Purdue University
Purdue e-Pubs

Charleston Library Conference

Discovery by the Numbers: An Examination of the Impact of a Discovery Tool through Usage Statistics

Jody Fagan
James Madison University, faganjc@jmu.edu

Meris Mandernach
James Madison University, manderma@jmu.edu

Follow this and additional works at: http://docs.lib.purdue.edu/charleston
An indexed, print copy of the Proceedings is also available for purchase at: http://www.thepress.purdue.edu/series/charleston.
You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences.

Jody Fagan and Meris Mandernach, "Discovery by the Numbers: An Examination of the Impact of a Discovery Tool through Usage Statistics" (2011). *Proceedings of the Charleston Library Conference.*
http://dx.doi.org/10.5703/1288284314967

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Discovery by the Numbers: An Examination of the Impact of a Discovery Tool through Usage Statistics

Jody Fagan, Director of Scholarly Content Systems, James Madison University
Meris Mandernach, Collection Management Librarian, James Madison University

Abstract:
In August 2010, James Madison University (JMU) implemented EBSCO Discovery Service (EDS) and placed its search widget front and center on the library home page. This paper will examine general usage trends over the tool's first two semesters, including changes in physical circulation, library catalog searches, home page traffic, and other database usage. Searches, sessions, and full-text downloads of subject-specific databases before and after the implementation of the discovery tool will be compared. Finally, the limitations of the data and our methods will be discussed in order to inform other libraries' work with similar data. The objective of the paper will be to share information for those considering a discovery tool or those preparing to evaluate a discovery tool that has already been implemented.

Introduction

Changes reflected in library use statistics can be attributed to many factors; the adage "correlation does not imply causation" definitely applies. Nevertheless, everyone wants to know how implementing a discovery tool affects the usage of library collections, both in print and online. Administrators may hope to reduce spending on resources made superfluous by the discovery tool. Academic subject librarians may worry that students, seduced by the discovery tool's Google-like appearance, will not use subject-specific databases. Web librarians wonder if the tool confuses online pathways to other library systems and services.

James Madison University is a comprehensive public university located in Harrisonburg, Virginia. With over 19,000 students, it is primarily undergraduate-focused. JMU also belongs to the Virtual Library of Virginia (VIVA) consortia, which provides access to a number of databases and journals.

JMU Libraries implemented EBSCO Discovery Service (EDS) during the summer of 2010, with full release for the fall 2010 semester. Previously, the search box on the libraries' home page forced users to choose between "Catalog" and "Articles & Journals." In fall 2010, the tabbed search box was replaced with "Quick Search," a single-input box entry into EBSCO Discovery Service (see Figure 1). Drop-down menus were added for the user to limit to keyword, title, or author searches, as well as by source type: Articles, Books, and More; Just Articles; Just Books; and so on.

Methodology

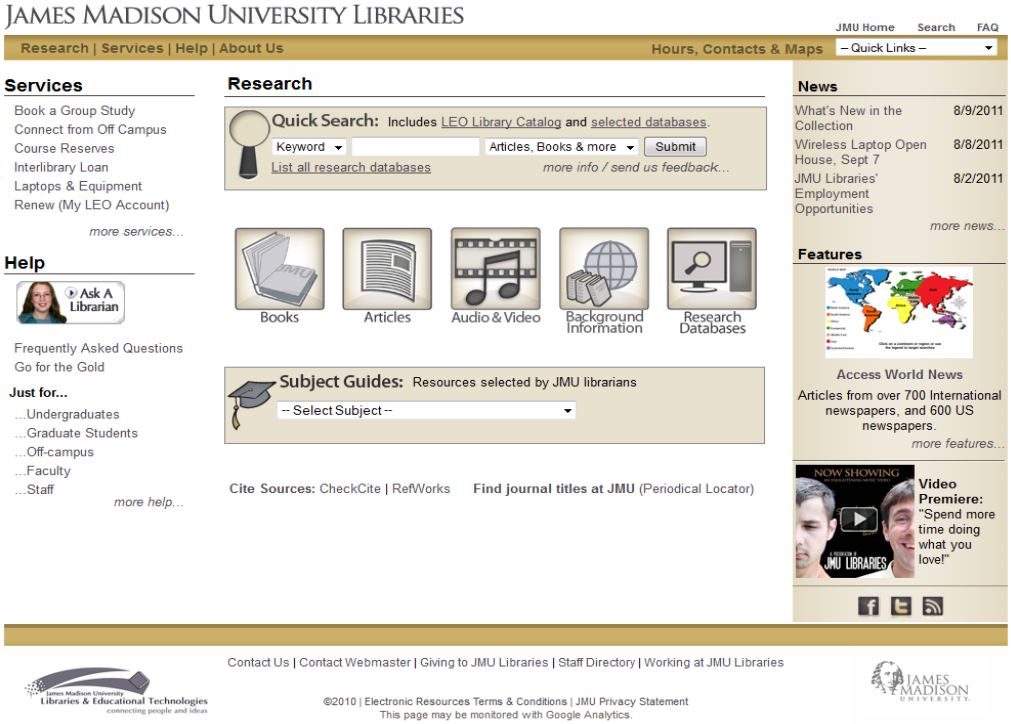
In order to get an idea of general library use patterns at JMU, we chose to look at the past five years of data, when available. The discovery tool was implemented in July 2010, and therefore we have only one year of usage data, July 2010 to June 2011. To assess how database use has changed since the implementation of the discovery tool, we chose to examine the percent change in searches, sessions, and full-text downloads between FY2010 and FY2011. Additionally, we noted that a number of databases changed platforms, were included in a federated search, or were taught in a different manner by librarians, all of which would affect use of specific databases.

Our intention was to use the most standardized information available, and we planned to select which databases to compare based on their level of COUNTER compliance. However, other factors intervened. First of all, EBSCO's statistical reports, for both the discovery tool and other databases, are problematic. Currently, searches in the discovery tool "count" as searches in other EBSCO databases, creating a problem libraries are familiar with from federated search experiences: wildly over-counted searches. Second, some databases that are COUNTER compliant experienced a major platform change in FY2010,

Copyright of this contribution remains in the name of the author(s).
DOI: http://dx.doi.org/10.5703/1288284314967

and we believe usage could have been significantly affected by that change. For example, JMU has subscribed to a variety of both CSA and ProQuest databases for many years; however, a major platform change across these databases created serious technical difficulties, which seem to have affected the usage statistics.

Figure 1: Quick Search at JMU Libraries (EBSCO Discovery Service)



Results

Library circulation, gate counts, home page visits, catalog searches

In the last five years, our circulation counts have hovered around 200,000 items per year (excluding e-book statistics) in a collection of 560,694 items. Seventeen percent fewer items circulated in FY2011 than in any of the previous four fiscal years (see Figure 2). Interlibrary loan requests, both book and article, steadily increased until FY2010 and dropped slightly in FY2011. The library home page, on the other hand, has continued a slow increase in use since 2009 (see Figure 3).

Figure 2: Total Circulation at JMU Libraries (excluding e-books) FY2007-FY2011

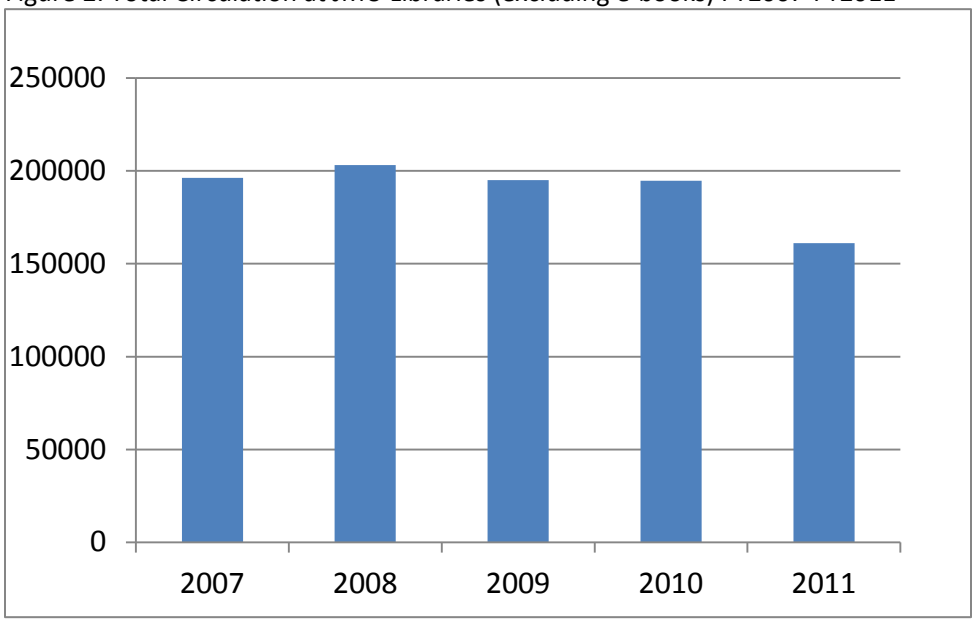
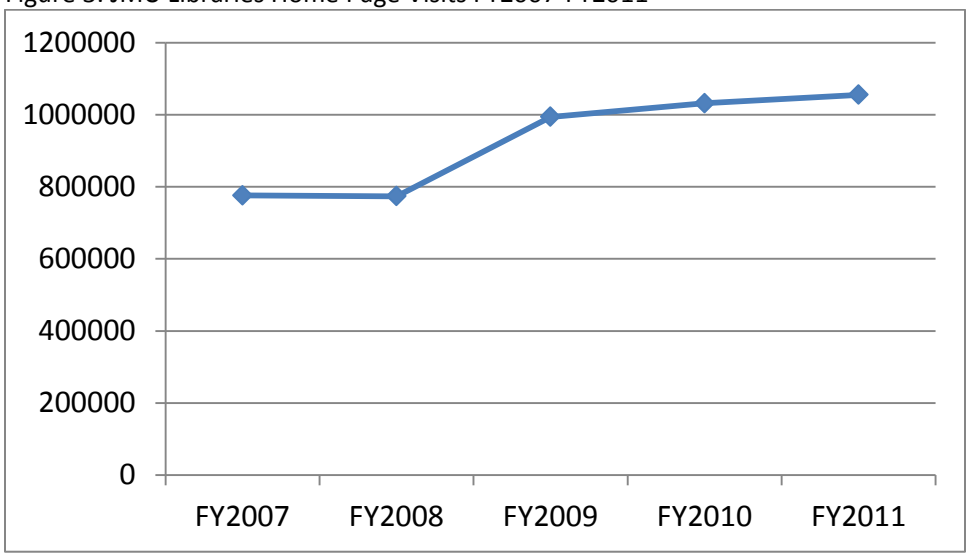
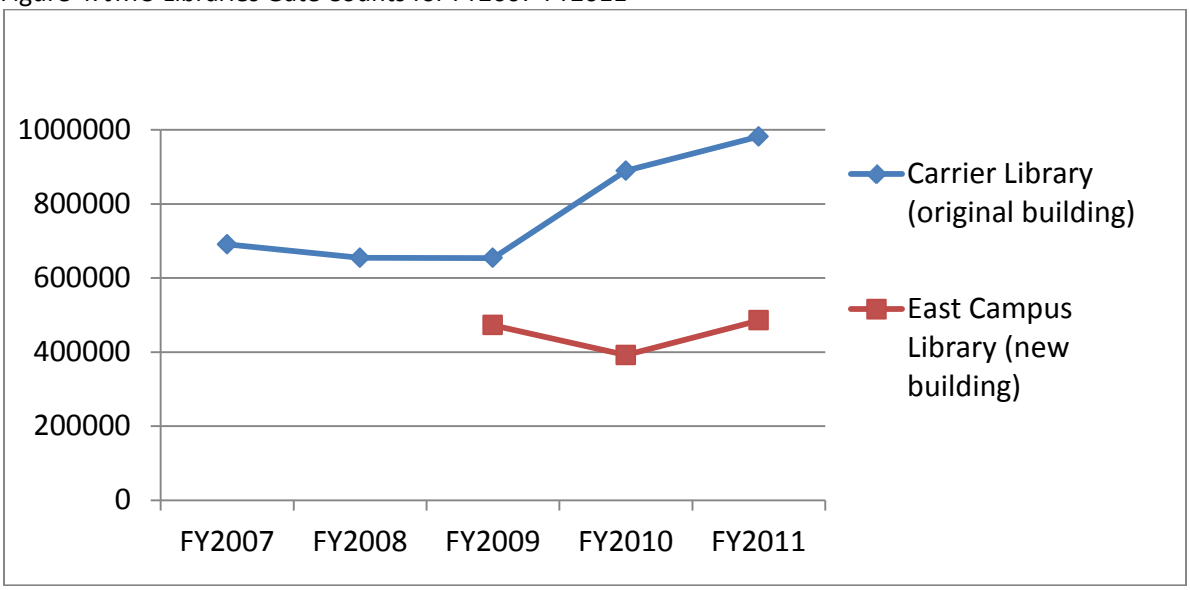


Figure 3: JMU Libraries Home Page Visits FY2007-FY2011



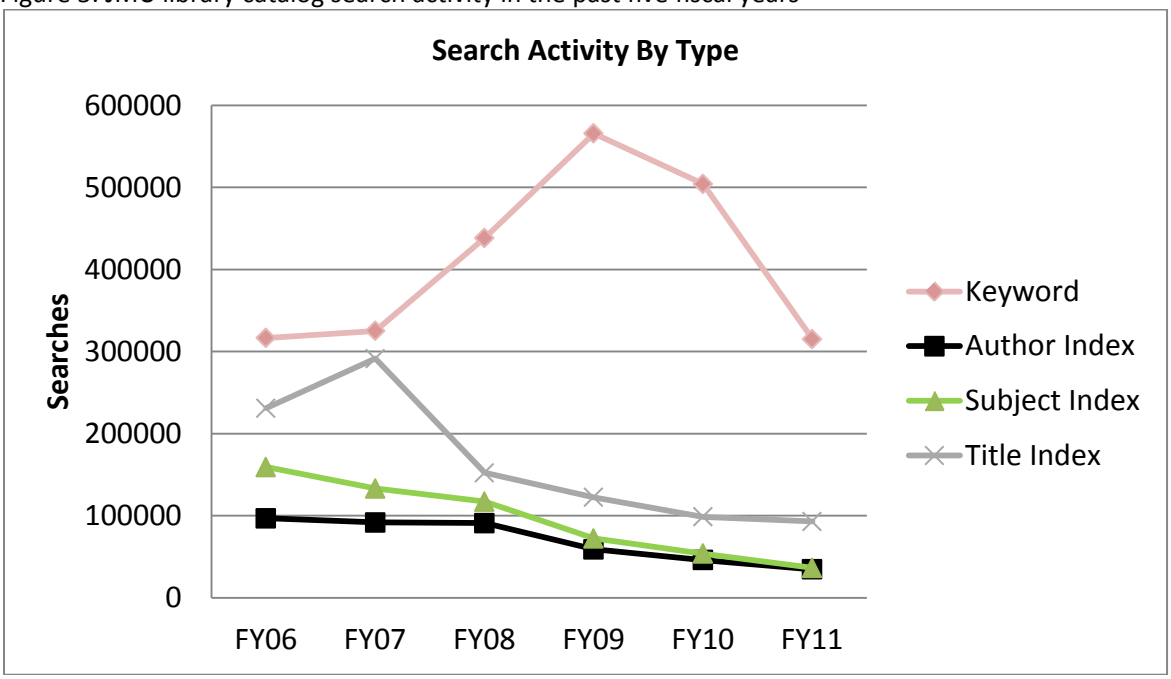
A few major changes at JMU are worth noting, although we cannot draw direct connections between the events and the use data. In 2008, JMU built a large, new library building, at which time one-third of the books were moved to the new library. During FY2009, a number of new librarians were hired and new programs were started within the Public Services department of the libraries. Also, in FY2010, JMU opened a Starbucks in the original library building; at the same time, building gate counts dramatically increased (see Figure 4).

Figure 4: JMU Libraries Gate Counts for FY2007-FY2011



Sessions, Searches, and Full Text Accesses
Looking at searches in JMU's library catalog for the past five years reveals two different trends (see Figure 5). All data include both staff and end-user searching in the public catalog interface. Index searches, including author, title, and subject searches, have slowly declined since FY2007, with all three types of searches netting fewer than 100,000 searches per fiscal year. On the other hand, keyword searches experienced dramatic growth from FY2007 through FY2009, for an approximate 75% increase. However, they then fell by the same amount from FY2009 to FY2011, with the biggest decline from FY2010 to FY2011, which coincides with the first year of the discovery service at JMU.

Figure 5: JMU library catalog search activity in the past five fiscal years

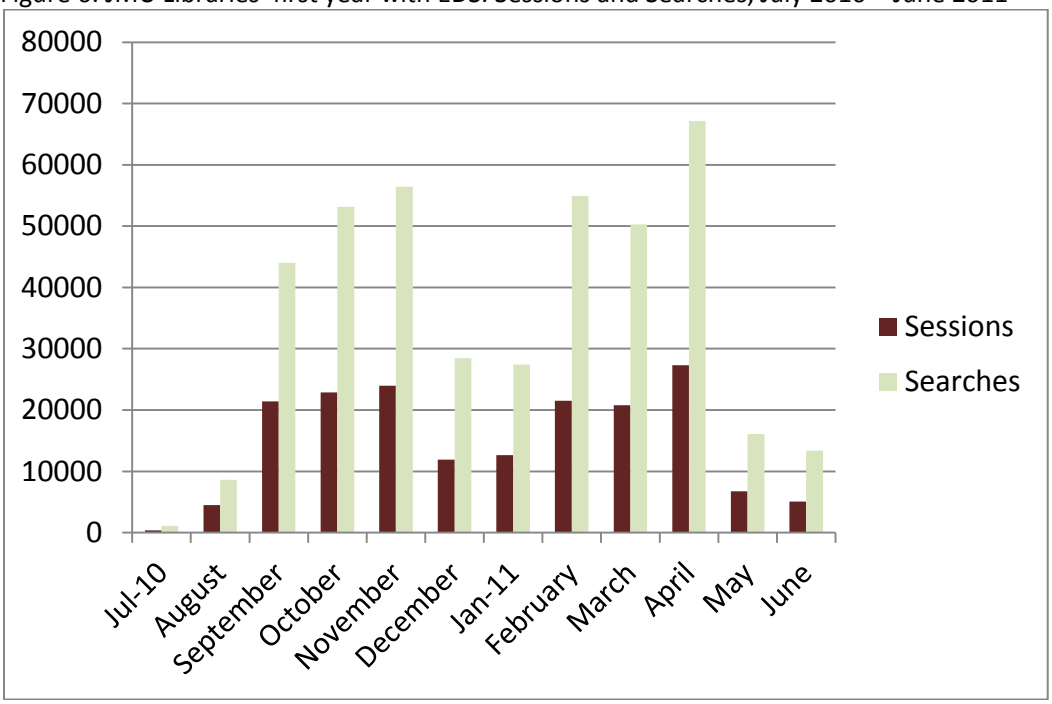


EBSCO Discovery Service

JMU's first semester with the discovery tool began in August 2010. Figure 6 shows it received over 20,000 sessions per month, with an average of over two searches per session. The average searches per session matches our previous experience with EBSCO's Academic Search Complete. Additionally, the use pattern over the course of the academic year matches the pattern for other research databases at JMU.

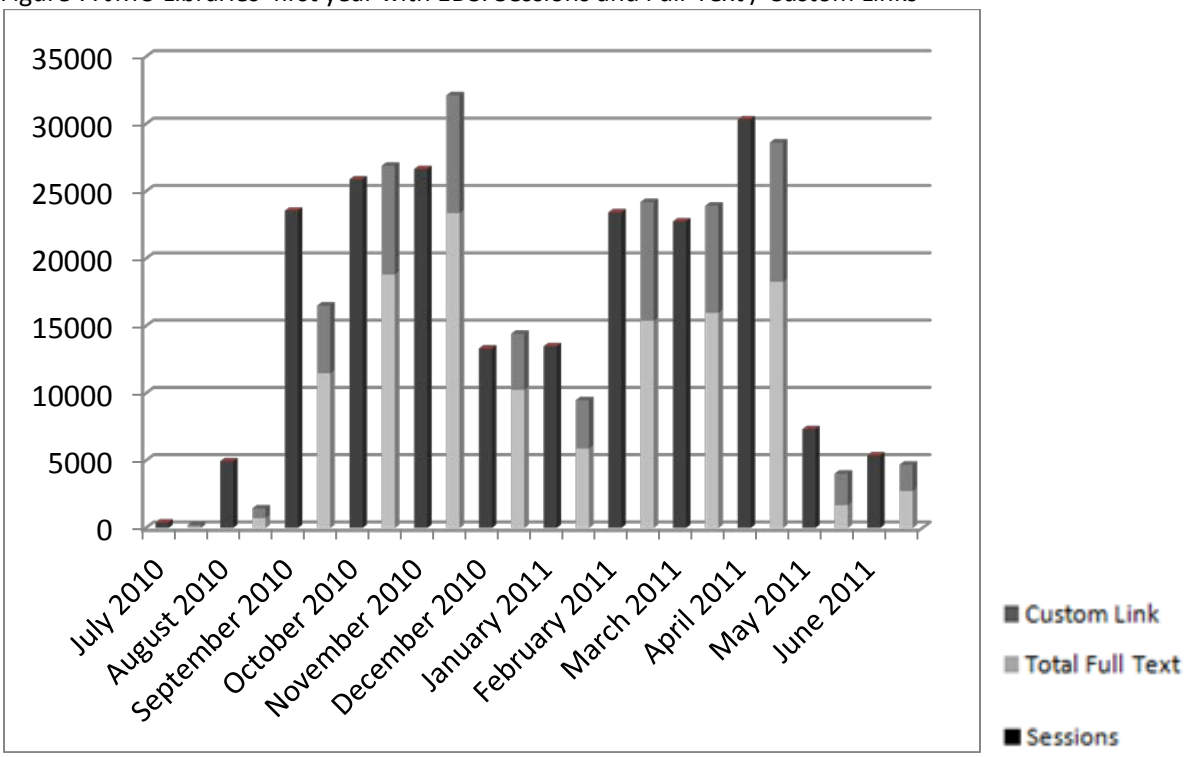
The EBSCO data reported in this presentation is from custom reports based on JMU profiles rather than from the COUNTER reports. Since JMU set up access to EDS through one set of profiles, and access to other databases using a different set of profiles, the data in this paper show just the searches and sessions for EDS or the specific databases in question.

Figure 6: JMU Libraries' first year with EDS: Sessions and Searches, July 2010 – June 2011



In EBSCO Discovery Service, "Custom Links" currently includes not only clicks on the link resolver button, but also links to library catalog records. For the entire fiscal year, about 11% of custom links from EDS were targeting the catalog, that is, a user was in EDS and clicked on the link to open the library catalog. There are links to the catalog for each catalog record in EDS as well as a link to the catalog in the header. The remaining custom link traffic currently represents access to the link resolver. So, the count of full-text downloads plus custom link use is one way to look at how relevant the results were to users. Looking at the full-text downloads and custom links over the course of FY2011, there were 1.26 full-text downloads plus custom links per session (see Figure 7). For comparison, in Academic Search Complete, there were 1.46 full-text downloads plus custom links per session. These numbers reflect full-text and custom link use from within these databases and exclude full-text accesses through the link resolver.

Figure 7: JMU Libraries' first year with EDS: Sessions and Full-Text / Custom Links



Usage of Full Text and Other Databases

Usage of Full Text

One of the more consistent usage reports JMU has is the report from Serials Solutions that shows the number of clickthroughs, or links to full text, that pass through the link resolver. JMU has used Serials Solutions as the e-journal portal and link resolver software provider for many years, and it has been a stable system, consistently implemented into new databases. The number of clickthrough requests dropped by 13.5% from FY2010 to FY2011. The larger trend can be seen in Figure 8, with usage dramatically increasing from FY2007 to FY2008, then holding fairly steady through FY2009. For most databases, the number of full-text downloads reported by the vendor include traffic coming from link resolver software. For JMU's EDS implementation, the number of full-text downloads via EDS does not include link resolver traffic. Figure 9 shows the number of direct full-text accesses in EDS. It is important to note that we were unable to add direct full-text accesses from other resources to this graph, as most of the full-text download statistics for a given database include link resolver traffic. Because of this constraint, we were unable to evaluate the actual proportion of additional full text accessed by our users through EDS.

Figure 8: Total Clickthroughs through Link Resolver FY2001-FY2011

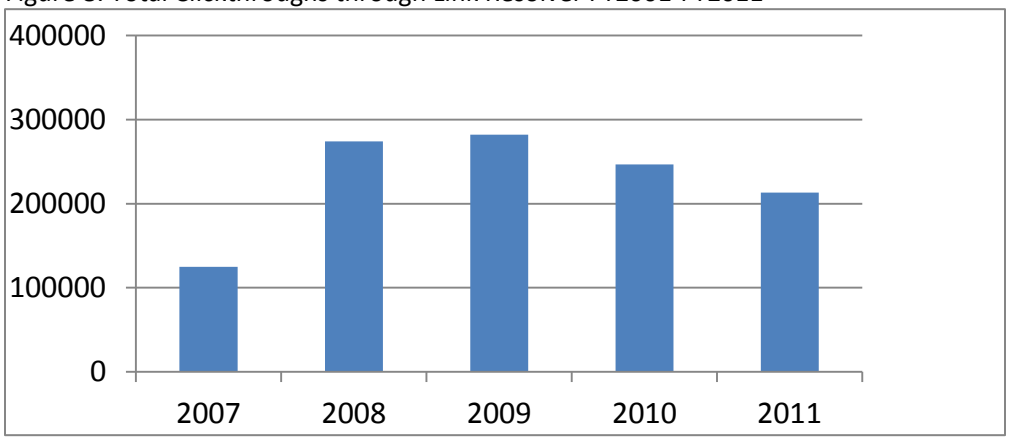
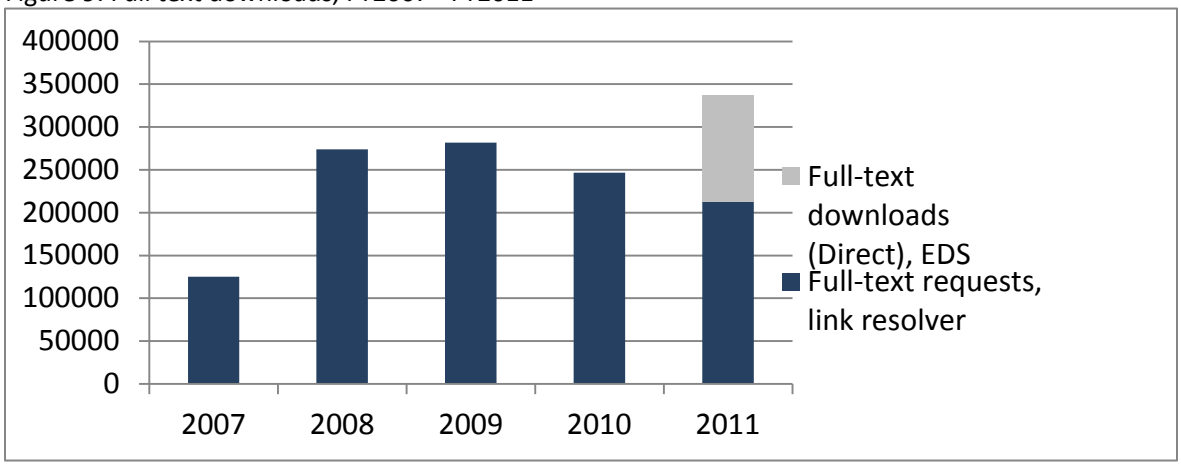
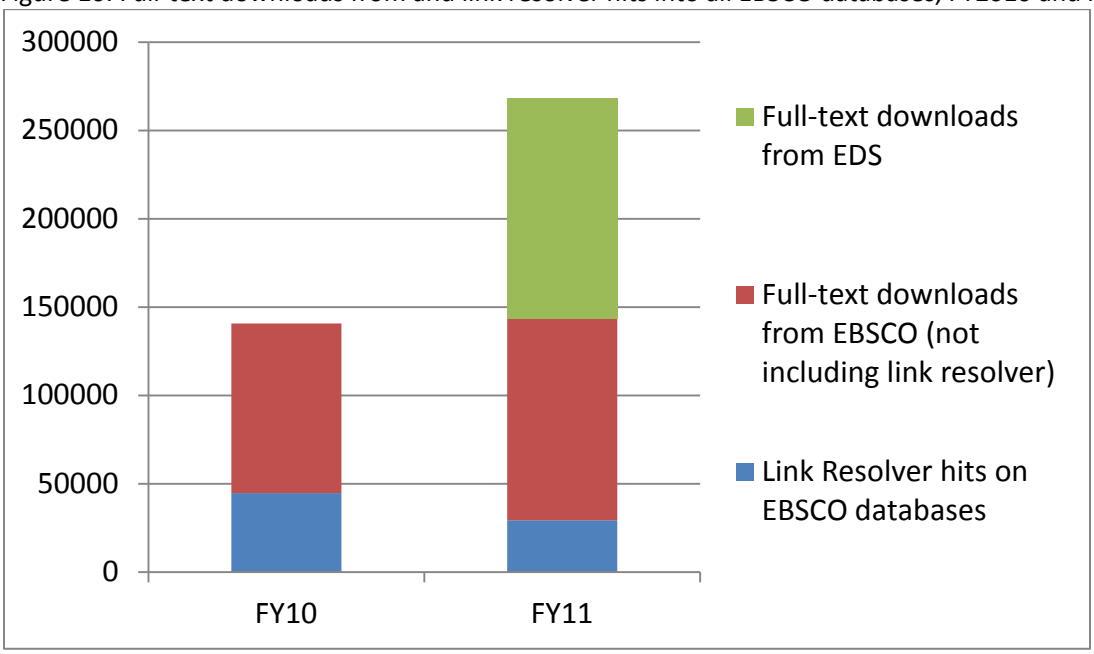


Figure 9: Full text downloads, FY2007 - FY2011



Looking at just EBSCO databases, the total full-text accesses have increased due to the large number of full-text downloads from EDS (see Figure 10). The proportion of direct full-text downloads has increased, while the link resolver hits on EBSCO databases have decreased.

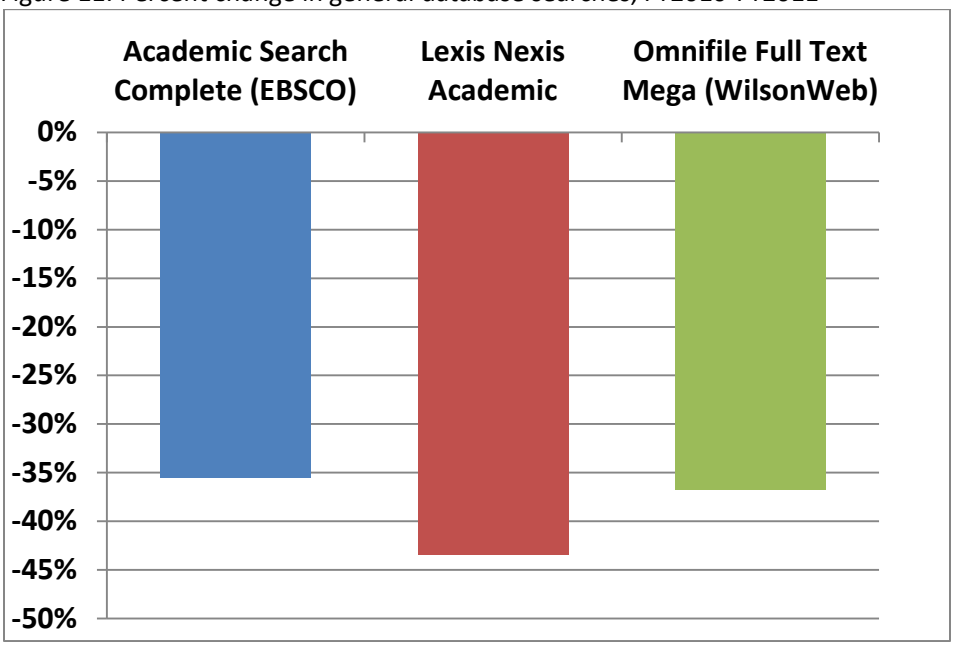
Figure 10: Full-text downloads from and link resolver hits into all EBSCO databases, FY2010 and FY2011



General Databases

JMU subscribes to three multidisciplinary article databases, Lexis Nexis Academic, Academic Search Complete, and Wilson OmniFile. Only Academic Search Complete's content was included in EDS during the time of this study (EBSCO will be adding Wilson OmniFile content in January 2012). All three experienced decreases in searches from FY2010 to FY2011: EBSCO Academic Search by 32%, LexisNexis Academic by 43%, and Wilson OmniFile by 37% (see Figure 11).

Figure 11: Percent change in general database searches, FY2010-FY2011

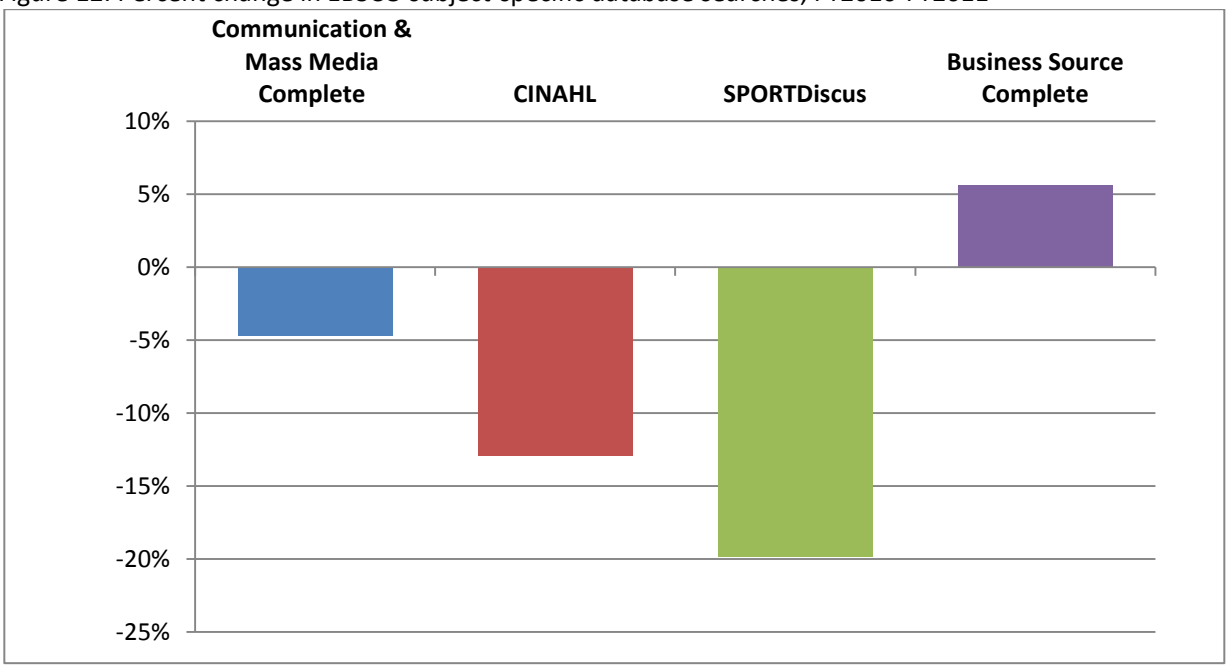


Looking more closely at the change in how Academic Search is used, there were 1.46 full-text accesses plus custom links per session in FY2011, compared to 1.27 full-text accesses plus custom links in FY2010. These figures exclude the traffic from the link resolver to avoid double-counting full-text and session accesses.

Subject-Specific Databases: Other EBSCO Databases

EBSCO Discovery Service includes all of JMU's EBSCO subscriptions in its search. Users can and do search EBSCO databases directly. However, looking at searches within subject-specific databases JMU subscribes to via EBSCO, Communications & Mass Media Complete decreased by 5%, CINAHL with Full Text usage dropped by 17%, and SPORTDiscus with Full Text dropped 39%. Business Source Complete increased in usage by 6% (see Figure 12).

Figure 12: Percent change in EBSCO subject-specific database searches, FY2010-FY2011

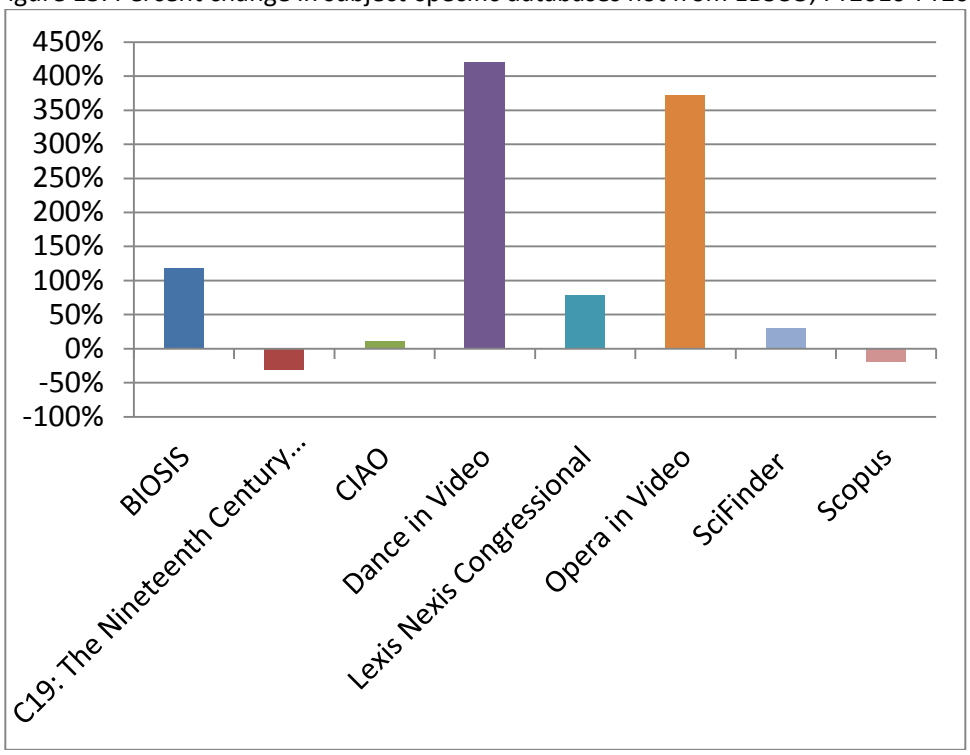


The full-text accesses plus custom links per session for Business Source Complete, Communications & Mass Media Complete, and SPORTDiscus with Full Text all decreased slightly, at 5%, 1.7%, and 3.7%, respectively. In FY2011, there were 1.76 full-text or custom link accesses per session in Business Source Complete, 1.31 full-text or custom link accesses per session in Communication & Mass Media Complete, and 1.56 full-text or custom link accesses per session in SPORTDiscus with Full Text. From FY2010 to FY2011, CINAHL Plus with Full Text increased 13.6%, reaching a ratio of 1.25 full-text or custom link accesses per session. These numbers also exclude link resolver traffic.

Subject-Specific Databases: Non-EBSCO Databases

Finally, we looked at several subject-specific databases that JMU subscribes to from vendors other than EBSCO. Results varied widely (see Figure 13). Several databases experienced dramatic increases in searches, including BIOSIS (118%), Dance in Video (420%), LexisNexis Congressional (78%), and Opera in Video (371%). Other increases were more moderate: Columbia International Affairs Online (CIAO) increased by 10% and SciFinder Scholar jumped 30%. Two databases decreased in usage: C19 decreased by 31%, and Scopus dropped in usage by 20% from FY2010; however, FY2010 uses were higher than FY2009 uses. With the exception of the "in Video" products, JMU has subscribed to these resources for at least five years.

Figure 13: Percent change in subject-specific databases not from EBSCO, FY2010-FY2011



Discussion and Analysis

A certain amount of natural variation occurs in all areas of life, so we do not think the changes—up or down—are necessarily signs of success or failure. Nor are we able to assign causes to changes. However, this analysis has enabled us to answer some questions that have arisen, at least on a preliminary basis. First, general library usage trends seem to be more affected by other factors than the advent of the discovery tool. Although there was a recent increase, then stabilization, of library Web site use and an increase in gate counts, circulation counts decreased. In JMU's case, changes in our building use patterns, including a new library building and a Starbucks added to the original library building, seemed to have more impact than the launch of the discovery tool. The decrease in circulation counts may suggest an increased reliance on the articles that are so prevalent in discovery tool results, but it could also be related to curricular changes, changes in how we count circulations, an increased use of e-books, or a combination of all these factors. Searches in the library catalog and general, multidisciplinary databases have dramatically decreased in the past year. Changes in our library Web site navigation did emphasize the discovery tool as a primary search pathway, and we suspect this is the major factor here, although we have no empirical basis for this conjecture. The library catalog used to be the engine behind our home page search box, and it was replaced by the discovery tool. Prior to the discovery tool, the "articles" tab provided a direct pathway to Academic Search Complete, which has since decreased in use. Such patterns should not be alarming; they seem to be the natural effects of discovery tool implementation, since some of the tool's purpose overlaps with these products.

The changing patterns in full-text access we observed were intriguing, but contextual factors prohibit conclusive answers. For example, JMU has continued to add full-text content through JSTOR and other sources, which may affect whether people are accessing documents through the link resolver or through direct links to full text. In addition, databases like SPORTDiscus with Full Text that have direct links to PDFs reduce link resolver usage. However, increasing the number of full-text sources does also increase the link resolver's knowledge base, which may actually increase link resolver usage. For example, JMU users of Google Scholar will be able to access more full text via the link resolver

as its knowledge base grows. In short, it is difficult to know if users are *discovering* more full text or if there is just more full text to be discovered.
Users clicked on full text or custom links more often per session when using Academic Search Complete, Business Source Complete, Communications & Mass Media Complete, and SPORTDiscus with Full Text than when using EDS or CINAHL. That could mean these databases were more satisfying than EDS or CINAHL. Or, it could mean EDS users were finding physical items like books. Changes in the full-text plus custom link ratio from year to year seem to reflect change due to something other than discovery. In the year since the discovery tool was implemented, the ratio increased with both Academic Search Complete and CINAHL, but slightly decreased with Business Source Complete, Communications & Mass Media Complete, and SPORTDiscus. One factor could be changes in instruction with CINAHL and Academic Search Complete. This area is ripe for further research.

It is also important to remember that vendors may not differentiate between direct full-text accesses and link-resolver accesses, making it challenging to determine where the user was when they "discovered" content. This factor makes it difficult to interpret the numbers related to full-text usage. However, it does seem like this area will be helpful in evaluating discovery routes to content.

Although at first blush it appears that libraries may not need to continue access to general, multidisciplinary databases as well as a discovery tool, the devil is in the details. Perhaps more importantly, the only reason JMU can include full indexing for many of the databases searched by EDS is because we subscribe to those products, such as Academic Search Complete, WilsonWeb OmniFile, JSTOR, and ScienceDirect. So, even here there are no easy answers.

The discovery tool seems to have had some impact on usage of subject-specific databases, although the magnitude varies widely. Searches in the native interfaces of EBSCO databases decreased, while some subject-specific databases from vendors other than EBSCO increased in usage. Changes in individual database use for non-EBSCO databases are clearly linked to other factors besides discovery, because they do not show any overall pattern: some increased dramatically, while others decreased. We suspect these changes are more related to library instruction, curriculum, and faculty research interests. The good news for subject librarians is that subject databases do not seem inherently threatened by discovery tools. Given the amount of increases in use of these subject-specific databases, it would appear that librarians' concerns over students being solely enticed by the promise of a single search are off the mark: students are savvy in their information needs and are able to locate subject-specific resources when the need arises.

We see several implications for collection development decisions. First, libraries need to understand the extent to which their discovery tool's inclusion of multidisciplinary databases and subject-specific content depends on their subscriptions to other resources. This can also relate to the extent to which the discovery tool indexes other vendors' databases. For example, Serials Solutions' Summon covers many of the journals included in EBSCO's Academic Search Complete and CINAHL, but it does not technically include Academic Search Complete or CINAHL. Therefore, if a library purchases Summon, it might be able to cancel Academic Search Complete or CINAHL without losing too much access. However, record retrieval from these databases may be affected. If the same library subscribed to EDS, it would want to enter into negotiations with a clear understanding of how its subscriptions to other EBSCO products would affect EDS pricing, indexing, and content, and vice versa.

Second, at least for the near present, libraries should continue to purchase subject-specific databases to support core curricular areas. At JMU, subject-specific databases are still heavily used. Discovery tools are not meant to reproduce the carefully cultivated content and specialized interface features of subject databases. Perhaps some savings could be realized by relying on the discovery tool for interdisciplinary coverage or to cover a subject area that does not have a major program to support it. For example, a library without a strong engineering program might be able to rely on the exposure of engineering journals in a discovery tool and avoid subscribing to Compendex.

The decrease in our circulation statistics makes us wonder if discovery has a significant impact on print collection use. Perhaps by making paths to online full text (both e-books and online articles) more obvious, users are quicker to rely on those sources instead of visiting the library to check out a book. Although JMU has a delivery service which allows patrons to route items to the most convenient library location, it does not deliver to dorms. It would be interesting to know how circulation and online full-text use by discipline varied, but this data will likely never be available from library system logs due to privacy concerns. Surveying users is another way to approach this question, but it too would have shortcomings due to the limitations of self-reporting.

This project underscored the importance of looking at many years of data when making collection development decisions, not just two or three. A dramatic change in use from one year to the next may reflect a return to normal rather than a new aberration. In future examinations of how discovery tools (or anything else) affect collection usage, it will be important to try to look at multiple years of data. Because discovery tools are new, it will take some time for usage trends to fully reveal the impact these tools have on library collections.

Conclusion

Wouldn't it be nice if there was a simple sentence that encapsulated the impact of discovery? However, discovery tools are so multi-faceted that there is no single approach to evaluating their effect on library use. Looking at usage trends over many years from many sources is important to shape a more complete view and provide more confidence for decision-making.

Identifying research questions is critical to avoid information overload when it comes to usage data. The more information gathered, the more overwhelming it can be. However, approaching massive spreadsheets with a purpose in mind can help quickly narrow the scope of analysis for a given project. Taking a big-picture view initially can illuminate the most interesting areas for future inquiries. For JMU, these will be full-text downloads and physical circulations.

Understanding usage trends of the physical and virtual collections and systems improves libraries' ability to plan strategically. Future systems decisions, such as an ILS migration, could relate to discovery trends. Usage information can also provide a common ground for organizational discussions related to a library's virtual presence. Finally, usage data should inform future allocation of resources for staff and materials.