



Driving results in higher education with data

Maryville University uses data lake to improve student success and retention

Maryville University recently launched an ambitious, multi-phased project to collect and analyze data to improve student outcomes. To make its vision immediately actionable and impactful, the university decided to focus on improving student retention first. University leaders recognized that online learners who successfully complete their first semester are far more likely to stay enrolled and graduate than those who do not. The goal is to use data and predictive analytics to quickly identify at-risk students, proactively intervene to help them stay on track, and continuously learn from the data being collected.

At the center of the university's vision is a data lake, built on Amazon Web Services (AWS), that serves as a central repository for student-related data. Data silos are a common challenge in higher education, and Maryville was no different. In just six weeks, IT staff implemented the data lake and achieved the mission-driven goal of beginning to use data to improve student success.

Maryville University

Private university located in
St. Louis, Missouri

90+ degree programs
in Arts & Science, Business,
Education, and Health Professions

Nearly 11,000 students enrolled:
50 percent undergraduate;
50 percent graduate and doctoral

Approximately **7,000 online students**

“Working with AWS gives us a lot of flexibility and freedom. We trust that we’ll have on-demand space and performance to fit our needs as we grow. And we can fine-tune our usage so we’re making the right choices and putting the right spend behind those choices to support our needs.”

Doug Glaze, Chief Technology Officer, Maryville University

Leveraging data from the data lake, the university can run data analytics to identify at-risk students, and then automatically nudge those students via text messaging to do things that will get them back on track.

“We’re providing students with resources they need before they know they need them,” says Jonathan Schlereth, executive director of the Office of Strategic Information.

Using predictive analytics based on data lake data, university leaders expect to provide a better student experience and see higher student retention rates, which will ultimately lead to higher graduation rates. These outcomes will reinforce the university’s mission of turning students’ dreams into reality and help the institution maintain fiscal stability.

A centralized source for all real-time and historical data

The university’s data lake centralizes multiple sources of data (e.g., the Ellucian Colleague student information system (SIS), which tracks data on student grades, attendance, and financial aid, and the Canvas learning management system (LMS)) for analysis and enables the university to store different versions of the same data in a single location. These capabilities are critical to get full visibility into data as well as observe changes in data over time.

“One of the main goals for the data lake was centralizing all of our data sources from across the institution so that we would have a single source of truth for building different analytic packages and providing strategic insights. The data lake helps ensure we have an up-to-date, accurate understanding of as many areas of the institution as we can,” says Schlereth.

The data lake also solves important technology challenges. Data from the LMS and the on-premises SIS database were critical components of the envisioned solution that would identify ways to improve the student experience.

However, the university’s existing manual extract, transfer, and load process was labor intensive, making it slow and extremely costly to process and prepare data for analysis and reporting. In addition, the on-premises storage platform and other systems didn’t have the level of scalability and flexibility needed for the types of analysis and data volume the university anticipated.

The team chose the prebuilt Amazon Simple Storage Solution (Amazon S3) data lake solution because it addressed these challenges.

“With support from the AWS team, migrating data from our storage environment and into the AWS data lake was significantly easier and faster than it would have been with other vendors’ solutions,” says Josh Tepen, a cloud engineer at the university.

The university IT team didn’t have to install components on premises and could quickly move raw data into the environment and set up the data warehouse. Within six weeks, the project team had launched a rich data lake solution that pooled LMS and SIS data.

“Building our student success strategy on our on-premises solutions would be limiting. Working with AWS gives us a lot of flexibility and freedom. We trust that we’ll have on-demand space and performance to fit our needs as we grow. And we can fine-tune our usage so we’re making the right choices and putting the right spend behind those choices to support our needs,” says Doug Glaze, chief technology officer for the university. Another advantage is the technology’s easy integration with other tools.

Answers to complex questions

AI/ML services are a core component of the university’s high-level vision and multi-phased implementation plan for using data. Maryville leaders aim to analyze as many areas of the institution as possible and find every opportunity to improve



In fall 2021, automated text notifications will remind students their online course is about to start. If a student doesn't log into the course by Wednesday of week one, the university will send another reminder, and so on.

outcomes. Now that the IT team has seeded the data lake with LMS and SIS data, they can begin to focus more closely on the next phase, which includes implementing and using AI and ML services for predictive modeling, analysis, and fine-tuning of intervention strategies, and other projects aimed at supporting student success.

The university is also looking at demographics such as financial aid status and known risk indicators such as whether the student is the first in their family to go to college. Each investigation involves creating a series of scripts that ask the right questions in the right way and in the right order. One benefit of the AWS machine learning service, AWS SageMaker, is that it will be easier for less experienced users to learn how to run these steps automatically. Ultimately, this will enable an end-to-end automated pipeline so analysts no longer have to run scripts and perform other data analysis steps manually.

"We can't have our data scientist running scripts manually on a daily basis, especially as we expand. That end-to-end pipeline through AWS SageMaker will be crucial as we bring in more students, add new use cases, and study new models. Our data scientist is really looking forward to being able to run different algorithms and do things that were not easily accomplished with our on-premises system," says Tina McQuie, data engineer for the Office of Strategic Information.

Automated notification — when time is of the essence

The solution's automated notification system enables the university to intervene with students and help them as soon as

analysis detects counterproductive behavior or other risks, or as a proactive measure that applies to all online learners.

Currently, the university is focused on automatically identifying students who haven't activated their Maryville student account and then triggering a text reminder. Account activation is a vital step, enabling students to access important resources they need to successfully start classes. For example, resources like the online student orientation introduces students to the learning management system, financial aid information, tutoring services, and more. Preparing students before classes begin allows them to focus on their coursework from day one and potentially prevent them from falling behind.

"One of our strategies has been to ask, 'What can we provide to help students be successful right now?' and then look for easy opportunities to do so," says Schlereth. With that mindset, the university set up a workflow to identify students who are not taking steps to engage and send a text reminder to them. To handle the workflow, a third-party custom application identifies these students through data from the data lake, and then sends that data to a Salesforce application that handles notifications and other interactions with students.

The system has sent more than 5,800 text messages during this initial phase of automation. "It's early in our analysis, but we're also starting to dig into how soon they actually activate their account after they receive the text message," says Schlereth. In fall 2021, automated text notifications will remind students their online course is about to start. If a student doesn't log into the course by Wednesday of week

“It’s been a great accomplishment to look back and see that something that was a large — and pretty overwhelming — idea has come together and is helping us make progress toward our larger goals.”

Doug Glaze, Chief Technology Officer, Maryville University

one, the university will send another reminder, and so on. Other near-term plans include automating email messages that offer additional resources to help students get oriented to online learning.

In the future, the university will leverage Amazon AppFlow — a fully managed integration service — to securely transfer results of those messages from the Salesforce application back into the data lake so the university can assess the efficacy of various notification and intervention approaches. The university has already begun tracking open and click-through rates and is starting to test the content and mode of communications to optimize messaging and message delivery. An important advantage of Amazon AppFlow is that non-IT staff can easily configure data flows and enrich data with filters, validation, and formulas in just a few steps.

Path to self sufficiency

In addition to AWS’s full-featured cloud technologies and services, Amazon’s partnership and support were key factors in Maryville University’s procurement decision and have been critical for the project’s success. AWS’s Data Solutions for Education Guided Lab offered the project team step-by-step instructions for building the data lake. Weekly consultations with AWS Solution Architects and education consultants helped answer questions and made sure the solution aligned with the

university’s vision. In the end, Maryville’s team was able to build the data lake on their own and develop the core competencies they need to continue to add more data sources to their data lake.

Looking to the future

While the university is currently focused on rolling out the solution to identify critical success factors and proactively serve up student support services to improve student outcomes, the long-range vision is to also leverage the solution to improve the institution’s business and operations departments. Project leaders are working with data scientists to better understand the university’s current position and identify future opportunities.

“We’re really excited about the possibilities that the data lake solution lends us. Now, when we discuss any implementation, we ask how we’re going to get it to the data lake and how we’re going to use the new data to build the next level of student service and support. It’s been a great accomplishment to look back and see that something that was a large — and pretty overwhelming — idea has come together and is helping us make progress toward our larger goals,” says Glaze.

This piece was developed and written by the Center for Digital Education Content Studio, with information and input from AWS.

PRODUCED BY:  CENTER FOR
DIGITAL
EDUCATION

FOR:  aws

The Center for Digital Education is a national research and advisory institute specializing in K-12 and higher education technology trends, policy and funding. The Center provides education and industry leaders with decision support and actionable insight to help effectively incorporate new technologies in the 21st century. www.centerdigitaled.com

AWS works with institutions around the world including Arizona State University, the University of Oxford, and Ventura County Community College to modernize all aspects of their mission delivery. Our customers rely on AWS to support the continuity of teaching and learning, connect the campus community to systems and tools, and store institutional knowledge and data. They use AWS to make data-driven decisions to save money and resources, enable researchers to accelerate time to science, and power their educational technology systems. Contact us to learn how AWS can help you with your biggest IT challenges. <https://aws.amazon.com/government-education/contact/>