Journal of Online Learning Research (2017) **3**(2), 197-221

Usage Data as Indicators of OER Utility

MARCIA A. MARDIS
Florida State University, United States
mmardis@fsu.edu

CHANDRAHASA R. AMBAVARAPU
Florida State University, United States
ca13@my.fsu.edu

A key component of online and blended learning content, open educational resources, (OER) are heralded in a global movement toward high-quality, affordable, accessible, and personalized education. However, stakeholders have expressed concern about scaling OER use due to a lack of means to ensure a fit between learner, resource, and task. Usage data, or "paradata," such as reviews, ratings, views, downloads, favorites, and shares, may yield insight into the fit. We examined paradata from National Science Digital Library (NSDL), the largest extant accessible corpus, for the extent to which K-12 science, technology, engineering, and mathematics (STEM) resource fit can be determined from user- and system-generated data. We conducted sentiment analyses of user reviews and correlations between the sentiment scores and data elements. Some relationships between NSDL paradata elements suggested aspects of resource fit. Despite prior research indicating that user reviews tended to be strongly positive or strongly negative, the results of this study indicated that educators left feedback that contained a blend of sentiments and that users usually downloaded resources they viewed. The results of this study suggest that while it is unlikely that educator feedback can currently be used to assess resource quality, with larger and more robust usage data sets, this area is a fertile area for further research into nuanced sentiment. We conclude with observed data trends and further research directions to inform online learning.

USAGE DATA AS INDICATORS OF OER UTILITY

Horn and Stake (2014) noted that “one of the most significant ways that online learning has improved is by leaning more heavily on in-person, bricks-and-mortar experiences to provide support and scaffolding for students learning online” (p.32). This shift, driven largely by the need for instructors to personalize learning experiences more precisely to learners’ needs, has been accompanied by an increased need for a wealth of resources that address varied interests, abilities, and learning paths (U.S. Department of Education, 2017b). Efforts to promote learning personalization through blended and online learning, common standards, competency-based education, and K-12 digital textbooks are rapidly increasing open educational resource (OER) demand (Ash, 2012; Mickey & Meaney, 2013). The vision for these efforts is for educators to curate and create high-quality openly licensed educational resources in order to provide students with more personalized learning opportunities than learning environments based in traditional print materials can support (U.S. Department of Education, 2017a).

Personalization depends on an adequate supply of high-quality digital learning resource—there must be enough resources to meet a range of learning tasks and needs and the resources must be effective and accurate. Policymakers and educators are concerned about an adequate supply of affordable, readily available, high quality learning resources (Lagoze, 2010; McMartin et al., 2008; Tonks, Patrick, & Bliss, 2013), particularly in the areas of science, technology, mathematics, and engineering (STEM) learning (Porcello & Hsi, 2013). To determine resource quality and appropriateness, educators may need to rely on more than resource descriptions (i.e., metadata) to judge suitability. Descriptive metadata have been a primary means to support OER selection and implementation (Abramovich & Schunn, 2012). While descriptive metadata may help to determine what a resource is about, educators also benefit from information about how other educators have used the resource and whether they found it to be to be effective (U.S. Department of Education, 2012, 2013).

Metadata are data about the data; paradata are interaction data, that is, the data around the data (MacNeill, Campbell, & Hawksey, 2014). Learning analytics practitioners have the potential to use these data generated by the educational resource repository users to recognize patterns in digital resource use and impact (Abramovich, Schunn, & Correnti, 2013). Because paradata include system generated and social contributions such as ratings, reviews, comments, favorites, shares, downloads, and other activity data relating to resource use, their analysis can give insight into educators’ and learners’ perceptions and preferences (Campbell & Barker, 2013; U.S. Department of Education, 2012).

Complementary work has been done to assess OER quality from metadata based on expert reviews (e.g., Bethard, Wetzler, Butcher, Martin, & Sumner, 2009); however, little work has been done to determine ways in which user-contributed, and user-generated paradata might reflect perceptions of resource quality and use.

A popular way of deriving meaning from free text user-provided feedback is opinion mining via sentiment analysis. Sentiment analysis has grown out the need for providers like shopping site Amazon.com (Liu, Yu, An, & Huang, 2013), travel planning site Trip Advisor (Lak & Turetken, 2014), and restaurant review site Yelp (Ganu, Kakodkar, & Marian, 2013; Seaman & Allen, 2014) to be able to be able to recommend products and services to their users. This emerging area of text mining and computational linguistics may also provide a useful, but underexplored, approach to examining learning resource usage data.

Purpose and Research Questions

The purpose of this study is to explore the relationship between paradata elements and user perceptions of resource quality. To meet this purpose, we addressed the following research questions (RQ):

RQ1. To what extent are users' assessments of resource quality positive, negative, or neutral?

RQ2. What are the relationships between user sentiments and other learning resource usage data?

In this study, we first examined user reviews OER paradata with sentiment analysis to explore how useful affective feedback may be for developing and curating collections of quality assured web-based resources useful for teaching and learning. Then, we attempted to determine the possible relationships between sentiment and other user activities like sharing and favoriting.

Whether encountered in a fully online or blended learning environment, OER can generate data that can allow for tailoring of resources to learners' needs and determining resources' usefulness, thus closing the loop between what is taught and what is learned (Essa, 2016). Once a problem of interest only to researchers interested in learning analytics, the need to understand which resources will help learners grasp concepts most effectively and equalize learning opportunities is a national imperative and an issue of social justice (U.S. Department of Education, 2017a).

LITERATURE FOUNDATION

Teaching and learning with digital resources can engage, challenge, and motivate students with compelling experiences that encourage independent and collaborative learning. A personalized mix of aesthetic, technical, and educational resource designs enhances learning engagement and motivation.

When learners have the ability to use and create digital resources to extend their repertoires, they can realize learning outcomes; support their agency; develop metacognitive skills and higher order thinking skills; and participate in reflection and collaboration (Chen & Sun, 2012; Project Tomorrow, 2012a). While effective teaching is an important element in fostering learner outcomes, in the immediate future, so is the availability of high-quality, engaging, personalized learning content (Hanover Research Council, 2011; Maull, Saldivar, & Sumner, 2010; New Media Consortium [NMC], 2014).

Learning personalization results from data-informed decisions made at key points in the instructional process: determination of teacher and student characteristics; guidance of a coherent, rigorous curriculum; execution of appropriate student-centered teaching activities; and application of meaningful formative assessments (U.S. Department of Education, 2012). Whether in an elementary or secondary online course, blended learning environment, or instructional improvement system, this process rests on a foundation of high quality, plentiful, multimodal OER (Collins & Levy, 2013; Hewlett Foundation).

Open Educational Resources

The Hewlett Foundation has defined OER as:

> [T]eaching, learning, and research resources that reside in the public domain or have been released under an intellectual property license that permits their free use and re-purposing by others. [OER] include full courses, course materials, modules, textbooks, streaming videos, tests, software, and any other tools, materials, or techniques used to support access to knowledge. (para. 3)

This definition indicates that OER may be used in a variety of learning settings, including those that are entirely face-to-face, those that are entirely online, or those that represent a combination or blend of the two.

OER originated in and are often created for K-12 STEM learning (Hanson & Carlson, 2005; Mardis, 2003; Mardis & Howe, 2010; Mardis & Zia, 2003). While the benefits of using OER versus costly commercial instructional materials are evident, problems of determining OER quality have persisted for over two decades (Hewlett Foundation, 2013; Leutkemeyer & Mardis, 2016; Okerson, 2000; Project Tomorrow & Blackboard, 2016). In an unpublished report of a 2013 meeting of statewide education officials hosted by the National Science Digital Library (http://nsdl.org), participants stated that the operationalization on their digital learning initiatives relied on an adequate supply of high-quality K-12 STEM OER. Participants cited threats such as a lack of curation strategies to manage collections; insufficient resources to identify and vet OER; and a strong desire to automate the selection, validation, and management processes as essential issues to resolve if their delivery of curricula based on the Common Core State Standards (CCSS) and Next Generation Science Standards (NGSS) were to be successful (National Science Digital Library [NSDL], 2013).

Many researchers have noted that descriptive metadata standards are neither necessary nor sufficient to address issues to expressing learning resource quality to users (Bethard et al., 2009; Porcello & Hsi, 2013; Wetzler et al., 2013) because aspects of OER quality often include resource curation issues that are not captured in descriptive metadata schema. OER curation is discussed in the following section.

OER Curation

Resource selection, management, and promotion comprise the curatorial enterprise (Rosenbaum, 2011). Researchers have demonstrated that teachers have tended to prefer learning materials that have been reviewed by experts and used with positive effect by educators in nearby and in similar circumstances (Abramovich & Schunn, 2012; Williams & Coles, 2003). In recent years there has been a growing awareness that usage data, such as ratings and user reviews, are needed to properly curate the content in repositories and determine resource suitability for teaching and learning tasks. As Griffin (2013) pointed out, "Subjective reviews submitted from highly qualified educators as well as independent reviewers are valuable elements that could be, and should be, included as unique data..." (para. 3). Features that capture other aspects of the user experience include incidences of favoriting, sharing, viewing, and downloading (Campbell & Barker, 2013).

Traditional instructional materials are primarily reviewed and selected for content accuracy, particularly in the STEM fields (Spiegel, 1989; Stern & Roseman, 2004). As materials have become more digital and curriculum standards more influential, media quality and standards alignment issues must also be routinely considered (Hanson & Carlson, 2005; Mardis, ElBasri, Norton, & Newsum, 2012). However, in an autonomous search mode, when teachers assess quality as they collect resources for immediate use, sentiment and quality determination tend to vary widely (Perrault, 2007; Recker et al., 2011; Recker et al., 2007). Some teachers favorite assessments based on the resource's appearance (e.g., colors, font, web page format) and some prefer assessments based on content (e.g., number and currency of citations, content provider affiliation) (Price, 2007). For this reason, many teachers and education policymakers have expressed preferences for repositories of vetted resources with clear designations of quality, utility, and curriculum support (Griffin, 2013; Project Tomorrow & Blackboard, 2016; Sumner, Khoo, Recker, & Marlino, 2003).

However, encouraging teachers to contribute plentiful, complete, and accurate reviews of resources has proven to be difficult, despite the fact that many teachers have responded favorably to complete reviews left by other educators (Abramovich et al., 2013). Incentives are rarely effective because many teachers view feedback about their resource preferences as personal criticism. Also, teachers tend to be altruistically motivated and enthusiastic about sharing feedback when they feel the resource is valuable and would help other teachers (Van Acker, van Buuren, Kreijns, & Vermeulen, 2013). To gain insight into teachers' feedback behaviors, we looked to studies of contributions from other domains, including online news readers, shoppers, and service consumers.

Usage Data in Other Domains

As the number of websites that provide informational and commercial resources has grown, so has the availability of tools that capture user feedback in usage data. Like educational usage data, these data exist in two forms: 1) subjective usage data such as user reviews, comments, annotations, and recommendations; and 2) objective usage data that are system generated like numbers of ratings, views, downloads, and shares. In this section, we explore not only users' motivations to provide feedback, but also the types of feedback consumer-users tend to contribute.

Motivation to Contribute Ratings and Reviews

In a longitudinal study, Tenenboim and Cohen (2015) examined the association between users' clicking and commenting activities for online news items, as well as differences and similarities between items with a high number of views and items with a high number of comments. Significant differences were found between the two groups of items, especially regarding the news topics and elements that may arouse controversy or curiosity. Recently, researchers (Ziegele, Breiner, & Quiring, 2014) used mixed methods to study news story user commenting behavior, the results of which implied that commenters could trigger further response by including controversial or unexpected statements, personalizing comments, and expressing conflicting opinions in their postings and by avoiding incomprehensibility and negativity. Length, position, the news medium itself, and the news story topic further affected the probability of whether a comment received feedback. In another longitudinal study, Springer, Pfaffinger, and Engelmann (2015) investigated users' motives and inhibitors for commenting on news websites. When comments were visible and accessible beneath their respective articles, users reported deep engagement with other users' viewpoints, and suggestions for related stories.

Consumer Ratings and Reviews

Many content providers allow users to contribute their feedback about resources. User reviews, comments, and annotations can provide valuable information that other users may find helpful, but they must be systematically analyzed and integrated into a content provider's recommender system to have maximum effect. Mumdabi and Schiff (2010) analyzed Amazon.com reviews to determine whether sentiment extremity, review depth, and product type affected the perceived helpfulness of the review. They found that reviews with extreme ratings were perceived as less helpful than reviews with moderate ratings. Review depth had a positive statistical correlation with the perceived helpfulness of the review. The researchers also reported that Amazon.com reviews which contained positive and negative comments and moderate ratings (for example, three out of five stars) were perceived as most helpful. These findings were echoed by subsequent researchers working with ratings and reviews in the consumer and leisure domains (e.g., Kronrod & Danziger, 2013; Ludwig et al., 2013).

Recent research has also revealed possible problems with written user reviews. For example, while consumer research firm YouGov (2014) reported that most Americans used consumer reviews and that almost half who used them also contributed reviews, users were not likely to leave negative reviews unless they had a very bad experience with the product. Instead, most Americans opted not to leave negative reviews to avoid being critical. YouGov also reported that nearly a fifth of its adult survey respondents contributed false negative reviews for malicious or spiteful reasons.

Because the study of usage data in non-education domains is mature, this prior work may offer insight into phenomena that may be uncovered through the study of educational usage data. While there may be ways to increase the extent to which educators contribute reviews and ratings of learning resources, educational paradata may also be affected by issues observed in other domains: the difference between expert and community concepts and expressions of quality, difficulty in controlling review quality, and an absence of means for quality to be automatically characterized to enhance personalization (Massart & Shulman, 2013). Without further study, these issues will likely also remain barriers to scaling OER use.

METHOD

In this study, we analyzed usage data from resources in the National Science Foundation's (NSF) National Science Digital Library (NSDL), a leading provider of K-12 STEM-focused OER. NSDL housed OER metadata from nearly 400 NSF-funded content providers. NSDL was at the forefront of envisioning OER metadata integrated with usage data, and coined the term paradata in its proposed research agenda in 2010 (VanGundy, 2010a).

Description of the Data

NSDL began providing usage data for resources in its collections in January 2011 in the form of annotations about the resources and related information on usage (VanGundy, 2010b). NSF discontinued NSDL funding, and new data sets were no longer available after December 2013. To date, no other K-12 OER paradata sets are publicly available. In this study, we used NSDL paradata from January 2011-December 2013. NSDL paradata consists of usage data provided by the OER content providers listed in Table 1.

NSDL has also contributed two distinct usage data schema (comm_anno and comm_para) (NSDL, 2014a, 2014b; 2014c), which are considered among preferred paradata formats by many OER providers (Bienkowski, Brecht, & Klo, 2012; Campbell & Barker, 2013; Niemann, Scheffel, & Wolpers, 2012; U.S. Department of Education, 2014). The "comm_anno" XML sets contain review text, and "comm_para" XML sets contain non-textual usage data such as star ratings, favorites, and downloads. Table 1 lists the collections that provided these types of paradata to NSDL.

Table 1
NSDL Paradata Providers, Domains, and Learner Levels

Data Provider	STEM Domain	Level
CLEAN	Climate science education	All
ComPADRE	Physics and astronomy education	All
CTE Online	Career and technical education,	K-12
DLESE	Earth systems education	All
iCPALMS	All STEM education	K-12
Instructional Architect	All STEM education	K-12
PBS Learning Media	All STEM education	All
PRISMS	All STEM education	All
SMILE Pathway	Science and mathematics informal education	All
TeachEngineering	Engineering education	K-12
TeachSpatial	Spatial cognition, learning, and literacy education	All

The researchers harvested the XML usage data files directly from NSDL servers and then combined the comm_anno and comm_para files to integrate all available usage data for each resource. In sum, we extracted a total of 2,505 comm_anno and comm_para resource records from NSDL. Table 2 provides an overview of the usage data included in the integrated file.

Table 2
Types of Usage Data Present in NSDL Paradata Records (N=2505)

Usage data Type	Occurrence
Annotations (subjective)	752
Downloads (objective)	17
Favorites (subjective)	499
Features (subjective)	101
Ratings (subjective)	519
Recommended (subjective)	113
Views (objective)	504
Total	2505

As Table 2 shows, the sample included 752 comments. Records also included 1232 other occurrences of user-provided (subjective) data in the form of favorites, features, ratings, and recommendations as well as 521 system-generated (objective) contributions in the form of downloads and views, for a total of 2505 discrete usage data elements.

We used the Simple API for XML (SAX) Java API1 to parse the XML documents. The parser extracted data for usage data specific fields (such as user comments, views, and downloads in our case) and excluded content from other fields. The parsed usage data were redirected to an Excel spreadsheet using the Poor Obfuscation Implementation (POI) API.

Procedure

We performed our analysis in two phases to establish sentiment polarity (positive or negative) and its relationship to subjective rating (e.g., "two and a half stars") and system-generated usage data.

Phase I. Sentiment Analysis

We used sentiment analysis to determine whether user reviews and comments were positive, negative, and neutral and assigned a numerical weight according to each determination. As the base for our sentiment analysis, we used commonly accepted techniques (Padmaja & Fatima, 2013) including the Natural Language Toolkit (NLTK), an external library for Python, because it can perform a vast amount of text processing and analysis. We used an NLTK-trained text processing classifier derived from the data set created by Pang and Lee (2005). This classifier contains data from movie and Twitter reviews in which the content of reviews is categorized as positive, negative, and neutral based; these sentiments were used as the basis for comparison. This classification approach is the most widely accepted

1 Available from http://docs.oracle.com/javase/7/docs/api/javax/xml/parsers/SAXParserFactory.html

approach to classifying natural language in sentiment analysis (Piryani, Madhavi, & Singh, 2017). The classification process proceeded as follows:

1. Initially “stop words” (e.g., a, and, the) were filtered from user comments, by importing a corpus of stop words from NLTK library.
2. Featured words were then extracted from the comments into a list in which every word was ordered by the number of times it appeared.
3. The classifier then created a “feature extractor” to select the relevant featured words. The feature extractor returns a list of tuples where each tuple contained a dictionary entry and featured sentiment for each comment.
4. A Naïve Bayes Classifier2 was then used to classify of comments. It uses the probability of each label that is the frequency in the training set and the support from each featured word.
5. The trained classifier was then used to classify user comments and assign a sentiment output number. Possible sentiments range from -1 Fully Negative to 1 Fully Positive. Zero (0) is considered neutral.

For resources that had multiple reviews, we calculated an average (i.e., arithmetic mean) of sentiment classification scores. Once the sentiment analysis scores were assigned, both researchers independently reviewed random selections of the polarity assignments and achieved approximately 90% agreement with the machine assignment.

Phase II: Bivariate Correlation

Using Pearson product-moment correlation analysis, the annotation sentiment average values (N=501) were first correlated with subjective usage data (favorites, features, ratings, and recommendations), then with objective usage data (downloads, views). The researchers chose to use Pearson correlational approach because it is well suited to exploring correlational relationships (Green & Salkind, 2005). Finally, the subjective and objective usage data were correlated absent the average sentiment values. All correlation analyses were conducted using the Statistical Package for the Social Science (SPSS).

Validity Concerns

The limitation with sentiment analysis is that machine learning for this kind of data set is relatively new and has yet to undergo further research and testing to achieve 100% accuracy. Subtle language and domain-specific language can affect sentiment accuracy. As Pang and Lee (2008, p. 21) pointed out:

2 Available from http://nltk.googlecode.com/svn/trunk/doc/api/nltk.classify.naivebayes.NaiveBayes-Classifier-class.html

> Sentiment and subjectivity are quite context-sensitive, and, at a coarser granularity, quite domain dependent (in spite of the fact that the general notion of positive and negative opinions is consistent across different domains). Note that although domain dependency is in part a consequence of changes in vocabulary, even the same expression can indicate different sentiment in different domains.

Neutral reviews required internal parsing to determine polarity. There are certain user comments like “5 star” or “cool” and some comments that begin with strong positive sentiment and ends with weak negative sentiment, for which emotion calculated is neutral, although the review sentiments are considered positive. After consulting research by experts in sentiment analysis on how to address this issue (Wilson, Wiebe, & Hoffmann, 2009), we determined that manual review was needed for the analysis of such comments to determine true expressed sentiment and reclassify them accordingly.

Finally, it should be noted that the prior research done in this area suggests that automated sentiment analysis works best with big datasets and cannot be relied upon for smaller data samples (Pang & Lee, 2008). Despite these potential challenges, each researcher checked different random samples of the polarity assignments and found the sentiment analysis to be 90% accurate.

RESULTS

The NSDL file reflected the textual and non-textual usage data collected for 2,505 records. Of the records in that file, 757 had user-contributed annotations, i.e., comments or reviews. These reviews contained 13,353 words of resource review text, with an average of 18 words per annotation and approximately two sentences per annotation. Figure 1 depicts the distribution of comments across records.

Few records had two comments (n=135). The slope continues downward with 60 records having three comments, 28 records having a fourth comment, 15 records with five, nine records having six comments, six records having seven comments, two records having eight comments, and only one record with nine comments.

The 757 annotation records frequently included “video” (n=376), “students” (n=370), “great” (n=362), “use” (n=354), “good” (n=345), “Excellent” (n=337), “lesson” (n=328), “class” (n=320), “love” (n=312), and “information” (n=304) were the most frequently used terms in resource annotations. These frequencies suggested what we would find in the sentiment analysis.

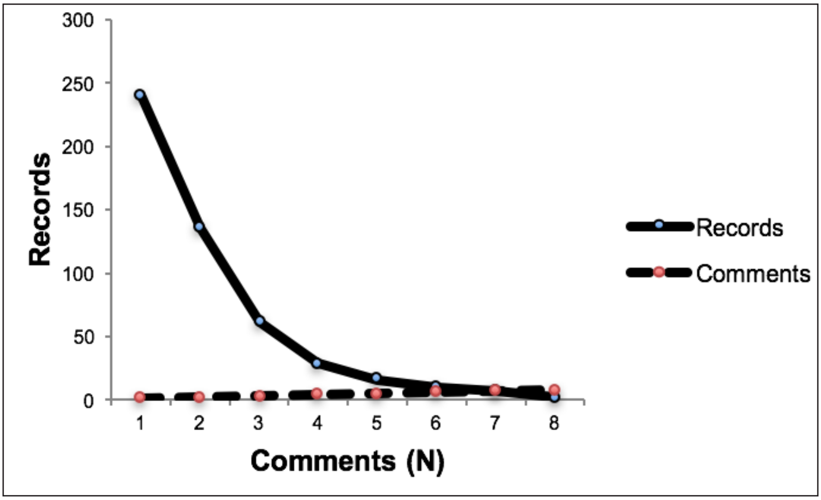

Figure 1. Distribution of annotation entries per record (N=757).

Phase 1. Sentiment Analysis Results

The first step of the sentiment analysis was to examine all of the annotation entries and determine their polarity. Then, we divided the range of sentiment scores into three equal ranges: negative, neutral, and positive. Figure 3 provides an overview of the distribution of the annotations across polarity.

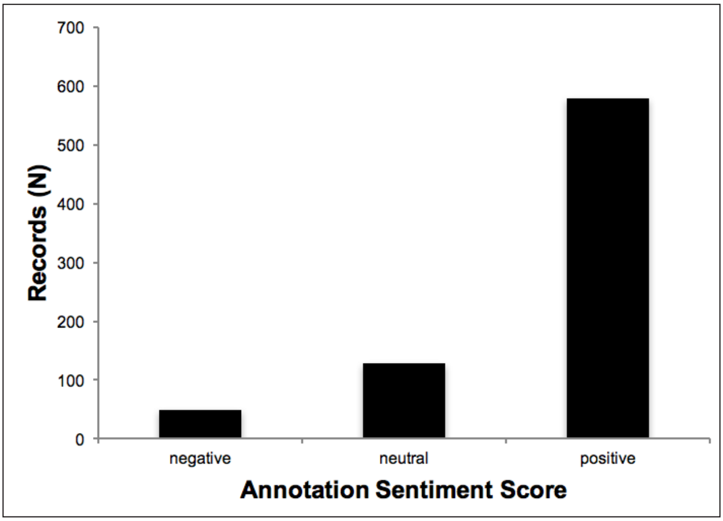

Figure 2. Distribution of annotation polarity (N=757).

As Figure 3 depicts, the data set included 135 neutral annotations with sentiment scores ranging from -.147 to .460, with a mean of .042 and a median of .000. The 49 negative annotations had sentiment scores ranging from -.754 to -.146, with a mean of -.364 and a median of -.404. Most of the annotations were positive (n=579). Positive sentiment scores ranged from .461 to 1.07, with a mean of .508 and a median of .500.

For resources that had more than one review (n=501), we then calculated the arithmetic means of the sentiments for those resources to enable further explorations. Figure 3 depicts the distribution of average annotation sentiment means relative to a normal curve.

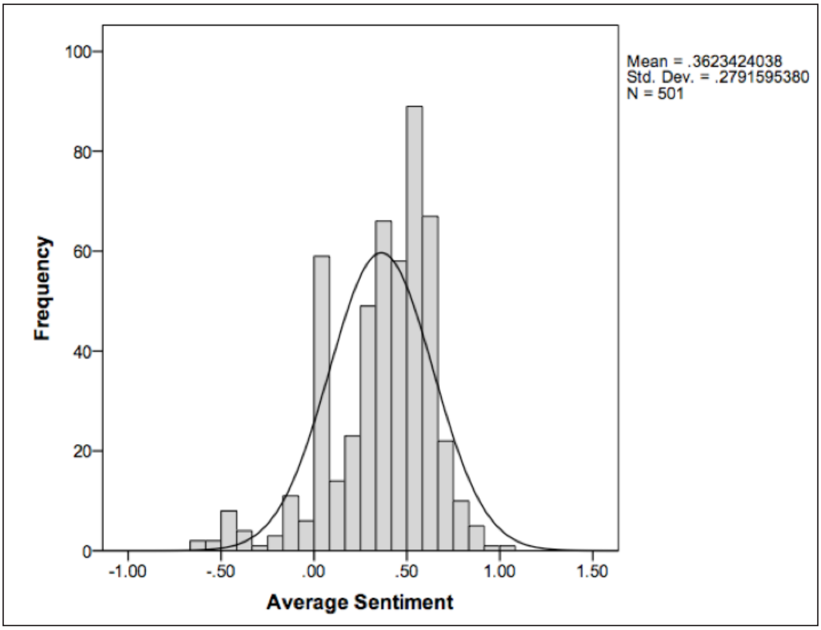

Figure 3. Distribution of average annotation sentiment scores (N=501).

As Figure 3 shows, the distribution of the average annotation sentiment somewhat fits the normal curve, with a skewness of -.956 and a kurtosis of 1.064. Negative values for the skewness indicate data that are skewed left, with the left tail is long relative to the right tail. A positive kurtosis indicates a "peaked" distribution, with few extremely high values.

Next, we plotted the individual sentiment scores versus the average sentiment scores, as Figure 4 depicts.

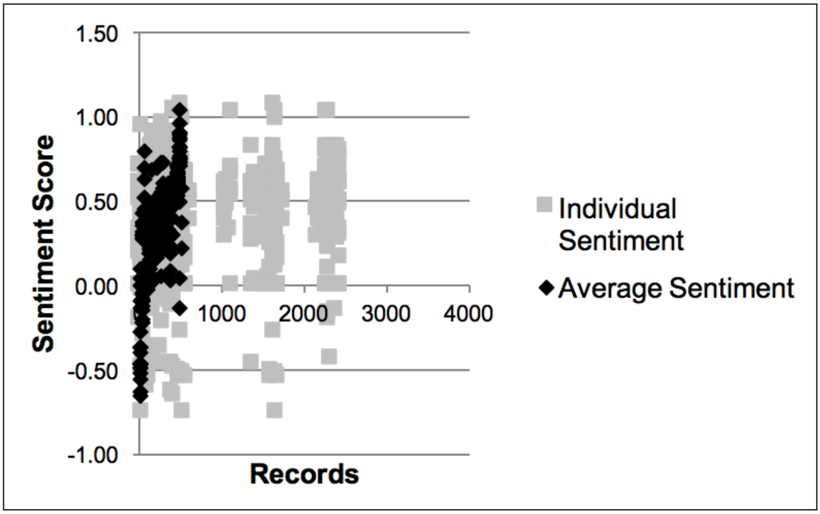

Figure 4. Plots of individual annotation sentiment scores and average sentiment scores.

When the average annotation sentiments were plotted against the individual sentiment scores, the results showed that, while the averaging muted very high and very low values, the majority of the average scores remained clustered similar to the pattern of the individual scores, thus suggesting that they would be viable for use in correlational analyses to detect linear relationships.

Phase 2. Bivariate Correlations

Next, we calculated Pearson Product Moment Correlation coefficients among the six non-textual system-generated usage data and average sentiment scores. Using the Bonferroni approach to control for Type I error across the bivariate correlations, a p value of less that .01 was required for significance (Green & Salkind, 2005). Table 3 depicts the results.

Table 3

Pearson Correlations Between Usage Data Elements (N=2254)

Usage data Element		Downloaded	Favorited	Featured	Rating	Recommended	Viewed	Average Sentiment
Downloaded	Correlation	1	.596	**1.000****			-.372	
	Sig. (2-tailed)		.158	.000			.141	
	N	17	7	10	0	0	17	0
Favorited	Correlation	.596	1	-.285	-.059	.071	**.384****	
	Sig. (2-tailed)	.158		.087	.781	.481	.000	
	N	7	499	37	25	101	326	0
Featured	Correlation	**1.000****	-.285	1			**.305****	
	Sig. (2-tailed)	.000	.087		.000	.000	.002	
	N	10	37	101	18	5	96	0
Rating	Correlation		-.059		1	.271	.148	.158
	Sig. (2-tailed)		.781	.000		.329	.427	.001
	N	0	25	18	519	15	31	471
Recommended	Correlation		.071		.271	1	.201	
	Sig. (2-tailed)		.481	.000	.329		.054	
	N	0	101	5	15	113	93	0
Viewed	Correlation	-.372	**.384****	**.305****	.148	.201	1	
	Sig. (2-tailed)	.141	.000	.002	.427	.054		
	N	17	326	96	31	93	504	0
Average Sentiment	Correlation				**.158****			1
	Sig. (2-tailed)				.001			
	N	0	0	0	471	0	0	501

** $p < .01$

The results of the analyses presented in Table 3 show that four correlations were statistically significant. The most significant correlation was between Featured (i.e., the number of times the resource was a featured resource on the website) and Downloaded at p=1.00. Other significant correlations were between Viewed and Favorited (p=.384) as well as Viewed and Featured (p=305.) and the Average Sentiment and Rating (p=.158).

Although not statistically significant, negative correlations were observed between Viewed and Downloaded (p=-.372), Featured and Favorited (p=-.285), and Rating and Favorited (p=-.059). Average Sentiment and Rating demonstrated very few notable relationships with objective or subjective usage data elements. The findings, when compared to the literature foundation, provided points for discussion about the research questions, which are discussed in the following section.

DISCUSSION

In this section, we examine research questions in light of the presented literature and the research questions.

RQ1. To what extent are users' assessments of resource quality positive, negative, or neutral?

Pang and Lee (2005) established standards for positive, negative, and neutral sentiments in neutral contributions; we used these standards to classify contributors' resource reviews. The literature relating to user reviews suggests that although users most frequently leave comments that express extremely positive or extremely negative sentiment, comment readers and users value comments that include both positive and negative assessment elements. NSDL commenters tended to use a balanced approach because the sentiments primarily contained resource description words and secondarily contained resource use words. Most reviews appeared to be about the format of the resource, i.e., "lessons" and "videos." The fewest annotation words related to the content of resources such as "water" and "Hemisphere."

Our usage data analyses suggested that while users tended to leave positive comments more frequently, those positive comments tended to be weakly positive, weakly negative, or neutral. These kinds of assessments suggest that NSDL collection users, presumably educators, blend positive and negative elements in their annotations, along the lines of the "feedback sandwich" form of teacher feedback in which a negative comment in placed between positive comments (Milan, Parish, & Reichgott, 2006; Parkes, Abercrombie, & McCarty, 2013) or the modulated language that consumers tend to employ when reviewing a product they have experienced (Kronrod

& Danziger, 2013; Mudambi & Schuff, 2010). Unlike to content of consumer reviews, very few of the annotations reflected a very high sentiment score, which few readers find helpful anyway (Ludwig et al., 2013).

RQ2. What is the relationship between sentiment assessments and resource use?

Bivariate correlational analysis revealed very weak or no correlations between annotation sentiment and indicators of resource use, such as incidences of downloading, viewing, or favoriting. This finding is in contrast with prior studies in which researchers reported a relationship between star ratings and downloads (Abramovich et al., 2013). However, the relationship between Featured and Downloaded was significant and strong, suggesting that when a collection provider promotes a resource, additional attention results in a willingness to use the resource. The power of promotion may also account for the significant relationship between Featured and Viewed.

The result that Viewed and Favorited had a significant correlation may be symptomatic of a Facebook-like behavior in which a user reviews quickly and likes (or, here, favorites) any items that seem to appeal (Gerlitz & Helmond, 2013; Sharifrazi & McCabe, 2014). It is difficult to determine the extent to which this relationship proxies an assessment of the resource's quality or usefulness. However, that this relationship is more significant than the relationship between Views and Average Sentiment may be telling of a new type of user feedback preference. From a curation perspective, it may be important for collection providers to consider targeted promotion of their collections to ensure broad use.

CONCLUSION

Driven by the common standards movement and federal education accountability requirements, education trends toward personalized learning are prompting an explosion of interest, and even anxiety, about securing a large supply of high quality open educational resources for K-12 learners (Hanover Research Council, 2012; Hewlett Foundation, 2013; Patrick, Worthen, Frost, & Gentz, 2016). Personalization thrives in an environment that enables teachers to adapt content in ways that allow students to explore, create, and demonstrate their knowledge (de los Arcos, Farrow, Pitt, Weller, & McAndrew, 2016).

In this study, we took an initial look at the extent to which subjective and objective learning resource usage data from the National Science Digital Library, a leading K-12 OER provider, can be used alone to determine resource usage, utility, and quality.

Overall, the sentiment analysis revealed only initial insight into users' resource quality perceptions. Reviews tended to be short with brief, declarative sentences that reflected a narrow range of sentiment. Bivariate correlations suggested that more could be learned from the relationship between non-textual usage data than from the relationship between sentiment text and non-textual usage data elements. These possible limitations may be due to sentiment analysis' underappreciated usefulness for predictions (Gayo-Avello, 2012) or NSDL's modest educational usage data corpus (Lak & Turetken, 2014). However, another possible interpretation of this study's findings is that sentiment analysis of educational usage data is not yet useful to detect valence or affect, and sentiment analysis alone may never be sufficient. This conclusion aligns with Gerlitz and Helmond's (2013) point that user engagement on the web is increasingly consists of quick, easy, measurable button-generated interactions such as ratings and "likes." Despite this uncertainty, the researchers sense that this type of research will yield much more interesting and helpful results as the usage data corpus grow in size and quality. Nonetheless, this study did produce some recommendations and directions for further research.

Limitations, Recommendations, and Directions for Further Research

We acknowledge several limitations to this study should be acknowledged and can be addressed with recommendations for content providers. These areas provide suggested directions for further research.

1. **Beware of "junk" annotations and misspellings.** The data set contained annotations such as "That was crazy Jeffrey wrote this," or "b," or "I really love this sight." Content providers can address this issue by providing a review scaffold with character limits and spell check features. Researchers should be sure to check for misspellings and inappropriate reviews and flag the annotations for further review.
2. **Be attuned to subtle language and overly declarative annotations.** Because research has suggested that users tend to soften their language when writing reviews, sentiment analysis may underrepresent subtle reactions such as "so educational" and "thank you for making this video." Likewise, sentiment analysis may over weigh hyperbolic comments such as "this is the coolest thing I've ever seen." These issues are especially problematic in the analysis of a mixed sentiment comment such as "Very simple. I used as a homework assignment. The students wanted more variety and more challenge, though." Content providers can scaffold with annotation prompts instead of free text boxes that will guide users' contributions with specific questions. Scaffolds can also prompt for comments relating to content and

teaching strategies. Content providers can also employ annotation-voting features to allow users to mark reviews they find helpful.

Researchers can improve and expand the dictionaries they use to detect sentiment and use word, phrase, and concept level analyses (Cambria, Schuller, Liu, Wang, & Havasi, 2013). Researchers can also consider a discipline-specific dictionary that reflects aspects of resource quality drawn from instructional materials selection and teaching literature.

3. **Factor in the user's identity.** User profiles were not part of this study and user job role, experience level, and work site may have a relationship with review content. Collection providers may wish to consider gathering minimal user profile information to gain a sense of the contributor. Researchers may explore differences between collections that contain discipline expert annotations and collections with teacher annotations as well as compare expert and teacher annotation ratings, downloads, and other non-textual usage data. Expert input shows great promise as a basis for machine learning techniques to automatically assess resource quality (Bethard et al., 2009; Wetzler et al., 2013). The extent to which sentiment analysis can factor into improving or extending automatic assessment is a fertile area for exploration.
4. **Factor in the context of contribution.** Many reviews appear to have been left during a professional development event or content provider presentation. To the extent possible, collection providers may wish to capture the context in which the review is contributed, perhaps with a checkbox or radio button selection in the annotation authoring dialogue. Researchers may analyze reviews for indicators of application such as "Students examined this video in small groups…"

The power of usage data to inform personalized learning ecosystems may not be ready to be fully realized, but as the field of learning analytics continues to grow as a result of, or along with, the rising use of OER and their social media features, this type of large-scale educational data analysis offers teachers, students, and researchers an unprecedented opportunity to better understand the role of instructional resources in learning and teaching.

ACKNOWLEDGMENTS

The authors wish to extend their heartfelt gratitude to D. McArthur, J. Weatherley, and D. Finke of the University Corporation for Atmospheric Research (UCAR) for their generous data sharing and support.

This work was supported in part by IMLS RE-04-09-0055, NSF #1304382, and NSF #1449622.

NOTE

A significantly condensed version of this paper was presented in the E-Learn 2016 World Conference on E-Learning virtual papers track.

References

Abramovich, S., & Schunn, C. (2012). Studying teacher selection of resources in an ultra-large scale interactive system: Does metadata guide the way? *Computers & Education, 58*(1), 551-559. doi: http://dx.doi.org/10.1016/j.compedu.2011.09.001

Abramovich, S., Schunn, C. D., & Correnti, R. J. (2013). The role of evaluative metadata in an online teacher resource exchange. *Educational Technology Research and Development, 61*(6), 863-883. doi: 10.1007/s11423-013-9317-2

American Association of School Librarians [AASL]. (2011). School libraries count! National longitudinal survey of school library programs. from http://www.ala.org/aasl/sites/ala.org.aasl/files/content/researchandstatistics/slcsurvey/2011/AASL-SLC-2011-FINALweb.pdf

Ash, K. (2012). Common Core drives interest in open education resources: Spurred by the adoption of common-core standards by nearly every state, the movement for open digital resources is growing as educators realign curricula. *Digital Directions, 6*(1), 42.Retrieved from http://www.edweek.org/dd/articles/2012/10/17/01open.h06.html?tkn=NTOFMxrZ1BSvAwIeHWdbG7nRcERE8nd99JI3&print=1

Bethard, S., Wetzler, P., Butcher, K., Martin, J. H., & Sumner, T. (2009). *Automatically characterizing resource quality for educational digital libraries.* Paper presented at the Joint Conference on Digital Libraries (JCDL) '09, June 15-19, 2009, Austin, Texas, USA.

Bienkowski, M., Brecht, J., & Klo, J. (2012). *The Learning Registry: Building a foundation for learning resource analytics.* Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada. http://dl.acm.org/citation.cfm?doid=2330601.2330651

Cambria, E., Schuller, B., Liu, B., Wang, H., & Havasi, C. (2013). Statistical approaches to concept-level sentiment analysis. *Intelligent Systems, IEEE, 28*(3), 6-9. doi: 10.1109/MIS.2013.68

Campbell, L. M., & Barker, P. (2013). Activity data and paradata: A briefing paper. Retrieved February 5, 2014, from http://publications.cetis.ac.uk/wp-content/uploads/2013/05/paradataBriefing.pdf

Chen, C.-M., & Sun, Y.-C. (2012). Assessing the effects of different multimedia materials on emotions and learning performance for visual and verbal style learners. *Computers & Education, 59*(4), 1273-1285. doi: 10.1016/j.compedu.2012.05.006

Collins, S., & Levy, P. (2013, March). Guide to the use of Open Educational Resources in K-12 and postsecondary education. Retrieved June 25, 2013, from https://www.siia.net/index.php?option=com_docman&task=doc_download&gid=4029&Itemid=318

de los Arcos, B., Farrow, R., Pitt, R., Weller, M., & McAndrew, P. (2016). Adapting the curriculum: How K-12 teachers perceive the role of open educational resources. *Journal of Online Learning Research, 2*(1), 23-40.

Digital Textbook Collaborative. (2012). The digital textbook playbook. Retrieved March 10, 2012, from http://transition.fcc.gov/files/Digital_Textbook_Playbook.pdf

Essa, A. (2016). A possible future for next generation adaptive learning systems. *Smart Learning Environments, 3*(1), 16. doi: 10.1186/s40561-016-0038-y

Ganu, G., Kakodkar, Y., & Marian, A. (2013). Improving the quality of predictions using textual information in online user reviews. *Information Systems, 38*(1), 1-15. doi: 10.1016/j.is.2012.03.001

Gayo-Avello. (2012, April 28). "I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper" -- A balanced survey on election prediction using Twitter data. Retrieved April 20, 2014, from http://arxiv.org/pdf/1204.6441v1

Gerlitz, C., & Helmond, A. (2013). The like economy: Social buttons and the data-intensive web. *New Media & Society, 15*(8), 1348-1365. doi: 10.1177/1461444812472322

Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh for analyzing and understanding data.* Upper Saddle River, NJ: Pearson Prentice Hall.

Griffin, A. (2013, July). Subjective and objective reviews of instructional materials. Retrieved April 1, 2014, from http://simra.us/wp/news/subjective-and-objective-reviews-of-instructional-materials-2/

Hanover Research Council. (2011, October). K-12 STEM education overview. Retrieved July 12, 2012, from http://www.hanoverresearch.com/wp-content/uploads/2011/12/K-12-STEM-Education-Overview-Membership.pdf

Hanover Research Council. (2012, October). Best practices in personalized learning environments (Grades 4 – 9). from http://www.hanoverresearch.com/media/Best-Practices-in-Personalized-Learning-Environments.pdf

Hanson, K., & Carlson, B. (2005). Effective access: Teachers' use of digital resources in STEM teaching. Newton, MA: Education Development Center, Inc.

Hewlett Foundation. Open Educational Resources. Retrieved May 5, 2014, from http://www.hewlett.org/programs/education/open-educational-resources

Hewlett Foundation. (2013, November 22). White paper: Open Education Resources: Breaking the lockbox on education. from http://www.hewlett.org/sites/default/files/OER White Paper Nov 22 2013 Final_0.pdf

Kronrod, A. N. N., & Danziger, S. (2013). "Wii will rock you!" The use and effect of figurative language in consumer reviews of hedonic and utilitarian consumption. *Journal of Consumer Research, 40*(4), 726-739. doi: 10.1086/671998

Lagoze, C. J. (2010). *Lost identity: The assimilation of digital libraries into the web*. (PhD Dissertation), Cornell University, Ithaca, NY. ProQuest Dissertations & Theses Database database. (AAT 3396228)

Lak, P., & Turetken, O. (2014). *Star ratings versus sentiment analysis: A comparison of explicit and implicit measures of opinions.* Paper presented at the 47th Hawaii International Conference on System Sciences (HICSS), 6-9 Jan. 2014 Waikoloa, HI

Leutkemeyer, J., & Mardis, M. A. (2016). Applying the quadratic usage framework to research on K–12 STEM digital learning resources. *School Library Research, 19*. Retrieved from http://www.ala.org/aasl/sites/ala.org.aasl/files/content/aaslpubsandjournals/slr/vol19/SLR_K-12STEMDigitalLearningResources_V19.pdf

Liu, Y., Yu, X., An, A., & Huang, X. (2013). Riding the tide of sentiment change: sentiment analysis with evolving online reviews. *World Wide Web, 16*(4), 477-496. doi: 10.1007/s11280-012-0179-z

Ludwig, S., de Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates. *Journal of Marketing, 77*(1), 87-103.

MacNeill, S., Campbell, L. M., & Hawksey, M. (2014). Analytics for education. *Journal of Interactive Media In Education.* Retrieved from http://www-jime.open.ac.uk/jime/article/view/2014-07%3C/div>

Mardis, M. A. (2003). If we build it, will they come? An overview of the issues in K-12 digital libraries. In M. Mardis (Ed.), *Developing digital libraries for K-12 education.* Syracuse, NY: ERIC Information Technology Clearinghouse.

Mardis, M. A., ElBasri, T., Norton, S. K., & Newsum, J. (2012). The new digital lives of U.S. teachers: A research synthesis and trends to watch. *School Libraries Worldwide, 18*(1), 70-86.

Mardis, M. A., & Everhart, N. (2013). From paper to pixel: Digital textbooks and Florida schools. In M. Orey, S. A. Jones & R. M. Branch (Eds.), *Educational Media and Technology Yearbook* (Vol. 37, pp. 93-118). New York, NY: Springer.

Mardis, M. A., & Howe, K. (2010). STEM for our students: Content or co-conspiracy? *Knowledge Quest, 39*(2), 8-11.

Mardis, M. A., & Zia, L. (2003, January/February). Leading the wave of science and mathematics learning innovation. *Knowledge Quest, 31*, 8-9.

Massart, D., & Shulman, E. (2013). Unlocking open educational resources (OERs) interaction data. *D-LIb, 19*(5/6).Retrieved from http://www.dlib.org/dlib/may13/massart/05massart.html doi:10.1045/may2013-massart

Maull, K. E., Saldivar, M. G., & Sumner, T. (2010). *Observing the online behavior of teachers: From Internet usage to personalization for pedagogical practice.* Paper presented at the Association for Computing Machinery Conference on Human Factors in Computing Systems, Atlanta, GA. http://communication.ucsd.edu/barry/chi-ws10/maull_positionpaper_chi2010ws.pdf

McMartin, F., Iverson, E., Wolf, A., Morrill, J., Morgan, G., & Manduca, C. (2008). The use of online digital resources and educational digital libraries in higher education. *International Journal on Digital Libraries, 9*(1), 65-79. doi: 10.1007/s00799-008-0036-y

Mickey, K., & Meaney, K. (2010). Simba Information's 2010 national textbook adoption scorecard and 2011 outlook. Stamford, CT: Simba Information.

Mickey, K., & Meaney, K. (2011). Simba Information's 2011 national textbook adoption scorecard and 2012 outlook. Stamford, CT.

Mickey, K., & Meaney, K. (2013). Getting ready for the Common Core 2013-2014. Stamford, CT: Simba Information, Inc.

Milan, F. B., Parish, S. J., & Reichgott, M. J. (2006). A Model for Educational Feedback Based on Clinical Communication Skills Strategies: Beyond the "Feedback Sandwich". *Teaching and Learning in Medicine, 18*(1), 42-47. doi: 10.1207/s15328015tlm1801_9

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly, 34*(1), 185-200.
National Science Digital Library [NSDL]. (2013). *OER digital content: A combined voices meeting, June 27-30, 2013.* University Corporation for Atmospheric Research [UCAR]. Boulder, CO.
National Science Digital Library [NSDL]. (2014a). Comm._anno. from https://wiki.ucar.edu/display/nsdldocs/comm_anno
National Science Digital Library [NSDL]. (2014b). Comm._para. from https://wiki.ucar.edu/display/nsdldocs/comm_para
National Science Digital Library [NSDL]. (2014c). General reference and overview. from https://wiki.ucar.edu/display/nsdldocs/NSDL+Metadata+Formats
New Media Consortium [NMC]. (2014). NMC Horizon Report 2014 K-12 edition. from http://cdn.nmc.org/media/2014-nmc-horizon-report-k12-EN.pdf
Niemann, K., Scheffel, M., & Wolpers, M. (2012). *A comparison of usage data formats for recommendations in TEL.* Paper presented at the 2nd Workshop on Recommender Systems in Technology Enhanced Learning 2012 in conjunction with the 7th European Conference on Technology Enhanced Learning (EC-TEL 2012), September 18-19, Saarbrücken, Germany.
Okerson, A. (2000). Are we there yet? Online E-resources ten years after. *Library Trends, 48*(4), 671-693.
Padmaja, S., & Fatima, S. (2013). Opinion mining and sentiment analysis –An assessment of peoples' belief: A survey. *International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC), 4*(1), 21-33.
Pang, B., & Lee, L. (2005). *Sentence polarity dataset.* Retrieved from: http://www.cs.cornell.edu/people/pabo/movie-review-data/rt-polaritydata.tar.gz
Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1-135. doi: 10.1561/1500000001
Parkes, J., Abercrombie, S., & McCarty, T. (2013). Feedback sandwiches affect perceptions but not performance. *Advances in Health Sciences Education, 18*(3), 397-407. doi: 10.1007/s10459-012-9377-9
Patrick, S., Worthen, M., Frost, D., & Gentz, S. (2016, October). Meeting the Every Student Succeeds Act's promise: State policy to support personalized learning. from http://www.inacol.org/wp-content/uploads/2016/10/iNACOL_MeetingESSAsPromise.pdf
Perrault, A. M. (2007). An exploratory study of biology teachers' online information seeking practices. *School Library Research, 10*. Retrieved from http://www.ala.org/ala/aasl/aaslpubsandjournals/slmrb/slmrcontents/volume10/biology.htm
Piryani, R., Madhavi, D., & Singh, V. K. (2017). Analytical mapping of opinion mining and sentiment analysis research during 2000-2015. *Information Processing & Management, 53*(1), 122-150. doi: http://dx.doi.org/10.1016/j.ipm.2016.07.001
Porcello, D., & Hsi, S. (2013). Curating and crowdsourcing online education resources. *Science, 34*(6143), 240-241. doi: 10.1126/science.1234722
Price, L. (2007). Lecturers' vs. students' perceptions of the accessibility of instructional materials. *Instructional Science, 35*(4), 317-341. doi: 10.1007/s11251-006-9009-y
Project Tomorrow. (2012a, April). Mapping a personalized learning journey – K-12 students and parents connect the dots with digital learning. Retrieved June 10, 2012, from http://www.tomorrow.org/speakup/pdfs/SU11_PersonalizedLearning_Students.pdf

Project Tomorrow. (2012b, May). Personalizing the classroom experience – Teachers, librarians and administrators connect the dots with digital learning. Retrieved June 28, 2012, from http://www.tomorrow.org/speakup/SU11_PersonalizedClassroom_EducatorsReport.html

Project Tomorrow, & Blackboard. (2016, June). Trends in digital learning: How K-12 leaders are empowering personalized learning in America's schools. from http://images.email.blackboard.com/Web/BlackboardInc/%7Ba9f75675-5977-477d-8a7e-8589e19aa052%7D_Bb_TrendsinDigitalLeaningReport_Final.pdf

Recker, M., Leary, H., Walker, A., Diekema, A., Wetzler, P., Sumner, T., & Martin, J. H. (2011). Modeling teacher ratings of online resources: A human-machine approach to quality. Paper presented at the American Educational Research Association (AERA) Annual Meeting, April 8-11, New Orleans, LA.

Recker, M., Walker, A., Giersch, S., Mao, X., Halioris, S., Palmer, B., Johnson, D., Leary, H., & Robertshaw, M. B. (2007). A study of teachers' use of online learning resources to design classroom activities. *New Review of Hypermedia and Multimethods, 13*(2), 117-134. doi: 10.1080/13614560701709846

Rosenbaum, S. (2011). *Curation nation: Why the future of content is context.* New York, NY: McGraw Hill.

School Library Journal. (2013, December 5). School technology survey: U.S. school libraries 2013. from http://www.slj.com/downloads/slj-technology-survey/

Seaman, J. E., & Allen, I. E. (2014). Listening to sentiment. *Quality Progress, 47*(2), 48-50.

Sharifrazi, F., & McCabe, M. B. (2014). Perils of click like, share or comment on social media networks. *ASBBS Proceedings, 21*(1), 739-745.

Simba Information, I. (2014). 2013 National Instructional Materials Adoption Scorecard and 2014 Forecast. Stamford, CT: Author.

Spiegel, D. L. (1989). Instructional resources: Evaluating instructional materials. *The Reading Teacher, 43*(1), 72-73. doi: 10.2307/20140207

Springer, N., Pfaffinger, C., & Engelmann, I. (2015). User comments: Motives and inhibitors to write and read. *Information, Communication & Society, 18*(7), 798-815. doi: 10.1080/1369118X.2014.997268

Stern, L., & Roseman, J. E. (2004). Can middle-school science textbooks help students learn important ideas? Findings from Project 2061's curriculum evaluation study: Life science. *Journal of Research in Science Teaching, 41*(6), 538-568. doi: 10.1002/tea.20019

Sumner, T., Khoo, M., Recker, M., & Marlino, M. (2003). *Understanding educator perceptions of "quality" in digital libraries.* Paper presented at the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries Houston (JCDL), Houston, TX.

Tenenboim, O., & Cohen, A. A. (2015). What prompts users to click and comment: A longitudinal study of online news. *Journalism, 16*(2), 198-217. doi: 10.1177/1464884913513996

Tonks, D., Patrick, S., & Bliss, T. (2013, June). OER and collaborative content development. from http://www.inacol.org/wp-content/uploads/2015/02/oer-and-collaborative-content.pdf

U.S. Department of Education. (2012, February). Expanding evidence approaches for learning in a digital world. from http://www.ed.gov/edblogs/technology/files/2013/02/Expanding-Evidence-Approaches.pdf

U.S. Department of Education. (2013, October). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. from http://www.ed.gov/edblogs/technology/files/2012/03/edm-la-brief.pdf

U.S. Department of Education. (2014). Learning Registry: Sharing what we know. Retrieved March 8, 2014, from http://learningregistry.org/

U.S. Department of Education. (2017a, March). #GoOpen district launch packet, 1.3. from https://tech.ed.gov/files/2017/03/GoOpen-District-Launch-Packet-2017-V-1.3.pdf

U.S. Department of Education. (2017b, January). Reimagining the role of technology in education: 2017 National Education Technology Plan update. from https://tech.ed.gov/files/2017/01/NETP17.pdf

Van Acker, F., van Buuren, H., Kreijns, K., & Vermeulen, M. (2013). Why teachers share educational resources: A social exchange perspective. In R. McGreal, K. W. & M. S. (Eds.), *Open educational resources : Innovation, research and practice.* (pp. 177-192). Vancouver, CA: Commonwealth of Learning and Athabasca University.

VanGundy, S. (2010a, February). The STEM Exchange: Enabling next generation approaches for community access to federally funded online materials. . from https://wiki.ucar.edu/download/attachments/236520324/STEM Exchange_NSDL Exec Summary_0.pdf

VanGundy, S. (2010b). What is paradata. from http://nsdlnetwork.org/stemexchange/paradata

Wetzler, P., Bethard, S., Leary, H., Butcher, K., Bahreni, S. D., Zhao, J., Martin, J., & Sumner, T. (2013). Characterizing and predicting the multifaceted nature of quality in educational web resources. *ACM Transactions on Interactive Intelligent Systems, 3*(3), 15:11-15:25. doi: 10.1145/2533670.2533673

Williams, D. A., & Coles, L. (2003). The use of research information by teachers: Information literacy, access and attitudes. A report for the Economic and Social Research Council. Aberdeen, Scotland: Robert Gordon University.

Wilson, T., Wiebe, J., & Hoffmann, P. (2009). Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational Linguistics, 35*(3), 399-434.

YouGov. (2014, January 22). Reviews: Fieldwork Dates: 17th - 20th January 2014. Retrieved May 5, 2014, from http://cdn.yougov.com/r/14/YouGov_Omnibus_Online_Reviews.xls

Ziegele, M., Breiner, T., & Quiring, O. (2014). What creates interactivity in online news discussions? An exploratory analysis of discussion factors in user comments on news items. *Journal of Communication, 64*(6), 1111-1138. doi: 10.1111/jcom.12123