# **Canvas Data Mining Project**

Project Lead: Daniel R Getty - Flint Institutional AnalysisUMF Student GR Researcher: Mason Turnner

UMF Student UG Researcher: Matthew Nickols

Faculty Advisor: Halil Bisgin

**Project Overview**

The Canvas Data Mining Project aims to leverage data mining techniques to extract valuable insights from Canvas, an online learning management system widely used in educational institutions. By analyzing Canvas data, we aim to identify students’ who are at risk of dropping a course. The Canvas data is warehoused by a company called Unizin. The data will be extracted from Unizin to the Institutional Analysis Production Data Warehouse (IAPROD). Then, a flat de-identified file will be generated for research purposes. The Canvas Data will often be referred to as Unizin data interchangeably. This project will be submitted to the Frontiers In Education (FIE) for peer review and potentially present at the October Conference in DC.

**Security Concerns and Measures**

There are no data being pulled from the University of Michigan – Flint (UMF) Banner directly. All ID’s are generated by the Unizin Data Platform (UDP). UMF Student Researchers involved with the project will not have access to any personally identifiable data housed in Banner (ie names, UMID, PIDM). The Data set is limited to Canvas course data with exception to deriving term codes from course dates which will utilize a function in IAPROD that generates term codes using the stvterm\_ext table. Students will have access to a flat data file on a College of Innovation and Technology (CIT) Windows 11 virtual machine (VM). Extracts are not performed by any students, the Unizin extract is pulled by Dan Getty from the office of Institutional Analysis

**Project History**

The beginning of this project arose after a conversation with Dawn Markell the Director of the Student Success Center (SSC) and the UMF Institutional Analysis Department. Dawn was interested in a way to develop an early alert system that could be used by advisors for pro-active advising. We came up with a plan to develop a dashboard that would identify students at risk of dropping a course and have them grouped by major. The SSC business model is to have advisors target advising interventions to students in specific majors. By ranking students by risk and major this would enable the advisors to engage in pro-active advising.

**(UDP) Entity Relationship Diagram (ERD)**

Below you will find the UDP ERD. You will see that the tables contain ID’s that are generated by UDP and used to relate the data in the ERD. **These ID’s are all generated by UDP and do not contain banner id’s**. The key tables cross walking the Unizin keys to banner ID’s will be housed in IAPROD and will not be provided in the flat file. Note that the author name in the UDP\_annotation\_ext is struck out. This variable will be removed in the de-identification process.

