The Rise of Big Data

Technology will transform higher education, just as it has many other industries. This will happen as information technology becomes embedded in more institutional processes, such as course enrollment, classroom instruction, and student services. In other industries, including healthcare and travel, this embedding of technology is generating what is being called “big data.” Big data is fine-grained information—about customer experiences, organizational processes, and emergent trends—that is generated as customers conduct normal business. The organization of this data can be a rich source of business analysis that improves performance and even points to new opportunities.

In higher education, the IT infrastructure supporting educational processes allows students to register for a course more quickly, take courses online, and/or connect with campus tutors through social media platforms. Yet higher education institutions gather data from these processes largely for the purposes of reporting to public policymakers. Evidence suggests that very little of it is used to create the data-driven enrollment, instruction, or student-support practices that could promote college completion and success. Meanwhile, emerging technologies not only are providing institutions with data that could facilitate the development of these practices but also are giving students the opportunity to see the data about their journeys, successes, and failures. When provided to students in useful ways, this data can allow students to become better managers of their own educational experiences and can also, perhaps, improve collective outcomes across all of higher education. In short, the era of big data has arrived in higher education.

Technology and Change

Today, we treat higher education as a “black-box” experience managed by the intuition of faculty and administrators. Consequently, students, families, and taxpayers pay a lot of money for an offering they (and we) know very little about. Once we start to get a better sense of what works and what it costs, we can begin to have a real conversation about the affordability and performance of colleges and universities. Tomorrow, information technology will provide more cost-effective ways to ensure that students enroll in and learn from the courses best suited to them and also better manage their student experience.

Beginning with, but not limited to, online education, students and institutions are interacting more with information technology. This interaction is producing ways for students to “personalize” college by using technology to register for and take courses and even to manage their time. In addition, we are beginning to see a rise in the big data that is produced from these interactions and that can be used to empower students to make even better choices as their journey continues. A similar process occurred in the travel industry. With the rise of services such as KAYAK and Priceline.com, consumers became more empowered and began structuring, for themselves, the best experiences for the best prices. Likewise in healthcare, the rise of big data is causing an explosion in customized medicine as diagnostic technology allows physicians to tailor drug therapy and permits patients to manage chronic disease. Similar personalization tools in higher education can improve student learning, course enrollment, course success, and student lifestyle management.

Student Learning

Perhaps the most exciting of the personalization tools in higher education are those emerging to enhance the instructional process. The Open Learning Initiative, or OLI, at Carnegie Mellon University (http://oli.web.cmu.edu/openlearning/) provides an excellent example. OLI brings together evidence-based research in learning, science, and technology to create web-based learning environments. Self-directed learners can use these offerings to achieve the same learning outcomes as students in traditional, instructor-led courses. All OLI courses are online and free of charge. They are offered in student-centered learning environments and have measurable learning objectives and built-in tools to support students in achieving those objectives.

The aspect of OLI that most expresses the big data ethos is its embedded “mini-tutors”: computerized learning environments that are designed around cognitive principles and that interact with students similarly to a human tutor. They provide corrective comments when students err, answer questions about next steps, and maintain a low profile when students are performing well. The mini-tutors have two features that help create more big data. First, they learn with the student—this is called adaptive instruction. Based on a student’s errors, the mini-tutors come to anticipate future challenges, and they provide problem-sets to assist the student in mastering the material. Second, the mini-tutors generate robust data on learning across all students—data that can be used to improve individual performance, enhance course design, and even begin to predict future performance.

Initial research on the learning results of OLI is extremely promising, with students from diverse backgrounds learning
as much as or more than students in traditional classroom settings. With the tools provided by OLI, students can have more complete knowledge about how they learn and thus can manage the instruction process to their benefit.

**Course Enrollment**

An example of personalization in the process of course enrollment comes from Saddleback College in the South Orange County Community College District of California. Saddleback, which enrolls nearly 40,000 students, has developed an application called SHERPA, or Service-Oriented Higher Education Recommendation Personalization Assistant, which works similarly to the recommendation services on Netflix and Amazon. Students’ preferences, schedules, and courses can be stored to create profiles that are responsive to student needs. SHERPA was conceived and shaped by the realization that today’s students are accustomed to receiving recommendations regarding things they are considering doing or buying. So why not build “nudges” and lifelines into the online academic experience? Lifelines are tutors, time-management tools, and life-planning resources than can help students manage competing priorities or access additional support. For example, instead of simply telling a student that a class is full, the program will suggest classes that are open. If students program in their work schedule, SHERPA will guide them to only those classes that are available when they are.

**Course Success**

Building on SHERPA’s course-selection tools, Purdue University developed an early-warning system to help students improve their coursework. Course Signals (http://www.itap .purdue.edu/learning/tools/signals/) monitors students’ behavior patterns and academic performance to determine if they are at risk of earning a low grade, and it allows faculty to intervene with suggestions on actions students can take to improve their grades. An intuitive stoplight dashboard provides indications to students, on their course homepage, if they are underperforming and prompts students to take corrective action.

Course Signals scrapes and analyzes data from gradebooks and activity log-files while also incorporating a student’s demographic information to create a student profile that can be compared with profiles of successful students. The result is that students are able to get a very fine-grained sense of how they are doing in the course overall and can make adjustments to produce better results or can reach out for help from available resources such as faculty or tutors.

**Student Lifestyle Management**

Research on learning communities suggests that helping students manage their academic lives can encourage them to persist in and complete college. Based on this research and also on insights from behavioral science on how people make decisions, developers are using technology to design adaptive software tools (similar to the mini-tutors noted above) to motivate students to persist and succeed in college. The software builds profiles of students’ behavior, academic life, and preferences into interactive tools that help them stay on track.

An early example of this technology is being introduced by a social enterprise called Persistence Plus (http://persistenceplusnetwork.com/). One can think of Persistence Plus as the “Weight Watchers of college completion”: in the same way that Weight Watchers helps transform lifestyles around nutrition, Persistence Plus fosters the behaviors and mindsets that lead to college persistence, completion, and success. Persistence Plus uses student success profiles and mobile platforms such as cell phones and iPads to “nudge” students to action and to engage and motivate them to complete college. This process includes interventions for time management and academic setbacks, web-based peer groups for work accountability for work, and software nudges for low test scores or missed study times.

**The Promise of Big Data**

Each of the four tools discussed above uses individual-level data to transform the way higher education is being done today and to provide new data on how it should be done in the future. The key is to allow students access both to their own data and to the user-generated data of their peers.

Big data holds the promise of better solving nettlesome higher education practice problems such as improving the performance of developmental education or transitioning from a community college to a four-year institution. It will allow institutions to tap a rich source of creating value: the students themselves, who will become more self-aware learners. Further, policymakers can use big data to craft evidence-based policies to support innovation and develop performance funding formulas that make sense institution by institution.

The limits on big data in higher education are not found in the technology. The limits lie in organizational cultures that view maintaining the “black box,” with its inaccessible and unknown inner workings, as more important than ensuring student success. But in the end, big data will win. Why? Because big data will help us to extract more value for every dollar we invest in higher education.

**Notes**


Louis Soares is Senior Fellow, Center for American Progress.

© 2012 Louis Soares