429	344	557	311	823
MATHEMATICS	109	123	130	178	153	157	126	92	87	85
MEDICINE	402	394	412	451	521	440	445	340	286	269
MUSIC	144	140	123	147	131	143	112	101	88	80
PERFORMING ARTS	98	105	99	93	123	137	113	100	116	90
PHILOSOPHY AND RELIGION	692	673	757	768	856	795	787	704	778	715
PHYSICAL EDUCATION AND RECREATION	111	92	97	132	124	127	103	126	120	95
PHYSICAL SCIENCES	217	211	206	222	226	205	167	152	120	117
POLITICAL SCIENCE	501	519	550	568	542	526	542	544	484	505
PSYCHOLOGY	162	156	128	167	185	167	158	129	113	81
SOCIOLOGY	684	791	756	851	899	915	880	795	750	765
Subject Analysis	**13,269**	**13,632**	**14,430**	**15,153**	**15,554**	**15,600**	**15,179**	**13,895**	**12,729**	**12,475**
No Call Number Present	9,451	9,656	9,483	8,828	8,332	8,020	7,314	5,735	3,912	3,791
Grand Total	**22,720**	**23,288**	**23,913**	**23,981**	**23,886**	**23,620**	**22,493**	**19,630**	**16,641**	**16,266**

Table 4-2d
Unique Records Research Libraries - All Titles Held by Date (All Divisions)

SUBJECT DIVISION	1998	1999	2000	2001	Other	Total
AGRICULTURE	202	139	126	11	72	16,199
ANTHROPOLOGY	87	92	54	6	34	3,915
ART AND ARCHITECTURE	389	304	220	11	271	21,652
BIOLOGICAL SCIENCES	133	112	125	22	62	9,661
BUSINESS AND ECONOMICS	1,130	718	464	60	310	65,374
CHEMISTRY	20	22	24	4	10	2,923
COMPUTER SCIENCE	69	99	93	17	3	2,819
EDUCATION	289	170	105	41	142	28,205
ENGINEERING AND TECHNOLOGY	532	435	344	46	191	37,796
GEOGRAPHY AND EARTH SCIENCES	225	192	227	17	61	12,193
HISTORY AND AUXILIARY SCIENCES	1,359	1,055	585	81	532	76,922
LANGUAGE, LINGUISTICS, AND LITERATURE	2,525	2,230	1,623	149	1,303	131,500
LAW	703	415	137	16	59	22,643
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	175	123	68	5	128	21,052
MATHEMATICS	69	55	58	18	15	5,773
MEDICINE	223	223	229	47	78	14,245
MUSIC	86	69	61	13	69	5,734
PERFORMING ARTS	78	85	75	9	23	4,068
PHILOSOPHY AND RELIGION	591	417	337	49	351	35,055
PHYSICAL EDUCATION AND RECREATION	118	137	140	20	21	4,330
PHYSICAL SCIENCES	65	94	77	16	51	8,415
POLITICAL SCIENCE	393	320	246	12	109	19,855
PSYCHOLOGY	88	111	117	14	43	5,432
SOCIOLOGY	555	340	278	47	121	24,163
Subject Analysis	**10,104**	**7,957**	**5,813**	**731**	**4,059**	**579,924**
No Call Number Present	3,214	2,494	485	2	7,694	528,338
Grand Total	**13,318**	**10,451**	**6,298**	**733**	**11,753**	**1,108,262**

Table 4-3
Total Number of Unique Bibliographic Records
WorldCat and Research Libraries

	Total Records	Total Unique	Percent Unique of Total Records	Unique Call Number Subject Analysis	Percent Call Number Present of Unique Total	Percent Unique Call Number Subject Analysis Records of Total Records	Unique No Call Number Present	Percent Unique No Call Number of Total Unique Records	Percent Unique - No Call Number of Total Records
WorldCat	3,378,272	1,794,913	53.13%	939,711	52.35%	27.82%	855,202	47.65%	25.31%
Research	1,745,034	1,108,262	63.51%	579,924	52.33%	33.23%	528,338	47.67%	30.28%

Table 4-4
Total Number of Unique Records by Decade
WorldCat and Research Libraries

	WorldCat Total Unique Records	Research Libraries Unique Records	Percent Research of WorldCat
1900-1909	60,971	34,720	56.95%
1910-1919	47,690	34,077	71.46%
1920-1929	55,717	39,060	70.10%
1930-1939	68,414	45,980	67.21%
1940-1949	66,267	44,931	67.80%
1950-1959	96,464	62,564	64.86%
1960-1969	166,569	101,936	61.20%
1970-1979	268,879	156,940	58.37%
1980-1989	355,955	215,817	60.63%
1990-1999	332,217	194,199	58.46%

Table 4-5
Unique Records by Subject Division
WorldCat and Research Libraries

		Unique Totals				Percent WorldCat held by		Totals
SUBJECT DIVISION		World Cat		Research		Research Libraries		Academic
AGRICULTURE		23,882		16,199	[4]	67.83%		3,809
ANTHROPOLOGY		5,711		3,915	[2]	68.55%		808
ART AND ARCHITECTURE		33,192		21,653		65.24%		5,831
BIOLOGICAL SCIENCES		17,486		9,662		55.26%		3,948
BUSINESS AND ECONOMICS	[3]	97,507	[3]	65,375	[6]	67.05%		14,933
CHEMISTRY		4,752		2,923		61.51%		1,372
COMPUTER SCIENCE		6,264		2,816		44.96%		1,924
EDUCATION		62,281		28,205		45.29%	[2]	28,160
ENGINEERING AND TECHNOLOGY	[5]	71,342	[4]	37,797		52.98%	[5]	12,890
GEOGRAPHY AND EARTH SCIENCES		22,834		12,193		53.40%		3,688
HISTORY AND AUXILIARY SCIENCES	[2]	114,394	[2]	76,936	[5]	67.26%	[4]	15,516
LANGUAGE, LINGUISTICS, AND LITERATURE	[1]	175,353	[1]	131,514	[1]	75.00%	[3]	27,625
LAW		35,900		22,642		63.07%		6,386
LIBRARY SCIENCE, GENERALITIES AND REFERENCE		35,058		21,052		60.05%		8,238
MATHEMATICS		9,075		5,773		63.61%		2,444
MEDICINE		26,174		14,245		54.42%		7,477
MUSIC		10,626		5,735		53.97%		2,932
PERFORMING ARTS		5,979		4,069	[3]	68.05%		1,224
PHILOSOPHY AND RELIGION	[4]	75,763	[5]	35,056		46.27%	[1]	32,838
PHYSICAL EDUCATION AND RECREATION		9,288		4,330		46.62%		2,250
PHYSICAL SCIENCES		13,961		8,415		60.28%		2,675
POLITICAL SCIENCE		33,570		19,855		59.15%		5,844
PSYCHOLOGY		10,206		5,432		53.22%		3,580
SOCIOLOGY		39,113		24,163		61.78%		8,162
Subject Analysis		**939,711**		**579,955**		61.72%		**204,554**

*** Highlighted bracketed numbers indicate ranking.**

Table 4-6
Unique Bibliographic Records From Subject Analysis by Library Grouping

	Total Records	Total Unique Records	Percent Unique of Total
Academic	1,323,165	204,554	15.46%
Public	727,643	76,054	10.45%
School	130,309	995	0.76%
Special	736,165	78,356	10.64%
Research	1,745,034	579,955	33.23%
WorldCat	2,199,165	939,711	42.73%

Table 4-7
Unique Records By Subject
By Library Groupings

	Libraries									
DIVISION		**Academic**		**Special**		**Public**		**School**		**Research**
AGRICULTURE		3,809		2,268		1,592		17		16,199
ANTHROPOLOGY		808		315		637		38		3,915
ART AND ARCHITECTURE		5,831		2,658		3,019		40		21,653
BIOLOGICAL SCIENCES		3,948		2,897		920		64		9,662
BUSINESS AND ECONOMICS	**[5]**	14,933	**[2]**	9,770	**[3]**	7,412		36	**[3]**	65,375
CHEMISTRY		1,372		347		107		3		2,923
COMPUTER SCIENCE		1,924		1,219		292		7		2,816
EDUCATION	**[2]**	28,160		3,039		2,726	**[1]**	157	**[6]**	28,205
ENGINEERING AND TECHNOLOGY	**[6]**	12,890	**[1]**	14,144	**[4]**	6,475		44	**[4]**	37,797
GEOGRAPHY AND EARTH SCIENCES		3,688	**[5]**	4,145		2,772		44		12,193
HISTORY AND AUXILIARY SCIENCES	**[4]**	15,516	**[3]**	7,057	**[1]**	14,860		87	**[2]**	76,936
LANGUAGE, LINGUISTICS, AND LITERATURE	**[3]**	27,625		3,808	**[2]**	12,274	**[2]**	153	**[1]**	131,514
LAW		6,386	**[4]**	4,591		2,217		70		22,642
LIBRARY SCIENCE, GENERALITIES AND REFERENCE		8,238		3,695		2,053		26		21,052
MATHEMATICS		2,444		399		445		14		5,773
MEDICINE		7,477		2,380		2,051		26		14,245
MUSIC		2,932		771		1,179		10		5,735
PERFORMING ARTS		1,224		276		411		4		4,069
PHILOSOPHY AND RELIGION	**[1]**	32,838	**[5]**	4,193		3,676		21	**[5]**	35,056
PHYSICAL EDUCATION AND RECREATION		2,250		1,091		1,586		32		4,330
PHYSICAL SCIENCES		2,675		2,177		665		30		8,415
POLITICAL SCIENCE		5,844		4,000	**[5]**	3,860		26		19,855
PSYCHOLOGY		3,580		204		979		12		5,432
SOCIOLOGY		8,162		2,912	**[6]**	3,846		34		24,163
TOTAL		204,554		78,356		76,054		995		579,955

*** Highlighted bracketed numbers indicate ranking.**

Table 4-8
All Libraries - Title Overlap Between Groups

	Academic	Public	Research	School	Special
Academic	1,232,222	575,263	1,036,807	127,487	569,168
	100.00%	43.50%	78.40%	9.60%	43.00%
Public	575,263	733,539	591,913	122,189	403,576
	78.40%	100.00%	80.70%	16.70%	55.00%
Research	1,036,807	591,913	1,745,122	121,423	599,099
	59.40%	33.90%	100.00%	7.00%	34.30%
School	127,487	122,189	121,423	132,379	111,400
	96.30%	92.30%	91.70%	100.00%	84.20%
Special	569,168	403,576	599,099	111,400	736,191
	77.30%	54.80%	81.40%	15.10%	100.00%

Table 4-9
Total Number of Shared Records by Library Grouping

Libraries	Unique	Shared by 2	Shared by 3	Shared by 4	Shared by 5
Research	579,955	425,731	342,175	268,883	101,378
Academic	204,554	403,602	341,453	272,235	101,378
Special	78,356	97,086	197,468	261,903	101,378
Public	76,054	93,887	193,118	269,102	101,378
School	995	2,970	7,091	19,945	101,378
Total Number of Records Unique or Shared	939,914	511,638	360,435	273,017	101,378

Table 4-10
Overlap Analysis by Library Grouping

RESEARCH	Unique	Shared by 2	Shared by 3	Shared by 4		Shared by 5
AGRICULTURE	16,199	9,014	6,370	6,299		1,478
ANTHROPOLOGY	3,915	3,089	2,355	1,932		1,018
ART AND ARCHITECTURE	21,653	18,104	17,442	16,153	[5]	4,988
BIOLOGICAL SCIENCES	9,662	7,653	8,005	6,594		3,245
BUSINESS AND ECONOMICS	65,375	41,996	30,865	27,542		4,513
CHEMISTRY	2,923	1,486	1,459	1,680		307
COMPUTER SCIENCE	2,816	1,780	1,693	3,189		770
EDUCATION	28,205	13,677	8,886	5,392		3,069
ENGINEERING AND TECHNOLOGY	37,797	21,433	19,226	21,795	[4]	5,354
GEOGRAPHY AND EARTH SCIENCES	12,193	10,074	9,151	8,752		2,406
HISTORY AND AUXILIARY SCIENCES	76,936	62,637	46,647	38,783	[2]	14,139
LANGUAGE, LINGUISTICS, AND LITERATURE	131,514	122,230	81,302	42,651	[1]	27,197
LAW	22,642	16,046	13,466	11,508		3,661
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	21,052	14,064	9,361	7,220		1,847
MATHEMATICS	5,773	4,064	3,074	2,782		926
MEDICINE	14,245	10,404	11,205	12,552		4,528
MUSIC	5,735	6,836	4,657	2,640		1,258
PERFORMING ARTS	4,069	3,464	2,698	1,864		1,298
PHILOSOPHY AND RELIGION	35,056	40,205	30,489	18,837	[3]	5,691
PHYSICAL EDUCATION AND RECREATION	4,330	3,994	2,939	2,726		2,614
PHYSICAL SCIENCES	8,415	5,475	5,750	4,823		1,563
POLITICAL SCIENCE	19,855	14,337	11,039	10,124		3,448
PSYCHOLOGY	5,432	3,946	2,792	2,668		1,580
SOCIOLOGY	24,163	16,723	11,304	10,377		4,480
TOTAL	**579,955**	**452,731**	**342,175**	**268,883**		**101,378**

*** Highlighted bracketed numbers indicate ranking.**

CHAPTER FIVE
LANGUAGE ANALYSIS

The monographic bibliographic records for WorldCat, the research libraries, and the academic libraries were analyzed by six foreign language groupings and English. The other three types of library groupings were not analyzed by language. Although a call number is not necessary to analyze records by language since the language tag can be used to extract bibliographic records by language grouping, the analysis was performed only on those records with call numbers in order to analyze language by subject by imprint year and type of library. The tables in this chapter were produced from the iCAS tables *All Libraries Language Analysis: Language Titles Held by Date (All Divisions, Select Language*); *Research Libraries Language Analysis: Language Titles Held by Date...*; and *Academic Libraries Language Analysis: Language Titles Held by Date....*

The total number of English language records in the sample including those records with no call number is over two million. Of all records (3,378,272), call number present and no call number present, approximately 65% are English language monographs. Given that the library membership of WorldCat is primarily composed of institutions from North American, Western European and other English speaking countries internationally, slightly over one-third of records for foreign language monographs are a considerable proportion of foreign language titles.

English, Non-English

The first three tables (5-1,5-2,5-3) for this chapter display the number of records by time period for WorldCat and the two types of library groupings analyzed by language. All three tables*, English and All Non-English as a Percent of Total Subject Records*, are divided into a total for English language monographs and a total for non-English language monographs by the date ranges defined for the study. The totals for non-English are derived by subtracting the numbers of English language records from the total number of

subject records (with call number). In other words the non-English numbers were not produced by the iCAS analysis but were calculated from the iCAS data. Throughout this analysis it must be borne in mind that the six foreign language groupings do not include *all* foreign languages in the sample. Non-English is *all* foreign language records because this category includes all records that are not tagged as English language.

Table 5-1, *WorldCat English and All Non-English as a Percent of Total Subject Records,* shows the percentages of English and non-English for those records with call numbers in the WorldCat sample by time period. The growth rate and number of publications according to the two language categories can also be seen. Table 5-1 was constructed to show the proportions that each of these two major language groupings comprise of the total number of records in the subject analysis dataset. It is interesting, but not an unexpected finding, that the early centuries have a very high proportion of foreign language records, with English being correspondingly low. The two centuries, 17th and 18th, have similar proportions at approximately 50% of total for both English and non-English. In the 19th century, the proportion of foreign language records decreases considerably from 50% to one/third of total. There are slight increases at the beginning of the 20th century, with the number hovering in the low 40% range until the 1950s. After the decade of the 60s, the percentage of total for foreign languages stays in the mid 30% range until 1997 when it reaches 38%, the same proportion as the 1960s. While there is a decline from the high in 1997, it is not drastic until the imprint year 1999 in which it is lower than any other time period at 31%. As the data were extracted in February 2001, shortly after the end of the 2000 publication year, it could be expected that the number of records would decline at that point.

For English language titles in WorldCat, beginning in the 1950s, the proportion of total reaches nearly 63% and stays in the 60% range from that point on until 2000, in which the number of foreign language records drop off drastically. Thus, it can be seen from Table 5-1, that the proportion of English language titles to foreign language titles in WorldCat is approximately two to one for most of the 20th century. It stands to reason that

there would be more foreign language titles in the earlier centuries as printing began in central Europe and production was in non-English speaking countries, for the most part, prior to the 18th century, in which the rate of publication in Great Britain soared. It was not until the 19th century that the United States began contributing significantly to the universe of publication worldwide. The records in WorldCat, as stated before, are a reflection of the library membership of OCLC, which is dominated by English speaking countries or countries in which a substantial proportion of publication is in English. The total number of records in the subject analysis dataset is relatively small prior to 1850. The analysis of total records in WorldCat in Chapter Two showed the same pattern. It as if a gun went off in 1900 and the race began, with a publication explosion that saw twice the number of publications in the first decade of the 20th century as in the last half of the 19th century. The high in 1992 is due to English language records as the totals for foreign language records peak in 1994. In spite of the absolute numbers for foreign language titles peaking in 1994, the *percentage* of total for foreign language records peaks in 1997, due to the decline in the number of English language titles from 1992 forward. This is interesting, as it would seem that the number of foreign language titles published in the1990s and cataloged into WorldCat would decline at a higher rate than the English language titles. In the 1990s, OCLC secured agreements with a number of foreign national libraries, which brought a significant number of current foreign language titles into the database. The contributions from these national libraries could be responsible for the strong foreign language showing in the 1990s.

Table 5-2, *Research Libraries English and All Non-English as a Percent of Total Subject Records,* presents the same analysis for the grouping of research libraries as Table 5-1 for WorldCat. The percentage of non-English records to English language records is higher for most of the time periods in the research libraries than in WorldCat. The percentage of non-English records hovers in the middle 30% range in WorldCat in the 1990s, still at 31% in 1999. In the research libraries the percentage of total for non-English records ranges from the mid 30% to an upper share of slightly less than 45% from 1800-

1998. The percentage share drops to below 30% at 28% in 1999, the first time period below 30% for non-English records in the time frame of the study for research libraries. It is interesting that while the absolute number of records declines beginning in 1992, in both WorldCat and the research libraries, the percentage shares of total for English and non-English remain relatively the same until 1999. The continued addition of retrospective materials by research libraries probably causes an Aevening out@ of the number of titles over time so that again looking at the number of titles from the vantage point of 2001, we are not seeing the actual numbers which were the contemporaneous totals by time period.

The same analysis was conducted for the grouping of academic libraries shown in Table 5-3 *Academic Libraries: English and Non-English as a Percentage of Total Subject Records*. As would be expected, the percentages of total for non-English are lower in the academic libraries than in the research libraries. There is approximately a 10 percentage point difference in the percentage share of the non-English records in the academic libraries from the research libraries. The range for percentage of share in non-English records in academic libraries is 23% for the low and 35.66% for the high until 1999. From 1900-1969, the percentage share hovers in the 30-35% range. Beginning in the decade of the 1970s, the percentage declines to 26% and hovers in the mid 20% range. In 1997, the percentage share is 26.31%, almost identical to the 1970s share of 26.23%. From 1990-1992, the total number of records is very close to the same, near 30,000. As we have seen in WorldCat and the research libraries, the peak in number of records is in 1992 and the number declines from 30,751 to 18,247 in 1999. Yet the decline in percentages of share in non-English titles is only 2 percentage points from 1997 to 1999. The rate of addition of foreign language records seems to remain fairly consistent. The collections of the academic libraries do not show the same amount of fluctuation as the research libraries.

One interesting comparison for all three analyses in Tables 5-1, 5-2, and 5-3 is in the number of AOther@ records, those without publication dates. In WorldCat the non-English records comprise only 18% of AOther@ records. In the research libraries grouping, non-English is 41% of AOther@ records. In the academic libraries it is 29% of AOther@

records. The higher percentage of non-English records without publication dates can be another indication of the variety in materials that the aggregated resources base of the research libraries add to the overall commons of the WorldCat database.

To make comparison easier between the academic and research library groupings by English and non-English language proportions and rate of increase/decrease, Table 5-4 *Increase/Decrease by Decades for English/Non-English Records, Academic and Research Libraries*, was constructed to analyze the time period from 1930-1999. These analyses depict the number of records for those decades from the vantage point of the end of the 20th century hence; the analysis applies to the present state of the WorldCat database. The profiles for those decades may have looked quite different at earlier points in time.

Looking at the number of records by decade for the two library groups, academic and research, in Table 5-4, the patterns are quite interesting between the two major language groupings. The decade of the 1940s with WWII shows only a slight increase in overall total records from the previous decade for both academic and research libraries. The explosion in publication in the 1950s can be seen in the nearly 60% increase in total records over the 1940s. While the increase in the number of English language records is nearly 61% in the academic libraries, the non-English is only slightly less at almost 58%. The research libraries have the same rate of increase in English language records at 60%. But the growth in non-English is lower than in the academic libraries at 44% for the research libraries. The total increase in the academic libraries is almost 60%, whereas the research collections only have a 53% increase. The largest increases by decade occur in the 1960s. The academic library collections double with a 100% increase. The research library collections double in non-English, with an 89% increase in English language records for an overall 95% increase. The decade of the 1970s shows increases again, but lower than the sixties. In the academic libraries English language records increase by 72%, but non-English only increase by 32%. Overall growth is at 59%, the first divergence from the neck and neck pattern between the two language groupings. In the research libraries

growth in English language records is similar to the academic libraries, but slightly less at 69 percent. The increase in non-English records in the research libraries is much higher at nearly 48 percent. The overall growth rate is nearly identical in both the academic libraries and the research libraries at nearly 60% in the 1970s.

Compared to the 1970s, the 1980s see a marked slowdown in the increase in the number of records over the previous decade. The increase is also much lower in the academic libraries than in the research libraries. Both English and non-English grow at a similar rate in the academic libraries, with the overall increase at 12 percent. In the research libraries, the non-English records have a larger increase than the English language records with the overall increase at 20.5 percent. While the absolute numbers for the 1990 totals are higher than the 1980s, the increases in the academic libraries over the previous decade are very small. There is really no increase at all in the non-English records and a modest 4% increase in English language records. The overall increase for the academic libraries is 3 percent. In the research libraries the overall increase is 5 percent. As in the academic libraries, the increase in non-English records is very small at less than 2 percent. The English language records increase 8 percent.

While the decade of the 1990s has slight increases in number of records over the previous decade, Table 5-5 *Increase/Decrease by Five Year Periods, 1980-1999 for English/Non-English Language Groupings, Academic and Research Libraries* was constructed by five year intervals beginning in 1980, to more closely examine that time period. When viewed this way, it can be seen that the latter half of the 1980s did experience a drastic decrease in the number of foreign language titles over the first half of the decade.

In the two halves of the 1980s, the number of English language titles increase in both the academic and research library groupings. But the declines in the number of foreign language titles are so steep that the overall title count is negative in both groups for the latter half of the decade. The research libraries have a larger gain in English language titles, but the decline in foreign language titles is twice the decline in the academic libraries.

In the first half of the 1990 decade the number of records increased over the last half of the 1980s, but this increase was based upon the decreased totals of the latter half of the 1980s. It is also interesting to note that both library groups have a similar overall percentage increase at 12-13 percent. In the latter half of the 1990s both library groups have substantial decreases in the number of records over the first half of the decade. The decreases are larger in the research libraries with an overall decline of 19 percent. Non-English declines more than 3 percentage points over English language records. In the academic libraries the English records have a greater decline than non-English.

In both library groupings the number of records in the first half of the 1990s is greater than the latter half of the 1980s. In the academic libraries the absolute number grand total is larger than the grand total for either half of the 1980s. But in the research libraries the grand total for 1990-1994 is less than the grand total for 1980-1984 as well as the second half of the 1980s. It is easy to see that the grand total for the research libraries in the first half of the 1990s is greatly affected by the drastic decline in the number of non-English language records.

In previous studies by this researcher using the OCLC/AMIGOS CACD product, the drastic declines in the number of records in the latter half of the 1980s were found.[1] But the period preceding the 1980s was a dark mystery. This WorldCat study has enabled us to see the differences in the number of records for earlier time periods. It is an extremely interesting finding that the number of records for the first half of the 1980s, even from the vantage point of 2001, still exceeds the number for non-English language records for each of the subsequent five year intervals up through 1999. And the drastic decline in the number of foreign language records affects the overall total for the research libraries so that it is still below the total for the first half of the 1980s even in the first half of the 1990s. The increase in English language purchasing brought the total number of English language records up higher in the first half of the 1990s, but the increase in the number of foreign language titles in the first half of the 1990s was virtually the same as for the English language titles. But the research libraries have over 5,000 fewer non-English language

records in the latter 1990s as in the latter 1980s.

In the academic libraries the totals for non-English are very close in the latter 1980s and in the latter 1990s. The academic libraries grand total increases in the first half of the 1990s over the latter 1980s such that the grand total for the first half off the 1990s exceeds the grand total for the first half of the 1980s. The English language titles likewise are higher for the first half of the 1990s. The total for non-English in the academic libraries is lower in the first half of the 1990s than the first half of the 1980s. But the decline in non-English is not as drastic proportionally as in the research libraries, giving the academic libraries a real net gain for the grand total, whereas in the research libraries the increase in the 1990s over the latter half of the 1980s does not result in a net gain over the early 1980s.

The pattern of decrease in the last interval in the study is clearly evident in the last half of the 1990s. We can see by Tables 5-4 and 5-5 that although the 1990s show an increase in the number of records over the 1980s, the increase is due solely to the first five years of the 1990s, with a drastic decline in the number of records in the second half of the decade. And the decline in non-English language records is not as drastic in the academic libraries as in the research libraries where the decline is 22 percent. It cannot be ascertained from these data the reasons that the first half of the 1990s have a larger number of titles than the latter half of the 1980s. It is possible that the difference is due to increased publication internationally in the 1990s. It is also possible that the budget crisis in the latter 1980s due to the rapid escalation in serial subscription prices is one of the reasons the totals are lower for the latter 1980s. Although the continued addition of cataloging records results in an evening out over time, it is possible that the budget crisis in the latter 1980s is still visible in a reduced number of titles for that time period.

In the next section collecting patterns for the six foreign language groupings defined for the study are analyzed.

Foreign Language Groupings

The 2,199,165 records with call numbers were analyzed by language groupings of English, Chinese, French, German, Japanese, Russian, and Spanish. Of these 1,399,246 were for English language materials and 536,361 records were for the six foreign language groupings. (Remember, the six languages do not make up all foreign language records which are shown as non-English in the previous tables.) Of the total number of records with call numbers, 63.6% are for English language materials and 36.3% are for non-English titles.

The six language groupings as percentage share of all non-English records in the study and as a percentage share of the total number of records in the subject analysis of WorldCat (with call number) are shown in Table 5-6, *Foreign Language Groupings as Percentage of Total Records*. In the WorldCat sample, French has the highest total with German a close second. Spanish is third. Chinese and Russian are together with 52,605 and 50,373 records each. Japanese has the lowest total records at 37,302. These are the top six foreign languages in number of records in WorldCat.

The percentage share of these language groupings of total foreign language records is shown. The six languages account for 67.5% of foreign language records, meaning all other languages comprise 32.5% of foreign language records. The percentage share each language comprises of all of the records with call numbers in the sample is much lower. French and German make of over 6% of total subject records, with Spanish right at 5% and the other three having 2% of total. We have seen in the first section of this chapter that the ratio of the total number of records in the study of English language records to non-English language records is 65/35. And in turn, the percentage share the six languages in the study comprise of the non-English records is 67.5 percent. Of the total records with call number, the six languages comprise 24.8% of that dataset.

The foreign language records are not just contributed by North American libraries, but are also contributed by the international membership of OCLC. The top position held by French language materials in the 2001 sample may be because many of the largest

universities in France are OCLC members. The Bibliothèque Nationale du Québec became a member in 1999. In 2000, the Université du Québec, a network of 10 libraries in Québec, Canada joined OCLC. The international membership of OCLC has been growing in recent years. The National Library of China, the largest library in Asia and the fifth largest in the world, became an OCLC member in 1998. A consortium of Catalan libraries in Barcelona Spain became an OCLC member in February 2001. Also in the past several years, universities in Brazil, Mexico, and Germany have become members. While the distribution of English language records across time periods and subject divisions in WorldCat can be considered to represent the universe of publication reflected by those records, the same cannot be claimed for the foreign language profile. Although WorldCat has considerable international membership from major national libraries, many of the records from those libraries may be in the dataset of ANo call number present@ and could not be analyzed by subject. Hence, the holdings of those national libraries are probably under-represented in the language analysis and thus the languages are under-represented in the subject language analysis.

The distribution of records within the six foreign language groupings is shown in Table 5-7, *Foreign Language Groupings Number of Titles by Time Period*. The same pattern of decline in the number of records as was observed in the overall analysis of WorldCat records and in the preceding table in this chapter are evident here. The highest annual total for each group is highlighted in yellow in Table 5-7. Two language groupings, Japanese and Russian, also the languages with the lowest number of records, have the highest annual total in 1990. Spanish peaks in 1992. French and German, the two language groupings with the largest total number of records, both have the highest annual total in 1994. Chinese has a higher total closer to the year the sample was extracted. Chinese peaks in 1997. The joining of the national Library of China in 1998 may have influenced the pattern of distribution of Chinese language records. The year in which the peak number of records occurs for all six language groupings added together is 1994, influenced by the peak in that year of the two largest language groupings, French and

German.

In these data by imprint year, French maintains the largest number of records until 1900, then German maintains the largest number of records throughout the 20th century until 1985. After 1985, French regains the lead and retains it throughout the remaining years in the study. Again, these data may not reflect publication patterns, as the number of German language imprints is usually second only to English language materials in international publication, although publication in German speaking countries has become increasingly in English. In the OCLC Annual Report 1999/2000, the table showing the number and ranking of foreign language in WorldCat, shows French ranking second to English in the number of records.[2] But in the 2000/2001 Annual Report, German is in second place with French third.[3] The patterns seen here are more a reflection of collecting by the set of libraries contributing records to WorldCat than global publication patterns. As the international membership of WorldCat grows the number of foreign language records should increase and the percentage share of total by language may shift.

The differences in collecting for the seven language groupings between academic and research libraries can be compared by looking at Table 5-8 *Language Records 1985-1999: Academic and Research Libraries*. The table shows the number of records for each of the seven language groupings by five-year intervals for the last fifteen years in the study. The decline in the number of records overall in the last 5 years of the 20th century can be seen quite easily in this table. French has the highest number of records after English in both the research and academic library groupings. It is the only language in the academic libraries that has a higher number of records in the last time period than the preceding five-year intervals. No language grouping has its highest number in the last five years in the research libraries. While French fares very well in the academic library grouping, the number of German records in the academic libraries is approximately half the number for the research libraries. For English language imprints, the differential between academic and research libraries for 1985-1989 is approximately 10,000 fewer records in the academic library grouping. The differential increases for 1990-1994 time frame, then

remains about the same for the last five years of the 1990s. Spanish maintains third place with the highest number of records in the 1990-1994 time period for both library groupings. Japanese and Russian have considerably fewer records in the academic libraries than in the research libraries. The number for Chinese in the academic library grouping maintains about the same one to two ratio that the three major languages maintain in relation to the research library numbers. While Chinese maintains the same number of records across the fifteen years, Japanese plummets in the last time period in the academic libraries. In the academic library grouping, Chinese, French, and Spanish have a larger number of records in the 1990-1994 time period than in other two time periods. In the research library groupings all of the languages have the highest number of records from 1990-1994 except Russian.

As seen in Table 5-5, there seems to have been an increase in the acquisition of foreign language materials in the first half of the 1990s, or just an increase in records being added to WorldCat by the additional libraries joining OCLC. As the number of records declines in the latter 1990s, attributed at least in part to cataloging lag, the data in this study are not conclusive as to whether the number of records will be higher or lower for that time period as time goes on.

Might these percentage shares of WorldCat total by language be a reflection of the rate of publication in North America, Western Europe and the English-speaking world? The contribution of records by the national libraries in Europe, as well as China, Taiwan, and Japan, in the 1990s may have contributed records for a large proportion of publication from the industrialized nations.

Subject Analysis for the Seven Language Groupings

If we look at the language groupings by the 24 subject divisions shown in Table 5-9, *WorldCatBLanguage Titles by Subject Division*, we can see from the highlighted numbers in brackets that for all language groupings, the LLL division has the highest number of titles across the board. The unanimity continues with the second highest total being in the

history division in all of the seven language groupings. The first differential occurs with third place. In French, Japanese, Russian, Spanish, and English, business and economics has the third highest total. German has the third ranked total in philosophy/religion and Chinese has the third ranked total in library science, reference, and bibliography. After the third ranking, no pattern obtains and the rankings for the fourth spot are different. Three languages, French, Japanese, and Spanish have philosophy/religion ranked fourth. In Chinese, business and economics rank fourth; in German art and architecture rank fourth. In Russian and English, engineering and technology rank fourth. Chinese and English have philosophy/religion in fifth place. Business and economics is in fifth place in German. Political science is fifth in Russian. Law is fifth in Spanish. While the collecting emphases by subject and language are interesting, the rankings are probably not a surprise to most subject bibliographers. The emphases are a reflection of national cultures as well as academic curricula.

If we read the subject division lines across Table 5-9, the differences in collection emphases by language can be seen. The language, literature, linguistics division enjoys first place in all the languages as the area with the highest rate of monographic publication. History is likewise an area in which publication is primarily monographic and the rate of publication is high. To subject bibliographers, it would not be surprising that a high number of titles in German, after these two top ranked subject divisions, are in philosophy/religion. Another subject with a high rate of publication is business and economics and it ranks third in four out of the seven language groupings. These are the four subject divisions which reading across Table 5-9 are ranked in all of the language groupings. After the big four, the rankings vary by language for the next most collected subjects.

In Tables 5-1 through 5-5, the analysis was of English language and non-English language records. In the subject analysis in 5-9, the number of English language records are seen in comparison with the numbers for each of the other six language groupings by subject division. Analyzed in this way, it can be seen that on a one to one basis, the number of English language records overwhelm the other six languages when number of

records are compared by subjects. In the LLL division, English has 220,712 titles. Both French and German which are second and third in totals for LLL have only 38,031 and 34,823 titles. As seen in Table 5-9, the overall total number of records for French and German each are 146,815 and 140,006. In English, the LLL division alone has 220,712 titles.

Summary

The subject analysis from the 10% sample of WorldCat by language groupings shows that, for those records with call numbers in the sample, the ratio of English language to non-English language records is 65/35. Of the 35% in non-English language records, 67.5% of those records are comprised of the six foreign language groupings analyzed in the study. The six foreign language groupings contribute 24.8% of the total number of subject records in WorldCat. Of the total subject records, the research libraries hold 88% of all non-English records in WorldCat, whereas they hold 75% of all English language records. The academic libraries only hold 47% of non-English records, but 68% of English language records. Just the differences between the academic and the research libraries groupings indicate the large role the research libraries play in setting the profile of the WorldCat database.

Of all of the records in the 10% sample, approximately 65% are English language materials. In the mid 1990s the proportion of foreign language records in WorldCat began to increase and reached nearly 39% in 1997. This increase could be a result of the number of libraries internationally becoming OCLC members in the 1990s. As OCLC continues to implement its Global Strategy, AExtending the Cooperative", the number of international members and thus foreign language records may continue to increase.

The problem of the decline in the number of records in the last five-year interval of the study is further examined in Chapter Six. Click here for Chapter 6.

References

1.Anna H. Perrault, "The Shrinking National Collection: a Study of the Effects of the Diversion of Funds from Monographs to Serials on the Monograph Collections of Research Libraries," Library Acquisitions Practice & Theory, 18, no.1 (1994): 3-22.; "The Changing Print Resource Base of Academic Research Libraries in the United States," JELIS (Journal of Education for Library and Information Science) 36, no.4 (Fall 1995): 1-16; ANational Collecting Trends: Collection Analysis Methods and Findings.@ *Library & Information Science Research* 21, no.1(1999): 47-67.

2.OCLC Annual Report, 1999/2000: 46.

3.OCLC Annual Report, 2000/2001: 16.

Table 5-1
WorldCat
English and All Non-English as a Percent of Total Subject Records

Years	English	% English	Non-English	% Non-English	Total Records
Pre-1500	7	3.55%	190	96.45%	197
1500-1599	407	10.15%	3,601	89.85%	4,008
1600-1699	5,290	48.31%	5,659	51.69%	10,949
1700-1799	12,071	49.38%	12,375	50.62%	24,446
1800-1849	30,127	66.02%	15,503	33.98%	45,630
1850-1899	77,089	64.10%	43,181	35.90%	120,270
1900-1909	38,509	63.12%	22,504	36.88%	61,013
1910-1919	29,606	60.74%	19,139	39.26%	48,745
1920-1929	35,634	57.66%	26,168	42.34%	61,802
1930-1939	42,557	58.42%	30,284	41.58%	72,841
1940-1949	44,457	59.76%	29,941	40.24%	74,398
1950-1959	71,704	62.77%	42,535	37.23%	114,239
1960-1969	137,509	61.78%	85,084	38.22%	222,593
1970-1979	243,527	65.89%	126,061	34.11%	369,588
1980-1984	140,469	64.32%	77,927	35.68%	218,396
1985	28,912	64.14%	16,161	35.86%	45,073
1986	29,254	65.13%	15,663	34.87%	44,917
1987	29,512	64.88%	15,977	35.12%	45,489
1988	30,932	65.46%	16,323	34.54%	47,255
1989	31,463	64.82%	17,075	35.18%	48,538
1990	33,469	64.74%	18,230	35.26%	51,699
1991	34,543	65.43%	18,248	34.57%	52,791
1992	35,080	65.24%	18,694	34.76%	53,774
1993	33,502	63.67%	19,113	36.33%	52,615
1994	32,506	62.58%	19,435	37.42%	51,941
1995	31,035	62.87%	18,329	37.13%	49,364
1996	29,317	62.72%	17,424	37.28%	46,741
1997	27,864	61.23%	17,642	38.77%	45,506
1998	25,991	64.05%	14,587	35.95%	40,578
1999	23,037	69.06%	10,322	30.94%	33,359
2000	11,935	85.89%	1,960	14.11%	13,895
2001	789	99.75%	2	0.25%	791
Other	21,142	82.19%	4,582	17.81%	25,724
Totals	1,399,246		799,919		2,199,165

Table 5-2
Research Libraries
English and Non-English as a Percentage of Total Subject Records

Years	English	% English	Non-English	% Non-English	Total Records
Pre-1500	6	4.03%	143	95.97%	149
1500-1599	322	9.88%	2,937	90.12%	3,259
1600-1699	4,646	51.44%	4,386	48.56%	9,032
1700-1799	10,591	50.86%	10,232	49.14%	20,823
1800-1849	25,107	65.52%	13,212	34.48%	38,319
1850-1899	62,340	62.65%	37,171	37.35%	99,511
1900-1909	25,804	58.41%	18,377	41.59%	44,181
1910-1919	24,299	59.12%	16,799	40.88%	41,098
1920-1929	29,501	56.11%	23,074	43.89%	52,575
1930-1939	34,288	56.32%	26,597	43.68%	60,885
1940-1949	34,702	56.54%	26,671	43.46%	61,373
1950-1959	55,618	59.21%	38,315	40.79%	93,933
1960-1969	105,127	57.42%	77,972	42.58%	183,099
1970-1979	177,575	60.69%	115,019	39.31%	292,594
1980-1984	100,279	58.81%	70,233	41.19%	170,512
1985	21,257	59.46%	14,492	40.54%	35,749
1986	21,655	60.88%	13,915	39.12%	35,570
1987	21,636	61.05%	13,806	38.95%	35,442
1988	22,740	61.18%	14,429	38.82%	37,169
1989	23,296	60.90%	14,956	39.10%	38,252
1990	24,279	60.53%	15,829	39.47%	40,108
1991	25,936	62.14%	15,804	37.86%	41,740
1992	26,008	61.65%	16,179	38.35%	42,187
1993	24,760	59.87%	16,597	40.13%	41,357
1994	24,153	59.18%	16,663	40.82%	40,816
1995	22,895	59.43%	15,630	40.57%	38,525
1996	21,660	59.35%	14,835	40.65%	36,495
1997	20,590	58.51%	14,603	41.49%	35,193
1998	19,435	62.86%	11,485	37.14%	30,920
1999	17,689	72.07%	6,854	27.93%	24,543
2000	10,323	92.35%	855	7.65%	11,178
2001	782	99.74%	2	0.26%	784
Other	4,554	59.43%	3,109	40.57%	7,663
Total	1,043,853		701,181		1,745,034

Table 5-3: Academic Libraries
English and Non-English as a Percentage of Total Subject Records

Years	English	% English	Non-English	% Non-English	Total Records
Pre-1500	4	4.49%	85	95.51%	89
1500-1599	252	14.52%	1,484	85.48%	1,736
1600-1699	3,773	59.16%	2,605	40.84%	6,378
1700-1799	7,676	56.64%	5,876	43.36%	13,552
1800-1849	20,822	73.51%	7,505	26.49%	28,327
1850-1899	52,734	71.30%	21,223	28.70%	73,957
1900-1909	23,196	69.50%	10,180	30.50%	33,376
1910-1919	19,925	67.62%	9,540	32.38%	29,465
1920-1929	25,104	64.34%	13,912	35.66%	39,016
1930-1939	29,559	66.67%	14,775	33.33%	44,334
1940-1949	31,127	67.80%	14,781	32.20%	45,908
1950-1959	50,089	68.26%	23,291	31.74%	73,380
1960-1969	101,272	68.38%	46,831	31.62%	148,103
1970-1979	174,163	73.77%	61,940	26.23%	236,103
1980-1984	95,254	73.64%	34,090	26.36%	129,344
1985	19,531	74.03%	6,852	25.97%	26,383
1986	19,662	74.85%	6,606	25.15%	26,268
1987	19,985	74.61%	6,801	25.39%	26,786
1988	20,971	75.12%	6,945	24.88%	27,916
1989	21,325	74.78%	7,192	25.22%	28,517
1990	22,517	74.78%	7,594	25.22%	30,111
1991	22,937	75.18%	7,571	24.82%	30,508
1992	23,048	74.95%	7,703	25.05%	30,751
1993	22,297	75.27%	7,327	24.73%	29,624
1994	21,892	74.28%	7,581	25.72%	29,473
1995	21,038	74.78%	7,097	25.22%	28,135
1996	20,138	74.56%	6,872	25.44%	27,010
1997	18,863	73.69%	6,736	26.31%	25,599
1998	17,263	74.24%	5,989	25.76%	23,252
1999	13,988	76.66%	4,259	23.34%	18,247
2000	4,625	88.82%	582	11.18%	5,207
2001	41	100.00%	0	0.00%	41
Other	4,449	70.97%	1,820	29.03%	6,269
Totals	949,520		373,645		1,323,165

Table 5-4
Increase/Decrease by Decades for English/Non-English Records, Academic and Research Libraries

	1930-1939	1940-1949	% Increase/ Decrease	1950-1959	% Increase/ Decrease	1960-1969	% Increase/ Decrease	1970-1979	% Increase/ Decrease	1980-1989	% Increase/ Decrease	1990-1999	% Increase/ Decrease
Academic													
English	29,559	31,127	5.30%	50,089	60.92%	101,272	102.18%	174,163	71.98%	196,728	12.96%	203,981	3.69%
Non-English	14,775	14,781	0.04%	23,291	57.57%	46,831	101.07%	61,940	32.26%	68,486	10.57%	68,729	0.35%
Grand Total	44,334	45,908	3.55%	73,380	59.84%	148,103	101.83%	236,103	59.42%	265,214	12.33%	272,710	2.83%
Research													
English	34,288	34,702	1.21%	55,618	60.27%	105,127	89.02%	177,575	68.91%	210,863	18.75%	227,405	7.84%
Non-English	26,597	26,671	0.28%	38,315	43.66%	77,972	103.50%	115,019	47.51%	141,831	23.31%	144,479	1.87%
Grand Total	60,885	61,373	0.80%	93,933	53.05%	183,099	94.93%	292,594	59.80%	352,694	20.54%	371,884	5.44%

Table 5-5
Increase/Decrease by Five Year Periods, 1980-1999 for English/Non-English Language Groupings - Academic and Research Libraries

	1980-1984	1985-1989	Percent Increase/ Decrease	1990-1994	Percent Increase/ Decrease	1995-1999	Percent Increase/ Decrease
Academic							
English	95,254	101,474	6.53%	112,691	11.05%	91,290	-18.99%
Grand Total	129,344	127,271	-1.60%	142,248	11.77%	116,477	-18.12%
Non-English	34,090	25,797	-24.33%	29,557	14.58%	25,187	-14.78%
Research							
English	100,279	110,584	10.28%	125,136	13.16%	102,269	-18.27%
Grand Total	190,800	157,342	-17.54%	177,977	13.11%	143,512	-19.36%
Non-English	90,521	46,758	-48.35%	52,841	13.01%	41,243	-21.95%

Table 5-6
Foreign Language Groupings as Percentage of Total Records

	Number of Records	Percentage of Total Non-English Records	Percentage of Total WorldCat Subject Records
Chinese	52,605	6.6%	2.4%
French	146,815	18.4%	6.7%
German	140,006	17.5%	6.4%
Japanese	37,302	5.0%	2.0%
Russian	50,373	6.3%	2.3%
Spanish	109,260	13.7%	5.0%
Totals	**536,361**	**67.5%**	**24.8%**

Table 5-7
Foreign Language Groupings
Number of Titles by Time Period

LANGUAGE	Pre-1500	1500-1599	1600-1699	1700-1799	1800-1849	1850-1899	1900-1909	1910-1919
Chinese	1	16	43	111	118	341	282	538
French	0	191	1,347	5,045	6,230	13,579	5,372	3,966
German	5	626	499	2,166	3,882	12,844	7,028	6,227
Japanese	0	0	29	57	102	424	452	420
Russian	0	0	0	26	211	1,478	1,220	999
Spanish	0	90	249	890	1,321	4,390	2,405	2,263
Total	6	923	2,167	8,295	11,864	33,056	16,759	14,413
	1920-1929	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1984	1985
Chinese	1,087	2,698	1,813	2,736	4,058	6,129	6,813	1,637
French	5,819	5,521	4,665	6,531	11,568	18,467	10,694	2,343
German	7,576	7,629	5,209	7,572	14,826	20,581	10,848	2,156
Japanese	807	1,702	2,101	1,535	3,395	5,942	5,496	1,014
Russian	1,269	1,280	1,176	3,505	7,345	9,576	5,830	1,246
Spanish	2,877	3,767	5,737	6,549	11,958	18,525	10,650	2,212
Total	19,435	22,597	20,701	28,428	53,150	79,220	50,331	10,608
	1986	1987	1988	1989	1990	1991	1992	1993
Chinese	1,397	1,766	1,568	1,674	1,640	1,697	1,792	1,996
French	2,544	2,594	2,772	3,012	3,245	3,120	3,398	3,422
German	2,036	1,942	2,137	2,114	2,312	2,392	2,317	2,424
Japanese	830	942	981	1,094	**1,199**	1,136	1,135	1,173
Russian	1,276	1,141	1,185	1,250	**1,277**	1,206	1,036	1,114
Spanish	2,246	2,353	2,362	2,404	2,613	2,649	**2,831**	2,815
Total	10,329	10,738	11,005	11,548	12,286	12,200	12,509	12,944
	1994	1995	1996	1997	1998	1999	2000	2001
Chinese	1,932	1,936	1,653	**2,224**	1,417	923	91	0
French	**3,523**	3,402	3,170	3,375	3,145	2,949	893	1
German	**2,507**	2,233	2,096	2,022	1,450	1,071	191	0
Japanese	1,139	995	1,009	888	748	434	22	0
Russian	1,098	1,107	1,113	1,011	823	467	33	0
Spanish	2,968	2,720	2,510	2,595	2,279	1,342	177	1
Total	**13,167**	12,393	11,551	12,115	9,862	7,186	1,407	2
	Other	TOTAL						
Chinese	478	52,605						
French	912	146,815						
German	1,088	140,006						
Japanese	101	37,302						
Russian	75	50,373						
Spanish	512	109,260						
Total	3,166	536,361						

Table 5-8
Language Records 1985-1999
Academic and Research Libraries

	1985-1989		1990-1994		1995-1999	
	Academic	**Research**	**Academic**	**Research**	**Academic**	**Research**
Chinese	3,469	6,626	3,864	7,420	3,391	5,790
English	101,474	110,584	112,691	125,136	91,290	102,269
French	7,152	11,135	9,869	12,728	10,351	10,532
German	5,722	9,586	5,651	10,960	4,372	7,998
Japanese	819	4,550	1,350	5,393	393	3,769
Russian	2,047	5,909	1,594	5,467	1,196	4,318
Spanish	6,588	8,970	7,229	10,873	5,484	8,836

Table 5-9
WorldCat - Language Titles By Subject Division

DIVISION		Chinese		French		German		Japanese		Russian		Spanish		English
AGRICULTURE		537		1,949		1,326		450		903		1,559		40,441
ANTHROPOLOGY		543		1,234		1,083		591		398		847		7,164
ART AND ARCHITECTURE		2,275	[5]	8,998	[4]	11,123	[5]	2,379		1,493		3,865		50,164
BIOLOGICAL SCIENCES		393		2,878		2,991		414		1,246		1,444		33,075
BUSINESS AND ECONOMICS	[4]	4,456	[3]	12,520	[5]	10,646	[3]	5,903	[3]	5,236	[3]	11,588	[3]	142,429
CHEMISTRY		89		512		750		55		480		122		7,743
COMPUTER SCIENCE		184		995		319		89		276		230		12,746
EDUCATION		771		2,685		2,969		926		561		2,856		82,255
ENGINEERING AND TECHNOLOGY		1,545		5,531		4,912		932	[4]	4,537		2,334	[4]	125,078
GEOGRAPHY AND EARTH SCIENCES		591		3,056		2,457		399		2,361		1,349		40,708
HISTORY AND AUXILIARY SCIENCES	[2]	9,019	[2]	24,580	[2]	19,434	[2]	7,088	[2]	7,332	[2]	19,165	[2]	164,371
LANGUAGE, LINGUISTICS, AND LITERATURE	[1]	16,330	[1]	38,031	[1]	34,823	[1]	8,254	[1]	14,374	[1]	34,451	[1]	220,712
LAW		1,290		4,384		5,086		919		835	[5]	5,842	[4]	60,138
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	[3]	5,064		3,680		3,689		1,162		1,456		2,497		44,705
MATHEMATICS		197		1,963		1,603		115		801		390		14,733
MEDICINE		1,164		3,283		2,807		452		545		2,209		54,941
MUSIC		353		1,912		3,834		216		622		854		15,747
PERFORMING ARTS		384		1,265		1,050		498		522		714		9,116
PHILOSOPHY AND RELIGION	[5]	3,319	[4]	12,000	[3]	17,432	[4]	2,523		1,783	[5]	6,464	[4]	103,784
PHYSICAL EDUCATION AND RECREATION		322		1,007		758		222		219		604		19,273
PHYSICAL SCIENCES		444		2,357		2,156		214		1,214		676		23,849
POLITICAL SCIENCE		1,708		5,316		3,782		1,554	[5]	1,834		4,200		48,948
PSYCHOLOGY		557		1,762		1,236		178		208		1,031		15,901
SOCIOLOGY		1,070		4,935		3,740		1,769		1,137		3,969		61,225
Totals		52,605		146,833		140,006		37,302		50,373		109,260		1,399,246

* **Highlighted bracketed numbers indicate ranking.**

Chapter Six

WorldCat ISBN Analysis

ISBN data were requested to further explore the nature of the bibliographic records in the sample when it became known that a high percentage of the bibliographic records did not have a usable call number for subject analysis.[1] It was thought that determining the number of records with ISBN numbers might shed more light on the nature of those records lacking usable call numbers. It seemed reasonable to expect that those records without call numbers might be foreign language records or records for older materials. Would the ISBN records show what proportion of the bibliographic records are for mainstream publications? Could the "No call number present" records be primarily foreign language publications?

Three tables were constructed: *OCLC iCAS—WorldCat ISBN Analysis All Titles Held by Date (All Divisions)*; *OCLC iCAS—WorldCat English ISBN…*; and *OCLC iCAS—WorldCat Foreign ISBN…* The data from these tables were used to construct Tables 6-1, 6-2, 6-3, attached to this chapter. For those records with both ISBN numbers and call numbers, subject analysis by publication year was performed. The records without call numbers were included by publication year as one "No Call Number Present" line. Table 6-4, *WorldCat, English, and Foreign ISBN Analysis by Time Period,* is a summary table of WorldCat records with ISBN numbers from the three main ISBN tables, with and without call numbers in the records. The data are also divided into records for English language monographs and foreign language monographs and are shown in the timeline defined for the study.

As can be seen in Table 6-4, the number of records with both ISBN numbers and call numbers (607,856) in WorldCat, far exceeds the number of records with an ISBN number (104,064) but without call numbers. This obtains no matter whether the records are

[1] The ISBN analysis was Ed O'Neill's idea.

for English language or foreign language materials. Seventy-one percent of the records in the sample with ISBN numbers also have call numbers. It is interesting to note that even though the idea of using ISBN numbers began in the prior to the 1960s there are not a considerable number of records with ISBN numbers until the 1970s. The records with ISBN numbers for imprint years before the inauguration of the ISBN system are probably for facsimiles or reprints that carry the date of the original publications.

It is also easy to see that the number of English language titles with both call numbers and ISBN numbers (381,428) in WorldCat are higher than the total number of foreign language titles (297,668) with ISBN numbers. For English language records, 92% of those records with ISBN numbers also have call numbers. Sixty-eight percent of ISBN records without call numbers are foreign language records. For the foreign language records, only 76% with ISBN numbers also have call numbers. The number of foreign language ISBN records without call numbers (71,240) is larger than the number of English language records without call numbers (32,824).

If we look at the increase/decrease over time in the number of records in the ISBN analysis, we see that for English language records with call numbers, the number of records increases annually from 1985 until 1992 and then remains nearly the same until 1996. English language records without call numbers increase gradually each year from 1985 forward with a peak of 2,622 in 1997, and only slight declines in 1998 and 1999. The declines are not as serious in the ISBN data for English language records as for foreign language records, especially in the latter half of the 1990s. If we use the ISBN number as an indication that the titles are from mainstream publishers, then it appears that titles from mainstream publishers are being consistently added to WorldCat within a current time frame. The decreases in numbers of records in the latter 1990s are heavier in foreign language titles and not in English language titles with ISBN numbers.

As with the English language records, in the foreign language ISBN analysis, the number of records with ISBN and call numbers is much larger than those without call numbers. Looking at the foreign language records with ISBN numbers by time period we

can see that the number of foreign language records starts with a low of 8,277 in 1985, and increases gradually to the low 20,000 range in 1997, and then trends downward in the last three years. The foreign language records with call number set the pattern for all of the foreign language records with ISBN numbers. The number of records with call numbers increases steadily from a low of 6,463 in 1985, to a high of 14,541 in 1994, with slight fluctuation to 14,324 in 1997, trending downward with a large decrease in 1998, and a precipitous decline in 1999 and 2000. The foreign language records without call number, likewise, begin with a low of 1,814 records in 1985 and trend upward until a high of 5,789 records in 1997.

From 1993 to 1997 the absolute numbers of ISBN records for English and foreign languages are very close to the same with the foreign language records actually having higher totals than the English from 1994-1997. For 1999 and 2000 there are large differences with English having substantially more records than foreign languages. If we compare the absolute numbers in the “without call number” columns for both, we can see that the number of foreign language records without call numbers is much higher than the number of English language records without call numbers. For both English and foreign without call numbers the decrease is not as severe as for those with call numbers. Those without call numbers do not decrease considerably until after 1998. The decline in the foreign language records with ISBN numbers is responsible for much of the overall decrease in the ISBN records in WorldCat, as the English language records have slight decreases in 1997 and 1998 with an actual 2% gain in 1999.

The overall decline in the number of records with both call numbers and ISBN numbers in English and foreign languages is not as steep as the decline in the overall WorldCat totals seen in Table 2-1. In the ISBN data, the total number of records in WorldCat declines 6% in 1998, 12% in 1999, and 56% in 2000. For all records in WorldCat, the decline is 11% in 1998, 18% in 1999, and 63% in 2000. It is not unexpected that records with ISBN numbers would have a lower rate of decrease than records without ISBN numbers. The ISBN analysis does show that the rate of decrease in the late 1990s is

lower for those records with ISBN numbers. This somewhat supports the speculation that the “No call number present” records are largely comprised of foreign language titles and titles which are not from mainstream U.S. and European publishers. This does assist in confirming the suspicion that the records without call numbers are not mainstream materials.

If the data for number of titles with ISBN numbers is compared to total numbers by date of publication for WorldCat, the percentage with ISBN numbers makes up a small proportion of total records. Only 27.6% of all records with call numbers in the sample have ISBN numbers. Less than 10% of all records without call numbers in WorldCat have ISBN numbers.

Table 6-5

Comparison of Total WorldCat Records with ISBN Totals

	WC Total With Call Number	ISBN With Call Number	WC Total Without Call Number	ISBN Without Call Number
WorldCat Totals All Years	2,199,165	607,856 (27.6%)	1,179,107	104,064 (9%)
WC Totals 1970-1999	1,041,849	592,401 (57%)	566,600	102,930 (18%)

Since the total number of records covers the entire time span of printed works, it would be a more useful comparison to calculate the percentage of ISBN records to the totals for the sample of titles published from 1970-1999. Fifty seven percent of all records with call numbers in the 1970-1999 timeframe have ISBN numbers. Those records without call numbers, but with ISBN numbers, make up 18% of the total number of records without call numbers for the 1970-1999 time period. Thus, there are still 43% of records in the sample in the 1970-1999 time frame that have call numbers, but not ISBN numbers. And

82% of records in the sample without call numbers from 1970-1999 also do not have ISBN numbers. These proportions indicate a considerable number of titles that do not appear to be from mainstream North American and European publishers.

For those bibliographic records having both call numbers and ISBN numbers we can look at the profile by subject division. Table 6-6 *WorldCat ISBN Tables: Records With Call Number Present by Subject* is a summary table by subject divisions showing the number of English language records and the number of foreign language records with the total subject records in WorldCat that have ISBN numbers. In all subject divisions the number of English records is larger than the number of foreign language records. Only in a few subject divisions are the number of English language records and foreign language records even close to the same number. In law there are only 206 records less in foreign languages than in English and in anthropology there are only 212 records difference. In art/architecture, the two totals are very close with a difference of 1,488 records.

For the majority of the subject divisions the number of English language records with ISBN numbers is considerably larger than the number of foreign language records. Those with an especially large differential are computer science in which the English runs almost 6 to one. The same for engineering with a difference of 22,693 titles between English and foreign. Medicine is another area with a large discrepancy, almost 4 to one. Sociology runs at almost two to one. Education has a large gap between 12,330 and 4,713 titles. The sciences and the social sciences are collected predominantly in English and this is reflected in the English to foreign language ratios.

In subject areas in which research libraries collect globally, the differential between the English language and foreign language totals reflects the foreign language collecting in those libraries. Both LLL and the history division have a considerable number of foreign language records. The differential in history is not large with the proportions being 54/46 percent or14 percentage points difference. For LLL, the difference is twice that of history at 28 percentage points. These two divisions have the largest number of records of all the subject divisions as has been shown in the preceding chapters. It is not surprising that

philosophy/religion has 38% in foreign language records.

The remaining subject divisions have large differences between the number of English and the number of foreign language records. The subject divisions with the highest publication rates, business and economics, history, LLL, appear to be the early adopters of ISBN and have the largest number of records with ISBN numbers.

As the foreign language collections are only represented in this analysis by records with ISBN numbers, the non-English representation is most probably low. It is possible that the proportion of records without call numbers may be largely foreign language publications, based upon the ISBN data and the unique titles and language analyses. There are many foreign language theses and other types of reports which typically are not classified, even in U.S. universities. These materials would have neither call number nor ISBN numbers. Older works, both English language and foreign languages, may also have neither call number nor ISBN numbers. While books considered to be juvenile literature have ISBN numbers if published by trade houses, as pointed out in Chapter Three, many juvenile works are not classified and thus, would be in the “No Call Number” with ISBN numbers category. And other types of works not published by trade publishers, such as local history and many works of literature that do not receive widespread distribution may not have ISBN numbers. In answer to the questions posed at the beginning of this chapter, WorldCat contains records for a very large number of titles which possibly are not from mainstream publishers and the number of foreign language titles with ISBN numbers is considerably lower than for English language titles. These are indications of the vast diversity of resources represented in the WorldCat database.

Click here for Chapter 7.

Table 6-1a
WorldCat ISBN Analysis
All Titles Held By Date
(All Divisions)

SUBJECT DIVISION	Pre-1500	1500-1599	1600-1699	1700-1799	1800-1849	1850-1899	1900-1909	1910-1919	1920-1929
AGRICULTURE	0	0	0	2	7	89	10	0	0
ANTHROPOLOGY	0	0	0	0	4	19	7	3	0
ART AND ARCHITECTURE	0	0	0	2	6	50	27	3	2
BIOLOGICAL SCIENCES	0	0	0	5	11	72	9	0	1
BUSINESS AND ECONOMICS	0	0	0	17	69	330	53	3	3
CHEMISTRY	0	0	1	0	0	8	5	0	1
COMPUTER SCIENCE	0	0	0	0	0	0	3	0	1
EDUCATION	0	0	0	0	23	143	23	9	1
ENGINEERING AND TECHNOLOGY	0	0	0	2	20	151	52	2	1
GEOGRAPHY AND EARTH SCIENCES	0	1	4	52	55	125	14	0	2
HISTORY AND AUXILIARY SCIENCES	0	0	8	142	265	942	153	36	13
LANGUAGE, LINGUISTICS, AND LITERATURE	1	0	1	12	81	513	211	36	44
LAW	0	0	0	10	57	184	45	0	1
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	0	0	0	4	25	100	45	4	1
MATHEMATICS	0	0	0	1	5	23	8	0	1
MEDICINE	0	0	0	1	14	88	28	2	2
MUSIC	0	0	0	0	14	70	18	1	1
PERFORMING ARTS	0	0	0	0	10	59	5	0	0
PHILOSOPHY AND RELIGION	0	1	2	34	255	1,992	641	339	8
PHYSICAL EDUCATION AND RECREATION	0	0	0	0	3	35	4	1	1
PHYSICAL SCIENCES	0	0	1	1	1	20	5	1	1
POLITICAL SCIENCE	0	0	0	29	80	264	23	0	4
PSYCHOLOGY	0	0	0	0	1	14	14	8	2
SOCIOLOCY	0	0	0	2	25	131	16	4	4
Subject Analysis	**1**	**2**	**17**	**316**	**1,031**	**5,422**	**1,419**	**452**	**95**
No Call Number Present	0	0	4	7	23	71	89	99	86
Grand Total	**1**	**2**	**21**	**323**	**1,054**	**5,493**	**1,508**	**551**	**181**

Table 6-1b
WorldCat ISBN Analysis
All Titles Held By Date
(All Divisions)

SUBJECT DIVISION	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1984	1985	1986	1987
AGRICULTURE	0	2	2	103	1,670	1,481	386	375	395
ANTHROPOLOGY	3	1	2	43	701	610	143	145	167
ART AND ARCHITECTURE	3	5	14	244	4,040	3,716	882	883	985
BIOLOGICAL SCIENCES	3	2	8	116	2,232	1,661	423	454	470
BUSINESS AND ECONOMICS	6	9	13	366	7,865	7,348	1,714	1,919	2,020
CHEMISTRY	0	0	3	40	637	392	106	87	113
COMPUTER SCIENCE	0	0	0	17	440	824	272	275	277
EDUCATION	4	5	5	117	3,060	2,263	517	594	608
ENGINEERING AND TECHNOLOGY	1	4	19	417	6,825	5,753	1,442	1,545	1,410
GEOGRAPHY AND EARTH SCIENCES	1	2	3	81	1,649	1,508	345	394	388
HISTORY AND AUXILIARY SCIENCES	17	20	59	707	10,481	8,002	1,976	2,041	2,359
LANGUAGE, LINGUISTICS, AND LITERATURE	60	76	235	1,957	24,525	17,963	4,378	4,505	4,755
LAW	0	2	4	92	2,323	2,807	719	778	831
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	7	4	7	111	2,272	2,020	443	437	725
MATHEMATICS	0	3	6	71	1,218	974	258	251	267
MEDICINE	1	2	8	192	4,294	3,682	864	934	959
MUSIC	4	6	12	75	1,243	964	264	245	247
PERFORMING ARTS	1	2	5	40	980	793	200	213	199
PHILOSOPHY AND RELIGION	16	18	76	480	6,794	5,393	1,295	1,325	1,389
PHYSICAL EDUCATION AND RECREATION	0	1	9	94	1,895	1,410	310	306	356
PHYSICAL SCIENCES	3	2	2	102	1,305	1,010	259	255	311
POLITICAL SCIENCE	1	1	9	148	2,732	2,276	572	627	726
PSYCHOLOGY	5	4	17	79	1,541	1,136	262	332	290
SOCIOLOCY	5	2	6	170	4,297	3,558	809	895	996
Subject Analysis	**141**	**173**	**524**	**5,862**	**95,019**	**77,544**	**18,839**	**19,815**	**21,243**
No Call Number Present	91	83	147	427	7,376	11,275	2,771	3,124	3,698
Grand Total	**232**	**256**	**671**	**6,289**	**102,395**	**88,819**	**21,610**	**22,939**	**24,941**

Table 6-1c
WorldCat ISBN Analysis
All Titles Held By Date
(All Divisions)

SUBJECT DIVISION	1988	1989	1990	1991	1992	1993	1994	1995	1996
AGRICULTURE	435	445	491	521	504	557	477	480	501
ANTHROPOLOGY	211	224	246	286	281	283	338	290	319
ART AND ARCHITECTURE	1,151	1,260	1,317	1,435	1,453	1,446	1,467	1,492	1,587
BIOLOGICAL SCIENCES	528	533	639	677	672	668	639	536	516
BUSINESS AND ECONOMICS	2,404	2,470	2,571	2,786	2,888	2,840	2,937	2,725	2,563
CHEMISTRY	117	138	132	157	154	128	116	111	109
COMPUTER SCIENCE	277	374	429	443	466	463	492	508	527
EDUCATION	692	744	789	793	823	888	848	830	823
ENGINEERING AND TECHNOLOGY	1,560	1,677	1,797	1,910	1,931	1,831	1,873	1,733	1,699
GEOGRAPHY AND EARTH SCIENCES	484	582	562	782	706	719	666	639	633
HISTORY AND AUXILIARY SCIENCES	2,726	2,976	3,186	3,434	3,628	3,575	3,845	3,882	3,734
LANGUAGE, LINGUISTICS, AND LITERATURE	5,297	5,951	6,291	6,995	7,039	7,481	7,685	7,566	7,473
LAW	847	1,009	1,197	1,387	1,483	1,529	1,704	1,506	1,523
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	589	578	593	555	605	536	519	473	514
MATHEMATICS	290	312	333	357	378	373	344	347	286
MEDICINE	1,174	1,124	1,157	1,254	1,326	1,220	1,312	1,202	1,262
MUSIC	272	285	277	331	314	304	343	348	281
PERFORMING ARTS	237	242	287	294	310	336	282	335	346
PHILOSOPHY AND RELIGION	1,505	1,669	1,786	1,781	1,930	2,020	1,987	2,026	2,128
PHYSICAL EDUCATION AND RECREATION	385	382	406	508	469	521	489	494	555
PHYSICAL SCIENCES	371	379	411	484	415	400	380	438	363
POLITICAL SCIENCE	799	863	924	988	1,002	982	982	976	942
PSYCHOLOGY	325	395	385	405	434	433	489	402	436
SOCIOLOCY	1,033	1,207	1,232	1,429	1,486	1,513	1,469	1,487	1,442
Subject Analysis	**23,709**	**25,819**	**27,438**	**29,992**	**30,697**	**31,046**	**31,683**	**30,826**	**30,562**
No Call Number Present	4,172	4,871	5,170	5,309	5,462	5,598	5,945	6,557	7,614
Grand Total	**27,881**	**30,690**	**32,608**	**35,301**	**36,159**	**36,644**	**37,628**	**37,383**	**38,176**

Table 6-1d
WorldCat ISBN Analysis
All Titles Held By Date
(All Divisions)

SUBJECT DIVISION	1997	1998	1999	2000	2001	Other	Total
AGRICULTURE	505	426	414	216	12	10	10,516
ANTHROPOLOGY	339	308	256	106	7	2	5,344
ART AND ARCHITECTURE	1,576	1,479	1,262	486	12	9	28,294
BIOLOGICAL SCIENCES	536	520	420	267	23	2	12,643
BUSINESS AND ECONOMICS	2,573	2,318	2,010	924	68	12	53,824
CHEMISTRY	88	80	80	59	5	1	2,868
COMPUTER SCIENCE	527	509	660	313	21	2	8,120
EDUCATION	770	760	587	286	47	0	17,052
ENGINEERING AND TECHNOLOGY	1,743	1,633	1,509	777	52	12	39,381
GEOGRAPHY AND EARTH SCIENCES	599	603	569	384	18	7	12,577
HISTORY AND AUXILIARY SCIENCES	3,722	3,621	3,067	1,212	83	27	69,939
LANGUAGE, LINGUISTICS, AND LITERATURE	7,278	7,090	6,305	3,140	154	41	145,139
LAW	1,573	1,472	1,154	366	16	3	24,622
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	606	418	327	140	5	6	12,669
MATHEMATICS	290	305	252	127	20	0	7,100
MEDICINE	1,256	1,130	1,176	692	52	3	26,411
MUSIC	331	305	319	139	13	4	7,030
PERFORMING ARTS	344	326	325	149	9	0	6,329
PHILOSOPHY AND RELIGION	2,029	1,936	1,750	824	51	18	43,498
PHYSICAL EDUCATION AND RECREATION	569	494	519	290	21	4	10,531
PHYSICAL SCIENCES	347	328	318	169	19	1	8,102
POLITICAL SCIENCE	901	872	789	443	13	1	18,969
PSYCHOLOGY	440	419	442	233	16	0	8,959
SOCIOLOCY	1,441	1,391	1,198	635	50	6	27,939
Subject Analysis	**30,383**	**28,743**	**25,708**	**12,377**	**787**	**171**	**607,856**
No Call Number Present	8,381	7,538	6,345	1,603	7	121	104,064
Grand Total	**38,764**	**36,281**	**32,053**	**13,980**	**794**	**292**	**711,920**

Table 6-2a
WorldCat English ISBN

SUBJECT DIVISION	1500-1599	1600-1699	1700-1799	1800-1849	1850-1899	1900-1909	1910-1919	1920-1929
AGRICULTURE	0	0	2	7	89	8	0	0
ANTHROPOLOGY	0	0	0	3	18	4	3	0
ART AND ARCHITECTURE	0	0	1	6	50	15	3	1
BIOLOGICAL SCIENCES	0	0	5	11	71	6	0	1
BUSINESS AND ECONOMICS	0	0	17	69	329	30	3	2
CHEMISTRY	0	1	0	0	8	3	0	0
COMPUTER SCIENCE	0	0	0	0	0	2	0	1
EDUCATION	0	0	0	23	143	17	9	1
ENGINEERING AND TECHNOLOGY	0	0	2	20	151	35	2	1
GEOGRAPHY AND EARTH SCIENCES	1	4	51	55	125	6	0	2
HISTORY AND AUXILIARY SCIENCES	0	7	142	262	934	101	35	11
LANGUAGE, LINGUISTICS, AND LITERATURE	0	1	12	79	505	133	35	39
LAW	0	0	10	57	184	18	0	1
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	0	0	3	25	100	19	4	1
MATHEMATICS	0	0	0	5	23	5	0	1
MEDICINE	0	0	1	14	88	19	2	2
MUSIC	0	0	0	14	69	13	1	1
PERFORMING ARTS	0	0	0	10	59	3	0	0
PHILOSOPHY AND RELIGION	1	2	34	255	1,992	616	337	8
PHYSICAL EDUCATION AND RECREATION	0	0	0	3	35	3	1	1
PHYSICAL SCIENCES	0	1	1	1	20	5	1	1
POLITICAL SCIENCE	0	0	29	80	263	17	0	3
PSYCHOLOGY	0	0	0	1	14	13	8	2
SOCIOLOCY	0	0	2	25	131	10	4	4
Subject Analysis	**2**	**16**	**312**	**1,025**	**5,401**	**1,101**	**448**	**84**
No Call Number Present	0	2	4	18	62	71	97	78
Grand Total	**2**	**18**	**316**	**1,043**	**5,463**	**1,172**	**545**	**162**

Table 6-2b
WorldCat English ISBN

SUBJECT DIVISION	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1984	1985	1986
AGRICULTURE	0	2	2	98	1,415	1,158	293	297
ANTHROPOLOGY	1	1	1	39	483	336	62	73
ART AND ARCHITECTURE	0	2	11	225	2,764	2,147	476	462
BIOLOGICAL SCIENCES	3	1	6	111	1,808	1,273	328	354
BUSINESS AND ECONOMICS	4	7	12	357	5,524	4,729	1,089	1,257
CHEMISTRY	0	0	1	40	551	314	90	72
COMPUTER SCIENCE	0	0	0	16	342	707	235	244
EDUCATION	4	5	5	116	2,311	1,613	383	454
ENGINEERING AND TECHNOLOGY	1	3	16	407	5,805	4,690	1,166	1,256
GEOGRAPHY AND EARTH SCIENCES	1	2	2	78	1,296	1,120	268	289
HISTORY AND AUXILIARY SCIENCES	12	17	49	672	7,365	4,586	1,096	1,177
LANGUAGE, LINGUISTICS, AND LITERATURE	48	60	164	1,782	17,305	11,044	2,707	2,809
LAW	0	2	3	88	1,479	1,459	349	387
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	5	3	1	102	1,700	1,311	288	285
MATHEMATICS	0	3	6	69	999	756	211	196
MEDICINE	0	2	4	188	3,413	3,020	715	774
MUSIC	3	6	9	66	832	582	165	146
PERFORMING ARTS	1	2	3	38	752	506	122	133
PHILOSOPHY AND RELIGION	12	17	59	439	4,640	3,338	793	799
PHYSICAL EDUCATION AND RECREATION	0	1	9	93	1,629	1,192	269	258
PHYSICAL SCIENCES	1	2	2	99	1,019	701	194	174
POLITICAL SCIENCE	1	0	7	136	1,938	1,414	345	402
PSYCHOLOGY	5	3	16	72	1,156	789	194	244
SOCIOLOCY	5	2	3	165	3,156	2,385	538	628
Subject Analysis	**107**	**143**	**391**	**5,496**	**69,682**	**51,170**	**12,376**	**13,170**
No Call Number Present	76	69	73	232	2,854	4,001	957	1,014
Grand Total	**183**	**212**	**464**	**5,728**	**72,536**	**55,171**	**13,333**	**14,184**

Table 6-2c
WorldCat English ISBN

SUBJECT DIVISION	1987	1988	1989	1990	1991	1992	1993	1994
AGRICULTURE	284	300	322	368	374	354	389	326
ANTHROPOLOGY	86	98	100	106	134	143	127	159
ART AND ARCHITECTURE	528	576	639	604	705	680	667	645
BIOLOGICAL SCIENCES	360	368	410	435	483	489	487	452
BUSINESS AND ECONOMICS	1,198	1,298	1,315	1,368	1,524	1,495	1,439	1,518
CHEMISTRY	94	82	103	97	127	113	102	100
COMPUTER SCIENCE	233	229	291	310	338	372	350	396
EDUCATION	435	479	527	549	558	566	615	558
ENGINEERING AND TECHNOLOGY	1,131	1,151	1,246	1,323	1,403	1,441	1,379	1,392
GEOGRAPHY AND EARTH SCIENCES	293	314	382	362	539	482	474	456
HISTORY AND AUXILIARY SCIENCES	1,289	1,418	1,461	1,514	1,591	1,621	1,526	1,616
LANGUAGE, LINGUISTICS, AND LITERATURE	2,912	2,886	3,201	3,329	3,840	3,782	3,757	3,657
LAW	446	423	460	531	602	736	668	777
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	312	337	310	323	313	309	260	250
MATHEMATICS	196	210	233	239	263	268	271	243
MEDICINE	784	957	908	910	975	1,040	934	988
MUSIC	143	157	171	151	195	179	180	195
PERFORMING ARTS	117	138	135	156	157	168	174	132
PHILOSOPHY AND RELIGION	828	889	937	1,002	996	970	997	1,019
PHYSICAL EDUCATION AND RECREATION	282	319	308	318	424	380	419	406
PHYSICAL SCIENCES	215	228	231	245	322	279	250	243
POLITICAL SCIENCE	421	432	483	509	565	514	459	453
PSYCHOLOGY	185	216	242	252	284	290	267	307
SOCIOLOCY	667	656	759	713	868	922	890	854
Subject Analysis	**13,439**	**14,161**	**15,174**	**15,714**	**17,580**	**17,593**	**17,081**	**17,142**
No Call Number Present	1,245	1,202	1,290	1,431	1,455	1,435	1,504	1,491
Grand Total	**14,684**	**15,363**	**16,464**	**17,145**	**19,035**	**19,028**	**18,585**	**18,633**

Table 6-2d
WorldCat English ISBN

SUBJECT DIVISION	1995	1996	1997	1998	1999	2000	2001	Other
AGRICULTURE	316	351	355	305	322	203	12	6
ANTHROPOLOGY	125	140	141	145	158	85	7	0
ART AND ARCHITECTURE	684	691	657	662	620	356	12	2
BIOLOGICAL SCIENCES	360	358	364	341	318	245	23	1
BUSINESS AND ECONOMICS	1,444	1,353	1,326	1,300	1,261	793	67	6
CHEMISTRY	78	79	50	52	54	47	5	1
COMPUTER SCIENCE	369	366	351	343	510	279	21	0
EDUCATION	579	568	536	557	424	257	47	0
ENGINEERING AND TECHNOLOGY	1,276	1,232	1,290	1,211	1,242	707	51	7
GEOGRAPHY AND EARTH SCIENCES	433	408	359	399	423	322	18	1
HISTORY AND AUXILIARY SCIENCES	1,630	1,630	1,577	1,641	1,586	969	83	10
LANGUAGE, LINGUISTICS, AND LITERATURE	3,688	3,623	3,408	3,386	3,586	2,606	153	22
LAW	659	714	719	760	600	266	16	0
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	219	219	210	206	180	78	5	0
MATHEMATICS	272	186	210	224	179	114	19	0
MEDICINE	913	964	921	879	965	642	51	2
MUSIC	179	151	186	167	207	119	13	1
PERFORMING ARTS	161	180	176	164	216	130	9	0
PHILOSOPHY AND RELIGION	1,051	1,122	1,030	1,053	1,055	647	50	9
PHYSICAL EDUCATION AND RECREATION	396	456	456	414	457	274	20	2
PHYSICAL SCIENCES	251	223	195	194	216	132	19	1
POLITICAL SCIENCE	451	469	439	455	446	360	13	0
PSYCHOLOGY	253	271	282	276	310	193	16	0
SOCIOLOCY	876	834	821	833	795	535	50	3
Subject Analysis	**16,663**	**16,588**	**16,059**	**15,967**	**16,130**	**10,359**	**780**	**74**
No Call Number Present	1,791	2,409	2,622	2,347	2,225	722	7	40
Grand Total	**18,454**	**18,997**	**18,681**	**18,314**	**18,355**	**11,081**	**787**	**114**

Table 6-2e
WorldCat English ISBN

SUBJECT DIVISION	Total
AGRICULTURE	7,958
ANTHROPOLOGY	2,778
ART AND ARCHITECTURE	14,891
BIOLOGICAL SCIENCES	9,472
BUSINESS AND ECONOMICS	32,134
CHEMISTRY	2,264
COMPUTER SCIENCE	6,305
EDUCATION	12,339
ENGINEERING AND TECHNOLOGY	31,037
GEOGRAPHY AND EARTH SCIENCES	8,965
HISTORY AND AUXILIARY SCIENCES	37,628
LANGUAGE, LINGUISTICS, AND LITERATURE	84,559
LAW	12,414
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	7,378
MATHEMATICS	5,401
MEDICINE	21,075
MUSIC	4,301
PERFORMING ARTS	3,842
PHILOSOPHY AND RELIGION	26,997
PHYSICAL EDUCATION AND RECREATION	8,825
PHYSICAL SCIENCES	5,466
POLITICAL SCIENCE	11,104
PSYCHOLOGY	6,161
SOCIOLOCY	18,134
Subject Analysis	**381,428**
No Call Number Present	32,824
Grand Total	**414,252**

Table 6-3a
WorldCat Foreign ISBN

SUBJECT DIVISION	Pre-1500	1600-1699	1700-1799	1800-1849	1850-1899	1900-1909	1910-1919	1920-1929
AGRICULTURE	0	0	0	0	0	2	0	0
ANTHROPOLOGY	0	0	0	1	1	3	0	0
ART AND ARCHITECTURE	0	0	1	0	0	12	0	1
BIOLOGICAL SCIENCES	0	0	0	0	1	3	0	0
BUSINESS AND ECONOMICS	0	0	0	0	1	23	0	1
CHEMISTRY	0	0	0	0	0	2	0	1
COMPUTER SCIENCE	0	0	0	0	0	1	0	0
EDUCATION	0	0	0	0	0	6	0	0
ENGINEERING AND TECHNOLOGY	0	0	0	0	0	17	0	0
GEOGRAPHY AND EARTH SCIENCES	0	0	1	0	0	8	0	0
HISTORY AND AUXILIARY SCIENCES	0	1	0	3	8	52	1	2
LANGUAGE, LINGUISTICS, AND LITERATURE	1	0	0	2	8	78	1	5
LAW	0	0	0	0	0	27	0	0
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	0	0	1	0	0	26	0	0
MATHEMATICS	0	0	1	0	0	3	0	0
MEDICINE	0	0	0	0	0	9	0	0
MUSIC	0	0	0	0	1	5	0	0
PERFORMING ARTS	0	0	0	0	0	2	0	0
PHILOSOPHY AND RELIGION	0	0	0	0	0	25	2	0
PHYSICAL EDUCATION AND RECREATION	0	0	0	0	0	1	0	0
PHYSICAL SCIENCES	0	0	0	0	0	0	0	0
POLITICAL SCIENCE	0	0	0	0	1	6	0	1
PSYCHOLOGY	0	0	0	0	0	1	0	0
SOCIOLOCY	0	0	0	0	0	6	0	0
Subject Analysis	1	1	4	6	21	318	4	11
No Call Number Present	0	2	3	5	9	18	2	8
Grand Total	1	3	7	11	30	336	6	19

Table 6-3b
WorldCat Foreign ISBN

SUBJECT DIVISION	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1984	1985	1986
AGRICULTURE	0	0	0	5	255	323	93	78
ANTHROPOLOGY	2	0	1	4	218	274	81	72
ART AND ARCHITECTURE	3	3	3	19	1,276	1,569	406	421
BIOLOGICAL SCIENCES	0	1	2	5	424	388	95	100
BUSINESS AND ECONOMICS	2	2	1	9	2,341	2,619	625	662
CHEMISTRY	0	0	2	0	86	78	16	15
COMPUTER SCIENCE	0	0	0	1	98	117	37	31
EDUCATION	0	0	0	1	749	650	134	140
ENGINEERING AND TECHNOLOGY	0	1	3	10	1,020	1,063	276	289
GEOGRAPHY AND EARTH SCIENCES	0	0	1	3	353	388	77	105
HISTORY AND AUXILIARY SCIENCES	5	3	10	35	3,116	3,416	880	864
LANGUAGE, LINGUISTICS, AND LITERATURE	12	16	71	175	7,220	6,919	1,671	1,696
LAW	0	0	1	4	844	1,348	370	391
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	2	1	6	9	572	709	155	152
MATHEMATICS	0	0	0	2	219	218	47	55
MEDICINE	1	0	4	4	881	662	149	160
MUSIC	1	0	3	9	411	382	99	99
PERFORMING ARTS	0	0	2	2	228	287	78	80
PHILOSOPHY AND RELIGION	4	1	17	41	2,154	2,055	502	526
PHYSICAL EDUCATION AND RECREATION	0	0	0	1	266	218	41	48
PHYSICAL SCIENCES	2	0	0	3	286	309	65	81
POLITICAL SCIENCE	0	1	2	12	794	862	227	225
PSYCHOLOGY	0	1	1	7	385	347	68	88
SOCIOLOCY	0	0	3	5	1,141	1,173	271	267
Subject Analysis	34	30	133	366	25,337	26,374	6,463	6,645
No Call Number Present	15	14	74	195	4,522	7,274	1,814	2,110
Grand Total	49	44	207	561	29,859	33,648	8,277	8,755

Table 6-3c
WorldCat Foreign ISBN

SUBJECT DIVISION	1987	1988	1989	1990	1991	1992	1993	1994
AGRICULTURE	111	135	123	123	147	150	168	151
ANTHROPOLOGY	81	113	124	140	152	138	156	179
ART AND ARCHITECTURE	457	575	621	713	730	773	779	822
BIOLOGICAL SCIENCES	110	160	123	204	194	183	181	187
BUSINESS AND ECONOMICS	822	1,106	1,155	1,203	1,262	1,393	1,401	1,419
CHEMISTRY	19	35	35	35	30	41	26	16
COMPUTER SCIENCE	44	48	83	119	105	94	113	96
EDUCATION	173	213	217	240	235	257	273	290
ENGINEERING AND TECHNOLOGY	279	409	431	474	507	490	452	481
GEOGRAPHY AND EARTH SCIENCES	95	170	200	200	243	224	245	210
HISTORY AND AUXILIARY SCIENCES	1,070	1,308	1,515	1,672	1,843	2,007	2,049	2,229
LANGUAGE, LINGUISTICS, AND LITERATURE	1,843	2,411	2,750	2,962	3,155	3,257	3,724	4,028
LAW	385	424	549	666	785	747	861	927
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	413	252	268	270	242	296	276	269
MATHEMATICS	71	80	79	94	94	110	102	101
MEDICINE	175	217	216	247	279	286	286	324
MUSIC	104	115	114	126	136	135	124	148
PERFORMING ARTS	82	99	107	131	137	142	162	150
PHILOSOPHY AND RELIGION	561	616	732	784	785	960	1,023	968
PHYSICAL EDUCATION AND RECREATION	74	66	74	88	84	89	102	83
PHYSICAL SCIENCES	96	143	148	166	162	136	150	137
POLITICAL SCIENCE	305	367	380	415	423	488	523	529
PSYCHOLOGY	105	109	153	133	121	144	166	182
SOCIOLOCY	329	377	448	519	561	564	623	615
Subject Analysis	7,804	9,548	10,645	11,724	12,412	13,104	13,965	14,541
No Call Number Present	2,453	2,970	3,581	3,739	3,854	4,027	4,094	4,454
Grand Total	10,257	12,518	14,226	15,463	16,266	17,131	18,059	18,995

Table 6-3d
WorldCat Foreign ISBN

SUBJECT DIVISION	1995	1996	1997	1998	1999	2000	2001	Other
AGRICULTURE	164	150	150	121	92	13	0	4
ANTHROPOLOGY	165	179	198	163	98	21	0	2
ART AND ARCHITECTURE	808	896	919	817	642	130	0	7
BIOLOGICAL SCIENCES	176	158	172	179	102	22	0	1
BUSINESS AND ECONOMICS	1,281	1,210	1,247	1,018	749	131	1	6
CHEMISTRY	33	30	38	28	26	12	0	0
COMPUTER SCIENCE	139	161	176	166	150	34	0	2
EDUCATION	251	255	234	203	163	29	0	0
ENGINEERING AND TECHNOLOGY	457	467	453	422	267	70	1	5
GEOGRAPHY AND EARTH SCIENCES	206	225	240	204	146	62	0	6
HISTORY AND AUXILIARY SCIENCES	2,252	2,104	2,145	1,980	1,481	243	0	17
LANGUAGE, LINGUISTICS, AND LITERATURE	3,878	3,850	3,870	3,704	2,719	534	1	19
LAW	847	809	854	712	554	100	0	3
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	254	295	396	212	147	62	0	6
MATHEMATICS	75	100	80	81	73	13	1	0
MEDICINE	289	298	335	251	211	50	1	1
MUSIC	169	130	145	138	112	20	0	3
PERFORMING ARTS	174	166	168	162	109	19	0	0
PHILOSOPHY AND RELIGION	975	1,006	999	883	695	177	1	9
PHYSICAL EDUCATION AND RECREATION	98	99	113	80	62	16	1	2
PHYSICAL SCIENCES	187	140	152	134	102	37	0	0
POLITICAL SCIENCE	525	473	462	417	343	83	0	1
PSYCHOLOGY	149	165	158	143	132	40	0	0
SOCIOLOCY	611	608	620	558	403	100	0	3
Subject Analysis	14,163	13,974	14,324	12,776	9,578	2,018	7	97
No Call Number Present	4,766	5,205	5,759	5,191	4,120	881	0	81
Grand Total	18,929	19,179	20,083	17,967	13,698	2,899	7	178

Table 6-3e
WorldCat Foreign ISBN

SUBJECT DIVISION	Total
AGRICULTURE	2,558
ANTHROPOLOGY	2,566
ART AND ARCHITECTURE	13,403
BIOLOGICAL SCIENCES	3,171
BUSINESS AND ECONOMICS	21,690
CHEMISTRY	604
COMPUTER SCIENCE	1,815
EDUCATION	4,713
ENGINEERING AND TECHNOLOGY	8,344
GEOGRAPHY AND EARTH SCIENCES	3,612
HISTORY AND AUXILIARY SCIENCES	32,311
LANGUAGE, LINGUISTICS, AND LITERATURE	60,580
LAW	12,208
LIBRARY SCIENCE, GENERALITIES AND REFERENCE	5,291
MATHEMATICS	1,699
MEDICINE	5,336
MUSIC	2,729
PERFORMING ARTS	2,487
PHILOSOPHY AND RELIGION	16,501
PHYSICAL EDUCATION AND RECREATION	1,706
PHYSICAL SCIENCES	2,636
POLITICAL SCIENCE	7,865
PSYCHOLOGY	2,798
SOCIOLOCY	9,805
Subject Analysis	226,428
No Call Number Present	71,240
Grand Total	297,668

Table 6-4a
WorldCat ISBN Analysis By Language

	Subtotal English	Percentage Increase/ Decrease	English ISBN		Subtotal Foreign	Percentage Increase/ Decrease	Foreign ISBN	
			With Call #	Without Call #			With Call #	Without Call #
Pre-1500	0		0	0	1		1	0
1500-1599	2	-200.00%	2	0	0	-100.00%	0	0
1600-1699	18	800.00%	16	2	3	300.00%	1	2
1700-1799	316	1655.56%	312	4	7	133.33%	4	3
1800-1849	1,043	230.06%	1,025	18	11	57.14%	6	5
1850-1899	5,463	423.78%	5,401	62	30	172.73%	21	9
1900-1909	1,172	-78.55%	1,101	71	336	1020.00%	318	18
1910-1919	545	-53.50%	448	97	6	-98.21%	4	2
1920-1929	162	-70.28%	84	78	19	216.67%	11	8
1930-1939	183	12.96%	107	76	49	157.89%	34	15
1940-1949	212	15.85%	143	69	44	-10.20%	30	14
1950-1959	464	118.87%	391	73	207	370.45%	133	74
1960-1969	5,728	1134.48%	5,496	232	561	171.01%	366	195
1970-1979	72,536	1166.34%	69,682	2,854	29,859	5222.46%	25,337	4,522
1980-1984	55,171	-23.94%	51,170	4,001	33,648	12.69%	26,374	7,274
1985	13,333	-75.83%	12,376	957	8,277	-75.40%	6,463	1,814
1986	14,184	6.38%	13,170	1,014	8,755	5.78%	6,645	2,110
1987	14,684	3.53%	13,439	1,245	10,257	17.16%	7,804	2,453
1988	15,363	4.62%	14,161	1,202	12,518	22.04%	9,548	2,970
1989	16,464	7.17%	15,174	1,290	14,226	13.64%	10,645	3,581
1990	17,145	4.14%	15,714	1,431	15,463	8.70%	11,724	3,739
1991	19,035	11.02%	17,580	1,455	16,266	5.19%	12,412	3,854
1992	19,028	-0.04%	17,593	1,435	17,131	5.32%	13,104	4,027
1993	18,585	-2.33%	17,081	1,504	18,059	5.42%	13,965	4,094
1994	18,633	0.26%	17,142	1,491	18,995	5.18%	14,541	4,454
1995	18,454	-0.96%	16,663	1,791	18,929	-0.35%	14,163	4,766
1996	18,997	2.94%	16,588	2,409	19,179	1.32%	13,974	5,205
1997	18,681	-1.66%	16,059	2,622	20,083	4.71%	14,324	5,759
1998	18,314	-1.96%	15,967	2,347	17,967	-10.54%	12,776	5,191
1999	18,355	0.22%	16,130	2,225	13,698	-23.76%	9,578	4,120
2000	11,081	-39.63%	10,359	722	2,899	-78.84%	2,018	881
2001	787	-92.90%	780	7	7	-99.76%	7	0
Other	114		74	40	178		97	81
Total	**414,252**		**381,428**	**32,824**	**297,668**		**226,428**	**71,240**

Table 6-4b
WorldCat ISBN Analysis By Language

	Total WorldCat	Percentage Increase/ Decrease	WorldCat ISBN	
			With Call #	Without Call #
Pre-1500	1		1	0
1500-1599	2	100.00%	2	0
1600-1699	21	950.00%	17	4
1700-1799	323	1438.10%	316	7
1800-1849	1,054	226.32%	1,031	23
1850-1899	5,493	421.16%	5,422	71
1900-1909	1,508	-72.55%	1,419	89
1910-1919	551	-63.46%	452	99
1920-1929	181	-67.15%	95	86
1930-1939	232	28.18%	141	91
1940-1949	256	10.34%	173	83
1950-1959	671	162.11%	524	147
1960-1969	6,289	837.26%	5,862	427
1970-1979	102,395	1528.16%	95,019	7,376
1980-1984	88,819	-13.26%	77,544	11,275
1985	21,610	-75.67%	18,839	2,771
1986	22,939	6.15%	19,815	3,124
1987	24,941	8.73%	21,243	3,698
1988	27,881	11.79%	23,709	4,172
1989	30,690	10.07%	25,819	4,871
1990	32,608	6.25%	27,438	5,170
1991	35,301	8.26%	29,992	5,309
1992	36,159	2.43%	30,697	5,462
1993	36,644	1.34%	31,046	5,598
1994	37,628	2.69%	31,683	5,945
1995	37,383	-0.65%	30,826	6,557
1996	38,176	2.12%	30,562	7,614
1997	38,764	1.54%	30,383	8,381
1998	36,281	-6.41%	28,743	7,538
1999	32,053	-11.65%	25,708	6,345
2000	13,980	-56.38%	12,377	1,603
2001	794	-94.32%	787	7
Other	292		171	121
Total	**711,920**		**607,856**	**104,064**

Table 6-6
WorldCat ISBN Tables
Records With Call Number Present by Subject

	English ISBN	Foreign ISBN	WorldCat Total ISBN
	With Call #	With Call #	With Call #
AGRICULTURE	7,958	2,558	10,516
ANTHROPOLOGY	2,778	2,566	5,344
ART AND ARCHITECTURE	14,891	13,403	28,294
BIOLOGICAL SCIENCES	9,472	3,171	12,643
BUSINESS AND ECONOMICS	32,134	21,690	53,824
CHEMISTRY	2,264	604	2,868
COMPUTER SCIENCE	6,305	1,815	8,120
EDUCATION	12,339	4,713	17,052
ENGINEERING AND TECHNOLOGY	31,037	8,344	39,381
GEOGRAPHY AND EARTH SCIENCES	8,965	3,612	12,577
HISTORY AND AUXILIARY SCIENCES	37,628	32,311	69,939
LANGUAGE, LINGUISTICS, AND LITERATURE	84,559	60,580	145,139
LAW	12,414	12,208	24,622
LIBRARY SCIENCE, GENERALITIES AND REFERENCES	7,378	5,291	12,669
MATHEMATICS	5,401	1,699	7,100
MEDICINE	21,075	5,336	26,411
MUSIC	4,301	2,729	7,030
PERFORMING ARTS	3,842	2,487	6,329
PHILOSOPHY AND RELIGION	26,997	16,501	43,498
PHYSICAL EDUCATION AND RECREATION	8,825	1,706	10,531
PHYSICAL SCIENCES	5,466	2,636	8,102
POLITICAL SCIENCE	11,104	7,865	18,969
PSYCHOLOGY	6,161	2,798	8,959
SOCIOLOGY	18,134	9,805	27,939
TOTAL	381,428	226,428	607,856

Chapter Seven

Summary and Conclusions

In February 2001, the Boston Public Library entered the 46 millionth record into WorldCat.[1] In that same month the 10% systematic random sample of monographic bibliographic records for this study was extracted from WorldCat. This study presents a Asnapshot in time@ of the WorldCat database as it existed in February 2001. The study is Ahindsight@ or *ex post facto*, in that we don=t know what the profile looked like in earlier time frames. A record is added to WorldCat every 15 minutes, so that literally, a database scan is only accurate for a few minutes. We are viewing the Auniverse of publication@ as it was reflected in WorldCat early in 2001.

The tables in this study were all constructed for the report from the iCAS analysis of the 10% systematic random sample of monographic bibliographic records from WorldCat extracted in February, 2001. The subject divisions, categories, and most detailed subject descriptors follow the construction of the iCAS product that is based upon the WLN Conspectus subject breakdowns. The unique title and title overlap, language, and the adult/juvenile audience analyses follow the iCAS product templates. The ISBN analysis is not a regular feature of iCAS, but was run especially for this study. The major findings from the study are summarized below by chapter.

Findings

<u>WorldCat</u> (Chapter Two) http://www.oclc.org/dummyaddress.html

$ **The 10% systematic random sample of monographic bibliographic records in WorldCat contained 3,378,272 usable records.**

---

[1]. "The Year in Review, July 1-June 30, 2001," *OCLC Annual Report, 2000/2001*, p.6. (The 46 million is not the actual total of records that is less, at 45.1 million when duplicate records are removed.)

- **2,199,165 of those records had call numbers and could be analyzed by subject.**
- **An additional 1,179,107 records did not have call numbers (shown as ANo call number present@ in the tables).**
- **A pattern of decline in the number of records annually beginning in the early 1990s is evident The number of records with call numbers declines 72 percentage points between 1992 and 2000, while the number of records without call numbers declines 88 percentage points in the same time frame.**
- **Subject analysis on the set of records with call numbers shows that the language, linguistics, literature division has the largest number of records comprising 21% of all records with call numbers. The history division is second with 13.3%. The business and economics division comprises 9.6%; philosophy/religion has 8.2% and engineering/technology comprises 6.8% of all records with call numbers.**

<u>Library Groupings</u> (Chapter Three) http://www.oclc.org/dummyaddress.html

- **The research libraries grouping has the largest number of records with call numbers in the dataset, 1,745,034.**
- **The profile of the research libraries by imprint year and subject divisions closely parallels the profile of WorldCat.**
- **The decline in the number of records annually in the 1990s is most acute in the research libraries aggregated collections.**
- **The total number of records with call numbers for the academic libraries grouping is 1,323,165.**
- **The patterns of increase/decrease over time in records for the academic libraries are similar to those of the research libraries.**
- **The major differences between the research libraries and the academic libraries are in the records without call numbers.**

- **The concentration of records by subject division in the academic libraries is the same for LLL and history, but the third and fourth ranked subject divisions are reversed in the academic library collections from the research library collections. In academic libraries philosophy/religion is ranked third, but it is fourth in the research libraries. In the academic libraries business/economics is fourth, whereas it is third in the research libraries.**
- **On a percentage basis, the decline in the number of records annually for the academic libraries is less severe than in the research libraries, but follows the same pattern.**
- **In the special/other libraries grouping, the pattern of decline in the number of records does not begin until 1993 and it is less severe than in the research and academic libraries.**
- **The subject concentrations in the special libraries are not as similar to the research and academic libraries by percentage of total for subject divisions. There are differences that can easily be attributed to the nature of the professional and discipline emphases in the special libraries. The large number of corporate, legal, medical, and technical libraries account for higher percentages of total in business, legal, and scientific subject divisions.**
- **The bibliographic records for the public and school library groupings were analyzed by adult and juvenile audience levels, a feature of the iCAS analysis. The public libraries, with adult collections comprising over 90% of total, have a ratio of 21 to one, adult to juvenile. For school libraries the ratio of adult to juvenile is 5 to one.**
- **In the school library grouping the percentage of juvenile titles steadily increases from 1985 to present, the only measure to show a pattern of increase in the number of records in the latter 1990s.**

<u>Unique Records</u> (Chapter Four) http://www.oclc.org/dummyaddress.html

- **The data analyses on number of unique records and title overlap demonstrate that the universe of materials under bibliographic control in WorldCat show a high level of diversity of resources with 53% of records in the analysis having only one library location symbol.**
- **The unique records in WorldCat are almost evenly divided between records with call numbers and records without call numbers.**
- **Both WorldCat and the research libraries have close to the same absolute numbers in the datasets for unique records, both those with call numbers and without call numbers.**
- **The research libraries have 63.51% of total records as unique records.**
- **The research libraries grouping contributes 62% of the unique records in the sample database, while the other four library groupings contribute the remaining 38 percent.**
- **The other four types of library groupings have low percentages of unique titles and higher overlap among the four.**
- **The academic library records naturally have a high overlap with the research libraries.**
- **The core set of records (101,378) shared by all of the library groupings (overlap) represents only 3% of the total records in the sample. These represent 5% of the records used in the subject analysis of which the 101,378 records is a subset.**
- **The number of records in common among the five types of library groupings (101,378) is governed by the school libraries grouping in that it has the smallest number of total records and thus the overlap cannot exceed the number of records the school libraries have in common with the other four groupings.**

<u>Language analysis</u> (Chapter Five) http://www.oclc.org/dummyaddress.html

- $ **Of all records in the study, approximately 65% are for English language monographs.**
- $ **Of the total number of records with call numbers, 63.6% are for English language materials and 36.3% are for non-English titles.**
- $ **The six foreign language groupings analyzed in the study B Chinese, French, German, Japanese, Russian, and Spanish B account for 67.5% of all foreign language records in the sample.**
- $ **In the analysis by subject divisions, the language, linguistics, and literature division has the largest number of records in all language groupings, followed by the history division. After these two subject divisions, the ranking of subjects varies by language grouping.**

<u>ISBN Analysis</u> (Chapter Six) http://www.oclc.org/dummyaddress.html

- $ **Only 27% of all records with call numbers in the sample have ISBN numbers.**
- $ **Fifty-seven percent of all records with call numbers in the 1970-1999 time frame have ISBN numbers.**
- $ **Less than 10% of all records without call numbers in WorldCat have ISBN numbers.**
- $ **For English language records, 92% of the records with ISBN numbers also have call numbers.**
- $ **For foreign language records, 72% with ISBN numbers also have call numbers.**
- $ **The number of foreign language records with ISBN numbers but without call numbers (71,240) is larger than the number of English language ISBN records without call numbers (32,824).**
- $ **Sixty-eight percent of ISBN records without call numbers are foreign language**

records.

- $ **It appears that titles from mainstream publishers are being consistently added to WorldCat within a current time frame, and the decreases in numbers of records in the latter 1990s are not in the English language titles with ISBN numbers.**
- $ **The decline in the foreign language ISBN records is responsible for the overall decrease in records with ISBN numbers in WorldCat, as the English language records have slight decreases in 1997 and 1998, with an actual 2% gain in 1999.**
- $ **Records with ISBN numbers have a lower rate of decrease than records without ISBN numbers and the rate of decrease in the late 1990s is lower for those records with ISBN numbers. This supports the speculation that the ANo call number present@ records are largely comprised of foreign language titles and titles which are not from mainstream U.S. and European publishers.**
- $ **In subject areas in which research libraries collect globally, the differential between the English language and foreign language totals reflects the foreign language collecting in those libraries. Both LLL and the history division have a considerable number of foreign language records with ISBN numbers.**

One of the most troubling of the findings is that of the decline in the number of records in the most current years of the analysis. This aspect of the findings is examined more closely in the next section.

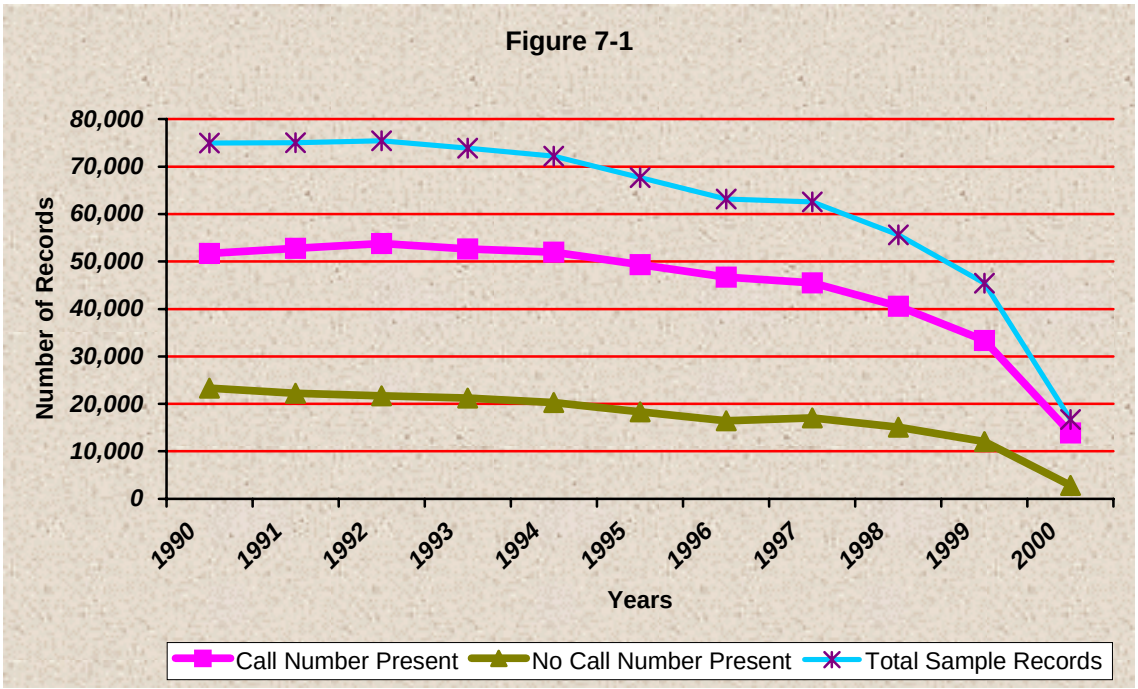
Decline in the Number of Records for Current Years

The findings by several measures show a decline in the number of records for the most current imprint years in the study. Beginning before 1995, in most instances, the absolute number of records decreases steadily and annually for

unique records, language records and certain categories of ISBN records. Tables 7-1 through 7-6 trace the decline by these different measures for the imprint years 1990-2000. Figures 7-1 through 7-6 are derived from these tables.

In Table 7-1 and Figure 7-1 the *Annual Increase/Decrease in Number of WorldCat Records 1990-2000* is shown. The rate of change is shown for both those records with call numbers and those records without call numbers and the two combined as total records. It can easily be seen that the number of records without call numbers is considerably less than the number of records with call numbers. But the decrease is much higher for the records without call numbers, causing the overall decrease to be higher than the rate of change for the records with call numbers.

Figure 7-1



The decline also begins to occur earlier for the records without call numbers. The peak year for total records and those with call numbers is 1992, after which the rate of change is negative. Oddly, there is a reversal in 1997 of the decline in records without call numbers, but the pattern reverts back to a decrease in the following year. The absolute number difference between 1997 and 1996 for the no call number records is only 600 records. In that same year, 1997, the records with call numbers also have less of a decrease than the preceding year. The total number of records is nearly the same from 1996 to 1997. We can only speculate as to the reason(s) for 1997 interrupting the downward pattern. 1997 is four years removed from the year of extraction of the sample. It is possible that it represents the point at which acquisitions and cataloging lag actually commence, whereas those years prior to 1997, the numbers of records are a closer reflection of true collecting patterns.

The downward trend is most severe for all three categories beginning in 1995. While the records without call numbers have the steepest percentage decline, as seen in Table 7-1, the lower number of those records shows as a flatter curve in Figure 7-1. It will be seen in the remaining tables and graphs that the other measures show similar downward trends. As Table 7-1 and Figure 7-1 show total records for WorldCat, the other measures have to follow the overall trend.

Table 7-2 and Figure 7-2, *Annual Increase/Decrease in Number of Records 1990-2000 Academic and Research Libraries*, show the totals for the two largest library groupings. From Table 7-2 and figure 7-2 it can be seen that the decline in records for both the academic and research libraries groupings follows the same pattern as the WorldCat dataset. The pattern of increase/decrease annually is nearly the same for both of these library groupings. The only difference is in the number of records for the two groups, with the research libraries having the higher total and the academic libraries following the same pattern and keeping the same differential between totals for the time period.

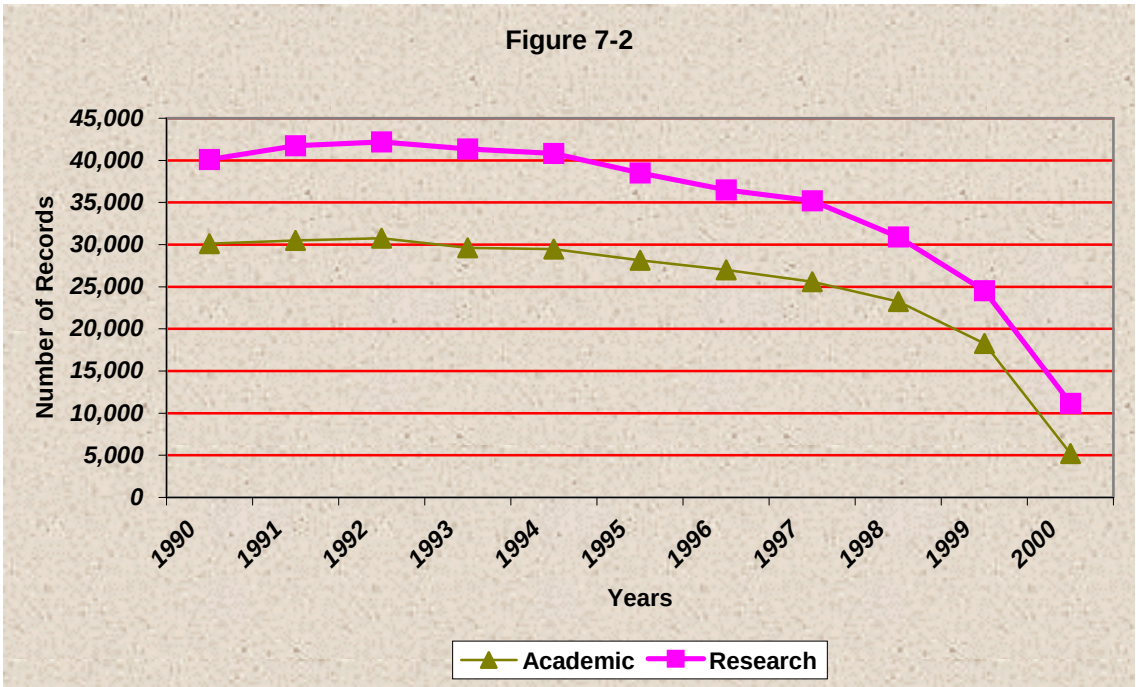
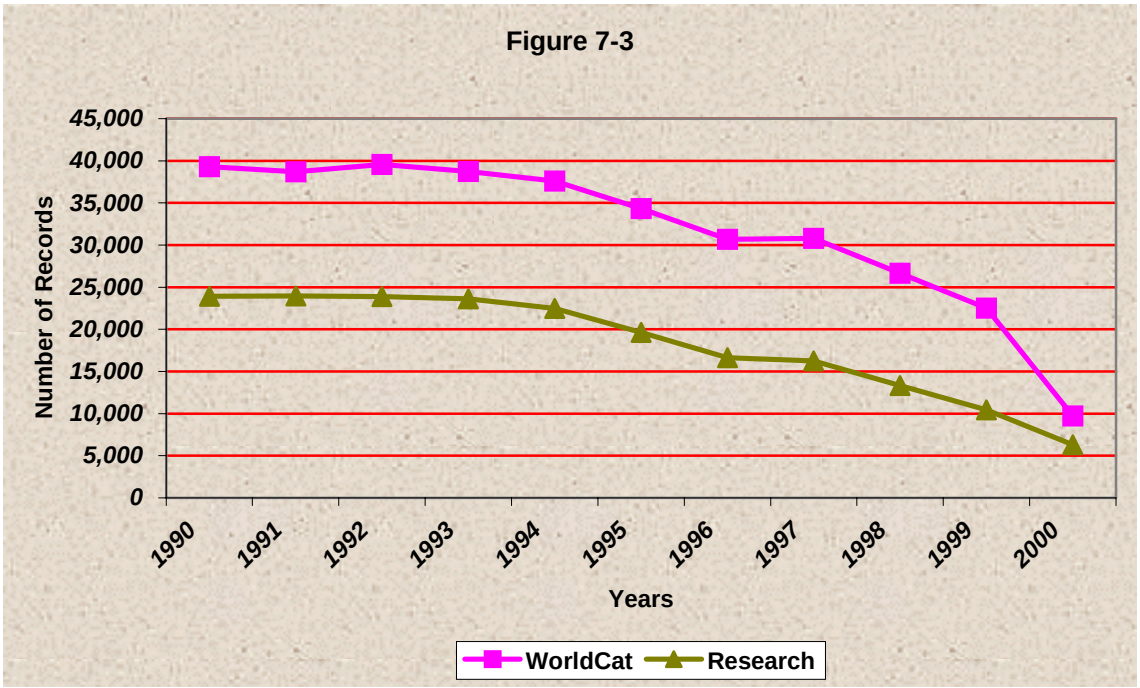


Table 7-3 and Figure 7-3 *WorldCat and Research Libraries Unique Records 1990-2000* show the pattern of decline in the number of unique records for both WorldCat and the research libraries for the 1990-2000 imprint years. For the majority of the years, the percentage decline in WorldCat is higher than the decline in the research libraries' records. Again, 1997 stands out in the unique records measure as an anomaly in the pattern for the decade. The number of unique records in WorldCat remains static from 1996-1997, but there is a slight decrease in the number of unique records in the research libraries. Both have considerable drops in the number of records from 1995 to 1996 and the decline continues in 1998. The other types of library groupings are not similarly analyzed here as the unique records analysis was performed only on WorldCat and the research libraries.

Figure 7-3



The series of tables 7-4a, 7-4b, and 7-4c and accompanying figures show the number of records from 1990-2000 imprint years for English language records and non-English language records. These are from the dataset of those records with call numbers used for the language analyses. In 7-4a and Figure 7-4a, WorldCat *English and Non-English Language Records 1990-2000*, it can be seen that the decline in the number of English language records and the total number of records begins in 1993, whereas the decline in non-English records commences in 1995. It is evident in table 7-4a that the decline is much steeper in the non-English language records, again a factor of the overall lower absolute numbers of those records.

Figure 7-4a

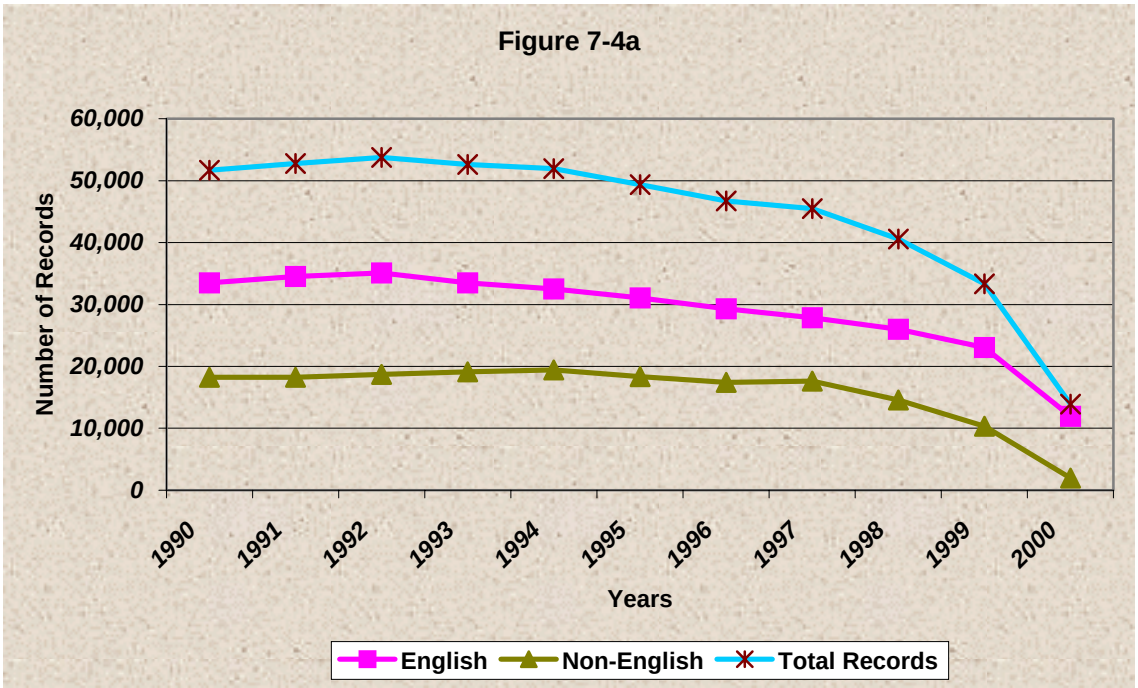


Table 7-4b and Figure 7-4b, *Research Libraries English and Non-English Language Records 1990-2000*, show the decline in English and non-English records for the research libraries. It is more obvious from figure 7-4b that the decline in non-English language records is much steeper than in English language records. For the years 1996 and 1997, the absolute numbers of records in the two language categories draw closer together. From 1995-1998, the number of English language records maintains a steady and modest decline around 5 percent. The number of non-English records plunges, dragging down the total number of records after 1997.

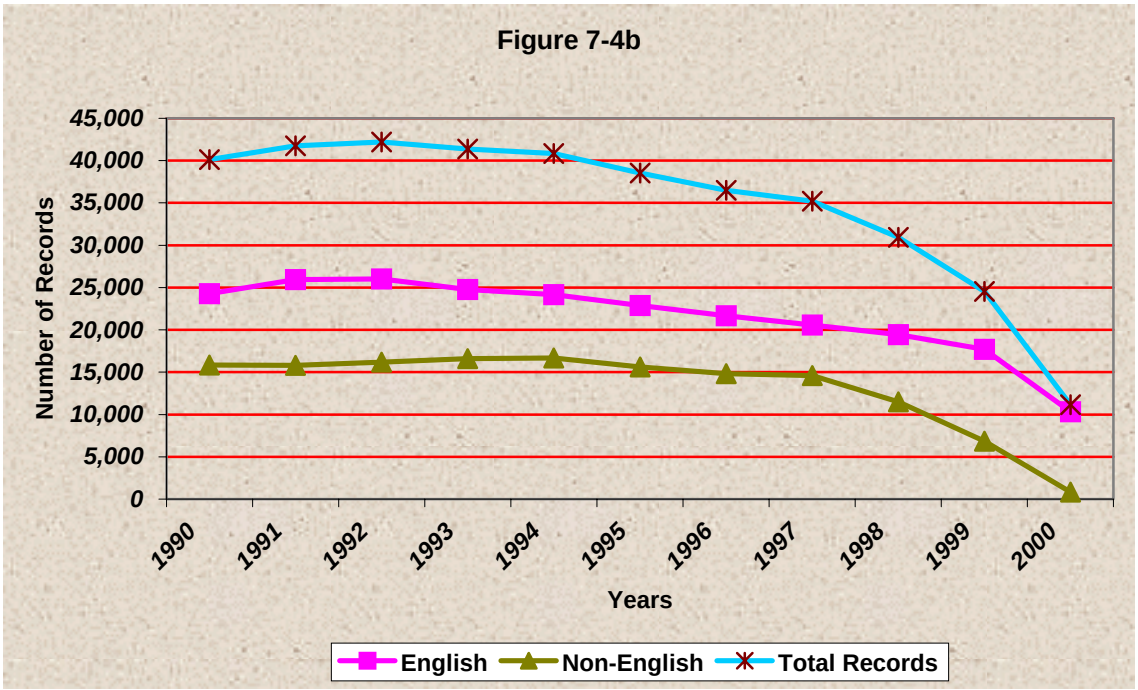
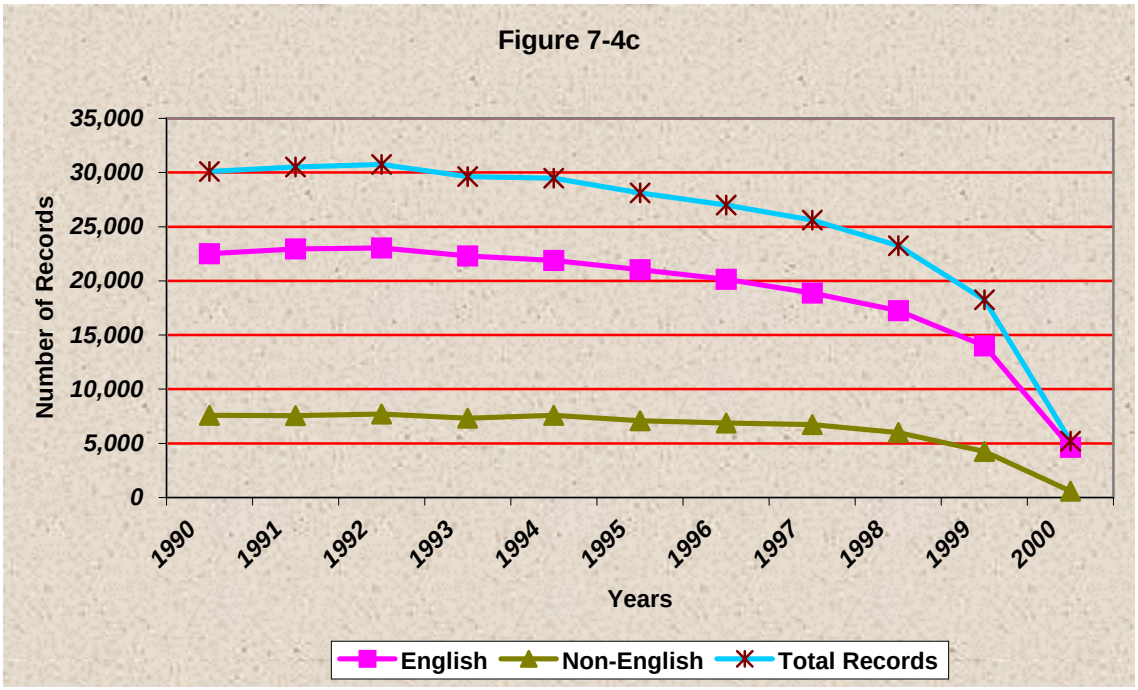


Table 7-4c and Figure 7-4c, *Academic Libraries English and Non-English Records 1990-2000* show the data for the same measures for the academic libraries. Figure 7-4c makes it obvious how low a number of non-English language records are owned by the academic libraries, and this grouping includes a considerable number of non-U.S. libraries. The trend line of non-English records is relatively flat until 1998. The international academic library membership of OCLC may be responsible for a relatively constant rate of acquisition and cataloging in foreign language records. Because of the predominance of records for English language publications in the academic libraries, the trend lines for the total and English language records have an identical pattern. The scale for Figure 7-4c is 10,000 records lower than in 7-4b. It was observed in chapter three that there is a differential of approximately 10,000 records between the research libraries and the academic libraries groupings in total records.



The last set of tables and figures depict the ISBN analysis from chapter seven. Table 7-5 and Figure 7-5 *WorldCat Records with ISBN Numbers*, shows the three ISBN analyses that were run for WorldCat. In the ISBN analyses, the downward trend seen in the other measures commences only a few years from the time of data extraction in 2001. The number of records with ISBN numbers and call numbers peak in 1996; the number of ISBN records lacking call numbers peak in 1997. The number of records with ISBN numbers but without call numbers is very low in comparison to those with call numbers which set the pattern for the total. The decline in these measures is much less severe than in the other measures. This stands to reason in that imprints with ISBN numbers are typically from mainstream publishers and large conglomerates internationally, and these would be titles which are added by libraries on a more current basis than more unique and esoteric titles which do not as readily come to attention and are not cataloged as close to the time of acquisition.

Figure 7-5

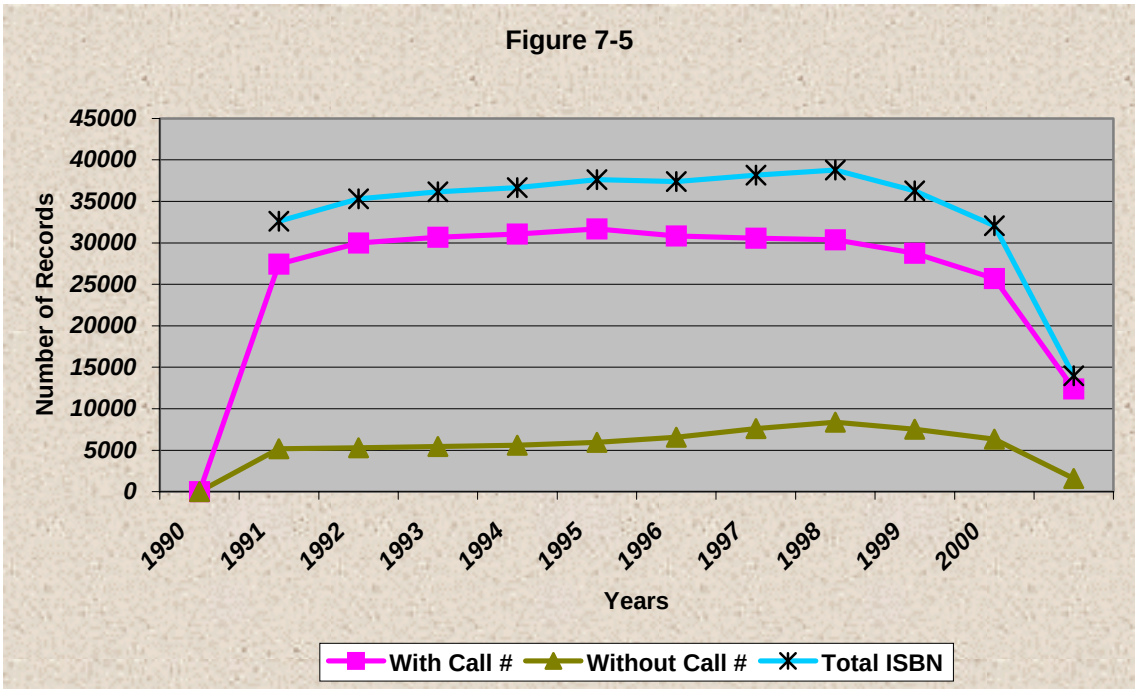
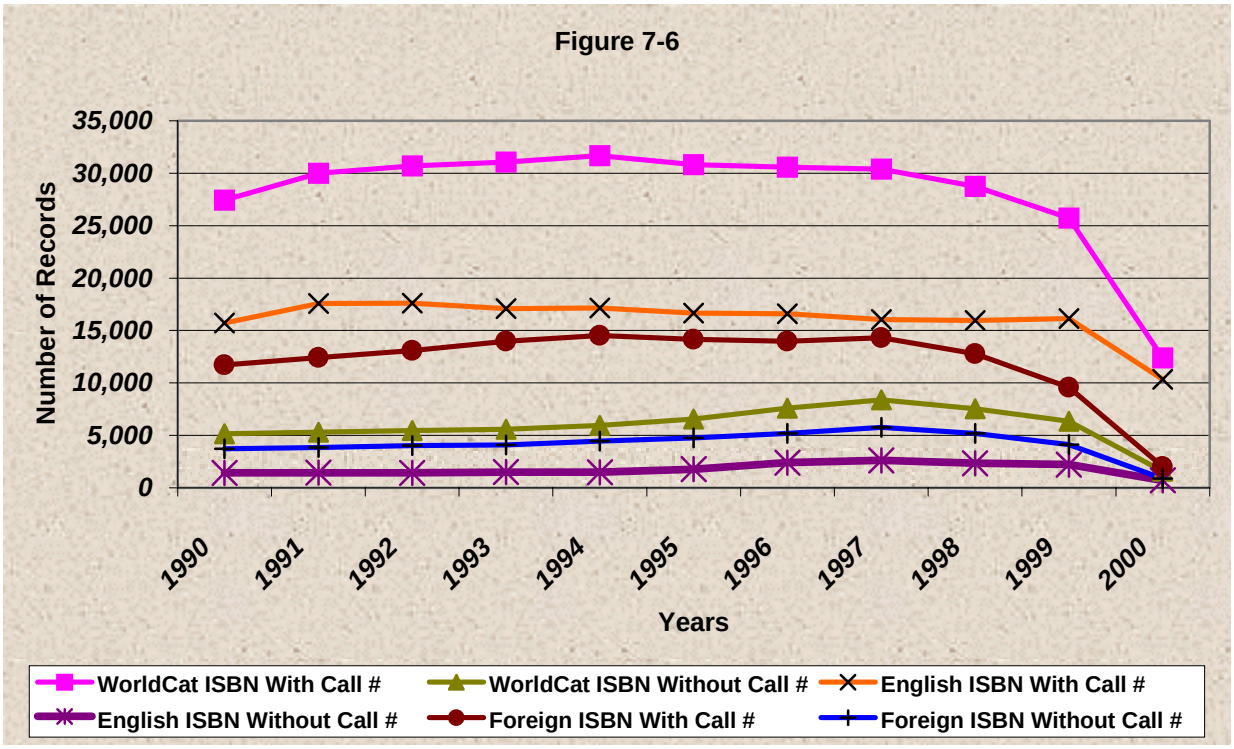


Table 7-6 and Figure 7-6 *WorldCat, English, and Foreign Language Records with ISBN Numbers*, show the ISBN analysis for WorldCat by both call number presence and language categories for the decade of the 1990s. The category with the lowest number of records with ISBN numbers is that of English language records without call numbers. As English language publications are those most likely to have ISBN numbers and also most likely to be owned by U.S. and libraries in other English speaking countries, it follows that most English language titles with ISBN numbers would most likely be cataloged and have call numbers.



The next to lowest category in the ISBN analysis is that of foreign (non-English) titles without call numbers. From the two lowest categories it can be seen that records without call numbers comprise the lowest number of records with ISBN numbers. This points to the type of materials which are not from mainstream publishers and tend to be local history and other titles of local interest, juvenile materials not classified, theses and dissertations, and reports of all kinds. The third line from the bottom in Figure 7-6 is the summation or total of the two bottom lines, those records in WorldCat with an ISBN number but not a call number.

The top three trend lines in Figure 7-6 are for records with both ISBN numbers and call numbers. The lowest of the three is foreign language records with call numbers, which are less numerous than English language records with call numbers. It can be seen that these two categories do not have exactly the same pattern for the imprint years shown. The English language records with call numbers begin to trend downward in 1992, but the decline is slight and the number or records in 1990 and 1998 are very close to the same number. The number of English language records with ISBN numbers and call numbers is actually higher in 1999 than in 1998. Whereas the foreign language records trend upward and remain in the same range until after 1997 but then show the same steep decline as other measures have shown. From Table 7-6 and Figure 7-6 it can be seen that English language records with both call numbers and ISBN numbers are added to the database in a more timely manner than the other categories of records analyzed. The decline is severe in foreign language records with call number and all records without call numbers. What this means is that mainstream publications from English speaking countries are the only category of materials added to the database in a timely manner. Those materials that are more unique and contribute to variety within the resources base are added in a slower manner over a longer period of time.

The implications of these findings and further interpretation are addressed in the next section.

Implications

The findings of this report have direct implications for resources sharing and coordinated cooperative collection development, both of which depend upon the availability of database bibliographic records. The profile of WorldCat revealed in the data analyses in this report provides a basis for discussions on the role of WorldCat in these activities.

OCLC WorldCat has facilitated resources sharing in a number of ways, first and foremost by developing a database of bibliographic records for shared cataloging with holdings information which forms the basis for resources sharing. The cataloging records are augmented by union lists of serials and newspapers with detailed holding information for volumes and issues. The interlibrary loan module of WorldCat revolutionized the procedural nature of resources sharing which went from a system of detective work and guessing to a certainty of which libraries hold a desired title.

The effect of WorldCat upon coordinated cooperative collection development has not been of the same magnitude as the interlibrary loan system upon resources sharing. Now it is more a matter of what capabilities of the database and existing products can be taken advantage of for cooperative collection development activities. ACAS (Automated Collection Assessment and Analysis Service) is a set of collection analysis and assessments tools developed by WLN to assist libraries and consortia in analyzing collections for resources sharing and CCCD. The iCAS product from which the data analysis in this study was produced is one module of ACAS. ICAS has been used in consortia in the United States since its inception and it is now being utilized internationally. As an example, CURL, the Consortium of University Research Libraries in the British Isles, began a collection analysis project using the iCAS in June 2001.[2] The ACAS collection analysis products can assist in

[2] *OCLC Annual Report, 2000/2001*, p.15.

formulating cooperative collection development arrangements and in providing the data necessary for the implementation and functioning of such agreements. Both the ACAS products and the WorldCat database have the potential to augment CCCD activities and to play a larger role in the measurement of cooperative activity.

The extent to which the bibliographic records in WorldCat reflect the global universe of publication has a bearing on the support for resources sharing and CCCD. In considering the title of this report, *Global Collective Resources*, to what extent does WorldCat reflect the global universe of publication? While this question cannot be completely answered with the data in this study, it is easy to see that the profile of WorldCat reflects the collections of the libraries that contribute the records to the database. The farther back we go in time, it would seem, the more accurately WorldCat reflects the universe of publication as libraries have contributed records for retrospective materials. As we have seen from the previous section examining the annual decline in records in the last decade of the 20th century, certainly for that time frame the universe of publication is not yet represented in WorldCat.

What the findings for the 1990s do show is that as the international membership of OCLC continues to grow, the database will come ever closer to reflecting the universe of publication worldwide. The addition of many foreign language records in the 1990s are the result of the addition of bibliographic records from research and academic libraries internationally, which joined WorldCat.

Another aspect the data from the sample show is that the variety of resources in WorldCat is due, in large part, to the unique records for titles which are not from Amainstream@ publishers. While the records for these titles may present some problems in analysis, i.e. the large number of records without call numbers that cannot be analyzed by subject, these records are present in the database and identify items, which may not be identified through any other venue, or at least any other electronic database. As with the study sample, and indeed as the study sample analysis has shown, the most unique titles appear to be added to the WorldCat

database much later and over an extended period of time as compared to current mainstream publications with ISBN numbers, both English and foreign language titles. Many of these unique titles are only Adiscovered@ then acquired and cataloged over an extended period of time. In order to be most effective for acquisitions and resources sharing, bibliographic records for current, new publications need to be added to WorldCat within the time frame of publication. The lower number of records in the study sample at the time of data extraction shows what was available through the database at the time of data extraction. For purposes of any current activity, it does not matter that records will be added later and they do, indeed enrich the database. It is doubtful that most CCCD agreements will focus just upon the most common titles with ISBN numbers which are the predominant type of records appearing within the most current years of imprint records.

Agreements were made with international book vendors in the1990s to contribute their bibliographic records to WorldCat. These book vendor records would not have been included in the study sample because there are no library holding symbols attached to those records. And at least some of them are brief cataloging and many are pre-publication cataloging. These records do provide bibliographic information for the most current publications available through those vendors, including many foreign language titles. Although these records are mostly for mainstream publications with ISBN numbers and do not assist in resources sharing, they are useful information for acquisitions and CCCD. And these records contribute to the universe of publication visible through WorldCat.

One of the steps taken in the OCLC effort to Aglobalize the cooperative,@ was the establishment of a jointly owned organization, Pica B.V. to provide cataloging, interlibrary loan, local system and end-user services to libraries in the Netherlands, France and Germany. Pica B.V. is developing a European Central Catalog that will contain bibliographic and holdings information from many European academic libraries and other important libraries in Europe. What is not reflected in the sample

for this study is the large number of European libraries from Germany, France, and the Netherlands in Pica B.V., which are now considered to be members of OCLC, but have their records in the Pica B.V. database that is separate from WorldCat.[3] At some point in the future these two databases will most probably be searchable as one and will truly reflect the universe of publication for North America and Western Europe.

WorldCat does represent an international universe of publications under bibliographic control with records for publications that are available for resources sharing and assistance in cooperative collection development activities. It seems certain that WorldCat will become more and more representative of the worldwide universe of publication as time goes on. It is doubtful, however, with a critical mass of over 45,000,000 records that the profile of WorldCat by imprint year and subject analysis for printed materials will change substantially for the foreseeable future. The snapshot in time, which is analyzed in this study, will form a picture of the database that will only gradually change, with the exception of the last decade of records.

These implications and questions can be explored in further research.

Further Research

The profile of WorldCat described in this report is the most detailed study to date of the monographic bibliographic records in WorldCat by multiple parameters of imprint date, subject, language, unique titles/overlap, and ISBN data. And yet, in some respects, the analysis is still very broad. The data analyses from the 10% random sample have not yet been Amined out.@ These suggestions are made for further research.

[3] *OCLC Annual Report, 1999/2000,* p.30.

- **Subject analysis on the 10% systematic random sample thus far has been only on the iCAS broad 24 subject divisions. More detailed subject analysis should be conducted to look for more specific collecting patterns within the 24 subject divisions.**
- **Further analysis of the records without call numbers in the sample might assist in formulating cooperative collection development agreements.**
- **Further research is needed into using key word and subject headings to perform subject analysis instead of using only call numbers.**
- **Further exploration of the differences in profile between the set of U.S. libraries in the OCLC/CACD and the WorldCat sample which represents the entire OCLC international membership of contributing libraries may show the variety of resources reflected in WorldCat not owned by U.S. libraries.**
- **Comparisons could also be performed between the WorldCat analyses and those of large consortia using the iCAS product such as CURL in Great Britain, and several state consortia in the United States.**
- **An interesting study would be a comparison of the profiles of WorldCat and the Pica.B.V. European Central Catalog.**

Conclusion

The profile formed by the findings of this study captures the nature of the bibliographic contents of WorldCat at the end of the 20th century when the printed book was the predominant vehicle for the dissemination of knowledge and the preferred format for educational and recreational reading. The point of critical mass seems to have been reached in the first 30 years of WorldCat=s existence, in that from the beginnings with a few bibliographic records from Ohio libraries, it has

grown to over 45,000,000 million records. In June 2001, the total number of records for books was at 37,787,280 records, comprising 84% of records in the WorldCat database.[4] Additions at this point will be for new publications and Afilling in@ with retrospective and foreign language records still not in the database. Undoubtedly, the profile will slowly change over time reflecting the broadening global membership of OCLC. There almost certainly will be an increase in records for foreign language materials. Libraries will continue to add records for retrospective materials and records for formats other than for printed monographs -- for web sites and electronic publications.

As the hegemony of the printed book wanes in the electronic environment, the 20th century will probably be regarded as the apex of print. As the WorldCat database continues to grow in retrospective records, it will increasingly reflect the profile of the universe of publication for the history of printed works. But the question remains as to the efficacy of the database in studying current library acquisitions patterns and as a basis for acquisitions and coordinated cooperative collection development. These concerns can only be addressed by the membership of OCLC.

[4] *OCLC Annual Report, 2000/2001*, p.18.