Discovering Utilization Patterns in an Online K-12 Teacher Professional Development Platform: Clustering and Data Visualization Methods

Javier Leung
University of Missouri-Columbia
Teacher Professional Development, Educational Data Mining
2605 Spanish Bay Drive
Columbia, Missouri 65202
850-777-6131
leungj@missouri.edu

Abstract

Teachers have many options for continuing education and professional development. Online platforms are one option that can provide 24/7 access to professional development resources and communities for practice and support. By examining how teachers use these sites, we can offer suggestions for improvement and continued engagement. This paper describes an analysis of utilization patterns of instructional resources and site features in one online professional development platform for K-12 teachers in the state of Missouri in the United States. A clustering algorithm was applied to a large dataset of web analytics records to find patterns of use of professional development materials and site features. Prominent patterns suggest that teachers seek concrete video examples of teaching strategies and examples of units of instruction, and school administrators review teachers' journals and perform administrative site functions. Teachers who are new visitors use the journal tool for initial reflection activities, but they do not access the journal tool in the future. Small patterns provide additional knowledge about the most utilized resources and the issue of referrer spam.

Keywords: teacher professional development, educational data mining, web analytics, data clustering

Introduction

With increased access to the Internet, teachers have access to multiple opportunities for teacher professional development (PD) online. Carter (2004) identified four characteristics of a successful online teacher professional development: (1) 24/7 access to online resources, (2) community of learners, (3) community of practice, and (4) professional support. Each of these characteristics is designed to support the immediate PD needs of teachers through online self-paced resources.

The EdHub Library is an online PD platform for K-12 teachers and school administrators maintained by the College of Education at the University of Missouri-Columbia. The EdHub Library is part of the Network of Educator Effectiveness (NEE). The Bill and Melinda Gates Foundation initially funded the EdHub Library in 2014 in partnership with the University of Missouri-Columbia. NEE is a comprehensive educator evaluation system that tracks multiple measures of educator effectiveness, including classroom observation of teachers, units of instruction, professional development plans, and student surveys (Network for Educator Effectiveness, n.d.).

Similar large online PD platforms exist, such as PE Central and MyTeachingPartner. PE Central is an online professional development platform for physical education teachers. With 162,000 visitors, PE Central provides lesson plans, best practices, classroom management, use of technology, and assessment ideas (Hanson, Pennington, Prusak, & Wilkinson, 2017). MyTeachingPartner (MTP) is another example of online self-paced resources where teachers have access to three types of resources that aim to improve teacher-student interactions: a video library of best practices of teacher-student interactions, a college course, and individualized web-mediated coaching sessions (MyTeachingPartner, 2017).

The purpose of the EdHub Library is to support K-12 teachers, principals, university teaching assistants, teacher prep students in the advancement of best practices in teaching and learning in PK-12 and higher education levels (EdHub Library, n.d.). School districts with a subscription to NEE have access to (1) a video library of best practices in classroom teaching, (2) a video library of examples for scoring classroom observations, (3) a catalog of self-paced online modules, (4) copyrighted assessment instruments, and (5) a yearly principal calibration training.

In the EdHub Library, teachers can browse online resources from a list of topics and search content organized by teacher standards. EdHub can be accessed directly on the homepage at *www.theedhub.org* or through the NEE online teacher evaluation tool that tracks multiple measures of teacher effectiveness. Online resources are presented in modules with several activities that prompt teachers to reflect on their teaching practices or generate new ideas on instructional strategies for classroom implementation using the journal feature. Teachers have control over sharing journal entries with school administrators and reviewing their activity history from the user dashboard. In addition to the site features available to teachers, school administrators have administrative permissions that allow them to review teacher activity and control access within the school group.

While we know about the purpose of the website and how many users are registered on EdHub, little is known about what users do once they are on the website. The purpose of this study is to discover the utilization patterns of content and site features using a clustering method and data visualization from 460,721 records available from October 2015 to February 2017.

Google Analytics (GA) data are collected in the form of dimensions and metrics using a script embedded in the site and online resources. Dimensions are qualitative attributes of the data such as the page, type of browser, and location. Metrics are quantitative measures of data such as time spent and number of views of a page. To make sense of the data collected, combining dimensions and metrics generates insights from user interactions with online resources and platform features.

Literature Review

This section outlines existing literature related to the components of effective teacher professional development, evaluation of in-person and online PD, and research design for investigating utilization patterns of professional development platforms.

What We Know About the Impact of Teacher Professional Development

Multiple studies have investigated the components of teacher professional development and the impact of PD on teaching practices and teachers' professional careers. Desimone (2011) argued that the core features of effective teacher PD lie in content focus, active learning, coherence, duration, and collective participation. Borko (2004) described teacher PD as a contextualized educational system that consists of a PD program, facilitators as the providers of PD, and teachers as recipients of PD. Guskey (2002) argued that the ultimate goal of staff development is to address the motivational aspect of teachers engaging in professional development. A change in teachers' beliefs and attitudes occurs when teachers participate in PD and change their classroom practices when teachers see evidence of increased student learning outcomes.

Bechtel and O'Sullivan (2006) described contextual aspects in schools that affect teacher change, including school culture, micropolitics of schools, level of support, and workplace conditions. Coldwell (2017) argued that highly engaged teachers in PD are very likely to stay in the profession because they are given opportunities to improve and validate their knowledge that makes them more motivated and confident as science teachers. Rice and Dawley (2009) surveyed 259 teachers, administrators, and instructional coaches from American K-12 virtual schools to understand professional development practices. Their findings suggest that online teacher professional development should be characterized by offering (1) multiple opportunities for professional development, (2) ongoing training sessions, (3) coach or peer mentoring sessions, (4) student-focused activities, (5) customized pieces of training based on teacher needs, and (6) trainings aligned to standards.

Evaluating Teacher Professional Development

While multiple studies investigated the characteristics of teacher professional development from the teacher perspective, the following literature describes evaluation studies of high quality in-person and online teacher professional development.

Lindvall, Helenius, and Widerg (2018) argued that high-quality teacher professional development exhibits a level of coherence or alignment between a teacher's current knowledge and the proposed teacher change from the PD that is suitable for his or her current knowledge and skills. Abell et al. (2007) constructed profiles of professional development programs that aimed at building contextual factors in designing and implementing effective teacher professional development.

With the advent of educational technologies, teachers have multiple venues to engage in professional development opportunities online that allow them to address their immediate classroom challenges (O'Brian, 2016).

The following literature identifies research in online teacher professional development with an instruction that takes place 80% to 100% online based on the *Online Report Card* by Allen and Seaman (2016).

Yoo (2016) investigated the effects of online professional development on teacher self-efficacy using the Teacher's Self-Efficacy Scale (TSES) developed by Tschannen-Moran and Woolfolk Hoy (2001). Participants reported increased self-efficacy regarding (1) enhancement of teaching skills, (2) attainment of professional goals, and (3) acquisition of new pedagogical knowledge.

Frey (2009) investigated the effects of project-based online professional development experiences on the performance of K-12 students with disabilities and teachers' instructional practices. The qualitative analysis of discussion boards and journal entries suggested that (1) teachers improved their skills related to PD activities, (2) students with disabilities exhibited improved classroom performance, and (3) improved student performance affected teachers' readiness to implement evidence-based instructional practices.

Understanding Resource Utilization with Clustering of Digital Libraries, Online Courses, and MOOCs

The researcher's approach to analyzing large web analytics data is to use a clustering algorithm to derive utilization patterns of resources and site features. Several studies have used the simple k-means algorithm to understand patterns of educational digital libraries, online courses in higher education, and massive open online courses (MOOCs). Khoo et al. (2008) examined user behavior by analyzing the session length web metrics produced by various educational digital libraries. Bollen and Luce (2002) assessed the user characteristics and retrieval behaviors in a digital library by generating clusters to describe research interests and preferences from document retrieval networks. Valsamidis et al. (2012) utilized the simple k-means clustering method to generate patterns of online course quality. Rodrigues et al. (2016) performed a cluster analysis of student engagement patterns in a massive open online course (MOOC) and identified levels of student engagement regarding activity completion, number of comments, and answers in discussion forums.

While existing research identifies positive effects of online teacher PD and success factors in high-quality online PD environments, only handful studies have documented user behavior in large online teacher professional development platforms. The need to analyze online behavior provides insights on how teachers use self-paced instructional modules and interactive site features in online teacher professional development environments. While web metrics are widely used in tracking user behavior and generate large amounts of data about resource utilization, the representation of web metrics data through a cluster analysis allows the observation of patterns from site features and resources when grouped based on similar characteristics.

Statement of Purpose

The purpose of this study is to discover the utilization patterns of the content and site features of a K-12 online professional development platform. By data mining web metrics records from Google Analytics, this study allows a comprehensive examination of the patterns between new and returning visitors by applying the simple k-means algorithm and visualizing natural groupings or cluster outputs (Jain, 2010).

The following research questions guided the design of the study: (1) What dimensions and metrics in Google Analytics are potential predictors of clustering? (2) What behaviors do new and returning visitors exhibit when using the platform? (3) What are the content and site feature patterns present from the predictor metrics and dimensions? The following section describes data extraction, manipulation, computation, and deployment of results.

Methods

Research Setting

The research context is a large online content management platform with 50,425 K-12 teachers and school administrators seeking professional development in the state of Missouri in the United States. A total of 460,721 web metrics records from the platform are available from October 2015 to February 2017. GA collects metric and dimension properties when a tracking script snippet is assigned to online resources and platform features. The web metrics reports of 17 months are available for extraction from the free edition of Google Analytics.

Data Mining Process

This study adheres to the cross-industry standard process for data mining (CRISP-DM) methodology. The CRISP-DM is a popular data mining methodology that brings organization into data mining projects (Shearer, 2000). This approach requires six systematic steps in any data mining project: (1) research understanding, (2) data understanding, (3) data preparation, (4) modeling, (5) evaluation, and (6) deployment. The following subsections describe each phase.

Research understanding phase. The EdHub Library is built on Plone, an open-source content management system, which provides a mechanism for controlling user and content access. Each school year, access tokens are provided to teachers and principals to control access to the platform. At a system level, the EdHub Library is integrated with GA to track user and content activity. In this data mining project, the analysis aims at examining the patterns among users, instructional materials, and site features.

Data understanding phase. To make sense of GA records, the site and content taxonomy was required to understand the overall structure of the EdHub Library. The taxonomy is derived from the web hosting indexing structure. Additionally, the researcher is responsible for creating and deploying instructional materials in the web repository in a structured manner.

The Google Analytics Dimensions and Metrics Explorer provides descriptions of dimensions and metrics extracted from the platform shown in Table 1 ("Google Developers," n.d.).

Table 1. Google Analytics Dimension and Metric Descriptions

Dimension	*Description*
Page	A page on the website specified by path or query parameters
User Type	A boolean, either new visitor or returning visitor, indicating if the users are new or returning
Browser	Users' browsers, for example, Internet Explorer or Firefox
City	Users' city, derived from their IP addresses or geographical IDs
Page Depth	The number of pages visited by users during a session
Day of the Week	Day of the week, a one-digit number from 0 (Sunday) to 6 (Saturday)
Metric	
New Users	Users who have visited the site for the first time
Users	The total number of users during the requested time period
Sessions	A session is a group of interactions lasting at least 30 minutes
Bounce Rate	The exit rate of a single page
Session Duration	The total duration (in seconds) of users' sessions
Pageviews	The total number of pageviews of a page
Time on Page	The total time (in seconds) users spent on a particular page

Data preparation phase. After extracting 460,721 records in a tabular format, descriptive statistical measures were performed to identify missing cases. This preparation work makes it possible to create data visualizations in Tableau software to showcase (1) patterns between new and returning visitors, (2) the most searched terms, (3) the most accessed resources, and (4) the most accessed features. Tableau is a data visualization software that produces interactive data visualizations (Tableau, 2017).

Modeling phase. A feature selection was performed to understand the potential predictors of site utilization among the dimensions and metrics. The simple k-means algorithm was used to develop similar groupings or patterns for online resources and platform features among the predictor variables. The simple k-means algorithm was performed by running the dataset using 2, 10, 20, 30 and 40 cluster parameters by standardizing Euclidean distances with ten folds of cross-validation and 50 iterations. Euclidean distance is a measure of similarity and dissimilarity among data points.

Evaluation phase. To validate the selection of the cluster parameter, the Davies-Bouldin validity index was used to evaluate the internal validity of clusters. The Davies-Bouldin validity index measures the average similarity between clusters. A lower validity index translates into a better cluster configuration regarding separation and compactness (Kovács, Legány, & Babos, 2005). After manipulating several cluster parameters, a cluster parameter of 20 produced the lowest validity index.

Deployment phase. The deployment phase involves presenting the data mining efforts to understand how teachers and school administrators use site features and resources. The data mining results are presented in data visualizations in the form of bar charts for searched terms, resources, and features. Also, visualizations are presented in heatmaps to show cluster classifications organized by user type and page.

Results

Seven out of twenty cluster patterns describe utilization patterns of resource pages and features for new and returning visitors. The remaining thirteen clusters overlap with the previous patterns. It is important to note that patterns do not represent a sequence of actions that occur in the site. Instead, patterns show the number of cases in which actions occur the most frequent and paint the 'picture' of how many events unfold in the site by examining prominent classifications or groupings.

Pattern 1: new visitors - teacher or school administrator - user registration through the online teacher evaluation tool. This pattern shows new visitors, either teachers or school administrators, accessing the site through the online teacher evaluation tool to sign up for an account through the registration page on the EdHub homepage. This cluster shows classified cases with the following utilization pattern: (1) accessing EdHub through the online teacher evaluation tool, (2) signing up for an account using the login form, (3) reviewing content subscriptions, (4) verifying group access, (5) accessing content list of topics, (6) viewing group membership, (7) receiving a successful password reset, (8) reviewing teacher standards descriptions, (9) accessing terms and conditions on the registration page, (10) logging into the site, (11) getting an error with the wrong group access token, (12) reviewing content subscriptions, and (13) resetting a password.

Pattern 2: returning visitors - teacher or school administrator - registered users through the online teacher evaluation tool. This pattern shows returning visitors, either teachers or school administrators, accessing the site through the online teacher evaluation tool. This cluster shows classified cases with the following utilization pattern: (1) accessing EdHub through the online teacher evaluation tool, (2) reviewing content subscriptions, (3) verifying group access, (4) signing up for an account using the login form, (5) viewing the group listing page, (6) logging out of the site, (7) reviewing teacher standards descriptions, (8) viewing token content access, (9) reviewing group subscriptions, (10) getting an error with the wrong access group token, (11) accessing online resources related to school administrators, (12) receiving a successful password reset, and (13) accessing online resources related to speech pathologist.

Pattern 3: new visitors - school administrators - training, user subscription, and password reset. This pattern shows new visitors, school administrators, accessing a recertification training available to school administrators through the EdHub site directly. The recertification training allows school administrators to practice scoring teacher interactions in the classroom. To access the recertification training and copyrighted materials, users need specific content access the EdHub Library. This cluster shows classified cases with the following utilization pattern: (1) accessing recertification training, (2) accessing terms and conditions on the registration page, (3) resetting a password, (4) entering content token access, (5) reviewing teacher standards descriptions, (6) verifying group access, (7) reviewing content subscriptions, (8) accessing the library content list, (9) viewing members in a group, (10) entering the wrong password, (11) logging out of the site, (12) signing up for an account using the login form, and (13) viewing the group listing page.

Pattern 4: new visitors - teachers - journaling and library content related to cognitive and affective engagement. This pattern shows new visitors, teachers, accessing journal entries and reviewing library content related to assessment topics, and teacher indicators related to cognitive engagement (1.2) and motivation strategies (5.1). This cluster shows classified cases with the following utilization pattern: (1) viewing journal entries, (2) accessing the library content list, (3) viewing the journal landing page, (4) reviewing teacher standards descriptions,

(5) reviewing content subscriptions, (6) reviewing NEE Teacher Standard 1 or content knowledge, (7) viewing the group listing page, (8) reviewing NEE Indicator 1.2 or cognitive engagement, (9) accessing assessment topics, (10) verifying group access, (11) reviewing NEE Indicator 5.1 or affective engagement, (12) signing up for an account using the login form, and (13) reviewing the about Edhub page.

Pattern 5: returning visitors - school administrators - administrative viewing of journal entries. This pattern shows returning visitors, school administrators, accessing the site directly to view teachers' journal entries and review library content related to instructional leadership topics and classroom observation videos. This cluster shows classified cases with the following utilization pattern: (1) viewing journal entries, (2) reviewing teacher standards descriptions, (3) reviewing user group preferences, (4) accessing assessment topics, (5) accessing online resources related to school administrators, (6) reviewing content subscriptions, (7) verifying group access, (8) viewing the group listing page, (9) accessing online resources related to instructional leadership topics, (10) reviewing teacher standard NEE Teacher Standard 1 or content knowledge, (11) accessing online resources related to administrator classroom observation training videos, (12) accessing formative assessment topics, and (13) accessing online resources related to speech pathologist.

Pattern 6: returning visitors - school administrators - administrative group and user preferences. This pattern shows returning visitors, school administrators, accessing the group and user settings pages through the online teacher evaluation tool. These returning visitors have access to administrative settings to control group settings and member access. However, teachers do not have access to administrative settings. This cluster shows classified cases with the following utilization pattern: (1) viewing group members, (2) reviewing user preferences, (3) accessing group preferences through the online teacher evaluation tool, (4) logging out of the site, (5) accessing content knowledge topics, (6) accessing assessment topics, (7) accessing school administrators topics, (8) accessing professional development topics, (9) accessing formative assessment connection topics, (10) managing user subscriptions, (11) accessing administrator classroom observation training videos, (12) reviewing content subscriptions, and (13) reviewing NEE Indicator 1.2 or cognitive engagement.

Pattern 7: returning visitors - teachers - group and user preferences. This pattern shows returning visitors, teachers, accessing the teacher group and individual user settings through the online teacher evaluation tool. In this particular pattern, site pages related to administrative settings are missing since teachers who are members of a group do not have access to group controls. Only school administrators have access to administrative settings. This cluster shows classified cases with the following utilization pattern: (1) reviewing user preferences, (2) accessing group preferences, (3) reviewing teacher standards descriptions, (4) reviewing content subscriptions, (5) viewing the group listing page, (6) accessing beginning teacher assistance topics, (7) accessing assessment topics for beginning teachers, (8) accessing assessment topics, (9) verifying group access, (10) accessing platform tutorials, (11) accessing assessment connection topics, (12) accessing modules in assessment connection topics, and (13) accessing professional development topics.

Discussion

In this study, 460,721 web analytics records from October 2015 to February 2017 were clustered using the k-means algorithm to develop content and site feature utilization patterns of an online K-12 teacher professional development platform. In this study, the free version of Google Analytics was used to extract thirteen variables (*page, user type, browser, city, page depth, day of the week, new users, users, sessions, bounce rate, session duration, pageviews,* and *time on page*).

A feature selection was performed to discover the predictors of the *page* variable, which tracks the content and features of the platform. Based on the feature selection, *city, browser,* and *users* were not significant predictors of the *page* variable. Nine predictor variables for the *page* variable were found: (1) *user type,* (2) *page depth,* (3) *day of the week,* (4) *new users, (5) sessions,* (6) *bounce rate,* (7) *session duration,* (8) *pageviews,* and (9) *time on page.* By selecting the lowest Davies-Bouldin internal validity index, a cluster parameter of 20 was performed on the dataset to form utilization patterns by user type and page The site taxonomy of the platform was obtained from the web server to interpret of clusters.

User Search Behavior

While prominent clusters do not show specific search behavior of users, exploring what users typed in search queries allows further understanding of teachers' needs for professional development. Users searched for examples of units of instruction, examples of professional development plans, and classroom training videos.

Users' search queries highlight the need for teachers to find practical and concrete implementation examples of teaching strategies that model the skills in their classroom context. Research shows that teachers face challenges in implementing new teaching strategies rather than merely learning the theory behind teaching strategies. Ermeling (2010) argued that teachers look for instructional solutions that lead to detectable improvements in student outcomes. Joyce and Showers (2002) stated that teachers are likely to explore training that models a new skill that is close to their classroom setting. Studies by Peery (2002), Blank and De Las Alas (2009), and Darling-Hammond et al. (2009) explored the positive outcomes between teacher practice and student learning when teachers focus on training with similar teaching skills and practices.

Deriving Use Cases from New and Returning Visitor Cluster Patterns

Seven patterns describe use cases for teachers and school administrators when seeking PD and performing administrative functions. Patterns 1 and 2 are related to new and returning visitors, either teachers or school administrators, obtaining a registration to EdHub through the online teacher evaluation tool, reviewing their group and content subscriptions, and resetting passwords.

Pattern 3 describes new visitors, school administrators, accessing the EdHub site directly for training, checking their group subscription access, and performing password resets. In this particular pattern, these new visitors are school administrators going through a recertification training that requires administrative access to EdHub to verify teacher members in a school group.

Pattern 4 describes new visitors, teachers, accessing the site directly to perform journal tasks, reviewing descriptions of teacher NEE Indicators 1.2 (cognitive engagement) and 5.1 (affective engagement), and accessing topics related to assessment or NEE Standard 7. The absence of administrative tools in this pattern indicates that these new visitors are teachers.

Pattern 5 involves returning visitors, school administrators, accessing the site directly to view teachers' journal entries, reviewing group and personal preferences, and viewing topics related to assessment, instructional leadership, speech pathology, and classroom observation training videos.

Patterns 4 and 5 explain how teachers and school administrators behave on the site by discovering what is available in the library through the NEE Standards descriptions page or the library topic listing before performing journaling activities. Users can review descriptions of standards and their related topics from the NEE Standards page. Alternatively, users can review descriptions of PD topics through the library homepage. Regardless of how users access topics, users are presented with the alignment of a topic and standard to aid in the selection of journal activities.

Pattern 6 involves returning visitors, school administrators, reviewing their group membership and personal preferences. In this particular pattern, school administrator access EdHub via the online teacher evaluation tool to review administrator PD materials, which is only accessible to school administrators, and manage administrative settings of a group. Regarding professional development pieces, this pattern contains topics related to assessment and professional development.

Pattern 7 describes returning visitors, teachers, reviewing user preferences, group membership, and content subscriptions. Similar to pattern 6, teachers access EdHub via the online teacher evaluation tool to review teacher standard descriptions, getting started guides of the platform, and topics related to assessment, professional development, and beginning teacher. In this pattern, however, administrative pages are not present in the cluster.

Examining Small Cluster Patterns

Although prominent classifications are useful in understanding which site features and content occurred most frequently for the majority of users, small cluster patterns provide insights about the most utilized instructional resources and the issue of referrer spam.

Small classifications for sessions. In the cluster output for *sessions*, small classifications with a high number of sessions appear among returning visitors in topics related to instructional strategies, examples of units of instruction, and critical thinking, whereas new visitors access topics about the process of evaluating units of instruction.

Small classifications for page depth. In the cluster output for *page depth*, small classifications with a high page depth appear among returning visitors who access group preferences, whereas new visitors access the journal tool.

Small classifications for pageviews. In the cluster output for *pageviews*, small classifications with a high number of pageviews are related to new visitors accessing video examples of indicators 1.2 (cognitive engagement)

and 7.4 (effects of instruction). Returning visitors have a high number of pageviews when accessing their user preferences page, and video examples of teacher indicator 4.1 (problem-solving and critical thinking) and 1.2 (cognitive engagement).

Small classifications for session duration. In the cluster output for *session duration*, small classifications with a high session duration can be observed in topics related to academic vocabulary for new visitors, and administrative group settings and recertification training for returning visitors.

Small classifications for time on page. In the cluster output for *time on page*, small classifications with a high time on page suggest that new visitors spend more time on resources related to speech-language pathology, recertification training, and formative assessment, whereas returning visitors spend more time on topics related to instructional strategies and examples of units of instruction.

Small classifications for referrer spam. Referrer spam is a troublesome issue in Google Analytics tracking that makes repeated website requests using fake links to the site being spammed for advertising purposes even though these spam websites never visit the target site. In Google Analytics, website administrators can turn on the bot filtering option to block web traffic from referrer spam. Even when the bot filtering option is checked in Google Analytics for this study, less than 3 percent of classifications show referrer spam websites for *bounce rate*, *new users* and *session* variables. The clustering algorithm was able to detect patterns of three referrer spam websites (traffic2cash, snip.to, and adf.ly) that sent advertising links to Google Analytics. Spammers aim to create backlinks for users to access spam sites inadvertently.

While deleting spam records from web analytics records is ideal, two challenges are present in dealing with referrer spam. First, identifying referrer spam among thousands of records is difficult because these spam websites generate new subdomains or links to disguise their domains as new visitors. Second, cleaning spam website records is time-consuming and inaccurate because not all referral sites generate new traffic for spam purposes. For example, search engines and other sources, such as school districts networks and other online PD platforms, may redirect their searches to the EdHub domain.

Even if the bot filtering option is enabled in Google Analytics, it is not guaranteed to eliminate all spam website traffic. It is important to note that referrer spam websites were not able to access and alter web analytic records because the platform requires login credentials to access site features and content. Since the clustering algorithm was able to detect very small classifications of spam websites while the bot filtering option was checked, proactive measures can be taken by creating specific filters to prevent these spam website domains from sending backlinks in the future.

Comparing Tasks from New and Returning Visitor Cluster Patterns

By comparing the seven prominent cluster patterns, it allows further exploration of differences in tasks and site access performed between teachers and school administrators. As shown in Table 2, in four patterns out of seven, teachers and school administrators access the EdHub site through the online teacher evaluation tool. In the three remaining patterns, teachers and school administrators access the EdHub site directly.

When accessing the EdHub site through the online teacher evaluation tool, new visitors who are teachers (1) register for an account and (2) verify group and content membership, whereas new visitors who are school administrators (1) verify group members, (2) check content membership, and (3) access training materials.

When accessing the EdHub site directly, new visitors who are teachers (1) access the journal tool, (2) review teacher standards descriptions, and (3) search for content related to NEE Indicators 1.2 (cognitive engagement) and 5.1 (affective engagement), whereas new visitors who are school administrators (1) verify group members and (2) check content membership.

A fundamental difference exists in the tasks for new teachers when accessing the EdHub site through a third party and directly. When accessing the site directly, new teachers tend to engage more with the site by accessing the journal feature, searching teacher standard descriptions, and seeking online PD based on teacher standards. When the site is accessed through the online teacher evaluation tool, new teachers perform the necessary tasks of registration, and checking their group and content memberships. For school administrators, either new or returning visitor, they perform administrative tasks regardless of access.

Returning visitors who are school administrators accessing the EdHub site, either through the online teacher tool or directly, show more engagement with reading teachers' journal entries, checking group preferences, and accessing various online PD topics. Returning visitors who are teachers, accessing the EdHub site through the online teacher tool, review assessment and beginning teacher topics. Unfortunately, returning teachers do not engage with the journal tool in the future.

Table 2. Comparison among cluster patterns by user type

Pattern	User Type	User Role	Point of Access	Prominent User Tasks	Most Utilized Instructional Resources
1	New	Teacher or School administrator	Teacher evaluation tool	1. Register for an account through the online teacher evaluation tool 2. Reset password 3. Verify group and content membership	
2	Returning	Teacher or School administrator	Teacher evaluation tool	1. Register for an account through the online teacher evaluation tool 2. Reset password 3. Verify group and content membership 4. Access teacher standard descriptions	1. Speech-language pathologist
3	New	School administrator	EdHub homepage	1. Register for an account directly on the site 2. Reset password 3. Verify group and content membership 4. Verify group members	1. Recertification training
4	New	Teacher	EdHub homepage	1. View journal tool	1. Teacher standards descriptions 2. Teacher indicator description for Indicator 1.2 (cognitive engagement) 3. Teacher indicator description for Indicator 5.1 (affective engagement) 4. Assessment topics
5	Returning	School administrator	EdHub homepage	1. View journal entries 2. Access group preferences	1. Teacher standard description for Standard 1 (content knowledge) 2. Classroom observation training videos 3. Assessment topics 4. Speech pathologist
6	Returning	School administrator	Teacher evaluation tool	1. Verify group members 2. Access group preferences 3. Access personal preferences	1. Administrator library 2. Assessment topics 3. Professional development topics 4. Classroom observation videos

| 7 | Returning | Teacher | Teacher evaluation tool | 1. Verify group and content membership
2. Access personal preferences | 1. Teacher standards descriptions
2. Beginning teacher topics |

Several online professional development topics are present across the patterns pertaining to (1) cognitive engagement, (2) formative assessment, (3) beginning teacher, (4) affective engagement, (5) administrator library, (6) speech-language pathologist, (7) getting started with EdHub, (8) classroom observation video examples, and (9) examples of units of instruction. However, the following topics are not present: (1) communication, (2) data analysis, (3) Common Core State Standards and Next Generation Science Standards, and (4) family and community involvement. All topics are available to teachers and administrators, except for the administrator library. Descriptions of these topics present in the patterns are shown in Table 3.

Table 3. Topic Descriptions

Topic	Description
Cognitive engagement	Strategies for engaging students with content and developing students' deep thinking
Formative assessment	Key concepts of ongoing assessment and processes of setting clear student targets
Beginning teacher	Curated instructional modules covering assessment, classroom environment, cognitive/affective engagement, professional communication, and ethical behavior
Affective engagement	Motivation strategies that teachers use to influence students' motivation
Administrator library	Training videos and copyrighted materials related to scoring teacher-student interactions in the classroom
Speech-language pathologist	Roles and responsibilities, professional knowledge, and referral documentation and process.
Getting started with EdHub	Tutorials related to registration, subscriptions, and journaling
Classroom observation video examples	A collection of exemplary videos showcasing teaching strategies by standard and grade level
Examples of units of instruction	A collection of units of instruction examples shared by Missouri teachers
Communication	Strategies for effective communication from teachers to students, parents, and administrators
Data analysis	Strategies for analyzing and building grade or building level data for instruction
Common Core State Standards and Next Generation Science Standards	Strategies for integrating Common Core State Standards and Next Generation Science Standards in the classroom
Family and community involvement	Strategies for fostering adult involvement into the classroom

Conclusion

The results of this study can be used to understand user behaviors of teachers and school administrators, prioritize the development of professional development topics, and improve the navigation of the EdHub Library. Users will benefit from having quick access to the most utilized professional development materials when resources are placed on the homepage. As the Network of Educator Effectiveness grows, stakeholders will need to address the increased access to online professional development through the teacher evaluation tool and develop online professional development materials that target specific teaching skills with several classroom contexts. Teachers and school administrators perform the essential tasks based on their role. Teachers obtain a 'birds-eye' view of what is available in the EdHub Library before completing journal activities. Even though cluster patterns do not show returning teachers accessing the journal tool for future activities, cluster patterns show returning school administrators reading teachers' journal entries and performing administrative functions.

References

Abell, S. K., Lannin, J. K., Marra, R. M., Ehlert, M. W., Cole, J. S., Lee, M. H., ... & Wang, C. Y. (2007). Multi-site evaluation of science and mathematics teacher professional development programs: The project profile approach. *Studies in Educational Evaluation*, *33*(2), 135-158.

Allen, I. E., & Seaman, J. (2016). Online Report Card: Tracking Online Education in the United States. *Babson Survey Research Group*.

Bechtel, P. A., & O'Sullivan, M. (2006). Chapter 2: Effective professional development—What we now know. *Journal of teaching in physical education*, *25*(4), 363-378.

Blank, R. K., & De Las Alas, N. (2009). *The Effects of Teacher Professional Development on Gains in Student Achievement: How Meta-Analysis Provides Scientific Evidence Useful to Education Leaders*. Council of Chief State School Officers. One Massachusetts Avenue NW Suite 700, Washington, DC 20001.

Bollen, J., & Luce, R. (2002). Evaluation of digital library impact and user communities by analysis of usage patterns. *D-Lib Magazine, 8*(6), 1-13.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher*, *33*(8), 3-15.

Carter, K. (2004). Online Training: What's Really Working? What Does Today's Successful Online Professional Development Look Like?. *Technology & Learning*, *24*(10), 32.

Coldwell, M. (2017). Exploring the influence of professional development on teacher careers: A path model approach. *Teaching and teacher education*, *61*, 189-198.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). Professional learning in the learning profession: A status report on teacher development in the United States and abroad. Washington, DC: National Staff Development Council.

Desimone, L. M. (2011). A primer on effective professional development. *Phi delta kappan*, *92*(6), 68-71.

EdHub Library. (n.d.). Retrieved from https://education.missouri.edu/outreach/edhub/

Ermeling, B. A. (2010). Tracing the effects of teacher inquiry on classroom practice. *Teaching And Teacher Education, 26377-388*. doi:10.1016/j.tate.2009.02.019

Frey, T. J. (2009). An analysis of online professional development and outcomes for students with disabilities. *Teacher Education and Special Education*, *32*(1), 83-96.

Google Developers. (n.d.). Dimensions & metrics explorer. Analytics Reporting API v4. Retrieved from https://developers.google.com/analytics/devguides/reporting/core/dimsmets

Guskey, T. R. (2002). Does it make a difference? Evaluating professional development. *Educational leadership*, *59*(6), 45.

Hanson, A., Pennington, T. R., Prusak, K., & Wilkinson, C. (2017). PE Central: A Possible Online Professional Development Tool. *Physical Educator*, *74*(3), 570.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, 31*(8), 651-666.

Joyce, B. R., & Showers, B. (2002). Student achievement through staff development.

Khoo, M., Pagano, J., Washington, A., Recker, M., Palmer, B., & Donahue, R. (2008). Using web metrics to analyze digital libraries. In *Proceedings of the joint conference on digital libraries* (pp. 375-384). New York, NY: ACM.

Kovács, F., Legány, C., & Babos, A. (2005, November). Cluster validity measurement techniques. In *6th International symposium of hungarian researchers on computational intelligence*.

Lindvall, J., Helenius, O., & Wiberg, M. (2018). Critical features of professional development programs: Comparing content focus and impact of two large-scale programs. *Teaching and Teacher Education*, *70*, 121-131.

MyTeachingPartner. (2017). Retrieved from https://curry.virginia.edu/myteachingpartner

Network for Educator Effectiveness. (n.d.). Retrieved from http://nee.missouri.edu/

O'Brian, M. M. (2016). Big data in the service of educator learning: what should be done with collected online professional learning information?. *Quarterly Review of Distance Education*, *17*(4), 39.

Peery, A. (2002). Beyond inservice. *Principal Leadership*, *3*(3), 22-28.

Rice, K., & Dawley, L. (2009). The status of professional development for K-12 online teachers: Insights and implications. *Journal of Technology and Teacher Education*, *17*(4), 523-545.

Rodrigues, R. L., Ramos, J. L. C., Silva, J. C. S., Gomes, A. S., Souza, F. F., & Maciel, A. M. A. (2016). Discovering level of participation in MOOCs through clusters analysis. In *IEEE 16th international conference on advanced learning technologies (ICALT)* (pp. 232-233). Austin, TX: IEEE.

Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing, 5*(4), 13-22.

Tableau Software. (2017). Technical Specifications Tableau Desktop. Retrieved from http://www.tableausoftware.com/products/desktop/specs

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and teacher education*, *17*(7), 783-805.

Valsamidis, S., Kontogiannis, S., Kazanidis, I., Theodosiou, T., & Karakos, A. (2012). A clustering methodology of web log data for learning management systems. *Journal of Educational Technology & Society, 15*(2), 154.

Yoo, J. H. (2016). The effect of professional development on teacher efficacy and teachers' self-analysis of their efficacy change. *Journal of Teacher Education for Sustainability*, *18*(1), 84-94.