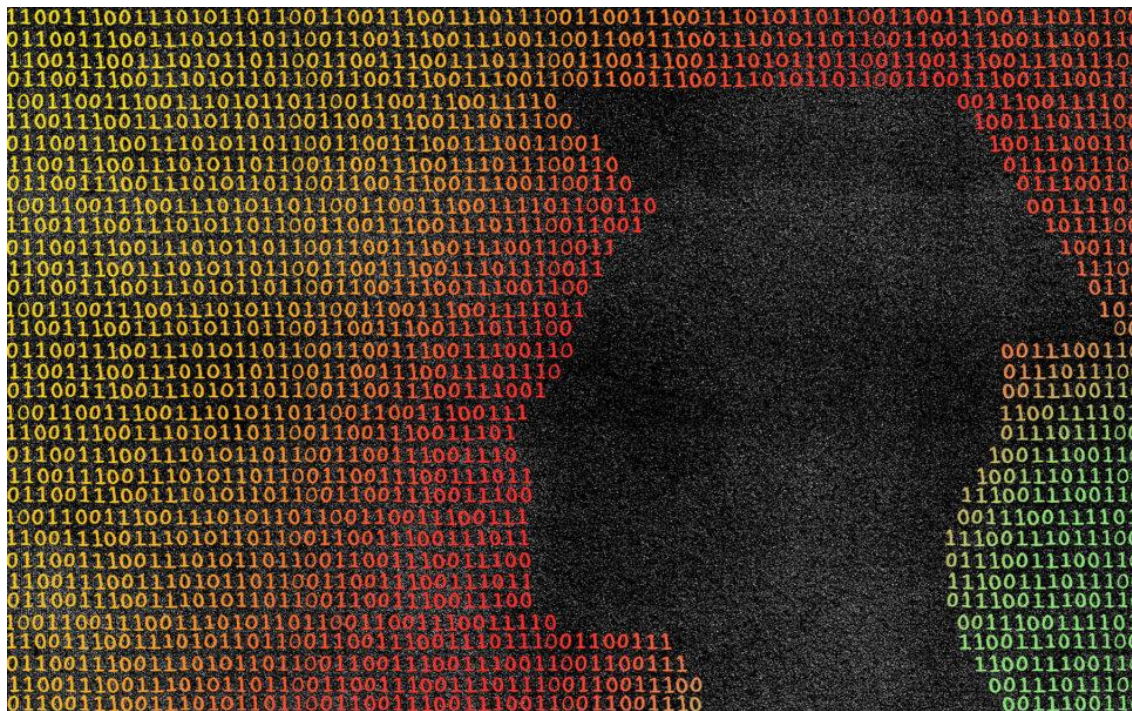


THE HECHINGER REPORT



Divided We Learn

Colleges are using big data to track students in an effort to boost graduation rates, but it comes at a cost

Under a Watchful Eye: critics question if predictive analytics at college invades privacy and reinforces racial inequities

by JILL BARSHAY and SASHA ASLANIAN
August 6, 2019

The Hechinger Report is a national nonprofit newsroom that reports on one topic: education. Sign up for [our weekly newsletters](#) to get stories like this delivered directly to your inbox.



Keenan Robinson works out at Georgia State University's fitness center. School advisers drew on predictive analytics to point him toward a degree in respiratory therapy. Photo: Jill Barshay/The Hechinger Report

This story was produced in partnership with [American Public Media](#).

ATLANTA — When Keenan Robinson started college in 2017, he knew the career he wanted. He'd gone to high school in a small town outside Atlanta. His parents had never finished college, and they always encouraged Robinson and his two older siblings to earn degrees. Robinson's older brother was the first in the family to graduate. "My parents always stressed how powerful an education is and how it is the key to success," Keenan said.

When Robinson arrived at Georgia State University in Atlanta, he wanted to major in nursing. "I always knew I had a passion for helping people," he said. Biology had been his best subject in high school. "My dad, my mom would always kind of call me like the king of trivia because I'd always have just like random science facts," he said.

During his freshman year, Robinson earned a B average. But the university was closely tracking his academic performance and knew from 10 years of student records that Robinson wasn't likely to make the cut for the nursing program.

Georgia State is one of a growing number of schools that have turned to big data to help them identify students who might be struggling – or soon be struggling – academically so the school can provide support before students drop out.

In meetings with his academic adviser during the second semester of his freshman year, Robinson said he learned that though his GPA was solid, the school's computer algorithm saw trouble. Georgia State's analytics system color codes a student's risk of dropping out and Robinson's file was showing yellow, a sign that his plan to go into nursing was risky. At Georgia State, students need to apply to the nursing program at the end of their sophomore year. His adviser told Robinson he would need at least a 3.5 GPA – a high B+ average – to be admitted. Robinson's grades were a little short of that.

“Once they know how they have options, it's not the end of the world,” said Joshua Reaves, an adviser at Georgia State. He sometimes steers students like Robinson into another healthcare major that accepts students with lower grades. “They stay on track and we still get them to graduation,” Reaves said.

Advisers may have crushed Robinson's nursing dream long before he ran into any obstacles but he doesn't see it that way. Robinson was pointed toward a related major, respiratory therapy, that he likes.

“I have asthma,” he said. “I know both my grandparents on my mother's side had lung issues. And once it kind of clicked, it was something I wanted to do.”

Georgia State is encouraging students like Robinson to change their majors in an effort to address one of the most debilitating problems in higher education: low graduation rates.

Nationally, only about half of students who start college actually earn degrees, according to the National Student Clearinghouse, a nongovernmental organization that tracks college enrollment. Many drop out mired in debt and lacking earning power. In fact, because of crushing debt, students who drop out are often worse off than if they'd never gone to college.



For the past four years, black students have graduated at the same or higher rates than white students at Georgia State University. Photo: Jill Barshay/The Hechinger Report

To address this problem, an estimated 1,400 colleges and universities like Georgia State are turning to a high-tech data solution called predictive analytics. The idea is to find trends and patterns in huge amounts of historical data and use those patterns to predict the future.

Companies like Amazon and Netflix have been using data tools like these for years to track our clicks and steer us to buy or watch more of their products. In higher education, colleges are using analytics to keep students enrolled and continue collecting tuition dollars.

Most of the schools using predictive analytics systems aren't elite universities, where low graduation rates aren't a big problem. Instead, analytics is reshaping the college experience for students at less selective schools, putting them on a narrow, data-driven path to graduation, with fewer dead ends and wrong turns. If the data can actually help them earn a diploma, students will have a better shot at successful careers and middle-class lives.

There are indications that analytics may be having an impact. In 2016, after years of declines, national college graduation rates started [ticking back up again](#) and have continued rising for the past three years.

But that progress may come at a cost. Critics say there are potential downsides to monitoring student data so closely. They worry about invasion of privacy

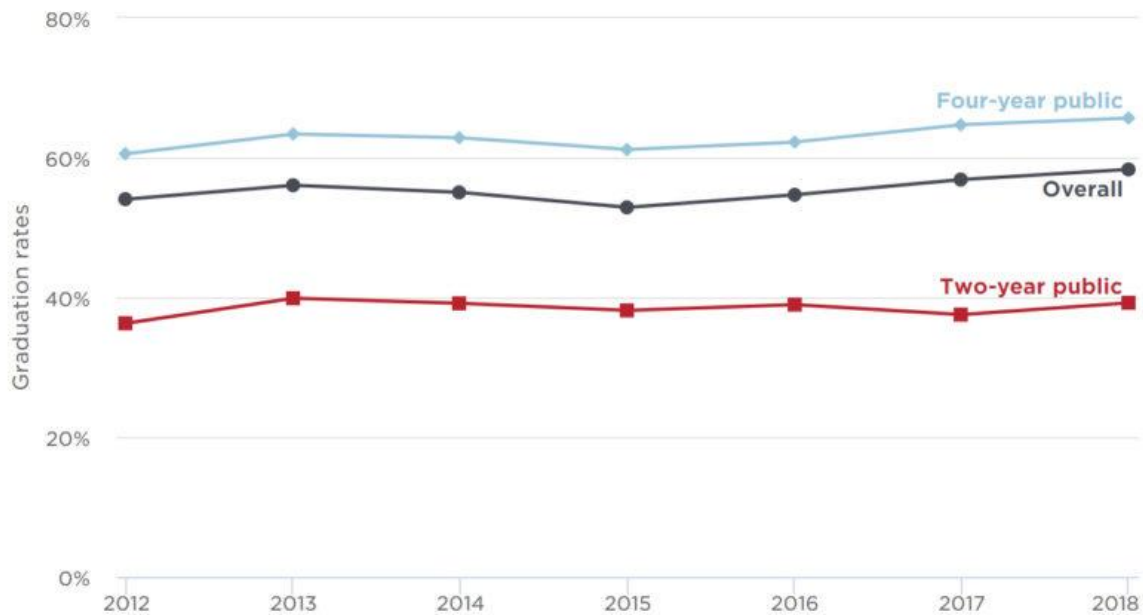
and surveillance of students. And they say that the algorithms might be reinforcing historical inequities, funneling low-income students or students of color into easier majors. There are also fears that the data could unintentionally discourage students and prompt some who might have otherwise stayed in school to drop out.

Earning a profit

What's clear is that supplying predictive analytics to higher education institutions has become a big business. James Wiley is a technology analyst with Eduventures, which does consulting work for companies in the predictive analytics industry. Wiley says that colleges are typically paying \$300,000 a year for these data dashboards and a third of all higher education institutions have bought them. It's grown into a \$500 million market with more than 30 for-profit companies selling predictive analytics tools to colleges.

The companies sell a data mining service that employs artificial intelligence to hunt for patterns in millions of past student records. Then it matches the patterns to students who ultimately dropped out. For an absurd example, if dropouts tended to take classes on Thursdays in their first semester at college, but students who completed their degrees didn't, then you might worry about current students who are currently taking classes on Thursdays. That's how the models make predictions of which students are currently at risk. Past behavior of former students drives predictions for current students. The companies claim this kind of pattern analysis can help colleges pinpoint which students are veering off course and help them when there's still plenty of time for them to get back on track to finish college.

Schools are buying these tools for a good reason – too many students are dropping out. Only about 55 percent of students who started at a four-year school in 2008 had bachelor's degree six years later. At two-year colleges, which more than a third of the nation's college students attend, it's much worse. Less than 40 percent were getting associate degrees after six years.



Percentage of students who completed their degrees within six years of starting college. After years of decline, the rates have begun to rise. Data source: National Student Clearinghouse Graphic credit: Andy Kruse/APM Reports

It wasn't always this way. The dropout problem got a lot worse in the 1990s when more people started attending college. Young adults who used to go straight from high school to the factory floor were suddenly on college campuses.

"There was a focus on getting students into college but not a focus on getting them through college," said Iris Palmer of the New America, a left-of-center think tank.

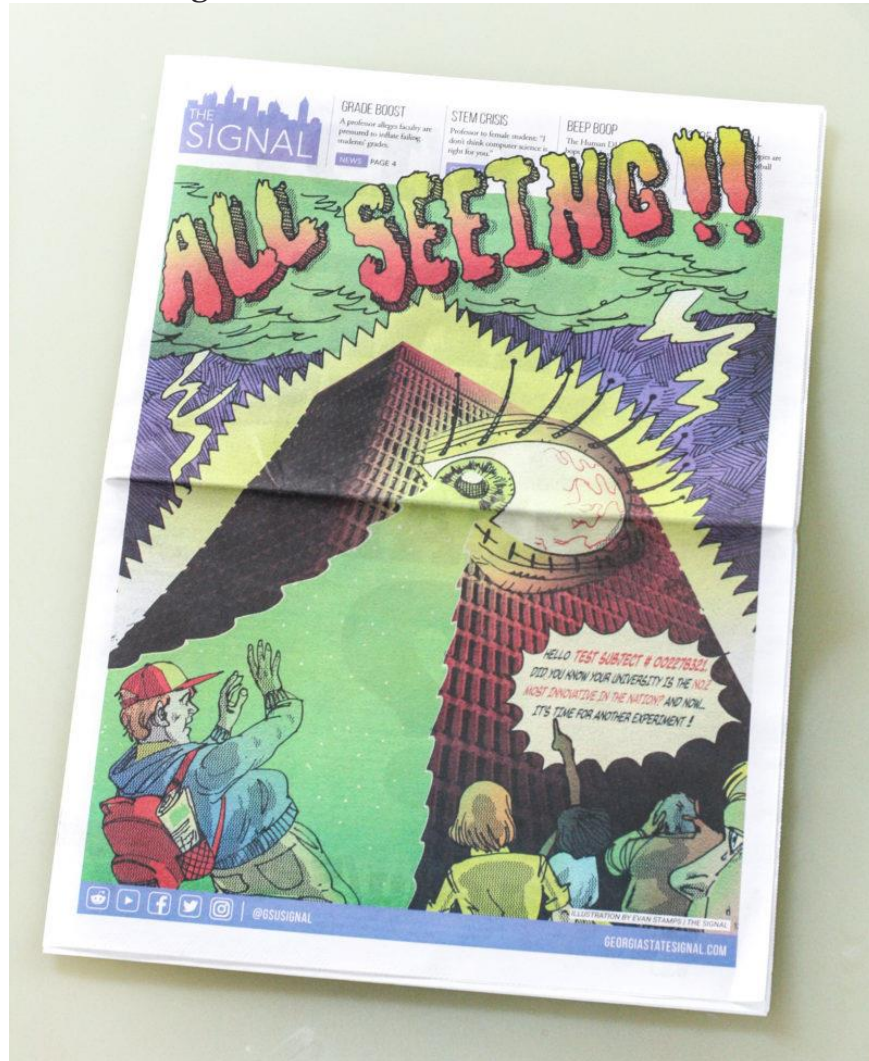
Low-income students were more often dropping out with debt than getting a degree. They were taking on thousands of dollars in student loans to pay for their college educations, but they were far less likely to get a degree than their wealthier peers.

Eventually, the Obama administration took colleges to task for their poor results. At the time, public budgets were reeling in the aftermath of the 2008 recession, and states were cutting funding for public universities. Tuition was rising. With the price of college going up, school officials were under pressure to increase graduation rates.

Around the same time, philanthropic foundations weighed in. Bill Gates called upon colleges to track student progress. "Without measurement, there is no pressure for improvement," he said in a 2009 speech to the National Conference of State Legislatures.

The Bill and Melinda Gates Foundation gave grants to colleges to buy data tools and software. Other foundations such as Kresge and Lumina did too. (The Gates and Lumina foundations are among the funders of the Hechinger Report, and APM Reports receives support from Lumina as well.)

Companies saw a business opening to soak up these foundation grants. In fact, one of Gates's foundation officers, Mark Milliron, who had been giving out these grants, left the foundation to found one of the first companies to sell analytic tools to colleges.



Georgia State's student newspaper, The Signal, devoted a special April 2019 issue to how administrators were monitoring student data. A poll by the student reporters found that many students didn't know Georgia State was using predictive analytics to measure their risk of dropping out. Illustration by Evan Stamps for The Signal

“They can’t afford to hire their own data scientists,” Milliron said. “They can’t afford to basically buy all the hardware to be able to run these data systems.

Let's launch an initiative to... really help them think about how they could use data to help students."

When Milliron pitched his start-up, Civitas Learning, to investors, he had to explain that he was going after the bottom half of the college market.

"I still remember having a conversation when people are saying yeah we should go get Stanford and Duke," Milliron recalled. "I remember saying, they're not going to be interested in this at all. They're basically a hospital who only took healthy patients. So they're not the ones we're going to focus on."

Other businesses had the same idea. Advisory Board – a consultancy to hospitals – was starting to think that it could do the same data crunching for higher education.

It created a spinoff called EAB, headquartered in Washington, D.C. It's now one of the biggest players in the market along with Austin-based Civitas Learning and Cincinnati-based Hobsons.

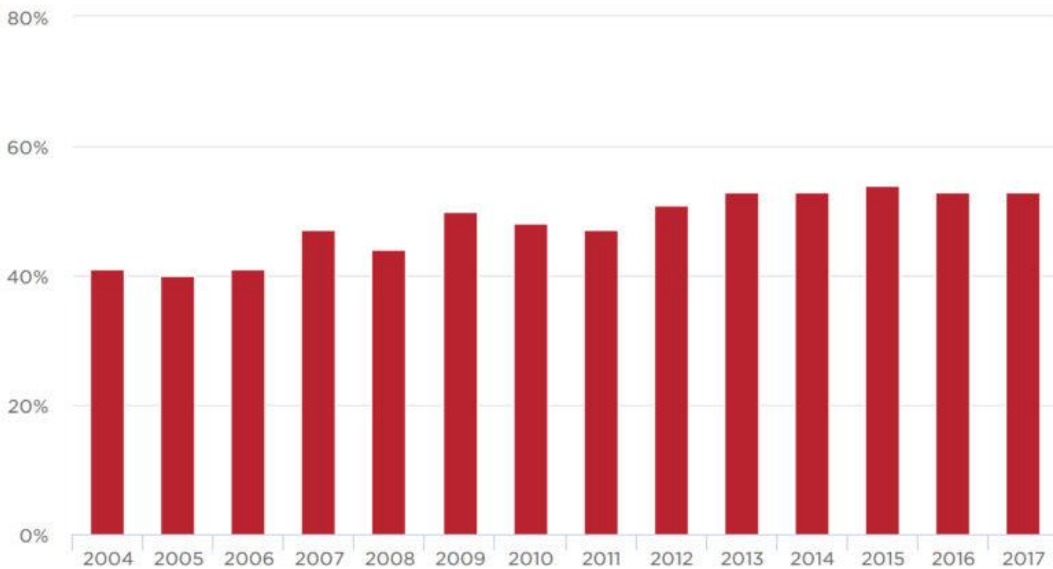
"We were just borrowing the technological resources from the healthcare side of the business," said Ed Venit, a managing director at EAB and one of its first employees.

By 2012 the conditions were ripe for the predictive analytics industry to take off on campuses around the country. Government had created pressure for accountability. Foundations had given seed money. And new companies were popping up to sell data analytics packages.

The poster child

One of the early adopters was Georgia State, which proponents of predictive analytics often hold up as a success story. Located in downtown Atlanta with more than 50,000 students, it's Georgia's largest university but not as prestigious as the Georgia Institute of Technology two miles away, or the state's flagship, the University of Georgia in Athens.

It was an all-white commuter school before black students were allowed to enroll in 1962. By 2006, the school was majority minority. As the college accepted less prepared students and students from poor backgrounds, many of whom were the first in their families to attend college, graduation rates dropped. Around that time, only 18 percent of the black men who enrolled at Georgia State graduated. White students were graduating at a higher but still comparatively low rate of 32 percent.



Percentage of Georgia State University students who completed their degrees within six years of starting college. Georgia State agreed to start using analytics in 2011. Data source: IPEDS Graphic credit: Andy Kruse/APM Reports

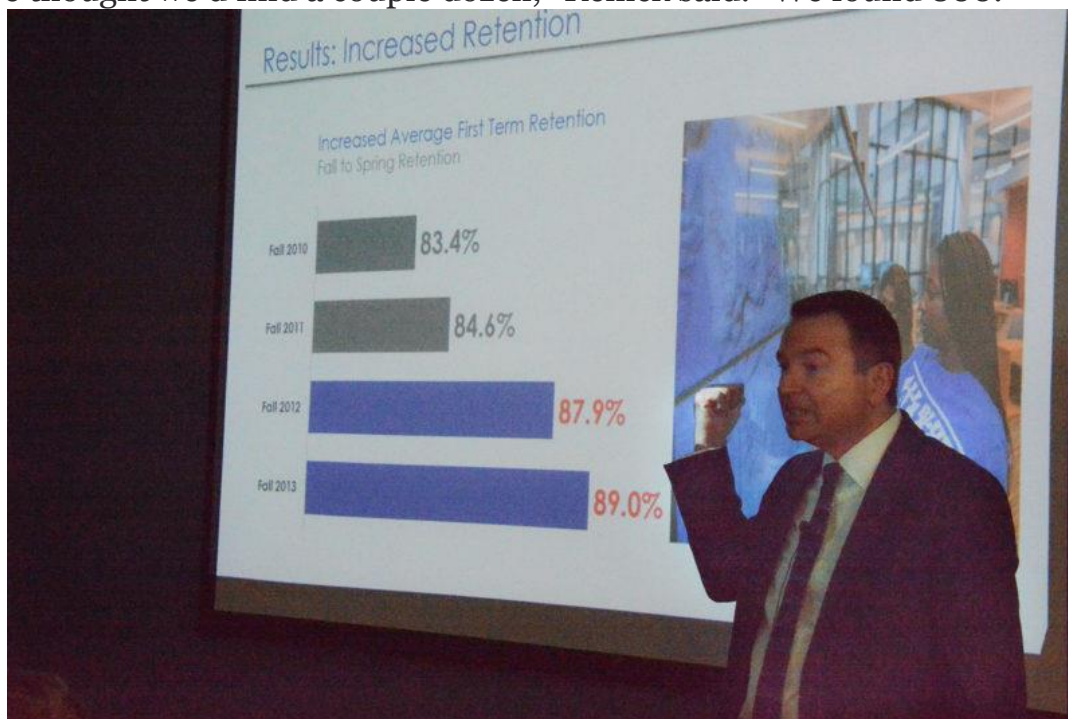
Georgia State was enrolling more and more students who were less likely to graduate, recalled Tim Renick, senior vice president for student success. Renick came to Georgia State as a religious studies professor in the 1980s. He eventually moved into administration, and his job was to improve Georgia State's bleak graduation numbers, especially for low-income and first-generation students.

Renick didn't have many resources. Georgia State had only one adviser for every 1,000 students. And the students who typically showed up in advisers' offices were either honors students, who were managing just fine, or failing students, who were often coming in too late.

Renick needed to reach the students in between. "What would an advising system look like that targeted the students in the middle that really served the B-minus or C-plus students who mostly are sailing under the radar screen?" he wondered. "Those are the students we realized who could move our graduation rate."

He suspected that data could help pinpoint exactly which middling students could be helped and when. Renick says he signed a contract with EAB 2011 before it even had a prototype for the software. The company began sorting through 10 years of Georgia State data and 140,000 student records.

“We ran this big data set to try to figure out what were the advancing indicators that a student might drop out six months, 12 months, a year later? We thought we’d find a couple dozen,” Renick said. “We found 800.”



Georgia State University senior vice president for student success Tim Renick tells a group of visitors how predictive analytics boosted retention and graduation rates. January 16, 2019. Photo: Jill Barshay/Hechinger Report

Renick is referring to 800 “marker courses.” These are classes that students take early in college that predict whether a student can actually pass the toughest required courses that will come later in their field of study.

Among the 800 marker courses are classes like introductory English, typically taken in the fall of freshman year. The data showed that students who graduated from the nursing program mostly had a B or better in that class. A B-minus might seem like a decent grade, but it’s actually a giant warning sign for an aspiring nurse. Students with B-minuses in their first English course often didn’t do well in biology, chemistry and anatomy requirements later on, didn’t make it in the nursing program and dropped out.

That same B-minus turns out to be just fine for accounting majors, who have graduated even if they got a C in intro to English. It’s not that an English grade is causing a student to succeed in accounting or fail in nursing, but it’s just a correlation in the historical data.

A human would likely never spot such complex patterns. But with the use of data analytics, advisers at Georgia State now can.

The human element

The analytics are just a tool, though. The other key element for boosting graduation rates at Georgia State are the advisers. These are the professionals who help students map out their college plans and navigate the university bureaucracy. Since 2010, the school has tripled the number of student advisers, Renick said. A decade ago, there was one adviser for every 1,000 students; now that ratio is down to one adviser for every 500 students.

Keenan Robinson, now majoring in respiratory therapy, met with his adviser one day last January. He was sitting in a small office of a 25-floor office tower in downtown Atlanta that used to be the SunTrust Bank building before Georgia State bought the property.

Joshua Reaves, his adviser, had Robinson's data on his computer screen. Robinson had a 3.1 GPA, a solid B student. It also showed Robinson's probability of graduating. The rating is based on how he's done in the marker courses, if he's taken the right courses for his major and whether he's gotten flagged for skipping class or flunking quizzes. The system sorts students into three risk categories and color codes them, like a traffic light: green, yellow and red.

Robinson could see his yellow light on the screen. "It doesn't really scare me," he said. Robinson said the yellow light reminds him to keep up with his work.

Before it turned to big data, Georgia State probably wouldn't have advised Robinson to abandon nursing. But the university's data analysis showed that many students who didn't get into the nursing major ultimately dropped out of college because they had wasted their first two years taking prerequisites that didn't count toward another major. It was critical to steer these students — many of them with above average grades — to something else early in their college careers.

It's up to each adviser whether to display these risk indicators during an advising session. Reaves says he usually doesn't show a high or moderate risk warning to a student initially. Instead he uses the information to draw the student into a longer conversation. "In my advising style, I just don't say you're here because you did poorly on your first exam," Reaves said. Instead, when he sees a student with a yellow light, Reaves asks more questions until the student starts talking. "I want you to open up to me, and then I'll kind of help you along the way," he said.

Georgia State's computers also scan for students who have good academic records but are late paying their tuition bills — an indication that the student might drop out for financial reasons. Just before the registrar would normally

drop a student from the rolls for nonpayment, the school now intervenes by giving the student a “grant” to pay the bill, effectively giving the student a last-minute tuition discount. It’s only for small outstanding bills; the average grant is \$900.

But the giveaway is financially wise for Georgia State. The university makes money on any student who stays in school and continues paying tuition. Georgia State has issued more than 13,000 of these “retention” grants over the past six years.

But the predictive analytics system isn’t catching everyone. Some students who were failing classes and taking longer in their majors than necessary said they didn’t get alerted. It was unclear if the algorithms simply overlooked them or if the students had overlooked emails from advisers.

Racial bias?

One complication is that race and ethnicity are closely intertwined with graduation rates. Historically, African-American and Latino students have lower graduation rates than white students.

Georgia State doesn’t enter a student’s race or other demographic information into its algorithms but many other schools do. Some observers worry that the predictive analytics tools might be reinforcing historical inequities.

“There is historic bias in higher education, in all of our society,” said Iris Palmer of the New America Foundation. “If we use that past data to predict how students are going to perform in the future, could we be baking some of that bias in?”

Palmer says that the algorithms can unintentionally target black and Latino students because they’re hunting for patterns of dropping out of college, such as low grades and missed assignments. Black and Latino students might have more of these dings in their records than white students. And they could disproportionately be flagged as high risk.

“And so what will happen is they’ll get discouraged, like why are they even trying?” she said. “What that could end up doing is being a self-fulfilling prophecy for those particular students.”

At Georgia State, Renick said he’s heard the fear that minority students could be shunted into easier majors. But he said the reality is more students than ever are graduating in the toughest majors, thanks to better support, like

tutoring. Low-income students who are the first in their families to attend college have “benefited the most,” he said.

He added that the number of students graduating with majors in STEM – science, technology, engineering and math – has doubled since the introduction of predictive analytics. But some STEM majors are tougher than others. In the competitive nursing program, only a third of the graduates are black even though they make up more than 40 percent of students at the university. In Keenan’s program, respiratory therapy, black students are overrepresented and make up almost half the students who graduate in the major.

Georgia State officials point out that predictive analytics aren’t determining who gets admitted into the nursing program. Faculty still make the selection. But the faculty selection occurs late in a person’s college career, often after two or three years, which increases the chances of them dropping out. The predictive analytics tool attempts to anticipate this faculty decision a year or two earlier, the university explained, so that a student who is not likely to be accepted into nursing still has resources to graduate in another program.

The graduation rate for students who are rejected from the nursing program is now much higher, the university says.

But while more black students are graduating overall, they may be getting degrees in fields with less earning potential. For instance, according to federal data, if Keenan Robinson becomes a respiratory therapist, he’ll be in a field that earns, on average, \$11,000 less per year than nurses.

Showing results

Georgia State pays \$150,000 a year to EAB, its predictive analytics vendor. That’s a heavily discounted rate because the school signed up early. Renick argues that the data system more than pays for itself. He calculates that every percentage point increase in the graduation rate is worth \$3 million dollars a year in additional tuition and fee revenues. Since 2003, Georgia State has raised its graduation rates by 23 percentage points, which adds up to more than \$60 million.

“These were students who were paying customers who were walking away from the university who now are actually getting their degrees,” Renick said. “So you talk about win-win situations.”

It might be a stretch to attribute every improvement in the graduation rate to mining student data. Perhaps the increase in advisers alone, without the data,

would have accomplished these improvements, though Renick says the data helps the advisers to be effective.

The school has greatly reduced the disparities in graduation rates. Blacks, Latinos, first-generation college students, and low-income students who qualify for federal Pell education grants now graduate at the same or even higher rate than the student body overall for the last four years, Renick says.

“This institution in the shadow of the Martin Luther King district that was segregated well into the 1960s ... now is conferring more bachelor’s degrees to African-Americans than any [nonprofit] college or university in the United States,” Renick said in a presentation to a group of visiting administrators from colleges around the country.

Other schools, including the University of South Florida and Arizona State University, have also seen their graduation rates rise after turning to predictive analytics.

But that success isn’t universal. The graduation rate at Montgomery County Community College outside of Philadelphia improved slightly with the use of analytics but it’s still only 20 percent. Meanwhile, at Strayer University, a for-profit school where 90 percent of the students take their courses online, graduation rates remain painfully low: 16 percent. The vast majority of Strayer’s students attend school part-time and it’s too soon to tell if their graduation rates are improving. But since implementing predictive analytics, Strayer says fewer students are dropping out from semester to semester.* At both schools, the student population is dominated by older, working adults. At Montgomery County Community College, administrators said many students have complicated lives and financial problems. Mental health issues are common.* There’s a limit to how much the data can help.

Students don’t know they’re being monitored

Though Georgia State has generated a ton of national coverage for its use of predictive analytics and banners around campus tout that it’s the second most innovative university in the nation, the use of big data isn’t on students’ minds.

“Most students are not aware that this is a thing and it’s taking place,” said Ada Wood, an editor at Georgia State’s student newspaper, *The Signal*.

The student journalists recently dug into what Georgia State was doing with their data in a special issue in April. On the cover is a cartoon drawing of a giant green eye bulging out of a campus building and staring down at students on the bottom of the page. It evokes a “big brother vibe,” Wood said.

A poll by the student reporters found that many students didn't know Georgia State was using predictive analytics to measure their risk of dropping out. Some students interviewed worried about their privacy.

Georgia State is bound by federal student privacy laws, which aim to safeguard grades and personal information from public view but it can share and discuss student data with EAB, its predictive analytics vendor, as long as it's for a legitimate educational interest.

In its contract with Georgia State, EAB promises not to disclose confidential student information to a third party. That means it cannot sell student data to marketers who want to tap into the college consumer market. There's a little wiggle room for EAB to disclose the data, as long as students cannot be identified. EAB compiles its customer data to report on national trends in college completion. Other vendors do, too.

Georgia State has recently expanded the kind of student data it monitors. In 2018 the university began tracking how often each student connects with campus wifi and logs into the school's computer system. It doesn't track what students are doing on their cellphones or which websites they're visiting, just how often they're connecting. Georgia State calls it "an electronic footprint." If a student's pattern suddenly changes – say the student stops coming to campus or logging into class websites – that can be an early alert that a student is stumbling.

The potential downside

It might seem intrusive to track digital behavior, but some researchers go further. At the University of Arizona, a professor looked back at the ID card swipes of first-year students when they were new on campus. The ID records included when students entered classroom buildings, returned to the dorm at night, checked out library books or bought a coffee.

A machine learning algorithm detected patterns of social engagement on campus and matched them to those who dropped out. The professor could predict with 85 to 90 percent accuracy who would drop out based on their first few weeks of activity on campus.

This was a one-off research experiment. But the results were tantalizing and show that personal data contains incredible clues of how well students will do.

But Kyle Jones, an assistant professor in the Department of Library and Information Science at Indiana University who's studying the ethics of

predictive analytics in education, says he's concerned about the implications of the Arizona study.

"They were doing this with good ends in mind," Jones said. "They did it to develop new retention models and figure out where they could potentially improve services to increase retention. That's a good thing but at the expense of creating a pretty significant surveillance system."

For Jones, the ethical problem is that college surveillance or monitoring of students' grades and computer clicks isn't being openly and clearly explained to students.

"Students aren't aware," Jones said. "There's really no culture of informed consent when it comes to educational technologies and higher education. These decisions are often made by intermediaries like chief information officers, instructors and advisers. Decisions are made on behalf of students."

The algorithms can generate a lot of false positives, signaling risk for a student who will be fine, or overlooking students who actually are at risk, he said. The for-profit analytics companies don't publish their models or open them up to scientific scrutiny.

Jones also worries what's being lost when algorithms are directing students through college. He argues that data can steer students too narrowly, limiting the mistakes that might have opened unexpected doors.

"I always say higher education is a time of exploration," he said, "where you start to figure out who you want to become for the later period of your life."

A growing trend

There's another element that the analytics can't measure — human ambition and the drive of an individual student to succeed. On a large scale, the algorithms will mostly be right in their forecasts. But with each individual student, there can be wide variance. Perhaps Keenan Robinson could have succeeded in nursing at Georgia State. We'll never know.

A growing number of schools have decided the risks and potential downsides are worth it. After years of declines, college graduation rates started ticking back up again for the first time in 2016 and continued ticking up for three years in a row, to 58 percent for students who started in 2012, according to the latest report from the National Student Clearinghouse. It's unclear if those improvements can be attributed to predictive analytics, but it could be a factor.

Regardless, the use of predictive analytics on college campuses will likely increase in the coming years. Driving the market is demographics. Americans stopped having so many babies after the 2008 recession and the fertility rate hasn't recovered since. As a result, the college-age population is expected to decline by 15 percent after the year 2025.

With a smaller pool of potential students, it won't be as easy for universities to replace each year's dropouts with new students. So schools will have a greater incentive to hold onto students.

Meanwhile, the analytics approach appears to have helped Robinson. He learned on June 25 that he was formally accepted into Georgia State's respiratory therapy program. He could have been another dropout casualty. Instead, he's on track to graduate.

**Update: This paragraph was updated with additional information from Strayer University after initial publication. Officials revised its calculation of its graduation rate for full-time students from 10 percent to 16 percent.*

**Correction: An earlier version of this story implied that students at Strayer also have drug addictions and mental health issues. Strayer officials said the school does not have such data about its students.*

Main illustration: Rachel Sender for APM Reports

This story about [predictive analytics](#) was produced by [APM Reports](#) and [The Hechinger Report](#), a nonprofit, independent news organization focused on inequality and innovation in education. Sign up for [Hechinger's newsletter](#).



Jill Barshay

Jill Barshay is a staff writer and editor who writes the weekly “Proof Points” column about education research and data. She taught algebra to ninth... [See Archive](#)

[@jillbarshay](#)

<https://hechingerreport.org/predictive-analytics-boosting-college-graduation-rates-also-invade-privacy-and-reinforce-racial-inequities/>