

**EXAMINING THE APPLICATION OF BIG DATA ANALYTICS TO INCREASE UNIVERSITY  
RETENTION AND PROMOTE STUDENT SUCCESS**

*Jody Strausser, Robert Morris University/Clarion University, jstrausser@clarion.edu*

**ABSTRACT**

*One of the driving forces for using big data analytics in education is the need to improve the retention rate among college students. Much of the student retention effort has been focused on providing support for students in their freshman year; however, the Educational Advisory Board (EAB) describes a series of students who successfully completed their freshman year with grade point averages between 2.0 and 3.0 but ultimately do not wind up graduating. These students perform well enough to not raise any flags, but often do not complete their major studies. Big Data tools can apply historical data and predictive analytics to educators in assisting students in completing their majors. As big data usage continues to evolve, it is hoped that historical data and predictive analytics can better support student success.*

**Keywords:** Analytics, Big Data, University Retention, Student Success

**INTRODUCTION**

The phrase “big data” has become an increasingly important term in the vernacular of many businesses and organizations. Big data is broadly defined as “an accumulation of data that is too large and complex for processing by traditional database management tools” [2]. The use of big data continues to influence the educational community. One of the driving forces for using big data analytics in education is the need to improve the retention rate of college students. This article addresses the retention rate issue, provides a history of big data, examines the analytic methodologies, and provides a short case study of the Educational Advisory Board’s Student Success Collaborative at Clarion University of Pennsylvania. By understanding the importance of big data and data analytics in education, it is hoped that more institutions will adopt technologies in order to better support their students.

**BACKGROUND**

**The Need to Improve Retention Rates**

Many colleges and universities are facing continued financial difficulties ranging from a smaller pool of high school applicants to decreased state funding. Colleges and universities are thus being evaluated using business success and profitability measures [24]. While recruitment continues to be a driving force for universities, many are addressing financial shortfalls by reexamining retention rates. Retaining students is not a new problem. According to Braxton, McKinney and Reynolds [4], over 75 years of research has been devoted to this subject with little change in dropout rates. Many universities simply are not engaged in campus student retention research. A 2009 College Board study discovered that “campus resources for initiatives aimed at increasing student persistence were minimal and inadequate and that the vast majority of retention coordinators on college campuses were given little or no authority to implement new program initiatives” [11]. Most of the higher educational institutions who responded indicated they were “engaged in assessments of the college environment” which very often includes items such as extracurricular activities and could even include campus housing or even the quality of the food. It is essential that colleges and universities look at their own populations and examine what can be done to increase retention.

Research by Cuseo and Farnum [11] have prescribed best practices by refuting a series of retention myths that create blocks to enhancing student learning and success. The first myth is that retention or attrition is a student problem, not a college or university issue. The authors indicate that student success must be based upon both what the institution does for students as well as what students can do for themselves. Institutions should provide clear lines of communication, allow students to participate in decisions, when applicable, and have student experiences with administration that assist in student development. A second myth addresses the common misconception that colleges and universities should just admit better prepared students. Again, the authors refute this myth by indicating academically well prepared students from college educated families are decreasing. Rather, low income

and first generation college students are on a rapid increase and colleges and universities should be prepared to support these students. Personal issues such as emotional inconsistencies, balancing work and school, and the social impacts of the college environment all play an important role in student dropout rates, as well. Institutions should have well established programs and counselors to assist students, but this is not saying it should ignore academic progress either. There is an ongoing final myth that it is not the faculty's job to retain students but to promote student learning. Studies [1] have shown that students who are actively engaged in the learning process and find personal meaning in their academic careers are far more likely to remain at a college or university. This does not imply that faculty should be concerned about being well liked or make courses easy in order to gain positive faculty evaluations. Academic rigor remains important in overall student success [11].

From the data above it is clear there are many areas to address with examining college and university retention rates. Many colleges and universities simply are not addressing any of the issues. According the Jones and Braxton, a survey of 54 college and university respondents indicated that only 33% were conducting analysis of more than one statistical outcome variable at a time [19]. More multivariable studies need to be performed, but starting with even one area, most likely academics, has great potential to show improvement.

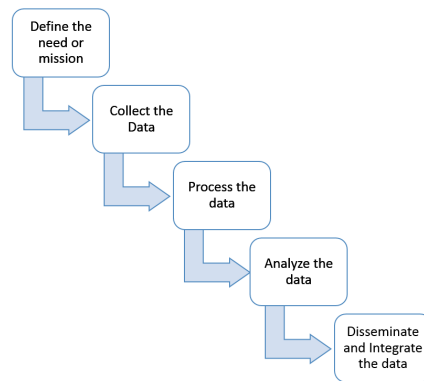
### **Big Data and Analytics**

The National Aeronautics and Space Administration (NASA) first defined the term "Big Data" in 1997 from an issue with computer graphics. NASA was working on a visualization project in which the data sets "provides an interesting challenge for computer systems: data sets are generally quite large, taxing the capacities of main memory, local disk, and even remote disk. We call this the problem of big data. When data sets do not fit in main memory (in core), or when they do not fit even on local disk, the most common solution is to acquire more resources" [10]. In 2008, computer scientists Bryant, Katz, and Lazowska [5] popularized the term by indicating that big data computing was going to revolutionize the computing industry and impact organizations. In their article, Bryant, Katz and Lazowska [5] indicate a number of organizations were collecting huge amounts of data beginning in the late 1990s. Wal-Mart was in the process of working with Hewlett-Packard to develop a data warehouse capable of storing over 4 petabytes of data (4000 trillion bytes) in order to make decisions regarding pricing strategies, advertising campaigns, and inventory. Scientists were gathering over 30 trillion bytes of image data from the Large Synoptic Survey Telescope (LSST) per day. The huge amount of data from diagnostic equipment and other sources was assisting medical researchers in diagnosing patients. Today (2015), the term big data has taken on multiple meanings, but Merriam-Webster defines big data as "an accumulation of data that is too large and complex for processing by traditional database management tools" [2]. The methods of storing data, in house or via "the cloud" are becoming company independent, but the amount of data being collected from many sources continues to grow exponentially [17].

Big data lends itself to the idea of business analytics or simply analytics. Often the term business intelligence would be used here, but because this paper addresses the analytics behind student success in higher education, the term analytics is more appropriate. Although MacNeill, Campbell, and Hawskey [20] indicate a broad lack of consensus over the term, like the previously stated authors, Cooper [9] develops a reasonable definition that will adopted in this paper. Cooper's definition of analytics states, "analytics is the process of developing actionable insights through problem definition and the application of statistical models and analysis against existing and/or simulated future data" [9].

Because analytics have been associated with addressing business problems with data, turning that data into insights and providing a business with knowledge to assist in the decision making process in order to gain a competitive advantage, it makes sense to turn to industry to examine the business analytics process.

While each organization develops a business analytics process to meet its needs and resources, according to Shapira [21], there are five steps which are considered critical to the analytics process. Figure 1 shows the steps that are used to examine the business analytical process [23].



**Figure 1.** The Intelligence Process

The following steps summarize the process shown in Figure 1.

Step 1: Define the need or mission. Typically this involves defining a gap or desire that will assist an institution in achieving its goal. It is necessary to define what data is required to achieve the goal of the mission and is sometimes broken down into smaller objectives [23].

Step 2: Collect the data. This step should involve a member of management who will need to revise data collection methodologies as time progresses. In the data collection method, it is essential that the data will assist in answering the mission statement or business problem or decision [23].

Step 3: Process the data. In this phase, raw collected data is converted into forms that can be used by decision makers at all levels. This process is often called cleaning the data which includes activities such as removing outliers and combining variables to form new variables. This can be a relatively time intensive process if done on its own; tools are generally used to assist in the processing of data [23].

Step 4: Analyze the data. During the analysis phase, intelligence is gathered from the processed data. Depending upon the environment, predictive analysis or what-if scenarios are often run so a comparison can take place with actual results. This process is often used to verify the models being used in the analysis. Techniques include using statistical methods such as regression testing and correlation analysis in order to determine which factors in the data best address the mission or business question. An analyst will then select the optimal solution based on lowest error and that which address the mission most effectively [23].

Step 5: Disseminate and Integrate the data. In this phase, the results of the analysis are transmitted to the decision maker. All limitations of the analysis must be provided to ensure the decision maker considers all factors. While recommendations may be made regarding the optimal solution, decision makers are often provided alternative solutions for their own evaluation purposes. Finally, evaluate the decision and determine if the decision was effective. Some business analytics models consider this a separate step, but following the intelligence study, it is included in this step because decision outcomes can often be rapidly evaluated in the intelligence field [23].

### **Using Big Data and Analytics to Increase Student Success and Retention**

The awareness of “big data” and analytics for education has increased over the past several years, mainly due to decreased funding and a poor economic climate, as discussed above. Colleges and universities are known for collecting large amounts of data including graduation rates and student survey results. Analyzing this data in order to make institutional decisions and changes has not been very successful. According to MacNeill, Campbell and Hawksey, “until recently, few institutions have been able to explore the wealth of information they routinely collect through the core business of teaching and learning” [20].

Higher education has integrated big data analytics into its processes through three cycles. During the first cycle, Learning Management systems such as Blackboard [3], were incorporated to support online learning, hybrid learning (partial online and partial face to face meetings), and to support traditional face to face classes. In the second phase, Web 2.0 technologies such as social networking, video transmission, and interactive simulations were

used. In the third phase analytics are now being incorporated. The term learning analytics “focuses on collecting and analyzing data from a variety of source to provide information on what works (and what doesn’t) with respect to teaching and learning” [18].

According to Buckingham Shum [6], there are three layers to learning analytics, the macro, micro, and meso levels. The macro level analytics enable sharing of data across a variety of educational institutions and are often used to define best practices. The meso level of analytics is used at the institutional level and generally guide business, or in this case, educational, decision making. The mirco level of analytics is used at the process level in businesses but can be related the learner level. The micro level would include items such as dashboards integrated into learning management systems such as Blackboard Learn [3]. According to MacNeill, Campbell and Hawksey [20], these tools have to potential to provide insight into the educational process for both the micro and meso levels. At the micro level, there is the expectation that such tools will provide rapid feedback to students, thus quickly identifying gaps in the knowledge base.

Higher levels of graduates increase the reputation of the institution and are expected to increase enrollments at the same time. Higher education is taking notes from a variety of analytic tools from Netflix, Amazon, and Pandora. The higher educational goal is to use familiar features of the tools such as recommendation systems into advising and registration systems. Tistan Denley [13], a member of the Tennessee Board of Regents, started research in 2012 to develop a tool, now called Degree Compass, which allows students and advisors the same suggestion capabilities as companies such as Netflix. “The model combines hundreds of thousands of past students grades with each particular student’s transcript to make individualized recommendations for each student” [13]. Instead of relying on previous classes taken by the student, this system uses predictive analytics based upon grades and rank courses according to factors that will assist the student in his or progress towards their selected major completion. An important factor of this product is its ability to interface with the degree auditing and associating it with available classes in the following semester in addition to making a decision based upon potential success of the individual student. One risk of introducing new system involves technology overload. New systems do not integrate very well or even at all into current systems and often duplicate information. Although, at the time of publication of this paper, it was early to make generalizations about the use of the Degree Compass program, the author indicated graduation rates increased from 33% to 37.4% in the spring of 2011 [13].

Additional examples of big data analytics in education include Santa Monica College’s Glass Classroom initiative where, at the macro level, student performance is evaluated by collecting large amounts of data. Based upon real-time assessment of a student’s performance, the courseware can adjust itself to provide more examples or practice to ensure the learning objectives are met. At the University of Wisconsin-Madison, engineers were developing an early warning system for advisors to alert them about potential risks for the individual student. “The system aims to intervene early, improve students’ academic success and bolster the campus’s retention and graduation rates” [18]. These analytic big data tools are part of an ever growing number of tools available to educators at both the learning management system and institutional levels.

## **BIG DATA ANALYTICS FOR STUDENT SUCCESS**

### **Big Data at Clarion University for Student Success**

In 2012, Clarion University became one of 27 institutions which joined the Student Success Collaborative’s founding cohort, now part of the Educational Advisory Board (EAB). In 2014, the university adopted and launched the Educational Advisory Board – Student Success Collaborative (EAB-SSC) tool which is “a “big data” initiative that analyzes the past decade of Clarion students’ performance. The SSC platform may be used to find pathways and support services that enable a student to be successful” [7].

The EAB-SSC tool costs vary depending upon the desired capabilities. Columbia College in Chicago, IL indicated an approximate cost of \$100,000 per year, which is a relatively low cost for technology according to the Columbia College vice president and provost, Stan Wearden [8]. Other universities such as Georgia State indicate costs of around \$70,000 per year [12] and Central Washington University indicated a yearly cost of \$95,000 per year [14]. The tool itself is web-based and can be hosted on or off site. In the case of Clarion University, the user logs on through a local system web interface and is then taken to the SSC.

Data regarding the EAB-SSC tool's success or failure at Clarion University is not available as of the writing of this article. The author has added anecdotal evidence of personal use at the end of the article. This evidence should not be interpreted as scientific or representative of the university.

Because universities lose most students during the first year, most of the college and university resources are devoted to freshmen. According to Tyson [22] and the Educational Advisory Board (EAB) [15], 45% of students finish their first year with a grade point average between 2.0 and 3.0, yet may still eventually drop out of college. EAB refers to this group as the "murky middle." These students generally spend between 4.5 and 5.7 semesters at the university. EAB indicates that some of these students will stay and some will leave regardless of any intervention. An advisor's intervention could provide enough support in order to retain those students who are considering exiting the university.

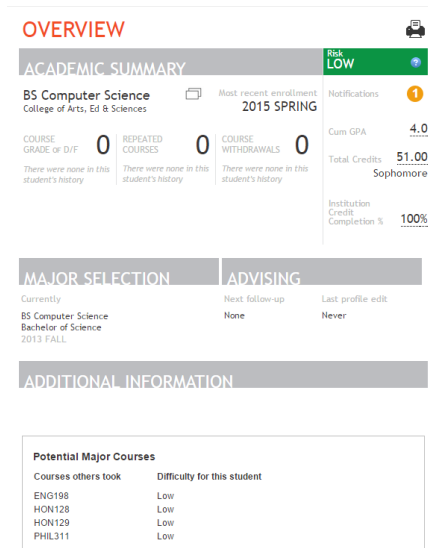
It is suggested that colleges might allocate some of the resources to the murky middle group. Shifting resources must be completed with care as experts, such as Alan Seidman, executive director for the Center for Study of College Student Retention, argue that students lacking in foundational expertise for a particular course are most likely to fail the course so early intervention (freshman year) is necessary [22]. Regardless from where the support comes, any intervention by advisors or instructors has the possibility of increasing the student's success, simply by demonstrating that student support exists.

The SSC tool was developed based upon the Educational Advisory Boards research into the retention of at-risk students. EAB conducted a study [15] that identifies obstacles for student success in higher educational institutions. Several of the top lessons from the study are provided here as they are incorporated into the resulting technology. The first lesson learned deals with collegiate processes. Conflicting or confusing processes in other areas of the institution could overwhelm at risk students and take time away from the true intent of an advising session with a student. For example, time that should be spent advising a student in terms of course paths and current difficulties is spent talking about the financial aid process or the various signatures required to gain entry into a class. The second lesson addresses the need for more faculty student interaction. In many cases, students are often allowed to withdraw from a class without being advised or fully understanding the consequences of this action. A withdrawal from a prerequisite course in a hierarchical program structure could delay the student's graduation for at least semester, if not more. These rigid programs can influence a student to transfer to another college or university.

To address some of the issues listed above, EAB suggests a concept called *choice architecture*. "Students choices are made within contexts that bias them toward one or more particular options (whether those contexts were designed intentionally or not). By rethinking those context and providing students with better information, stronger recommendations, and smarter options, intuitions can mitigate many of the navigational roadblocks mentioned above" [15]. By using technology to support basic advising and flagging at risk students, ensuring students and advisors are not distracted by confusing processes, and preventing delays in degree progress, an investment in even basic advising assistance can increase student retention.

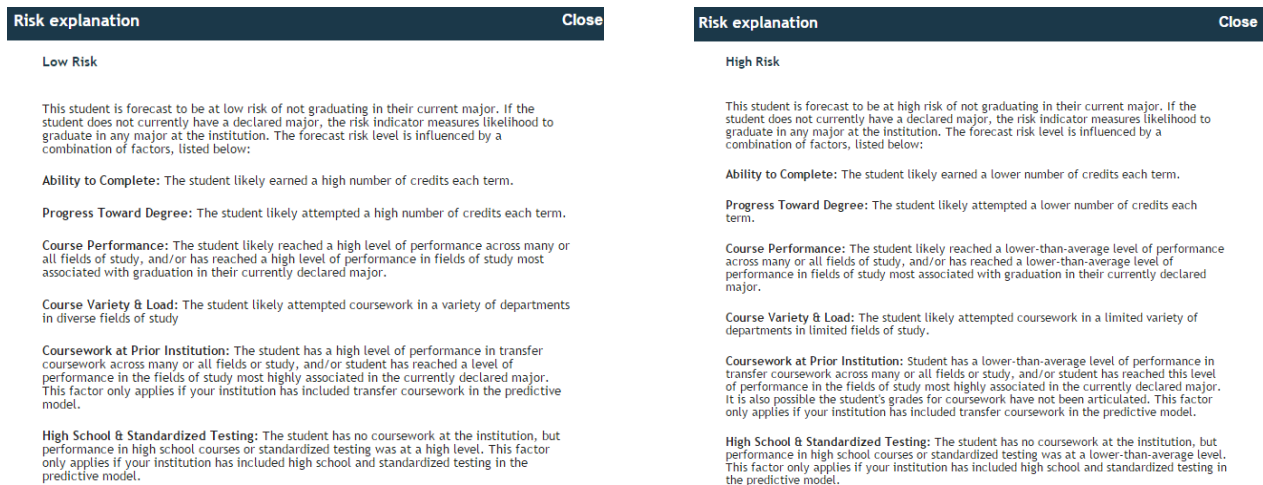
The Educational Advisory Board - Student Success Collaborative (EAB-SSC) tool assists the advisor in several ways. The first is what EAB calls the 30 second "gut check" student profile. While the student profile may contain a similar information to that in a Student Information System (SIS), it differs by proactively identifying students at risk of not graduating in their major and providing the tools to assist the student in his or her academic progression. GPA is the common mechanism in which a student's risk is often addressed. The EAB-SSC tool uses big data to incorporate a "full picture" of the student and compares it against historical data and then predicts the student's likelihood of graduating in that program. A sample student profile is shown in Figure. The academic summary quickly provides the advisor with the student's major, a color indicator risk level, current GPA, number of credits taken as well as a list of potential courses this student may take in the program. Advisors can quickly determine in which courses the students received a D or F, which courses were retaken, and from which courses the student withdrew. If the faculty is the student's advisor, he or she can determine when the profile was last visited and when the next follow up should occur. These alerts assist the advisor in keeping a high level of interaction with the student. In 30 seconds an advisor can see the student's major, when he or she was last enrolled and a color coded risk. By gathering this overview, the advisor can quickly begin to talk with the student.





**Figure 2.** Student Overview

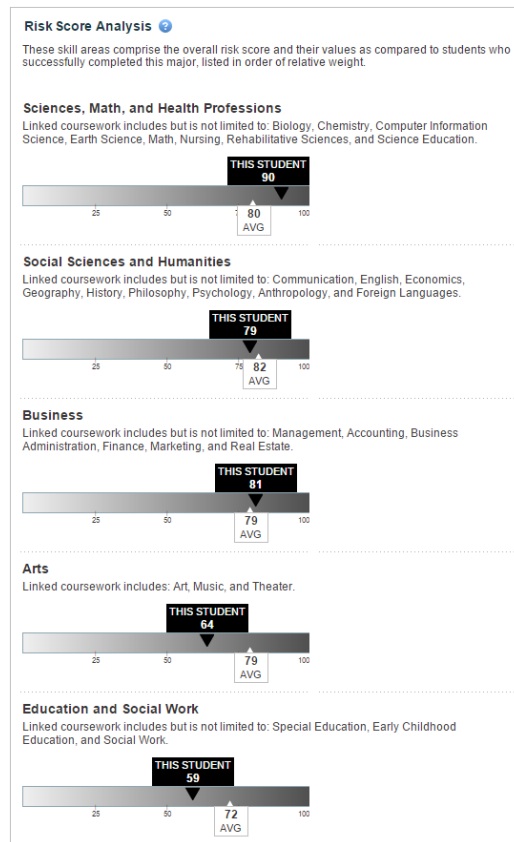
In Figure 2, there are two risk assessment examples provided. The left example is for the student overview in Figure. The right hand example is a risk assessment provided by a student at high risk. The assessments are gathered by clicking on the question mark in the Risk Color in the upper right hand corner of Figure 2. By doing so, the advisor can gain a better understanding of how the risk was assessed. The risk analysis uses the number of credits, performance in current courses, and course variety and load to determine the level of risk for the student.



**Figure 2.** Risk Analysis Explanation

In the student overview, a Risk Score Analysis is presented to the advisor shows how the student compares to students who successfully completed the major. Looking at risk area performance will also help the advisor understand why the student is or is not at high risk of not completing the major. The highest importance of a skill area for major is indicated first. In the example in Figure 3, below the skill name describes a collection of courses in which students' performance tends to correlate. The skills are also color coded. If an area shows red, it is usually because the student's performance has a high importance to the major and the student has performed below the historical average for successful completion of the major. The risk score analysis also provides suggestions for major completion. For example, course grades in mathematics and science are correlated because both require intensive quantitative skills. The student's performance is measured on a scale of 0 to 100. The historical average of

student performance in a particular skill area is provided. The advisee's average of classes taken thus far in the area of study is also provided. A student who performs below average in a particular area is at higher risk of not completing the major. In the example in Figure 4 below, the student is performing in the 90<sup>th</sup> percentile which means that the student is performing better than 90% of the students who declared the major and is at a low risk of completing the major. In the same example provided in Figure 4, the student is performing a bit below the average in the Arts. While it is essential that a student develop in all areas in a liberal arts education, the risk of completing the major is not as significant as the Arts are not as strongly associated with, in this case, a Computer Science major which is indicated by its position in the list.

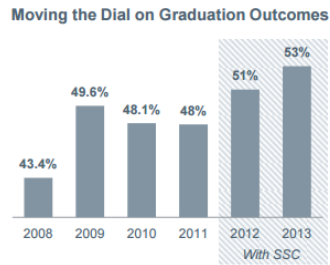


**Figure 3.** Sample Student Risk Analysis

There are many other tools that assist the advisor in promoting student success, including lack of registration notifications, credit accumulation and cumulative GPA, however as the topic of this article addresses the use of big data in student success. Interested readers are encouraged to access the Educational Advisory Board's website further details of the Student Success Collaboration tool [15].

As stated previously, the recent adoption of this tool combined with training faculty and staff to use it, has prevented any specific conclusions to be made at this time. Due to the proprietary nature of the software, access to internal equations and predictive analytics was impossible. Future work would include statistical analysis of the predictive capabilities and actual retention as the data becomes available.

There have been a number of case studies published on the EAB website which indicate the success of this product. For example, at Georgia State University, there was a three percent increase from 2012 to 2014. Advisors at the university credit, in part, the use of the SSC on daily basis to reduce time to degree and lower the risk to students. This increase in graduation rate, create an estimated three million dollars in revenue in 2014 [16]. Figure 5 shows the graduation rate increases with the use of the SSC at Georgia State University. In 2012, there was a three percent increase with the use of SSC and an additional two percent in 2013.



**Figure 5.** Graduation Outcomes at Georgia State University

A Georgia State Advisor is quoted as saying “SSC allows us to be hugely supportive of any student. We can encourage students that are on path with data; we can provide a visual to students who are off path; or if a student is right in the middle, we can teach them about the hill they have to climb” [16]. Finally, the case study indicates an approximate four million dollars in savings by students in the graduating class of 2014 by advisor interaction, thus allowing them to graduate on time.

#### **Author Comments on the Tool Usage**

The author has been using the tool over the past year with some success. In one particular instance, a student had registered for the Fall 2015 semester without coming to an advising session with me. When I was reviewing the student’s schedule I performed a comparison with the EAB-SSC tool and quickly realized that the student had not registered for a required prerequisite class which would have resulted in requiring an extra semester. This tool is particularly important as some required classes in the current curriculum are only offered every other year. If a student misses the class, there is a high potential that the student will have to enroll an extra semester. This tool has provided the capability to capture and address this issue.

I have not been as impressed with the predictive analytics. Many times the system will make predictive risks for a particular major that just do not make a whole lot of sense. For example, I had a student who was struggling in a Calculus course and the tool suggested that he might not be the best fit for a computer science degree. After review his grades in the computer science courses and talking to other faculty, I determined that the student just needed support in that one area; he was an exceptional at developing algorithms and programming. Finally, as of the writing of this document, the software does not necessarily assist in the advising of freshmen as it does not take into account any historical data regarding high school performance so faculty who advise freshmen are sometimes frustrated with the tool. It has been suggested that additional data will be incorporated into the model in the future. The EAB-SSC is considering the addition of non-articulated transfer data, financial data, learning management systems data and student engagement data to refine the predictive capabilities of their model [15].

### **CONCLUSIONS**

In this article, the use of big data was examined in its ability to promote student success. Big data is defined as “an accumulation of data that is too large and complex for processing by traditional database management tools” [2]. The use of big data continues to influence the educational community. One of the driving forces for using big data analytics is the need to improve the retention rate of college students. There are a number of ways in which big data can assist in student success at the macro, micro, and meso levels. Much of the effort in the past was focused on providing support for students in their freshman year; however, the Educational Advisory Board [15] describes a series of students who successfully completed their freshman year with grade point averages between 2.0 and 3.0 but ultimately do not wind up graduating. These students perform well enough to not raise any flags, but often do not complete their major studies. Big Data tools can be used to use historical data and predictive analytics to assist advisors and administrators in assisting students in completing their majors. As analytic capabilities expand, it is hoped that student success rates can continue. Further research is suggested in analyzing the predictive capabilities of a variety of tools and determine how they can influence student success.



**REFERENCES**

1. Astin, A.W. *What matters in college*. 1993, San Francisco: Jossey-Bass.
2. big data. 2015 [cited 2015 April 4]; Available from: [http://www.merriam-webster.com/dictionary/big data](http://www.merriam-webster.com/dictionary/big%20data).
3. Blackboard Inc. 2015 [cited 2015 April 3]; Available from: <http://www.blackboard.com/>
4. Braxton, J.M., J.S. McKinney, and P.J. Reynolds, *Cataloging Institutional Efforts to Understand and Reduce College Student Departure*. New Directions for Institutional Research, 2006. 2006(130): p. 25-32.
5. Bryant, R., R.H. Katz, and E.D. Lazowska, *Big-data computing: creating revolutionary breakthroughs in commerce, science and society*. 2008, December.
6. Buckingham Shum, S. *Learning analytics*. 2012 [cited 2015 April 3]; Available from: <http://iite.unesco.org/pics/publications/en/files/3214711.pdf>
7. Clarion University of Pennsylvania. *Office of the Provost--Intranet:EAB-SSC*. 2015 [cited 2015 April 8].
8. Columbia Chronicle. *College to invest in new advising software*. 2015 [cited 2015 July 11]; Available from: [http://www.columbiachronicle.com/campus/article\\_1bc884ee-e005-11e4-977e-6f7f1646ee0a.html](http://www.columbiachronicle.com/campus/article_1bc884ee-e005-11e4-977e-6f7f1646ee0a.html)
9. Cooper, A., *What is analytics? Definition and essential characteristics*. Cetus Analytics Series, 2012. 1(5).
10. Cox, M. and D. Ellsworth, *Application-controlled demand paging for out-of-core visualization*, in *Proceedings of the 8th conference on Visualization '97*. 1997, IEEE Computer Society Press: Phoenix, Arizona, USA. p. 235-ff.
11. Cuseo, J. and T. Farnum, *Seven myths about student retention*. Teresa Farnum and Associates, Inc.(TFA), Austin, TX. Accessed, 2011. 7(7): p. 2011.
12. Davis, J., *Georgia State University tracks data to boost graduation rates*, in *McClatchy - Tribune News Service*. 2015: Washington.
13. Denley, T., *How Predictive Analytics and Choice Architecture Can Improve Student Success*. Research & Practice in Assessment, 2014. 9(2): p. 61.
14. EAB Student Success Collaborative, *Consultancy involving both data analytics and staff training business case*. 2014, Central Washington University.
15. Educational Advisory Board (EAB). *Guiding student choice to promote persistence: Tools, technologies, and policies that support retention and timely completions*. 2015; Available from: [http://www.eab.com/~media/EAB/Research-and-Insights/AAF/Studies/2015/Guiding-Student-Choice-to-Promote-Persistence/AAF\\_30226.pdf](http://www.eab.com/~media/EAB/Research-and-Insights/AAF/Studies/2015/Guiding-Student-Choice-to-Promote-Persistence/AAF_30226.pdf)
16. Educational Advisory Board (EAB) *Improving student outcomes with data-driven advising and intervention: A case study from Georgia State University*. 2014.
17. Feinleib, D., *Big data and storage: Why you can't save everything forever*. 2012.
18. Fiaidhi, J., *The Next Step for Learning Analytics*. IT Professional, 2014. 16(5): p. 4.
19. Jones, W., A and J.M. Braxton, *Institutional level catalysts and constraints: Cataloging and comparing institutional efforts to increase student retention rates*. Journal of College Student Retention: Research Theory and Practice, 2009. 11(1): p. 123-139.
20. MacNeill, S., L.M. Campbell, and M. Hawksey, *Analytics for Education*. Journal of Interactive Media in Education, 2014: p. 1-12.
21. Shapira, G. *The seven key steps of data analysis*. Big Ideas [cited 2015 April 5]; Available from: <http://www.oracle.com/us/corporate/profit/big-ideas/052313-gshapira-1951392.html>
22. Tyson, C. *The Murky Middle*. 2014 [cited 2015 April 28]; Available from: <https://www.insidehighered.com/news/2014/09/10/maximize-graduation-rates-colleges-should-focus-middle-range-students-research-shows>
23. U.S. Joint Chiefs of Staff. *Joint publication 2.0: Joint intelligence*. Available from: [http://www.dtic.mil/doctrine/new\\_pubs/jp2\\_0.pdf](http://www.dtic.mil/doctrine/new_pubs/jp2_0.pdf)
24. Wright, R.E., *Student focused marketing: impact of marketing higher education based on student data and input*. College Student Journal, 2014. 48(1): p. 88-93.