The Pennsylvania State University

The Graduate School

College of Information Sciences and Technology

EVALUATING WEB REAL ESTATE VIA PIXEL EFFICIENCY ANALYSIS

A Thesis in

Information Sciences and Technology

by

Alex T. Brown

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 2016

The thesis of Alex T. Brown was reviewed* and approved by the following:

Bernard J. Jansen
Professor of Information Sciences and Technology
Thesis Adviser

Jessie Li
Assistant Professor of Information Sciences and Technology

Luke Zhang
Associate Professor of Information Sciences and Technology

Andrea Tapia
Associate Professor of Information Sciences and Technology
Director of Graduate Programs

*Signatures are on file in the Graduate School.

ABSTRACT

This research presents a quantitative web analytics approach tailored for academic libraries. Specifically, we introduce *pixel efficiency analysis*, with the associated measures of *pixel efficiency value* and *conversion efficiency value*, as a web analytic approach for evaluating potential website changes. Pixel efficiency analysis is the practice of relating screen real estate measured in pixels to the achievement of organizational goals and key performance indicators as indicated by quantifiable user behavioral interactions on a webpage. We employ both the concept and measures through a case study focusing on high traffic webpages of an academic library website for a major research university. An overarching web analytics investigation in combination with pixel efficiency analysis of four of the library's major webpages identifies the key areas of improvement in regards to real estate usage and provides quantifying numbers to support the improvement. Based on these results, we investigate changes to each of the four pages utilizing A/B testing of tens of thousands of library patrons and the measurements of pixel efficiency value and conversion efficiency value to examine the effect on user behaviour, demonstrating the value of pixel efficiency analysis. Our research findings show the capability of pixel efficiency analysis to provide insight not delivered by existing web analytics approaches for academic libraries. Namely, we emphasize the importance of page real estate by showing that components of a webpage can be optimized and that users overall prefer the optimized web layouts. Real estate usage is expected to be increasingly important given the trend towards mobile, and it is an increasingly important consideration within web analytics and design. While specifically tailored to academic libraries, pixel efficiency analysis has applications to all websites and has significant potential for future research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACKNOWLEDGEMENTS

I would like to thank Penn State University Libraries, where I completed a Research Assistantship during year one of my master's program. Since the completion of the assistantship, I have stayed on with PSUL and worked as a Web Analytics Specialist. My time with PSUL has allowed me to vastly expand my knowledge of digital analytics and user experience. It is my hopes that my contributions will allow PSUL to become a leader among academic libraries within the field of web analytics.

I would also like to thank Binky Lush, who has been my supervisor at the libraries for the past two years. Her management abilities allowed me to work in an efficient and effective manner, while her communication skills kept me up-to-date on everything going on with the libraries, which allowed me to stay on task.

In addition, I could not have completed this thesis without the help of Dr. Jim Jansen, who has served as my thesis adviser. Dr. Jansen continually gave prompt, helpful feedback throughout my stint at Penn State and held me to high standards, which pushed me to excel. But beyond keeping me on task to finish the Master's program here, he was also always concerned about how I was doing. While it may seem small, this meant a lot to me. Throughout the stresses of a graduate program, having an adviser who cares about you beyond a business standpoint makes things much easier.

Lastly, I would like to thank all of the staff and faculty who are part of the College of Information Sciences and Technology at Penn State. Everybody I have had the pleasure of interacting with during my time here has been nothing but helpful. Everybody here wants to see you succeed, and is well equipped to do so. I simply cannot say enough about the faculty and staff here.

Sincerely,

Alex Brown

CHAPTER 1: INTRODUCTION

Advances in technology are continually shifting the online landscape for academic libraries, as well as other organizations, especially in regards to offerings of online capabilities. Academic library sites offer resources, such as e-journals, e-books, enhanced search features, and virtual reference services (Aharony, 2012). Yet, these increased online services come at a cost. Increased electronic services have pushed expenditures of academic libraries at a growth rate 12% above the inflation rate, while decreasing the amount of physical assets offered (Regazzi, 2012). Libraries must, however, invest in these services to offer the high quality online services today's consumers (e.g. professors, students, and staff) have come to expect.

The increasing importance and investment in web services offered by academic libraries is apparent, but this increase in online offerings comes with challenges, namely high quality of service expectations by library patrons. These expectations formed by consumers have largely been a result of commercial non-library offerings, in particular services like Google Scholar (Kesselman & Watstein, 2005). These levels of service concerns by libraries are not unwarranted.

For example, Brophy and Bawden (2005) show that Google Scholar, in comparison to most library services, ranks superior in accessibility and coverage when conducting academic-related queries. While library services rank superior in quality of search results, the researchers (Brophy & Bawden, 2005) point out that the emerging generation of scholars is likely to prefer accessibility over, perhaps, a marginal quality increase.

It is therefore imperative that academic libraries engage in a process of continual improvement of their offered online services and site features to better support their customer base and bring their services in line with commercial expectations. Libraries must achieve this, while also attempting to better understand the specific information needs of academic consumers, which we propose differs from typical online information seeking.

To better understand the information seeking behaviors of academic users, some academic libraries have turned to the practice of web analytics. The reported results, though largely exploratory and descriptive in nature, are mixed (Betty, 2009; Black, 2009; Deschenes, 2014; Fang, 2007; Fang & Crawford, 2008; Ghaphery, 2005; Loftus, 2012; Memmott & deVries, 2010; Turner, 2010; Whang, 2007; Young, 2014). By the very definition of web analytics, some sort of optimization and/or enhancement is to be expected, as the definition of web analytics offered is "…the objective tracking, collection, measurement, reporting, and analysis of quantitative Internet data to optimize websites and web marketing initiatives" (Kaushik, 2007c). So, while the descriptive findings of past research have resulted in website improvements, the descriptive nature fails to advance a field of unique information seeking behavior: academic libraries. This notion should not necessarily be taken in surprise, as Kaushik (2007c) notes that web analytics is still in its "infancy" (p. 7).

While we initially mention that 'web analytics' has been used to optimize and enhance academic library websites, we find it important to note that future research in this area may incorporate phrases such as 'digital analytics' and 'web analytics 2.0' (Google, n.d.; Kaushik, 2007a, 2007c). To emphasize the importance of these differences, Kaushik (2007a) defines web analytics 2.0 as "the analysis of qualitative and quantitative data from your business and the competition to drive a continual improvement of the online experience that your customers and potential customers have which translates to your desired outcomes (both online and offline)" (para. 5). This definition differs from the original definition in that it is more complex and demands the need for a holistic approach to data analysis that is synchronized with organizational goals. This holistic approach has been lacking in most web analytic studies previously conducted for academic libraries. By adopting best practices of digital analytics, academic libraries stand to increase their competiveness relative to commercial offerings (e.g. Google Scholar), while also seeing an increase in return on investment (ROI). Adopting more comprehensive data-driven approaches will also inherently allow libraries to increase their expertise within the field of data science, which is particularly important as academic libraries seek to expand research data services to accommodate students within the shifting learning environment (Tenopir, Sandusky, Allard, & Birch, 2014). Increasing competiveness with commercial offerings and also increasing ROI is vital to the future success of academic library websites; but, these cannot be achieved without adopting strategic approaches that embody utilization of multi-methodological approaches that are accompanied by tactical web analytic measures.

Thus, the goal of this research is to propose a strategic framework for academic libraries that can increase the effectiveness and efficiency of academic library website presence. This strategic framework is supported by empirical data from a case study, an academic research library that is a part of a major U.S. university. Research findings highlight the impact of leveraging a holistic web analytics approach (i.e. web analytics 2.0 or digital analytics) for academic libraries by utilizing *pixel efficiency analysis*, an approach we developed specifically for academic libraries and similar organizations. Pixel efficiency analysis inherently directs an analyst to employ best practices of web analytics through a strategic analysis of combined real estate usage and user behavior. Within our research, we (1) employ an overarching web analytics investigation on four major webpages of the library, while making special note of real estate inefficiencies; (2) make iterative changes via A/B testing based upon identified weaknesses and record results; and (3) report the findings through tactical measures that tie back to organizational objectives and key performance indicators (KPIs).

In short, this research seeks to maximize the efficiency of screen real estate utilization, while achieving maximum effectiveness. Screen real estate is defined as, "the amount of space available on a display for an application to provide output" (Usability First, n.d.) and, in our research, is measured in pixels. Pixels can be thought of as a measurement similar to that of square feet or square meters for a room measurement, just in this case the measurement is screen real estate.

Our research findings show that pixel efficiency analysis proves useful beyond that of current measures associated with web analytics typical within academic libraries, while having applicability to websites and organizations in general.

CHAPTER 2: LITERATURE REVIEW

2.1 Web Analytics Overview

The foundation of web analytics can be traced to the field of behaviorism. In contrast to the original field of behaviorism that strictly accounted for outer behaviors, the field of web analytics takes into account both inner experiences and outer behaviors (Jansen, 2009). The way one thinks and feels, then in turn influences how one interacts with the surrounding world. This notion not only exists within the physical world, but it is transferable to the virtual world. People interact with technology and the web in unique manners. Each individual therefore leaves unique behavioral patterns that can be referred to as cognitive footprints (Guidorizzi, 2013). These cognitive footprints can then be viewed via unobtrusive data collection means and allows for a sample size far larger than what a researcher can traditionally capture in a traditional laboratory setting (Jansen, 2009). This data can be used for a multitude of purposes, including: e-marketing, website optimization, personalization, and active authentication (Abramson & Aha, 2013; Brown & Abramson, 2015; Jansen, 2009; Srivastava, Cooley, Deshpande, & Tan, 2000).

Originally, one of the main web analytics techniques for the optimization of websites was transaction log analysis (TLA) (Jansen, 2006; Kaushik, 2007c). As the web became a bigger part of online business, web analytics companies with analytics dashboards, such as Accrue, WebTrends, and CoreMetrics, arose shifting the responsibility of data collection from IT departments to web analytic vendors, thus reducing the need for TLA (Kaushik, 2007c). In 2005, Google purchased Urchin and released Google Analytics (Google, 2014). This introduction caused a shift in web analytics because, in contrast to its competitors, Google Analytics was offered for free, in most cases. The introduction of Google Analytics brought forth rapid growth within the field of web analytics, as various disciplines quickly began to acknowledge its usefulness (Kaushik, 2007c).

Examples as to the specific uses of web analytics is documented by Jansen (2009) in an exhibition of a hypothetical scenario involving a retailer. In the scenario, questions are raised such as, “How do potential customers find our online store? Do they find us via major search engines or from other sites? … If a customer starts to make a purchase but then leaves before completing the order, should we look at a site redesign” (p. 1)? Web analytics provides the capability to answer these questions (Ferrini & Mohr, 2009; Jansen, 2009; Kaushik, 2007c; Plaza, 2009). For example, in reference to the last question, if one of the transaction pages is a top exit page (i.e. a page that many visitors depart the site from without taking the expected action), this may indicate that content on the page is confusing and can steer developers to the page of issue.

Taking web analytics in the context of academic libraries, consider bounce rate (i.e. a single page session that is so short in duration that no reasonable action could occur) on the homepage of a library’s website. The homepage serves as the gateway to all of the library’s electronic resources. A high bounce rate on the homepage indicates that many of the users coming to the site, are perhaps not making use of the resources offered. From a business perspective, this can correlate to low ROI and, for an academic library, is detrimental given increasing expenditures on electronic resources.

Although utilizing web analytics has potential benefits, it does not come without its limitations. For instance, while web analytics may tell how a user is interacting with a website, it is difficult to tell why a user is engaging in a particular behavior. Nor can web analytics identify underlying needs of users or user satisfaction throughout their webpage engagement (Conyers & Payne, 2011; Jansen, 2009). Moreover, web analytics has inherent data accuracy issues and data error margins exist within the 5-10% range (Ferrini & Mohr, 2009; Jansen, 2009).

2.2 Web Analytics in Academic Libraries

Despite the caveats of web analytics, Kumar, Singh, and Kaur (2012) find organizations that utilize web analytics see a significant positive correlation with consumer satisfaction. Several studies have emerged utilizing web analytics in the context of academic libraries. Specifically, academic libraries have utilized web analytics to better understand how the website is being used, increase visitors, improve loyalty, enhance navigation, and advance marketing efforts (Betty, 2009; Black, 2009; Deschenes, 2014; Fang, 2007; Fang & Crawford, 2008; Ghaphery, 2005; Loftus, 2012; Memmott & deVries, 2010; Turner, 2010; Whang, 2007; Young, 2014). The results within academic libraries have been promising, but web analytics is a process of continuous improvement and should build off the past research to evolve and tailor towards the needs of academic libraries.

In other prior work, Fang (2007) utilizes Google Analytics in an attempt to enhance a law library website to better suit visitor needs. The study evidences the usefulness of Google Analytics for academic libraries, as increases in traffic and loyalty are shown. In a similar study, Black (2009) uses TLA to analyze user behavior of a major academic library website. This study exhibits how a web analytic approach can be used to identify what content is of value to consumers. While both Fang and Black are successful in using a web analytic approach for their respective objectives, these studies are primarily descriptive in their findings.

Other studies have informally adopted best practices and encompassed strategic approaches. Loftus (2012), for example, notes that a strategic web analytics program can lead to informed decision making and proceeds to identify KPIs and utilization of multiple technologies and methods. Young (2014) provides another example of strategic web analytics, with emphasis on A/B testing. While these studies are accomplishments in moving towards strategy-based analytics within academic libraries, overarching frameworks and precise best practices are still lacking.

2.3 Best Practices and Application to Academic Libraries

As web analytics transitioned from TLA to enhanced visualized reports offered by vendors, standards and best practices began to emerge to make “analytics professionals more effective and valuable through professional development and community” (Digital Analytics Association, n.d.). In line with that mission, Jansen (2009) offers Figure 1 based upon best practices identified by the DAA.

Fagan (2014) specifically mentions the process outlined by Jansen (2009) and the lack of its adoption by academic libraries, echoing the call for a more strategic approach to website redesign made by Manuel, Dearnley, and Walton (2010). Failing to do so moving forward risks more findings similar to that of Paul and Erdelez (2013), who report that web analytics is underutilized by library management and insist, "[m]ore studies involving the latest analytics solutions in a library setting are required …." (p. 130).

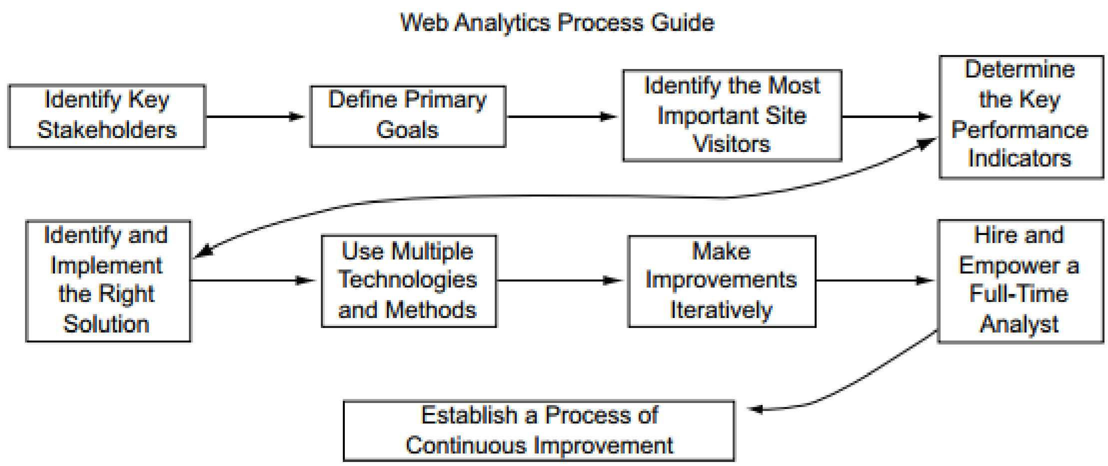

Figure 1. The 'Web Analytics Process Guide' (Jansen, 2009).

2.4 Need for a Quantitative Metric

As requests for new and strategic web analytic approaches are made for academic libraries (Fagan, 2014; Manuel et al., 2010; Paul & Erdelez, 2013), it is also important to understand where potential frustrations may lie. Often, the largest frustration of implementing web analytics is its inability to provide actionable intelligence (Kaushik, 2007c). Rather, the reporting of data itself does not necessarily inform what changes should or should not be done to a webpage. This is potentially exacerbated within academic libraries in contrast to businesses, as businesses have the luxury of making changes and using monetary values to measure KPIs. Turner (2010) discusses this very issue and details potential ways to configure and customize Google Analytics for academic libraries, although they offer no metrics unique for libraries.

While it is true that Google Analytics and similar tools are powerful and customizable, especially when combined with a tag management system (e.g. Google Tag Manager), academic libraries are still left without a monetary value or anything similar in nature to provide a quantitative measure related to KPIs. Looking outside the context of traditional web analytics brings about potential solutions, while placing particular emphasis on the best practice of utilizing multiple technologies and methods.

Specifically, in this research, we explore the aspect of using pixels to replace monetary units of the ecommerce domain and provide a quantitative metric for libraries and other non-commercial organizations wishing to implement changes to their websites while evaluating the potential effect of these changes on organizational KPIs.

Perhaps surprisingly, pixels have had limited use in past research examining webpages or webpage usage. Nicholson et al. (2006) utilize a pixel approach to analyze the space used by the paid placement of search engine ads. By using pixels, they were able to quantify ad-based real estate within search results pages. In an attempt to quantify user behavior, Buscher, Cutrell, and Morris (2009) utilize an eye-tracking approach to highlight page real estate issues. We believe that pixel analysis geared towards quantifying user behavior on a webpage has potential to provide a quantitative metric for academic libraries, as well as other organizations, to measure the effect of webpage changes on achievement of KPIs, and the use of pixels as a quantifiable measure provides a needed element to analytical processes for organizations such as libraries.

The use of pixel analysis as a metric is founded in the notion that every pixel serves as an opportunity to convert a user to a resource or provide a resource to a user. This assertion is not to be confused for suggesting that every portion of a webpage should be clickable. Rather, we suggest that webpage components, along with webpages overall, theoretically speaking, have an optimal size in terms of achieving effectiveness and efficiency. Seeing as academic libraries have various resources available, measuring webpage real estate in terms of efficient usage of pixels, links screen real estate to academic library goals and KPIs. If quantities of real estate on a page are not being efficiently and effectively utilized, this indicates that the space may be better used towards achieving some other organizational goal, allowing academic libraries to provide high value content to users.

Pixel usage can also provide a method for evaluating suggested changes to a webpage, replacing "gut feel" and heuristics. We can tie pixel usage back to Fitts' Law. In order to increase usage of specific components, it may be compelling to enlarge items to make them easier to find and click on, but the relationship between usability and target size is not linear (Karafillis, 2012). Rather, while there are instances where usability can be increased by enlarging an item, bigger is not always better and there reaches a point where the increased usability levels off. Therefore, there will be a certain point that enlarging an item just becomes wasted real estate, while potentially becoming less aesthetically appealing to the user at the same point.

Hence, the motivating questions for this research are: *How can web performance of an academic library website be measured similar to that of a monetary value in ecommerce? What would pixel analysis look like in alignment with a larger web analytics strategy? Can pixel analysis be used to measure performance of a webpage? How might pixel analysis be intertwined with other methodologies? Can pixel analysis be leveraged for analysis of KPI achievement within academic libraries? What improvements might pixel analysis lead to on an academic library's webpage?*

CHAPTER 3: RESEARCH QUESTIONS

Pixel analysis being of value for academic libraries is founded on the assumption that every pixel serves a potential purpose. Purposes on a webpage can range from converting users (e.g. getting users to click on a link) to serving an aesthetically pleasing presence that subconsciously impacts user acceptance. Effectively and efficiently doing so without causing confusion can be cumbersome, as library staff can struggle to appropriately identify what resources should receive respective webpage real estate within a given page. In a sense, the snowball effect applies here.

An unofficial observation made during the time of this study, is that whenever library website stakeholders meet and discuss content management for a webpage, the session often revolves around the compilation of an exhaustive list of resources that should be made available on a page with no metrics to justify the inclusion or exclusion. This creates an overburden of information that comes at the cost of potential confusion to the user. If we assume that an overburden of information causes confusion to the user, then the value of webpage real estate becomes of greater importance, since one needs to then ensure that the webpage does not become "crowded" and determine what components are most important. By using pixel analysis, we can determine what components should receive the most importance and how to optimize the size of individual components, which would then correlate with greater optimization of the webpage and website as a whole.

Concept/ Measure	Acronym	Definition	Usage	Benefit
Key Performance Indicators	KPIs	"A measureable expression for the achievement of a desired level of results in an area relevant to the entity's activity" (The KPI Institute, 2016).	Project and Program Management	Enhances Data Driven Decision Making and Installs Overarching Strategy
Pixel Efficiency Analysis	PEA	The practice of relating screen real estate measured in pixels to the achievement of organizational goals and *KPIs* as indicated by quantifiable user behavioral interactions on a webpage.	Any page that seeks to streamline users and maximize the proportion of high value content to low value content.	Inherently promotes the web analytics best practice of utilizing multiple technologies and methods.
Pixel Efficiency Value	PEV	An accompanying measure of PEA that seeks to maximize the proportion of moderate-heavy usage within a given area of real estate.	Employable at either a componential or page level and is particularly useful whenever large areas of real estate are not being utilized.	Provides a tactical measure that allows performance measurement accounting for effectiveness and efficiency.
Conversion Efficiency Value	CEV	An accompanying measure of PEA, which is weighted in a manner that seeks to maximize conversions and minimize real estate usage.	Employable at a componential level whenever seeking to increase conversions.	Provides a tactical measure that allows performance measurement accounting for effectiveness and efficiency.

Table 1. An overview of concepts and measures important to the practice of PEA.

We label the process of using pixels for webpage evaluation as *pixel efficiency analysis*. Important concepts and measures to the analysis can be viewed in Table 1. As a formal definition, *pixel efficiency analysis* (PEA) is the practice of relating screen real estate measured in pixels to the achievement of organizational goals and *KPIs* as indicated by quantifiable user behavioral interactions on a webpage. The following research questions seek to explore the novel approach of PEA for academic libraries.

3.1 Research Question 1 (RQ1): Does a web analytics investigation of a library's website support the utilization of PEA as a methodology to improve webpage design?

We begin with a comprehensive web analytics investigation because inherent within the best practices is to utilize multiple technologies and methods. This is because each technology and methodology used helps to tell a story with the data. Rather, if we solely rely upon one tool, such as Google Analytics (which several past academic library studies have done), we only get part of the story. Hence, within the examination of RQ1, we will employ Google Analytics, CrazyEgg (a heatmapping tool), and a user survey into order to thoroughly evaluate our use of PEA. Doing so will give us a high level depiction of the user experience that occurs within the library's website, while eliciting existent weaknesses. Will those weaknesses provide validation for the employment of PEA?

3.2 Research Question 2 (RQ2): What impact will the attempt to increase pixel efficiency have on user behavior?

While RQ1 attempts to better understand the website through the insights of the user and identify weaknesses, with RQ2, we seek to correct those weaknesses. Understanding what is going on will not be enough to know for certain what changes should or need to be made. As detailed in Figure 1, an iterative approach is vital to the field of web analytics. We follow that approach here, while attempting to encourage the utilization of data-driven decision making within academic libraries. Specifically, the iterative approach that will be taken within this research will utilize A/B testing in attempt to enhance pixel inefficiencies elicited within RQ1. Though RQ1 may elicit weaknesses and areas for improvement, simply making changes "assumed" to be better is risky. A/B testing helps reduce risk and test changes, while, in conjunction with PEA, relating said changes back to KPIs.

3.3 Research Question 3 (RQ3): How can we report findings from PEA in a manner that enhances the decision making process in alignment with KPIs?

Leading up to RQ3, we seek to depict an overarching image of the library user, while also testing potential website design changes that potentially make more efficient and effective usage of webpage real estate. While we will be taking various measures into account up to this point, no traditional or existing measure takes the value of webpage real estate into account. In turn, for RQ3, we will offer two different measures capable of showing more efficient usage of pixel efficiency. Particularly, we aim to provide measures that combine conversion/engagement levels with the amount of screen real estate being measured by PEA.

CHAPTER 4: METHODOLOGY

4.1 PEA Within the Framework of Web Analytics

This research takes a similar approach to that of Coughlin, Campbell, and Jansen (2013), who modify web analytic techniques to fit the needs of academic libraries for purchasing online content. More specifically, we mold the Web Analytics Process Guide (Figure 1) to a three phase strategic framework, as seen in Figure 2.

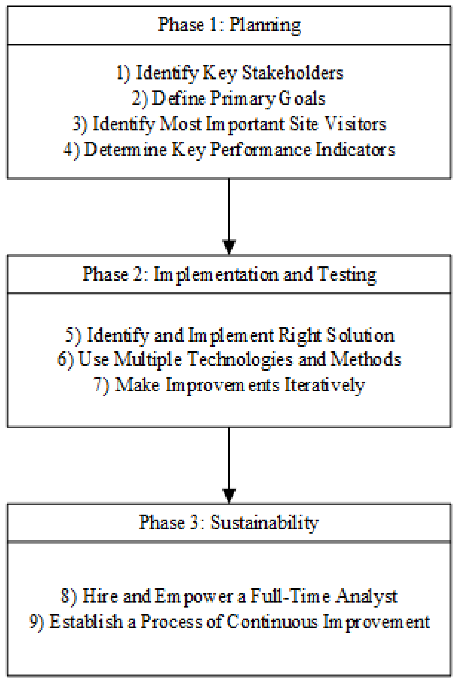

Figure 2. A strategic web analytics framework modified from the Web Analytics Process Guide introduced by Jansen (2009).

This case study focuses on the homepage, databases page, research guides page, and interlibrary loan (ILL) page of a major academic library. These four pages have significant monetary and content value to the libraries and are chosen for three reasons: (1) they are gateway pages that serve as entry points to other library resources, (2) the pages receive a lot of traffic relative to the other pages on the site, and (3) the pages collectively allow for the identification of organizational goals and KPIs (Phase 1 of the analytic model we are following) that are generalizable to nearly all academic libraries. Upon identification of key stakeholders as students, faculty, staff, and alumni and our subsequent selection of students as the most important site visitors for this research, we outline webpage goals, KPIs, and associated metrics, as shown in Table 2.

	Website Goals	KPIs	Associated Metrics
1	Improve Content Value	Engagement Patterns with Page	Pageviews Bounce Rate
2	Streamline Users to Key Tasks	Visits that Correlate with User Satisfaction	Entrances Returning Visitor Rate

Table 2. Identified goals, KPI's, and associated metrics for the selected pages.

4.2 Pixel Efficiency Analysis, Pixel Efficiency Value, and Conversion Efficiency Value

Again, we define pixel efficiency analysis as *the practice of relating screen real estate measured in pixels to the achievement of organizational goals and KPIs as indicated by quantifiable user behavioral interactions on a webpage*. Effectiveness is measured by achievement of library KPIs, where efficiency is based upon user behavioral patterns on the webpage. Efficiency is measured by usage of pixels. In this case study, user behavioral patterns are identified by a click analysis and heatmap analysis, which we translate into pixel real estate.

To perform PEA, we employ a heatmap on the pages previously identified as the focus of the case study. While various heatmapping tools are available, CrazyEgg is chosen for the study due to its wide set of features, ease of implementation, and relative low cost. Heatmaps allow for quick qualitative analysis of user behavior and adds substance to the quantitative metrics. Brighter colors indicate heavier usage, whereas darker colors indicate lesser usage. Based upon the color distribution, we identify three levels of usage: heavy (red and brighter), moderate (green), and little (blue and darker). Gray represents no usage. Where conflicts exist (e.g. a link receives heavy usage in the center, but little to no usage on the edges), the highest level of usage is awarded.

Pixel quantification is achieved by using the Google Chrome extension Page Ruler, which allows for the measurement of the webpage, webpage components, and behavioral usage (based upon the identified levels) in terms of square pixels. It should be noted that all pixel measurements are taken on a 1366 x 768 screen resolution, as this is by far the most common screen resolution of users for the library site in question, based on analysis of logs data.

Leveraging PEA into a measurement, at this moment, is not meant to be one exact measurement. Refining a specific measurement to be built upon is the subject of future research. For our purposes, we introduce two measurements. Our intention, at this stage of the research, is to utilize these measurements to advance and spur research utilizing pixels as an analytic tool.

The first measurement we introduce is achieved by summing the areas of moderate usage with areas of heavy usage, and dividing this by total real estate of the page. We refer to this measure as pixel efficiency value (PEV), capable of being employed at the page or component level (i.e., a subsection of a page, widget on a page), as shown in equation 1.

The test Conversion Efficiency Value (CEV) is shown in equation 3 and is achieved by taking the number of conversions within the respective area of measurement and adding this to the difference between the test variation conversion rate with the control conversion rate[1], then multiplying by

[1] The control CEV, shown in equation 2, is simply the ratio of conversions per square pixel multiplied by 100.

100. The difference in conversions between the test and baseline is taken into account to create a reward if conversion rate increases or a penalty if the conversion rate decreases. This is to be compared to the control CEV found in equation 2.

1. $$\text{PEV}_{\{\text{Page, Component}\}\ \text{Level}} = \frac{\text{Moderate Usage} + \text{Heavy Usage}}{\text{Total Real Estate of}\{\text{Page,Component}\}}$$

2. $$\text{CEV}_{\text{Control}} = \frac{\text{Control Conversions}}{\text{Total Real Estate of Component}} * 100$$

3. $$\text{CEV}_{\text{Test}} = \frac{\text{Test Conversions} + (\text{Test Conversions} - \text{Control Conversions})}{\text{Total Real Estate of Component}} * 100$$

PEV is a measure of the efficiency of webpage design that seeks to maximize the proportion of moderate-heavy usage within a respective area. CEV is weighted in a manner that takes conversions within a given area of real estate, while limiting the amount of pixel space allocated to achieve that goal. Both measures are ultimately geared towards similar concepts, exhibiting that pixels can be leveraged in a variety of manners to elicit the added value to KPI measurement.

The utilization of pixels provides libraries a numeric value for statistical testing of webpage design, similar to what monetary value provides for ecommerce sites, and aligns website changes directly to achievement of KPIs. Explicitly, improving PEV and CEV equates to a higher density of valuable content, thereby aligning with the organizational objectives and KPIs to increase overall content value and streamline users to key tasks. Even though the highest possible PEV or CEV is the end goal, keep in mind that not every portion of a page or component is clickable, nor is every portion of a webpage equal value (i.e. eyes get drawn to certain areas). For example, the upper left corner of a webpage is typically considered the most valuable. With this taken into account, we seek to steer academic librarians towards limiting the amount of content offered on a given page so that consumers do not become overwhelmed and are provided with the brief quality results that today's information era has embedded within users.

4.3 Data Collection and Analysis

4.3.1 High Level Analysis of the Academic Library Website User

To reiterate, our first analysis (RQ1) seeks to help us better understand the library website through the eyes of the user. Subsequently, we first collect measures of webpage goals and KPIs identified in Table 2. Measures are drawn from Google Analytics and are representative of academic year 2014/15 (AY 14/15). Next, heatmaps are employed through CrazyEgg and are set to collect data for 10,000 pageviews for the homepage and 5,000 pageviews for the remaining pages.

We then also distribute a survey to our users, which seeks to get a general perspective from the user's, while also specifically exploring our intuition that academic library consumers feel overburdened with the amount of information offered, which is an underlying argument for the employment of PEA. The survey collects basic demographic information, asks users to rate responses regarding the website on a Likert scale, and also asks for open-ended qualitative feedback. The survey is distributed through word-of-mouth, email, and is also placed on the homepage of the library in question.

Our first analysis ends by collectively analyzing the measures, heatmaps, and survey responses to identify webpage ineffectiveness and inefficiencies, within the webpages identified for analysis. Our aim here is not necessarily precision, but rather to enable our ability to tell a story about the user experience of academic library website users.

4.3.2 Testing of Inefficiencies Identified in High Level Analysis

To conduct our second analysis (RQ2), which focuses on behavioral patterns as a result of pixel inefficiency identification and subsequent recommendations, we employ the Optimizely A/B testing platform. Optimizely provides the capability to make changes to a webpage and then test those changes real-time against the control page by disseminating half of the traffic to the control page and half of the traffic to the variation page. This provides a low-risk method to evaluate attempted improvements of identified pixel inefficiencies. If improvement, based upon KPI achievement, is achieved with a more efficient usage of real estate, then we have clearly achieved a better variant. Even if KPI performance remains the same with more efficient usage of real estate, then we have still identified a better variant. Simply put, we suggest the main reason a more efficient usage of page real estate not be adopted is if KPI identified performance decreases. Following this model conceptually leads to utmost efficient balance that can be found between content and white space on a given page.

4.3.3 Creation of Pixel-based Values for Purposes of Reporting

Within the field of web analytics, common reporting practices utilize metrics (e.g. pageviews, bounce rate, exit rate, returning visitor rate, etc.). While investigating RQ2, we note differences in screen real estate allocation in pixels; however, a change in pixel allocation alone is not enough to report a difference. Rather, what value (or lack thereof) results from the change in pixel allocation? The third analysis (RQ3) seeks to take the results from RQ2 and turn them into the tactical measures of PEV and CEV.

CHAPTER 5: RESULTS

5.1 High Level User Experience Analysis

Table 3 represents an overview of the identified measures associated with goals and KPIs, noted in Table 2 of the current library website. Generally, the measures indicate mixed performance of KPIs, which indicates room for improvement in terms of effectiveness of the webpages. In addition, if we consider Figures 3-6, we can note large areas of real estate that are unused, which indicates, possibly, page real estate inefficiencies. Specifics of the inefficiencies can be seen in Table 4. Based upon the metrics and heatmaps, it appears the purpose of PEA, which is to increase effectiveness and efficiency of webpages with special emphasis on pixel sizing of components, is well suited for these pages. Yet, we are still left with a lack of understanding from users and are left wondering what the users are feeling, thinking, and experiencing. Overcoming this gap is achieved by employing a survey.

Page	Pageviews	Bounce Rate	Entrances	Returning Visitor Rate	Avg. Time on Page
Whole Site	7,581,761	50.86%	3,575,399	48.70%	2:49
Homepage	1,765,478	31.85%	1,124,262	71.49%	4:45
Databases	374,723	47.06%	107,106	76.00%	5:16
Research Guides	64,943	12.93%	3,176	74.67%	0:32
ILL	37,025	30.55%	8,628	83.56%	4:42

Table 3. Measures associated with identified goals and KPIs of the selected pages.

Figure 3. A heatmap image of the homepage produced by CrazyEgg.

Figure 4. A heatmap image of the databases page produced by CrazyEgg.

Figure 5. A heatmap image of the research guides page produced by CrazyEgg.

Figure 6. A heatmap image of the ILL page produced by CrazyEgg.

Page	Page Goals	Pixel Inefficiencies
Homepage	(1) Increase Content Value (2) Streamline Users to Tasks	Search Area Much Larger in Size Than Usage
Databases	(1) Navigate Users to Desired Database as Quickly as Possible	Large Areas of White Space in Navigational Area
Research Guides	(1) Streamline Users to Desired Guide (2) Simplify Content Offerings	Poor Page Balance with Large Portions of White Space
ILL	(1) Have Users Login to Their ILL Accounts	Users Attracted to Small Login Link Over Large Logo

Table 4. Overview of potential pixel inefficiencies based upon the CrazyEgg heatmaps shown in Figures 3-6.

Turning to the survey results displayed in Table 5, which reflect the views of 100 respondents that answered Likert-scale questions, several significant correlations exist. Notably, among the significant correlations, seven obtain $r > .55$ (i.e. seven correlations are moderate-high correlations). (3) Believing the site in question is easy to use is significantly and positively correlated with: (5) believing the site has easy navigation, $r = .656$, $p < .01$; (7) believing the organization of the site streamlines the information seeking process, $r = .720$, $p < .01$; and (8) believing the homepage is easy to navigate, $r = .694$, $p < .01$. (5) Believing that the site has easy navigation is significantly and positively correlated with: (7) believing the organization of the site streamlines the information seeking process, $r = .681$, $p < .01$; and (8) believing the homepage is easy to navigate, $r = .568$, $p < .01$. (6) Believing much of the content on the website is distracting is significantly and positively correlated with (9) believing that terminology on the homepage is confusing, $r = .615$, $p < .01$. Lastly, among the significant and moderate-high correlations, (7) believing that the organization of the site streamlines the information seeking process is significantly and positively correlated with (8) believing that the homepage is easy to navigate, $r = .783$, $p < .01$.

	1	2	3	4	5	6	7	8	9
1. Too Much Info = Confusion		-.158	**-.308****	.163	**-.267****	**.246***	**-.247***	**-.200***	**.289****
2. Aware of Resources	-.158		**.370****	.003	**.439****	-.130	**.272****	**.294****	-.009
3. Ease-of-Use	**-.308****	.370		.008	**.656****	**-.301****	**.720****	**.694****	-.132
4. Photographs	.163	.003	.008		-.060	.106	.074	.090	.037
5. Easy Navigation	**-.267****	**.439****	**.656****	-.060		-.151	**.681****	**.568****	-.140
6. Distracting Content	**.246***	-.130	**-.301****	.106	-.151		-.194	-.150	**.615****
7. Efficient Organization	**-.247***	**.272****	**.720****	.074	**.681****	-.194		**.783****	-.149
8. Homepage Easy Navigation	**-.200***	**.294****	**.694****	.090	**.568****	-.150	**.783****		-.138
9. Homepage Terminology Confusing	**.289****	-.009	-.132	.037	-.140	**.615****	-.149	-.138	

*$p < .05$. **$p < .01$ for $n = 100$ (two-tailed).

Table 5. A Pearson (two-tailed) correlations matrix representing the variables within the survey.

We will highlight two correlations that are particularly important for the foundational assumptions of PEA. First, among our findings is the notion that believing the site is easy to use is significantly and positively correlated with believing that the organization of the site streamlines the information seeking process (3 & 7, $r = .720$, $p < .01$). Rather, if users believe the site is not organized in a manner that allows them to quickly find what they are looking for, then it is statistically probable that the users will also believe the site has low usability. Second, noteworthy of reiterating, is that believing much of the content on the website is distracting is significantly and positively correlated with believing that terminology on the homepage is confusing (6 &9, $r = .615$, $p < .01$). Hence, if we assume that academic library website users are more likely to find homepage terminology confusing than on a traditional website (at least at first), then we can also determine that users will likely find much of the content on the website distracting. Confused and distracted users are unlikely to believe they are engaging in a streamlined information seeking process, and subsequently, will likely believe low usability exists. So if we rely upon PEA to help effectively and efficiently utilize page real estate, then we will lower the amount of distracting and confusing content, while also streamlining the information seeking process and increasing usability of the site.

The utilization of existent web analytic methodologies to explore issues within the academic library website in question (RQ1) has allowed us to find ineffectiveness and inefficiencies. Moving forward we could simply utilize A/B testing and measure engagement levels, yet we seek to take that a step further by incorporating pixel measurements in order to aid in webpage redesign. Doing so provides a way to rank and value components on a webpage, which has the potential to eliminate an overburden of information that has been identified in the survey as an issue within this academic library website.

5.2 A/B Testing of Webpage Inefficiencies

Given the identified pixel inefficiencies from the high level analysis, we move to RQ2 and seek to analyze what impact attempts to improve the pixel inefficiencies have on user behavior. For this research, we identified one area or component of weakness for each selected page.

- For the **homepage**, we observe that the search box tends to be relatively large in comparison to the amount of real estate that obtains clicks.
- On the **databases page**, the navigational area appears to have a low density of clicks that are widely dispersed.
- The **research guides page** contains a large amount of white space and offers no promotion of content other than a link list.
- Lastly, the **ILL page** seems to detract users from its main goal, which is to get users to login to their ILL account.

In essence, we strive to show that each of the four mentioned issues can condense information and utilize less page real estate, while also maintaining or increasing KPI effectiveness.

Based upon these identified weaknesses, Table 6 presents the A/B testing results on the components of interest. Notably, in 9 out of 10 analyses, we are able to increase the efficiency at which pixel space is utilized and also increase engagement within the targeted componential area. Following, we will briefly overview each test that was done and interpret the results.

Page	Control Visitors	Variation Visitors	Component/Area	Control Real Estate	Variation Real Estate	Control Engagement	Variation Engagement
Homepage	5,481	5,459	Search Box	123,200	106,232	108.48%	**111.63%**
Homepage*	*0*	*0*	*Search Box*	*123,200*	*49,006*	*108.48%*	***111.63%***
Databases	3,478	3,491	Navigational Area	302,670	224,475	**28.32%**	26.30%
Databases	3,582	3,490	Navigational Area	302,670	206,025	29.59%	**31.78%**
Databases	2,831	2,856	Navigational Area	302,670	206,025	31.61%	**32.04%**
Research Guides	776	779	Featured Guides	826,937	826,937	11.86%	**18.23%**
Research Guides	741	756	Featured Guides	826,937	826,937	1.35%	**4.63%**
Research Guides	904	880	Featured Guides	826,937	826,937	6.19%	**7.61%**
Interlibrary Loan	314	327	Login Area	32,040	27,813	54.78%	**65.44%**
Interlibrary Loan	339	290	Login Area	32,040	24,455	65.19%	**75.52%**

**Was not an A/B test. Users were asked to specify what search box they preferred in a survey, as the change was to risky for real-time implementation. Consequently, we carried over our findings from the real-time A/B test given that 61/93 users indicated that they preferred a search box of 49,006 square pixels over the orginal.*

Table 6. An overview of A/B testing results.

5.2.1 Homepage Analysis

An examination of current library homepages exhibits a tendency for focused emphasis on the search box, that permits access to the libraries extensive collection of online academic databases via a web search engine like search widget. This trend is intuitive, seeing as it is often times the quickest and easiest way to access an online library resource. However, we maintain that increased importance does not necessarily mean that more page real estate should be allocated to a respective component, particularly in the case of a search box. Market trends of major websites (e.g. Google, Amazon, etc.) dictate streamlined search areas that, in comparison, do not take up much page real estate. Hence, we explore what impact streamlining a library search box has on user behavior.

The A/B test we employ on the search area (Figure 7) removes the black border that exists around the current search area, which is shown in Figure 8. This decreases the size of the search area from 123,200 square pixels to 106,232 square pixels. Doing so results in an engagement level of 108.48% across 5,481 users for the control and a variation level engagement of 111.63% across 5,459 users, where engagement is defined as a click within the identified componential area divided by the number of visitors. Given the significant importance of the search area to the actual user base, we were limited in how much live testing we could do and how radically we could change the search box. Nonetheless, we are able to decrease the amount of real estate taken up by the search area, while even slightly increasing engagement levels.

Figure 7. Search area variation of the library in question.

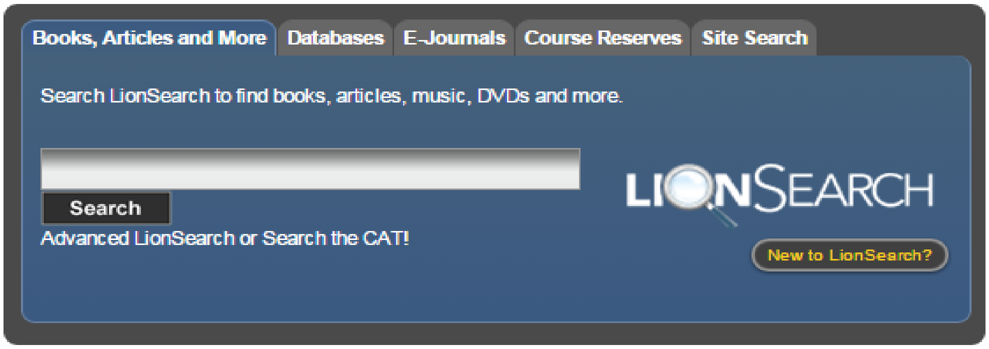

Figure 8. Original search area of the library in question.

Yet, we still believe that an even smaller search area would perform better. To pursue this notion, we offer two search areas, Figures 8 and 9, in the survey and ask users to indicate their preference. Of the responses, two-thirds of respondents prefer Figure 9, which is 49,006 square pixels in size. That difference in size represents a 60.22% decrease in page real estate, which would free up valuable real estate above the fold on the homepage, while apparently also offering better performance and better alignment with user preference.

Figure 9. A variation of the search box preferred by 61/93 users surveyed.

5.2.2 Database Page Analysis

Next, we turn to the databases page of the library's site. Like the homepage, the databases page is a "gateway" page, in that it serves as a portal to browse various other resources. Specifically, databases within the academic library world refer to bundles of electronic resources. The library in question spends millions of dollars per year on these databases (Coughlin & Jansen, 2015), so

Figure 10. The navigational area of the databases page.

conversions (i.e. clicking on a database) and having visitors return are KPIs for the page. The databases page consists of a navigational area, then a long link list. Figure 10 depicts the navigational area. We focus on the navigational area because, as noted earlier, it is an area of weakness upon analysis of the heatmap from our data collection analysis. In particular, we notice that the clicks within the navigational area tend to have relatively low density and are widely dispersed, which is not what we would expect from an efficient and effective navigational area. Consequently, we seek to streamline the navigational area that would allow for more efficient and effective utilization of page real estate.

Figure 11. A screenshot of the first A/B test for the databases page.

We conduct three A/B tests on the databases page. The first A/B test, as shown in Figure 11, is the only test of all of our A/B tests that does not achieve better engagement levels. The decrease in engagement, which comes after 3,478 control users and 3,491 variation users, should be considered with the notion that only an iterative change (a vital step within the methodology of A/B testing) is first made to examine user behavior. The decrease in engagement that we see is slight and therefore continue on with A/B tests 2 and 3 in order to measure the impact on user behavior.

Within A/B tests 2 and 3 (e.g. Figure 12) we aim for better utilization of the pixel space within the navigational area. In particular, we take the "Try These First" link and offer links under it within previously unused space, while also rearranging some elements in order to condense and simplify the information offered within the navigational area. The result is a reduction in pixel space from 302,670 square pixels to 224,475 square pixels. We also note slight increases in engagement levels. For test 2, with 3,582 control users and 3,490 variation users, engagement increases from 29.59% to 31.78%. For test 3, with 2,831 control users and 2,856 variation users, the engagement levels increase from 31.61% to 32.04%. Similarly, to the homepage, we are able to increase pixel efficiency, as well as increase effectiveness relative to KPIs.

Figure 12. A screenshot of the second A/B test for the databases page.

5.2.3 Research Guide Page Analysis

Next, under consideration of RQ2, we focus on the research guides page. The research guides page (Figure 13) is also a gateway page. The content produced within the research guides is intended to help the users of the website conduct research and is maintained by staff members of the library. While the databases page contains some sort of navigational area to help assist users in locating the appropriate database, this is not the case with the research guides page; outside of a search box, the research guides page offers no navigational area. In taking an empathetic stance with the users, if you are seeking a guide for help to do research, coming to a page full of links is likely not preferable as an overabundance of information can be overwhelming to begin with, let alone when one is confused and looking for help. We subsequently seek to experiment with the idea of "featured guides." The featured guides are displayed above the fold in what was previously white space. Three different iterations of five featured guides are tested. Each test increases the collective engagement levels of the featured guides. The approach to improving pixel efficiency for the research guides page is different than the approach taken for the homepage and databases page. Rather, instead of condensing componential areas, the focus here is to utilize white space to feature content in attempt to enhance the ability of users to quickly navigate to the appropriate resource, a KPI of the research guides page. Again, collective engagement levels of the featured guides increase in all three A/B tests.

Similarly to the search area, we offer a more drastic variation (Figure 14) within the survey. Within this variation, all area "above the fold" on the research guides page is dedicated to assisting users in finding a research guide. Notably 86 of 93 (92%) respondents prefer the variation over the original variation. Hence, a trend seems to be occurring that within this academic library website, users tend to prefer tools that help find resources quickly rather than offering the resources outright. A possible explanation for this observation is that libraries use terminology that may often seem confusing to users, although this needs research to confirm. Hence, searching for a resource using personal language over librarian language increases usability.

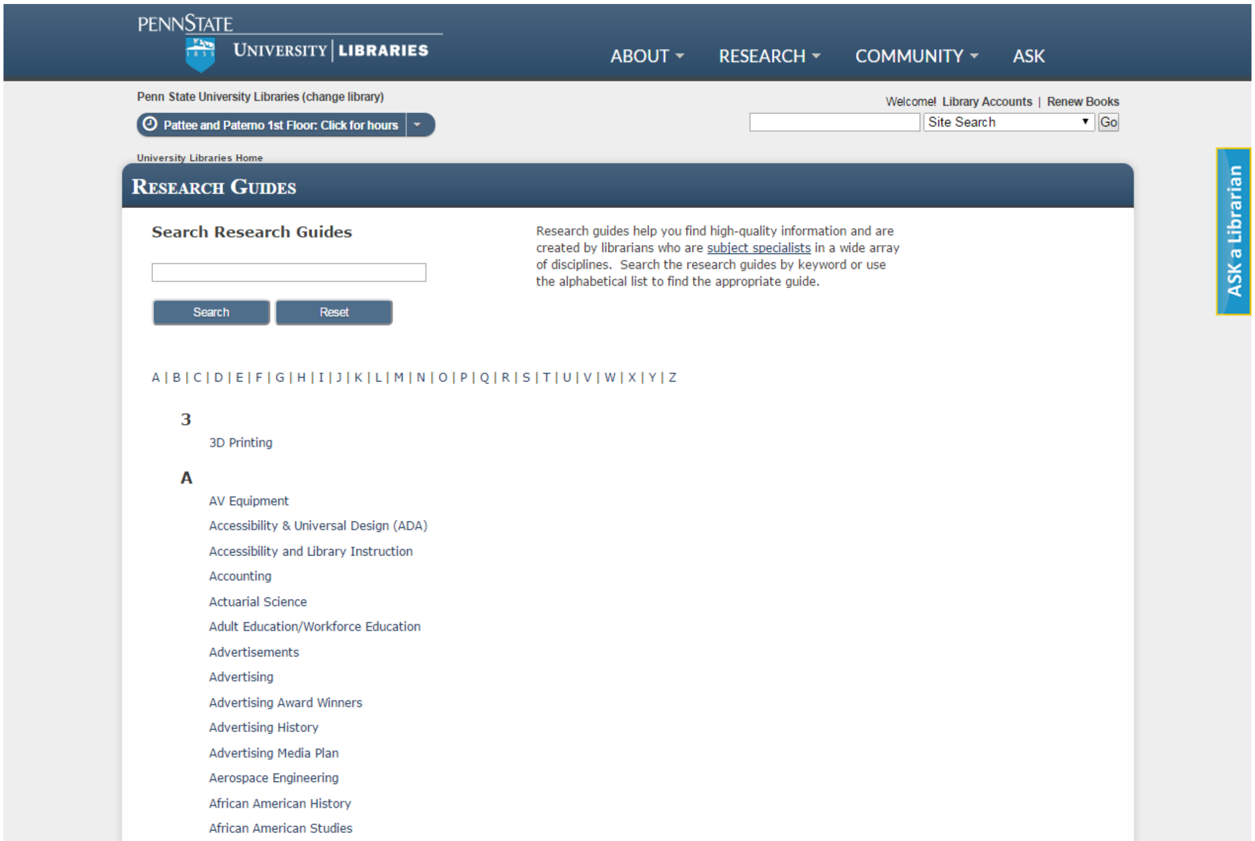

Figure 13. The research guides page of the library in question.

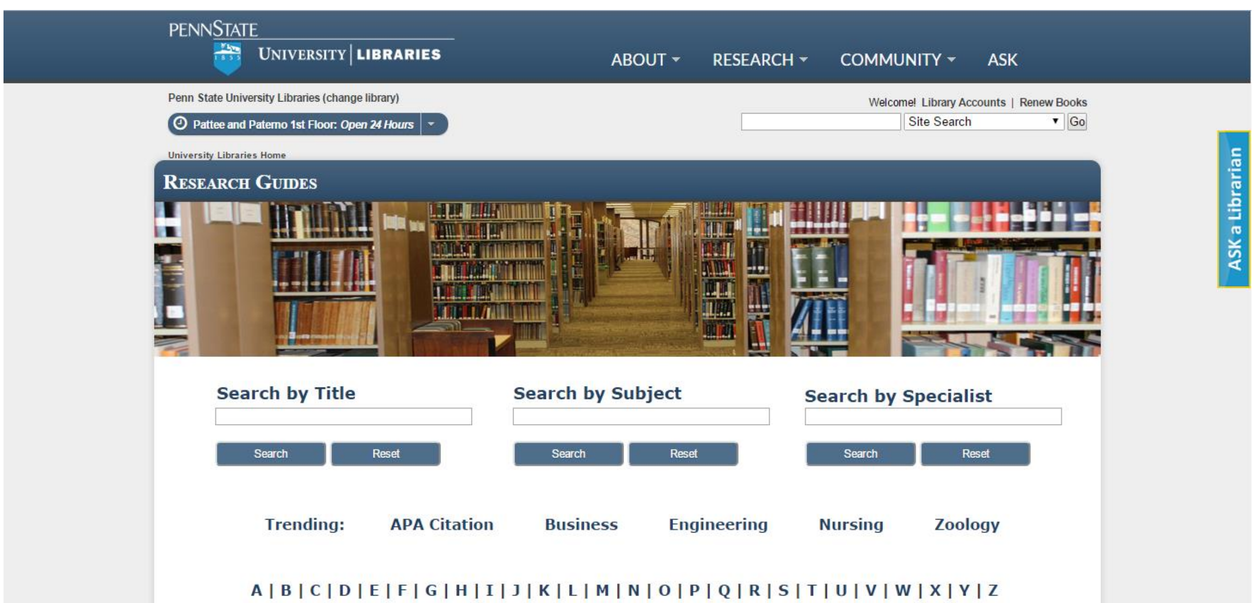

Figure 14. A variation of the research guides page offered to users in the survey. 86/93 participants selected this option over Figure 13.

5.2.4 ILL Page Analysis

The last page examined with A/B testing is the ILL page (Figure 15). The ILL page serves as a way for users to request books, articles, and other materials that the library in question may not have access to, but that another library does. This results in users saving money on subscriptions and is arguably an underutilized resource within the libraries. Yet as noted earlier, we have concern that the existing page is not getting users to login to their ILL accounts in an efficient and effective manner. Namely, this is because users are utilizing a small login link just as frequently as the central login feature. While this may not be a major problem, we still find it surprising that users are drawn to a small login link over instead of the login logo that is the central focus of the page. We therefore make two iterations of the central login feature to see if we can increase the number of logins (i.e. effectiveness), while also increasing pixel efficiency. In both tests, we are able to increase effectiveness by getting more users to login to their ILL accounts, while also making more efficient usage of webpage real estate by using less webpage real estate with the central login feature.

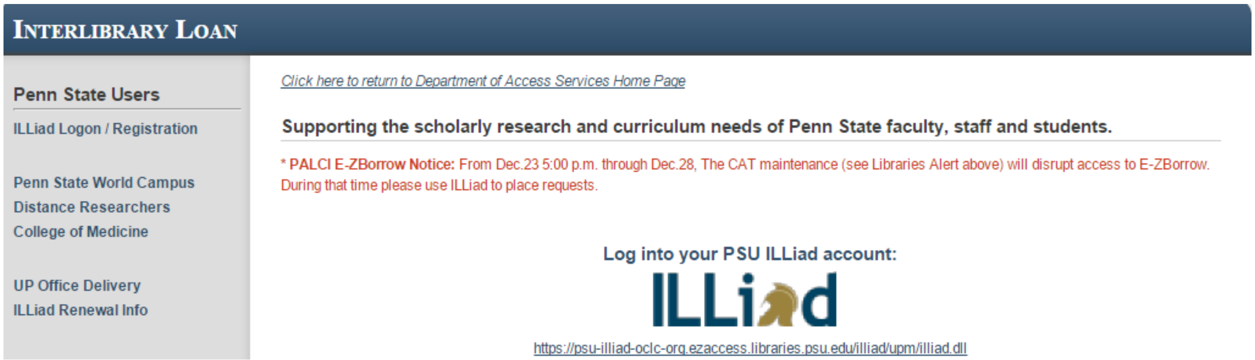

Figure 15. The control version of the ILL page.

Specifically, ILL test 1 (Figure 16) achieves an increase in logins of 54.78% to 65.44%, while also seeing a 13.19% reduction in pixel space allocation from 32,040 square pixels to 27,813 square pixels. In ILL test 2 (Figure 17), we note logins increase from 65.19% to 75.52% with pixel space allocation decreasing 23.67% from 32,040 square pixels to 24,455 square pixels.

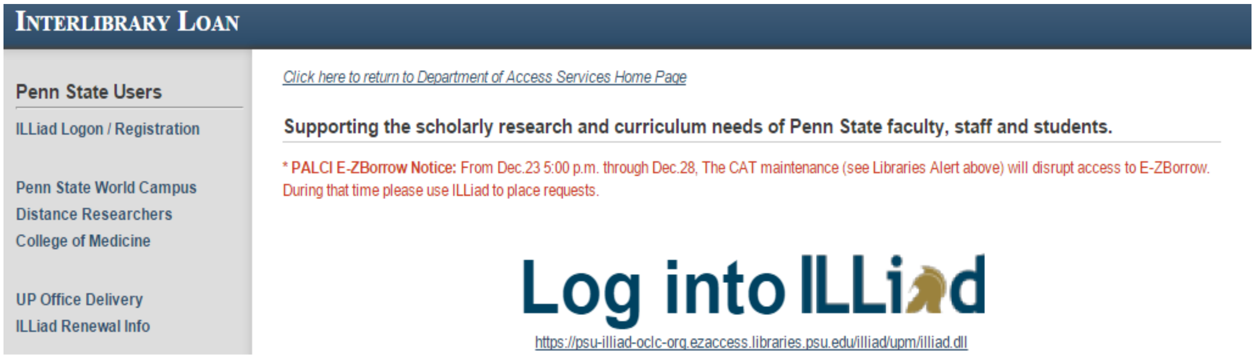

Figure 16. A/B test 1 of the ILL page.

Figure 17. A/B test 2 of the ILL page.

5.3 Pixel Efficiency Measurements

We now turn to RQ3 and attempt to turn our findings thus far into the tactical web analytic measures of PEV and CEV. Through PEA, we find that the library homepage contains 1,245,127 square pixels (1,349 x 923). Of the roughly 1.25 million square pixels, 0.87% (i.e. 10,876) of square pixels obtain moderate usage and 2.63% (i.e. 32,788) of square pixels obtain heavy usage based on the measurements taken, where green usage is moderate and red and brighter is heavy usage.

$$3.\ \text{PEV Homepage} = (10{,}876+32{,}788)/1{,}245{,}127 = 3.51\%$$

We cannot for certain say where a PEV of 3.51% lies on the efficiency spectrum of academic library homepages, given the lack of comparative data among research libraries. While we do not suggest that every square pixel of a webpage should be clickable[2], we are compelled to believe if 96.5% of an academic library homepage is achieving little to no usage, even when there are clickable elements, then the real estate can be better utilized. It is noteworthy to mention that certain homepages (e.g. Google) will undoubtedly have low PEV levels of the overall page. However, academic libraries differ given the number of services offered; hence it is seemingly important to maximize screen real estate efficiency on the homepage. However, even Google, which occasionally introduces additional widgets on its homepage, could use PEA for evaluation purposes.

We find that 6,368 square pixels, or 5.17%, of the search area obtain moderate usage and 16,623, or 13.49%, square pixels obtain heavy usage, combining for 22,991 square pixels of the search area obtaining moderate to heavy usage. Given that 43,664 square pixels obtain moderate to heavy usage in total on the homepage, we can note that the search area accounts for 52.65% of moderate to heavy usage on the homepage. While the proportion of moderate to heavy usage achieved by the search area (in comparison to the rest of the page) bodes well for ROI, we are still left with low efficiency. PEA reveals that the search area on the homepage takes up a total of 123,200 square pixels, yet is that size necessary to maintain effectiveness? Rather, we maintain that there is potential to improve efficiency of the search area, while maintaining and perhaps even increasing effectiveness. Equation 4 shows a search area PEV of 18.66%.

[2] *The Million Dollar Homepage* is an interesting example of pixel real estate usage where every square pixel is clickable: https://en.wikipedia.org/wiki/The_Million_Dollar_Homepage.

4. PEV Search Area = (6,368+16,623)/123,200 = 18.66%

5. PEV A/B Search Area = (6,368+16,623)/106,232 = 21.64%

A potential solution to the lacking efficiency is to offer a streamlined search area, as tested earlier. We can see the results in equation 5 and note better performance than equation 4. Notably, we argue that we could streamline the search area even further. For example, Amazon's search area (Figure 18) is 29,341 square pixels, which is a 76.18% decrease from this library's search area. This also ties back to Fitts' Law, which tells us that the increased size does not necessarily bring about increased usability (Karafillis, 2012).

Figure 18. A screenshot of Amazon's search box at the time of this study.

Though, the original A/B test achieves only slight improvement (18.66% vs. 21.64%), this must be taken into account with the notion that this represents a real-time A/B test, which hinges on slight variations to existing and operational system interfaces. The point that the area maintained relatively heavy usage even when manipulating the search area in real time indicates that the same or better organizational KPI achievement is attainable via a more efficient utilization of the page real estate. While the A/B test presented does not necessarily represent the envisioned streamlined search area (i.e. Google, Amazon, etc.), the results show the viability of the pixel efficiency method.

We were able to employ a real-time manipulation to the search area of the homepage that permitted a more efficient usage of page real estate, with no loss of KPI effectiveness. In particular, when utilized in combination with a survey, we exhibit how you can drastically improve the efficiency of real estate usage and that bigger is not always better. The modified search area within the survey was much smaller (i.e. 49,006 square pixels) than both the original search area and the first A/B test search area. Despite being tested through a survey instead of an A/B test, which we determined to be too big of a risk, we report that 2/3 of users prefer the smaller search area in the survey. The streamlined search area shown in the survey, equated to PEV, is shown in Equation 6.

6. PEV Survey Search Area = (6,368+16,623)/49,006 = 46.91%

For the purposes of converting our other results into measures, we utilize CEV as the other pages we are primarily concerned with increasing conversions to the resources on the page rather than increasing the pixel efficiency of the page as a whole. Rather, with these pages we are more concerned with page real estate "above the fold." As the nature of these pages requisites that many resources be offered. If streamlining users to resources is a KPI, then effectively and efficiently getting users to engage in a conversion above the fold is important. One database test and one ILL test are selected for the purposes of prototyping. We elect not to provide a quantitative measure for the research guides testing, seeing as we only utilized previously unused white space. Results can be seen in Table 7.

Component	Conversions	Component Real Estate	CEV Equation	CEV
Database Baseline	1,060	302,670	$\lvert (0 + (0 – 1,060) \rvert / 302,670) * 100$	35.02%
Database A/B 2	1,109	205,200	$\lvert (1,109 + (1,109 – 1,060) \rvert / 205,200) * 100$	**56.43%**
ILL Baseline	123	32,040	$\lvert (0 + (0 – 123) \rvert / 32,040) * 100$	38.39%
ILL A/B 1	160	27,813	$\lvert (160 + (160 – 123) \rvert / 27,813) * 100$	**70.83%**

Table 7. An overview of CEV results regarding the Databases and ILL pages.

In both cases, we find that the A/B variation achieves better performance. What is particularly noteworthy about these results is that comparatively these values are able to elicit the value of the more efficient usage of webpage real estate, while also achieving better effectiveness in support of KPIs.

Without CEV, the databases A/B test would be presented as: a reduction of 96,645 square pixels and an increase in engagement levels from 31.61% to 32.04%. With CEV, we can say that we relatively high increase in performance from 35.02% to 56.43%, which represents an increase in efficiency and effectiveness. The value in essence rewards more efficient usage of pixel space, if engagement levels are also increased. In contrast, if engagement levels decrease with decreased pixel allocation, then a penalty is given to the value indicating that more pixel space should be allocated to the component.

We have established PEV and CEV as additional measures that can be used to work towards optimizing the goals of providing high value content and streamlining user behaviors. In fact, PEV and CEV provide quantitative measures no other current measure is capable of producing. Namely, a measure that links the aesthetic and organizational presence of a webpage to measurable outcomes of user behavior.

CHAPTER 6: DISCUSSION

The purpose of this research is to present a web analytics approach for academic libraries and a novel approach in which PEA is used for evaluating webpage changes. This approach offers the ability to show areas for more efficient and effective utilization of a webpage that cannot be identified and evaluated through traditional web analytic measures. PEA allows for quantification of pixel efficiency at the page and component levels.

From a practical standpoint, PEA is not intended to replace traditional web analytics methodologies, such as the use of measures common in packages such as Google Analytics. Instead, PEA can be used in combination with other technologies and methods, as we did in this case study by using Google Analytics, CrazyEgg, Optimizely, PEA, and market analysis via surveys in order to evaluate possible courses of action identified through these techniques. Via the use of multiple technologies and methods, we can more effectively understand user behavior and then meet their needs on a library website via effective and efficient webpage real estate usage.

Another key takeaway worth noting is the story that can be told from this case study. Telling a story with the data should be a goal of any web analyst (Kaushik, 2007c). Here, we start with an assumption that users of an academic library website may feel burdened with too many options of resources to choose from. Our survey results provide support for that assumption. If we further examine our testing results, we can find further support. Take the search area for instance. It appears users prefer the streamlined version.

In essence, it appears that within the academic library context users feel easily confused and overwhelmed when provided with too much information options on a library webpage. There are likely many reasons this trend seems to be exacerbated within an academic library, such as terminology that differs from typical everyday usage of library patrons.

As such, our recommendation is that academic libraries simplify the information seeking process as much as possible, despite the richness of research that can result from complex information seeking tasks within an academic library. PEA can support such simplification by employing PEV and CEV, or other similar measures that come as a result of research that extend PEA. Namely, streamlining the information seeking process with particular emphasis on high value content, which is the main emphasis of PEA, stands to enhance usability and competiveness of academic library websites.

CHAPTER 7: CONCLUSION

Given the vast amount of electronic resources libraries contain and have access to, webpage real estate stands as the equivalent of monetary value in an academic library. The analysis conducted in this research is based upon that concept. To look further into this concept, we took a three-pronged approach. First, we seek to understand the website from the eyes of the user. In turn, we distribute a survey, analyze heatmaps, and identify weaknesses of the website. The second research question seeks to increase pixel efficiency based upon identified weaknesses. We conduct ten A/B tests across four major library webpages. In 9 of 10 tests, we increase the pixel efficiency of the components, while also increasing effectiveness. The third research question pursues the capability of PEA to provide actionable intelligence. Here, we elicit how the creation of the web analytic measures of PEV and CEV can be used to enhance interpretation of PEA results.

As with any research, there are limitations with the current use of PEA. For instance, some pages that are intended solely to provide information are not likely to have high engagement levels. Employing PEA that rely on clicks on such a page is likely to indicate that the page serves of little value, when in reality, users may find the content on the page very useful. However, this should be taken in lieu of the notion that different pages serve different objectives and therefore will require different levels of measurement and is no different than selecting a metric from Google Analytics. An example of this can be seen from the APA in-text citation page within this library's website. It has a bounce rate close to 70%, which typically is a bad thing. However, this page also achieves the second most entrances (next to the homepage) and most users access this page via a search engine. Hence, the page is of value and, in this case, bounce rate is not a useful measure in support of KPIs.

Worth mentioning as well, is that we find the use of a heatmap tool to be somewhat limiting in that users look at more content than they click on. With that said, eye tracking, in combination with a heatmap, will likely allow for even greater analysis of the efficiency and effectiveness of webpage content for the purposes of PEA, which also addresses the issue regarding the evaluation of information pages. For example, the upper left corner of a webpage is typically considered the most valuable. Hence, what impact does shifting different content to more salient webpage regions have on overarching goals and KPIs? Aside from utilizing more precise eye tracking methods, we also suggest that future research analyze different mediums of technology. For instance, focusing on PEA within the realm of mobile computing stands to be particularly powerful, given the limited screen real estate. Another area that we believe is worth pursuing with PEA is that of computational advertising. Future work may explore webpage changes based upon PEA recommendations from a lab-based setting or combined with more complex web analytics methodologies, such as exploring PEA in relation to various browsers and/or time of day (Zhang, Jansen, & Spink, 2009).

REFERENCES

Abramson, M., & Aha, D. W. (2013). *User authentication from web browsing behavior*. Paper presented at the The Twenty-Sixth International FLAIRS Conference.

Aharony, N. (2012). An analysis of American academic libraries' websites: 2000-2010. *The Electronic Library, 30*(6), 764-776.

Betty, P. (2009). Assessing homegrown library collections: Using Google Analytics to track use of screencasts and flash-based learning objects. *Journal of Electronic Resources Librarianship, 21*(1), 75-92.

Black, E. L. (2009). Web Analytics: A picture of the academic library web site user. *Journal of Web Librarianship, 3*(1), 3-14.

Brophy, J., & Bawden, D. (2005). Is Google enough? Comparison of an internet search engine with academic library resources. *Aslib Proceedings, 57*(6), 498-512.

Brown, A., & Abramson, M. (2015). *Twitter fingerprints as active authenticators.* Paper presented at the 2015 IEEE International Conference on Data Mining (ICDM).

Buscher, G., Cutrell, E., & Morris, M. R. (2009). *What do you see when you're surfing?: using eye tracking to predict salient regions of web pages.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Conyers, A., & Payne, P. (2011). Library performance measurement in the digital age. In P. Dale, J. Beard, & M. Holland (Eds.), *University Libraries and Digital Learning Environments* (pp. 201-214).

Coughlin, D. M., Campbell, M. C., & Jansen, B. J. (2013). *Measuring the value of library content collections.* Paper presented at the Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries, Montreal.

Coughlin, D. M., & Jansen, B. J. (2015). Modeling journal bibliometrics to predict downloads and inform purchase decisions at university research libraries. *Journal of the Association for Information Science and Technology*.

Deschenes, A. (2014). Improving the library homepage through user research - without a total redesign. *Weave: Journal of Library User Experience (e-journal), 1*(1).

Digital Analytics Association. (n.d.). About us. Retrieved April 8, 2015, from http://www.digitalanalyticsassociation.org/about

Fagan, J. C. (2014). The suitability of web analytics key performance indicators in the academic library environment. *The Journal of Academic Librarianship, 40*(1), 25-34.

Fang, W. (2007). Using Google Analytics for improving library website content and design: A case study. *Library Philosophy and Practice (e-journal), Paper 121*, 1-17.

Fang, W., & Crawford, M. E. (2008). Measuring law library catalog web site usability: A web analytic approach. *Journal of Web Librarianship, 2*(2-3), 287-306.

Ferrini, A., & Mohr, J. J. (2009). Uses, limitations, and trends in web analytics *Handbook of Research on Web Log Analysis* (pp. 122-140).

Ghaphery, J. (2005). Too quick? Log analysis of Quick Links from an academic library website. *OCLC Systems & Services: International digital library perspectives, 21*(3), 148-155.

Google. (2014). Our history in depth. Retrieved April 8, 2015, from http://www.google.com/about/company/history

Google. (n.d.). Digital analytics fundamentals: The importance of digital analytics. *Google Analytics Academcy*. Retrieved from: https://analyticsacademy.withgoogle.com/course01/assets/pdf/DigitalAnalyticsFundamentals-Lesson2.1TheimportanceofdigitalanalyticsText.pdf

Guidorizzi, R. P. (2013). Security: Active authentication. *IT Professional, 15*(4), 4-7.

Jansen, B. J. (2006). Search log analysis: What it is, what's been done, how to do it. *Library & Information Science Research*, *28*(3), 407-432.

Jansen, B. J. (2009). Understanding user-web interactions via web analytics. *Synthesis Lectures on Information Concepts, Retrieval, and Services, 1*(1), 1-102.

Karafillis, A. (2012). When you shouldn't use Fitts's Law to measure user experience. Retrieved July 24, 2015, from http://www.smashingmagazine.com/2012/12/fittss-law-and-user-experience/

Kaushik, A. (2007a). Rethink web analytics: Introducing web analytics 2.0. Retrieved from http://www.kaushik.net/avinash/rethink-web-analytics-introducing-web-analytics-20/?utm_source=analytics%20academy&utm_medium=text%20lesson&utm_campaign=lesson%202.1

Kaushik, A. (2007c). *Web analytics: An hour a day*. New York: Wiley Publishing.

Kesselman, M., & Watstein, S. B. (2005). Google Scholar™ and libraries: point/counterpoint. *Reference Services Review, 33*(4), 380-387.

Kumar, L., Singh, H., & Kaur, R. (2012). *Web analytics and metrics: A survey*. Paper presented at the ICACCI, Chennai, India.

Loftus, W. (2012). Demonstrating success: web analytics and continuous improvement. *Journal of Web Librarianship, 6*(1), 45-55.

Manuel, S., Dearnley, J., & Walton, G. (2010). Continuous improvement methodology applied to united kingdom academic library websites via national survey results. *New Review of Information Networking, 15*(2), 55-80.

Memmott, S., & deVries, S. (2010). Tracking the elusive student: opportunities for connection and assessment. *Journal of Library Administration, 50*(7-8), 798-807.

Nicholson, S., Sierra, T., Eseryel, U. Y., Park, J.-H., Barkow, P., Pozo, E. J., & Ward, J. (2006). How much of it is real? Analysis of paid placement in Web search engine results. *Journal of the American Society for Information Science and Technology, 57*(4), 448-461.

Paul, A., & Erdelez, S. (2013). Implementation and use of web analytics for academic library websites. *World Digital Libraries, 6*(2), 115-132.

Plaza, B. (2009). Monitoring web traffic source effectiveness with Google Analytics. *Aslib Proceedings, 61*(5), 474-482.

Regazzi, J. (2012). Constrained? An Analysis of U.S. Academic Library Shifts in Spending, Staffing, and Utilization, 1998-2008. *College & Research Libraries, 73*(5), 449-468.

Srivastava, J., Cooley, R., Deshpande, M., & Tan, P.-N. (2000). Web usage mining: Discovery and applications of usage patterns from web data. *ACM SIGKDD Explorations Newsletter, 1*(2), 12-23.

Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research, 36*(2), 84-90.

The KPI Institute. (2016). Key Peformance Indicators Infographic (pp. 1).

Turner, S. J. (2010). Website statistics 2.0: Using Google Analytics to measure library website effectiveness. *Technical Services Quarterly, 27*(3), 261-278.

Usability First. (n.d.). Glossary. Retrieved July 23, 2015, from http://www.usabilityfirst.com/glossary/screen-real-estate/

Whang, M. (2007). Measuring the success of the academic library website using banner advertisements and web conversion rates. *Journal of Web Librarianship, 1*(1), 93-108.

Young, S. (2014). Improving library user experience with A/B testing: Principles and process. *Weave: Journal of Library User Experience (e-journal), 1*(1).

Zhang, Y., Jansen, B. J., & Spink, A. (2009). Time series analysis of a Web search engine transaction log. *Information Processing & Management, 45*(2), 230-245.

APPENDIX: EXAMPLES OF OPEN-ENDED SURVEY RESPONSES

Describe what you dislike about the Penn State Library's site navigation.

"Can be cumbersome to navigate where you want to go at times"

"complicated"

"I can never find ILL when I need to."

"I dislike that it is difficult to search for the different things needed. Sometimes I can't figure out how to loan a book from another campus or where the book is. Very very confusing!"

"if search fails, hard to find stuff"

"it can be difficult to find resources"

"It can be obtrusive and overwhelming"

"It is very easy to get confused while on the site. It can also be hard to find exactly what you need because there is so much information."

"The pages are rather busy"

"too much info and too many different ways to get to the same places"

What frustrations do you have about the current homepage of the Penn State Library site?

"A lot of information in one place."

"hard to find what I need. Have to filter all the content to get what I want."

"Hate the slideshow. I'm not sure it's WCAG accessibility compliant. It would be nice of some features like hours, contact & some key reference pages were above the fold instead of events."

"It is difficult to figure out where to find information."

"It's a little busy."

"It's sort of cluttered."

"Needs a little revamping, made simpler, makes it easier to figure out what you're looking for"

"Not enough visuals on the current homepage. It is too wordy!"