Deputy Vice-Chancellor (Academic) - Papers

Deputy Vice-Chancellor (Academic)

2014

Unlocking the potential of library generated data to assess value, impact, and influence

Margie Jantti
University of Wollongong, margie@uow.edu.au

Brian Cox
University of Wollongong, bcox@uow.edu.au

Unlocking the potential of library generated data to assess value, impact, and influence

Abstract

The University of Wollongong Library (UWL) has initiated an innovative and collaborative research project to unambiguously demonstrate the value that can be provided by academic libraries. The project centres on the integration and interrogation of a series of discrete datasets, e.g. student performance, student attrition, student demographic data, and borrowing and electronic resources usage data. The project has enabled UWL to identify whether a correlation exists between usage of Library resources and academic performance (e.g. grades). Findings reveal a strong and sustained correlation; providing a new facet through which to view and understand the student academic experience.The project is different to other institutions' efforts to link usage to student outcomes, in that the Library Cube is not a one-off research project, but is now an ongoing part of UOW's systems and performance reporting. The most recent development centres on the Marketing Cube. This cube will demonstrate whether electronic resources are being used to their optimum. Updated weekly, it's now possible to assess the uptake of uptake of products (e.g. databases), how often they are being used and to test the impact of marketing and promotion initiatives as to whether user behaviour can be influenced through targeted marketing.

UNLOCKING THE POTENTIAL OF LIBRARY GENERATED DATA TO ASSESS VALUE, IMPACT AND INFLUENCE.

Margie Jantti and Brian Cox
University of Wollongong Library, Australia
margie_jantti@uow.edu.au bcox@uow.edu.au

Abstract
The University of Wollongong Library (UWL) has initiated an innovative and collaborative research project to unambiguously demonstrate the value that can be provided by academic libraries. The project centres on the integration and interrogation of a series of discrete datasets, e.g. student performance, student attrition, student demographic data, and borrowing and electronic resources usage data. The project has enabled UWL to identify whether a correlation exists between usage of Library resources and academic performance (e.g. grades). Findings reveal a strong and sustained correlation; providing a new facet through which to view and understand the student academic experience.The project is different to other institutions' efforts to link usage to student outcomes, in that the Library Cube is not a one-off research project, but is now an ongoing part of UOW's systems and performance reporting.

The most recent development centres on the *Marketing Cube*. This cube will demonstrate whether electronic resources are being used to their optimum. Updated weekly, it's now possible to assess the uptake of uptake of products (e.g. databases), how often they are being used and to test the impact of marketing and promotion initiatives as to whether user behaviour can be influenced through targeted marketing.

Key Words: analytics, value, evaluation

Introduction
The desire to assess the value or impact of libraries through the measurement of library use is not a new phenomenon. Over the past decade and beyond, a number studies have been undertaken to understand the significance of the role of the library in the student experience and to ascertain whether correlations exist with student engagement with a library's resources, services and space and of their performance, i.e. academic grades (Association of College and Research Libraries, 2010. (De Jager,2002) and(Wells,1995) describe methodologies used in the then print dominated Library setting. The increasing sophistication of data capture and warehousing capability within contemporary tertiary education institutions has offered new approaches for the examination of data generated from library management systems and other data sources to assess the impact of both print and electronic resources and in some instances, the use of Library space (Stone and Ramsden, 2013). This case study describes how the University of Wollongong Library (UWL) in partnership with the University's Performance Indicator Unit partnered to build new data and reporting models to join library usage data with students' demographic and academic performance data to test correlations and to ultimately create a new narrative; a discourse on the contribution of the library to the student experience and their academic success.

Constructing the problem statements
The proliferation of information via services external to the library, e.g. the internet, has resulted in a seismic shift in how libraries are used. Students have unprecedented choice in the source, content and format of learning materials they use; and can effectively bypass the library. These wide-ranging choices have in effect repositioned the student from being an active, though still highly dependent learner, into a consumer of information. This shift in behaviour comes with a learning cost, and it has become a battle that is fought daily in tutorial classes and lectures, as academics and librarians try to encourage students to make better use of high quality sources of information, acquired or subscribed to by the library (at considerable cost). For these reasons, it is more

important than ever for libraries to demonstrate to students and stakeholders the value of using the library's resources and services.

For the academic library, a key focus is on the transformative power of information; and the question to be answered is: does a student's academic performance improve as a result of using information resources made available by the library; what does the student receive for their investment of time and energy spent in using library resources and services (Neal, 2011)? Using data to demonstrate the worth and value of libraries has, up until recently, been heavily reliant on discrete evaluation projects, the use of surveys and other qualitative assessments. While these activities continue to be worthy, many are questioning as to whether they are sufficient for recognition of the library as an indispensable partner to the university; that through a library's endeavours there is a demonstrable benefit to the university's current and aspirational aims. For example, many of us wish to believe that through the utilisation of Library resources and services that there is a positive effect on student performance; the challenge lies with the availability and visibility of evidence to underscore such claims. This is despite the fact that libraries generate and/or collect lots of data.

Prior to the development of the Cube, UWL lacked ongoing valid and reliable data collection from both library and enterprise systems, which prevented us from making supported assertions about the value provided by the library. What was needed was a cost-effective, reliable, and sustainable way of collecting information on the library's impact on student outcomes. The genesis of the UWL Cube was centred on the examination of two significant interrelated problems facing academic libraries: Firstly, many libraries are under mounting pressure to demonstrate value in the context of economic adversity (and accountability) and, the intensification of competition for students. Libraries are increasingly being challenged to provide compelling evidence that directly links information resources and engagement activities to positive outcomes for their clients. Libraries that do not provide such evidence will be at an increasing risk of having their funding reduced or eliminated (Jantti & Cox, 2013) and (Soria, Fransen & Nackerud, 2013).

Another problem facing academic libraries is the lack of ongoing access to up-to-date business intelligence that can be used to identify which segments of their client base are not using their resources and services (faculty awareness and knowledge of student use of library information resources is equally limited). Without this information, academic libraries cannot identify in an accurate and timely manner who to target through communications and engagement initiatives, or whether such activities have increased their market share of information consumers (Soria et al. , 2013). Sound marketing can improve a library's value proposition, and being able to articulate strong evidence of value enhances a library's ability to effectively market itself. Consequently, a lack of evidence linking clients' library usage to positive client outcomes can lead to a cycle of devaluation; as poor evidence of value leads to weak marketing impact, leading in turn to poorer usage, which then reduces the value offered by the library as the cost per client increases. Breaking this cycle is paramount to the long term survival of academic libraries. For both problems the issue is a lack of ongoing valid and reliable data being collected from which generalisations can be made about the value provided by the Library.

UWL, through the Library Cube, has produced the information it needs to unambiguously demonstrate the contribution it is making to institutional learning, teaching and research goals. The Cube centres on the integration and interrogation of a series of discrete datasets, e.g. student performance, student attrition, student demographic data, and borrowing and electronic resources usage data. The Cube illustrates the correlation between usage of Library resources and academic performance and demonstrates that those students who are non or low users of UWL information resources are at an academic disadvantage

Creating the Library Cube

The University of Wollongong (UOW) has a Performance Indicators Unit (PIU) to provide senior staff with accessible and integrated reporting and analysis solutions through a secure, web-based Performance Indicators Portal. PIU collects and leverages critical data through scorecard, reporting, and analysis capabilities, providing a one-stop data source for institutional data. In 2009, UWL began working with PIU to develop a tailored database and reporting function — the *Library Cube*— to join library usage data with students' demographic and academic performance data.

The main requirement for joining any two data sets together is that each data set contain a common unique identifier. In our case, all of the systems required to create the Library Cube do contain a unique personal identifier — the student number.

The two UWL data sources fed into the Library Cube include loans data and data relating to usage of electronic resources. Due to the limitations of UWL's Library Information Management System, the only usable data that can be exported that includes the student unique identifier is the total number of items borrowed to date. To build time series data, an export of data occurs weekly, and the difference between two weeks is the borrowing activity that occurred each week.

UWL uses ezproxy logs (sourced from authentication) to extract information on usage of the library's electronic resources. Each time a user accesses certain library resources, an entry is generated in the log file. These resources include subscription databases, e-books, and e-readings materials, which can be accessed via the UWL catalogue or through links provided on UOW's learning management system.

Once exported, the data is contained within an Oracle Data Warehouse with access managed through the IBM Cognos 8.4 business intelligence toolkit. The entire system is developed and maintained by PIU. Users can simply drag and drop the dimensions of the Library Cube, such as gender, faculty, and country of origin, either as a field in a cross-tab table or as a filter value. Users can also select what they wish to measure, such as student head count or total number of borrowings. The web-based system lets users easily manipulate the available dimensions and measures in a myriad of ways, as well as save their views.

Limitations

Before looking at the relationship between usage of our resources and student performance, it is important that the authors acknowledge the following:

- **Borrowing a book does not automatically translate into learning**. Even if the student read the book it does not mean they understood, or used the book. The same logic applies to electronic resources.
- **There are many other factors besides the Library that contribute to students' academic success**, not least of which is academic teaching skills and students' attitudes and aptitude.
- **Correlation does not prove cause**. For example, good teachers might encourage students to use the Library more frequently, and therefore the correlation may be a product of teaching skills, rather than engagement with the collection per se.
- **Other variables that may contribute to students' academic success**, such as attendance, either cannot or have not been captured in the Library Cube due to technical and resource limitations.

There are, however, a number of factors that increase UWL's confidence the validity of the findings, including:

- Sampling error is not a problem as the data is a census, and a census that is updated weekly for loans, and daily for electronic resources.
- There is very little variability within the categories over time.

- There is a very strong relationship between the average marks for each level of resource usage and student grades.

Finally, with a few notable exceptions, the relationship still holds for many views of the data.

How the Library Cube operates:

Although PIU hosts a large historical data set for UOW student demographics and grades, UWL began collecting resource usage data for integration into the Library Cube only in mid-2009 for loans and in January 2010 for online resources. Consequently, the Library Cube only contains time series data starting from 2010.

The logs contain useful information, such as the students' unique identifiers, but they also contain a lot of extraneous data. Counting the number of log entries proved futile as they vary wildly depending upon the online resource being accessed. UWL decided to use the time stamp in the log as a de facto time session measure, using the following business rules:

- The day is divided into 144 10-minute sessions
- If a student has an entry in the log during a 10-minute period, then 1/6 is added to the sum of that student's access for that session (or week, in the case of the Marketing Cube).
- Any further log entries during that student's 10-minute period are not counted.

Using this logic, UWL measures how long students spent using its electronic resources with a reasonable degree of accuracy due to small time periods (10 minutes) being measured.

The primary ethical and legal issue is privacy. The UOW's Privacy Information Sheet outlines the 12 principles to which the university must comply regarding the collection, storage, access, use and disclosure of personal information. Fortunately, there are no legal barriers, as UOW has consent to use personal information for the project through its privacy consent framework (see http://www.uow.edu.au/about/privacy/index.html), which students must agree as part of their enrolment).

At an ethical level, the additional privacy risks potentially posed by the project have been eliminated by managing the personal information in a particular way. Privacy is an issue only to the extent that it involves the use and disclosure of personal information. UWL will not use the Cubes to drill down to see a specific individual's personal information. That is, the data viewed in a Cube will always be aggregated. In all cases, the personally identifiable data that users can glean from a Cube is significantly less than that which can already be ethically and legally obtained through the library management system (LMS), logs, and access to student management systems. Moreover, access to the Cube will be even more restricted than is the case for the other systems that contain the same information.

Resource Use and Student Performance

Examination of the earliest captured raw data showed that there were students who use UWL electronic resources and fail, and there were some non-users that performed strongly. However, all else being equal, the more students used UWL's electronic resources, the more likely they were to perform better. At the time of conducting this level of analysis in 2010 only half of one percent of the high users failed (0.47%), whereas 19% of non-users failed. In other words, non-users were 40 times more likely to fail than high users of library electronic resources. The story was similar for loans, though not as marked.

Further analysis of the aggregated data does reveal a strong correlation between the use of UWL information resources and student grades. For example, the average mark for students who never used UWL electronic resources in 2013 was 59. The average mark for students who spent up to one hour a year accessing UWL electronic resources per year was 64. The following chart shows a

very strong nonlinear relationship between average usage of resources and average student marks (R-squared = 0.82).
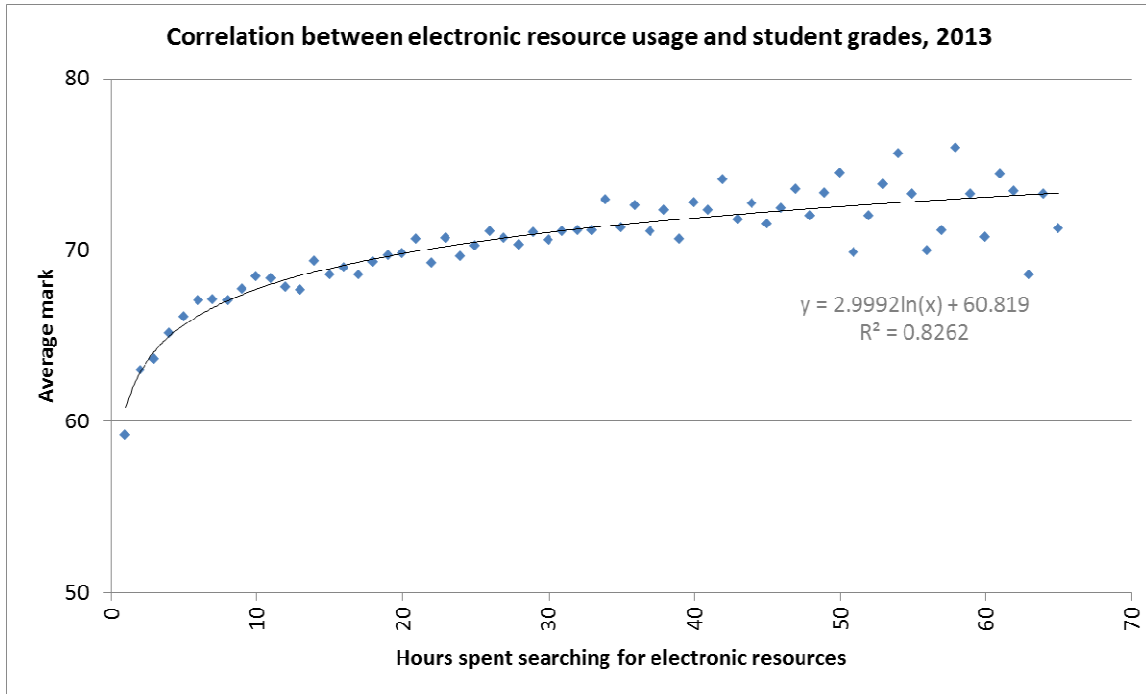

Figure 1. Correlation between Electronic Resource Usage and Student Grades

We defined the first frequency to contain less than 20 students as a cut-off point (to exclude the long tail). These outliers constitute 1.8 percent of the total (579 out of 32,322 students). To apply a logarithmic line of best fit, we incremented each of the frequencies by one. This has no impact on the correlation, the line of best fit's shape, or the points' relationship to each other.

In 2013, the correlation between borrowing and student marks was less striking (R-squared = 0.64). The correlation was not as tight, and the increase in marks with usage was not as steep.
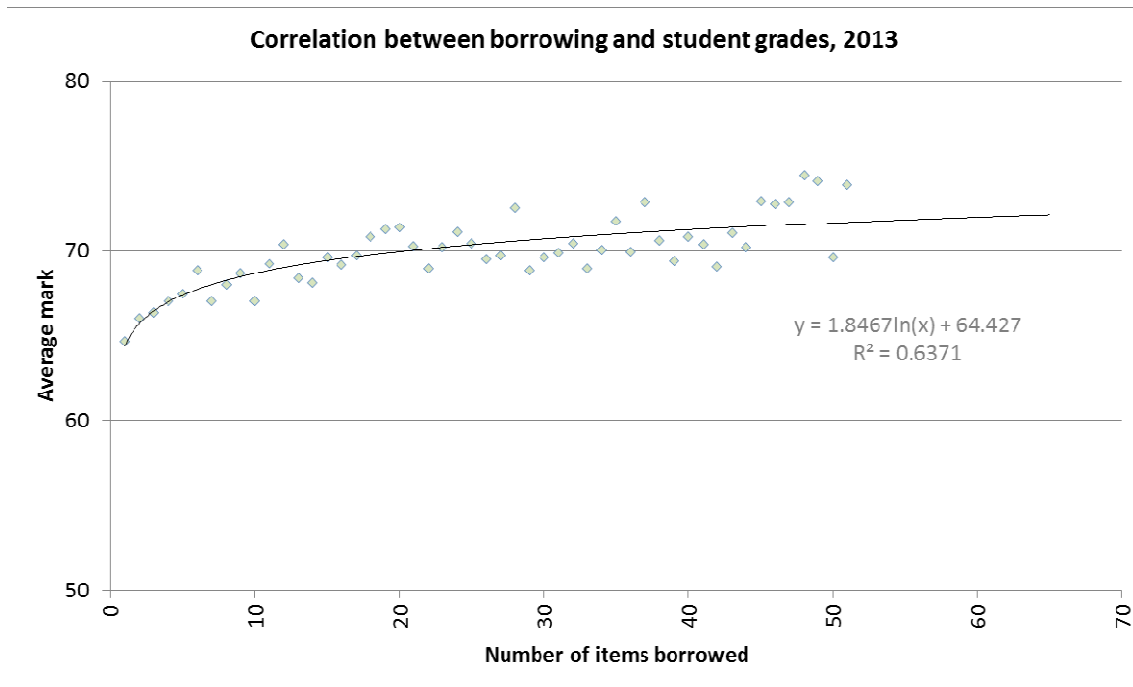
Figure 2. Correlation between Borrowing Materials and Student Grades

Table 1 illustrates the types of reports that both UWL and faculty can access and produce. The views outline the usage trends of undergraduate students from the Faculty of Business as well as the distribution of usage and impact in terms of the weighted average mark (WAM).

Hours spent searching electronic resources by marks and number of students for Business Faculty undergraduates, 2013		
	Marks	Number of Students
No Usage	57.1	1,651
1 - 5	63.6	2,652
6 - 10	67.1	824
11 - 20	68.6	627
21 - 40	70.0	280
41 - 80	76.5	42
Outliers		3
Annual Electronic Resource Usage Frequency (Hours)	**63.6**	**6,079**

Table 1. Weighted Average Marks (WAM) for Undergraduate Students (Faculty of Business) Against Annual Electronic Resource Usage Frequency (Hours)

The Library Value Cube is now part of the UOW enterprise reporting system and went live in May 2012. Representatives from UWL and PIU have met with key stakeholders within the faculties to communicate the purpose, scope, and functionality of its reports and views. Faculty members have

welcomed access to another dataset that helps them better understand the student experience and the impact on student success.

Introducing a new conversation
Although the data illustrates the extent and impact of the Value Cube's utility, it also poses many questions that will require further investigation, such as why some student groups are more successful than others when usage patterns are similar.

For example, domestic[1] female students gain significantly more traction from using resources than their international[2] counterparts. Domestic female students who spent between 81 and 160 hours searching for electronic resources in 2013 scored 21 marks higher than non-users. In comparison, however, the traction for international female students was only 7 marks.

When looking at the difference in academic performance of students who do not use our resources, compared to the heaviest users, the Faculty of Science, Medicine and Health gets the most out of electronic resources. The improvement for using electronic resources is striking, with non-users obtaining an average weighted mark of 45, and steadily rising up to a distinction (76) for students who spent between 81 and 160 hours searching for electronic resources in 2013. The students of the Faculty of Law, Humanities and the Arts receive the highest benefit from books, while the Faculty of Business students receive the least benefit.

Another interesting finding is that the benefit postgraduate students receive from using the UWL collection relative to undergraduates. Postgraduates receive significantly less benefit than undergraduate students. Increased usage of the collection is not having anywhere near the same impact for our experienced students. This finding poses questions on the relevance of current approaches to information literacy support and how these skills could be promoted throughout a student's academic career to improve traction.

The 'Marketing' Cube
The utility of the Library Value Cube is designed to be manifold to provide the information needed to further support continuous improvement in three areas: collection development; academic relationships; and marketing.

The Library spends a significant proportion of its budget subscribing to electronic databases. We are able to obtain third party information (e.g. Counter Statistics) on the number of downloads associated with subscriptions, and we combine this with cost data, to create indices, such as cost per download. The Library uses this information, in consultation with academic staff, to continually improve and develop its collection. There are, however, some major limitations of this data: it is not linked to academic performance; it is highly aggregated, thus not tied to user segments; and it takes too long to get the data.

For example, we will be able to provide academics with the evidence they need to effectively promote the Library to their students. We will also be able to draw on this information in our own instruction activities, to convincingly demonstrate the research behaviours that led to academic success. We will know which specific group we should target to improve take-up. Most importantly, we will know almost immediately whether our marketing and engagement efforts succeeded, which in turn will help us to make informed decisions about whether to change tack, or continue with more of the same.

---

[1] An Australian Citizen, Australian Permanent Resident or a New Zealand Citizen.

[2] A student who is not a domestic student. This includes students on temporary residence visas.

The Marketing Cube will be updated weekly, which will allow us to view in a much more timely fashion how our electronic resources are being used. We will also be able to see at the end of each session, which resources had a significant impact on academic performance, and which resources did not. We will be able to use this information to make more informed decisions about electronic resource collection development and to identify and replicate the processes that led to specific resources facilitating higher academic performance.

Conclusion

The ability to demonstrate the value and impact of libraries and their collections is becoming all the more important. Not only do we need to convince the university executive and faculty of the value of libraries; our most challenging audience is increasingly that of the student body. We now have a compelling story to share: based on the data generated to date, students who access library resources *do* outperform students who don't. The Library Cube, therefore, provides a new facet to better understand the total student experience.

The time required to establish the problem statement, business rules and reporting requirements has been lengthy, but ultimately worthwhile. The UWL Cube, the University of Huddersfield Library Impact Data and the University of Minnesota projects all provide insight on the types of datasets that can be considered and collected, identify key collaborators (e.g. PIU) and the business rules or algorithms that can be applied. Libraries interested in pursuing a similar project are encouraged to talk to their enterprise data centres or units to identify the potential for linking discrete datasets. Enquiries and requests for further information on the UWL Cube can be directed to the authors.

References

Association of College and Research Libraries, (2010). *Value of academic libraries: A comprehensive research review and report*. Researched by Megan Oakleaf.
Chicago: Association of College and Research Libraries, 1-182. Retrieved April 28, 2014 from http://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value/val_report.pdf

Counter – Counting Online Usage of Networked Electronic Resources. Retrieved April 14, 2014, from http://www.projectcounter.org/

Jantti, M., & Cox, B. (2010). Measuring the value of library resources and student academic performance through relational datasets, *Proceedings of the Library Assessment Conference: Building Effective, Sustainable, Practical Assessment*, Baltimore, Maryland, 25-27 October 2010. Retrieved December 1, 2011 from http://libraryassessment.org/bm~doc/proceedings-lac-2010.pdf

Jantti, M., & Cox, B. (2013). Measuring the Value of Library Resources and Student Academic Performance through Relational Datasets. *Evidence Based Library and Information Practice*, 8(2), 163-171. Retrieved April 28 from http://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/19574/15244

Neal, J.G. (2011). Stop the madness: the insanity of ROI and the need for new qualitative measures of academic library success, *Proceedings of the ACRL 2011 Conference, A declaration of Interdependence*, Philadelphia, March 30 - April 2. Retrieved December 1, 2011 from http://www.ala.org/ala/mgrps/divs/acrl/events/national/2011/papers/stop_the_madness.pdf

Soria, K.M., Fransen, J., & Nackerud, S. (2013). Analyzing Demographics: Assessing library use across the institution, *Libraries and the Academy*, 13(2), no. 2, 131-145. DOI: 10.1353/pla.2013.0017

Soria, K.M., Fransen, J., & Nackerud, S. (2013). Library Use and Undergraduate Student Outcomes: New evidence for student's retention and academic success, *Libraries and the Academy*, 13(2), 147-164. DOI: 10.1353/pla.2013.0010

Stone, G., Ramsden, B. (2013) Library Impact Data Project: Looking for the Link between Library Usage and Student Attainment, *College and Research Libraries*, 74(6), 546-559.

UOW (2010) Privacy Information Sheet – General. Retrieved April 28, 2014 fromhttp://www.uow.edu.au/content/groups/public/@web/@fin/@lcu/documents/doc/uow061010.pdf