Joining the Borg: Creating the Library Cube

Dr. David Evans • Erik Bowe • Dr. Linda Golian-Lui
Joining the Borg: Creating the Library Cube

Kennesaw State has attempted to replicate the Library Cube initiated at the University of Wollongong (Cox and Jantti). The Library Cube is based on a multidimensional data warehouse that joins library usage data and student demographic data and students marks.
Discovering the Impact of Library Use and Student Performance

Brian Cox and Margie Jantti  Wednesday, July 18, 2012

*Brian Cox* is Manager Quality and Marketing, and *Margie Jantti* is University Librarian at the University of Wollongong Library.
What the Collective is doing today

• Case study of how Kennesaw State replicated the Library Cube at University of Wollongong
• Our ongoing efforts at building our unique data cube
• What software was used to develop the data cube
• Results and comparisons
• Future efforts
Some Comparisons

The Borg

The *Borg* is defined as the collective intelligence comprising all members linked together.

The Data Cube

The *data cube* is a three dimensional array of values, commonly used to describe a time series. The array can be linked to other data arrays.
Collecting Reliable Data

Difficult to quantify how the library adds value to:

Student Progression Rate
Student Retention Rate
Student Graduation Rate

Difficult to collect reliable data

Difficult to validate the data we did have

Difficult to demonstrate the intrinsic value and the ROI of the library
KSU and SPSU Pre-Consolidation

KSU:
- Founded in 1964
- Headcount: 25,714 (as of Fall 2014)
  - Graduate: 2,122
- FTE: 22,971
- Carnegie Classification: 4-Year Public, Master's Colleges and Universities (larger programs)
- 21:1 Student to Faculty Ratio
- Large suburb setting
- 58% female to 42% male

SPSU:
- Founded in 1948
- Headcount: 6,786 (as of Fall 2014)
  - Graduate: 815
- FTE: 5,935
- Carnegie Classification: 4-Year Public, Master's Colleges and Universities (medium programs)
- 21:1 Student to Faculty Ratio
- Small city setting
- 21% female to 79% male
- STEM focused
The New KSU

• Headcount: 33,253 (as of Fall 2015)
  – Graduate: 2,742
• Comprehensive State University with light research activity
• Multi-campus service environment, including four library facilities
• Approximately 50% female and 50% male
• Just over 50% of programs are STEM focused
Uses of the Cube

• Facilitates Business Intelligence
• Aggregates data
• Allows for flexible analysis of diverse data sets
Some Comparisons

The Borg

The Data Cube
What our data looked like - Legos
Initial Library Efforts
Library Data Cube
A Collective Waiting to be Assimilated

• Horace W. Sturgis Library at KSU utilizes EZproxy for authenticating and tracking access to online resources/databases by NetID
• EZproxy logs contain fields quintessential to creating usage analytics and analyzing the “hive”
• KSU maintains the log files for up to 30 days
• 60% of the students log into EZproxy each semester allowing for a variety of potential data mining initiatives
  – This includes integrating EZproxy log data using multiple data sources
How the Borg Built the Cube

Step 1: Assimilation

• A data mart in an Oracle database was built to maintain the EZproxy logs over time
  – The data mart is part of a larger data warehouse
• EZproxy logs are loaded into the mart nightly using SAS Data Integration Studio
  – KSU utilizes the SAS Enterprise Intelligence Suite for Higher Education
  – This occurs after the EZproxy daily log roll
  – Data is appended for performance reasons
EZproxy Process

The EZproxy hosted service provides on- and off-campus access to your e-content and online materials. Remote servers, maintained by OCLC and managed in a secure data centre based in the UK, make it possible. This frees your technical staff from maintaining hardware and software, and ensures your library is always up-to-date with the latest security protocols.

-OCLC
How the Borg Built the Cube

Step 2: Establishing a Collective Conscience

- The library data mart is joined to other data warehouse tables with fields containing student demographic and progression data
  - KSU uses Ellucian Banner ERP
  - The NetID serves as the primary key initially

- SAS OLAP Cube Studio was used to define dimensions, hierarchies, and measures
  - Examples include institutional GPA ranges and grade distributions
  - The cube is built at the end of the process
How the Borg Built the Cube: Understanding the Cube

- EZProxy Logs
- Circulation
- Card Swipe Data
- Banner Data

Diagram showing a cube divided into regions and products over days.
How the Borg Built the Cube

Step 3: Conquering the Universe

• The SAS cube is then surfaced for end-user consumption in SAS Web Report Studio
  – SAS WRS is a thin client web-based reporting tool
  – Thick client SAS tools can also use the cube for statistical analysis

• Trend analysis and assessments from the cube’s data help support institutional and library effectiveness
  – Use of data visualization tools like SAS Visual Analytics is possible
The purpose of this report is to show the number of students using online library resources by Cumulative KSU.

To limit results, right click on the data table and select Filter and Rank.

Applied filters: Faculty Indicatory Hierarchy equal to N

<table>
<thead>
<tr>
<th>Subject</th>
<th>Number Of Students</th>
<th>Number Of Logins</th>
<th>Average Cumulative KSU Adjusted GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADS–African Diaspora Studies</td>
<td>3</td>
<td>4</td>
<td>2.28</td>
</tr>
<tr>
<td>ACCT–Accounting</td>
<td>317</td>
<td>879</td>
<td>3.45</td>
</tr>
<tr>
<td>AMST–American Studies</td>
<td>50</td>
<td>171</td>
<td>3.30</td>
</tr>
<tr>
<td>ANTH–Anthropology</td>
<td>116</td>
<td>488</td>
<td>3.33</td>
</tr>
<tr>
<td>ARED–Art Education</td>
<td>30</td>
<td>105</td>
<td>3.88</td>
</tr>
<tr>
<td>ARH–Art History</td>
<td>18</td>
<td>42</td>
<td>3.27</td>
</tr>
<tr>
<td>ART–Art</td>
<td>37</td>
<td>132</td>
<td>3.53</td>
</tr>
<tr>
<td>ASIA–Asian Studies</td>
<td>6</td>
<td>21</td>
<td>3.64</td>
</tr>
<tr>
<td>BIOL–Biology</td>
<td>140</td>
<td>355</td>
<td>3.17</td>
</tr>
<tr>
<td>BLAW–Business Law</td>
<td>33</td>
<td>113</td>
<td>3.08</td>
</tr>
</tbody>
</table>
The purpose of this report is to show the number of students using online library resources by classification.

*To limit results, right click on the data table and select Filter and Rank.*

Applied filters: Faculty Indicatory Hierarchy equal to N

<table>
<thead>
<tr>
<th>Classification</th>
<th>Number Of Students</th>
<th>Number Of Logins</th>
<th>Average Cumulative KSU Adjusted GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–Freshman</td>
<td>16</td>
<td>39</td>
<td>3.31</td>
</tr>
<tr>
<td>2–Sophomore</td>
<td>16</td>
<td>30</td>
<td>3.17</td>
</tr>
<tr>
<td>3–Junior</td>
<td>6</td>
<td>12</td>
<td>3.67</td>
</tr>
<tr>
<td>4–Senior</td>
<td>186</td>
<td>752</td>
<td>3.19</td>
</tr>
<tr>
<td>5–Graduate</td>
<td>10</td>
<td>30</td>
<td>3.46</td>
</tr>
<tr>
<td>Subtotal: CHEM–Chemistry</td>
<td>234</td>
<td>863</td>
<td>3.22</td>
</tr>
<tr>
<td>Total</td>
<td>234</td>
<td>863</td>
<td>3.22</td>
</tr>
</tbody>
</table>
Understand our Limitations

◆ Use of electronic resources does not translate into learning

◆ Many factors other than the library contribute to student’s academic success

◆ Correlation does not equal cause

◆ Other variables that contribute to success such as showing up and attending class have not been captured
Findings:
Use of Electronic Resources and Student Performance

Wollongong (UOW)  Kennesaw (KSU)
Table 2. Library Usage by Term GPA and the Number of Students

<table>
<thead>
<tr>
<th>Frequency of Usage</th>
<th>No. of Students</th>
<th>%</th>
<th>WAM</th>
<th>Frequency of Usage</th>
<th>Headcount</th>
<th>%</th>
<th>Term GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>no usage</td>
<td>1,707</td>
<td>7.9%</td>
<td>58</td>
<td>1,450</td>
<td>740</td>
<td>4.1%</td>
<td>&lt;1.00</td>
</tr>
<tr>
<td>1-5</td>
<td>7,316</td>
<td>33.9%</td>
<td>63</td>
<td>2,160</td>
<td>1,044</td>
<td>5.9%</td>
<td>1.00-1.99</td>
</tr>
<tr>
<td>6-10</td>
<td>4,058</td>
<td>18.8%</td>
<td>66</td>
<td>3,552</td>
<td>1,656</td>
<td>9.3%</td>
<td>2.00-2.49</td>
</tr>
<tr>
<td>11-20</td>
<td>4,556</td>
<td>21.1%</td>
<td>69</td>
<td>5,285</td>
<td>2,383</td>
<td>13.4%</td>
<td>2.50-2.99</td>
</tr>
<tr>
<td>21-40</td>
<td>2,923</td>
<td>13.5%</td>
<td>71</td>
<td>11,360</td>
<td>4,944</td>
<td>27.7%</td>
<td>3.00-3.49</td>
</tr>
<tr>
<td>41-80</td>
<td>923</td>
<td>4.3%</td>
<td>73</td>
<td>17,351</td>
<td>7,073</td>
<td>39.6%</td>
<td>3.50-4.00</td>
</tr>
<tr>
<td>81-160</td>
<td>122</td>
<td>0.6%</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>161-320</td>
<td>7</td>
<td>0.0%</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21,612</td>
<td>100%</td>
<td></td>
<td>17,840</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Findings

UOW

Frequency distribution of student marks by level of usage of electronic resources, 2010

KSU

Frequency of usage by semester GPA

Students Cumulative Grade Point Average
Frequency of Usage by Semester GPA

$y = 2181x$

$R^2 = 0.77163$
Resistance is futile: Further Assimilation

At UOW Value Proposition
engagement, access, discovery, and collection development

At KSU Value Proposition
services, facilities, organization, and collections
Resistance is futile: Further Assimilation

- First Year experience
- Progression of students
- Retention of students
- Graduation rates
- Within degree programs usage
- Between colleges
- Gender, age, and race differences
- International students
- Low/poor usage
- Progression from
Further reading while waiting to be assimilated

- *Return on Investment: Libraries and Student Retention* by E. M. Mezick
- “Libraries need to be able to demonstrate how expenditures for resources and services result in significant contributions to academic and social environments that positively impact institutional outcomes such as student persistence.”
- “The strongest relationships found were those between student retention and total library expenditures…”
- Investment in the library leads to increases in student retention.
• *Academic Library Non/low Use and Undergraduate Student Achievement: A preliminary report of research in progress* by D. Goodall and D. Pattern

• Non/low use defined as <5 visits, checking out <5 books, logging in to electronic resources <5 times

• “…the correlation with print and electronic resource usage is present and appears to repeat over time.”

• “Although the actual library usage varies, the graphs nearly all show that students attaining lower grades made less use of the library than students reaching higher grades…”
Must Read

• Capturing Business Intelligence Required for Targeted Marketing, Demonstrating Value and Driving Process Improvement by B. L. Cox and M. Jantti

• Used EZProxy logs to estimate how long system users spent searching for electronic materials

• “…very strong nonlinear correlation between average usage of resources and average student marks (R squared =0.87).”

• “…only half of one percent of the high users fails (0.47%) whereas 19% of nonusers fail.”
Some additional contacts:

Dr. David Evans, Dean and AVP Library Services
devans@kennesaw.edu
Dr. Erik Bowe, Director of Special Projects
ebowe@kennesaw.edu
Dr. Rob Smith, AVP for Institutional Research &
Chief Data Officer rsmit429@kennesaw.edu
Dr. Linda Golian-Lui, Director of the Sturgis Library
lgolian@kennesaw.edu
Mr. Michael Luther, Assessment Librarian
mluther1@kennesaw.edu