How are the tools of predictive analytics being used, how can they be better used, and how can institutions understand their progress with analytics?

The Unpredictability of Predictive Analytics 2.0

If I were into scrying (the art of predicting the future by gazing into a crystal ball), I would prophesy that EDUCAUSE Review readers will have two equal and opposite reactions on seeing an issue devoted to predictive analytics. The first reaction might be: “Are we still talking about how to use predictive analytics?” And the second reaction might be: “I wonder what predictive analytics we are using on our campus.” We are all accustomed to tracking technologies that are emerging or that may seem to be more hype than substance, but what do we make of technologies like analytics? Here is a combination of tools and practices whose fundamental value is rarely questioned but that have not achieved the traction we might have expected by now. This issue of EDUCAUSE Review is a timely consideration of the state of predictive (and other) analytics across higher education: How are these tools and practices being used, how can they be better used, and how can institutions understand their own progress?

First, in “Predictive Analytics: Nudging, Shoving, and Smacking Behaviors in Higher Education,” Kevin C. Desouza and Kendra L. Smith explore the use of predictive analytics in “nudge theory”—the concept that nudging individuals into making better decisions can be the key to improving institutional effectiveness and student success outcomes, high priorities both locally and nationally. The authors imagine the value of not just gathering data of all kinds but also bringing together and analyzing nonacademic behaviors such as a student’s meal-consumption and gym-attendance patterns. We may know that a student is in trouble sooner if we are paying attention to when he or she starts eating less/more or exercising less/more. Proactively, predictive analytics points the way to harmless and noncoercive nudges to help a student be positioned for success.

However, Desouza and Smith point out that the deployment of predictive analytics is hardly straightforward. After all, without careful attention, noncoercive nudging can become “shoving” or “smacking”—efforts that are “more coercive, restrictive, or punitive” in order to change student behavior and outcomes. Nudging, shoving, and smacking can limit student privacy in the interest of developing interventions, an especially unfortunate outcome if the line is crossed because a correlation is considered causation. For example, it may be that strong academic students eat three meals a day, so perhaps we should charge students more for their meal plan if they eat only two. Or because academically strong students tend to go to the gym, we should shove or smack underperforming students into adding a regular gym routine into their already challenging schedules.

As Desouza and Smith explore potential cautionary concerns, they go even deeper, asking a fundamental question related to what they call “automation of the academic enterprise”: Who gets to decide which interventions should be used for which students? Although algorithms are the “secret sauce” of predictive analytics and automation, “ethical issues come into play,” Desouza and Smith note: “Algorithms are designed by humans and can be programmed to capture biases or make judgments within those biases, either on purpose or accidentally.” Some of these thoughts are reflected in a recent New York Times op-ed, where Kate Crawford notes that artificial intelligence and algorithms reflect the values of their creators. A very real threat is the subtle embedding of human bias in the automation code that we will increasingly rely on, in ways most of us can’t yet even imagine. This is a “data problem,” Crawford concludes. “Predictive programs are only as good as the data they are trained on, and that data has a complex history.”

The source of data is a focus for Chris Dede, Andrew Ho, and Piotr Mitros as well. As they explain in “Big Data Analysis in Higher Education: Promises and Pitfalls,” many of the pitfalls for big data analysis stem from failing to ask the question “Where does data come from?” Conventional digital assessments

(continued on page 6)
also fail to capture the extent to which students have mastered complex skills. As a result, although big data is increasingly being used for decision making in higher education, the authors note that “practical applications in higher education instruction remain rare.”

Dede, Ho, and Mitros emphasize how MOOCs (massively open online courses) provide a promising opportunity for data-intensive research and analysis in higher education instruction: “MOOCs illustrate the many types of big data that can be collected in learning environments. Large amounts of data can be gathered not only across many learners (broad between-learner data) but also about individual learner experiences (deep within-learner data).” Not surprisingly, all of this data collection leads back to predictive analytics: “The most common questions being asked of digital learning data involve prediction.” In one sentence, the authors ground this work in a crucial, student-centered context: “The criterion for prediction is not accuracy, as measured by the distance between predictions and outcomes. Instead, the criterion is impact, as measured by the distance between student learning with the predictive algorithm in place and student learning had the algorithm not been in place.”

We have data, and we have tools for analyzing that data, and we have reasons for using that analysis. To what extent are our institutions collecting the data, adopting the tools, and deploying the analytics? In “Moving the Red Queen Forward: Maturing Analytics Capability in Higher Education,” Eden Dahlstrom illuminates a confounding picture in which interest in analytics is high but deployment lags. As she explains, new tools from EDUCAUSE may help shed some light. We developed our first stand-alone maturity index in 2012, and we are currently beta-testing eight maturity indices and five deployment indices. The strength of these indices is that they offer institutions the chance to answer a few dozen questions and to see, at a glance, the maturity of a specific initiative (i.e., the maturity index) or the stage reached by a given technology deployment (i.e., the deployment index). Institutions can then compare their results with those of other institutions or groups of institutions. In other words, using these new benchmarking tools, institutions can “pop the hood” and see exactly what is going on in eight dynamic topic areas, including analytics. They can also track numerous dimensions—32 factors in the case of analytics, for example.

Not all of what makes an analytics initiative “mature” consists of predictable technology issues such as technology infrastructure and data efficacy. For example, the analytics maturity index also considers the resources and investment dedicated to analytics, the decision-making culture on campus, data-related policy sophistication, and collaboration between IT and IR professionals. It is a complete picture of the complex array of factors that make such initiatives successful. Unfortunately, the dimension of investment and resources remains the least advanced. Dahlstrom notes: “Despite widespread interest, analytics is still not regarded as a major institutional priority at most institutions.” And in the world of analytics investment and interest, learning analytics lags even the lagging institutional analytics—as Dede, Ho, and Mitros also observe. The EDUCAUSE deployment index for analytics shows most of the deployments happening at the experimental level, “with fewer than 21% of institutions reporting institution-wide deployment.” When it comes to longer-term predictions, EDUCAUSE strategic technology research finds that five years from now, big data use by colleges and universities will still be emergent.

It may well be that all three of the conversations about analytics in this issue of EDUCAUSE Review are, well, predictable for a technology in version 2.0 or higher. As we settle into the role that analytics can, will, or should play on our campuses when fully deployed and fully matured, we naturally move from expressions of hype to more realistic, balanced, and cautionary conversations. Tackling the quandary of why higher education hasn’t seen the traction that we expected may not be a separate venture from tackling the tactical, practical, and ethical concerns. Both ventures are necessary to fully realize the promise—dare I say, the predictability?—of predictive analytics in higher education.

Note

John O’Brien (jobrien@educause.edu) is President and CEO of EDUCAUSE.

© 2016 John O’Brien. The text of this article is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0).