



# FLORIDA STATE UNIVERSITY BOOSTS STUDENT SUCCESS THROUGH BIG DATA

40K

Students reported on

## Impact

- Utilizing predictive analytics to identify patterns in LMS data, to predict if students will pass a course or not
- Make sense of trends, create custom reports to help several departments make key decisions
- Single environment for data collection and storage, processing, analytics, and forecast. This simplifies the stack, eliminates silos, and lowers cost
- Now proper analysis and prediction of seasonal test demand leads to more efficiency and lower operational costs for the testing center

One of the nation's elite research universities, Florida State University (FSU) preserves, expands, and disseminates knowledge in the sciences, technology, arts, humanities, and professions, while embracing a philosophy of learning strongly rooted in the traditions of the liberal arts and critical thinking. Founded in 1851, it is located on the oldest continuous site of higher education in the state of Florida.

With the responsibility for Learning Management System (LMS) reporting on over 40,000 students and 3,000 instructors, the FSU Office of Distance Learning's (ODL) limited data technology staff was looking for a way to address challenges and automate processes that had traditionally been manual. For public universities, metrics such as retention rate and graduation rate are important indicators for standing out in the competitive landscape. These success metrics are paramount to bringing in more students, making them successful, and continuing to grow a strong alumni network.

In 2016, FSU's Office of Distance Learning began its digital transformation journey. ODL was experiencing challenges with data loss, an inability to do advanced analytics due to limited resources, and needed more information about how to process data from various sources. Data was siloed and the team didn't have a robust and security compliant environment to properly collect, process, and analyze the data.

There was also an issue of space constraints - the LMS application servers and database generated vast amounts of data files, causing capacity limits to be reached on a recurring basis, so decisions had to be made regularly on what could be kept and what had to be sacrificed. This led to loss of data insights and the inability to make preventive decisions. As a result of manual processes, data loss, inefficient systems, and fragmented insight into the individual student's needs and lifecycle, meant ODL knew it needed to find a better way to increase student success.

## Boosting success through big data and machine learning

FSU has an impressive track record of success for its students, including a 93% retention rate for freshmen and a 72% graduation rate within four years (which compares favorably to the industry averages of a 81% retention rate and 60% 6-year graduation rate via the National Center for Education Statistics). In order to be able to continually sharpen these successes, there was a need for deeper data insights. The university's goals are to attract students based on its strong track record, ensure their retention through success in their studies, and help them to graduate in an appropriate timeframe so the metrics as well as the alumni themselves can reinforce the university's accomplishments.



“Cloudera offers us the convenience of integrating all the tools and capabilities into one platform. This means I don't have to hire multiple providers for multiple functions. Other universities try to add more and more technologies, but still struggle. We needed a solution that was easy, inexpensive, and offered good value in the long-run - and Cloudera makes this possible.”

Mary Eichin, Assistant Director for Technology, FSU Office of Distance Learning

FSU has many services available to drive each student’s achievements: mental health, tutoring, and disability services, among others. As students don’t often self identify, the challenge was being able to proactively offer the right support to the right student to ensure their success.

The need to provide better data and help improve student graduation rates as well as maximize retention, pushed ODL to utilize predictive analytics with a more robust and complete big data stack. The university uses the Canvas Learning Management System (LMS) to track student activity in a particular course. With Cloudera’s platform, ODL can now identify patterns in the LMS data, helping them identify students that are at risk of falling behind or failing in a course. Algorithms such as Logistic Regression in Apache Spark ML are used to predict if students will pass a course or not. Similar to financial service fraud analysis, the university can now also monitor cheating - helping to pinpoint and investigate it as needed, thereby also ensuring the quality of its graduates.

“By analyzing the specific records that could potentially produce a negative outcome for students, we can intervene and work to improve the situation before it’s too late. This helps us identify which students need more care. You can have all the services in the world, but if they don’t come in the door, it doesn’t help them,” said Mary Eichin, Assistant Director for Technology at FSU’s Office of Distance Learning.

### Powering decision-making across the university

Several departments across the university rely on reporting from the ODL technology team to power decision making. Previously, ODL was having to look through spreadsheets, compressed server logs, database output, and manually piece data together to create these reports. With Cloudera’s tools, the data is prepared so it can be utilized with Apache Impala and Apache Hive to build insightful use cases and make sense of trends as well as create custom reports that help the departments make decisions. Each morning, 85 tables are collected from a single data source, while concurrently more than 20 historical tables are fully replicated in the cluster. All ingested tables are joined to create clean baselines ready for reports, a process difficult to emulate with other technologies. ODL is able to use this data to collaborate with other departments regarding how the students are performing. This helps them answer questions they couldn’t answer before, faster and more reliably.

ODL was looking for a platform that could combine various tools all into one place. It needed a single environment for data collection and storage, processing, analytics, and forecast. Not only did this help simplify the stack and eliminate silos, but also lower cost.



"Cloudera offers us the convenience of integrating all the tools and capabilities into one platform. This means I don't have to hire multiple providers for multiple functions. Other universities try to add more and more technologies, but still struggle. We needed a solution that was easy, inexpensive, and offered good value in the long-run - and Cloudera makes this possible," said Eichin.

## Creating a more efficient testing center

Another primary use case where there was a need to store and analyze massive amounts of data was for the testing center. Previously, the testing center hired and scheduled staff without knowing how many students were going to show up. This guesswork created issues and inefficiencies. At times, there were too many staff members, and other times there were long lines of students with no one to assist them.

It was clear the testing center needed a more efficient approach to the problem. By utilizing Cloudera's platform, proper analysis and prediction of seasonal test demand could be conducted. With these insights, the testing center could improve processes and efficiency with better scheduling. Now more students can be served more efficiently, with fewer, shorter lines at the testing center, and operational costs for staffing have decreased.

## Looking to the Future

Based on their experience and success thus far, the ODL Data Technology Team has drawn up a list of more than 30 use cases where they can leverage the power of data to drive efficiency and insight. In some cases, the planned projects are refinements of the existing implementations, driving even more value; in others, they are innovations that will drive students as well as organizational success. One area they will tackle in this respect is the alignment of scoring rubrics to student proficiency for departments like the Liberal Studies group.

Whichever the use case, ODL is confident that with Cloudera, they have the data platform to help drive both their students' as well as organizational success.