

# **Big Data, Big Challenges: Data-Driven Decision Making as a Core Institutional Practice**

*By Corrie L. Commisso  
University of Illinois Urbana-Champaign*

***“Data! Data! Data!” he cried impatiently. “I can’t make bricks without clay.”***

*Sherlock Holmes, in *The Adventure of the Copper Beeches*, by Sir Arthur Conan Doyle*

Much like Sir Arthur Conan Doyle’s famous character Sherlock Holmes, many leaders and administrators in higher education are calling for their institutions to adopt a data-driven mindset in establishing policies, programs, services, and long-term strategies for institutional growth. While some higher education institutions (HEIs) have already introduced data analysis activities in various decision-making capacities, recent research has shown that adopting data-driven decision making as a core institutional practice is not only proving difficult for most HEIs — but that data science in general is by far the most underutilized resource in the management culture of HEIs (Fong 1).

In an industry where competition for resources (students, faculty, funding) is high, and where institutional success is built or broken on the cornerstone of student success, utilizing data to inform the decision-making process is more than just a good idea: it’s become an imperative. But the collection, and even analyzing, of data in and of itself is not enough to promote the kind of informed decision making that can influence and drive change. Implementing data analytics is only the beginning — data only tells part of the story. The process of moving from data to information, from information to insight, and from insight to action requires collaboration, dedicated resources, and the ability to engage others in understanding and interpreting the data.

This paper seeks to examine some of the barriers for HEIs in moving data-driven decision making from theoretical concept to core practice, and to explore ways HEIs can begin to take initial steps towards making data responsiveness a vital part of institutional culture and identity.

## **Background: Institutional Research and the Rise of Data-Driven Decision Making**

The 21<sup>st</sup> Century is an era of challenge for higher education. With fewer public resources, increasing community needs, and more stringent demands for institutional accountability, HEI administrators are facing complicated decisions that have the potential to build — or burn — their academic communities. Many are turning to their institution’s Institutional Research (IR) departments (or in many cases, their singular IR staff member) to seek information and data that can help them navigate the decision-making process.

While the role of IR at colleges and universities has developed and changed, particularly over the course of the past 20 years, its origins are linked to the explosive growth of higher education in the post-World War II era (Calderon 81). As the government sought to ensure that resources were managed and distributed efficiently, it required HEIs to develop a system of accountability,

one that would require them to establish standards for admission, develop a fee and tuition structure, and assess course viability, among other things — and thus, Institutional Research as a practice was born (Calderon 81).

Today, state and federal governments are pressing HEIs for even more of this statistical data. Beyond just an evaluation of resources and efficiency, government institutions are more frequently requiring demonstrated evidence of institutional performance when making decisions related to allocation of resources. Accreditation agencies are also asking colleges to provide more data related to student learning and achievement (Smith Morest 4).

Government compliance reporting and preparation for accreditation processes are by far the most common focus of college and university IR; however, this data is not always the most helpful or useful when it comes to strategic decision making, and there is a growing demand for more relevant data and analysis at the policy and decision making level (Smith Morest 4). As a result, modern IR is currently experiencing a shift in its institutional role — moving away from a focus on assessment and reporting and into a more active role in policy development, strategic management and planning, and visionary decision making:

IR professionals will need to innovate, multitask, and connect multiple pieces of information together in a coherent manner (foresight) providing a way forward. The likely emphasis of IR professional skills will be more on interpretation and drawing conclusions, rather than on data processing and managing skills. In short, there will be an increasing need for IR units to combine quantitative data analysis with policy and strategic planning. (Calderon 85)

Data-driven decision making is not a new concept for most HEIs — but what *is* new is the nature and volume of available data, as well as the technological tools available to process and communicate it. For many HEIs, the discussion of how to manage data processes comes down to a question of capacity — balancing the available technologies and solutions with the time and resources needed to implement them (Fong 11). The benefits of finding this balance, however are evident in the student and institutional success rates among HEIs that have adopted a data-driven mindset.

### **Data-Driven Decision Making in Practice: Case Studies**

One case study demonstrating the impact of data-driven decision making on policy change comes from the National Collegiate Athletic Association (NCAA). Since the early 1980s, the NCAA has collected data related to injuries of student athletes. The data is collected and maintained by a special committee (the Committee on Competitive Safeguards and Medical Aspects of Sports) dedicated to promoting health and safety for student athletes.

Since the committee began collecting and analyzing injury data, it has used this data to help the NCAA understand contributing factors to athletic injuries and legislate new policies and guidelines designed to mitigate risk factors for student athletes. This is evident in rule changes that have recently been added to NCAA rule books governing 19 sports (Klossner 59). In general, the NCAA rule books only address factors that officials can reasonably control during a

game; however, after incorporating the committee's injury data into their discussions, the NCAA made significant changes related to student athlete safety (Klossner 59). Some of the more significant changes include:

- Data on concussion injuries in ice hockey leading to new rules governing hitting from behind and contact to the head
- Evaluation of track and field injuries leading to changes in pole vaulting, including an extended landing area and padding around standards
- Data on catastrophic eye injuries in women's lacrosse leading to the required use of appropriate protective eye wear in games and practices
- Data on head and neck injuries in football leading to new rules regarding spearing and head-down contact. In addition, the data was concerning enough for the NCAA to launch a major educational initiative for athletes, coaches, officials and administrators regarding the prevention of head and neck injuries. (Klossner 59)

Beyond its immediate impact for student athletes, the NCAA's data has also been used to influence the policy discussions of other organizations with an indirect impact, including youth sports associations, equipment manufacturers and facility architects. The NCAA data has been used, for example, to study the impact of the playing surface on athlete safety, and has been used as a basis for longitudinal studies tracking student athlete safety from high school through college (Klossner 60).

Another example of data-driven decision making being used to impact policy and strategy development comes from Eastern Connecticut State University (ECSU), which launched a data-driven approach to helping low income, minority, and first generation students succeed. ECSU was one of six New England undergraduate universities participating in Project Compass, a five-year, grant-funded initiative designed to use a data-intensive approach to improve retention and graduation rates among these special populations. More than 50% of ECSU's student cohort is part of this target research group, and so using data to understand student experience was a critical part of the project (Cid 4).

While many different departments at ECSU were collecting student data — Student Affairs and Housing, the library, Institutional Research — the school did not have a process for compiling and analyzing the data across different departments, nor were they shaping their data collection practices around the core questions that needed to be answered; for example, no first-generation or high school GPA data was being routinely collected (Kelly 4). Carmen Sid, Dean of the School of Arts and Sciences director of Project Compass at ECSU, said, “All this data are there, but we wanted to know: Who's at risk of leaving in the first year? Who's at risk for leaving after the second year? Can we use the information on the students coming in and develop a model that will predict who's more likely to leave?” (Kelly 4)

The university began to focus its data collection to support the development of a predictive model to identify at-risk groups of students. This included extending data collection beyond numbers and checkboxes on forms by conducting in-person student focus groups to collect more qualitative data about the student experience. In addition, the university's approach to data analysis was based on a “community of practice” model that involved faculty and staff from across the university working together to process data findings and develop works plans based on

them (Kelly 4) “While in some respects it was difficult,” said Margaret Martin, Title III director and sociology professor at ECSU, “one of the things that did happen was that lots of people from different parts of the university got to grapple with the raw data as it was coming forward and really participated in the analysis. Through that process, they got a chance to really identify what the group thought was important as well as what the researchers thought was important” (Kelly 4).

ECSU was able to construct a predictive model using this data and subsequent community analysis that uses high school GPA, engagement, and other factors to place each student into one of five levels of risk for withdrawal (Kelly 4). In identifying these five risk categories, the university has been able to develop highly targeted interventions that have resulted in greatly increased graduation rates across every student cohort at the university — and the highest graduation rates in the Connecticut State University system (Kelley, et al. 13). The data-driven approach to improving retention and graduation was so successful, in fact, that ECSU has incorporated it into its long-term strategic planning process, both to keep the issue at the forefront of strategic discussions and to ensure that funding, resources and accountability measures would be allocated to keep the momentum of the project moving forward (Kelly 4).

These are just two examples of how HEIs have benefitted from data-driven approaches to policy, strategy, and decision making. But despite the fact that HEIs have been collecting and reporting data for decades, studies show that a surprisingly small number of HEIs are actually using collected data to inform strategic planning and program improvement (Smith Mostest 12). Given the enormous potential for data to shape decisions that improve student experience, improve student retention, and improve use of campus resources, among other things, the question is: why aren’t more institutions implementing data-driven decision making as a foundational management practice?

### **Challenges in Data-Driven Decision Making: Too Much Information, Not Enough Resources**

In order for data-driven decision making to move from theoretical concept to practical implementation and outcome, there are two critical criteria in play: 1) data must be systematically collected, and 2) data must be accessible to decision-making stakeholders (Klossner 53).

One of the biggest challenges in data-driven decision making is excavating the insights that matter, especially in a world of Big Data and information overload — where every digital process, every social media exchange, every system, sensor, and mobile device is generating data of some type (“Big Data Analytics”).

The term Big Data is exactly what it sounds like: datasets that are big. And not just big in terms of volume; common factors that are used to identify Big Data include:

- Volume — the scale of the data
- Variety — different forms of data
- Velocity — the rate at which data is collected
- Veracity — the trustworthiness of the data
- Verification — the security of the data
- Value — the significance of the data (Daniel 908)

Data is considered “big” when these factors change or accumulate at a pace that exceeds the processing capacity of conventional database systems (Daniel 907). “With the growing diversity in the nature of data, managing and analyzing diverse data set is becoming a very complex process,” says Daniel. “Analysis needs to include linking, connecting correlating different data sets to be able to grasp the information that is supposed to be conveyed by these data. This situation is, therefore, termed as the ‘complexity’ of Big Data” (908).

HEI administrations are facing a heavy burden to collect and publish this data; so much so that IR is quickly becoming “a distinctive information technology-based industry providing performance-related data” (Browne 293). This growth carries with it implications for added staff workload, professional development, and financial and technological resources that are challenging in the face of an overall decline in the size of the HEI administrative workforce (Browne 293).

Most HEIs have neither the staff, financial, nor the technological resources to collect and analyze Big Data at this level, although there is a growing trend towards HEIs tapping into corporate-academic partnerships that allow them to take advantage of corporate advancements in Big Data mining (Daniel 906).

Martin noted that before the launch of ECSU’s Project Compass, the volume of data the university was collecting and the lack of resources for analysis were problematic. “I think it’s hard for the general faculty population or administrator population to really have a handle on the data that is really driving decisions,” she said. “They don’t get a chance to see it or they just get very infrequent information about it. So there may be too much data, but it’s often not communicated effectively to people in ways that are both understandable and useful to them” (Kelly 4).

While the professional field of data analytics is growing — according to a recent report from Linked In, data analysis is currently one of the most in-demand professional skills (Dykes) — there is still an overwhelming lack of expertise available for HEIs to draw from (Picciano 18). In addition to experienced database administrators who can collect and integrate big data, HEIs also need support from instructional designers who can understand and translate insights, as well as institutional researchers who are knowledgeable about the kinds of information the HEI needs to collect for its decision-making processes and strategies (Picciano 18). This may require hiring consultants or investing in time-consuming and extensive professional development for existing staff (Picciano 18).

## **Challenges in Data-Driven Decision Making: Institutional Silos**

In addition to the complexities of having adequate resources to mine Big Data for insights and translate them into actionable insights, HEI organizational structures can also create challenges in moving from data to action. Successful implementation of big data in HEIs is dependent upon collaboration, but a gap still exists between the data collectors/analysts and non-technical staff who know what data is needed and how it should be used to inform decision making (Daniel 916). In one recent study on decision making in higher education institutions, for example, only 38% of the institutions surveyed reported any collaboration among marketing, enrollment management, and academic advising teams (Fong 2).

In a recent report, the U.S. Department of Education indicated that collaboration among various institutional departments would be a critical factor in the success of any strategic data initiatives (Daniel 916). IR and Information Technology (IT) departments tend to make up the core of an institution's data-driven decision making processes, but even within that core, there is often a lack of clarity around the roles each department should be playing (Glover 3). In a report focused on strengthening IR and IT capacity in HEIs for the *Achieving the Dream* program, this lack of clarity is identified as a common culprit in reducing an institution's ability to make data-driven decisions:

Another challenge for IR offices is the confusion that exists at colleges about which department should provide what type of data. Some requests sent to IT might better be completed by IR — for example, the percentage of first-time-in-college students enrolled in the fall semester. The opposite occurs when IR receives requests for class rosters or grade input sheets, which may typically be produced by IT. On occasion, college employees forward the same request to both departments (without their knowledge) because it is unclear which should be the appropriate responder. Such confusion surrounding the respective reporting roles of IR and IT wastes time and reduces productivity. (Glover 3)

In addition to lack of clarity around roles, IR departments are also frequently working to overcome the perception that their contribution is limited to compliance reporting. Other institutional departments are often unaware of (or discount) IR's potential to contribute valuable data that could have a positive impact on their decision-making processes (Glover 3). In discussing the importance of data-driven decision making for the Student Affairs department of an institution, for example, Community Engagement Manager Kayley Robsham says that, "Interpretation and analysis of data on college campuses are the bread and butter of understanding student behavior. We can reveal untold stories of the college student experience by grounding student behaviors in data." (Robsham) But lack of open access to that data continues to be problematic for HEIs, says Robsham: "Sometimes intentionally or unintentionally we lock data behind departmental and technological silos," she says. "It's frustrating for student affairs professionals to see data or know that data is being collected without having access to it or being able to apply it."

## **Recommendations: Moving from Theory to Practice**

In light of the potential impact of data-driven decision making, and given the challenges presented by information overload, limited resources, and institutional silos, how might an institution begin to incorporate data and analytics into its strategic decision making processes?

- **Establish an organizational environment that encourages data literacy and fluency.**

Not every faculty or administrative staff member needs to be an expert in data management or know how to parse Big Data. But they should know how to access institutional data and conduct basic data analysis and research — especially if they are in a decision-making role. HEIs can invest in the best tools and technology for capturing and analyzing data, but without working to develop a “data fluent” culture across the entire organization, impact is still limited. As noted in *Decision-Making in the Era of Accountability*, “Until cultural norms that support the regular engagement with data and other forms of evidence to routinely monitor performance are in place, it is hard to imagine improvements in organizational decision-making” (Hora 420).

Establishing a culture of data fluency begins with administrative leadership setting expectations and priorities around the collection and assessment of data. As observed in ECSU’s Project Compass, one way to communicate a commitment to data-driven decision making is to incorporate data collection and analysis as a core priority in the institution’s overall strategic plan. This helps to ensure that resources and funds can be allocated to improving the school’s data literacy and fluency; for example, providing faculty and staff with opportunities to build data analysis and research skills through professional development.

According to one study, outside of accreditation policies that mandated data collection and reporting, most institutions offered staff no compelling reasons to invest in data analysis as means to improve programs, services, or learning (Hora 419). A culture that encourages data fluency may also incorporate incentives and opportunities for engaging in data-driven decision making processes, including offering online instruction, certificates for completing a certain number of hours of instruction, or committee participation credit for involvement in a data analysis project.

- **Create a culture of communication and collaboration.**

Departmental silos are common in higher education, but institutional data should belong to everyone — from IR to Student Affairs to faculty. It’s critical to ensure that roles and responsibilities between data collection units (i.e., IR and IT) are clearly defined, and that metrics and data are being shared regularly across departments. While many HEIs are short-staffed in institutional data collection and analysis, if your IR department has the capacity, consider departmental IR liaisons that can streamline communication and ensure that the data needs of the department are being met.

Another means of fostering communication and collaboration is to invite faculty members and departments that wouldn’t normally be a part of a data analysis project to participate — sometimes outside input can bring a fresh perspective. Faculty are often an

overlooked source of expertise in data analysis; engaging faculty as problem solvers in institutional data projects has the double benefit of creating a sense of ownership in the project while also potentially helping faculty members produce publishable research.

Collaboration does not have to be limited to institutional departments — in fact, collaborative efforts may benefit from partners completely outside the institution itself. As evidenced in the NCAA case study, an institution's collected data can be shared with organizations outside the institution that may lead to indirect impacts and outcomes. In addition, HEIs with limited capacity for building in-house data collection and management capabilities may find that outsourcing some of this work to a corporate entity is not only a cost-effective solution, but also allows them continual access to the most state-of-the-art resources with minimal investment required on their part.

- **Break down Big Data by visualizing the patterns and connections.**

Encouraging data fluency institution-wide also includes empowering individuals to communicate their data findings:

With the shift towards more self-service capabilities in analytics and business intelligence, the pool of people generating insights will expand beyond just analysts and data scientists. This new breed of data tools will make it easier for people across business functions to access and explore the data on their own. As a result, we're going to see an unprecedented number of insights being generated... However, unless we can improve the communication of these insights, we will also see a poorer insight-to-value conversion rate. If an insight isn't understood and isn't compelling, no one will act on it and no change will occur. (Dykes)

Basic data visualization skills should be encouraged as part of data literacy. There are many options for visualizing data and helping people — especially decision-makers — see the patterns and connections in the numbers. From simple Excel pie charts to more complicated infographics, bar charts, animations, and videos, you don't have to be an expert in coding or have extensive statistical knowledge to create compelling visuals that communicate the impact of your data. There are numerous online tools that even a novice can use to create excellent data visuals, including:

- **[Gapminder](#)**  
Gapminder is a global organization that produces data visualization tools that help people explore global statistics. They also develop teaching materials and tutorials for data visualization practice.
- **[Plotly](#)**  
Plotly offers open-source data visualization tools for composing, editing, and sharing interactive data visualization via the Web. No coding knowledge is necessary.

- **Tableau**  
Tableau is a software platform designed to help people “see and understand data.” Products include self-service analysis tools, data visualization tools, and user-centered design solutions.
  
- **Highcharts**  
Highcharts allows users to create interactive (and embeddable) online charts from their data, optimized for mobile browsing.
  
- **Make your data go “the last mile.”**  
While Big data may hold a great deal of potential value, data alone can’t facilitate good decision making at the administrative level. In his examination into the craft of research and data, professor of management at Texas A&M Richard Daft posits that the *why* — not the data — is the real contribution to knowledge (Daft 541). “Data collection and analysis are integral parts of the research process, but they are intermediate points between an initial hunch and the final story about the organizational world... The data alone are not enough, no matter how sophisticated the techniques for collection and analysis” (Daft 541).

Getting to the *why* requires data, but it also requires *response*. Visualizations help with clarity and understanding, but your data shouldn’t stop there. “Last Mile” skills are a marketing phrase that describe the ability to translate insights into actions, and any presentation of data insights should include recommendations or requests for some type of action. These might include:

- Recommendations for further research indicated by the data
- Suggestions for policy changes the data points to
- Identifying opportunities for improving programs of services
- Next steps in the data analysis process

## **Conclusion**

While it is clear that most institutions see the value of data collection and analysis as part of the decision-making process — and in fact, are already collecting vast amounts of data about their institutions, programs, students, and services — it is also clear that implementing data-driven decision making as a core institutional practice is hampered by accessibility of resources, information overload, and a lack of institutional collaboration and cooperation. While addressing these challenges at an institutional level may take time, there is no time like the present to begin taking small steps towards large-scale application of data-driven decision making practice.

With access to more data and more technology than any other time in history, academic institutions have an unprecedented opportunity to tap into the power of data-driven decision making as a platform to inform strategic decision making and policy development, and create a lasting impact on institutional culture.

## Works Cited

- “Big Data Analytics.” *IBM.com*, <https://www.ibm.com/analytics/us/en/big-data/>. Accessed 4 Nov. 2017.
- Browne, Liz, and Steve Rayner. “Managing Leadership in University Reform: Data-Led Decision-Making, the Cost of Learning and Déjà Vu?” *Educational Management Administration & Leadership*, vol. 43, no. 2, 2015, pp. 290–307, [10.1177/1741143213494890](https://doi.org/10.1177/1741143213494890). Accessed 4 Nov. 2017.
- Calderon, Angel, and Charles Mathies. “Institutional Research in the Future: Challenges within Higher Education and the Need for Excellence in Professional Practice.” *New Directions for Institutional Research*, vol. 157, 2013, pp. 77–91, [10.1002/ir.20040](https://doi.org/10.1002/ir.20040). Accessed 6 Oct. 2017.
- Cid, Carmen R. “Best Practices in Diversifying the Student Body and Faculty: The Dean’s Role.” *Adobe Acrobat PDF file*, <http://www.ccas.net/files/public/HOME%20PAGE%20LEFT%20MARGIN/SessionFF-EastCTStateU.pdf>. Accessed 10 Dec. 2017.
- Daft, Richard L. “Learning the Craft of Organizational Research.” *The Academy of Management Review*, vol. 8, no. 4, 1983, pp. 539–546, [www.rc.usf.edu/~jdorio/ORM/Daft, R. L. \(1983\)..pdf](http://www.rc.usf.edu/~jdorio/ORM/Daft,%20R.%20L.%20(1983).pdf). Accessed 6 Oct. 2017.
- Daniel, Ben. “Big Data and Analytics in Higher Education: Opportunities and Challenges.” *British Journal of Educational Technology*, vol. 46, no. 5, 2015, pp. 904–920, [10.0.4.87/bjet.12230](https://doi.org/10.0.4.87/bjet.12230). Accessed 6 Oct. 2017.
- Dykes, Brent. “Data Storytelling: The Essential Data Science Skill Everyone Needs.” *Forbes.com*, 2016, <https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/#22c4e32952ad>. Accessed 19 Oct. 2017.
- Fong, Jim, and Danielle Caldwell. “Leading by Gut or by Data: The Data-Driven State of Higher Ed Decision Making.” *University Professional and Continuing Education Association*, Oct. 2016, pp. 1–12., <https://www.helixeducation.com/wp-content/uploads/2016/10/The-Data-Driven-State-of-Higher-Ed-Decision-Making.pdf>. Accessed 10 Nov. 2017.
- Glover, Rhonda. “Strengthening Institutional Research and Information Technology Capacity through Achieving the Dream: Principles and Practices of Student Success.” *Lumina Foundation for Education*, 2009.
- Hora, Matthew T., et al. “Data Driven Decision-Making in the Era of Accountability: Fostering Faculty Data Cultures for Learning.” *The Review of Higher Education*, vol. 40, no. 3, 2017, pp. 391–426, <https://doi.org/10.1353/rhe.2017.0013>. Accessed 4 Nov. 2017.

- Kelley, Craig, et al. "Predictive Analytics as a Tool to Improve Retention and Success." *Association of American Colleges and Universities Conference*, Jan 2017. Microsoft PowerPoint file.
- Kelly, Rob. "Data-Driven Student Success." *Academic Leader*, vol. 5, no. 2, 2011, pp. 4–5.
- Klossner, David, et al. "Data-Driven Decision Making in Practice: The NCAA Injury Surveillance System." *New Directions for Institutional Research*, no. 144, 2009, pp. 53–64, [10.1002/ir.313](https://doi.org/10.1002/ir.313). Accessed 10 Nov. 2017.
- Picciano, Anthony G. "The Evolution of Big Data and Learning Analytics in American Higher Education." *Journal of Asynchronous Learning Networks*, vol. 16, no. 3, 2012, pp. 9–20.
- Robsham, Kayley, "The Importance of Data Storytelling in Student Affairs." *Presence.io*, 11 Oct. 2016, <http://www.presence.io/blog/the-importance-of-data-storytelling-in-student-affairs/>. Accessed 19 Oct. 2017.
- Smith Morest, Jessica, and Davis Jenkins. "Institutional Research and the Culture of Evidence at Community Colleges," *Community College Research Center*, 2007, [ccrc.tc.columbia.edu/publications/institutional-research-culture-of-evidence.html](http://ccrc.tc.columbia.edu/publications/institutional-research-culture-of-evidence.html). Accessed 6 Oct. 2017.