2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering
Education for 130 Years
Paper ID #39838

A tag-based framework for collecting, processing, and visualizing student learning outcomes

Tonghui Xu, University of Massachusetts Lowell

PhD student

Dr. Hsien-Yuan Hsu, University of Massachusetts Lowell

Dr. Hsien-Yuan Hsu is an Assistant Professor in Research and Evaluation in the College of Education at the University of Massachusetts Lowell. Dr. Hsu received his PhD in Educational Psychology from Texas A&M University and has a background of statistics

Dr. Melissa Nemon, University of Massachusetts Lowell
Dr. Christopher Hansen, University of Massachusetts Lowell
John Hunter Mack, University of Massachusetts Lowell
Dr. David J. Willis, University of Massachusetts Lowell

David Willis is an Associate Professor of Mechanical Engineering at UMass Lowell. His interests are in aerodynamics and engineering education.

A Tag-Based Framework for Collecting, Processing, and Visualizing Student Learning Outcomes

Abstract

The Mechanical Engineering faculty at a public four-year, comprehensive university in the Northeast region are developing and piloting a tag-based framework to systematically identify, collect, process, and visualize student learning data for course- and program-level outcomes assessments and improvements. Student learning outcome identifier tags are used to link the questions on assignments, quizzes, projects, and exams to course outcomes and overall program outcomes. The goal of this pilot effort is to inform improvements to instruction, course design, course objective alignment, and program delivery. The tag data collected from grading a given assessment is de-identified, cleaned, and entered in a SQL server database. This data is then processed in a Python-based visualization platform.

Background

Course- and program-level assessments help determine student achievement of learning outcomes and support data-driven decisions about instructional and learning improvements in courses and curricula. At the course level, direct assessment focused on student knowledge and skills provides insight into course efficacy, student learning, and opportunities for instructional intervention. At the program level, student formative and summative assessment can lead to a deeper understanding of the overall program and curriculum efficacy. Compilations of course assessments can provide insight into the efficacy of curriculum structure, pre-requisite knowledge, course design, and opportunities for program improvement.

Several prior efforts have examined larger volumes of data used in outcomes assessment. The Individual Development and Educational Assessment (IDEA) center has collaborated with Kansas State University to use the campus labs software for building a student ratings of instruction (SRI) system (Ideaedu.org). The SRI system enables various stakeholders (e.g., faculty, staff, students, and administrators) to use the system to monitor their goals. Amos et al. (2021, 2023) also use Gradescope tagging to identify student assessments associated with key course and program outcomes. These tags are used to provide instructors and students feedback into the individual student and overall class performance and provide targeted feedback for growth. Tello and Motiwalla (2010) used blackboard data, MySQL, and Gradebook software to develop a web-interface system "eOutcome" to conduct program and course-level assessments. The "eOutcome" system includes faculty course reporting, administration assessment reporting, and student assignment and course information outcomes. They concluded that the system could help to collect, organize, and report campus progress toward course & program learning outcomes in higher education.

This paper has several objectives. First, in the methods section, we present the data sources, tagging framework, and the approach for data processing and visualization. Second, in the results section, we present examples of student assessment data visualization and discuss how these visualizations can enhance class-level and program-level assessments. Third, in the discussion section, we briefly provide an overview of the participating faculty's perceptions of the tagging framework and their reactions and levels of adoption and discuss the strategies for obtaining their support.

Methods

An electronic data mining/learning analytics (EDM/LA) knowledge discovery cycle model (Romero & Ventura, 2020) is applied. EDM/LA is one type of knowledge discovery (KDD) in a databases process model. KDD is a process model that applies data mining methods to discover the knowledge pattern behind large databases (Fayyad et al., 1996). According to Romero and Ventura (2020), the EDM/LA process model typically includes the (a) education environment (e.g., face-to-face classroom education, and information system used (e.g., blackboard or gradescope), (b) education data (e.g., school, course, student),

(c) processing the data (e.g., remove irrelevant variables, imput missing values), (d) method and techniques (e.g., data mining or learning analysis), (e) interpretation and application of new knowledge. Figure 1 illustrates the EDM/LA in our study. Andreas Holzinger (2013) proposed an integration of Human–Computer Interaction (HCI) and KDD to enable end users (e.g., instructors) to discover previously unknown and potentially useful information from interactive visualization data.

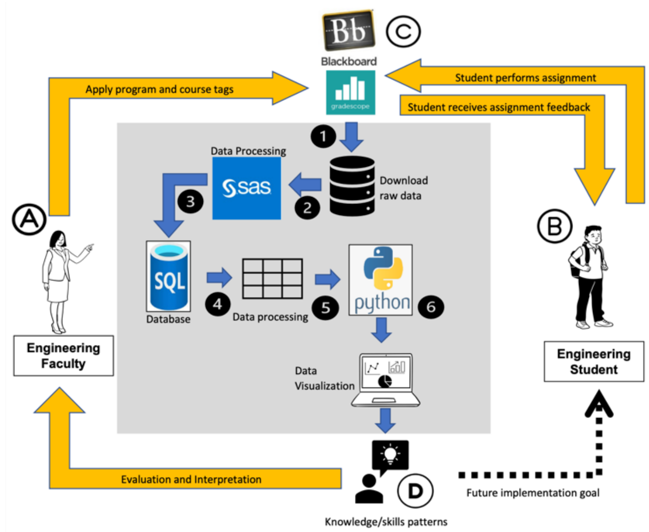


Figure 1. *EDM/LA Process of Our Study*

Data Sources: Here, we pilot this approach with undergraduate engineering students' assignments from a subset of courses in the Mechanical Engineering Department.

ABET-Tagging: Based on our process model, the initial pilot uses rubrics-based tagging (assessment criteria) and Gradescope (virtual environment) to apply the identifier tags. Singh et al. (2017) indicated that the primary benefits of Gradescope include (a) grading speed-up, (b) grading consistency, and (c) rubric modification flexibility.

To support program-level outcomes assessment, the department has developed pilot rubrics aligned with ABET's EAC (Accreditation Board for Engineering and Technology, Inc., ABET.org, Engineering Accreditation Commission) student outcomes. These student outcomes have been divided into sub-outcome performance indicators which are divided into four performance levels (see Table 1). The rubrics are used to apply program level tags to course performance. Our pilot framework also allows course specific content and skills tags to identify course outcomes that instructors may wish to track.

There are two approaches we are exploring for applying tags in Gradescope. The first uses the built-in rubrics functionality in Gradescope and the second uses the Gradescope's post-grading tags functionality.

1. **Grading Rubrics:** program and course tags are identified in the Gradescope grading rubric using square brackets, e.g., [ABET1c2]. Allocating assignment points to the rubric items is optional.
2. **Question Tagging:** Gradescope has a tagging function that allows the grader to create tag associated to a particular assessment item after grading is complete (Atwood & Singh, 2018). The student's score on the sub-assessment is linked to the assigned tag.

Once an assessment is tagged and graded, the assessment data can be downloaded from the LMS, cleaned, processed, visualized, and evaluated.

Table 1: example hierarchical tagging structure for Student Outcome # 2 (ABET EAC).

Student Outcome Tag	***Sub-Tag***	***Tag***
[ABET 2] ***Description:*** *an ability to apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare,*	**[ABET2c]** ***Description:*** *Present mandatory design considerations* *Performance indicator: Students have the ability to present (*consider) mandatory engineering*	**[ABET2c4]** ***Description:*** Presents if and how each mandatory consideration is relevant to the solution and includes a d*eep analysis for those which are relevant.*
		[ABET2c3] ***Description:*** Presents *if and how each mandatory consideration* is relevant to the solution.

as well as global, cultural, social, environmental, and economic factors. From ABET.org.	*design considerations including public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors in Engineering Design.*	**[ABET2c2]** ***Description:*** Neglects to or inappropriately/incorrectly determines if and how 1-3 considerations that legitimately appear in the solution.
		[ABET2c1] ***Description:*** Neglects to or inappropriately/ incorrectly determines if and how $\geq$4 appear in the solution

***Raw Data Pre-Processing, Gradescope*:** Because the raw data from Gradescope is not formatted for data management and processing, it is processed in SAS version 9.4 before being imported to the SQL server. The raw data pre-processing includes (a) encrypting student identity and deleting all personal identifying information, (b) reorganizing the data file structure, and (c) extracting and integrating similar data content.

SQL Database: After the raw data pre-processing step, Microsoft SQL Server is used to store the data files. This study followed the scheme from the Open University Learning Analytics Dataset (OULAD) (Kuzilek et al., 2017) to design an SQL database. Table 1 shows the data structure schema used. Using an SQL database allows integration with powerful data analytics software or languages (e.g., Python, SAS, R).

Table 1: *SQL Database Schema Description*

Table Name	Variable	File Descriptions
Course_Information	AID, Course_ID, Semester, Type, Grader, Total_Score, Description, Post_Time, Deadline_Time	Contains the general information of course
Assignment_Score	AID, SID, Score, Submission_Time, Adjustment, Comments	Contains the score of each student assignment
Assignment_Tag	AID, Tag_description, Tag_ID, Tag_Level, Value,	Contains the tag information of each assignment
Student_Assignment	AID, SID, Tag_Level	Contains each student assignment sub-tag levels.
Contents	AID, Tag_ID, Tag_contents	Contains the tag contents of each assignment.

Data Transformation: This step is used to prepare the data for data description, and interactive data visualizations. We used Python 3.8 version libraries (e.g., Numpy and Pandas) to process the data in visual studio code software. Table 2 lists the Python libraries used in this project.

The pre-processing step addressed the following aspects of the assessment data in our study:
- **Data frame combination and creation:** We read the original datasets from the SQL server and create new data frames to store the data.
- **Data transformation:** We transform the dataset for interactive data visualization.

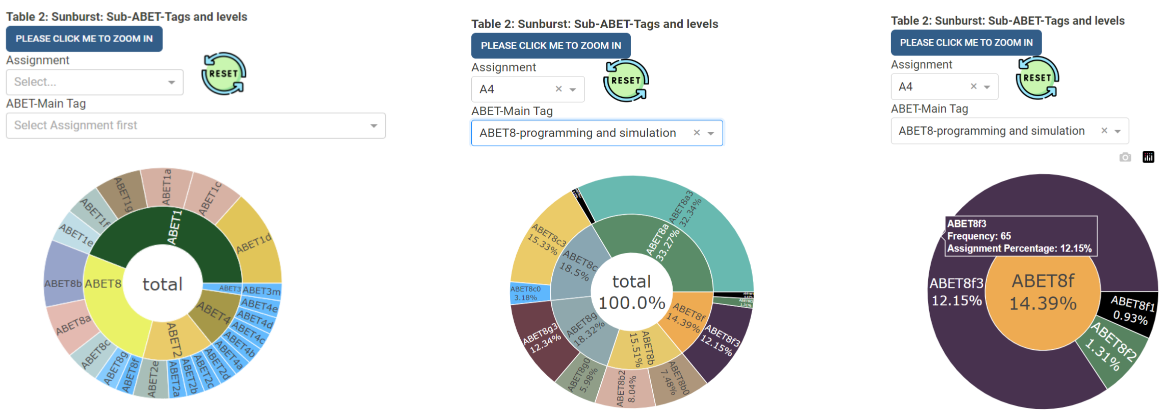
Table 2: *Python Library Description*

Library Name	Description
Numpy	Support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays
Pandas	Data structures and operations for manipulating numerical tables and time series
Plotly	Interactive, open-source plotting library
Dash	Build full stack web app with interactive data visualization application
Sklearn	Machine learning and statistical analytics

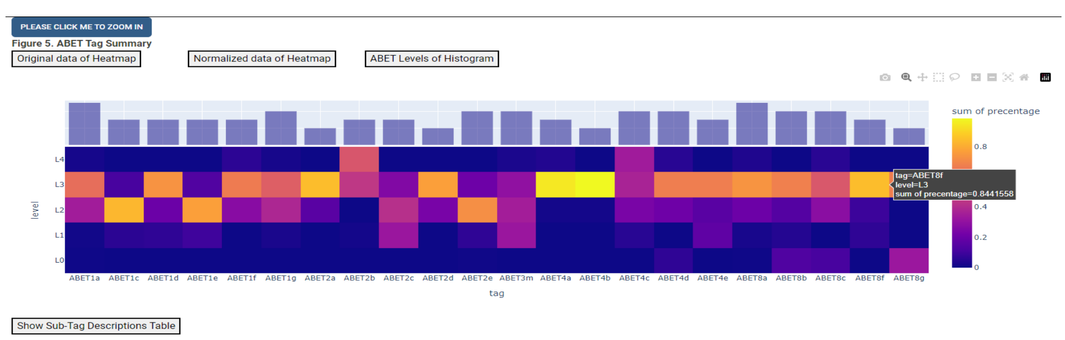
***Data Visualization Approach*:** Although many different visualization approaches were considered, interactive figures were deemed most useful by the project team including heatmaps, sunbursts, and pie charts. In the final process, the Dash library is used to develop an interactive dashboard application.

The dashboard implements interactive data charts, as shown in Figure 2. The sunburst chart presents hierarchical data from the center to the outer ring, with interactive functions to zoom and explore sectors. The visualization shows the rubric-level performance rather than assignment scores. An instructor can use this visualization explore data by selecting specific tags and assignments. In addition, a heatmap interactive chart shows course performance by tag. The top vertical bar chart indicates the tag appearance frequency in the course. Tag descriptions are also summarized in a table view within the dashboard.

Figure 2a-left shows how the interactive dashboard can be used to visualize overall tag data distribution for a class. In Figure 2a-middle, the same outcome is shown with sub-outcome and performance level outcomes. Finally, in Figure 2a-right, the ABET8f outcome tag is shown with the performance distribution on assignment components shown. This interactive visualization provides the instructor with a global view of assignment focus as well as a detailed view of each assignment outcome. Figure 2b shows the same data presented in a heatmap. This view provides a view of the overall course outcomes performance and can quickly indicate how students are performing on individual tags.



(a) Sunburst Chart



tag	Short Description	Performance Indicator Description -- student will demonstrate ability to...
ABET8f	Simulation grid and geometry (Could include CAD?)	Develop appropriate geometry/grid representation
ABET8g	Simulation boundary conditions	Identify and implement appropriate boundary conditions

(b) Heatmap Chart

Figure 2: *Selected Charts from the tag visualization dashboard: (a) Sunburst and (b) Heatmap.*

Discussion: Faculty Acceptance: A subset of faculty (n=5) provided feedback via an open-response survey that was deployed to understand adoption benefits and barriers. As faculty, many saw the advantage of being able to see program-level progress and perspectives beyond the individual class; and over time, it is clear to some faculty that this will be important to continued program and departmental growth.

The amount of time it takes to set up assessment components varied considerably across faculty, with some (n=2) noting that it added about 10 to 15% more time to their preparation process, while others (n=2) noted that it adds about 10 to 14 hours to their preparation per assignment. On average, it appears that most faculty currently take about 10 hours to set up per assignment, and an additional 10 to 15 hours of analysis. It is important to note that this is also the pilot version of this project, with no prior content to build off-of, and therefore some faculty noted that a lot of this process was a lot of trial and error. A few faculty (n=2) members noted specifically that it does increase the workload in the beginning and end of the course with creating the tags and managing the analysis, however this also helped with intentionality as it forced some faculty to rethink their assignments and ensure that they were addressing particular ABET outcomes. All faculty surveyed (n=5) agreed that the effort was worthwhile although immediate impacts were not always observed. The survey asked faculty to consider what they felt would encourage other faculty – whether in their program or not – to adopt such an initiative. Three respondents noted that one of the key elements lies in the messaging. Currently, the process looks and feels cumbersome and time-consuming, and the current messaging is too confusing for someone to readily embrace. Dedicated workshops, hands-on technical assistance, and knowledgeable mentors in the development and onset of the process would be helpful. Several faculty (n=3) mentioned time release or financial incentives for faculty, particularly at the early implementation stage and when creating assignments and tags for the first time.

The faculty were asked to give some final thoughts about the program within their own department as well as some words of consideration to other departments who may want to adopt this model. A few (n=2) noted that it is important to keep in mind that this process is in its earliest form and over time will likely become much more manageable, however the clear benefits are there and offer some revealing insights into student and program outcomes. Two other faculty noted that for this to really work, there needs to be a concerted, unified effort by all members of the faculty and leadership to dedicate both the time needed to establish this process and the engagement in the training necessary to develop the assessment skills (creating, deploying, and analyzing). One other faculty member added that it is important to recognize that this is not a radical idea, and perhaps emphasis on the fact that this is a tool to help faculty grow and become key opinion leaders in the field of learning outcome assessment.

Conclusions

A pilot effort to deploy a tag-based framework for course- and program-level outcomes assessments is described. Although still early in the development and deployment phase, faculty see the potential benefits of the approach despite the challenges associated with low process maturity and the time required to implement the tagging-based approach.

Bibliography

Amos, J. R., Angra, S., Castleberry, C., & Stadie, O. (2021, March). *Using Gradescope to Facilitate Tag-Enhanced Student Feedback* [Conference presentation abstract].52nd ACM Technical Symposium on Computer Science Education. https://doi.org/10.1145/3408877.3432498

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, *17*(3), 37-37. https://doi.org/10.1609/aimag.v17i3.1230

Ideaedu.org. (n.d.). *Idea's history*. IDEA. Retrieved April 12, 2023, from https://www.ideaedu.org/about-idea/ideas-history

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3), e1355. https://doi.org/10.1002/widm.1355

Tello, S. F., & Motiwalla, L. (2010). Using a learning management system to facilitate learning outcomes assessment. In *Learning management system technologies and software solutions for online teaching: Tools and applications* (pp. 138-156). IGI Global. https://doi.org/10.4018/978-1-61520-853-1.ch008