Research Note

Introducing Undergraduate Students to Writing and Algorithms: Understanding Analytics and Measuring Content on Personal Capstone Websites

Laura A. Palmer, *Kennesaw State University*

Structured Abstract

- **Aim:** Little is understood about how students learn to create content at the intersection of algorithmic functions, rhetorical constructs in digital spaces, and user engagement as measured through basic analytics on their personal capstone websites. Towards providing some insight in this area, this research note will discuss how students in a technical communication class on search engine optimization (SEO) and analytics began to re-situate their conceptions of what it means to be findable on the Internet. This note provides initial findings from the literature and presents several top-level outcomes from teaching TCOM 3245: SEO and Analytics within the Bachelor of Science in Technical Communication degree at Kennesaw State University. Specifically, these generalized findings may help instructors deepen their teaching expertise about content development for digital platforms such as blogs or websites; additionally, this research note engages others with ideas about pedagogy or more formal research in the areas of teaching SEO and/or website analytics.
- **Problem Formation:** If students are to work in digital spaces and places with the goal of publishing information across the Internet, it is important they learn how to drive traffic to a digital asset and measure what the traffic does

DOI: 10.37514/JWA-J.2019.3.1.10

with an eye towards strategic iteration and re-deployment. Many classes engage students with "digital writing"—developing written content for blogs or websites. However, the countless blogs and websites created each semester enter the black hole of the Internet because they lack basic findability via search engine optimization practices. The content of a blog or a website must first appeal to an algorithm, the proprietary machine crawling the Internet and indexing pages. To miss the algorithm is to never reach a human reader; thus, developing a properly leveraged site and assessing tactics and strategies via website analytics is paramount. Towards providing insight into this area, this research note seeks to understand what students know and discuss how they turned mostly inert capstone websites into findable assets that could help them reach employers and ultimately find a post-college job.

- **Information Collection:** This research note is based on teaching TCOM 3245: SEO and Analytics within the BS in Technical Communication program at Kennesaw State University. In addition to the observations and interactions from teaching TCOM 3245, studies on pedagogical practices and students were collected from the literature. Together, this information frames important ideas for assessing students' knowledge and developing some recommended practices for teaching SEO and analytics. Emerging from the work are five key areas seen as important for teaching students to develop content for digital platforms: terminology/vocabulary development, proficiency "reading" HTML and understanding its function, using simple tools to conduct site audits, developing SEO-leveraged content for public-facing pages and for HTML, and deploying and learning Google Analytics.
- **Conclusions:** Based on the literature, little work exists on the pedagogical practices associated with teaching SEO and/or analytics. Some preliminary work appears in disciplines such as public relations or marketing, but there is insufficient scholarly input at this time to truly define disciplinary boundaries as lying with any one area. This research note, however, situates SEO and analytics as falling soundly within a humanities focus centered on new writing practices and including analytics as a way to view, refine, and iterate new strategies for driving traffic. From teaching TCOM 3245, a case can be made for including website development concepts in writing programs, as students need solid HTML, CSS, and general technological literacies to effectively publish on digital platforms. Reframing audience is also a key requirement, as writing for the web means writing for a search engine first and a human second. Thus, content (text, images, video) must be leveraged for search engines at both a public-facing level and behind the scenes as HTML. Including analytics helps students see if their sites attract traffic, what visitors do on their sites, and provides insight into what iterations may be necessary to

increase traffic. However, analytics need to be delivered with caution in classes as the many features/options can be overwhelming to students. Additionally, students need practice in translating analytics activity into actionable strategy on a website or blog. Finally, in the case of a capstone portfolio, students need to re-situate their thinking about class artifacts. A capstone portfolio functions as a digital business asset; it must work for the students. However, unlike a paper, an online portfolio is an organic artifact, and, for SEO purposes, it must be deployed, measured, and iteratively updated in order to be effective.

- **Directions for Further Research:** With respect to developing effective content and assessing that content's effectiveness via analytics, more work is needed to understand how to best teach students to function at the intersection of humanities and technology. Currently, baseline measurements of student competencies indicate that while they understand the importance of learning SEO and analytics, most have no exposure to hands-on implementation. While some students may develop certain technological literacies in programs such as technical and professional communication, many embark on digital writing projects, in class and beyond, with insufficient skills. Finally, little is known about students' understanding of digital ethics—in particular, what they know about many of the undesirable or outright unethical work that can occur in digital platforms.

Keywords: digital content development, experience-based learning, findability, search engine optimization, website analytics, writing analytics

1.0 Aim

How students learn to develop and analyze content at the intersection of algorithmic functions, rhetorical constructs in digital spaces, and engagement as measured through basic analytics on their personal capstone websites is not well understood. Towards providing some insight in this area, this research note will discuss how students in a technical communication class on search engine optimization (SEO) and analytics began to re-situate their conceptions of what it means to be findable on the Internet. Typically, students understand the terms "SEO" and "analytics" at their most general; their anecdotal knowledge leads them to describe analytics, especially analytics like Google's, as a platform that tracks all user activity on web-based assets. They believe keywords—especially paid keywords—drive traffic to a website, and that content, beyond being keyword rich, is neither dynamic nor rhetorical. Public-facing writing is perceived as a flat dimension serving little purpose on webpages beyond informational content. Code-based content—content that will be read exclusively by search engines—was not integral to findability

in students' assessment. Challenging deeply embedded assumptions was a significant early hurdle in the class.

The capstone website for graduating seniors is the vehicle through which they present themselves to the professional world. Their sites contain samples of their work and provide general information about them, their degree programs, and their competencies. Yet, sadly, most student sites are not findable; thus, the sites fail to fulfill their primary purpose of introducing students to companies and talent acquisition professionals. Though small in scale, the data students gleaned about their findability and the success of their content in attracting both search engines and users, while retaining users as seen through engagement metrics, provided them with new insights about the function of digital writing as it is measured, iterated, and redeployed.

2.0 Problem Formation

Every semester, countless instructors assign the creation of blogs or the development of websites with the goal of teaching students to publish on digital platforms. The blogs may include standard class writing assignments, while the websites typically function as capstone or portfolio showcases designed to put students' work out on the Internet so they have a digital presence (URL) and can ultimately be found through searches done via the common engines, such as Google, Bing, or DuckDuckGo. The goal for all these digital publishing enterprises is as follows: to increase students' abilities in developing content for digital spaces and places and, in turn, enhance their proficiencies as digital authors. Ultimately, competencies in content development for digital spaces will help students with their post-college careers. Students who publish on digital platforms have trouble, though, with putting findability—the ease with which users can locate information on a website—into action (Morville, 2005). Most instructors tell students to use keywords in the copy; some may do exercises on generating multiple keywords and, in turn, develop what they consider to be SEO-attractive headings. In many cases, this represents the beginning and the end of SEO training.

As a result, the countless blogs and websites created each semester enter the black hole of the Internet and will almost never be found through a common search. Students will have effectively published into the digital abyss, and, while meeting the requirements of a class assignment, their content will never reach another human reader. Because these sites all lack findability, as implemented through SEO and measured through analytics, students miss the techno-rhetorical goal of reaching their primary audience—the search engines—and in turn, miss the opportunity to see how searchers interacted with their site via website analytics. And, regrettably, in order to rank—to gain standing with the search engines and appear in results—and produce user data from analytics, a site first has to be found.

In *The Art of SEO: Mastering Search Engine Optimization*, Enge, Spencer, and Stricchiola (2015) state that the goal of search engines is to return the most relevant results in the least time possible while reducing hits that reflect intentional, deceptive manipulations, or what we commonly refer to as spam. Making the work of the search engines more complicated are users themselves as they seek information. The actual construction of most searches or "queries" is

less about well-formed questions and carefully-considered word choices; most searches, per Enge, Spencer, and Stricchiola are "performed by users simply entering word combinations" usually with four or fewer words, and then relying on the search engine to determine user intent.

Adding to the complexity of findability is the intersection of content and technology. Search engines, as they scour the Internet for relevant content (video, text, images, etc.) for a human searcher, are essentially "reading" and evaluating the public-facing content and the tag-based HTML behind the scenes. Predicated on the words/phrases entered in the search bar (the query), a search engine algorithm uses a complex set of proprietary processes to sift through the contents of the Internet, rate the quality/relevance of each site/page, and return useful results to the human searcher. Pages then appear in order on the Search Engine Results Page or SERP through what are known as Page Rankings (PR). These rankings are the result of complex calculations, but as Enge, Spencer, and Stricchiola explain, the historical calculation behind page ranking was built on the number of links to a page. The links acted as a "vote," but this vote was also mitigated by a weighting process that ranked certain types of links (e.g., a federal government website) more highly than others. Today, the PR models are extremely complex and continue to be refined. In the same vein, search algorithms are not static; they change regularly, from major overhauls to minor updates, thus complicating how to develop effective SEO strategies. For example, Google has had several major releases over the last ten years with each release bearing a unique name. Two early game changers for SEO were Panda (released in 2012) and Penguin (five versions released from 2012 through 2014), with Fred launched in 2017 (History of Google Algorithm Updates). The changes help Google return better results to human searchers while downplaying poor content or spam. Bing, Yahoo, and others follow suit, but through their own proprietary algorithms.

Search engines function as stern critics of the pages they review. For example, Google's algorithm resembles a first-year composition instructor in its assessment: spelling, grammar, readability, and cohesion are graded. Sites whose pages do not meet established standards are then ranked lower by Google due to poor quality. At the same time, Google is also assessing the accuracy, currency, and compliance of the HTML underlying each page. HTML 5 arrived in 2015 when the World Wide Web Consortium (W3C) approved and launched this new version (W3C, 2019). Among the many changes to HTML 5, including new tags and the deprecation of older tags, was an increased emphasis on accessibility through the addition of new semantic tags. At this point, the search engines began to increase rankings of sites that helped persons with disabilities who use assistive technologies to access the Internet. Thus, a stronger hand-in-hand relationship was born between the search engines and HTML.

HTML signals Google through important tags such as the page title or <TITLE> tag, which tells Google what the page is about, but Google also approaches each page to "read" the HTML tags and follow what should be a well-structured digital pathway to information seeking and indexing. Unfortunately, Google often stumbles down poorly constructed paths or veritable dirt roads of HTML and ultimately determines the value of the page (site) to be such that it is ranked

extremely low. Thus, sites/pages will not appear in searches, as the lack of quality has been equated to low relevance, a key criterion for the results a search engine returns.

Without traffic, a website analytics program has no user data to register with respect to visitors on a site. It is important to know that website analytics programs are not collecting, at a basic level, anything new. Web servers have for years had raw data about activity, but website analytics programs, such as Google Analytics (GA), "packaged" the data in a way that was highly visual through its dashboard displays. The simplicity of products like GA make it infinitely easier to understand the path users take through a site, their geolocation, the pages they view, the time they spend, and more. This view of visitors' actions and behaviors on a website, even for student sites, functions as website developers' first glimpse into the potential of a large, active dataset that can inform the next round of SEO development and recursive measurement for a site.

While the above is only a top-level synopsis of SEO and analytics, creating and deploying findable content, and attracting sufficient traffic to a website can prove to be quite challenging. Most well-meaning instructors who assign a blog or website assignment in a class do not possess the technical knowledge to adequately include true best practices in SEO. Thus, this research note seeks to provide initial findings from the literature and present several top-level outcomes from teaching TCOM 3245: SEO and Analytics within the Bachelor of Science in Technical Communication degree at Kennesaw State University. Specifically, these generalized findings may help instructors deepen their teaching expertise about content development for digital platforms such as blogs or websites; additionally, this research note may engage others with ideas about pedagogy or more formal research in the areas of teaching SEO and/or website analytics.

3.0 Information Collection

For nearly 25 years, the Bachelor of Science in Technical Communication has been the only degree program of its kind at a US public university in the Greater Atlanta area. The degree was launched in 1995 as part of the offerings at Southern Polytechnic State University (SPSU), a small STEM-focused institution serving approximately 6,500 students. In 2015, SPSU became part of Kennesaw State University (KSU) via a state-initiated consolidation process; as a result of this merger, KSU's student population rose to just over 30,000. In late 2018, KSU achieved Carnegie R2 status and has since consolidation seen solid enrollment growth each year. At approximately 36,000 students today, KSU is one of the 50 largest institutions in the country (About KSU). For general admission to KSU, freshmen must have a GPA of 2.5, an SAT Evidence-Based Reading score of 500, and an SAT Math score of 490 for both first-year and transfer first-year students (Freshmen and Transfer Freshmen).

The Department of Technical Communication and Interactive Design (TCID) is a stand-alone department within the College of Humanities and Social Sciences at KSU. At the time of consolidation, the department had a different name: Digital Writing and Media Arts (DWMA) and belonged to the School of Arts and Sciences at SPSU. The department's mission was also

significantly different, as DWMA housed not only the Technical Communication degree, but also General Education English, Communication Studies, and Art/Drama/Music appreciation classes, in addition to offering a BA degree in English and a BA degree in New Media Arts. Charged with creating a department that did not contain offerings duplicative of English or Communication/Media, DWMA realigned its personnel and offerings by moving faculty to existing KSU schools/departments, retaining/restructuring the Technical Communication degree curriculum, and by deploying a BS in Interactive Design/UX. Thus was born the new department name: Technical Communication and Interactive Design (TCID). The Technical Communication (TCOM) degree was revised to strongly align its focus with "tech" competencies such as video, XML, accessibility, and other skills—reflecting a modernized vision of the field. TCOM 3245: SEO and Analytics functions as an elective in the Bachelor of Science degree. The course introduces students to search engine optimization and analytics and prepares them for positions in a profession undergoing rapid change.

Information was collected from the literature and from informal observations while teaching TCOM 3245; together, this information frames important ideas for assessing students' knowledge and designing effective assignments for SEO and website analytics. In terms of academic information, limited literature exists about SEO and analytics with respect to students and learning about search algorithms, large data sets, and the iterative and recursive processes necessary to attract search engines first and human searchers second. This gap may be, as Enge, Spencer, and Stricchiola (2015) claim, because almost no college courses address SEO or, for that matter, website analytics, especially in the humanities and with a humanities-based perspective. Thus, two questions emerge when discussing SEO and website analytics: (1) what discipline has primarily laid claim to these practices in terms of educating students, and (2) what literature exists on students, their understanding of SEO and analytics, and how classes have been structured to teach these subjects?

At this time, there have been no significant attempts by any one discipline to claim SEO and analytics as their area of expertise. Some areas, such as marketing and public relations, cover SEO and/or analytics in the literature, but only a minimum of scholarly work exists, and that work does not cover the actual content development as the central core of good SEO. Rather, the current available work—discussed in more detail below—examines basic but admirable SEO implementations in classroom situations through a narrow lens and demonstrates the ways in which students can grasp salient ideas. Missing though are multiple important considerations for digital spheres and what it means to work in born-digital spaces. To bridge this gap, *The Journal of Writing Analytics* sees the importance of understanding writing processes that include born-digital practices and platforms; the journal itself is actively establishing connections from writing to technology, but with a humanistic focus at the core. In "Transforming Text: Four Valences of a Digital Humanities Informed Writing Analytics," published in the inaugural issue, Palermo (2017) engages in a critical conversation about digital methods and media as writing studies begin to include more digital work that represents large, dynamic datasets. In particular, the

inclusion of ethics in the study sets a precedent for how humanities-based thinking should inform the work done in the digital domain.

A review of the literature indicates that public relations and marketing, as disciplines, have produced some work on SEO from a teaching perspective and, at the very least, attempt to lay some claim to SEO as their area of expertise. It is important to note that the following articles provide teaching frameworks for SEO, but do not situate SEO as intersecting with GA or big data in general. From a public relations (PR) point of view, Moody and Bates (2013), in "PR Students' Perceptions and Readiness for Using Search Engine Optimization," examine the critical nature of SEO for PR work and the preparedness of students to conduct that work. Reiterating the lack of relevant literature about SEO and students, the authors note that hundreds of trade articles have been written about SEO and its implementation; however, few scholarly studies have been published about courses and pedagogy related to SEO. Moody and Bates indicate that, as a disciplinary area, PR functions as a likely home for SEO, and that such knowledge would be extremely beneficial for the current workplace and beyond. Moody and Bates determined more about what students know about SEO through a survey designed to assess baseline knowledge, determine how competencies were learned, and gauge the accuracy of knowledge. Moody and Bates confirm students' knowledge of SEO is limited, but that student awareness of the importance of SEO for their future professions is high.

Clarke and Clarke (2014) tackle SEO from a marketing (business) perspective in their article, "A Competitive and Experiential Assignment in Search Engine Optimization Strategy." The authors echo the findings from Moody and Bates regarding the dearth of literature and lament that pedagogical research regarding SEO remains negligible. Attending to the public relations students of Moody and Bates, Clarke and Clarke found that marketing students understand, in the same general sense, the importance of high organic search engine rankings to an enterprise, but they are unsure how to actually implement techniques that could increase rankings. In their study, Clarke and Clarke had students work on a website or blog they control; during the semester, students completed a pre-test on the SERP (Search Engine Results Page) status of their site. They then used an assigned keyword phrase to increase their rankings, and concluded the semester with a post-test. Findings indicate that students gained considerable knowledge about implementing SEO. Clarke and Clarke see courses such as e-Commerce, Social Media Marketing, and Public Relations as logical homes for teaching SEO; however, they did not mention areas such as technical or professional communication as possibilities. Overall, Clarke and Clarke conclude that an SEO technical skillset is one that will make students highly marketable.

As with SEO, there is scant academic literature on the use of GA in a classroom setting. Meng, Jin, Lee, and Kim (2018) discuss their study on implementing the GA Certificate course in a public relations research class. As with the literature on SEO, Meng, Jin, Lee, and Kim found that little public relations pedagogy research has focused on the implementation of GA in the PR curriculum. Reflecting the results in the SEO articles, students in the Meng, Jin, Lee, and Kim study saw the value of learning more about GA for their future careers, but had low levels

of knowledge about the inner workings of analytics. Adding the GA Certificate course to the research class increased students' understanding about analytics and the data it delivers.

Based on the literature review, this research note is forging new ground. As noted above, very little academic literature exists regarding teaching students about SEO and analytics and, in particular, situating this practice around the idea of large data sets and iterative work. Petersen and Martin (2015), while not discussing pedagogy in their article "Misuse, Play, and Disuse: Technical and Professional Communication's Role in Understanding and Supporting Website Owners' Engagement with Google Analytics," situate technical and professional communication as a likely epicenter to work with the big data generated by (GA). Petersen and Martin found in their study that most of their participants did not use or fully understand the data GA generated. As a result, they emphasize that a growth-area role for technical communicators would be working with site owners to unpack the complexities of the data, understand the business intelligence generated by GA, and what it means for modifications to their online assets such as websites.

In sum, the research discussing how to teach students to leverage websites for SEO, measure the outcomes via website analytics, and iterate changes is limited. No articles exist that discuss how students can leverage SEO and website analytics as they work on their own websites as real-time laboratories. TCOM 3245: SEO and Analytics for Technical Communicators thus presents a one-of-a-kind opportunity to engage students with the ideas of content development for algorithms and website analytics as a large dataset that can be used to measure the success or failure of SEO implementations. Following are top outcomes related to what students needed to know as they worked on their capstone websites in preparation for the showcase and their upcoming job searches.

3.1 Student Learning: Where Do Students Begin?

Students in TCOM 3245: SEO and Analytics create content—both as public-facing and as embedded within the architectures of HTML and CSS—to attract a machine, otherwise known as an algorithm. The machine then directs human users to a website, and analytics compile user activities into a large dataset that, when analyzed, can inform strategic changes to a website that in turn attract more traffic. For the purpose of the class itself, the description is less lofty: students' take existing personal websites, leverage them for SEO, measure and assess visitor traffic to the site via analytics, make any necessary refinements, and prepare the site for display at the capstone event, the Student Showcase (capstone). From teaching TCOM 3245 each fall since 2016, it is clear that students arrive in class with varying degrees of certainty and trepidation about the course and its content. Some believe they innately *know* how to leverage SEO and, because of some limited exposure to analytics, have a solid understanding about the role of analytics for digital platforms. Others have arrived in the class with a bit of curiosity and considerable amounts of fear—fear that the "technical" nature of the course will be overwhelming. Students are generally pleased that they've managed to build their own website in the pre-requisite course (TCID 3400: Front-End Development); others, who were less happy

with the website outcomes from the pre-requisite, come to class with sites they've developed on popular blogging platforms (i.e., WordPress) or in one of the purportedly "easier" template-driven webdev (website development) platforms (i.e., Square Space). Because the class is, at the outset, defined as not a WordPress class, those students have a unique opportunity to re-hone their competencies from TCID 3400 and build a site. Students who choose to stay with a template-driven site create a unique set of challenges for themselves as many template-driven platforms keep end users away from the nuts and bolts operations that are actually key to good SEO.

Congruent with the findings of both Moody and Bates, as well as Clarke and Clarke, most students in TCOM 3245 know the value of learning SEO and website analytics, but very few have knowledge beyond very basic implementations; fewer, of course, have ideas that are completely correct. A small percentage have gleaned some knowledge through their own explorations (i.e., YouTube videos). Most students begin the class believing paid ad words are the secret to SEO success. While money can buy one a position on the first page of the search engine results page (SERP), search engines still prefer sites to seek organic traffic.

3.2 Five Key Areas for Learning

TCOM 3245, as noted earlier, introduces students to two important ideas: (1) SEO is about reaching a machine (algorithm) first and a human second, and (2) user actions/behaviors on a website are measurable and, as a large dataset, can be analyzed and used to improve rankings with the algorithms. While it is not the intent of this piece to cover all the content of TCOM 3245, five main areas are significant, especially for those who use blogs or websites as part of teaching. These areas emerged during the teaching of TCOM 3245 for three consecutive fall semesters, 2016 through 2018:

1. Terminology/vocabulary: Defining the field
2. Reading HTML: Gaining a new literacy
3. Using simple tools: Achieving big improvements
4. Developing content: What search engines want to see
5. Google Analytics: Deploy and learn

3.2.1. Terminology/vocabulary: Defining the field. Learning the terminology/vocabulary surrounding SEO marks a key step in increasing students' understanding of the practices and processes. Two of the earliest definitions, direct search and organic search, open the door for students to see that their sites are not findable. When asked, "How would a potential employer find your site and view your work?" most students reply with "They can type in my URL." However, when asked what an employer would do if they did not know the personal URL but were looking for a recent technical communication graduate with knowledge of SEO and

analytics, the students have that initial glimmer of understanding about SEO and findability. Learning that an organic search occurs when a user is entering keywords or phrases they hope will lead them directly to the site they want or to a site they might consider a serendipitous find emphasizes the importance of findability on the Internet. An organic search for universities in Georgia may be "large universities in GA," and from the search engine results page (SERP), the user will browse and select from the various links returned. A direct search occurs when a user knows the exact URL of the website they seek. For example, a user who wanted to visit the website for Kennesaw State University and who knew the URL would type "Kennesaw.edu" as their query. Anchoring these key concepts forms a cornerstone of the class and prepares students to adapt to countless unfamiliar terms and acronyms.

Students also enjoy learning some of the terminology related to disruptive practices within SEO; especially, they are eager to absorb information about practices from the "dark side." Discussion on the hats of SEO—black hat, white hat, and grey hat—are always animated. Students sit with rapt attention to explanations about *black hat* SEO work—work that intentionally sets out to fool the algorithm and drive more traffic through deceptive practices, such as hidden text. The application of *white hat* to quality SEO work and *grey hat* to work that is dubious but passable always resonates with classes and carries across the semester. Terms like *hijacking* (making the search engine believe a website is at another URL or *Google Bowling* (black hat sabotage practice of pointing questionable links to a competitor's site) again led to lively discussions about ethics and digital spaces.

3.2.2 Reading HTML: Gaining a new literacy. Reading HTML—a technological literacy—has proven to be a critical skill for developing websites that work well for SEO. Reading the HTML underpinning web pages is a first step in understanding if the HTML is current (HTML 5), compliant (all valid tags, no deprecated tags), and well-formed (proper structure/hierarchy). Reading HTML is a matter of knowing what information should occur on the page, where it should occur, and if it is structured correctly. This "read" functions as a first step to assess the SEO health of a page/site and determine the first phase of revisions. Students need to be able to read to ensure the correct document type declaration has been made, i.e., <!DOCTYPE HTML>, the <META> elements are complete, the <TITLE> of the page is unique, and the major headings <H1 – H6> are orderly.

However, the ability to read HTML is a competency many students lack. One reason comes from the purpose of the class in which they are creating blogs/websites. Not every class is a technical class—one focusing on the nuts and bolts aspect of digital platforms. Blogs make the concept of publishing on the Internet palatable because of the simplified interface; as a result, blogs can be very popular in classes focused on content development. These platforms are favored especially by faculty who are not technically inclined. A second reason for limited student proficiencies can actually stem from website development classes themselves. The web development process itself can hinder students' ability to read HTML well, especially if they were not required to hand code for at least part of their web classes. Many website development classes use site builder programs such as Adobe's Dreamweaver. Thus, while students learn to

build their website and upload it to a host server, they do not always learn to initially read HTML nor heed the requirements to use HTML cleanly and appropriately and avoid what Enge, Spencer, and Stricchiola (2015) refer to as HTML bloat (formatting that could be moved from HTML to CSS). It can be more important to merely have the site "work"—as in *function* for the purposes of the class—than for it to be correct for SEO purposes.

A third reason students may not be proficient at reading HTML originates from the use of template-driven sites. Students who purchase a template-based website from the slew of companies proffering beautiful designs and no need to learn HTML put themselves at an immediate disadvantage. Granted, the beauty of the templates is seductive for students; the refined images and gentile typography give way to magical thinking about beauty and attraction. But search engines do not rank on looks, so even the most sublime design will not organically result in more hits to a page. The ease-of-use factor in template platforms keeps end users from exploring the HTML code (though it is available) and the overly-featured nature of these platforms means much of the functionality underlying each page remains interwoven between complex scripts and code which can be difficult to parse and understand, especially for novices. All of the students who worked with a commercial template-driven website experienced difficulty in TCOM 3245. While they thought the platform would be a distinct advantage, it proved to be challenging, if not impossible, to do many of the activities required in the class. With lesser-known packages (read: less expensive), considerable amounts of functionality were not available.

3.2.3. Using simple tools: Achieving big improvements. There are a variety of free tools available to assist students in identifying areas for improvement that will, when implemented, make their sites more attractive to the algorithms. These tools produce helpful inventories that students can then use to guide them through changes that will make their sites more compliant. In TCOM 3245, students did an initial review of their websites for compliance with accessibility requirements (W3C-WAI or Website Accessibility Initiative) and for compliant HTML 5 (W3C). Most were taken aback by the number of problems, both great and small, that these free tools found. The WAVE web accessibility evaluation tool from WebAim finds accessibility problems such as images with no ALT-TXT or alternative text. For users who require assistive technologies to support or supplement their visual abilities, screen readers are the window to the world of the Internet. The screen reader uses voice to narrate the page to the end user; thus, each image, graphic, table, or chart needs rich, descriptive text written to be about 125 characters (including spaces) in total. ALT-TXT also functions as content for SEO; search engines read this HTML tag and use the ALT information in their overall indexing of a site. The report generated by WAVE provides a comprehensive listing of accessibility issues and color-codes them from critical to minor; students generated the report multiple times over the semester as they worked on their sites to ensure they were cleaning up legacy errors and not introducing additional problems.

Using the W3C's Markup Validation Service for HTML (W3C), students were able to check the HTML of their sites to ensure compliance with the new standard (HTML 5). Most were

somewhat surprised to learn that their HTML needed work—empty tags, unclosed tags, deprecated tags, and unnecessary tags were only some of the common problems. Interestingly, students were, as noted above, unaware of the tag-level problems, as they assumed the site development software made everything correct. In actuality, most students had hierarchical errors with their use of heading tags. For example, students had used multiple <H1> tags on their pages when, for SEO purposes, a page can only have one <H1> tag. Students also applied the <H2 – H6> tags in random order when, in practice, the numbering system represents a hierarchy from major headings to sub-sub headings. Like the WAVE tool, the Markup Validation Service provided a detailed report highlighting problems areas that needed attention in order to conform to the best practices for HTML 5. Students ran this checker throughout the semester as they improved their sites. It is important to note that both the WAVE tool and the Markup Validation Service require students to understand ("read") HTML and be able to make changes directly to the tags. This underscores the necessity for literacy with HTML.

3.2.4. Developing content: What search engines want to see. Content development for the students' sites proved to be one of the bigger challenges in TCOM 3245. Most had, prior to arriving in the class, developed very little content (text, video transcripts, etc.) and instead had focused on the ephemera of design which, as stated earlier, search engines are unable to judge. In addition to design (color, typography, logos), students had posted images of their class projects on pages, but provided almost no descriptive narrative about their projects, scope, goals, and so on. In instances where narrative existed, students were pleased to point out that they had used keywords to increase SEO, but they admitted to just guessing at what words may have traction and used the same words repeatedly on the page. As discussed earlier, students assume that beyond paid ad words, keywords are the holy grail of being found on the Internet by search engines. However, keywords are only part of the findability equation, and if the same keywords are used multiple times on one page, the problem of *keyword stuffing* rears its head and the site will rank lower. In a similar fashion, if the same keywords appear throughout all of the site's pages, the issue of *keyword cannibalization* occurs. As the search engine cannot discern which page is more important when all the keywords are the same, the relevancy of the keywords becomes diluted (Enge, Spencer, and Stricchiola, 2015). Finding good keywords is important, and in TCOM 3245, the Explore feature of Google Trends proved to be a rich starting point for students. Here, they could enter terms, see which terms were on the rise, and begin the process of brainstorming for their written content. Brainstorming also consisted of developing key phrases that could be included in content—this was a technique students had not considered before. With the focus of their sites centered on finding a post-college career, exercises in reverse-engineering searches were important. For example, the question posed to the class was "What would a talent acquisition professional enter as a query in order to find you?" From this point, students began to generate detailed knowledge of course goals such as "technical communication graduate with experience in search engine optimization."

Keyword research remains important to SEO, but those keywords need to live within written content that meets specific requirements. At the outset, there must be content and, as noted

above, many students had only created a bare minimum of text—sometimes just two or three sentences for a page. While there is no required minimum or maximum word count for SEO and content, web pages do need an amount appropriate to the business goals of the site's owner and to the topic. This proved to be a novel idea to students who had not thought of their capstone sites as a type of business asset that needed to work for them as they sought jobs. However, what gave students additional pause came from the actual requirements for the content with respect to best practices for SEO.

To begin, a website's public-facing content must be well written; this includes text that is grammatically correct and free from common problems such as spelling or diction errors. Search engines function as arbiters of good content and view content that does not meet accepted grammatical standards as less reliable (and possibly as spam). Second, the content must also be at the correct readability level for the topic. A site about children's craft activities should not be written at the level of an engineering journal, for example. Doing so will again trigger the algorithm that the site may not be reliable. Finally, the content and the topic of the site must match. Google's algorithm checks content carefully to ensure not just quality, but to ensure a match (cohesion) between the content and the topic of the site. As an example, a site about soap making should not reference pastry recipes; such mismatches can function as an indicator that the site may be spam. For students, developing written content had already been a challenge; however, learning that they needed to increase both the quantity and quality of their content resulted in unexpected work.

In terms of other types of content, students learned to optimize images (reduce the resolution) for the web and include high quality ALT TXT that would also work for SEO. For those uploading video projects, captions and transcripts were required, as both are findable by search engines and increase SEO. An exercise in developing the <META DESCRIPTION> content also proved to be useful. This HTML tag contains ~150 characters that form the first two sentences that display under an entry on the SERP. While this content isn't read by the search engines, it is a vital piece of content for the human readers and can make or break a site's click-thru rate. The last piece of content students developed was their résumé. All students wanted a résumé on their site so it could be downloaded by a potential employer. As such, students were required to develop an accessible PDF (compliant for screen readers) and complete the Properties section with rich, descriptive metadata. This technique was new to every student, but proved to be worthwhile as, for several, the PDF résumé was the first artifact that was indexed and appeared on the SERP or search engine results page.

3.2.5. Google Analytics: Deploy and learn. GA, as a free program providing website analytics, is extremely powerful. Designed for marketing and e-commerce, there are myriad sophisticated features that enable web masters to track and analyze users' behaviors on websites. As mentioned earlier, the success of GA comes from its visual dashboard displays of the actions and transactions on a website. Pre-Google, server logs provided lists of unformatted data that were neither easy to obtain nor available in an easy-to-read format. GA changed that for the better and for everyone.

Setting up GA was a vital part of TCOM 3245. Students had to create an Analytics account, add their property (website URL), and include the generated snippet of Javascript-based analytics code on all their pages to capture user actions. This task momentarily constituted one of the earlier identified areas of technical trepidation for students—they had never set up analytics before and were uncomfortable. However, in a few minutes, when they determined the process to be easy by following the instructor's directions, all was well. One of the biggest assumptions about GA was immediacy: results would now be streaming into the analytics dashboard within mere minutes of deployment. That however, is not the case; a general crawl—the reading of a websites' pages for indexing—takes time. Even with the submission of an XML site map through Google Search Console, a free Google service that, among other features, confirms that Google can find and crawl a site, indexing may take several weeks. Following the students in the Meng, Jin, Lee, and Kim (2018) study, students in TCOM 3245 completed the GA for Beginners course from GA Academy (2019) as part of the course requirements. Students reported enjoying the certificate course and many noted that they would add the certificate credential to their LinkedIn profile and their résumé.

Once the capstone sites were indexed and data appeared in GA, students were very much like the website owners described in the Petersen and Martin (2015) article; they tended to "play with the data as a curiosity or an entertainment" rather than viewing it as real-time, actionable business intelligence that could drive iterative changes to their sites' SEO. Students reported that part of their inaction came from the overwhelming amount of information GA provides and the seemingly endless data that can be generated through myriad customizable metrics and dimensions. It caused them to freeze. Thus, walking through major reports and drilling into the results were an important part of the class, as was asking, "So, you see [X] in your analytics data. What would you do to your website?" Students were quite unsure about how to translate findings from analytics into actions. As a result, Q and A sessions were important for helping students make sense of the dataset analytics provided.

Students appreciated that they could see what visitors looked at on their capstone sites. While they made some gains in understanding that the data should lead to actions, analytics functioned more as a "window" through which to view activity than as a lens to actively shape and re-shape SEO techniques on the capstone sites. This finding aligns with the findings of Petersen and Martin in which the site owners studied did little with the analytics data and were overwhelmed by it, almost to the point of inaction, but fascinated by what they saw.

4.0 Conclusions

Based on the outcomes from TCOM 3245, having students work on their own websites to improve SEO and measure those improvements through analytics constitutes an important new direction for teaching and learning. Rhetorical appeals have shifted from those made to a human audience to those made to a machine. The data returned from analytics provides students with actionable insights into the success of their websites, from attracting users to the sites to seeing user interaction and engagement with the material. Thus, students can iteratively refine their

SEO tactics and strategies. First, though, it is important to note that this research note covers only some of the implementations from TCOM 3245; the class, in its 16-week iteration, dove deeply into a variety of technical areas (XML sitemaps, custom 404 pages [page not found], etc.), backlinks, social signals, and more. However, several outcomes from this paper can inform teaching practices for those who engage students with digital publishing and platforms. We know that students are knowledgeable about the importance of SEO and analytics; however, it is paramount to keep in mind that this type of class changes students from consumers, those who search for information on the Internet, to producers of searchable content. Much of what this class covered required students to engage in new practices, technologies, and ways of thinking about their work. The content of the class prepared students to think about a type of profession that varied considerably from their expectations for technical communication; however, students did not expect the level of work, both in terms of detail and volume, that would be required to take a beginner's level site to one that was reasonably well-optimized for SEO. The experience of teaching this class also brings to the fore several considerations that are significant for delivering a similar offering. These four areas for consideration are discussed in the following sections.

4.1 Website Design Class: A Necessary Prerequisite

Students without a pre-requisite course in basic website development may have more difficulty with a class focusing on SEO and analytics. TCOM 3245 students were presumably in a better position than most, as they had taken the pre-requisite webdev class (TCID 3400), an offering common to many technical and professional communication programs. They had some exposure to HTML, CSS, and the overall technical functionality of websites in general. This meant they were able to make changes directly to their own sites and gain critical real-time experience upgrading for SEO, deploying analytics, and viewing results. Classes that wish to engage students with digital platforms and publishing should include some basic web publishing principles that focus on technical definitions/functions and on HTML, with an emphasis on learning what the tags do and how to "read" them. This competency will be invaluable as students learn to engage actively with search engine optimization practices.

4.2 Shifting Ideas of Content Development

Reframing the idea of audience and continually reinforcing that across the semester proved to be important for students' understanding. Developing SEO-leveraged content for a website requires a radical shift in thinking, as what is read and by whom has changed significantly with digital platforms. Search engines perform complex "reads" of the pages on the Internet. They examine the public-facing content as well as the code and tag-based content underlying each page with the goal of determining, at a basic level, overall quality and match to the user's query. It is, however, easy for students to forget their "audience" in SEO; unintentional recidivism towards the human-reader-first model happens because students have no previous classes that have focused on reaching the algorithm first. Beyond the machine, there are other important

implications for students and their content development work, many of which were unexpected. TCOM 3245 students did not expect their websites would need so much content review, revision, and re-checking at all levels in order to be properly leveraged for SEO. Students also had knowledge gaps across a variety of technical areas that influence content. As a result, filling those gaps and gaining proficiency with new techniques and technological practices formed an integral part of the class.

4.2.1. Public-facing content. Students who had produced little public-facing content to begin with were at somewhat of an advantage as they were not deeply invested in their work; others, who had created significant amounts of content and felt they had written a personal *magnum opus*, were dismayed to learn it would need a total reconsideration. However, for everyone, revamping the rhetorical approach proved to be challenging as it was a matter of working backwards from "What would a searcher enter into the search bar (query) to find you?" Brainstorming user searches in three- to five-word strings, researching keywords, and then constructing the actual content of the pages—without keyword stuffing, keyword cannibalization, and with good natural language phrases that would appeal to a search engine—was not an intuitive model for college students. It is important to remind students that their content must function at the intersection of what a human enters as a query and what the search engine can find across the entirety of the Internet.

4.2.2. HTML-level content. Developing quality content that functioned on the HTML side, such as the <META DESCRIPTION>, <TITLE>, and the <ALT TEXT> proved to be another activity that students had not expected; they were somewhat reticent to admit their lack of familiarity with these tags. In fact, the idea that strategic SEO content lived within the HTML components of a page was new to most. As noted earlier, the <META DESCRIPTION> appears under the SERP result and is the two-sentence description of the page. While it is not read by a machine, it is read by a human and must be sufficiently persuasive for the human searcher to decide to click on the page. When asked, most students knew the location of the text on the SERP but did not know where in the HTML the tag belonged, how to add the tag, or how to write a meaningful description within approximately 160 characters for the tag. Class time was spent developing this snippet and subjecting it to peer testing through simple mock-ups and comparisons. Students needed several revisions to create the appropriate "hook" for a human reader, but each student site had a reasonable <META DESCRIPTION> that functioned to invite a reader to their site at the end of the class.

Other tags, such as <TITLE> were not well understood or properly implemented. The <TITLE> tag tells the search engines the title of the page, and each page requires a unique title of approximately 50 to 60 characters. Multiple pages with identical information in the <TITLE> do not help search engines differentiate content between pages, and duplicative tag information can actually hinder SEO. Most students, if they had completed the <TITLE> tag, had used their name in the form *Firstname Lastname | Portfolio* across each page in their site. A few had an About page with the title, *Firstname Lastname | About*. Unless a query was based on the individual's name, these formats did little to help the search engines glean page-level

information about students' work, competencies, and other content in their portfolios. Here, it was again critical to spend time workshopping best practices in tag development for findability.

In contrast to <META DESCRIPTION>, <ALT TEXT> functions as a machine-readable element on a webpage; it serves to make an image, chart, or table, accessible for screen readers and is read by search engines. Students had a passing familiarity with this tag, but most had not deployed the text, as it did not seem important in the context of the pre-requisite class. In TCOM3245, students practiced writing <ALT TEXT> with a picture of the Eiffel Tower and a prompt that the picture would be on a French bakery site. Most wrote "*Picture of Eiffel Tower*" on their first attempt and felt this was adequate. It was necessary to articulate the idea of "rich text"—descriptive text that would help a visually impaired person "see" the image, but still serve SEO purposes via good keyword placement and correct writing. Teaching students to remove "Picture of" or "Image of" was another important step with the <ALT> tag exercise. Where a complex chart or table would require a longer description, students were introduced to the <LONGDESC> or long description tag for extended content. From the experience of teaching TCOM3245, providing ample time to merge technical issues with content development issues would be vital for any class.

4.3 Teaching Analytics: Narrowing the Focus

Web analytics, such as GA, have revolutionized how we understand user behavior and interactions on websites. The data—as an aggregate of all users' interactions over time—provides actionable insights for tactical and strategic SEO redeployments. It truly is writing analytics in real time. However, GA, while free and easy to deploy, provides an overwhelming amount of data not just for page acquisition and user behavior, but for ecommerce functions such as conversions (converting from a searcher to a buyer). As student sites are informational and not ecommerce based, the defaults from GA can be confusing and/or distracting if not overwhelming. Students required a curated tour through analytics to understand what functions are relevant to their sites and to briefly learn some of the other functions. Unfortunately, it is impossible to fully understand all that analytics can offer; there is simply so much functionality in Google's product as to be beyond the scope of any one-semester class. Additionally, students can, as new users, neither process nor apply all the features in analytics. It is beyond their abilities and the scope of their websites—not enough time passes in one semester to collect a robust body of data for in-depth analyses. This fact—the time required to generate a suitable amount of traffic—surprised students. They thought that the data would be instant with traffic accruing on their sites after only a few days; unfortunately, indexing and attracting real traffic can easily take weeks.

4.4 Class Artifact or Digital Business Asset

"What is a website and what is its purpose?" proved to be a compelling question to insert into a casual class discussion. With most 20-somethings having grown up in a world where websites have always existed, the answers varied greatly, especially as we were not including

ecommerce/retail sites. From the informal discussion, websites were defined as places to post hobbies/interests, personal creative projects, or passion projects (e.g., environmental issues); however, social media channels and, in particular, YouTube, appeared to have more currency and/or authority. Websites or blogs built in high school or other college classes had been done for specific assignments, but served no other function. For the most part, students admitted they had not fully conceived of the greater rhetorical purposes of their portfolio website. They had not considered that their capstone sites needed to function at a level beyond a digital billboard or shingle (seen by chance) and that, as a digital business asset, the website had to work for them to achieve their "business" goals (attracting potential employers, self-promotion, etc.). Shifting mindsets that saw capstone sites as just one of many artifacts from college classes—created solely for the purposes of a grade—proved to be challenging. Not only is a website a business asset, but it functions as an organic asset; it requires updates and regular "care" to work well. While they had taken a class on SEO and analytics, the nature of the class with its ongoing cycle of deployment, measurement, and iterative development meant the artifacts from the class were never truly finished—finished in the way a paper or other project might be. Rather, SEO practices required continual implementation, testing, and adjustment with tracking logs of what was done and when, plus the results. A fatigue was evident among some students who wanted a real sense of completion versus a project that never seemed to end.

5.0 Directions for Further Research

Classes that engage students at the intersection of technology and humanities-based thinking can contribute significantly to competencies necessary for the future of work. No longer can content be uploaded with the hope that some far and away reader may find it. It is vital, especially for any business dependent on online traffic, that their digital assets can be found and that user behavior and engagements can be measured. In addition to including SEO basics, both as a technical and content component in classes, students also need more exposure to and practice with datasets such as those delivered by GA. Specifically, humanities researchers should explore the best ways and means to teach how real-time, actionable business intelligence can be derived from datasets. Learning to "read" large-scale data and make determinations about what it means will be a necessary competency for students but a valuable one, according to Petersen and Martin, who conclude in their 2015 study that training and hands-on practice to maximize the benefits of using analytics platforms will be an excellent competency to possess, especially for technical and professional communication students.

Are students, however, adequately prepared to work in areas that merge new and old practices into an emerging profession such as SEO and analytics analysis? Students across the three sections of TCOM 3245 left the class with an increased confidence for both learning unfamiliar material and working in real-time with technology. They gained experience with implementing SEO and learning about how visitors interact with their websites. However, were they truly prepared for a 3000-level class that merged technology and content development for what was effectively some truly heavy lifting of new skill sets? With that in mind, further

research should consider more assessment of students' baseline knowledge of working in digital places and spaces. Beyond the tech and content components, what knowledge of digital ethics do students possess, and do they understand the ramifications of spam, black hat work, and more in the digital arena? In an era where social platforms promulgated false information that, in turn, had national influence, we are obligated to ask what frameworks students have for assessing ethical practices and identifying without hesitation dubious requests and other intentional online deceptions. Ultimately, we are only just beginning to understand what students will need to know for their future digital careers.

Author Biography

Laura A. Palmer is a Professor and Chair of the Department of Technical Communication and Interactive Design (TCID) at Kennesaw State University. She is currently serving a two-year term on the Board of Directors of the Society for Technical Communication (STC). Her research interests include issues in digital accessibility in addition to teaching and learning practices for search engine optimization and website analytics.

References

About KSU. (2019). Kennesaw State University. Retrieved from https://www.kennesaw.edu/about/

Clarke, T. B., & Clarke, I. (2014). A competitive and experiential assignment in search engine optimization strategy. *Marketing Education Review, 24*(1), 25

Enge, E., Spencer, S., & Stricchiola, J. (2015). *The art of SEO: Mastering search engine optimization* (3rd ed.). Boston, MA: O'Reilly.

Freshmen and transfer freshmen. (2019). Kennesaw State University. Retrieved from https://admissions.kennesaw.edu/apply/freshmen.php

Google. (2019a). Mobile friendly test. Retrieved from https://search.google.com/test/mobile-friendly

Google. (2019b). Google Search Console. Retrieved from https://support.google.com/webmasters/answer/9128668?hl=en

Google. (2019c). Google Trends. Retrieved from https://trends.google.com

Google Analytics Academy. (2019). Google Analytics for beginners. Retrieved from https://analytics.google.com/analytics/academy/

History of Google algorithm updates (n.d.) *Search Engine Journal*. Retrieved from https://www.searchenginejournal.com/google-algorithm-history/

Meng, J., Jin, Y., Lee, Y.-I., & Kim, S. (2018). Can Google Analytics certification cultivate PR students' competency in digital analytics? A longitudinal pedagogical research. *Journalism and Mass Communication Educator*. https://doi.org/10.1177/1077695818816916

Moody, M., & Bates, E. (2013). PR students' perceptions and readiness for using search engine optimization. *Journalism and Mass Communication Educator, 2*, 166.

Morville, P. (2005). *Ambient findability*. Sebastopol, CA: O'Reilly.

Palermo, G. (2017). Transforming text: Four valences of a digital humanities informed writing analytics. *The Journal of Writing Analytics, 1*. Retrieved from https://wac.colostate.edu/docs/jwa/vol1/palermo.pdf

Petersen, E. J., & Martin, B. M. (2015). *Misuse, play, and disuse: Technical and professional communication's role in understanding and supporting website owners' engagement with Google Analytics*. 2015 IEEE International Professional Communication Conference (IPCC), 1.

WebAIM. (2019). Web Accessibility Evaluation Tool. Retrieved from https://wave.webaim.org/

W3C. (2019). Markup Validation Service. Retrieved from https://validator.w3.org/