

SUNY Business Intelligence Initiative (SBII) Library Data Project

by Maureen Zajkowski (OLIS)



Cover Story

[Library Resources via Angel \(Fredonia\)](#)

Features

[Library Resources via Angel \(Onondaga\)](#)

[OLIS Library Resources Starting Point](#)

[Welcome Chancellor Zimpher](#)

[Brockport and Aquabrowser](#)

[SUNY Business Intelligence](#)

[Focus on SUNY Libraries](#)

[SUNYConnect Updates](#)

[How to Contact Us](#)

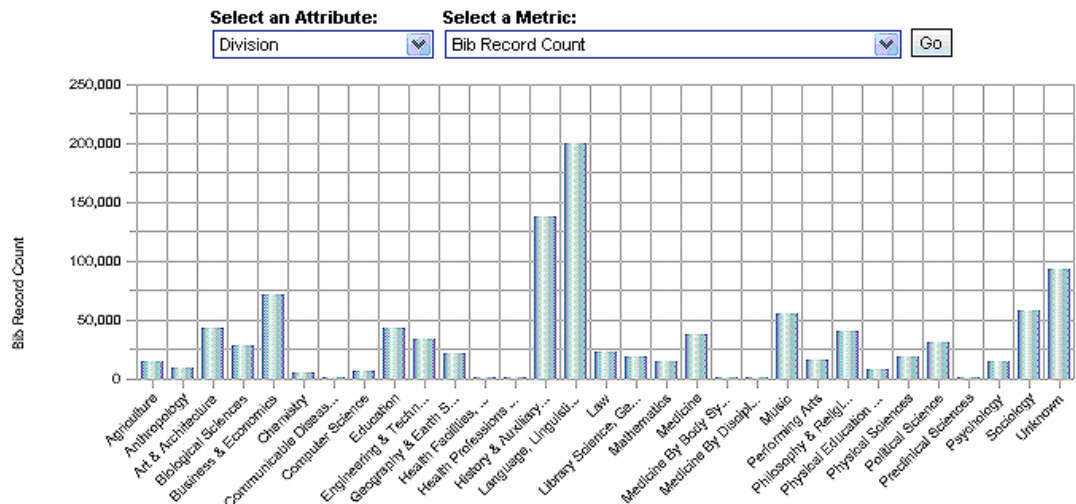
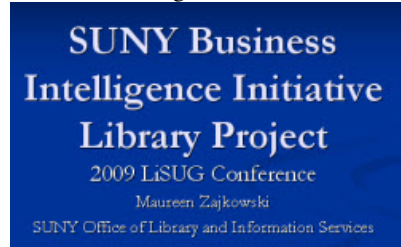
[Linkable Links](#)

The SUNY Business Intelligence Initiative (SBII) is a SUNY-wide effort to provide data analysis and reporting tools for a wide range of university information. SBII is using the Oracle Business Intelligence Suite Enterprise Edition (OBIEE) software tools.

The Library aspect of the SBII project is to provide for a SUNY-wide collection analysis reporting tool in support of program review and resource sharing. The Proof of Concept (POC) phase includes bibliographic and circulation data from a subset of SUNY campuses.

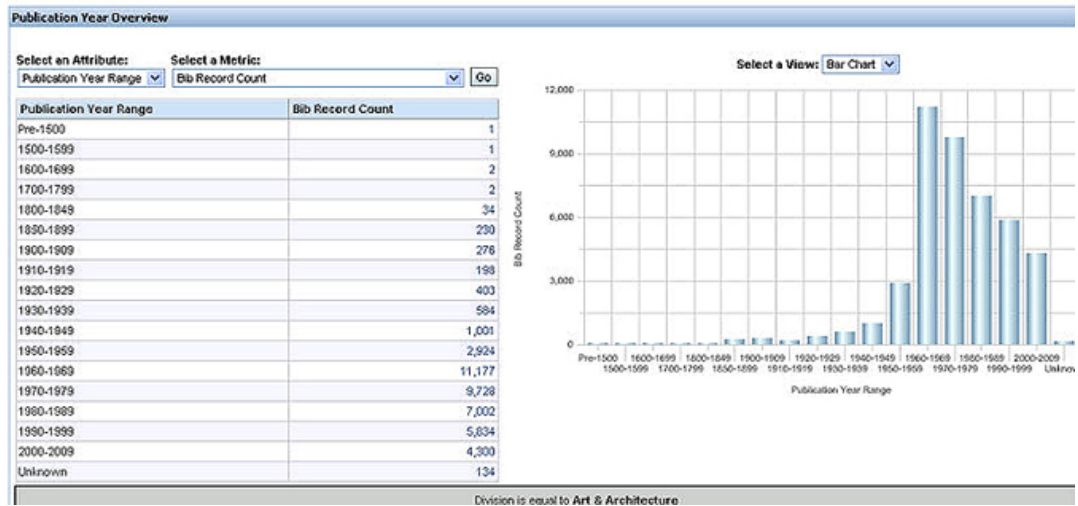
Goals of the POC:

- Analyze and compare aggregate holdings
 - SUNY-wide
 - by SUNY sectors, by regions, and campus defined peer groups
- Compare data between/among selected campuses and sectors/regions
- Access collection usage using circulation information
- Analyze circulation activity by patron status
- Analyze collections, circulation activity and patron status usage by subject, publication date, language, format, other qualifiers (secondary formats, material description, target audience, form of item, government publication)
- Determine collection age and rate of growth
- View overlaps and gaps of the SUNY library collection
- Identify unique campus holdings



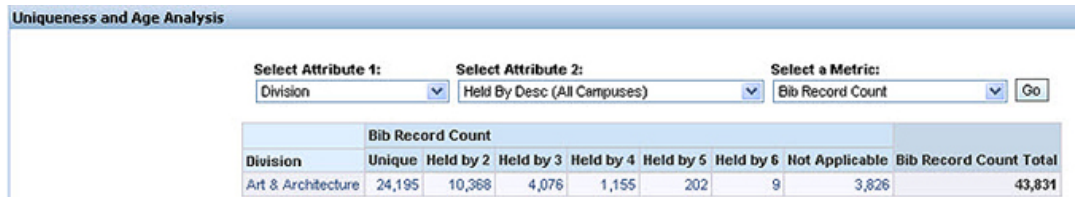
Activities

Over the spring and summer of 2009, the OLIS staff worked closely with staff from the System Administration Office of Information and Technology to analyze data and build the system schemas that would become the basis for the library dataset. Training was provided to the OLIS on developing the web-based “Dashboards” which provide library users the ability to access and manipulate the library collections and circulation data. During July and August 2009, a consultant worked with the OLIS staff to develop the Library Dashboards. Four Dashboards were created. Each Dashboard is made up of several “Pages.” The Pages provide the ability to focus on specific aspects of the library data analysis.



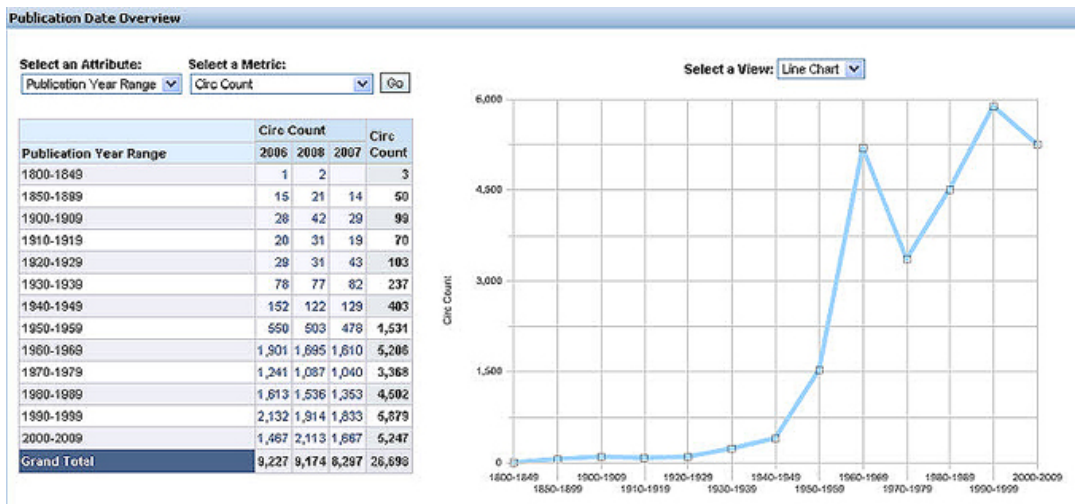
Collection Analysis Dashboard

- Seven Dashboard Pages were created to help analyze and compare bibliographic holdings SUNY-wide, by sectors and regions, and within selected campuses. The analysis includes identification of unique holdings and overlap of holdings among campuses; review of collections' ages and rate of growth based on publication dates; and the ability to filter by a variety of attributes including subject areas, language, and format.



Circulation Analysis Dashboard

- Seven Dashboard Pages were created to help analyze and compare circulation activity by most of the attributes identified in the Collection Analysis Dashboard. In addition, circulation activity can be analyzed by year of circulation, patron type and by lending policy.



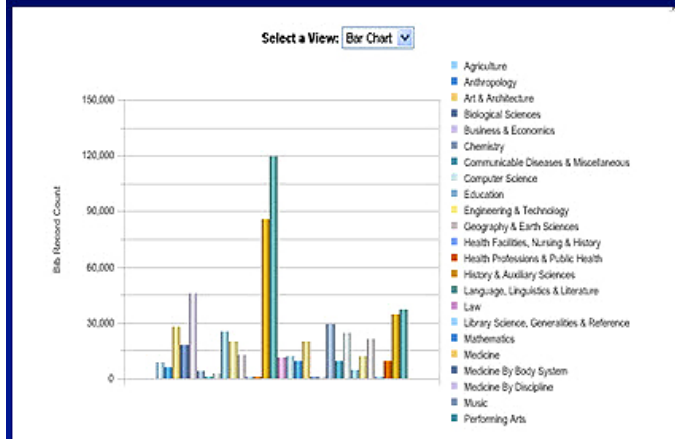
Education Overview Dashboard

- The Proof of Concept aspect of this Dashboard was to experiment with the creation of reports that would compare program data from SUNY Institutional Research with library data. Two Pages were created with side by side reports with program information pulled from the Institutional Research database and subject/circulation data pulled from the Library database.

Executive/Financial Overview Dashboard

- The Proof of Concept aspect of this Dashboard was to build queries from data sources (such as spreadsheets) outside of the physical database in the Oracle Business Intelligence library subject database. Two Pages were created to experiment with the development of reports based on data from the 2002-2008 biennial National Center for Education Statistics (NCES) Academic Library surveys. The pages include a review of expenditures for one time and continuing resources and for interlibrary loan transactions. All data in these surveys are self reported by the campuses.

Subject Division Breakdown by Publication Dates 1960 - 1989



Findings

- There was no history of standard SUNY-wide library collection development reports. The structure of the POC was based on a combination of earlier SUNY collection analysis investigations, local campus activities, and functionality review of available commercial services.
- Data from six libraries were used (2 Comprehensive Colleges, 2 Community Colleges, and 2 Colleges of Technology--Geneseo and Potsdam, Orange and Suffolk, Alfred and Cobleskill). Data from additional campuses would need to be added before any reliable conclusions can be made on SUNY-wide collections trends.
- Data anomalies were known prior to the start of the POC. The main ones included the following. These are data issues that would need to be addressed locally at each campus.
 - Not all bibliographic data contained call number information that could be mapped to divisions and categories for subject analysis of holdings (8.9% of POC data)
 - Not all bibliographic data contained an OCLC#, a unique number that would be used to determine whether one or multiple campuses owned the same bibliographic title (9.2% of POC data)
 - Not all bibliographic records contained valid dates of publication (1% of POC data)
 - Overall, the correctness of values representing attributes such as language, format, type of reproduction, are dependent on local campus input.
- Unanticipated data anomalies were also found. They have been documented and will need to be considered as the project moves forward.
 - The 035 field that contains the OCLC# was found to include values other than the OCLC#. Data cleanup was done by OLIS staff on the six POC campus databases to ensure that values in the 035 field were true OCLC values.
 - Circulation events data was in the data extracts even though corresponding bibliographic and item data was not extracted because those records had been marked for deletion. This will be addressed in future data extracts.
 - Circulation events and bibliographic records were found that did not have item level details (contains circulation policy information). These findings need additional investigation as how to address.
- The Executive/Financial Facts were created from 2002-2008 biennial National Center for Education Statistics (NCES) Academic Library surveys. The data is self-reported. Use of this data would require many caveats.
- Within the data identified for this POC, there was no data available to accurately identify in which year a given bib record was added to a campus's collection. Therefore, the publication year was chosen as a method to estimate the year in which a bib record entered the collection.

Next Steps


Demonstrations of and access to the Library Dashboards have been provided to staff at the six POC campuses. They have been asked to [review and comment](#) on the four Dashboards. Demonstrations are also scheduled for the SUNY Council of Library Directors October Fall Meeting and at the [2009 LiSUG Conference](#). The review period will be held through December 2009. Based on the POC evaluation, a determination will be made regarding the direction of the project.

Considerations for moving forward include:

- feedback from campus review
- determine possible changes to the data model, available metrics, and updates to the Dashboards
- determine costs for staff resources and hardware
- review data anomalies to determine what changes might be needed in extract parameters
- determining types of data from Institutional Research to compare with library data and how to best integrate the data from these different sources
- determine the best source of data and types of metrics to be used with the Executive/Financial Dashboard

In order for this project to be sustainable, a strong support model needs to be developed to ensure that required updates to the Library database and Dashboards can be maintained, to provide necessary training and support to the campuses, and to ensure that campus access to the Dashboards can be provided. If the POC is considered successful and the decision is made to go forward additional considerations could include the development of an E-Resources Usage Dashboard to analyze database usage trends and the development of an Interlibrary Loan Dashboard (as an aid to collection development and resource sharing) to evaluate the types of materials not available within the SUNY system.

The [Oracle Business Intelligence software](#) proved to be a powerful and flexible application. The role of the consultant was critical in the rapid development of Dashboards and in providing a strong foundation for using best-practices. Further, a solid working relationship that developed among staff from System Administration's Office of Library and Information Services and Office of Information and Technology significantly contributed to the success of the initial phase of the Library POC and will be a great asset as the Library project moves forward.

What is 'Business Intelligence' software?

A set of concepts and methods to improve business decision making by using fact-based support systems. BI is sometimes used interchangeably with briefing books, report and query tools and executive information systems. Business Intelligence systems are data-driven DSS(decision support services).

Cover Story

[Library Resources via Angel \(Fredonia\)](#)

[Library Resources via Angel \(Onondaga\)](#)



Features

[OLIS Library Resources Starting Point](#)

[Welcome Chancellor Zimpher](#)

[Brockport and Aquabrowser](#)

Features

[SUNY Business Intelligence](#)

[Focus on SUNY Libraries](#)

[SUNYConnect Updates](#)

How to Contact Us

Linkable Links