

ACADEMIC ANALYTICS

A New Tool for A New Era

By John P. Campbell, Peter B. DeBlois, and Diana G. Oblinger

In responding to internal and external pressures for accountability in higher education, especially in the areas of improved learning outcomes and student success, IT leaders may soon become critical partners with academic and student affairs. IT can help answer this call for accountability through *academic analytics*, which is emerging as a new tool for a new era.

John P. Campbell is an EDUCAUSE Learning Initiative (ELI) Scholar in Residence and Associate Vice President for Teaching and Learning Technologies at Purdue University, where he is responsible for all aspects of instructional technology. Peter B. DeBlois is Director of Programs and Media Relations at EDUCAUSE, where he coordinates member-engagement programs and media relations. Diana G. Oblinger is Vice President of EDUCAUSE, where she is responsible for the association's teaching and learning activities and for the EDUCAUSE Learning Initiative.



As populations become concerned for their financial well-being and economic security, pressures increase on those individuals and institutions that might influence the outcome.

A New Era

Although *A Nation at Risk* was released by the U.S. Department of Education in 1983, in many ways it remains an apt description of current conditions.¹ Consider the following statistics:

- Western nations are looking over their shoulders at China and India. As a popular YouTube video highlighted, if you are “one in a million” in China, there are 1,300 other people just like you; in India, there are 1,100 others just like you. The 25 percent of the Chinese population with the highest IQs is greater than the total population of North America; in India, the 28 percent of the population with the highest IQs is greater than the total population of North America.²
- Economies depend on a well-educated population. As a result, many countries are investing in strengthening their educational systems, with an emphasis on increasing the proportion of their population that has a postsecondary degree.³
- In spite of the importance of postsecondary education, the percentage of the U.S. population with a postsecondary degree has slipped relative to that of other countries. Rather than leading the world, as it has for most of the past fifty years, in 2003 the United States ranked ninth among industrialized nations in the percentage of 25-to-34-year-olds who have completed at least an associate’s degree.⁴ If current trends are left unchanged, the percentage of the U.S. workforce with a bachelor’s degree will decline from 17.1 percent (as of 2000) to 16.4 percent (in 2020).⁵
- Minorities are an increasing percentage of the population, as well as of the college-going population. For example, by 2015, Hispanics will be the

second-largest student group (by race/ethnicity), growing to 15.4 percent of the nation’s campus population.⁶ Yet the educational attainment for minorities is below that of other groups. High school graduation rates in the United States are around 70 percent but fall to 50 percent for black, Native American, and Hispanic students.⁷ Although the percentage of students who graduate from four-year institutions after six years is 57 percent overall, the rate is 41 percent for African Americans, 47 percent for Hispanics, and 39 percent for Native Americans.⁸

- Just being able to hold a job may not be enough. Each year, more jobs are outsourced to countries with lower labor costs. The good jobs that remain in countries with higher labor costs are those that require education and more refined skills. For example, the U.S. Department of Labor projects that twenty of the thirty fastest-growing jobs in the United States will require education beyond high school and that 40 percent will require at least an associate’s degree.⁹
- If the current educational gaps remain, U.S. per capita income is projected to decline 2 percent from 2000 to 2020. By contrast, per capita income has increased 41 percent in the past two decades. And trailing other developed countries on education may reduce U.S. economic growth by as much as half a percent a year.¹⁰

As populations become concerned for their financial well-being and economic security, pressures increase on those individuals and institutions that might influence the outcome. In the information age, one of the most influential institutions is education. And in an era of accountability and liability, organiza-

tions that resist pressures for results, accountability, and action are suspect. With economic security at stake, how long will society accept that the percentage of the population with a college/university education is stagnant even as the demand has risen, that retention rates have not significantly improved in decades, and that graduates may not have mastered even basic competencies?¹¹

A greater proportion of the U.S. population is being educated, but overall college/university graduation rates have remained relatively unchanged for decades. And even though the United States is holding its own in the percentage of the population with college/university degrees, other countries are moving ahead (e.g., Sweden, the United Kingdom).¹² Blaming K–12 for inadequately preparing students does not divert attention from the root problem, since K–12 teachers are the products of higher education.

But thanks to enterprise-wide systems that generate massive amounts of data, data warehouses that aggregate disparate types of data, and processing power that sifts, sorts, and surfaces patterns, academic analytics is emerging as a new tool that can address what seem like intractable challenges.

A New Tool

Analytics marries large data sets, statistical techniques, and predictive modeling. It could be thought of as the practice of mining institutional data to produce “actionable intelligence.” Just as Amazon.com knows when to send someone an e-mail notice of a new book that he/she might be interested in buying, so does an admissions office know whether to invest in the printing and postage necessary to send a high school junior a glossy campus viewbook.

Today, analytics is most often used in higher education for administrative decisions—from delivering the targeted number and quality of a freshman class to cultivating likely donors. But the use of analytics will likely grow in high-stakes areas such as academic success. Whether the catalyst for adoption is a call for accountability from outside of higher education or the need for scorecards or decision-making models from within, analytics is in higher education's future. To prepare, IT and institutional leaders need to begin to understand analytics—as well as the changes that may be required in data standards, tools, processes, organizations, policies, and institutional culture.

Many institutions have implemented analytics to improve enrollment management. Institutional researchers collaborating with admissions staff have created complex formulas—based on standardized exam scores, high school coursework, and other information—to determine which applicants will be

admitted. The “actionable intelligence” generated from statistical analyses of these diverse data sources can guide a more efficient use of limited admissions budgets and staff time. For some institutions, analytics means that the institution can provide applicants with an immediate response to their admissions applications. Analytical models and decisions have been refined over the years to produce fairly predictable enrollment rates, as well as balance in areas such as in-state/out-of-state students, enrollments within programs, and other demographic factors. Based on current data, models are refined annually to improve enrollment decisions.

Beyond enrollment management, analytics is increasingly used to inform fund-raising. By building a data warehouse containing information about alumni and friends, institutions can use predictive models to identify those donors who are most likely to give. Aside from academic and co-curricular history, information may



include an individual's response to past solicitations; interest in particular college/university initiatives; employment in, contributions to, and honors received in fields related to institutional programs; and participation in institutional events.

With the increased concern for accountability, *academic* analytics has the potential to create actionable intelligence to improve teaching, learning, and student success. Traditionally academic systems—such as course management systems, student response systems, and similar tools—have generated a wide array of data that may relate to student effort and success. Early academic analytics initiatives are seeking to predict which students are in academic difficulty, allowing faculty and advisors to customize learning paths or provide instruction tailored to specific learning needs.

Sample Academic Analytics Initiatives

Baylor University, the University of Alabama, Sinclair Community College, Northern Arizona University, and Purdue University are among the pioneers in higher education analytics.

Using Enrollment Predictive Modeling at Baylor University

Since the late 1990s, Baylor University has pioneered and refined the gathering and analysis of massive amounts of data on prospective students to support a sophisticated admissions strategy. The traditional admissions cohort-action funnel (moving from inquiries to applications to acceptances to deposit payments through enrollments) analyzes factors from the application stage forward to the enrollment yield. Baylor, however, looks at a broad array of variables from the *inquiry* stage forward.



IT and institutional leaders need to begin to understand analytics—as well as the changes that may be required in data standards, tools, processes, organizations, policies, and institutional culture.

Out of all the variables available, Baylor identified eight that result in the best predictive model for Texas residents:

- Attendance of a premier event
- Campus visit
- Extracurricular interest
- High school attended
- Mail qualifying score (Baylor level of interest)
- SAT score (for non-Texas residents, this variable was replaced by the number of solicited contacts)
- Number of self-initiated contacts
- Telecounselor score (Baylor level of interest)¹³

The inquiry pool is updated once a week with additional prospective students and with additional data on existing inquirers. Scores from the predictive model are added to the student database, which admissions staff can then query to identify those students most likely to be admitted. Specific scores trigger who will receive various types of follow-up. For example, the top 75 percent of scored inquirers are sent the expensive-to-print-and-mail campus viewbook, whereas the bottom 25 percent receive only a reply card and the application form. The goal over time has been to refine the model so that no mailings go to those lower-scoring inquirers who are least likely to enroll.

The same model can be used to select the most likely enrollees for phone calls by telecounselors. Admissions counselors making high school visits use the model to identify students to seek out personally. At all points of contact, admissions staff and telecounselors enter new or updated interest scores into the prospect database; the system populates missing data with averages from past cohorts. The net effect has been the creation of information that lets

Baylor segment its prospect pool, target the most likely enrollees, and more efficiently use human and financial resources to deliver the desired freshman class. One measure of this use of actionable intelligence was the increase in new student applications from approximately 15,000 for the fall semester of 2005 to approximately 26,000 in fall 2006.

Predicting and Improving Student Retention at the University of Alabama

As part of an effort to improve student retention from the freshman to sophomore year, the University of Alabama (UA) experimented with analytics. Graduate students in a data-mining course were given access to the data files of enrolled freshmen (identities were concealed) from 1999, 2000, and 2001 and were asked to develop predictive models of at-risk students. Using such statistical techniques as logistic regression, decision trees, and neural networks, the students developed a single refined model with eight significant variables:

- UA cumulative GPA
- English course
- English course grade
- Distance from UA campus to home
- Race
- Math course grade
- Total earned hours
- Highest ACT score (ACT or ACT-converted SAT score)¹⁴

Using the retention model along with pre-enrollment data and an aggregate cut-off score—a score developed in consultation with the university registrar and one below which students are considered in need of intervention—UA is able to identify 150–200 freshmen each year who are not likely to return for their sophomore year. The information is then shared with faculty and academic advisors for outreach

efforts, counseling, or other action. Consistent with best practices in deploying analytics, a UA admissions research team is presently assessing and refining the retention model and will introduce a new one next year.

Developing a Student Success Plan and Early Alert System at Sinclair Community College

Since 2004, the Student Success Plan (SSP) at Sinclair Community College (SCC) has won seven national awards for its innovative and effective data-gathering for student advising and retention.¹⁵ The SSP is a Web-based counseling records management, reference, and reporting system that uses an SQL database to integrate demographic and admissions information from a data warehouse; real-time course registration, grades, and financial aid status from the student information system; and counselor risk-assessment notes and faculty-initiated early alerts.

Analytics generates a system alert for advisors to initiate an Individual Learning Plan (ILP) whenever any one of the following four criteria appears in a new student's profile:

- Placement-test referrals into two or more developmental courses numbered below "100"
- Individual or family income level below the federal poverty level
- Full-time work
- Undecided major

As the student and the counselor develop an ILP, additional information is captured, including Myers-Briggs Type Indicator, Learning and Study Strategies Inventory, personal life challenges, student satisfaction surveys, course enrollment planning, study plan and tutorial referrals, student progress markers, and counselors' notes.

Aggregate factors that signal student success are achieving a GPA of 2.0 or better, passing all developmental courses, deciding on a major and a career, resolving child-care and transportation issues, and attending class regularly.¹⁶ The fall 2006 impact assessment of the ILP program from the SCC Office of Research, Analytics, and Reporting shows that first- to second-quarter new student retention was 93.3 percent for ILP completers, 76.0 percent for ILP-active students, and 65.7 percent for non-ILP students.¹⁷

The SSP, funded in part through a five-year federal Title III “Strengthening Institutional Programs” planning grant, has future plans to develop a student success course, create secure faculty-access modules, integrate the early-alert system into the campus portal, and develop new analytics reports to support prediction and analysis of student success.

Connecting Resource Utilization, Risk Level, and Outcomes at Northern Arizona University
Northern Arizona University (NAU)

is in the third year of an initiative to use multiple data sources to identify at-risk first-year students and to assess which proactive interventions have the best influence on their academic success and retention. Like most other colleges and universities, NAU has a robust set of academic and student-life support resources, but members of a support task force knew that students and, by extension, staff were using services in a reactive way—after problems had surfaced and students were already at greater risk of failure-withdrawal. The task force set as its goal a predictive model that would identify which students would benefit from which resources.¹⁸ The model comprised three critical elements:

- *Resource/Service Utilization*: a distillation of NAU’s numerous resources and services into five categories that would make the analysis manageable but not too broad-brush: (1) academic services (academic advising and tutoring); (2) recreational resources (recreation cen-



ter usage, fitness program, intramural sports, and wellness education); (3) social resources (student organization membership, after-hours events, and social activities); (4) academic referrals (by centralized advising center staff to academic departments for curricular and degree program assistance); and (5) advising/career sessions (with centralized advising center staff and resources). Although gathering the data for some categories of utilization involved blending online records and manual rosters, much of the data gathering was facilitated by usage records created from student ID card swipes by service offices and at campus activities and events.

- *Levels of Risk*: established by admissions test scores, high school GPAs, and psychosocial factors as measured by NAU’s deployment of the ACT Student Readiness Inventory
- *Outcomes*: measured by first-year student GPAs and enrollment retention status

Among the telling results: the GPAs of students who used one to three academic services increased 0.192 points, on average; those who used four services increased GPAs by 0.280 points; and students who were high-risk and used four services increased GPAs by 0.460 points. The utilizations that had the greatest impact on retention were academic referrals and advising/career sessions.¹⁹

The NAU researchers found that the most efficient use of advising and support resources occurs when interventions are focused on high-risk students who are engaged through academic referrals and advising/career sessions. They also recognized an important aspect of “intrusive advising”: that despite the positive gains



Analytics can be used to examine key institutional issues, such as enrollment or retention, but the decision to move forward with analytics depends on knowledgeable champions among senior administrators.

in performance and retention, the way in which students learn about institutional efforts on their behalf may affect their perceptions of privacy; consequently, the timing and the content of communications require careful planning.

In the future, NAU hopes to use prediction information to connect resources and services with students as early as possible and to continue tracking both resource use and student success to refine the predictive model. One possibility is to add new data sources to the model, such as UCLA's Cooperative Institutional Research Program (CIRP) Freshman Survey (<http://www.gseis.ucla.edu/heri/freshman.html>) and the National Survey of Student Engagement (<http://nsse.iub.edu/redirect.cfm?target>).

Using Course Management System Data to Identify At-Risk Students at Purdue University

Purdue University is extracting data from the course management system (CMS) to build models that predict which students may be struggling academically and to provide proactive intervention. Purdue's premise is that student academic success is the result of the student's aptitude (as measured by standardized test scores and similar information) and the student's effort (as measured by participation within the CMS).

The CMS was selected as the initial focus because of broad campus adoption and the automatic collection of more than twenty activity variables, ranging from the time spent within the CMS to the number of discussion postings. Since it can be difficult to understand how each faculty member utilizes the CMS, an individual's data is compared with that of his/her peers. The resulting comparison provides a standardized value of student effort (e.g., the number of discussion

postings is shown as the standard deviation from the class mean). If the faculty member decides not to utilize a tool, everyone in the course is "average."

Using factor analysis and logistic regression, a model was developed to predict student success within a course. Six variables were found to be significant:

- ACT or SAT score
- Overall grade-point average
- CMS usage composite
- CMS assessment composite
- CMS assignment composite
- CMS calendar composite

Two models were developed and validated: for freshmen and for the overall campus population. Both models included the same significant variables, but the variables differed in importance. In each model, the CMS data significantly contributed to predicting academic success. However, the freshman-only model was able to correctly classify nearly 80 percent of the students, whereas the all-student model was able to correctly classify only 67 percent of the students.

Purdue's next step is to connect the results of the predictive models to existing student-intervention programs including help desks, supplemental instruction, and similar programs.

Building an Academic Analytics Initiative

Academic analytics relies on the extraction of data from one or more systems, such as the CMS or a student information system. The data, which may be stored in a data warehouse for ongoing use, is analyzed using statistical software, and a mathematical model is generated. Based on the model and predetermined values, a particular action may be triggered, such as sending the student an electronic noti-

fication or initiating a personal intervention by college/university staff.

For example, data extracted from a student information system provides baseline student demographic, academic performance, and aptitude information. The CMS provides a snapshot of the student's course-related efforts by providing real-time interaction information that allows for comparison with peers. The two sources of data are combined to predict the probability of student success. Using this probability, the institution can decide whether to take certain actions such as inviting a student to a help session via e-mail or calling a student with an invitation to meet with an advisor.

Three characteristics of successful academic analytics projects are worth highlighting:

1. Leaders who are committed to evidence-based decision-making
2. Administrative staff who are skilled at data analysis
3. A flexible technology platform that is available to collect, mine, and analyze data²⁰

Any academic analytics effort begins with leaders who are committed to decision-making based on institutional data. Analytics can be used to examine key institutional issues, such as enrollment or retention, which by their nature are complex and often sensitive, but the decision to move forward with analytics depends on knowledgeable champions among senior administrators.

The second critical component to building an academic analytics initiative is staffing. Staff members involved in analytics efforts often include database administrators, institutional researchers, educational researchers, programmers, and domain specialists (e.g., student

services, retention, development/ advancement). Academic computing staff may be needed to collect information from various academic systems such as the CMS. The team must have the skill to build predictive models based on institutional data guided by educational research. Other staff may be needed to focus on policy development and clarify who has access to the data, how the data can be used, and which data-security models are required. Since analytics requires data analysis, institutions will need to invest in effective training to produce skilled analysis staff. Obtaining or developing skilled staff may present the largest barrier and the greatest cost to any academic analytics initiative. Whether such staff are added to existing institutional research units or are cultivated in the IT organization, student affairs divisions, or academic units will depend on the organizational culture and the locus of resources.

The third element in any academic analytics project is technology. A data

warehouse is the key component of the technology infrastructure, housing information from a variety of sources in a common structure that enables data analysis. To populate the data warehouse, the institution will need to build a “bridge” between the application and the warehouse. For some applications, standard interfaces facilitate the transfer of data. For other applications, the interface development requires significant programming effort.

Piecing together a coherent academic analytics effort can be difficult, requiring support from many units: enrollment management, institutional research, IT, the registrar’s office, academic divisions, student affairs, and more. Standards must be agreed upon for the data (e.g., is enrollment based on headcount on day seven after the start of the semester or on day ten?). Extracting information from academic systems requires careful analysis and programming effort. Building the appropriate models requires staff with statistics and educational research backgrounds. Creating interventions requires



domain knowledge (e.g., advising, retention) and advising/counseling skills. For institutions to be successful in academic analytics projects, IT leaders must build a coalition of people.

Considerations and Concerns

With analytics being used to address complex institutional issues, concerns are likely to arise that the complexity has been reduced to “a number,” potentially resulting in oversimplification or insensitivity. In particular, analytics projects that focus on teaching and learning need to be approached deliberately. As these projects are developed, refined, and implemented to support academic success, several concerns must be addressed:

- *Big Brother:* The notion that a person or institution can track the actions of individuals within a software application will be welcomed by some and will be threatening to others. Who determines which data is collected? What obligation does the institution have to inform faculty and/or students that their behavior within an application is being tracked? Does an individual need to provide formal consent before data can be collected and/or analyzed? Does an individual have an option to “opt out” of an analytics project?
- *Holistic View:* Although analytics produces a prediction based on the data available, no prediction can take into account all the possible causes of success or lack of success (problems at home, financial difficulty, and so on). In addition, some will be skeptical of the ability of “a number” to account for the interpersonal relationships and personal growth that come from

attending a college or university, irrespective of grades or graduation.

- *Faculty Involvement*: As analytics enters the academic realm, ensuring faculty involvement is critical—in the measures as well as the actions that address at-risk students' needs. Faculty are key to “interventions” such as inviting students to office hours, providing additional practice quizzes, or encouraging participation in tutorial programs. For some faculty, analytics may provide a valuable insight into which students are struggling or which instructional approaches are making the greatest impact.
- *Profiling*: One potential use of analytics is to create a profile of successful, or unsuccessful, students. The profile may be used to prompt interventions or to predict student success. Does the profile bias people's expectations and behaviors? Should the institution even create profiles that lead to generalizations about students? Are there profile uses that should be prohibited?

- *Data Privacy*: The data that is collected and analyzed may be protected by federal, state, and institutional privacy regulations. For example, the Family Educational Rights and Privacy Act (FERPA) of 1974 ensures privacy except in cases of “legitimate educational interests.”²¹ Does the institution need approval before data is used? Who has access to the data during model development and implementation? Will the information be shared?
- *Data Stewardship*: The data for any academic analytics project may derive from a wide range of sources. How is the data preserved, secured, and shared? Once a data warehouse has been established, can anyone use it for any purpose? If not, how are use decisions made?
- *Information Sharing*: Initial academic analytic models have produced a probability of student success. Should the results be shared with the student, faculty, or other staff? Who makes the determination of what and

how information is shared?

- *Obligation to Act*: If the academic analytics model provides a probability of student success, what is the obligation of faculty, students, and institutions to act on that information? With whom does the obligation to act lie? How is the responsibility shared among different groups?
- *Distribution of Resources*: With quantifiable prediction models, the distribution of resources to those who most need them may emerge as an issue. Will access to support services be limited to those with the greatest need, or will anyone who has interest be able to receive help? Who receives priority if resources are limited?

Analytics can be a powerful tool for higher education, but analytics can also magnify existing value conflicts and introduce new ones. Higher education will need to balance the expectations of faculty (are faculty required to intervene in all situations where a student is at

risk?), federal privacy laws (who can view student information or have access to predictions of success?), and the institution's own philosophy of student development (should the institution endorse a sink-or-swim or a nurturing environment?).

Potential Impact

Richer data sets, new ways of extracting and organizing data, more sophisticated predictive models, and additional research will drive the evolution of analytics. As the practice of analytics is refined, colleges and universities can place more and better information into the hands of a greater number of people, enabling informed decision-making.

With the public demand for documented learning outcomes and increased retention, academic analytics can contribute to institutional action. Data from the CMS, e-portfolios, student response systems, course podcast downloads, and similar applications can be used in academic analytics. And the focus of future analytics efforts can

shift from predicting who is going to be successful to customizing learning environments so that the most effective instructional approaches are used for each student. Eventually, institutions may be able to provide unique learning paths, matching instructional activities to a student's learning needs.

As higher education continues to implement systems that collect a wide range of data, IT units will be called on to support analytics efforts. As a result, IT leaders will find that new expectations are being placed on their units. Staff will be required to have more than the traditional IT skills. They will need to be adept at mining data, understanding the nature of the data, creating metadata to provide long-term data management, analyzing data from multiple sources, and developing models that can be used for decision-making and action. By sharing data that is collaboratively interpreted and acted on, IT can help institutions bridge academic affairs and student affairs. As colleges and universi-

ties respond to the demand for greater accountability in higher education, the emerging practice of academic analytics is likely to become a new, highly useful tool for a new, highly demanding era. **e**

Notes

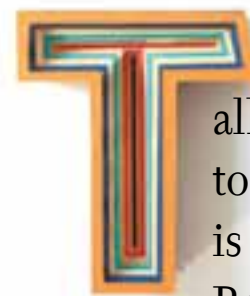
1. U.S. Department of Education, by the National Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform*, a Report to the Nation and the Secretary of Education, April 1983, <<http://www.ed.gov/pubs/NatAtRisk/index.html>>.
2. "Did You Know?/Shift Happens," created by Karl Fisch, modified by Scott McLeod, <<http://www.youtube.com/watch?v=xHWTLA8WecI>>, sources: <<http://www.lps.k12.co.us/schools/arapahoe/fisch/didyouknow/sourcesfordidyouknow.doc>>; blog post: <<http://thefischbowl.blogspot.com/2006/08/did-you-know.html>>.
3. Thomas L. Friedman, *The World Is Flat: A Brief History of the Twenty-First Century*, 1st updated and expanded ed. (New York: Farrar, Straus and Giroux, 2006).
4. "Education at a Glance, 2005," <<http://www.oecd.org/dataoecd/41/13/35341210.pdf>>.
5. Anne K. Walters, "Minority Education Should Be a Priority, Report Says," *Chronicle of Higher Education*, November 18, 2005.
6. Debra Humphreys, "Achieving Equity as Generation Y Goes to College: New Data," *Diversity Digest* (Spring/Summer 00), <<http://www.diversityweb>>

- .org/Digest/Sp.Sm00/geny.html>
7. G. Orfield, D. Losen, J. Wald, and C. Swanson, *Losing Our Future: Our Minority Youth Are Being Left Behind by the Graduation Rate Crisis* (Cambridge, Mass.: The Civil Rights Project at Harvard University, 2004), p. 2, <http://www.urban.org/UploadedPDF/410936_LosingOurFuture.pdf>.
 8. Kevin Carey, "One Step from the Finish Line: Higher College Graduation Rates Are Within Our Reach" (The Education Trust, January 2005), <http://www2.edtrust.org/NR/rdonlyres/12656449-03_FD-4F3F-A617-920E58F009C0/one_step_from.pdf>.
 9. Daniel E. Hecker, "Occupational Employment Projections to 2014," *Monthly Labor Review* (November 2005), <<http://www.bls.gov/opub/mlr/2005/11/art5full.pdf>>.
 10. Patrick J. Kelly, "As America Becomes More Diverse: The Impact of State Higher Education Inequality" (National Center for Higher Education Management Systems, November 2005), <<http://www.higheredinfo.org/raceethnicity/InequalityPaperNov2005.pdf>>; June Kronholz, "Economic Time Bomb: U.S. Teens Are among the Worst at Math," *Wall Street Journal Online*, December 7, 2004, <<http://www.mathnasium.com/press/wsj1204.pdf>>.
 11. "Fact Sheet: The National Survey of America's College Students" (American Institutes for Research, January 19, 2006), <<http://www.air.org/news/documents/collegeliteracyfactsheet.htm>>.
 12. "Education at a Glance, 2005," <<http://www.oecd.org/dataoecd/41/13/35341210.pdf>>.
 13. Kathleen Morley and Tom Bohannon, "Applications of Data Mining in Institutional Research," presentation, Association for Institutional Research Annual Conference, Cincinnati, May 2000.
 14. Cali M. Davis, J. Michael Hardin, Tom Bohannon, and Jerry Oglesby, "Data Mining Applications in Higher Education," unpublished book chapter (2007).
 15. Awards include the 2004 EDUCAUSE Information Technology Solutions Award, the 2005 National Council for Student Development Exemplary Practice Award, and the 2007 Bellwether Award for Instructional Programs and Services.
 16. "Student Success Plan and Early Alert System: Individual Learning Plan, Holistic Counseling, and Intervention Model," Sinclair Community College Presentation, March 5, 2007, <<http://www.sinclair.edu/stservices/edu/pub/Generic%20SSP-EAP%20Pres.ppt>>.
 17. Sinclair Community College, Office of Research, Analytics, and Reporting, "Fall 2006 ILP Results," March 2007, <<http://www.sinclair.edu/stservices/edu/pub/FA06%20ILP%20Results.pdf>>.
 18. Rebecca Pollard Cole, Margot Saltonstall, and Paul A. Gore Jr., "Assessing Student Readiness to Promote Student Success: A Campus Collaboration," unpublished manuscript developed from a paper presented at the 25th Annual Conference on The First-Year Experience, Atlanta, Georgia, February 24–28, 2007.
 19. Steven Robbins, Jeff Allen, Alex Casillas, Adaeze Akamigbo, Margot Saltonstall, Rebecca Cole, Eileen Mahoney, and Paul Gore, "Associations of Resource and Service Utilization, Risk Level, and College Outcomes," unpublished paper (2007).
 20. Philip J. Goldstein, with Richard N. Katz, "Academic Analytics: The Uses of Management Information and Technology in Higher Education," *EDUCAUSE Center for Applied Research (ECAR) Study*, vol. 8 (2005), key findings publicly available: <http://www.educause.edu/ir/library/pdf/ecar_so/ers/ers0508/EKF0508.pdf>.
 21. The specific FERPA provision (§99.31) under which a student's prior consent is not required to disclose information is as follows: "An educational agency or institution may disclose personally identifiable information from an education record of a student without the consent required by §99.30 if the disclosure meets one or more of the following conditions: (1) The disclosure is to other school officials, including teachers, within the agency or institution whom the agency or institution has determined to have legitimate educational interests" <http://www.ed.gov/policy/gen/reg/ferpa/rights_pg19.html#17>. This would seem to clear the anticipated federal legislative hurdle. However, whether there is an institutional will and constituent buy-in to create and share "actionable information" from diverse sources is something else altogether. Because each institution is obligated to interpret and publicly declare what "legitimate educational interests" means, this is an aspect of FERPA that has beguiled registrars and academic administrators for some time in aligning local information-sharing policies.

This is the first
in a three-article
EDUCAUSE Review
series exploring
analytics.

By Diana G. Oblinger

Let's Talk...



Talk about analytics seems to be everywhere. Everyone is talking about analytics. Presidents are talking. CIOs are talking. Corporate representatives are talking. Government representatives are talking. Media pundits are talking. Yet even with all the talk, many in higher education have questions about—and objections to—using analytics in colleges and universities.



ANALYTICS

“What is analytics?”

Analytics is the use of data, statistical and quantitative methods, and explanatory and predictive models to allow organizations and individuals to gain insights into and act on complex issues. In colleges and universities, analytics is used to improve operational efficiency and student success. The term *big data* is often used interchangeably with *analytics*, but the scientific community uses *big data* to describe research that uses massive amounts of data. The use of analytics to improve administrative functions is often called *business intelligence*; similarly, *academic analytics* is used to help run the business of the higher education institution. Finally, *learning analytics*

focuses specifically on students and their learning behaviors, gathering data from course management and student information systems in order to improve student success.¹ Although the labels can be confusing, overall the term *analytics* refers to an approach that can be used to explore a broad range of questions. The emphasis should be on the question to be addressed rather than the tool.

“We’ve been doing this for years. What’s new?”

Yes, we have been compiling data and generating reports in higher education for decades, but several things are new about analytics today. One is the volume of data being collected. Some of it is

automated; other data comes from “data trails,” whether left in social media, online purchases, or course management systems. Our computational capacity is radically different—in terms of speed, complexity, and throughput. And much of the use of analytics today goes beyond reporting—it is predictive and prescriptive. Analytics can provide a report on a query or a hypothesis. Analytics can predict which majors or courses match student profiles, which interventions are likely to help at-risk learners, and thus which actions should be taken next. Visualization is also an important element of analytics, revealing patterns that would have been impossible to discern by looking at the data alone. Lastly,

demand has risen significantly. Public pressure for accountability, efforts to increase efficiency, and concerns about college affordability are driving up the demand for analytics.

“Faculty and students don’t trust us with their data.”

People like to control data, often preferring to set context or provide interpretation. There are very legitimate concerns about data privacy and security. However, the lack of “trust” may be a fear that information will be used against someone. For example, a faculty member might fear that students’ performance data will be used to judge his or her effectiveness as an instructor. Analytics enables comparisons that are not typically performed today—such as expenditures across colleges or cost-to-degree across majors. This transparency can be threatening. Data governance is an essential issue for institutions to address; without it, distrust

“We can’t use student data because of FERPA.”

Institutions must be concerned about the privacy of student information. Yet though the Family Educational Rights and Privacy Act (FERPA) is an often-cited concern, it isn’t necessarily a roadblock. FERPA stipulates that institutions must not disclose personally identifiable information from a student’s education records without consent. But there are recognized exceptions for use by college officials. Using data about students to make decisions within the institution, such as to help advisors monitor students’ academic progress or decide which courses to suggest for a particular student, is acceptable. Students may also consent to the use of their data (e.g., for plagiarism detection through services like Turnitin) or may voluntarily disclose information about their educational record (e.g., as part of the college admission process). Transparency—whether or

To be without the tools that allow an institution to improve graduation rates or reduce unnecessary costs is a risk. Pressures on accountability, affordability, and effectiveness are making analytics essential. Very few colleges and universities can afford not to use analytics.

“Analytics is all hype and will go away.”

Although higher education analytics might be in a hype cycle, analytics has become a staple of other industries: retail, finance, insurance, health care, government. Even in higher education, few believe that analytics will go away. A growing emphasis on institutional performance, including calls for cost containment and outcomes-based funding tied to measures of college completion, requires that institutions have the capacity to understand what drives not only their performance but also that of their students. The trend toward more compliance and reporting at the federal and state levels is unlikely to change; in fact, most in higher education expect that the demand will only increase. The importance of accreditation is also a constant driver. Still, it is important to note that analytics is not the answer for everything. Sometimes analytics does not apply or is not practical. Even when analytics does apply, the best decisions are made by combining analytics with sound reasoning.

The pressures on accountability, affordability, and effectiveness are making analytics essential. Very few colleges and universities can afford not to use analytics.

grows. Knowing the purposes for which data is being collected and how sensitive data will be handled is important.

“Numbers aren’t enough to dictate what to do.”

Analytics is a tool—one that must be used for the right purposes. The key is asking good questions. But making incorrect assumptions and failing to test them can undercut even good questions. Any numbers, unwisely used, can be misleading. When using analytics, institutions need to let the facts be the guide to the right answer. Incorrect conclusions can be reached by failing to take the time to understand alternatives or interpret the data correctly. There are different decision-making styles, of course. Some people trust analysis over judgment. Others prefer to rely exclusively on intuition. Most seek a balance between intuition and analysis.

not it is required by law or regulation—is the key consideration. Letting students know what information about them will be used for what purposes, by whom, and for what benefit is critical. In addition, students should know what they can do about any data-privacy concerns. Being clear about data, its use, benefits, and safeguards may allay fears.

“We can’t afford to use analytics.”

Colleges and universities cannot afford *not* to use analytics. Yes, there is a cost to establish the infrastructure as well as to obtain the right talent for analytics. However, few institutions will be able to cope with today’s information demands or decision making without analytics. State reporting requirements call for an increasing level of sophistication. Boards, trustees, and legislative groups are demanding information, analysis, and projections that hinge on analytics.

“Analytics will dehumanize education.”

Rather than dehumanizing education, analytics is being used to personalize education and improve student success. The behavior of individual students can be monitored, and alerts can be sent if a student is lagging. The kind of assistance provided to the student can be targeted to his or her specific needs. Without tools such as analytics, students’ progress—and problems—may be hard to detect. Rather than allowing students to fall through the cracks or become “just a number,” colleges and universities can—through analytics—see and act upon individual students’ needs.

“What’s in it for me and my institution?”

Better insight. Clearer choices. Faster decision making. Analytics allows faculty to see patterns of success or disengagement in student populations. Analytics allows students to understand how their behaviors are linked to learning outcomes, not just to improving their grade in a single course. Analytics allows institution leaders to make better decisions about energy utilization, financial models, and staffing. Large data sets, powerful analytics engines, and skillfully designed visualization techniques can send alerts about matters requiring attention and can extrapolate to the future using predictive modeling and optimization techniques.

“What’s the next step?”

Individuals and institutions must engage in a dialogue about analytics. While the

technical issues might seem straightforward, many of the policy and ethical issues are complex. For example, should students be able to opt out of having their learning analytics data collected? What is the recourse for individuals who have had their data misused or inappropriately shared? Who owns the data mined from a learning process?

Analytics is being used in, and all around, higher education. Policies are being established. Statewide longitudinal data systems are being implemented. Massive amounts of data about students are being collected by corporations. If colleges and universities do not begin talking about how to leverage analytics, others will. Those of us in higher education can afford to ignore analytics only if we are willing to accept what others decide. If we aren’t willing to do that, we must be part of the solution—so that we

will have a voice in, and an ability to influence, the future directions of the use of analytics in higher education. ■

Note

1. Angela van Barneveld, Kimberly E. Arnold, and John P. Campbell, “Analytics in Higher Education: Establishing a Common Language,” *ELI Paper 1: 2012* (January 2012), <<http://net.educause.edu/ir/library/pdf/ELI3026.pdf>>.

© 2012 Diana G. Oblinger. The text of this article is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).



Diana G. Oblinger
(doblinger@educause.edu)
is President and CEO of
EDUCAUSE.

This is the second in a three-article *EDUCAUSE Review* series exploring analytics.

By Diana G. Oblinger

ANALYTICS

What We're Hearing

Over the last few months, EDUCAUSE has been focusing on analytics. As we hear from experts, meet with association members, and watch the marketplace evolve, a number of common themes are emerging. Conversations have shifted from “What is analytics?”¹ to “How do we get started, and how do we use analytics well?” What we are hearing from experienced implementers can benefit those who are getting started.

“Analytics requires a culture of inquiry, and inquiry creates an analytics culture.”

Analytics is a matter of culture—a culture of inquiry: asking questions, looking for supporting data, being honest about strengths and weaknesses that the data reveals, creating solutions, and then adapting as the results of those efforts come to fruition. Analytics not only relies on a culture of inquiry but can promote such a culture. Analytics, used well, is more like a flashlight than a hammer, illuminating where progress has been made and where

improvements are needed. There's a synergy between investing in analytics and developing a culture that embraces analytics.

“Ask good questions; use good data.”

Analytics is a problem-solving process: asking strategic questions, identifying relevant data, conducting analyses, exploring options, communicating results, and taking action. The more clarity there is around the problem statement, the higher will be the probability of success. The questions that analytics seeks to address are complex, the assets dispersed. Asking

good questions requires understanding the context, which means that leaders must gather inputs from all sides to obtain the perspective needed to address strategic questions. Questions must drive analytics, not the data at hand. Analytics requires collaboration among different campus divisions such as academic affairs, student affairs, institutional research, information technology, and finance. Analytics, as a discovery and decision-making process, requires stakeholders to think strategically about the role of data: inventorying what is available, assessing what is needed, and considering how it will be used.

“Analytics is a journey from ownership to stewardship.”

“Data hawks” will advocate for control of their data, whereas “data doves” believe that “data just wants to be free.” Both approaches create difficulties. Shifting the conversation toward stewardship helps overcome polarization. Campus leaders should consider reframing ownership

with the concept of “campus data custodians” whose role is protecting the longevity, quality, and relevance of data. Data governance can challenge political, technical, and organizational assumptions. Establishing data governance is a complex but necessary undertaking. It is likely to be multiyear project, whether at a large or small institution.

“Communication is essential to any analytics effort.”

Experts are clear that analytics demands “culture change.” Culture change requires communication and a shift of beliefs. Even so, few institutions have a comprehensive communication and training program in place to help their campus community understand analytics. Ongoing training, engagement, and communication are also essential aspects of rolling out new dashboards and other visualization tools.

“Analytics requires new skill sets.”

One of the biggest barriers to analytics can be the shortage of necessary skills. Recent research suggests that by 2018, the United States could lack 140,000 to 190,000 workers with “deep analytical” expertise. Additionally, the U.S. talent gap will require retooling or hiring 1.5 million data-literate managers.² A new field may be in the process of being defined—data science—which reflects the centrality of data as a resource. Data science blends multiple fields such as statistics, applied mathematics, and computer science. It blends not only disciplines but also other skill areas such as technical skill, teamwork skills, communication skills, and tool mastery.³ Other segments will compete for these skills—not just higher education.

“Risks are inherent in any decision—even in no decision.”

Risks are inherent in making any decision. Not adopting analytics does not avoid risk—it risks missed opportunities. When decisions are made with insufficient information, risks increase. Analytics adopters must accept the risk of exposing institutional weaknesses, perhaps obligating action based on new knowledge. Carefully

selected questions, quality data, sound practices, and prudent processes mitigate risks. Thoughtful analytics yields better processes and decisions. Risk management is important; risk avoidance is not a solution.

“The IT organization should neither drive nor own analytics.”

Analytics must be a shared responsibility across all parts of the higher education institution, with high-level executive ownership of the analytics agenda. The CIO is the executive most directly charged with developing infrastructure and increasing capacity for analytics. CIOs are most likely to have knowledge of where data resides and how business practices work across the institution. CIOs are accountable for data security and integrity. Although the IT organization has an important role, strong partnerships across the institution are critical, especially for information technology and institutional research as they negotiate data governance and ownership. Finance, admissions, student services, and academic affairs must become effective partners in adoption as well. Even within the technical support realm, the responsibility for analytics should not lie with the IT organization alone.

“Analytics is an investment.”

Analytics is an investment for the future. The ultimate goal of analytics is to make more-intelligent decisions. Those decisions provide the ROI. For example, analytics efforts to improve student outcomes often begin by identifying barriers. Approaches are tested; results are compiled; improvements are made. Improving pass rates and reducing course re-takes save resources—for students and the institution. Plus, the documented benefits of such analysis and decisions help convince others. When it becomes clear that the resulting redesigns are leading to success, the investment rationale is clear. Dashboards track progress toward strategic objectives. From a financial perspective, building and using a data warehouse may be a much better investment than designing custom reports and transactional reporting. Ana-

lytics can be an investment in building the quality and effectiveness of the institution.

Finally, in getting started with analytics, IT and institutional leaders may want to heed the following words of wisdom:

- Transitioning an institution from a culture of anecdote to a culture of evidence is a proactive process. Leaders must build structures—both technical and governance—and develop data sources, data flows, and data accessibility to support the initiatives.
- When people are in control, have ownership of quality data, and can trust the data, everything changes.
- Technology makes education more personal, not less. Systems don't replace people; they empower people—both advisors and students—to make better decisions.

Analytics is a powerful tool, one that we are just beginning to understand. ■

Notes

1. Diana G. Oblinger, “Let's Talk ... Analytics,” *EDUCAUSE Review*, vol. 47, no. 4 (July/August 2012), <http://www.educause.edu/library/ERM1240P>.
2. James Manyika et al., “Big Data: The Next Frontier for Innovation, Competition, and Productivity,” McKinsey Global Institute, May 2011, http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation; Steve Lohr, “The Age of Big Data,” *New York Times*, February 11, 2012, <http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>.
3. See, for example, Dan Woods, “What Is a Data Scientist? Michael Rappa, Institute for Advanced Analytics,” *Forbes*, March 5, 2012, <http://www.forbes.com/sites/danwoods/2012/03/05/what-is-a-data-scientist-michael-rappa-north-carolina-state-university/>.

© 2012 Diana G. Oblinger. The text of this article is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).



Diana G. Oblinger
(dblingle@educause.edu)
is President and CEO of
EDUCAUSE.

This is the third in a three-article *EDUCAUSE Review* series exploring analytics.

By Diana G. Oblinger

ANALYTICS

Changing the

We live in an information culture. We are accustomed to having information instantly available and accessible, along with feedback and recommendations. We want to know what people think and like (or dislike). We want to know how we compare with “others like me.” Just as analytics powers Amazon.com, Netflix, TheRestaurantFinder.com, and Yelp, it can support tools for student empowerment.

Empowerment hinges on having access to information, which can help students make better choices to navigate college more successfully. As we have explored student empowerment in higher education, several guidelines have emerged from what we are learning.

“Use data to change the conversation.”

Students, faculty, staff, and the public—we all make assumptions easily. Having data at hand can change the conversation by informing questions and providing concrete answers. In a highly complicated

environment, assumptions and anecdote aren't enough. Higher education has a responsibility to provide guidance to help students make good decisions based on data. But using data well requires the development of “data literacy.” We must help

students understand the data and what it means to them.

For example, too many course choices can be confusing for students, and making poor course selections could delay graduation. Degree Compass, a course-recommendation system developed at Austin Peay State University, uses predictive analytics to help students find the courses that best fit their degree program, overlaid with a model predicting the students' success in the course. The system guides students and advisors in deciding on a pathway to graduation. Recommended course lists, the role each course plays in a degree program, and class availability information are readily available to students.

“Move from the past to the predictive.”

Much of our data use has revolved around reporting on what happened—in the past. Data use is moving to the

and predictive algorithms, system alerts trigger individualized interventions that can help students, advisors, and/or faculty tap resources to avert failure.

“Empower choice, don’t restrict it.”

For almost any decision, there are more options than we may realize—sometimes so many that decision-making can become paralyzed. As noted above, choice—without good information and guidance—can be the enemy of student success. Analytic systems allow us to personalize recommendations that help students make better-informed choices. With good information, students might ask: “What should I do differently?” And they might use data to find alternative paths to their goal.

Knowing whether they are “on track” makes a difference to students. The

learners to personalized learning pathways tailored to their needs, aspirations, abilities, and timelines. Although data has historically flowed in only one direction, serving to validate compliance or trigger funding, data is actually most useful to inform thinking, questioning, planning, and next steps.

Rapid feedback enables accelerated learning for students. Carnegie Mellon University’s Open Learning Initiative uses feedback loops based on analytics. Students receive feedback in corrections, suggestions, and cues tailored to their performance. Student learning is accelerated, along with equal or better retention and performance compared with traditional instruction. Instructors receive feedback about students’ knowledge, how students are using course materials, and students’

Conversation

predictive—to what is likely to happen. Today’s online systems collect more data and more detailed information about how students learn, capturing inputs, problem-solving sequences, number of attempts, and time spent on task. Detailed learning activity can be used to predict how a student will perform in a future context, providing useful feedback to the student, the instructor, and the institution. The data is key to adaptive systems, diagnostics, and alerts that can make students aware of risks and can help students avoid problems. These systems are designed not to replace personal interaction but to make education more personal through timely alerts that assist individual students.

Purdue University and Rio Salado College are among the pioneers of predictive analytics, which allows institutions to make predictions and anticipate problems. Based on personalized data

eAdvisor system at Arizona State University and STAR at the University of Hawaii assist students in selecting appropriate courses and tracking progress toward their major. eAdvisor was designed to take the guesswork out of how to earn a degree so that students would stay on track to graduate. For example, it helps students choose a major based on their interests and career goals, and it then highlights appropriate course sequences. Because much of the degree requirements tracking is handled online, in-person advising services can be expanded, allowing more time to explore degree and career choices and address individual needs.

“Use data for gateways, not just gatekeeping.”

We have been conditioned to think of data as summative—as an end point. A report, a grade, or a test either opens a door or closes it. Thanks to analytics, data can point

use patterns. Instructors have better information on which to base course refinements and inform theories of learning.



Few transactions are as complicated as college. By informing questions and providing concrete answers, analytics can empower students to make good choices. Analytics can thus change the conversation, both for students and for institutions. ■

© 2012 Diana G. Oblinger. The text of this article is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).



Diana G. Oblinger
(doblinger@educause.edu)
is President and CEO of
EDUCAUSE.