Towards Requirements for Supporting Course Redesign with Learning Analytics

Robert K. Morse
Indiana University, School of Informatics & Computing at IUPUI,
535 West Michigan Street, Indianapolis, Indiana 46202
(317) 916-7865
rkmorse@iupui.edu

ABSTRACT
With increased focus on quality, many Colleges and Universities run large enrollment online courses by maintaining course masters or frameworks. A framework enables curriculum managers to instantiate hundreds of course sections by keeping under control the quality of common elements of instruction. Current research in learning analytics tools suggests their applicability for the purposes of course design however, the current set of tools were built to support either course or student interventions. Little work has been done to examine the challenges of developing tools for the purpose of supporting the course redesign process. This research is the first phase of a larger project aimed at developing a “Best Practice Finder” which would aggregate section level data to help curriculum managers improve the design of their frameworks. This phase of the research seeks to investigate key requirements for the support of the use of learning analytic systems for the purpose of course redesign. Responses from 71 program chairs of Ivy Tech Community College of Indiana were collected to gauge the relative importance of curricular, institutional, and statistical knowledge they would need to make sense of learning analytics. Three main approaches to sense making emerged: a course centered, an institution centered, and an information centered approach. This work paves the way for an articulation of a full set of requirements for supporting course design focused learning analytics.

Categories and Subject Descriptors

General Terms

Keywords
Learning analytics, analytics, distance education, online course development, education administration.

1. INTRODUCTION

This research attempts to establish a solid foundation for applying learning analytics to the task of course redesign. Instead of focusing exclusively on the presentation of the analytic data, this research will first identify the knowledge demands for making design decisions based on learning analytic data.

Imagine for a moment a design scenario where a course is designed according to the Quality Matters (QM) design standards. A major concept presented by QM literature is the principle of alignment where by course materials, activities, and assessments align to session level objectives which in turn align to course level objectives. [6] For precision in referencing this mapping later, Ivy Tech has called it the Session Alignment Matrix (SAM). In this scenario alignment which is captured in the SAM is evaluated across two conditions. Horizontal alignment is achieved when the material, activities and assessment align within a session. Vertical alignment is achieved when there is progression in the elements of instruction from one session to the next session. For instance, if the student is asked to write a culminating assignment like a presentation of a final project to class mates have they learned all the skills needed to be successful? Do students have the resources they need or know how to access the resources (in this scenario they might need guidance on PowerPoint or where they can get help with academic presentations)?

This research is a case study on the design and support of a learning analytic system for Ivy Tech Community College of Indiana. Ivy Tech is the largest singly accredited Community College system in the United States. It covers the entire state of Indiana. It is the largest college in the state and has an annual enrollment of 125,000 students. One in four enrollments is for online courses. The college is creating a large data warehouse which pulls together institutional data from the student information and the learning management system. In 2006 the college adopted Blackboard as its learning management system. This research aims at designing and supporting a learning analytics front-end for the data warehouse which can be used by curriculum owners to make improvements to courses semester to semester.

Furthermore the decision making process of curriculum owners when reacting to learning analytic systems to improve course designs from semester to semester is not well understood.

2. BLACKBOARD DATA REPORTING OPTIONS

Reporting from the Blackboard System is currently handled through one of several mechanisms. Blackboard can be extended by developing java based packages known as building blocks. Kunnen & Nucifora [3] were awarded a grant through the Blackboard Greenhouse Initiative for their building block called ASTRO. Advanced System Tracking and Reporting. The tool
allows the system administrator to easily run reports on system usage across groups, schools, divisions, or other categories supported by course identifier naming conventions. From a technical perspective the tool was a massive step forward from the direct querying that was required of system administrators to get information out of the advanced system reporting (ASR) tool.

Although ASTRO makes the data of the ASR more accessible by summarizing key patterns of activity, the results are still only accessible to a system administrator role.

John Fritz [1] of University of Maryland-Baltimore County (UMBC) has developed a building block called Check My Activity. This allows students to compare their activity in the course against other students in the same class. It categorizes the grade book to allow students to check their activity against an anonymous set of their peers so students can self-regulate their learning behavior. There are a large number of tools that make the learning data more visible for the purposes of monitoring and changing behaviors.

Until recently only the ASR schema has been documented well for product developers. Starting with release 9.1 Blackboard announced an effort they are making to document the full learning system schema through a project known as Open Database. In 2011 Blackboard acquired an analytics company, iStrategy, and began developing its own data warehouse product. By partnering with schools already invested in learning analytics like UMBC, Blackboard continued to develop its reports in the Analytics for Learn product.[5] Analytics for Learn provides limited support for the contextual knowledge needed to make sense of learning analytic data which this research attempts to define more fully.

### 3. METHOD

Program chairs at Ivy Tech Community College of Indiana who have decision making responsibilities for courses or course content were surveyed to identify the sources of knowledge that support their use of eLearning analytics. All chairs of programs with statewide online course were surveyed. It was expected the categorization of the responses would match to proposed knowledge sources groupings, which includes knowledge of the statistics, knowledge of the domain, and knowledge of the institution.

A web based survey was created presenting real data but for a hypothetical course. The hypothetical course was IVYC 301 Designing an Online Course. The IVYC prefix is used for a certification track at Ivy Tech so, the scenario although made up for the sake of this study was still plausible. At the top of survey was a short video walkthrough of a new analytic tool which guides the program chair through available learning analytics to identify the “best practice” in implementing the course.

Therefore this draws on three available data sets that are regularly reported to program chairs. Each semester since Fall 2008 the end of term grade data from our student information system has been prepared in an Excel Pivot Table and distributed on CD college-wide. The second area of data presented in the survey was a bar graph of tool use collected from the ASTRO building block. The second source of data for the survey was a chart of summary statistics based on the Student Evaluation of Instruction. This presented the question number and the responses for one section of the contrived IVYC 301 course.

The survey consisted of two parts. First, three samples of data as described above from an actual course, but presented as the IVYC301 course, were presented along with the prompt to rank the importance of the five sources of knowledge. Refer to Table 1 for more information on the five sources of knowledge. For each of the sets of rankings participants were asked to provide a rationale for their ranking. The second part of the survey asked participants for additional demographic information including years teaching, years teaching online, years teaching at Ivy Tech, and the certifications they had completed.

A list of 350 program chairs from across Ivy Tech Community College who are responsible for the 150 online statewide course frameworks that are distributed each semester was generated from the Outlook email lists for each program. An initial survey invitation was sent to the combined list of 350 program chairs. Two weeks following the initial email one reminder email was sent to all participants who had not submitted a survey.

### 4. SURVEY RESULTS

71 responses or about a 20% response rate to the survey was collected. There was representation from across the college. In fact the representation of those responding follows the same pattern as the school representation in the statewide library of courses.

The main content of the survey asked respondents to rank order the importance of the five sources of knowledge already proposed to support their making sense of the data presented in the graphs. The mean of the ranking was calculated. For all cases, additional knowledge of the course design was the most important source of additional knowledge and statistical knowledge was the lowest.

<table>
<thead>
<tr>
<th>Domain Knowledge (Content)</th>
<th>Domain Knowledge (Design)</th>
<th>System Knowledge</th>
<th>Institutional Knowledge (Institution)</th>
<th>Statistical Knowledge (Stats)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curriculum Knowledge</td>
<td>Pedagogical Knowledge</td>
<td>Technical Knowledge</td>
<td>Institutional Knowledge</td>
<td>Statistical Knowledge</td>
</tr>
<tr>
<td>Course Objectives</td>
<td>Teaching Methods</td>
<td>System Definition of a Click</td>
<td>Data Refresh Rate</td>
<td>Meaning of Statistics</td>
</tr>
<tr>
<td>Course Content</td>
<td>Point Distribution</td>
<td>Categories of Tools, ie. Content Files, Communication Tools, Assessments</td>
<td>Default Rate</td>
<td>Vocabulary</td>
</tr>
<tr>
<td></td>
<td>Instructional Modality</td>
<td></td>
<td>Schools/Departments</td>
<td>Meaning of Graphs</td>
</tr>
</tbody>
</table>

Table 1: Categorization of the Contextual Knowledge Needs for Sensemaking
The emergent themes were coded, grouped, and recoded so that a full set of requirements could be connected back to the feedback collected from actual users.

For the grade data, the second most important knowledge source was knowledge of the course content and objectives. The Content category had a mean score of 3.71. For the learning management system activity data, the second most important knowledge source was the knowledge of the tools used in the course. The Tools category had a mean score of 3.52.

Beyond the simple calculation of means a correlation coefficient was calculated within the knowledge types showing the strength of correlation between an application of the knowledge type to another application.

There were only two knowledge types that showed significant correlations, r > .5. For the Content and Objectives knowledge type, there was a significant correlation between Grades and Activity data. (r = .72)

For the Institutional knowledge type, there was a significant correlation between the Grades data and the Activity data. This represents a moderate positive correlation showing that for those that ranked institution data high for making sense of grade data also rated it high for the making sense of the Activity data. (r = .53)

Each of the demographic factors was examined systematically for its effect on the rankings. There were only two conditions that seem affected at all by demographic factors. Years at Ivy Tech appeared to affect the need for additional content knowledge. This might be due to the role that program chairs play in developing curriculum at Ivy Tech. The program chairs are responsible for writing the course objectives in a document known as the course outline of record. The other demographic factor that affects the rankings of knowledge types was a significant difference between groups with Quality Matters (QM) Training and without QM Training. Quality Matters is an internationally recognized quality course improvement program whose main instrument is rubric of design standards for online courses. [6]

Given QM’s strong focus on course design, it is not surprising that those who participated in training expressed a much stronger need for knowledge about the course design than was expressed by those without training. What is surprising is that this effect seems to be concentrated around the course design knowledge for the Activity data only.

As far as the survey is concerned the remaining area of analysis was to explore the themes that emerged from the rationale of the rankings. The emergent themes were coded, grouped, and recoded so that a full set of requirements could be connected back to the feedback collected from actual users.

4.1 Emergent Requirements for Analytic Tools Supporting Course Redesign

1. Knowledge Support Requirements
   1.1 Suggest support with course design for those seeking support with content and objectives.
   1.2 Suggest support with content and objectives for those seeking support with course design
   1.3 Suggest support with institutional definitions for those seeking support with statistics.
   1.4 Suggest support with statistics for those seeking support with institutional definitions.

2. Data Requirements
   2.1 Broad access to data
      2.1.1 Visibility of aggregate information to a broad audience
      2.1.2 Visibility of course specific data to appropriate instructors or chairs only
   2.2 Integration with data from external tools
   2.3 Connection to research literature for external validation
   2.4 Availability to query historical data
   2.5 Availability to query multiple ratings by multiple instructors

3. Interface Requirements
   3.1 Provide clear graphs
   3.2 Provide clear and related descriptions
   3.3 Include chart legends
   3.4 Allow for a variety of support paths.
   3.5 Allow for a variety of data display options.

4. Functional Requirements
   4.1 Ability to Display Historical Trends
   4.2 Ability to Calculate Interater Reliability
   4.3 Ability to Conduct an Item Analysis of any Assessment
   4.4 Ability to Make Comparisons at Multiple Levels

5. DISCUSSION

Emerging out of the rankings and rationales were three groups of sensemakers. These groups were best defined by the area of knowledge that they felt was most important. The largest group was course centered. They indicated that the course design, course objectives and course content was the most important. A secondary group was a group that wanted to make sure that we did not lose sight of the general contextual knowledge of the Ivy Tech environment. Several folks pointed to the overall importance of a statistical foundation to understand what was being presented. These tended to be the veteran members of the program committee who were indicating what they felt their colleagues needed. The final group was much smaller but still worthy of mention. This group could not decide on the relative importance of all the knowledge sources and therefore indicated in their rationale that all the information they could get about a course was important.

In addition to the groupings of sensemaking, the largest contribution of this study is the validation of the proposed knowledge support typology. Emerging out of the data was one additional knowledge source. Program Chairs indicated that they want to know demographic information about the instructors and students from each section. How many of the sections are taught by adjuncts? How long has the instructor taught the course both at Ivy Tech and elsewhere? What is the male/female ratio in the section? What is the academic background for students taking the course?

In 2012, at the Educause Spring Focus session on Learning Analytics, John Fritz said that the main task of learning analytics right now is comparison. [1] Students want to compare their

![Figure 1: Means by Contextual Knowledge Type](image-url)
behavior to their classmates. Instructors want to compare their section’s performance to the performance of the statewide average. Several program chairs specifically mentioned the need for comparing their sections to the other sections in their region. So, it should remain as a key requirement for applying learning analytics to course redesign that comparisons between sections or groups of sections should be supported. The section to section, section to campus average, campus average to campus average, campus average to statewide average as well as other organic groupings like instructor average to campus average are all needed in a tool that aims at the improvement of the design of course frameworks.

In order to fully support the comparisons described above the presentation of the data in both tabular and visual representations must have clear contexts which are both evident and decipherable without a lot of cognitive effort. It must be clear if a graph represents data from a single section, or if it is an aggregation of data at the campus or statewide level. Although this has already been mentioned in the section of emergent requirements the importance of context was raised by several program chairs. As a rationale for the additional context they wanted for the Student Evaluation of Instruction one respondent said, “Absent this additional contextual information counting responses or assessing percentages is meaningless.” (Respondent 32)

Beyond a clear articulation of context it is further evident from the data that program chairs must be able to clearly see the difference between sections of the course. There are two primary areas of difference in course sections that will need to be highlighted by a tool built to improve course frameworks. First, the course settings such as extra credit or feedback options for assessments will need to be highlighted to help explain some of the variance between sections. Second, the instructor behaviors such as speed of returning assignments or the nature of the announcements that they post are indicators of differences in communication styles between instructors. For multiple section courses this will be a unique design challenge as it may not be possible to show the settings of hundreds of sections.

6. CONCLUSION
The results of this research are expected to contribute significantly both to the local decision making process of supporting course design frameworks at Ivy Tech Community College of Indiana as well as contributing to the theory of sensemaking and the design of visual analytic systems. The results of the survey produced the requirements for an analytic system from the perspective of user knowledge types. The identification of a user knowledge typology to categorize the knowledge sources will serve the college well in conceptualizing the types of users to support with this and other information technology system. Other colleges and universities with similar structure for academic leadership may find the typology useful for describing their users’ needs. It is an attempt to provide evidence for which sources of knowledge are the most important when exploring eLearning data for the purpose of course redesign.

7. ACKNOWLEDGEMENTS
The author would like to acknowledge Dr. Nathaniel Samba and Dr. Davide Bolchini for their careful review of this manuscript.

8. REFERENCES


DOI:10.1145/2090116.2090140 [link]
