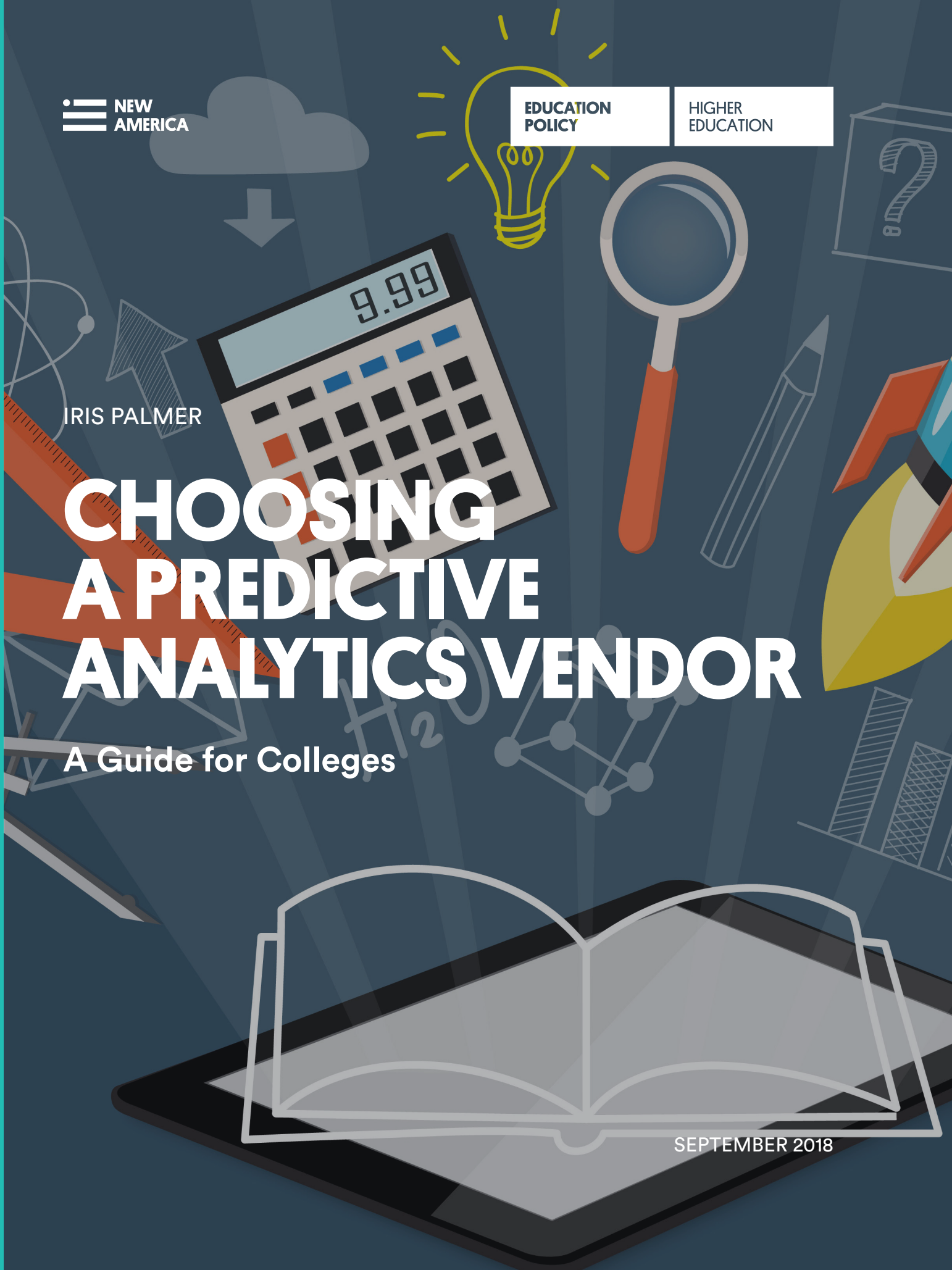
NEW AMERICA

EDUCATION POLICY | HIGHER EDUCATION



IRIS PALMER

CHOOSING A PREDICTIVE ANALYTICS VENDOR

A Guide for Colleges

SEPTEMBER 2018

About the Author(s)

Iris Palmer is a senior policy analyst with the Education Policy program at New America. She was previously a senior policy analyst at the National Governors Association.

About New America

We are dedicated to renewing America by continuing the quest to realize our nation's highest ideals, honestly confronting the challenges caused by rapid technological and social change, and seizing the opportunities those changes create.

About Education Policy

We use original research and policy analysis to help solve the nation's critical education problems, crafting objective analyses and suggesting new ideas for policymakers, educators, and the public at large.

About Higher Education

New America's higher education program works to make higher education more accessible, innovative, student-centered, outcomes-focused, and equitable.

Acknowledgments

This paper benefited from the insights of many, including the interviewees listed in the appendix. I am especially grateful for the review and feedback from of Sarah Zauner of the Ada Center, Karen Vignare of the Association of Public and Land-grant Universities, Gina Johnson of the Association for Institutional Research, and Ellen Wagner of George Mason University. Thanks to New America colleagues Steve Bird and Sabrina Detlef for their editorial insight. Maria Elkin and Riker Pasterkiewicz provided layout and communication support. This work would not have been possible without the generous support of the Kresge Foundation and the Bill & Melinda Gates Foundation. The views expressed in this paper are the author's alone.

Contents

Contents Cont'd

Introduction

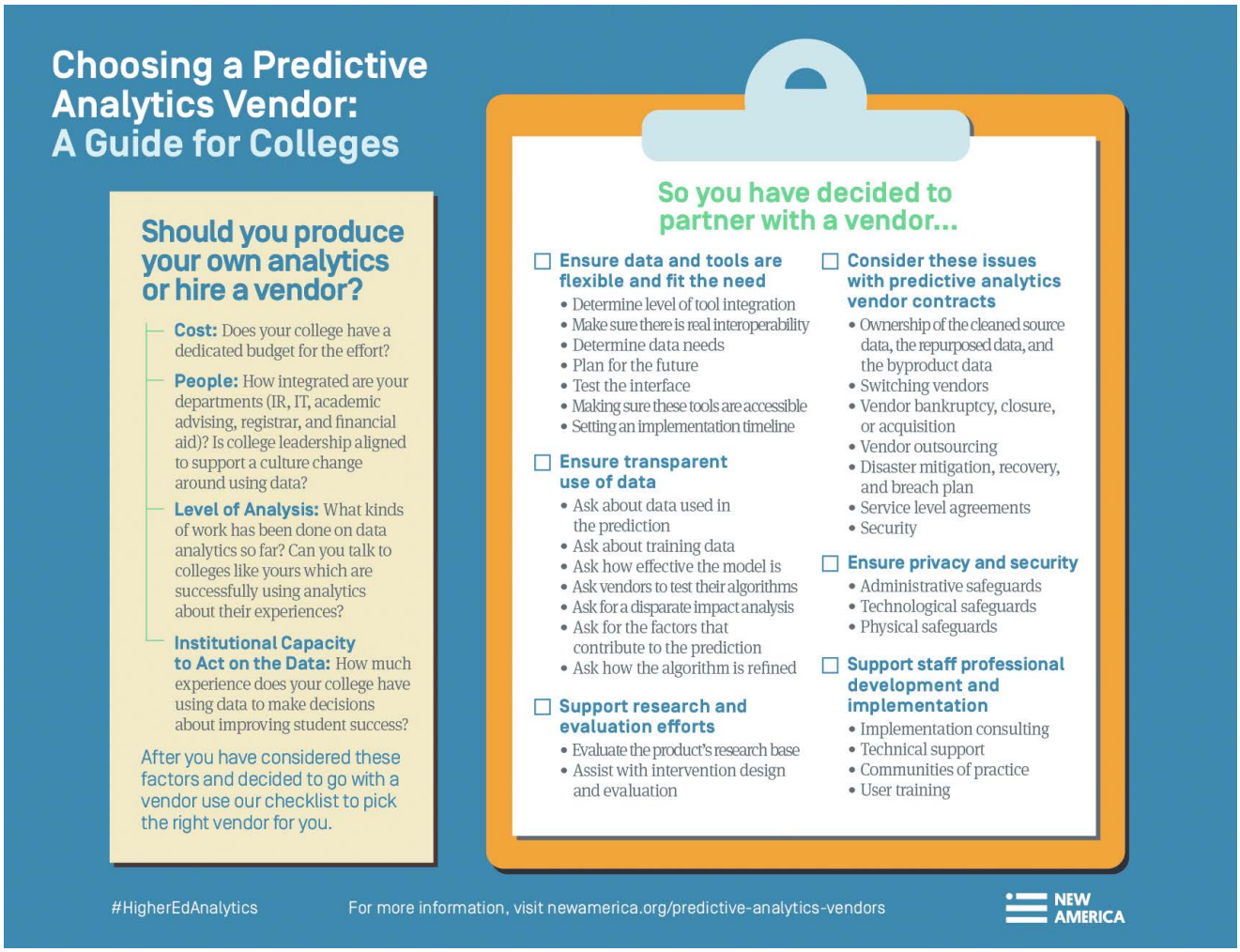
Colleges are increasingly using models to predict student behavior and intervene to change that behavior. Because of this, when projects involve partnering with a vendor, it is more important than ever to make the right choice about which vendor. In some ways, partnering with a vendor to use predictive analytics is similar to procuring any other technology product. But the complexity of the algorithms—and the predictions they produce—add another layer to the decision-making process. As we talked to a number of colleges, we heard from some whose vendors had algorithms that showed all of their students were "at risk." And we heard from others who were told that student GPA was the most significant predictor of student success. Neither of these are useful outcomes or productive partnerships.

Currently, many colleges choose vendors based on word of mouth. Colleges should certainly talk to other schools about their vendor experiences and ask vendors for their customer retention rates. But in this era of big data, colleges have additional ethical considerations.

Ideally, predictive analytics vendors can facilitate the ethical use of data all of the way through the student life cycle. Vendors can help ensure that data are complete and integrated correctly to diminish the chances of misidentifying students. They can be transparent about their algorithms and test them for disparate impact on student populations. They can be flexible with permissions and use reasonable security protocols to help preserve student privacy and security. They can help evaluate interventions to ensure they are not harming certain groups of students. And they can train staff on the correct interpretation of data and on the dangers of implicit bias.

Not all vendors, however, take this approach. Many do the bare minimum with data integration. Some consider their models and algorithms proprietary. Some are not focused on student privacy and security. Others do not support training and evaluation.

This guide will give administrators the tools to ask the right set of questions of predictive analytics vendors and provide a sense of what kind of answers they should expect. It will focus on ensuring that vendors use predictive analytics tools like early alert systems in an ethical manner.



Institutional Readiness

Partnering with a vendor will not solve all problems around data and analytics. Instead, colleges must first determine what they want out of these tools, put the capacity in place to maintain and act on the information these tools provide, and get buy-in from key members of the school community for using these tools. In 2017, we published ***Predictive Analytics in Higher Education:***, a guide for colleges looking to properly implement these tools with ethics in mind.[2] Before colleges think about partnering with a vendor, we recommend reading our guide, as well as exploring other useful resources, such as the paper **Putting Data to Work**, the book ***Creating a Data-Informed*** , the **Data Maturity Framework** from the University of Chicago Center for Data Science and Public Policy, and **The Ada Center**, which is dedicated to helping community colleges implement technology solutions.[3]

In the end, no vendor is going to do the hard work for your college. Many of the stories we have heard about unhappy partnerships between colleges and vendors come from misunderstanding what vendors and their products can unilaterally change on campus.

Produce Your Own Analytics or Hire a Vendor?

As colleges begin to look at using increasingly sophisticated analytic systems and see the price of partnering with a vendor, they inevitably think about using their existing institutional research (IR) and information technology (IT) resources to do the analysis themselves. But only a small number end up following this path. A recent estimate found that colleges spent about $22 million on in-house technology solutions in 2017 out of a total market of $858 million. Still, only between 1 and 3 percent of colleges have chosen this path.[4] Ivy Tech[5] and San Jacinto College[6] are examples of community colleges that have created their own systems. Arizona State University,[7] the University of New Mexico,[8] and the University of Texas system[9] are examples of public four-year colleges that have created their own in-house solutions.

By their nature, these systems are highly individualized, customized to the needs of the colleges that created them. And even these schools seek some outside help. The University of Texas system, for example, is partnering with Salesforce to create a solution for its schools. San Jacinto College is using SAS Enterprise Miner. Ivy Tech used an open source algorithm. The University System of Georgia is working to code its system in the coding language R.

→ COMMUNITY COLLEGE DATA ARE DIFFERENT

Colleges interested in making use of predictive analytics should look for vendors that are familiar with their sector. Many vendors have worked predominantly with four-year colleges. The community college context is different. For instance, community college students tend to be older, are more likely to have families, and are more likely to be low-income. These students need
different kinds of interventions than the typical student at a selective four-year institution, and they often have different goals. Some are just coming to get a few skills to support progress in their career, while others are aiming to transfer to four-year colleges. While many vendors use the same data elements, their predictive algorithms are based on outcomes for students that align more with goals of students at four-year universities. Any vendor should have models that are flexible enough to allow community colleges to advise this diverse set of students.

And community college data are also different than those of four-year colleges. Community colleges offer many more types of credentials including diplomas, certificates, associate degrees meant to transfer to four-year colleges, and

terminal associate degrees which are not designed to transfer. Some community colleges also offer stackable credentials that allow students to earn multiple credentials in the same academic year. These schools also enroll many more part-time and developmental education students than does the typical four-year college. Community colleges also typically have large non-credit course offerings.

All of these characteristics make community college data different and highlight the importance of finding a vendor that has experience working with schools in this sector.

So how do you know if your college should build a system in-house or partner with a vendor?

We have created a framework to help colleges think through this question. First, officials need to determine what level of experience their institutions have using data. To do that, consider the following questions:

- **How integrated are your departments (IR, IT, academic advising, registrar, and financial aid)?** The more connected and integrated the key departments are on campus, the more advanced the institution is in using data flexibly. Collaborative and healthy working relationships among these departments make a substantial difference in this work.
- **How much experience does your college have using data to make decisions about improving student success?** A college with regular meetings and committees charged with examining and making decisions based on data has a greater chance of achieving its aims.
- **Does your college have a dedicated budget for the effort?** Colleges with more resources to dedicate for this purpose are in a superior position to those that are trying to make it work with existing resources.
- **What kinds of work has been done on data analytics so far?** Colleges that have started to experiment with more sophisticated analytics have a higher capacity to use data than those just starting out.
- **Can you talk to colleges like yours which are successfully using analytics about their experiences?** A college that has access to a community of similar schools that have already gone through the process will be in a better position to experiment with analytics.
- **Is college leadership aligned to support a culture change around using data?** Having leadership on board to support the changes needed to use data will make implementation much simpler.

Once colleges have decided on their capacity for using data, they should work through the following considerations.

Cost

One of the main things that makes internal solutions attractive is their relatively low price. Even obtaining a license for an enterprise solution colleges can customize can be much cheaper than partnering with one of the large vendors. According to the Ada Center and Achieving the Dream, the cost of using a vendor can range from $30,000 to $100,000 annually, with a one-time implementation fee sometimes approaching $75,000.[10] The cost depends on a variety of factors including the number of students in the college, the number of licenses, the number of deployed products, and the sophistication of the data integration.[11] It is not surprising, then, that colleges would be attracted to doing this with their existing staff, particularly because existing staff will generally need to partner closely with a vendor to make any solution work. However, in-house solutions are not as low-cost as they may seem. Colleges need to take into account staff time, salaries (particularly if they need to hire new staff), and hardware needed to make an internally built solution work. This effort can be costly no matter what solution the college chooses so it is important to take all cost factors into account when deciding whether or not to hire a vendor.[12]

People

Another consideration for colleges is the capacity of staff and the culture of their IR and IT offices. Some IR offices are more focused on compliance reporting than using data to support student success efforts. And some IT offices concentrate on procurement and tech support rather than using data systems to help administrators serve students better. These offices need to be entrepreneurial and able to collaborate to create an internal system and make it work. They also need to focus on anticipating the next institutional needs and generating answers to those questions in user-friendly ways. In addition, the staff in these departments need to have the necessary expertise and training to be able to generate and interpret models. Finding individuals with this type of expertise can be difficult because statisticians and data scientists are in high demand in the private sector. And this understanding of data cannot just be in the IR and IT offices. Understanding how to use data needs to be a skill set across departments, including student affairs, academic affairs, financial aid, and the registrar.

This demand can lead to another problem: staff turnover. If a college relies on key staff to create a customized solution and those individuals leave, the school may not have anyone who knows how to maintain and update the internal system. To avoid this problem, colleges must ensure that the internal analytics system they

are creating is not too customized, that the customizations are documented, and that other staff members are trained on how to maintain it. A sign of how difficult it is to sustain these customized solutions is the number of them that have been spun off or sold, which is done for complicated reasons including financial gain and challenges involved with maintaining a customized system in the face of staff turnover. Sometimes companies acquire promising new models to accelerate new opportunities and sometimes they acquire them to kill new competition. Some examples of spin-offs include Purdue's Course Signals, which was licensed to **Ellucian** and the University of Maryland University College analytics tool, which was **spun off** into HelioCampus.[13]

Level of Analysis

A key question for colleges to consider is what they need the tool to do. If the college wants to evaluate interventions that take discrete, nimble analysis, it can be easier to keep that analysis in-house rather than trying to work through a third party. If the college wants to undertake a more sophisticated analysis with a significant visual component, the college might want to consider partnering with a vendor. It is common for colleges to experience big student success gains when these interventions are deployed that level off over time. More sophisticated algorithms, comparison data, and additional data sources can help boost student success if it hits a plateau.

Capacity to Act on Data

It might make sense for some colleges to start with simpler, in-house analytics and build the capacity to act on the data from there. Partnering with a vendor is not a silver bullet for those with limited experience using data analytics to make decisions. Instead, colleges must do the hard work of making their institutions ready to support and act on analytic insights. Practicing with their own data analysis can help improve institutional readiness. For instance, one useful finding from analyzing the colleges' own data is discovering all of the interventions the college already has in place and being able to assess those outcomes. To support this work, the Association for Institutional Research has created a **Statement of Aspirational Practice for Institutional Research** which sets a roadmap for developing analytics.[14] Of course, partnering with a vendor is not all or nothing. Colleges can outsource certain aspects of their predictive analytics and keep others in-house. This "smart-sourcing" strategy may be more complex to implement but may also allow colleges to have more flexible solutions.

So You Have Decided to Partner with a Vendor...

Once a college has decided on the functionality it needs and has done the work to ensure it is ready, it may decide to partner with a vendor. To ensure the predictive analytics solution is implemented in a sustainable, ethical way, colleges should consider the following when selecting a vendor:

- Ensure that data and tools are flexible and fit the need
- Ensure transparent use of data
- Issues with predictive analytics vendor contracts
- Ensure privacy and security
- Supporting research and evaluation efforts
- Supporting staff professional development and implementation
- Return to introduction

Ensure that Data and Tools are Flexible and Fit the Need

Depending on their goals and profile, colleges have different needs when partnering with a predictive analytics vendor. Those needs will also change over time. Schools need to think through the following set of technical and ethical considerations to ensure that the vendors' data and tools are flexible enough to meet those needs.

Determine Level of Tool Integration

It is important to look at the vendor landscape and decide how much integration the college should be looking for in the product. It is tempting to partner with a vendor that has all of the capabilities the college may ever need so that it would not have to go through another procurement process if it needs additional functionality. As a result, the trend in the marketplace is to build more horizontally integrated solutions. But a solution tailored to your current problem often works better. That is what colleges indicated in a recent Tyton Partners' survey, which found that the majority of schools believe less integrated solutions perform better than solutions with more integration.[16] Colleges need to assess the pros and cons of partnering with a vendor of integrated solutions.

Make Sure There is Real Interoperability

Vendors all say they are interoperable with major university data systems, but what that means can vary widely. Is there out-of-the-box data integration or does the vendor need to create a program interface? Creating and maintaining an interface can mean a lot of time and money.[17] A good product must be able to receive data from many different sources, possibly through a robust set of application programming interfaces (APIs).[18] One way to ensure that a vendor will truly integrate with a college's data systems is to check if that company is **compliant** with the **Learning Tools Interoperability** (LTI) and the **Caliper Analytics** standards set out by IMS Global, a non-profit member collaborative that develops open interoperability standards.[19] Another possibility is to partner with a vendor that specializes in supporting data integration. These products, like **Dxtera**, **Jitterbit**, and **AcademyOne**, facilitate secure information exchange between various technology tools and the institution's data and can help support long-term integration.[20]

Determine Data Needs

Chances are colleges, particularly institutional research offices, will have to play a large role in cleaning college data so they are ready to put a predictive analytics model in place. Each college is unique in the way it codes and stores its data and college staff will most likely have a hands-on role in helping vendors accurately interpret institutional data. When choosing vendors, colleges should ask for a realistic picture of how much staff time this will take and ask to talk with similar colleges to check these claims.

When a college lays out a plan for using analytics, it should consider the types of data the school community is comfortable using to predict student outcomes. For instance, will the college be comfortable using location or student financial aid data or is that information too sensitive to include in these systems? When searching for a vendor, the college should make sure that it is comfortable with the types of data the system uses. Most systems include data from the student information system (SIS), the customer relationship management (CRM) software, and the learning management system (LMS). Some also integrate data from other early alert or case management systems; location data from swipe cards or wireless internet connections; student membership in clubs and other organizations; tutoring/support systems; attendance records; financial information; and adaptive learning platforms.21

Colleges should also ask whether additional types of data increase the predictive power of the algorithms the vendor is using. It is particularly important to acknowledge the tradeoff between how closely the data reflect a student's current situation and how intrusive it might seem. While historical data like high school GPA and demographic information might seem less sensitive then say location data, that information also might fail to predict a student's current circumstances and/or lead to profiling while also failing to provide enough information to guide effective intervention.

It is also important for colleges to look at how often the data are updated in the system. The more the data reflect the current actions of students, the better the chance to intervene and put students on the right track. In addition, some data elements may need to be updated more often than others. Colleges should check to see if the vendor they are considering has the capacity to update data on different schedules.

Plan for the Future

Colleges will most likely find multiple vendors that have products with features that align with what the college currently needs. But they should also be looking to the future. While colleges hire vendors based on an initial set of needs, those needs will change over time. Colleges should create a plan for how they expect to use predictive analytics over the next three to five years. Then, they should ask prospective vendors about their future upgrade plans and compare the vendors' plans to their own to help determine which company might be the best fit. At the same time, colleges should ask about vendors' flexibility in building out new functionality. Is there an opportunity to build more functionality if there is aggregated demand among colleges? How much demand is enough to trigger building new functionality? A successful partnership will have flexibility built in to adapt to those changing requirements.[22]

Colleges should also check how the vendor system affects their current work processes. The faculty members and administrators on campus who will be the frontline users of the predictive analytics product should be on the purchasing committee to flag where the processes of the college may not line up with the requirements of the software. For instance, if a college does transcript review and degree planning with students before they apply and the predictive analytics tool requires a student ID to work, that tool will not be useful.

Test the Interface

The software interface can be one of the most difficult things to get right. If the tool is clunky or does not answer frontline staff questions effectively, no one will use it. As a result, it is critical for colleges to ask about the vendors' processes for testing the interface. Did they test it with real-world practitioners? Vendors should also allow colleges to test the tools with their own faculty and staff. As more and more of us do our jobs on the run, colleges should also ask about access via mobile device. If the system contains sensitive information and it is available on mobile devices, how is it protected in the mobile environment? Lastly, as more of these systems contain learning artifacts, does the system interface with open education resources?

Make Sure These Tools Are Accessible

Colleges have an ethical and legal responsibility to make sure these systems are accessible for people with disabilities. This is particularly true for testing compatibility with screen readers.[23] The vendor should be able to articulate exactly how the system supports accessibility requirements for staff who need accommodations under the Americans with Disabilities Act. Ideally, accessibility support should be a focus rather than an afterthought.[24] And colleges should not take the vendor's word for it. Instead, they should test the systems with their staff to ensure it is accessible to all. Colleges should also include accessibility in their contracts with vendors.[25]

Set an Implementation Timeline

Implementation tends to take longer than expected because of data cleaning and integration. It also takes a while to get frontline staff members comfortable with using the tool in their everyday work. Colleges should ask vendors about their typical implementation timeline but know that the process will most likely take longer. If the vendor runs a community of practice, reach out to other clients to get a realistic sense of the effort and time it took to get the tool up and running. Colleges can get a sense of some of the typical implementation timelines in the vendor profiles in the publication ***Crossing the Finish Line: Vetting Tools that Support Student Success***.

Ensure Transparent Use of Data

One of the biggest frustrations for colleges in choosing a predictive analytics vendor is trying to peer into the companies' algorithms, which are all too often a black box. Administrators want to know how the system comes up with its predictions but do not know which answers are reasonable to expect. Confusion over what vendors should be able to explain about the algorithm is particularly true for administrators who do not come from a technical background when they talk to salespeople who also do not entirely understand the system. The information asymmetry between the technical staff at the vendor and college can lead to misunderstandings and bad decisions. Colleges have an ethical obligation to know a certain amount about how the algorithms they are applying to their students work to ensure some accountability. Here are the things colleges should insist on knowing about the vendors' algorithms:

Ask about Data Used in the Prediction

Ask the vendor to provide a list of all the data elements it uses in the predictions. With this information, the college can decide the sensitivity of the data elements and document where that data is stored. And the documentation of data elements also provides an important baseline for what the predictions are based on. Vendors should commit to providing the school with an updated list of variables they use and alert the institution if they change.

Ask About Training Data

Colleges should ask vendors if they use data from other colleges to inform their models. Some vendors use data from across multiple schools to make their models more robust. But it is important to note that one way bias creeps into algorithms is through non-representative training data. Ask vendors to document the diversity of the types of data, the colleges, and the students that helped to train their models. Colleges should also ask how vendors plan to customize the algorithm to their specific school. Many vendors create the model with only the college's data. In that case, administrators should ask how far back the training data goes. It is important to balance having the data be as recent as possible to reflect the current college population with ensuring there is a valid sample size to train the algorithm. College administrators should also ensure that the makeup of the college population has not changed significantly over that time.

Ask How Effective the Model is

Colleges should know that overall accuracy rates can be misleading and easily manipulated. Instead, colleges should ask vendors for an AUC (area under the curve)/C statistic. A C statistic indicates how accurate the model is. For most models, colleges should be looking for a C statistic of .75 or greater.[27] But watch out. If the C statistic is .99 or higher, there is most likely a problem with the model like leakage (where the data used to train the algorithm includes the data you are trying to predict) or overfitting (where the model is so tuned to the sample data that it fails to generalize to the larger population). Colleges can also ask for a Brier Score.[28] A Brier Score is a way to measure the accuracy of a probability forecast, like a weather forecaster saying there is an 80 percent chance of rain. On this scale a lower score is better. Anything under .22 is a solid model. Administrators should also bring colleagues with technical experience into these conversations. Vendors should also commit to presenting on these measures and the overall results of the model at regular intervals, such as every time they rebuild the model.

Ask Vendors to Test their Algorithms

Some vendors will agree to conduct a pilot to test their algorithm with the college's data. This can help indicate if the system is discriminatory by overidentifying as "at risk" certain groups of students on campus or if it identifies all of a college's students as "at risk." A pilot project can also help highlight where the integration of the institution's data may need to be improved. Before signing a long-term contract with vendors, college administrators should ask if they are willing to conduct a pilot. Vendors may not be willing to conduct a pilot because they may need to do all the work with the data up front to show what they are capable of. In that case, administrators should ask for a de-identified example of analysis they did for a comparable school.

Ask for a Disparate Impact Analysis

If vendors need to agree that they will conduct a disparate impact analysis of their tool output after getting the contract. If a vendor does over identify a certain group as "at risk", that may be because we live in an imperfect world where certain groups are actually more at risk of failing. With the knowledge that a group of students is being overidentified, the college should come up with a plan to focus support on that population. The vendor also needs to ensure that its algorithm is useful in the college's particular context. For example, is the algorithm identifying every student as "at risk" and if so, how can the vendor ensure this will be fixed?

Ask for the Factors that Contribute to the Prediction

It can be a tough sell to convince staff members who have worked with students for years to trust the predictions of an algorithm. This is particularly true if there is no transparency into why it is making its prediction. Having vendors list the most weighted factors that went into the prediction for a particular student on the dashboard or interface can bring transparency that helps users act on the prediction. However, it is also important to note that these factors are not causal. For instance, if one reason for the prediction is that the student did not enroll in very many credits that is not necessarily the reason the student is at risk of dropping out. It may just be an indication of something else going on in the student's life.

Ask How Often the Algorithm is Refined

As we noted in *Predictive Analytics in Higher Education: Five Guiding Practices for Ethical Use*, what colleges learn from their use of predictive analytics should be used to refine the algorithm. There should be a cycle of improvement. This can be difficult if the vendor does not support this type of systematic refinement. Models should generally be rebuilt at least once a year and as often as once a term. More frequently than once a term should not be necessary if it is a solid model and more than a year risks degradation of the model.

→ TOOLS FOR EVALUATING THE ALGORITHM

Aequitas (dsapp.uchicago.edu/aequitas): An open source bias audit toolkit to audit machine learning models for discrimination and bias, and make informed decisions around developing and deploying predictive risk-assessment tools.

Algorithmic Impact Assessments (ainowinstitute.org/aiareport2018.pdf): A Practical Framework for Public Agency Accountability. See page 15 for how to conduct a disparate impact analysis.

Issues with Predictive Analytics Vendor Contracts

As with any vendor, colleges need to be sure that a contract with a predictive analytics company protects their students' privacy and maintains the college's ownership of that data. This is particularly critical when working with vendors that use cloud-based computing. All vendors must, at a minimum, meet the legal requirements of FERPA but there are other things to consider above and beyond that legal standard.[29] Although colleges' legal and IT departments probably know what contracts with vendors should look like, here are a few things administrators should keep in mind. This list is not exhaustive but it highlights important components to consider.

Ownership of the Cleaned Source Data, the Repurposed Data, and the Byproduct Data

In some cases, the vendor will not agree to return the cleaned source data to the college, instead destroying it at the end of the contract. The contract should clearly lay out who owns what data and under what terms.[30] Of course, it should also prohibit the vendor or any partner of the vendor from selling student data.

Switching Vendors

To prevent vendor lock in, colleges need to have a plan for how they will switch vendors, if they need to, with the minimum amount of pain. To facilitate any transition, the contract should spell out the college's right to access the data and lay out the process by which the vendor will return the data to the college at the end of the contract. The agreement should include the time frame for the data's return as well as the format in which the data should be returned. The contract should also obligate the vendor to destroy and verify the destruction of any personally identifiable information when the contract ends.[31] Colleges should consider including transition assistance in the contract to ensure a smooth transition to a new solution.[32] The contract also needs to address the end of the relationship, laying out timelines and each parties' obligations beyond the ones listed above.[33]

Vendor Bankruptcy, Closure, or Acquisition

In the relatively new space of predictive analytics, there are bound to be acquisitions and closures. The contract needs to address that possibility. Colleges can find model language in the mergers and acquisitions section in this article ***If It's in the Cloud, Get It on Paper***.[34]

Vendor Outsourcing

Many vendors work with subcontractors or other partners. The contract should require them to disclose those partners and their roles in providing services. It should also clarify that the vendor remains responsible for executing all aspects of the contract with the college.[35]

Disaster Mitigation, Recovery, and Breach Plan

The contract should also outline how the contractor would respond to a disaster and outline a recovery plan, including how the vendor would respond if data are lost or corrupted. The section should detail how the vendor would notify the college, how it would correct the problem, and how it would continue to provide service. The contract should also outline the vendor's plan of action if there is a data breach. And it should lay out who is responsible for what action and where the liability sits for the breach.

Service Level Agreements

Service level agreements (SLAs) lay out the acceptable levels of service for each element of the service provided.[36] For example, if a product has an unacceptable outage, the penalty would be addressed in this section. The SLAs should also outline what the vendor will do if the expected level of service is not met.[37] Examples of SLAs for the University of California system can be found in this **University of California Cloud Computing Service Agreement Template** (starting on page 11).[38]

Security

The contract should clarify the level of security expected by the college. See the next section for more detail.

→ EXAMPLES OF CONTRACT LANGUAGE

Texas State University Contracting Issues Checklist (bit.ly/2nTtD8B)

University of California Cloud Computing Service Agreement Template (bit.ly/2Prn61m)

Ensure Privacy and Security

Ensuring that privacy and information security protections are in place is fundamental to the ethical use of data and thus to choosing a vendor. Most CIOs and IT departments have a good handle on what standards vendors should meet to merit access to student data. In fact, in the **Higher Education Cloud Vendor Assessment Tool**, there are over 300 information security and privacy questions for vendors to answer about their products.[39] However, there are a few things for non-technical decision makers to keep in mind. Vendors should be able to easily answer questions about the administrative, technological, and physical security and privacy safeguards they employ. If they struggle to answer these questions, that could be a signal to dig deeper into their security and privacy practices.

Administrative Safeguards

Administrative safeguards are the management statements that an organization puts into place to protect information security and privacy. Colleges should ask vendors about who can view the institution's data in the vendor's product and stored in its IT systems and how the companies control that access. As human error is one of the biggest sources of security breaches, it is important to check that vendors are training all of their staff, including technical staff, on security procedures and ethics. Colleges should ask if the vendor keeps a log of who has access to records so that an audit can be conducted to ensure that employees are not inappropriately viewing records to satisfy their curiosity about a friend or family member. Keeping these logs shows how the vendor views verifying how its staff is handling data and allows for forensic analysis if something goes wrong. Vendors should also meet all appropriate data standards like those from the National Institute of Standards and Technology. The Higher Education Cloud Vendor Assessment Tool shows how all of these standards apply to security practices.

Technological Safeguards

Technological safeguards are the controls implemented in a college's or vendor's IT systems that protect those systems and the data located in them. There are a number of good technological practices for ensuring student data security, including industry-standard encryption and secure data connections. Ask if the data are encrypted using an industry-standard encryption both when stored at the vendor and when transferred between the institution and the vendor. Ask whether the vendor has passed an industry-standard security scan and if so, which one. Also ask if the vendor's employees have remote access to the data and check if the tool can mask certain sensitive data from certain types of users, which allows colleges to present data only to staff members with a compelling need to know.

Physical Safeguards

Physical safeguards are the controls put into place to keep unauthorized individuals out of controlled areas (e.g., data centers) or from accessing IT systems or data. These include controls designed to protect physical media like laptops, servers, storage media, and mobile devices. For example, Washington State University found that a hard drive containing personal information on over 1 million students had been stolen from a locked storage container in spring 2017.[40] Incidents like this show how securing the location of your data is an important

consideration. Ask vendors if data will be stored in a physically secured location. It is also worth asking if employees are allowed to take data home or on travel; both increase chances of theft or carelessness that could lead to a data breach.

→ SECURITY AND PRIVACY RESOURCES

University of California: Learning Data Privacy Practices and Learning Data Privacy Practices (bit.ly/2nVQMqU)

EDUCAUSE: Higher Education Cloud Vendor Assessment Tool (bit.ly/2tfTs6o)

Vendor Security Technical Assessment of Risk (V-STAR) (bit.ly/2LiscKk)

Data Privacy Addendum (bit.ly/2BwlkJt)

Protecting Student Privacy While Using Online Educational Services: Model Terms of Service (bit.ly/2nUFSBO)

EDUCAUSE: Information Security Guide: Effective Practices and Solutions for Higher Education (bit.ly/2w39Esz)

Supporting Research and Evaluation Efforts

To use predictive analytics tools well, colleges will need to continuously improve their intervention efforts. At the same time, the tools themselves need to be externally validated through research.

Evaluate the Product's Research Base

Many vendors say that their products are evidence-based. But it is important for administrators to look critically at that research and see if it is independent and high quality. For instance, has the research been published in a peer reviewed journal? Organizations like **Teachers College** at Columbia University partner with vendors to increase and strengthen their evidence base.[41] There are also a series of tools, mostly designed for the K–12 environment, which could help colleges evaluate the evidence (see box below).

→ **TOOLS FOR EVALUATING ED TECH RESEARCH**

Evaluating Studies of Ed-Tech Products (bit.ly/1WEIpwz)

Ed Tech Rapid Cycle Evaluation Coach (edtechrce.org)

Ed Tech Developer's Guide (tech.ed.gov/developers-guide)

Using Research in Ed Tech (bit.ly/2aZQPPM)

Assist with Intervention Design and Evaluation

Once colleges have access to a predictive analytics tool, what should they do to change the trajectory for students? It can be helpful for vendors to support the design and evaluation of a college's efforts intervene with students. Many vendors have experience with messages and interventions that worked in other college contexts. And they can also help evaluate how interventions perform by helping create well-matched control groups and processing the data. Since evaluation is an important part of using predictive analytics well, colleges should ask vendors how they can support these efforts.

Supporting Staff Professional Development and Implementation

Changing the way your faculty and staff do their jobs to incorporate a new tool can be a heavy lift. It is important to assess how the vendor will help overcome that hurdle by supporting implementation and providing technical support.

Implementation Consulting

Some vendors offer implementation consulting as part of the solution package. Colleges should consider whether they need this type of assistance given their experience level with data-led-decision-making because including it will increase the price of the product. Some vendors even assign the college a dedicated consultant to help them implement the tool. If that is the case, colleges should ask about the consultant's experience with schools in their sector. A consultant that has worked with only small private liberal arts colleges may not be helpful to a large public flagship, for example.

Technical Support

It may seem like a small detail but establishing who is responsible for which aspects of technical support is crucial for a successful rollout. Colleges need to know what technical support the vendor provides and who will be responsible for day-to-day troubleshooting. For example, when a high level administrator loses her password, who is responsible for quickly getting her access? If users do not have access to support for their questions, they will not use the system. It is also important to establish how the vendor will help after the launch with adjustments to the system.

Communities of Practice

Many vendors bring their active users together, forming a community of practice. These communities are cultivated online and at annual conferences or regional meetings. If a college hires a vendor that has such a community, the school should explore it and see if it will be useful for the frontline users and others at the institution. These communities can provide helpful insight into what it is actually like to implement the tool and use it daily. While participants in a community of practice are self-selecting, they can still give colleges a sense of what using the vendor's products is actually like.

User Training

Well-designed training is key to helping faculty and staff adopt a new tool. Most vendors will train a group of people on the campus and have those people train others. This model scales more quickly and can get everyone on board in an efficient way. If the vendor has agreed to participate in a pilot project, a core group of people with knowledge and buy in has a head start on training. That initial training should also include staff members from the school's institutional research office, so they can get a better sense of how the tool works and how it will fit into the college's information ecosystem. Training and assistance should also be embedded in the system so users can get help as they face challenges with using the new software. Ideally, the training should go beyond just the technical and should touch on how these tools can change the way administrators do their jobs and also show how to interpret data appropriately, with implicit bias in mind.

Conclusion

Colleges are continuing to take advantage of the increases in student success resulting from both home-grown and purchased predictive analytics systems. Our framework should help guide them through the complex choices they face as they try to decide which type of system to implement. Colleges need to consider the cost, their people, the ability to act on data, and the level of analysis needed when considering building a system in-house. If partnering with a vendor, colleges need to consider the flexibility of the tool, transparency, contract terms, privacy and security, research and evaluation, and professional development. Each of these considerations is important to creating a functional vendor/college relationship.

Appendix: Interview List

Laurie Heacock, vice president of data, technology and analytics, Achieving the Dream

September 27, 2017

Fiona Hollands, associate director and senior researcher, Center for Benefit-Cost Studies of Education, Columbia Teachers College

October 2, 2017

James Wiley, principal analyst, Eduventures

October 3, 2017

Ellen Wagner, former vice president of research, Hobsons

October 18, 2017

Scott James, data scientist, Hobsons

October 18, 2017

Lige Hensley, chief technology officer, Ivy Tech

October 18, 2017

Karen Vignare, executive director, Personalized Learning Consortium, APLU

October 20, 2017

James Cousins, senior statistical analyst, Rapid Insight

October 20, 2017

Scot Henley, vice president of Sales and Marketing, Rapid Insight

October 20, 2017

Amelia Parnell, vice president for Research and Policy, NASPA

October 23, 2017

Alexis Wesaw, director of data analytics, NASPA

October 23, 2017

Christine Keller, executive director, Association of Institutional Research

October 30, 2017

Darlena Jones, director of assessment & research, Association of Institutional Research

October 30, 2017

Dror Ben Naim, founder and CEO, Smart Sparrow

October 30, 2017

Laurie Fladd, associate dean of science and mathematics, Trident Technical College

October 31, 2017

Celeste Schwartz, vice president for information technology and college services, Montgomery County Community College

November 6, 2017

Sarah Zauner, founder, The Ada Center

November 8, 2017

Rahim Rajan, senior program officer, Gates Foundation

November 15, 2017

Vicki T. Tambellini, president and CEO, The Tambellini Group

November 20, 2017

Ed Venit, managing director, EAB

November 27, 2017

Jason Elwood, president, Elwoodway Consulting

December 7, 2017

Ellen P. Goodman, professor, Rutgers University Law School

December 8, 2017

Karen Ferguson, vice-provost, CSU Global

January 4, 2018

Dale Allen, president and co-founder, DXtera

January 9, 2018

Brendan Desetti, senior manager, Government & Stakeholder Relations, D2L

January 17, 2018

Shady Shehata, head of AI Platform & Data Science, D2L

January 17, 2018

Martin Kurzweil, director of Educational Transformation, Ithaka S+R

January 19, 2018

John Whitmer, learning analytics and research director, Blackboard

January 23, 2018

Derrick Tillman-Kelly, UIA fellow, Ohio State University

January 29, 2018

Timothy Harfield, senior product marketing manager for Analytics, Blackboard

February 8, 2018

Tristan Denley, executive vice chancellor for academic affairs and chief academic officer, University System of Georgia

March 14, 2018

Michelle Callaway, lead research programmer analyst, San Jacinto College

March 29, 2018

George Gonzalez, director of Institutional Research & Effectiveness, San Jacinto College

March 29, 2018

Amanda Vance, director of the Education Privacy Project and a Policy Counsel, Future of Privacy Forum

April 18, 2018

Ross Schulman, senior counsel and senior policy technologist, Open Technology Institute, New America

April 18, 2018

Nat Meysenburg, technologist, Open Technology Institute, New America

April 20, 2018

Glenda Morgan, research director, Gartner Research

May 1, 2018

Joanna Grama, director of the Cybersecurity Initiative, EDUCAUSE

May 9, 2018

Notes

1 Manuela Ekowo and Iris Palmer, Predictive Analytics in Higher Education: Five Guiding Practices for Ethical Use (Washington, DC: New America, 2017), https://www.newamerica.org/education-policy/policy-papers/predictive-analytics-higher-education/.

2 Ellen D. Wagner, "Putting Data to Work," Community & Resources for eLearning Professionals, The E-learning Guild, https://www.elearningguild.com/insights/224/putting-data-to-work/; Brad C. Phillips and Jordan E. Horowitz, Creating a Data-Informed Culture in Community Colleges: A New Model for Educators (Cambridge, MA: Harvard Education Press, 2017), https://www.amazon.com/Creating-Data-Informed-Culture-Community-Colleges/dp/1682530876; University of Chicago, Center for Data and Public Policy, "Data Maturity Framework," https://dsapp.uchicago.edu/home/resources/datamaturity/; and The Ada Center, https://www.theadacenter.org/.

3 Gates Bryant, Jeff Seaman, Nicholas Java, and Kathryn Martin, Driving Toward a Degree: The Evolution of Academic Advising in Higher Education (Boston, MA: Tyton Partners, 2017), http://drivetodegree.org/report-archive/driving-toward-degree-evolution-academic-advising-higher-education/.

4 "How Ivy Tech is Using Predictive Analytics and a Data Democracy to Reverse Decades of Entrenched Practices" (PowerPoint presentation, Ivy Tech Community College, Indianapolis, IN,) https://cdn.oreillystatic.com/en/assets/1/event/261/Learning%20from%20higher%20education_%20How%20Ivy%20Tech%20is%20using%20predictive%20analytics%20and%20data%20democracy%20to%20reverse%20decades%20of%20entrenched%20practices%20Presentation.pdf.

5 George González and Michelle Callaway, "Leveraging Software for Predictive Analytics" (PowerPoint presentation, Enterprise IT Summit, Orlando, FL, March 5–7, 2018), https://events.educause.edu/~/media/files/events/user-uploads-folder/ent18/sess15/leveraging-software-for-predictive-analytics--slides.pptx.

6 "Arizona State University Puts Student Needs First with its [sic] eAdvisor program," YouTube, June 4, 2014, https://www.youtube.com/watch?v=MZb8OniNZbk and Elizabeth D. Phillips, "Improving Advising Using Technology and Data Analytics," Change: The Magazine of Higher Learning 45, no. 1 (2013): 48–55, https://www.wsac.wa.gov/sites/default/files/2014.ptw.%2844%29.pdf.

7 Steve Carr, "UNM's Institute for Design and Innovation Fueled by Technology-Driven Innovation," University of New Mexico, Office of Provost, February 19, 2016, https://news.unm.edu/news/unms-institute-for-design-and-innovation-fueled-by-technology-driven-innovation.

8 Carl Straumsheim, "A Force in the Software Market," Inside Higher Ed, October 3, 2016, https://www.insidehighered.com/news/2016/10/03/u-texas-system-salesforce-team-build-out-learning-platform

9 Case Management and Early Alert Technology Evaluation Resource, Achieving the Dream and The Ada Center, https://static1.squarespace.com/static/58d9c2826b8f5bb9ec49ce86/t/5aa453139140b73db67c35a1/1520718616366/AdaCenterTechnologyEvaluationResource.pdf.

10 Alex Sigillo, Crossing the Finish Line: Vetting Tools that Support Student Success (Burlingame, CA: EdSurge, March 2017), https://go.edsurge.com/rs/590-LFO-179/images/EdSurge_StudentSuccessReport_March2017.pdf.

11 See page 12 of Thomas B. Cavanagh, The LMS Selection Process: Practices and Considerations (Louisville, CO: ECAR, July 8, 2014), https://library.educause.edu/~/media/files/library/2014/7/erb1406-pdf.pdf.

12 Ellucian (website), "Course Signals Solution Sheet," https://www.ellucian.com/Solution-Sheets/Ellucian-Course-Signals/; and University of Maryland, "UMUC

to Spin Off Data Analytics Division Into New Company Providing Business Intelligence Products and Services To Universities Nationwide," press release, September 18, 2015, https://globalmedia.umuc.edu/2015/09/18/umuc-to-spin-off-data-analytics-division-into-new-company-providing-business-intelligence-products-and-services-to-universities-nationwide/.

13 Randy L. Swing and Leah Ewing Ross, Statement of Aspirational Practice for Institutional Research (Tallahassee, FL: Association for Institutional Research, 2016), http://www.airweb.org/Resources/ImprovingAndTransformingPostsecondaryEducation/Pages/Statements-of-Aspirational-Practice-for-Institutional-Research.aspx.

14 Brad Wheeler, "Who Is Doing Our Data Laundry?" EDUCAUSE Review, March 13, 2017, https://er.educause.edu/articles/2017/3/who-is-doing-our-data-laundry.

15 See page 12 & 13 of Gates Bryant, Jeff Seaman, Nicholas Java, and Kathryn Martin, Driving Toward a Degree: The Evolution of Academic Advising in Higher Education (Boston, MA: Tyton Partners, 2017), http://drivetodegree.org/report-archive/driving-toward-degree-evolution-academic-advising-higher-education/.

16 See page 2 of "Checklist for iPASS Predictive Analytics Technology," EDUCAUSE and Achieving the Dream, n.d. https://library.educause.edu/~/media/files/library/2017/4/ipasschecklistpasystem.pdf.

17 Dian Schaffhauser, "Higher Ed Analytics Market Is Growing in Complexity," Campus Technology, March, 30, 2017, https://campustechnology.com/articles/2017/03/30/higher-ed-analytics-market-is-growing-in-complexity.aspx.

18 "IMS Certified Product Directory," IMS Global Learning
Consortium, n.d., https://www.imsglobal.org/cc/statuschart.cfm; "IMS Certified Product Directory Learning Tools Interoperability Certified Products," IMS Global Learning Consortium, n.d. https://www.imsglobal.org/cc/statuschart/lti; "IMS Certified Product Directory Caliper Analytics Certified Products," IMS Global Learning Consortium, n.d. https://www.imsglobal.org/cc/statuschart/caliper-analytics.

19 Robert Graham, William Liddick, David Weil, and Jeffrey Newhart, "Tying It All Together: Integration PaaS in the Next-Gen Enterprise," EDUCAUSE Review, February 12, 2018, https://er.educause.edu/articles/2018/2/tying-it-all-together-integration-paas-in-the-next-gen-enterprise; see page 22 in Gates Bryant, Jeff Seaman, Nicholas Java, and Kathryn Martin, Driving Toward a Degree: The Evolution of Academic Advising in Higher Education (Boston, MA: Tyton Partners, 2017), http://drivetodegree.org/report-archive/driving-toward-degree-evolution-academic-advising-higher-education/.

20 See the integration section of the company profiles (pages 18–51) in Alex Sigillo, Crossing the Finish Line: Vetting Tools that Support Student Success (Burlingame, CA: EdSurge, March 2017), https://go.edsurge.com/rs/590-LFO-179/images/EdSurge_StudentSuccessReport_March2017.pdf.

21 Dian Schaffhauser, "Higher Ed Analytics Market Is Growing in Complexity," Campus Technology, March 30, 2017, https://campustechnology.com/articles/2017/03/30/higher-ed-analytics-market-is-growing-in-complexity.aspx.

22 See page 13 in Thomas B. Cavanagh, The LMS Selection Process: Practices and Considerations (Louisville, CO: ECAR, July 8, 2014), https://library.educause.edu/~/media/files/library/2014/7/erb1406-pdf.pdf.

23 Lindsay McKenzie, "An IT Accessibility Watchdog?" Inside Higher Ed, November 15, 2017, https://www.insidehighered.com/news/2017/11/15/universities-mull-creation-it-accessibility-group-review-vendor-products.

24 See contract language in EDUCAUSE, "Accessibility Contract Language," n.d., https://

library.educause.edu/resources/2016/5/accessibility-contract-language.

25 Alex Sigillo, Crossing the Finish Line: Vetting Tools that Support Student Success (Burlingame, CA: EdSurge, March 2017), https://go.edsurge.com/rs/590-LFO-179/images/EdSurge_StudentSuccessReport_March2017.pdf.

26 "C-Statistic: Definition, Examples, Weighting, and Significance," Statistics How To, http://www.statisticshowto.com/c-statistic/http://www.statisticshowto.com/c-statistic/.

27 Statistics How To (website), "Brier Score: Definition, Examples," http://www.statisticshowto.com/brier-score/.

28 See the first issue in the checklist in Texas State University, "Key Issues in Contracting for Information Technology Resources and Services," Instructional Technologies Support, http://gato-docs.its.txstate.edu/jcr:6127a874-149f-4c05-a89c-9ef2a69715db/Contracting_Issues_Web_Checklist.pdf.

29 See the fifth issue in Texas State University, "Key Issues in Contracting for Information Technology Resources and Services," Instructional Technologies Support, http://gato-docs.its.txstate.edu/jcr:6127a874-149f-4c05-a89c-9ef2a69715db/Contracting_Issues_Web_Checklist.pdf; also see "Ownership of Data" section in Thomas Trappler, "If It's in the Cloud, Get It on Paper: Cloud Computing Contract Issues," EDUCAUSE Review, June 24, 2010, https://er.educause.edu/articles/2010/6/if-its-in-the-cloud-get-it-on-paper-cloud-computing-contract-issues.

30 See "Disposition of Data" section in Thomas Trappler, "If It's in the Cloud, Get It on Paper: Cloud Computing Contract Issues," EDUCAUSE Review, June 24, 2010, https://er.educause.edu/articles/2010/6/if-its-in-the-cloud-get-it-on-paper-cloud-computing-contract-issues.

31 See page 17 for language example in University of California, "Cloud Computing Services Agreement Template," December 8, 2011, http://sites.uci.edu/cloud/files/2014/07/UC-Cloud-Computing-Services-Agreement-Template-for-Distribution.pdf.

32 See page 20 for language example in University of California, "Cloud Computing Services Agreement Template," December 8, 2011, http://sites.uci.edu/cloud/files/2014/07/UC-Cloud-Computing-Services-Agreement-Template-for-Distribution.pdf.

33 Thomas Trappler, "If It's in the Cloud, Get It on Paper: Cloud Computing Contract Issues," EDUCAUSE Review, June 24, 2010, https://er.educause.edu/articles/2010/6/if-its-in-the-cloud-get-it-on-paper-cloud-computing-contract-issues.

34 See "Vendor Outsourcing" section in Thomas Trappler, "If It's in the Cloud, Get It on Paper: Cloud Computing Contract Issues," EDUCAUSE Review, June 24, 2010, https://er.educause.edu/articles/2010/6/if-its-in-the-cloud-get-it-on-paper-cloud-computing-contract-issues.

35 Ibid.

36 Ibid.

37 See page 11 in University of California, "Cloud Computing Services Agreement Template," December 8, 2011, http://sites.uci.edu/cloud/files/2014/07/UC-Cloud-Computing-Services-Agreement-Template-for-Distribution.pdf.

38 "Higher Education Cloud Vendor Assessment Tool," EDUCAUSE, October 25, 2017, https://library.educause.edu/resources/2016/10/higher-education-cloud-vendor-assessment-tool.

39 WSU Stolen Hard Drive Contained Over One Million People's Personal Data, Campus Safety Magazine, https://www.campussafetymagazine.com/university/wsu-stolen-hard-drive-personal-data/

40 Teachers College at Columbia University, TCEdTech (website), http://www.tc.columbia.edu/tcedtech/.

NEW AMERICA