Mining EZProxy Data: User Demographics and Electronic Resources

Ellie Kohler and Connie Stovall
Virginia Tech, USA

Abstract
After a mandate to utilize data to demonstrate impact on student success, Virginia Tech University Libraries began diving into previously untapped data sources. Given that the collections budget makes up 48% of the total library budget, roughly 90% of which streams to electronic resources, it was deemed necessary to make more direct connections between electronic resource usage and student success.

Usual practices prior to the charge involved analyzing usage from Counter reports and cost data, such as frequency and cost per use, primarily for the purposes of serials budgeting and negotiations. Due to these past data collection analysis practices, university libraries could only create basic inferences about its electronic resource users. In order to create more robust user inferences, the university libraries turned to EZproxy logs as well as university-collected student data and began a multiphase research project based on the connection between the two data streams.

The long-range purpose of the research project is to create better understanding of student user demographics by connecting electronic resource usage information with university-held student demographic information. Ultimately, plans include impact measurement of the university libraries on Virginia Tech's overall success and constitutes the start of a broader systematic study of the impact of university libraries' dollars spent on electronic resources. Development of this study includes research into encryption and anonymization techniques, as well as current best practices in security of personal information. Discussion will include challenges, including on- and off-campus usage access and meeting resistance to utilizing personally identifiable data. The discussion will also include tools utilized in the study, which include EZproxy, Graylog, Python, and Tableau.

Background and Purpose
Total collections spending typically makes up 37% of an academic libraries' total library expenditures.[1] Virginia Tech's collection spending consumes even more than the average at 48%; electronic resources consume 90% of that collections budget. Given the sheer proportion of funding devoted to electronic resources, it is not surprising that administrators need more data to demonstrate the effectiveness of investments. Libraries, like all other university units, need to map their outcomes to the university's and demonstrate value and impact, and doing so with data is imperative. "More than 2,500 institutions worldwide are currently using Ezproxy," and, for many universities, utilizing usage data from EZproxy creates opportunities to demonstrate value and impact.[2]

Literature Review
Libraries use a variety of methods to demonstrate the impact of their products and services. One university library analyzed the following service points where they also collected corresponding user identification at each: all types of reference questions, circulation transactions, instruction sessions, delivery requests, interlibrary loan requests, and EZproxy logins for off-campus users. While that university had already performed a cost-benefit analysis for most services points, it had not utilized EZproxy data but decided to do so after more pressure to demonstrate impact on student success. After collecting user data, they obtained demographic data corresponding to each user via campus institutional research. Like many investigating impact, they obtained the following: academic standing, academic level, academic program, age, sex, ethnicity, enrollment status, and GPA. Notably, the service demonstrating most use was the EZproxy login data.[3]

McCarthy studied data of 4,803 distance learners enrolled in at least one online class by means of off-campus EZproxy logins and from the college registrar records using the Banner system. The researcher utilized

EZproxy data to determine how frequently distance learning students accessed electronic materials and combined that data with GPAs from Banner. McCarthy found that students who accessed library resources remotely via EZproxy on at least one occasion outperformed those students who never accessed library resources remotely by 15%.[4] Only data from off-campus users of EZproxy was examined in this study.

Allison collected anonymized student identification numbers, their GPA, and class standing for all graduate and undergraduate students for two consecutive academic years. After matching these data to off-campus EZproxy login data and library circulation data, the researcher found that students using library resources demonstrated a usage pattern that increased as students advanced upward through ranks, and that ultimately students who used library resources more exhibited higher GPAs and retention rates. With regard to retention, lowest usage corresponded with those students who left the university prematurely. Further, the researcher discovered a correlation between GPA changes and variations in library activity for both undergraduate and graduate students.[5] Much like McCarthy's study, this investigation did not include on-campus use of electronic resources.

While McCarthy and Allison could not provide data for on-campus users of electronic resources, Davidson et al. included both on-campus and off-campus use of electronic resources as they studied correlation to student GPAs. They chose to include on-campus usage particularly because another study showed that two-thirds of electronic resource usage originated on campus. As with other studies, they retrieved student GPA data via Banner, and in their case, the Office of Institutional Effectiveness anonymized the data for them. Their study showed the percentage of the population who logged in at least once during the semester and the average number of users in that population. Subsequently, the researchers broke down the population by student by college, department, and major; student by level, class, and gender; undergraduate GPA by department and class; student by race/ethnicity; athletic status; Greek status; faculty by department; and staff by department. The analysis of students by college, department, and major revealed that, within the humanities, the classics, English, and religion constituted the three highest usage areas but art students were less likely to log in. In the sciences, students in health sciences used databases more than biology, computer science, and mathematics students. When the researchers looked at undergraduate use of electronic resources and GPA correlation, the data showed that a higher percentage of students logging in and a higher number of logins led to higher GPAs.[6]

Beyond using EZproxy to study impact on student success, at least one study writes of their use of EZproxy and Google Analytics to analyze against COUNTER usage reports and to understand user pathways to electronic resources.[7]

Protecting student privacy continues to be a topic when studying impact of services on students, and libraries exhibit several different methods of ensuring student privacy and disclosing the study to participants. Thorpe et al. revealed that they provided both digital and print patron consent forms and informed all patrons that their participation was optional. All staff were trained to describe the study in the event patrons inquired about it.[8]

Methodology or Study Design

Project Scope

This project involved the collection of data about how Virginia Tech licensed content is utilized by the Virginia Tech student body. The purpose of this program is twofold. The first is to discover whether information about off-campus use can allow reasonable assumptions to be made about on-campus use by comparing study demographics to enrollment demographics. The second purpose is to learn about off-campus student users. This project uses data collected from July 2, 2018, to August 11, 2018, which corresponds with a full academic course of Summer II classes.

Data Retrieval

This study involves database usage through the collection of EZproxy logs. EZproxy is an industry-standard access and authentication software that allows users to gain secure access to web-based, licensed content

that users discover in libraries.[9] This means that Virginia Tech users can access online library materials from off campus. EZproxy is always running, so if the IP address is acknowledged as coming from the university, EZproxy does not request authentication. If the student is off campus, the IP address does not match with the registered range of university IPs. EZproxy then will authenticate the user by requesting their university username and password. Once authenticated, EZproxy passes the user's request to the content provider using an accepted university IP.

At Virginia Tech University Libraries, EZproxy data is automatically collected in a log and held in a server with a Graylog interface. Graylog (www.graylog.org) is a log management system that overlays log data and aids in collection, visualization, and system alerts. This server and its interface has restricted access, which is housed and controlled through the university library's IT division. Graylog is useful to this project as it was relatively simple to use the interface to export various elements of EZproxy data logs in a common .csv format. This overlay also allows the programming of certain parameters to allow for an easier data cleaning and organization process.

From each login, the following variables were collected: date and time of access, database accessed, IP address, geolocation from IP address, personal identifier (PID) if available, and affiliation. On-campus users were screened out, and then unique PIDs were collected and connected with student demographic information via the university's Banner system. This step also eliminated non-students, as PIDs of people who were not enrolled in classes could be reasonably assumed to be faculty, staff, alumni, or another non-student designation. The demographic information collected included: residency (in- or out-of-state), major, college, race/ethnicity, completed hours, gender, age, and overall GPA. This data was completely de-identified.

Python and Relevant Libraries

Thanks to the organization through Graylog, the initial dataset extracted was generally straightforward and clean. But the dataset comprised several gigabytes in size, which made it a large dataset by the library's Data Analytics Team's standards. Because of its large size, this study utilized Python to implement all of the data pre-processing. In addition to being able to handle large datasets, this object-oriented language has a relatively shallow learning curve and excellent online learning resources. The Python programming language was chosen for several reasons: it is one of the top industry-standard languages for data analysis, the researchers in this study had familiarity with it, and, because it can run on a local machine and is non-cloud based, it is ideal for securely handling data with personally identifiable information.

The libraries used within Python for data wrangling for this project consisted primarily of NumPy and pandas. NumPy is a fundamental package needed for data analysis when working with arrays, more specifically arrays of different shapes.[10] Pandas assists with practical manipulation of relational and labeled data in Python.[11] Both libraries are considered core libraries for data analysis and statistics. In order to do some initial visualizations for early data analysis, this project employed the data visualization library Matplotlib.[12] Seaborn, a library based on Matplotlib but affording additional functionality, was also utilized.[13]

Data Pre-processing

One of the useful features of Graylog consists of its ability to automatically pull geolocation data in from non-private IP addresses. However, due to a systems error, the location was not available for the time period needed. To rectify the problem, the data analyst exported only unique IP address information and then used an alternative method to gather the geolocation information. This data was then reintegrated into the original data frame.

Based on the IP address, 24.9% of the over 4 million records were identified as automated system checks, and these were removed before analysis of the data. Other transformations of the data include adding converting timestamp information so that it is a recognized time object, and adding week numbering in addition to months and days.

Once this was complete, the next step involved removing on-campus and admin data, exporting remote users by their university personal identifier, and sending the information to be merged with university student demographic data. The student demographic information was returned completely de-identified by Academic Services Analysis & Reporting, and for reasons of security and privacy, all original datasets that were kept were similarly removed of this personal identifier.

Visualization and Analysis

All visualizations and final analysis utilized Tableau, an end-to-end analytics data platform. Many academic libraries have moved to using Tableau, and it is an industry standard in business and analytics. Tableau integrates querying, exploration, and visualization of data into a single process.[14] Tableau is a favored tool for library analytics at the university libraries at Virginia Tech for a variety of reasons, including the ability to connect to a variety of file and system types; drag and drop technology that requires no coding experience (although it can be helpful); visualizations and dashboards that are publishable to the web via Tableau Public; and dynamic dashboards rather than a static product that allows users to drill down into the data. In addition, the authors of this paper had experience with using Tableau. One author had prior experience and training before her experience at Virginia Tech, and the other was able to attend the Tableau conference with grant money for learning data visualization tools.

Tableau is not the best tool in all circumstances, however. It is fairly expensive, and because of all the capabilities, there is a steep learning curve, especially when it comes to specific customizations. In order to share protected data that should not be made public, there is an additional expense for software such as Tableau Server or Tableau Online. Although this institution makes extensive use of Tableau, there are many other reliable, cost effective options available.

Findings

Overview

For the 31-day time period of Summer II classes, patrons accessed electronic resources approximately 3.4 million times from both on and off campus. While this may sound impressive, automated system checks inflate the total. Consider, too, that research is not a linear process, and often researchers often access resources again and again in a single session.

Of the off-campus users, 5,386 unique users accessed resources with a total accession rate of 1.8 million hits, and an average 329 instances per login. Again, this number seems high, and only reflects each resource page accessed and not actual resources downloaded; it does not reflect time spent on a page. The vast majority of unique off-campus users were students, with 1,166 unique non-student users who are assumed to be faculty, staff, or faculty emeritus. The 4,220 unique off-campus student electronic resource users are in comparison to the 3,741 students who were enrolled in Summer II classes at Virginia Tech.

Initial observations

Off-campus use during the summer was expected to be higher than during the fall and spring academic semesters, and this was the case with 52.5% of users logging in from off-campus locations, 43.7% of users coming from on-campus, and the remaining 0.2% consisted of administrative checks on systems. Fall and spring semester off-campus use averages between 35% and 40% of all users.

During initial observations, the geographic information from IP addresses of both on- and off-campus users was mapped in Tableau. This helped to verify the data was accurate. The expected result was that on-campus users would be located on the Virginia Tech campuses in Blacksburg, Roanoke, and Arlington, Virginia, and off-campus use would be scattered. This indeed was the case; however, there were a few instances of on-campus uses from IP addresses in other parts of the United States and other countries throughout the world. Some of this anomaly can be explained by VPN use by faculty and staff.

On- and off-campus use comparison

One of the main reasons for this pilot study was to investigate whether off-campus users provide an accurate representation of enrolled students during Summer II. Although systems exist to collect off-campus electronic resource use, some library faculty resisted efforts to have users authenticate through EZproxy without regard to the IP location being on or off campus. In order to begin this investigation, it was necessary to compare on- and off-campus use to see where there were similarities and differences that occurred. It is important to note that, while this paper refers to on- and off-campus users, this is intended to mean on- and off-campus usage, without distinguishing whether a user is exclusively off-campus, exclusively on-campus, or regularly uses resources from both on and off campus.

An initial comparison included a look at usage through date and time. The periods of heavy use for both groups occurred late mornings, afternoons, and early evenings. Similarities in patterns of use showed up as well, including less use over the weekends and overall frequency increasing from mid- to end of semester. One notable difference indicated that off-campus users exhibited heavier usage later into the evening than on-campus users.

After looking at date and time, the next focus centered on the actual electronic resources. Little variety presented itself with regard to the number of publishers utilized during the study. The top 24 publishers off-campus users accessed comprised 99.43% of all publishers used during the Summer II session. Similarly, the top 24 publishers that on-campus users accessed comprised 99.89% of all publishers accessed. In addition, a high degree of similarity appeared when looking at the publishers accessed from both on and off campus. All of the top 24 publishers from the off-campus list appeared in the on-campus list, although the rankings could be quite different.

The average rank difference for the top 24 publishers was 5 places, the median was 3.5, and the mode was 3. The top publishers for off campus were: EBSCO (19.9%), JSTOR (11.3%), Science Direct (11.66%), Web of Knowledge (6.27%), and ProQuest (0.58%) (Figure 1). The top publishers for on-campus users were: Web of Knowledge (13.55%), Science Direct (11.6%), EBSCO (9.99%), Ancestry (Library Edition) (9.43%), and ProQuest (7.42%) (Figure 2).

Figure 1: Top 24 Publishers Accessed from Off-Campus

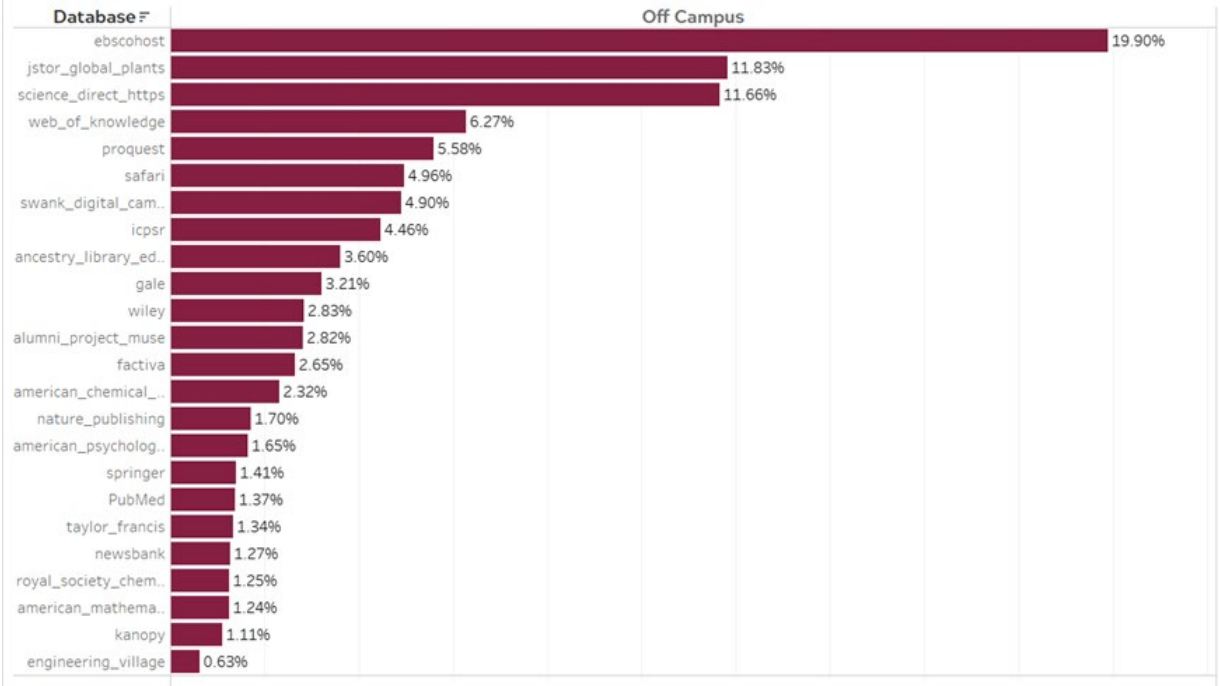
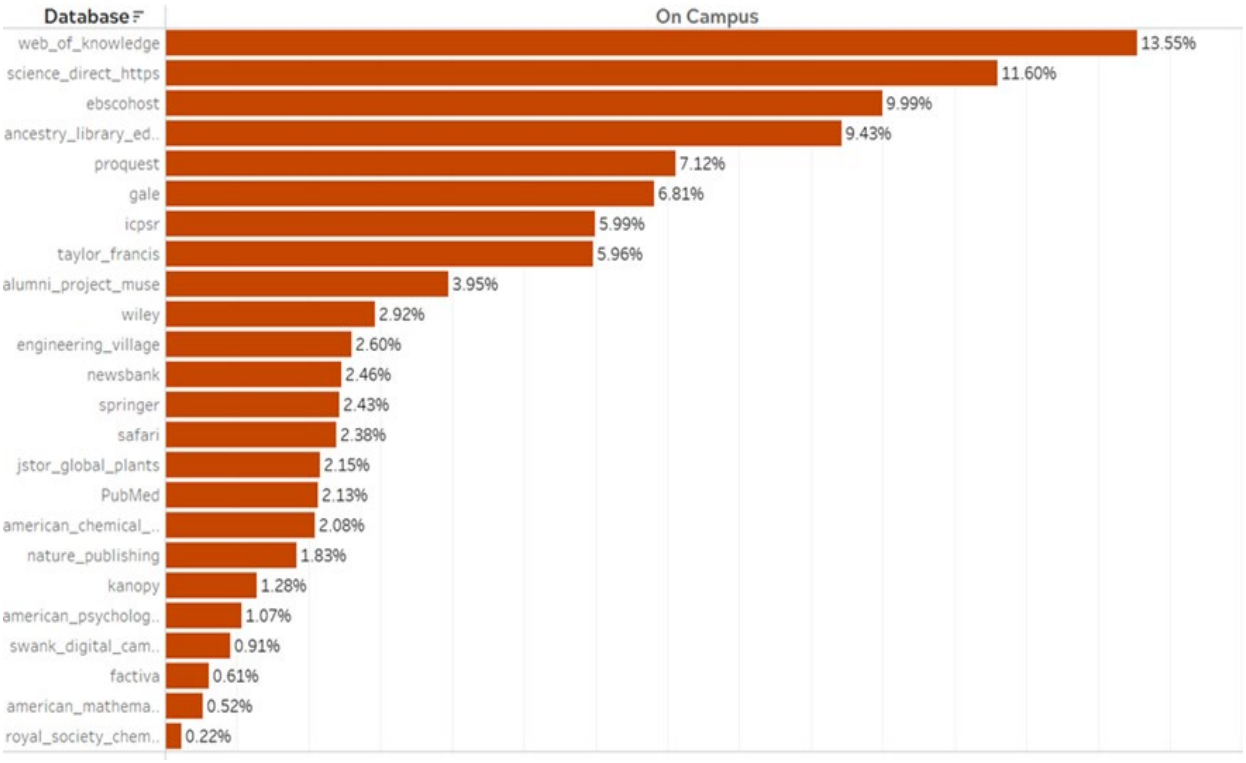


Figure 2: Top 24 Publishers Accessed from On-Campus



Where off-campus users are located: includes students and non-students

An overwhelming number of off-campus student and non-student users, 95.5%, connected to the databases from the United States. The next largest locations of connection include Denmark (0.6%), Japan (0.6%), Germany (0.5%), India (0.4%), and the United Kingdom (0.3%) (Figure 3). In the United States, 73.4% of connections occurred in the same state as the university location, Virginia. The rest of the connections came from every state in the Union, none of which comprised even 5% of the total, even from adjacent states like Maryland and North Carolina or the adjacent District of Columbia (Figure 4).

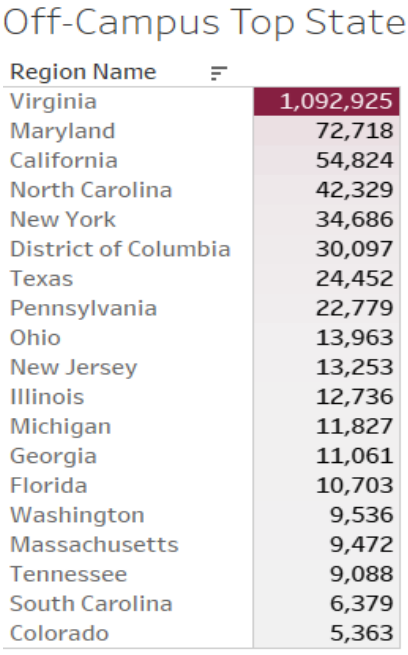
Off-campus and non-student users were most likely taking or teaching classes near campus locations. The highest percentage of off-campus use overall, 58.4%, comes from the city of Blacksburg, Virginia, where the main Virginia Tech campus is located (Figure 5). Surrounding towns such as Christiansburg and Radford are also in the top 20 cities. Other cities and towns in the top 20 include areas where there are satellite campuses such as in the National Capital Region with Arlington, Falls Church, and Alexandria all in the top 20.

Figure 3: Top Countries Where Off-Campus Users Accessed Electronic Resources

Off-Campus Top Countries

Country	
United States	1,566,231
Denmark	9,787
Japan	9,500
Germany	8,087
India	6,513
United Kingdom	4,854
Republic of Korea	4,483
Switzerland	4,353
China	3,981
France	3,848
Canada	3,218
Turkey	2,613
Saudi Arabia	2,274
Peru	1,828
Egypt	1,764
Ecuador	1,483
Bangladesh	1,311
Iran	1,242
Australia	1,220
Kuwait	624

Figure 4: Top States Where Off-Campus Users Accessed Electronic Resources



Off-Campus Top States

Region Name	
Virginia	1,092,925
Maryland	72,718
California	54,824
North Carolina	42,329
New York	34,686
District of Columbia	30,097
Texas	24,452
Pennsylvania	22,779
Ohio	13,963
New Jersey	13,253
Illinois	12,736
Michigan	11,827
Georgia	11,061
Florida	10,703
Washington	9,536
Massachusetts	9,472
Tennessee	9,088
South Carolina	6,379
Colorado	5,363

Figure 5: Top Cities Where Off-Campus Users Accessed Electronic Resources

Off-Campus Top Cities

City	
Blacksburg	572,414
Roanoke	75,301
Richmond	41,957
Washington	29,825
Midlothian	28,743
Radford	27,196
Christiansburg	25,351
Oxon Hill	23,179
New York	22,334
Annandale	21,701
Alexandria	21,580
Virginia Beach	18,788
Arlington	17,889
Ashburn	10,639
Fairfax	9,998
Falls Church	8,148
Raleigh	5,323
Williamsburg	3,141
Blacksburg (Farmvie..	1,244

Off-campus student user profile

Off-campus student electronic resource users had an overall GPA of 3.49 and an average age of 28. The higher age does reflect the higher number of graduate student users, who had an average age of 32. Undergraduate student users averaged slightly older than usual at 22 with a median age of 21. Summer off-campus student users were older than traditional students. Commonly, students range in age from 18 to 21. While Virginia Tech does not publish an average student age or age range, according to collegefactual.com, 64.5% of Virginia Tech students are in the age 18 to 21 bracket, compared to the national average of 60%. Undergraduate students had an average GPA of 3.24. This was then improved by the graduate students' average GPA of 3.74.

Off-campus unique student users and enrollment

The data show that the unique student users who logged into online databases from off campus are closely aligned with Summer II enrollment figures in some areas. The numbers are unequal, with there being more unique student users than enrolled students. This is not alarming, as many students who are not officially enrolled in summer classes may be working on individual projects. In order to make fair comparisons, the data are presented as percentages of the whole, such as percentage of total enrollment versus percentage of total unique student users.

One area where unique student users and enrollment figures are in opposition is in the composition of students. Undergraduates comprise 76.7% of Summer II enrolled students, but only account for 34.5% of unique electronic resource users (Figure 6). Similarly, Summer II enrollment consisted of 16.7% graduate students, and in the same time period, 62% of unique electronic resource users were graduate students (Figure 7). Among electronic resource users, the large percentage of graduate students is roughly equal across all colleges. This is not an unexpected result. Logically, it can be posited that graduate students are called upon to do more frequent and more in-depth research than undergraduate students. They have also had a longer time to familiarize themselves with the library and doing research for electronic resources. Institutional interlibrary loan statistics for the past three years corroborate this assumption; graduate students order over five times more articles than undergraduate students.

Figure 6: Student Use to Enrollment Comparison by Level

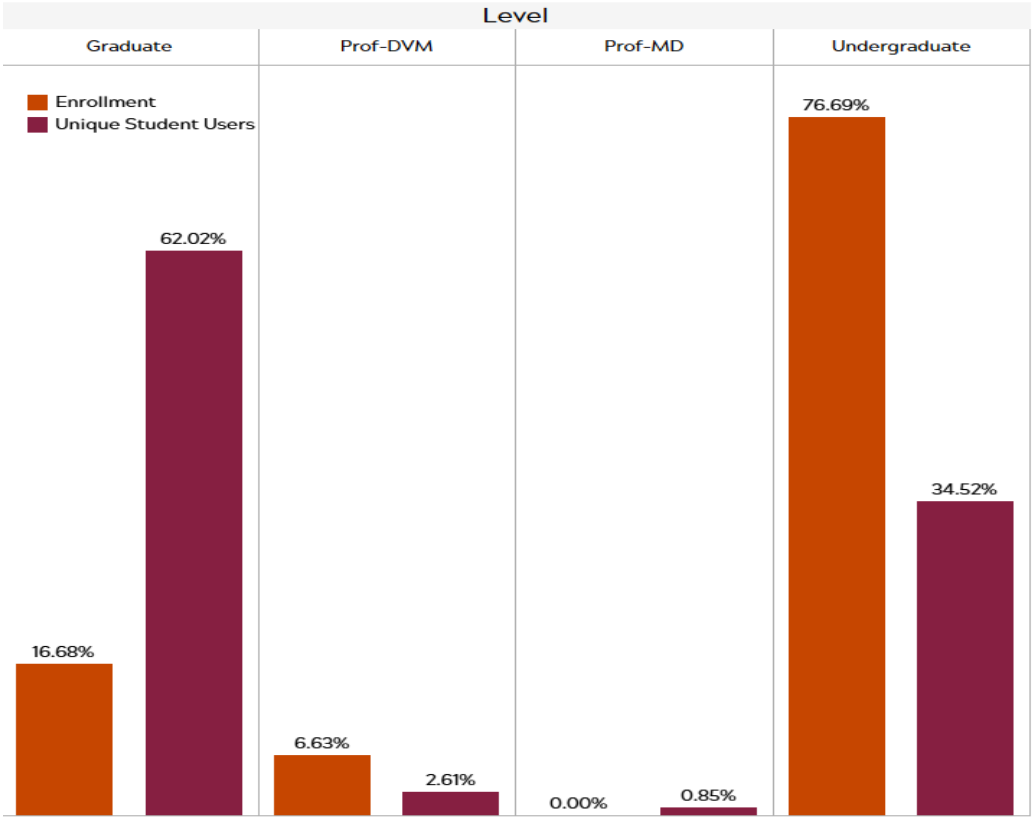
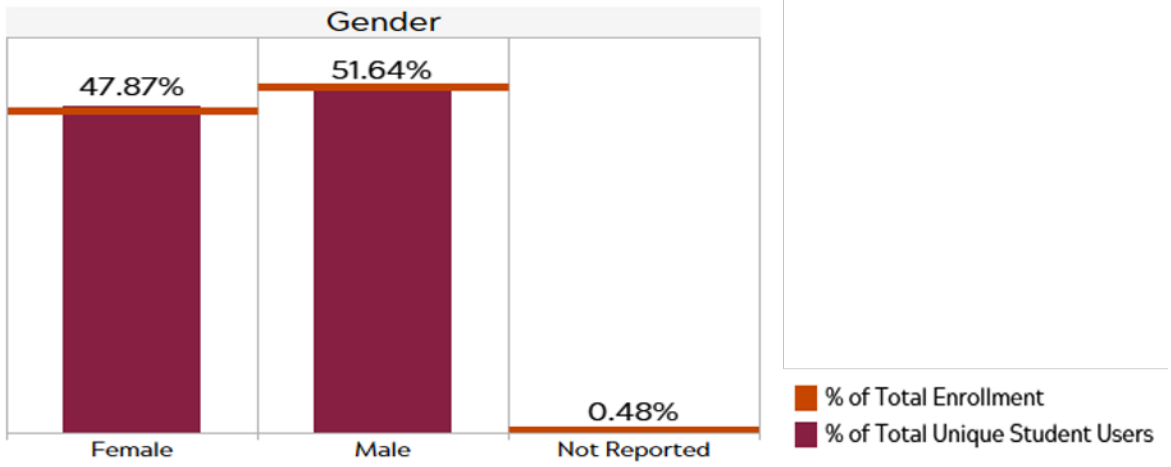


Figure 7: Distance between Student Use and Enrollment

Level	% of Total Enrollment	% of Total Use	Difference Between
Graduate	16.7%	62.0%	-45.3%
Prof-DVM	6.6%	2.6%	4.0%
Prof-MD	0.0%	0.9%	-0.9%
Undergraduate	76.7%	34.5%	42.2%

Both gender and ethnicity are voluntarily self-reported by students at Virginia Tech. Gender figures among unique electronic resource users correlate very closely to enrollment percentages, with surprisingly small percentage differences of (unique-enrolled) 0.87% female, 0.6% male, and 0.27% not reported (Figure 8). In the area of race and ethnicity, every category of unique electronic resource users fall short of the enrollment percentages, with the exception of students who identify as Asian. Because the library traditionally does not gather ethnicity information on its users, it is impossible to see if this is consistent with past library use, much less speculation on the implications.

Figure 8: Gender Comparison of Unique Off-Campus Students and Enrollment Figures



Comparing enrollment to usage by college, only two colleges exceed a 3% gap between enrollment and usage (Figures 9, 10). One of these colleges is engineering, which is overrepresented by electronic resource users, with 23.1% of all students enrolled in classes and 28.4% of all unique library users who are engineering students. The biggest discrepancy is in the College of Business. It comprised 15.9% of all enrollment, yet unique students only accounted for 6% of the total electronic resource users who were aligned with the College of Business. Average GPAs were similar between the two colleges, as well, although slightly higher for engineering (3.57) than for business (3.49). As noted before, graduate students account for the majority of off-campus electronic resource use. Comparing students from the Pamplin College of Business to the College of Engineering, the percentage of graduate to undergraduate students who used electronic resources were both roughly 60% undergraduate to 40% graduate.

Figure 9: Off-Campus Student Usage by College

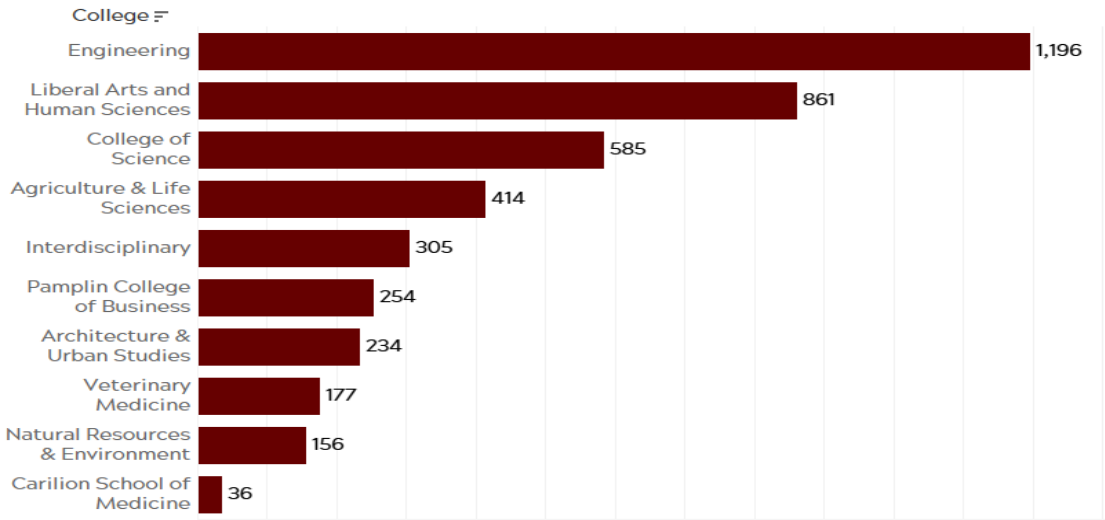


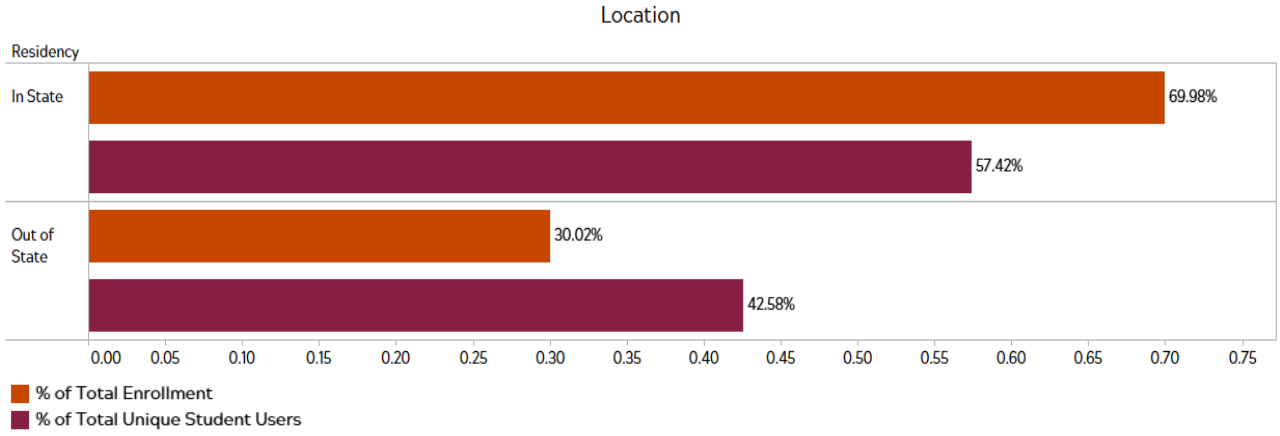
Figure 10: Distribution of Difference between Off-Campus Usage and Enrollment by College

College	% of Total Enrollment	% of Total Unique Student Users	Difference
Agriculture & Life Sciences	7.5%	9.8%	-2.3%

College	% of Total Enrollment	% of Total Unique Student Users	Difference
Architecture & Urban Studies	6.1%	5.5%	0.6%
Business	15.9%	6.0%	9.9%
Carilion School of Medicine	0.0%	0.9%	-0.9%
Engineering	23.1%	28.4%	-5.2%
InterCollege	5.9%	7.2%	-1.3%
Liberal Arts and Human Sciences	17.9%	20.4%	-2.5%
Natural Resources and Environment	2.7%	3.7%	-1.0%
Science	13.8%	13.9%	-0.1%
Veterinary Medicine	7.0%	4.2%	2.8%

Both off-campus student users and enrollment figures had a majority of students with a primary residence within the state of Virginia (Figure 11). In-state enrollment percentages are much higher at 69.98% than off-campus users at 57.42%. The higher percentage of off-campus out-of-state users may reflect the number of students who are not enrolled in summer classes yet still engage in research and academic study or are distance learning students.

Figure 11: Comparison of Location



The evidence is inconclusive to determine, based on this information, whether off-campus users are an accurate representation of on- and off-campus student electronic resource use in Summer II. While the gender balance is accurate, and the data look promising, without additional information about on-campus use, it would be overreaching to make a determination.

Recommendations based on findings

Further studies

As stated previously, without access to on-campus usage demographic information, it is difficult to come to a conclusion as to whether off-campus electronic resource usage data is an acceptable proxy for all electronic resource usage data and to create actionable recommendations. Therefore, further study is recommended.

In absence of on-campus EZproxy electronic resource usage data, a particularly useful study would be an in-depth, qualitative study that follows how students use electronic resources regardless of location. Pairing a qualitative study with the data from this quantitative study would allow a more holistic view of how students use library electronic resources. Coupled with the off-campus usage data, it would create a broader overview of patterns of use.

Although quite a bit of demographic information about unique users exists, it would also be beneficial to discover if connections exist between frequency of use and markers of student success, such as GPA, retention, and/or graduation rates. This extension of the original study would come closer to the ultimate goal of linking library use to student success.

Discussion

Limitations of the Data

It is important not to overstate the results of this study, due to a demographic analysis that consists sole of data from off-campus users. We can guess that the off-campus users are also on-campus users, but to what extent? In addition to the number of users who access electronic resources both on and off campus, it is difficult to know how a student's behavior changes when they are off campus versus when they are on campus.

Another limitation of the data is that, once the resource user information is connected with student demographics, the data is completely anonymized upon return. While this is beneficial when adhering to privacy protocols, it means that, at the current time, there is lacking a way to connect frequency or location information to demographic data. For instance, while the data show that unique engineering students make up the largest percentage of users, do they also access the resources more often? Having this information would lead to a clearer picture of our student users.

Patron privacy and ethical data collection

During the course of this project, there was much concern from library faculty at Virginia Tech about this project and the implications on patron privacy. Due in part to these apprehensions and other logistic practicalities, the full project involving requiring on-campus users to log in through EZproxy is not likely to move forward in the near future. However, this expressed unease underscores the importance of communicating with library faculty and all interested parties about what protocols and extensive security measures were used for collection of data and the security of personal information.

The original encryption procedures to ensure patron privacy involved sending student personal identifiers through two separate cryptographic processes. The first was to subject the identifiers to a random cryptographic process called "salt" that uses randomly generated additional input for each identifier. This new character string would have then been "hashed" to further de-identify the character string. Instead, when working with the Virginia Tech Office of the Registrar, there were university-level processes already in place regarding student information and anonymization protocols. Researchers underwent a thorough vetting process to receive permission to submit unique personal identifiers through Academic Services Analysis and Reporting in order to connect them with student demographic information. The information that came back was completely de-identified, as mentioned previously. While slightly frustrating at the time, placing the onus of anonymization protocols on the Registrar's Office actually did much to alleviate library faculty concerns.

This study also brought to light how much information about electronic resource use is available through EZproxy logs. Library IT, now fully aware of all the data the EZproxy logs collect, are in talks about whether to retain this information and, if so, what security protocols are needed for this data.

Further Analysis

Steps have already been made to collect electronic resource usage through EZproxy data that has occurred during the fall 2018 semester. At Virginia Tech and most universities, summer semester sessions are different from fall and spring semester sessions in a variety of ways. Using fall data will give a more accurate representation of an average user's experience. It will also be valuable to see how fall usage and summer usage data differ and at what points they converge.

Additionally, initial explorations have been made into how EZproxy information can be compared with traditional COUNTER reports. Knowing how those two data streams can inform each other would help with active decision-making in collection assessment.

Endnotes

1. Matt Enis, "LJ Study: Electronic Resources Continue Steady Gains in Academic Libraries," ***Library Journal*** (March 3, 2018), https://www.libraryjournal.com/?detailStory=lj-study-electronic-resources-continue-steady-gains.
2. M. Baikady, A Jessy, and K. S. Bhat, "Off Campus Access to Licensed E-resources of Library: A Case Study," ***DESIDOC Journal of Library & Information Technology*** 34 no. 6 (2014): 486–490, https://doi.org/10.14429/djlit.34.6.7509.
3. Angie Thorpe, Ria Lukes, Diane J. Bever, and Yan He, "The Impact of the Academic Library on Student Success: Connecting the Dots," ***portal: Libraries and the Academy*** 16, no. 2 (2016): 373–392.
4. Sandra Calemme Mccarthy, "Exploring Library Usage by Online Learners with Student Success," ***Community College Enterprise*** 23, no. 2 (2017): 27–31.
5. DeeAnn Allison, "Measuring the Academic Impact of Libraries," portal: Libraries and the Academy 15 no. 1(2015): 36–39, https://doi.org/10.1353/pla.2015.0001.
6. Karen S. Davidson, Stephanie Havron Rollins, and Ed Cherry, "Demonstrating Our Value: Tying Use of Electronic Resources to Academic Success," ***Serials Librarian*** 65, no. 1 (2013): 74–79, https://doi.org/10.1080/0361526X.2013.800630.
7. Jamane H. Yeager, "Using EZproxy and Google Analytics to Evaluate Electronic Serials Usage," ***Serials Review*** 43, no. 3/4 (2017): 208–215, https://doi.org/10.1080/00987913.2017.1350312.
8. Thorpe et al., "The Impact of the Academic Library on Student Success," 279.
9. "EZProxy: Connect Your Users to E-Content with a Single Sign-On," OCLC, accessed December 1, 2018, https://www.oclc.org/en/ezproxy.html.
10. "About NumPy," NumFocus, accessed December 1, 2018, http://www.numpy.org/#.
11. "Python Data Analysis Library version 0.24," NumFocus, accessed January 10, 2019, https://pandas.pydata.org/.
12. "Matplotlib Version 3.0.2," Matplotlib Development Team, last modified November 11, 2018, https://matplotlib.org/.
13. "Seaborn: Statistical Data Visualization," Michael Waskom, accessed December 1, 2018, https://seaborn.pydata.org/.
14. Sarah Anne Murphy, "Data Visualization and Rapid Analytics: Applying Tableau Desktop to Support Library Decision-Making," ***Journal of Web Librarianship*** 7, no. 4 (2013): 465–476, https://doi.org/10.1080/19322909.2013.825148.

Bibliography

Allison, DeeAnn. "Measuring the Academic Impact of Libraries." ***portal: Libraries and the Academy*** 15 no. 1(2015): 36–39. https://doi.org/10.1353/pla.2015.0001.

Baikady, M., A Jessy and K. S. Bhat. "Off Campus Access to Licensed E-resources of Library: A Case Study." ***DESIDOC Journal of Library & Information Technology*** 34, no. 6 (2014): 486–490. https://doi.org/10.14429/djlit.34.6.7509.

Davidson, Karen S., Stephanie Havron Rollins and Ed Cherry. "Demonstrating Our Value: Tying Use of Electronic Resources to Academic Success." ***Serials Librarian*** 65 no. 1(2013): 74–79. https://doi.org/10.1080/0361526X.2013.800630.

Enis, Matt. (2018) "LJ Study: Electronic Resources Continue Steady Gains in Academic Libraries." ***Library Journal***. Last modified March 3, 2018. https://www.libraryjournal.com/?detailStory=lj-study-electronic-resources-continue-steady-gains.

Matplotlib Development Team. "Matplotlib Version 3.0.2." Last modified November 11, 2018. https://matplotlib.org/.

McCarthy, Sandra Calemme. "Exploring Library Usage by Online Learners with Student Success." ***Community College Enterprise*** 23, no. 2 (2017): 27–31.

Murphy, Sarah Anne. "Data Visualization and Rapid Analytics: Applying Tableau Desktop to Support Library Decision-Making." ***Journal of Web Librarianship*** 7, no. 4 (2013): 465–476. https://doi.org/10.1080/19322909.2013.825148.

NumFocus. "About NumPy." Accessed December 1, 2018. http://www.numpy.org/#.

NumFocus. "Python Data Analysis Library version 0.24." Accessed January 10, 2019. https://pandas.pydata.org/.

OCLC. "EZProxy: Connect Your Users to E-Content with a Single Sign-On." Accessed December 1, 2018. https://www.oclc.org/en/ezproxy.html.

Thorpe, Angie, Ria Lukes, Diane J. Bever and Yan He. "The Impact of the Academic Library on Student Success: Connecting the Dots." ***portal: Libraries and the Academy*** 16, no. 2 (2016): 373–392.

Waskom, Michael. "Seaborn: Statistical Data Visualization." Accessed December 1, 2018. https://seaborn.pydata.org/.

Yeager, H. Jamane. "Using EZproxy and Google Analytics to Evaluate Electronic Serials Usage." ***Serials Review*** 43, no. 3/4 (2017): 208–215. https://doi.org/10.1080/00987913.2017.1350312.