

Findability, Accessibility, and Usability of Data Portals in Education

Deborah Anne Banker

Tarleton State University

1333 W Washington St, Stephenville

TX 76401, United States

Abstract

There is a need for stakeholders, at all levels of education, to be able to locate, access, and use the preponderance of data available about schools and students' demographic information and academic performance. The reason for findability, accessibility and usability of the data is for the stakeholders to make data-driven decisions based on trends the data bring to light. However, in many cases, data are located in portals that are cumbersome and difficult to navigate. This article explores the complexity of educational data portals; the wide gap in technology skills of the user of the data; and, interface design challenges.

Keywords: Data Analytics, Data Accessibility, Data Portals

One of the most exciting elements of data analytics in education is the ability it can afford educators to identify and predict trends in many different areas (e.g. academic achievement, student enrollment, student retention, career interests, pedagogical preferences, etc.) for data-driven decisions to improve student learning. The successful use of data analytics starts with educational data mining (EDM). EDM is the process of using computational approaches (statistical algorithms) to analyze different data sets collected by various educational institutions (Romero & Ventura, 2010). In *A Roadmap for Educational Technology*, 2010, there are seven information technology areas that hold great promise for the future to better help students learn. Those areas are: 1) user modeling; 2) mobile tools; 3) networking tools; 4) serious games; 5) intelligent environments; 6) educational data mining; and, 7) rich interfaces (Computing Community Consortium, 2010, p. 43). As one of the seven information technology areas, EDM is knowledge discovery in databases (Siguenza, Saquicela, Avila-Ordonez, Vandewalle, & Cattrysse, 2015). Knowledge discovery occurs when EDM uses statistical methods that search for new and generalizable relationships rather than testing a hypothesis (Slater, Joksimovic, Kovanovic, Baker & Gasevic, 2017). The new and generalizable relationships can then be used to make decisions that will help learning professionals better design learning for students (Masie, 2013). In other words, the use of EDM can produce knowledge, such as trends and patterns, that was not known before, about the relationship between learning and students from the statistical analysis done on large sets of data. (Siguenza et al., 2015). Due to advances in technology, a large volume, a wide variety, and high speeds of data about students are being generated and collected (Wang, 2017). These data are being stored in computer databases (Kumar & Chadha, 2011). The "gateway" or "entrance-way" to all this information contained in the computer databases is known as an information, or data, portal (Tatnall, 2005). For all stakeholders to be able to use EDM analytics effectively and in a meaningful manner, the data portal must have three distinctive design characteristics, i.e., findability (Wilkie & Azzopardi, 2013), accessibility (Smith, 1994), and usability (Berzal, Cubero, Marin, Serrano, & Blanco, 2003).

The purpose of this article is to explore the elements of findability, accessibility, and usability as they relate to pedagogical and interface design issues of data portals. Before those elements are explored, it is important to provide foundation information as to why findability, accessibility, and usability are becoming increasingly important. Since we are looking at the findability, accessibility, and usability of data portals in education, a good place to begin is to look at what constitutes a data portal, what is contained in a data portal, i.e., "big data" as termed in the literature, and who traditionally works with the data portals. It is important to note that "big data" is not about the number of pieces of data, "it is about the *analytics*" of the data (King, 2016; Wang, 2017, p. 6). In addition, the article delves into data quality and why data need to be useful as well as in formats available and usable to educational stakeholders and learning professionals.

Data Portals

A portal is a single point of access to information (Tatnall, 2005). In the context of this article, a data portal is a single point of access to data sets located on the internet. Another way to think of a data portal is that it is a gateway or entry point (door) to where data are kept or housed. Sullivan (2004) describes the data portal as the interface between the user and the data. An example of a portal would be the opening web page of a website like what you experience when

you go to AOL.com or Amazon.com. For an additional viewpoint, Shivakumar (2016) states that a portal can provide an information aggregation platform. He further explains that, “portals provide a convergence of various technologies, platforms, and disparate applications and services (p.4).

Various design considerations of a data portal would include

- users of the data portal must be able to search for, acquire, interpret, store, retrieve, and re-use data;
- system developers of the data portal must provide portals that are collaborative workspaces that foster creation, sharing, and re-use of data; and,
- managers of the data portal must be able improve the flow and exchange of information for users. They also must establish protocols to allow information and data to be securely managed (Deltor, 2000).

What has been evident in the literature about data portals is that there are some challenges. An example would be that there are no common standards with respect to a semantic schema (Cope & Kalantis, 2015; Wolf, 2010). In other words, there is no shared educational data dictionary, i.e., common terms to describe the different search (keyword) possibilities, platforms, profile specific filters, etc. (Zezula, 2014; Morville & Sullenger, 2010). What this lack of a shared educational dictionary can cause is a difficulty in data retrieval. Another challenge mentioned in the literature is that one organization could have several different data portals that are designed differently, and the different portals are not often compatible with one another (Woolf, 2010). A third challenge mentioned is that too many portals are designed for a particular audience who have a relatively sophisticated set of computer/technical skills. However, the audience who are the actual users of the portal, usually do not have the same technical skills as the designer and, in fact, the members of the audience have various levels of technical skills (Morville & Sullenger, 2010). This could create the impression for the user audience that the data portal is not the most user-friendly.

Big Data in Education

According to Nasie, 2013, in education, enormous amounts of data are collected by learning professionals, usually from classrooms or programs, i.e., school or school district levels. The data are for the most part, inexpensive and easy to acquire. Although the term “big data” is used frequently in the literature, there is no exact number or definition associated with what separates “big data” from “small data” (Wang, 2017; Henry & Venkatraman, 2015)). However, what is found in the review of the literature are parameters of large volume, wide variety, and high velocity to distinguish big data from small data (boyd & Crawford, 2012). According to Wang, 2017, the big data schools and school districts provide are information on students, teachers, administrators, schools, and even communities. Big data allows education professionals to analyze, compare, contrast, and evaluate “enormous streams of data.” (Nasie, 2013, p. 33). This analysis/evaluation can help “better design learning, better evaluate the impact of learning, better fuel an evidence-based approach to experimentation, and better create personalization.” (Nasie, 2013, p. 33). An example of big data are the results of surveys administered to large groups of students. With an increased emphasis on the movement for personalized student learning, instruments are used, such as the CDE Personalized Learning Survey of 2015 and data analytics to determine what the personalized learning paradigm might look like by predicting trends and/or patterns found in the results (Dunlap-Kahren, 2015).

The use of data analytics can predict trends that can pinpoint a flaw in instructional delivery or modality. An example would be test question analytics used to determine how many learners consistently miss a concept tested on an exam. Analysis of the data may point to issues such as the way questions are phrased and/or, determine that questions are not aligned with the mandated learning outcomes. Learning professionals, through data analysis, could compare convenience samples of their populations to larger educational community samples to validate an instructional methodology variance that needed to be changed to meet the needs of the localized learners. In a literature review, Chiappe and Rodriguez, 2017, found that EDM and the learning analytics involved offer the possibility of educational stakeholders being “better informed and thus make better and timely decisions” (para. 3). An example of this statement would be in the case of a learner who is struggling with a particular unit of study. Through learning analytics, a search could be performed through a database of similar learners who struggled but then eventually succeeded with the same instructional unit. In the successful learner profiles would be the supporting unit of instruction. The supporting unit of instruction could then be given to the struggling learner in an effort to create personalized learning for him/her (Spector, 2013). Klačnja-Milićević, Ivanović, and Budimac (2017) offer that learning analytics can: make predictions about future uses of learning sequences; learn the relationship between course materials and a learner’s academic performance; and, determine significant relationships between learner knowledge level, learner study time, and learner academic achievement. With the growing need to collect and analyze huge sets of data, there are two major issues: 1) the abilities of educators to find and make use of the available data, i.e., *findability* and *usability*, and 2) educators do not always have the knowledge and tools necessary to easily access and accomplish successful data analysis, i.e., *accessibility* and *data (learning) analytics skills* (Rolleston, Howe & Sprague, 2015).

When discussing the successful integration of technology tools and data analysis to improve learning and instruction, Chiappe and Rodriguez (2017) comment that making better and timely decisions will not be possible if the “skills related to learning analytics continue to be restricted to highly technical profiles” (para. 3). Highly technical profiles in this case would refer to individuals who have a set of unique technical skills and who collect and analyze data, i.e., data scientists (Kim, Zimmermann, DeLine, & Begel, 2016). To illustrate the unique technical skills an individual would need to possess and the complexity of tools an individual may encounter in employing data analytics, the following was presented by Slater, *et. al.* (2017) in reference to analyzing data from a Massive Open Online Course (MOOC) to improve student learning. The different program tools are in parentheses:

A researcher may have data on 60 million system transactions in a popular MOOC. From this data set, he or she wishes to select only data of a particular year (*SQL*), then refine that data set to calculate total student time in the system (*Excel*) before fitting a predictive model (*Rapid Miner*) that analyzes the relationship between forum posts and replies (*NodeXL*) and overall textual quality of posts and replies by that student (*CohMetrix*). Finally, this researcher may wish to visualize the most interesting clusters of students found with the social network data (*Gephi*). (Slater *et al.*, 2017, p. 102).

As illustrated above, although not impossible to find, but rare is the educational stakeholder who would have the skillset to know and use all the different and ever-evolving data analysis program tools listed above (Rolleston *et al.*, 2015). The individual who does have that skillset is the data scientist (Kim, *et al.* 2016). The Global Institute predicted that just for the United States alone, by 2018 there will be shortage of 1.5 million data scientists who are fully trained to meet the job market demand for those individuals (Asamoah, Sharda, Zadeh, & Kalgotra, 2017). Therefore, data portals need to be designed so that authorized users who “have different expertise and know-how on using electronic media” can find, access, and use the data portal (Wimmer & Holler, p. 167).

Data Scientist

As defined by Kim *et al.* (2016), a data scientist is a professional who builds data platforms, collect data, build predictive algorithms, and analyze data. This individual also may use other analytic tools to spot trends, patterns, and relationships in data sets. Since the term was just coined in 2008 by Patil and Hammerbacher to describe their job functions at Facebook and LinkedIn (Kim *et al.*, 2016), the data scientist is the newest specialization in the information management field and is in high demand (Strawn, 2016). The fact that the specialization of a data scientist is a recent occurrence is corroborated by Asamoah *et al.* (2017) as they describe a design for a big data analytics course for their university’s business school in the Management Information System program. Strawn (2016) further speculates that the data scientist is “not only in demand in Silicon Valley, but also in company headquarters around the world.” (para. 18). Further, investigation found that data scientists are not only employed in business/industries, but also healthcare facilities, and government entities which produces a huge demand for these individuals (Mills, Chudoba, & Olsen, 2016). There is a profound shortage of data scientists (Power and Hermacinski, 2013) and they are extremely well-paid positions (Columbus, 2014). Due to this shortage and the salary demands, educational institutions may not have access to data scientists. This could create problems as less skilled individuals, i.e., education stakeholders, conduct analyses and potentially produce faulty data that other stakeholders may use to make decisions. Therefore, when designing web portals, data scientists play a major role in how to organize information on a page, logically categorize and group related content to provide a context for users, and use taxonomies that aid in navigation (Sullivan, 2004). First and foremost, a data scientist must create a data portal with findability.

Data Portals Findability

Findability is as it sounds when searching the literature. Operationally, in the context of data portals, the portal user can find the data portal itself, and the portal user can find what he/she needs once in the portal. (Sullivan, 2004). In a 2013 study to measure findability, Wilke and Azzopardi found that if a user cannot find what they are looking for he/she will become frustrated and leave the website. To measure findability, they used *navigability* as one measure and *retrievability* as another measure. They defined the navigability measure as the “ease with which a user can traverse a website by clicking through the link structure” (p. 809). They also defined the retrievability measure as “how easily a document can be retrieved using a search engine” (p. 809). Since we now live in a world with ever increasing amounts of information, the information is of little use if an individual cannot find it (Morville & Sullenger, 2010).

A challenge for findability is the lack of a common nomenclature for educational data terminology to be used in searching for data and data portals (Woolf, 2010). A great deal of time is given over by the user of the data as he/she searches the internet for the data portal, and then for the data category itself once in the portal, using different keywords because there is no common lexicon (Wilkie & Azzopardi). Sullivan (2004) suggests the use of well-defined content taxonomies in portal design so that there is an organizing framework to ensure findability of data and data portals. Once the data portal is found and the data within it, the next step would be accessibility of the information.

Data Portal Accessibility

For the purposes of this article, *accessibility* will be operationally defined as the process of requesting, sending, and receiving data (U.S. Patent No. 6,233,618, 2001). Therefore, data portal accessibility would entail the requesting, sending and receiving data from the data portal. It is important to keep in mind that data portal accessibility has been the responsibility of the data portal content developers, *i.e.*, the data scientist (Henry, 2006). As a relatively new profession, according to Asamoah, *et al*, 2017 the data scientist curriculum consists of “advanced analytical, information technology, business, and technical communication skills necessary to harness the power of big data” (p. 164). There was no mention of clear, consistent design and navigation of data portals as Henry (2006) states is necessary for one aspect of accessibility. This is not to say that a data portal for an educational institution containing private and sensitive information on students, teachers, administrators, and others should be open to all possible stakeholders. One would expect this sensitive data to have security measures, such as passwords, firewalls, or encryption coding as part of the portal access protocol (U.S. Patent No. 6,219,706 B1, 2001). However, accessibility would indicate that any *authorized* stakeholder is not hindered by the technology or portal design from accessing and using data portals to make data-driven decisions. Any authorized stakeholder such as a classroom teacher, administrator, instructional specialist, or program assessment professional, should be able to access the data portal with little difficulty. It is important to note that a data portal may have stringent and redundant security to protect student privacy, yet to meet accessibility, the interface needs to be easy to understand and navigate (U. S. Department of Education, 2017).

According to Sullivan (2004), the design of the data portal is key to accessibility to the information and content in the portal. He states that a portal should be designed to consider how to organize information so that it makes sense to a wide variety of users for ease of navigation from webpage to webpage. In addition, content pages should follow a pattern with landmarks and links back to a home page, as well as “labels should be applied consistently through the portal” (p. 13). The links should be named logically so that the “user’s experience is consistent across the portal” (p. 12).

According to *A Roadmap for Education Technology* (2010), another challenge for accessibility of a data portal is the lack of a “shared data dictionary that prevents duplication of efforts and streamlines the use of nomenclature and categorization” (p. 65). This is corroborated by West (2012) as he points out that the inconsistent terminology can cause a problem with school system administrators to aggregate and analyze information. As a side note, it is important to think about the validity of any analytics that may be done. “If access to the original data from published studies was routine, then faking results would become more tricky” (Smith, 1994, pg. 1519). When a major decision must be made, *e.g.*, a curriculum overhaul, data and datasets should be widely accessible to many individuals, who may have excellent analytical skills, but not have complex data portal skills, so that the replication of the results could be accomplished to verify the determinations (Leetaru, 2017). However, before the analytics can be validated, a user must be able to interact with, and use the portal, *i.e.*, usability.

Data Portals Usability

EDM requires a data portal that is usable. To review, “usability refers to the quality of a user’s experience when interacting with products or systems” (U.S. Dept. of Health & Human Services, 2017). Further from this report, usability can be a combination of factors depending on the user. Those factors include intuitive design and ease of learning the interface (how well a novice user could tell by looking at the data portal where to click to start a data mining project), efficiency of using the interface (whether users can get the information they desire quickly and without wasting time navigating through complex protocols), memorability (how easily a user can reestablish proficiency (Nielsen, 2012, para. 3)), error frequency and severity (how many times the user runs into a problem and if they he/she is able to overcome the problem, or does the problem persist (Nielsen, 1995)).

An additional element to usability is whether or not a data portal is presented in a format that is highly unusual, unfamiliar, or inconvenient. Also, the database within the data portal may be system-logged and not reachable (Slater, *et al.*, 2017). What is meant by system-logged is that there are layers upon layers of protocol or complex commands that must be used to reach the data within the database in the data portal. Those protocols or complex demands are not intuitive nor easily obtainable from a master framework or orientation as the master framework or orientation may not exist (Woolf, 2015).

Usability is how effectively, efficiently and satisfactorily a user can interact with a user interface (U.S. Dept. of Health & Human Services, 2018). Oftentimes, the terms *accessibility* and *usability* are used interchangeably, but the terms have distinctly different meanings.

Rubin and Chisnell (2008) define usability as, “the user can do what he/she wants to do the way he/she expects to be able to do it, without hindrance, hesitation or questions” (p. 4). Another way to think of usability is to consider what users are able to do when they access the data portal. Will they be able to do anything without a lot of training? (Nielsen Norman Group, 2017). How the data are displayed, so it brings meaning to people, is also an important element of usability (Masie, 2013). After finding, accessing, and using the portal and the user derives a data-driven decision, the decision could be skewed based on the quality of the data.

Data Quality

An extremely important concept for consideration in direct relation to accessibility and usability is the quality of the data. “Ensuring data quality is widely recognized as a valid and important activity” (Tayi and Ballou, 1998, p. 54). Once a user finds the data, are they able to use it? Are the data what they need? An early IBM engineer and programmer by the name of George Fueschel is credited with coining the phrase, “Garbage in-garbage out” (GIGO) (Clark, 2009). Essentially this phrase refers to a computer program only processes what is put into it. If incorrect data are input into a computational program, then questionable results are the output of the program calculations. This outcome underscores the necessity of high-quality data. According to Strong, Lee, and Wang (1997), a simplistic definition of high-quality data is its usefulness and usability, *i.e.*, “fit for use” by data consumers and researchers. They go on to comment that some users may find the data to be exactly what they need, while other users might find the same data to be completely useless.

Summary

Researchers and other educational stakeholders often rely on computer programmers or other technical professionals such as data scientists to overcome the barriers between them and the data and data portals they need to use (Woolf, 2010). The reason for this is that the data portals are either not findable, accessible, nor usable for the majority of the stakeholders without complex technical expertise needed to use them. This points to the importance of the data portal’s usability and accessibility.

“Schools face a situation where they need to improve the overall accountability of their operations. In an environment of considerable public, media, and policymaker scrutiny and scarce resources, educational institutions must get better at data collection, record-keeping, analysis, and reporting (West, 2012, pg. 9).” As ever-growing sets of data are collected in the educational field, and as databases and data portals become more sophisticated to house the datasets, designers and programmers of data portals must remember the variability of the skills of stakeholder users and design data portals that have findability, accessibility and usability. To illustrate the journey a stakeholder will embark upon to derive a data-driven decision beginning with the data within a data portal, please see Figure 1 below.

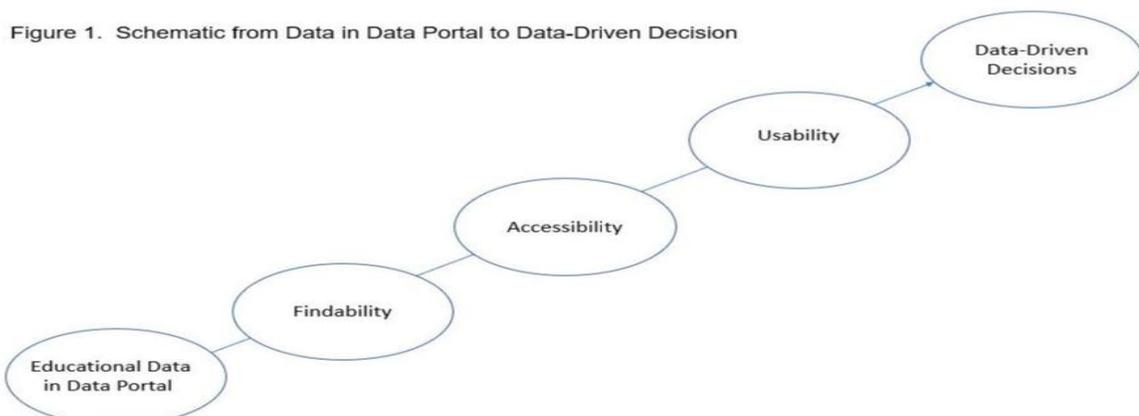


Figure 1. Starting with the Educational Data in a Data Portal, an education stakeholder will go through the stages of *findability*, *accessibility*, and *usability* in order to arrive at a Data-Driven Decision

To put the above illustration into a corollary format, and to summarize the essence of the information presented beforehand in this chapter:

1. In order for a data portal to be of use, those who need to use it, must be able to find it and the data within it, *i.e.*, *findability* (Wilkie & Azzopardi, 2013);
2. In order for a data portal to be of use, those who need to use it, must be able to get to it and the data within it, *i.e.*, *accessibility* (Smith, 1994); and,

3. In order for a data portal to be of use, those who need to use it, must be able to use it and the data within it, i.e., *usability* (Berzal, Cubero, Marin, Serrano, & Blanco, 2003). *Next Steps: Visualizations*

It is to be acknowledged that currently data analytics are already being used to “predict which students are likely to become disengaged or drop out of school, then to intervene accordingly, among other purposes” (Herold, 2016, para. 20). It also needs to be acknowledged that the business sector of the education field already collects large streams of data from students’ use of their digital educational resources. The purpose of this data collection is to determine how students interact with the digital content. Unfortunately, many schools have a myriad of different databases that do not connect with one another (West, 2012). Attendance is housed in one portal, academic assessment is housed in another portal, if discipline records are kept digitally, there is a good likelihood that those are kept in another portal, none of which communicate with one another. Creation of data sharing networks is crucial. In addition, consistent terminology or coding need to be developed so that the resulting information is easily understood. (West, 2012). It additionally needs to be acknowledged that there is a preponderance amount of data being collected in education. A need for simplification becomes a necessity in order for the data to be meaningful. There is also a certain level of interpretation depending on what you leave in, what you leave out, or what you fix - known as ‘cleaning the data’ (Communications of the ACM, 2014, pg. 19). To clean the data, first and foremost a user must be able to access and use the data in a simplified manner. Simplified data results do not necessarily mean charts, graphs, and spreadsheets. There are already individuals providing easy-to-use interfaces that create data visualizations such as the Facebook Timeline (ACM, 2014). Imagine the future of data analytics with data visualization tools accessible and usable to a wide audience of variable skilled users.

References

- Asamoah, D.A., Sharda, R., Zadeh, A.H., & Kalgotra, P. (2017, April). Preparing a data scientist: A pedagogic experience in designing a big data analytics course. *Decision Sciences Journal of Innovative Education*, 15(2), 161-190.
- Berzal, F., Cubero, J., Marin, N., Serrano, J., & Blanco, I. (2003). *Usability issues in data mining systems*. Paper presented at the Fifth International Conference on Enterprise Information Systems, Angers, France. Paper retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.96.4640&rep=rep1&type=pdf>
- d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication and Society*, 15(5), 662– 679. <https://doi.org/10.1080/1369118X.2012.678878>
- CACM Staff. (2014, June). Visualizations make big data meaningful. *Communications of the ACM*, 57(8), 19-20.
- Chiappe, A. & Rodriguez, L. P. (2017). Learning analytics in 21st century education: A review. *Ensaio: Avaliação e Políticas Públicas em Educação*, 25(97), 971-991.
- Clark, T. (2009, February 17). *Fuzzing for software vulnerability discovery* (Technical Report No. RHUL-MA-2009-04). Surrey, England: University of London. Retrieved from http://bxi.es/Reversing-Exploiting/Fuzzing_for_software_vulnerability_discovery.pdf
- Columbus, L. (2014). Where big data jobs will be in 2015. *Forbes*. Retrieved from <http://www.forbes.com/sites/louiscolombus/2014/12/29/where-big-data-jobs-will-be-in-2015/>.
- Deltor, B. (2000). The corporate portal as information infrastructure: towards a framework for portal design. *International Journal of Information Management*, 20(2), pp. 91-101. Retrieved from [https://doi.org/10.1016/S0268-4012\(99\)00058-4](https://doi.org/10.1016/S0268-4012(99)00058-4)
- Dunlap-Kahren, A. (2015). K-12 schools say personalized learning is priority #1.[Website]. Retrieved from <http://www.erepublic.com/blog/K12-schools-say-personalized-learning-is-number-1-priority.html>
- Fan, S. & Truong, S. (2001). *U.S. Patent No. 6,219,706 B1*. Washington, DC: U.S. Patent and Trademark Office.
- Henry, R. & Venkatraman, S. (2015). Big data analytics the next big learning opportunity. *Journal of Management Information and Decision Sciences*, 18(2), 17-29.
- Henry, S. L. (2006). Understanding web accessibility. In C Mills (Ed.). *Web accessibility: Web Standards and Regulatory Compliance* (pp. 1-52). New York, NY: Springer-Verlag.
- Herold, B. (2016). The future of big data and analytics in k-12 education. *Education Week*. Retrieved from <http://www.edweek.org/ew/articles/2016/01/13/the-future-of-big-data-and-analytics.html>
- Khan, B. (2017). e-Learning Framework and Models. [Website]. Retrieved from *Market Insights for state & local government and education*. e.Republic <http://www.erepublic.com/blog/K12-schools-say-personalized-learning-is-number-1-priority.html>http://asianvu.com/bk/framework/?page_id=171

- Kim, M., Zimmermann, T., DeLine, R., & Begel, A. (2016). *The emerging role of data scientists on software development teams*. Paper presented at the 38th ACM International Conference on Software Engineering, Austin, TX, USA. DOI: <http://dx.doi.org/10.1145/2884781.2884783>
- King, G. (2016). Preface: Big data is not about the data! In R. M. Alvarez (Ed), *Computational Social Science: Discovery and Prediction* (pp. vii-x). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CB0978131625730.001>
- Klašnja-Miličević, A., Ivanović, M., & Budimac, Z. (2017, June 6). Data science in education: Big data and learning analytics. *Computer Applications in Engineering Education*, 25(6),1066-1078. Retrieved from <https://doi.org/10.1002/cae.21844>
- Kumar, V. & Chadha, A. (2011, March). An empirical study of the applications of data mining techniques in higher education. *International Journal of Advanced Computer Science and Applications*, 2(3), 80-84.
- Leetaru, K. (2017). A case study in big data and the replication crisis. *Forbes*. [Website]. Retrieved from <https://www.forbes.com/sites/kalevleetaru/2017/09/01/a-case-study-in-big-data-and-the-replication-crisis/#6b7f14e55105>
- Masie, E. (2013, December). The uncharted territory of big learning data. *T-D*. [Website]. Retrieved from Association for Talent Development <https://www.td.org/Publications/Magazines/TD/TD-Archive/2013/12/The-Uncharted-Territory-of-Big-Learning-Data>
- Mills, F. J., Chudoba, K. M., & Olsen D. H. (2016). IS programs responding to industry demands for data scientists: A comparison between 2011-2016. *Journal of Information Systems Education*, 27(2), 131-140.
- Morville, P. & Sullenger, P. (2010). Ambient findability: Libraries, serials, and the internet of things. *The Serials Librarian*, 58, 33-38. Retrieved from <https://doi.org/10.1080/03615261003622999>
- Nielsen, J. (1995). Severity ratings for usability problems. [Website]. Retrieved from Nielsen Norman Group <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>
- Nielsen, J. (2011). Top 10 mistakes in web design. [Website]. Retrieved from Nielsen Norman Group <https://www.nngroup.com/articles/top-10-mistakes-web-design/?lm=strategic-design-faqs&pt=report>
- Nielsen, J. (2012). Usability 101: Introduction to usability. [Website]. Retrieved from Nielsen Norman Group <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- Power, D. & Hermancinski, L. (2013). Big data and analytics. *MASSTLC*. Retrieved from <http://www.masstlc.org/?page=2013BigDataReport>. Rehabilitation Act of 1973, Pub. L, No. 93-112, 29 U. S. C. § 701.
- Romero, C.& Ventura, S. (2010, November). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 40(6), 601-618.
- Rolleston, R., Howe, R. & Sprague, M. A. (2015). Educational reports that scale across users and data. *Proceedings of the 8th International Conference on Educational Data Mining, Madrid, Spain*, 468-471. Retrieved from http://www.educationaldatamining.org/EDM2015/proceedings/edm2015_coverpages.pdf
- Rubin, J. & Chisnell. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests*. Indianapolis, IN: Wiley Publishing, Inc.
- Siguenza-Guzman, L., Saquicela, V., Avila-Ordóñez, E., Vandewalle, J., & Cattrysse, D. (2015). Literature review of data mining applications in academic libraries. *The Journal of Academic Librarianship*, 41, 499-510.
- Slater, S., Joksimovic, S., Kovanovic, V., Baker, R., & Gasevic, D. (2017). Tools for educational data mining: A review. *Journal of Educational and Behavioral Statistics*, 42(1), 85-106. DOI: 10.3102/1076998616666808
- Smith, G. D. (1994, June 11). Increasing the accessibility of data. *BMJ*, 38, 1519-1520.
- Smith, S. (2001). *U.S. Patent No. 6,233,618*. Washington, DC: U.S. Patent and Trademark Office.
- Spector, J. M. (2013). Emerging Educational Technologies and Research Directions. *Educational Technology & Society*, 16(2), 21-30.
- Strawn, G. (2016, May 25). Data scientist. *IT Professional*, 18(3), 55-57. DOI: 10.1109/MITP.2016.41
- Strong, D. M., Lee, Y. W. & Wang R. Y. (1997). Data quality in context. *Communications of the ACM*, 40(5), 103-110.
- Sullivan, D., (2004). *Proven portals: Best practices for planning, designing, and developing enterprise portals*. Boston, MA: Addison-Wesley.
- Tatnall, A. (2005). *Web portals: The new gateways to internet information and services*. Hersey, PA: Idea Group Publishing.
- Tayi, G.K. & Ballou, D.P. (1998, February). Examining data quality. *Communications of the ACM*, 41(2), pp.54-57. DOI: 10.1145/269012.269021
- U.S. Department of Education, Privacy Technical Assistance Center. (2017, January). *Integrated data systems and student privacy*. Retrieved from https://studentprivacy.ed.gov/sites/default/files/resource_document/file/IDS-Final_0.pdf

- U.S. Department of Health and Human Services. (2017, October). *Usability & evaluation basics*. [Website]. Retrieved from <https://www.usability.gov/what-and-why/index.html>
- U.S. Department of Health and Human Services. (2018, March 2). *What & why of usability* [Website]. Retrieved from <https://www.usability.gov/what-and-why/index.html>
- U. S. Department of Justice. (2003, June). *Accessibility of state and local government websites to people with disabilities*. [Website]. Retrieved from Civil Rights Division, Disability Rights Section: <https://www.ada.gov/websites2.htm>
- Wang, Y., (2017). Education policy research in the big data era: Methodological frontiers, misconceptions, and challenges. *Education Policy Analysis Archives*, 24(94), 1-21.
- West, D. M. (2012, September). Big data for education: Data mining, data analytics, and Web dashboards. *Governance Studies at Brookings*. Washington, D.C.: The Brookings Institute. Retrieved from <http://www.insidepolitics.org/brookingsreports/education%20big%20data.pdf>
- Wilkie, C. & Azzopardi, L. (2013). An initial investigation on the relationship between usage and findability. In P. Serdyukov, P. Braslavski, S. O. Kuznetsov, S. Ruger, E. Agichtein, & I. Segalovich, (Eds.), *Lecture Notes in Computer Science: Vol.7814. Advances in Information Retrieval* (pp. 808-811). doi: 10.1007/978-3-642-36973-5
- Wimmer, M. A. & Holler, U. (2003). Applying a holistic approach to develop user-friendly, customer-oriented e-government portal interfaces. In N. Carbonell & C. Stephanidis (Eds.), *User interfaces for all* (167-178). Berlin, Germany: Springer-Verlag.
- Woolf, B. P. (Ed.). (2010). *A roadmap for educational technology*. Retrieved from Computing Community Consortium: <http://cra.org/ccc/wp-content/uploads/sites/2/2015/08/GROE-Roadmap-for-Education-Technology-Final-Report.pdf>
- Zezula, P. (2014). Similarity searching for the big data: Challenges and research objectives. *Mobile Network Applications*, 20, 487-496. Retrieved from DOI: 10.1007/s11036-014-0547-2