Visualization Data For Learning Analytics

by

Halimat Ireti Alabi
Master of Arts, San Diego State University, 2007
Bachelor of Science, Purdue University, 2001

A Project Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF EDUCATION

in the Faculty of Education
Department of Curriculum and Instruction

Supervisory Committee

Visualization Data For Learning Analytics

by

Halimat Ireti Alabi
Master of Arts, San Diego State University, 2007
Bachelor of Science, Purdue University, 2001

Supervisory Committee

Dr. Jillianne Code, Department of Curriculum and Instruction
Supervisor

Dr. Valerie Irvine, Department of Curriculum and Instruction
Departmental Member

Abstract

Supervisory Committee

Dr. Jillianne Code, Department of Curriculum and Instruction
Supervisor

Dr. Valerie Irvine, Department of Curriculum and Instruction
Departmental Member

Learning analytics tools are used primarily by educational administrations to extract valuable information from learners' trace data to make institutional-level decisions. With the growing prominence of massive open online courses (MOOCs), the time has come for learning analytics tools designed for use by students to regulate their learning. To design effective learning analytics visualizations that provide formative feedback to learners, this project involves a pre-design study to explore the types of data collected by prominent learning management systems, how this data is visualized and the context of their use.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank Dr. Jillianne Code, Dr. Valerie Irvine and Dr. Yvonne Coady for their advice, inspiration and support. I am honoured to have worked with you.

Many thanks to Jillianne for reaching out to me when I needed it the most, for believing in me, trusting in my ability and trying to teach me patience. I'm still working on it…

My continued thanks goes to my mother, for teaching me early on the value of "applying the seat of my pants to the seat of my chair." Just like you promised, persistence has unlocked more doors for me than intelligence alone ever would have. Thank you for dealing with my distraction, and seeing my vision – even when it is just a picture painted with waving hands, rather than words.

Thank you to my golden child, who has been with me for every bit of this windy educational path. Please know that there is no one else whom I would choose to make this journey with. I cannot wait until it is your turn!

I cannot find the words to express my gratitude to HumB, so for now I will keep drawing my ideas, waving my hands, and planning until all of these dreams take flight.

I know that what I do is a privilege, not a right, and I am grateful for the ability. This work is dedicated to you, along with the many non-traditional students I have had over the years. Thank you for inspiring me every day.

Dedication

To my wonderful family,

thank you for the love, laughter,

long conversations and endless lentils.

Chapter 1 Introduction

Massively Open Online Courses (MOOCs) are the latest distance education trend, offering a level of convenience and zero dollar price point that has drawn an unprecedented number of learners from around the globe. Housed online in learning management systems (LMS), the value of MOOCs to educational research lies in the ability to collect fine-grained educational performance data through the learning management systems' infrastructure as every action in a learning management system may be recorded. The records of these logs, also known as trace data or log files, chart the path of learners through knowledge acquisition and the demonstration of their subject-matter proficiency. Though educational institutions often use this data for administrative purposes, it is seldom directly shared with learners to support their academic needs. To aid in the achievement of their educational goals, learners need access to their data, presented in ways to help them extract meaningful behavioural and performance patterns from it (Sadler, 1989).

Learning analytics is an interdisciplinary field that leverages academic performance and behavioural data by combining knowledge and techniques from educational data mining and psychometrics (Baker, Duval, Stamper, Wiley, & Buckingham Shum, 2012), This forms the foundation of learning analytics tools (LATs) that are seldom seen, informing the types of data collected and the algorithms used to manipulate this data. The organization and visual representation of the data by learning analytics tools is informed by human computer interaction techniques and best practices from the field of information visualization. LATs are impacted by the educational climate, their context of use and pedagogical assumptions about the nature of knowledge acquisition. Put into the hands of learners, LATs may empower learners "to bring the full power of data to bear on their learning related decisions" (Baker et al., 2012).

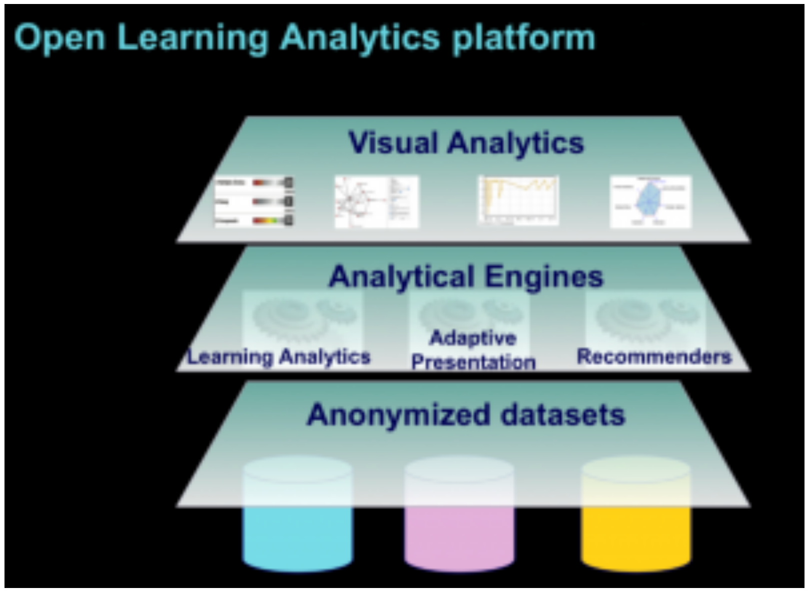

Figure 1. Open Learning Analytics structure (Greller, 2012)

Educational Climate

The “Iron Triangle” is what experts call the three interrelated problems of institutions of higher learning – limited access, rising costs and highly variant quality (Stengel, 2012). At over one trillion dollars, education is a large sector of the U.S. economy, one that has so far not been impacted much by information technology (Vardi, 2012). Motivated by economic, cultural and societal factors, this is beginning to change.

Knowledge workers currently comprise just over one-third of the American working population and their numbers are expected to increase (U.S. Bureau of Labor Statistics, 2012). Reflecting a cultural shift in the workplace, one of higher education’s biggest challenges is the need to efficiently prepare students for careers that require technical prowess and higher critical thinking skills, and then get them into the workplace quickly. There is no end of global competition for these top-performing individuals either – so much so that some sectors, particularly the fields of Science Technology Engineering and Mathematics (STEM), are experiencing shortages of trained workers. Meanwhile, more than a third of undergraduates must take remedial courses to begin collegiate study, indicating an enormous skill gap between what is required to graduate from high school and the skills necessary for collegiate success (Webley, 2012a). Without an adequately trained workforce the U.S. will have a difficult time staying globally competitive. Former IBM CEO Louis Gerstner Jr. summed the problem up well when he said

> "[Y]ou will never make it if the supply coming in is deficient… [t]hree million kids graduate high school every year. Half of them are unprepared for life. Does that create a sense of urgency for anyone?" (Webley, 2012a)

For some individuals the cost of higher education presents a barrier, especially during an economic recession. The American unemployment rate hovers at 8% (Webley, 2012a). When the American minimum wage is $15,080 but the average annual college tuition tops $27,000 (Department of Labor, 2013) making the cost of a college education a heavy financial burden. It is not surprising then that "[o]nly 3% of learners at the top 146 colleges are from families in the bottom fourth of household income" (Stengel, 2012). Author Richard Stengel describes higher education as an "engine of prosperity, innovation and social mobility" (Stengel, 2012), yet thus far it has proven to be a vehicle to success only for those who can afford the initial cost of entry. As a nation, there is a need to make higher education more accessible to the public, rather than only for those who can afford it.

Educational institutions are struggling to innovate and to provide better services to learners, even as financial contributions from government and private sources diminish. The support MOOCs are receiving from some of the world's most prominent universities such as MIT, Stanford, and Harvard reflect the pressure for increased accountability and performance. Universities and colleges are also competing with for-profit educational organizations that can make policy changes faster, rapidly integrating industry education and training needs into their curricula enabling for-profit organizations to offer a high degree of flexibility so students may obtain their education in a way that fits their personal needs, getting into the workforce faster than the traditional four-year degree. This degree of flexibility and personalization are important factors for today's “traditional” non-traditional learners.

Today’s “Traditional” Learners

Today’s incoming freshman are more likely to be holding full time jobs, balancing family, school and work obligations more than any other generation in the history of higher education . The traditional undergraduate, the model used for current course scheduling, is someone who enrolls full-time in college immediately after high school and does not work. The majority of today's undergraduate population – the current “traditional student” – would technically be defined as nontraditional by this standard. Part-time students, those working full-time, students older than 25 years of age, caretakers for children, parents or other family members, and disabled students who cannot attend brick and mortar schools or who choose not to compose the majority of those enrolling in undergraduate education programs today (National Center for Educational Statistics, 2002). These non-traditional learners are taking advantage of the convenience of online learning – as fully 62% of online learners are nontraditional students (Radford, 2011). A November 2011 report by the Babson Survey Research Group found that more than 6.1 million students took at least one online class during the fall of 2010, a 10% increase over the previous year and nearly four times the number of students taking online courses a decade ago (Webley, 2012b).

Massively Open Online Courses

Distance education programs have historically offered access to education for students requiring more flexibility than traditional brick and mortar classrooms could offer. The "massive" part is the most obvious reason MOOCs are so different from any other distance-learning course offered online; and since they are free, their price is another although most courses taken this way are not for college credit. Though time and cost savings are two major advantages to MOOCs, the calibre of the educators and their associated universities (i.e. Harvard, Stanford, University of California Berkeley, University of Toronto), their close ties to industry, and students' networking opportunities are also major benefits.

Cousera, Udacity, and EdX, the top 3 MOOC providers (Webley, 2012b), offer a wide range of collegiate level material to hundreds of thousands of learners from around the world. All three have ties to prominent universities. For example, Udacity was founded by three roboticists, including a former Stanford professor. Coursera has established partnerships with big names such as Princeton, Stanford, and recently, the University of Tokyo. EdX, a joint venture between MIT and Harvard, now also includes the University of Texas and the University of California, Berkeley.

EdX is the smallest of the big three providers with 350,000 enrolled learners (Seiffert, 2012), followed by Udacity with 400,000 enrollees since the fall of 2012 (Wallard, 2012). In the two years it has been in operation, Coursera has enrolled 2.4 million learners, offering 214 courses from 33 universities, including eight international institutions (Friedman, 2013). To put these numbers in perspective, the largest for-profit four-year university in the United States, the University of Phoenix, has been in operation since 1976 (Webley, 2012a) and has 328,000 students currently enrolled across all of their courses; the school has student body alumni of approximately 700,000 (Webley, 2012a).

The history of MOOCs, as with the history of distance learning as a whole, is checkered with both opportunities and challenges. Distance learning offered by radio, television and mail correspondence were all expected to revolutionize education, but none of them enjoyed the mass popularity of MOOCs. MIT's OpenCourseWare and Stanford's Engineering Everywhere project – predating MOOCs by a decade (Our History, 2013) – by offering free access to online course content (Cooper & Sahami, 2013). These projects have set the stage for today's MOOCs, along with technological advances such as cloud computing, increased Internet access, and the proliferation of mobile devices, and a lowering cost of personal computing.

Though MOOCs can vary in quality, they have also been a test-bed for innovative educational practices. For example, faculty members at participating institutions use MOOCs as a platform to develop and test unique course content online, without having to leave the tenured positions they already have. Further, MOOCs can be used in a 'flipped classroom' learning model, as is the case with San Jose State University's for-credit

MOOC pilot program. Learners enrolled in the pilot MOOC courses get transferable college credit in introductory courses such as Elementary Statistics, College Algebra, Intro to Psychology and Intro to Programming. This may open the door for later, direct MOOC accreditation – especially since the schools that prepared their learning content are already accredited.

MOOCs excel at meeting industry standards and employer competency requirements, in part due to their direct industry partnerships. For example, companies like Google and Microsoft are sponsoring Udacity classes in skills that are in short supply, from programming 3-D graphics to building apps for Android phones (Ripley, 2012). Though some of these partners are using MOOC platforms to educate their own employees, the public at-large benefits.

As MOOCs make course material more accessible, learner success in MOOCs is a complex, multidimensional, dynamic phenomenon (Subotzky & Prinsloo, 2011). A common criticism, as discussed in Subotzky and Prinsloo, is that learners get lost in massive learning environments; another is that learners suffer from reduced access to educators. A high degree of learner autonomy is necessary to meet the traditional model of success in MOOCs – leaners must self-motivate and self regulate (Anderson & Dron, 2011). Learners must hold themselves accountable for watching lectures, completing coursework, interacting with peers and fulfilling all the requirements necessary for the course. Though data collection in education is not new, the shear amount of data collected in MOOC environments may offer a way to solve these issues in the long term. Presently distance learners, particularly those in MOOCs, need tools to mitigate these challenges.

Focus of Inquiry

This research project is motivated by real world problems. Since nontraditional students are at risk to depart in their first year at over twice the rate of traditional students (Horn & Carroll, 1996), educational technologies that help learners transition into the expectations of collegiate communities are particularly important. Learners may benefit from educational technologies that give them the ability to understand and reflect on their evolution as life-long learners. Learning analytics tools (LATs) may be used to make students aware of themselves as evolving beings; LATs may also help them adapt to learning in online learning systems. Based on the hypotheses that learning analytics tools designed for learners' use will benefit their learning, this project seeks to explore the following questions:

- What data and data models have been used thus far in previous learning analytics tools, and do the resulting visualizations make sense for use by learners?
- What learning phenomena should analytics track to benefit learners in MOOCs, mitigating some of the issues experienced in online learning including isolation, disorientation and lack of motivation?

This pre-design study (Isenberg, Tang, & Carpendale, 2008) will help to ensure that the learners' context of use is understood before the LAT is fully developed. Pre-design studies in the field of information visualization are similar to exploratory data analyses or content analysis. LATs and the data used to compose them will be reviewed to better understand their users, goals, underlying data and context of use to be able to infer appropriate data to visualize in a LAT designed to aid learners in MOOC environments. The LAT is intended to be a form of formative feedback for learners to use to regulate their learning. This study is the first in a series of studies that will be used to inform the design of a LAT employing self-regulatory strategies. The goal of this study is the contextual understanding of LATs, their current and present usage to understand the learner, the learning process and the learning environment. Data gathered for this study is from qualitative and quantitative secondary sources. Data from instructor and learner experience surveys, LMS trace data, data models and the researcher's own experience as both a MOOC learner and an online instructor will be reviewed and analyzed. The results of this study will inform the design of a learning analytics tool (LAT) to assist learners by (1) detecting pedagogically important patterns in learners' assessments and classroom behaviours, (2) alerting individual learners to maladaptive behaviours, and (3) supporting their goal identification, monitoring and achievement through the provision of interactive visual feedback.

Chapter 2 Literature Review

The goal of this literature review is to present a historical review of analytics tools used in academic contexts and to identify the theories of learning (if any) that these tools support. The learning context, including the data models that conceptually framed the analytics tools, the data collected for and visualized in them will be explored in an effort to understand how these variables impact the many stakeholders invested in their use. Discourse will follow on the prevalence of LATs in current educational technologies, as well as the findings from the current study. The findings from this study will be used to establish a design framework for a learning analytics tool for use in online learning environments such as MOOCs.

Learning Analytics

Analytics may be defined by their area of impact, users, goals, specific topic or object of interest, for example, Facebook analytics (Barneveld, Arnold, & Campbell, 2012). The 2013 Horizon Report describes learning analytics as the "[F]ield associated with deciphering trends and patterns from educational big data, or huge sets of student-related data, to further the advancement of a personalized, supportive system of higher education" (*NMC Horizon Report, 2013 Higher Education Edition*, 2013). In online environments, learning analytics support the activity of learning through the collection, analysis, and interpretation of educational trace data (Ferguson, 2012). Siemens' definition of learning analytics as the "measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" (Siemens & Long, 2011) will be extended for this study, such that it is focused specifically on the process of learning, for use by learners.

The First International Conference on Learning Analytics and Knowledge took place in Banff in 2011. Learning analytics is a new field, formed after a split from academic analytics. In general, analytics tools are used to generate insight through obtaining and sustaining users' attention, showing user progress over time, highlighting areas that need attention, and possibly recommending future action. Thus, learning analytics tools lend students and instructors the ability to visually connect and analyze a wide variety of seemingly disparate learning data and enable the generation of new insights on the learning context in which to base future remedial action. The discipline draws on a broad array of academic disciplines that include academic analytics and educational data mining as well as concepts and techniques from information science and sociology, human computer interaction, social network analysis, latent semantic analysis, statistics, psychology, education and education technology. Epistemology, pedagogy, data sources, data models and assessment methods for LATs are all largely influenced by two of its predecessors in particular, academic analytics and educational data mining.

Academic & Action Analytics

Academic analytics, learning analytics and educational data mining are all data-centric, using large amounts of data to advance educational practice. Phil Long and George Siemens describe learning and academic analytics as “nascent fields that draw off of the success of business analytics’ influence on business intelligence” (Siemens & Long, 2011). Like business intelligence, the field of academic analytics uses massive data sets and predictive modeling to drive decisions (Siemens & Long, 2011). Academic, or action analytics can be thought of as a five-step process of 1) capture, 2) report, 3) predict, 4) act, and 5) refine, all conducted at the institutional level (Campbell, 2007). Unlike learning analytics used in classrooms, policymakers in educational institutions and governing agencies use action analytics. One of the most popular uses of analytics in an educational setting is for the tracking of student progress (*NMC Horizon Report, 2013 Higher Education Edition*, 2013). In the following case studies, academic analytics tools are used to respond to economic, political and social drivers.

Case studies in academic analytics

The most prevalent qualitative research methods in educational technology are ethnography, case study and design-based research (Luo, 2011); they also represent common educational technology analysis and reporting techniques. Case studies are appropriate research methodology when the focus of the study is a "how" or "why" question, when contextual conditions are highly relevant to the studied phenomenon, and when the boundaries between the phenomenon and context are not clear. The deeply contextualized focus and subjective reflection are the unique strengths of the following case studies (Luo, 2011). The following case studies from the University of Maryland Eastern Shore (UMES) and Bowie State University – both part of the University System of Maryland – shed light on the drivers that often lead universities to implement academic analytic tools used to support decision making in traditional brick-and-mortar universities.

The University of Maryland Eastern Shore (UMES) and Bowie State University have similar enrollments and implemented analytics systems for similar reasons – to aid in student recruitment, retention, and graduation rates. Retention rates are a significant issue in American universities because they are directly tied to federal funding, especially for state schools. The fallout from declining retention rates and the associated decrease of federal funds impacts resources available for current students. Though Forsythe et al. (Forsythe, Chacon, Spicer, & Valbuena, 2012) called the tools used in these studies learning analytics, they may arguably be classified as academic analytics tools given the institutional-level goals motivating their usage.

University of Maryland Eastern Shore (UMES)

The University of Maryland Eastern Shore implemented academic analytics after a three-year decline in the average SAT scores of applicants, retention and graduation rates. Due to the decline in enrollment, the school began registering more students for lower-level classes, resulting in an imbalance in their course offerings for upper-level courses. UMES responded to this imbalance by accepting more transfer students in the upper level courses. In addition financial, social, and academic challenges were impacting a large number of learners at the university (Forsythe et al., 2012). Fully 93% of the student body receive financial aid to attend the school with 53% of the student population being the first generation of their families to attend college (Forsythe et al., 2012). To attend the school, 70% of the enrolled students had to leave urban homes and to transition to the rural environment where the university is based (Forsythe et al., 2012).

The entire University of Maryland system of colleges and universities used annual dashboards to monitor graduation and retention rates at the time of this study. UMES implemented academic analytic dashboards using existing business intelligence tools, Microsoft PerformancePoint (Microsoft PerformancePoint, 2013), and Microsoft SQL Server (Microsoft SQL Server, 2013) to allow for daily student performance monitoring.

Many educational institutions have adopted the use of business-analytic like dashboards to increase measurement and evaluation transparency – even the United States Department of Education uses them to provide snapshots of public school information at the federal level (*NMC Horizon Report, 2013 Higher Education Edition*, 2013)..

Academic analytics, also known as action analytics, collect data at the institutional level and often use existing business analytics tools. Current action analytics tools and practices could be conceptualized as a niche of business analytics for use by organizations whose business happens to be education then repurposed for academic reporting needs. Academic analytic tools typically utilize information from internal and external institutional surveys, standardized exam scores, high school coursework and extracurricular activities, institutional enrollment, learner demographic and attendance data since this information is already collected.

The UMES action analytics tool was used to automatically alert academic support staff when a learner was not making adequate progress toward institutional goals. Faculty, staff, or students could also generate alerts, but it was not clear from the case study description how individuals would generate these alerts. Support staff is responsible for reviewing, prioritizing and acting on these alerts. When the number of flagged students overwhelmed the available staff, the alert system was revised to target specific issues, namely behaviors thought to lead to attrition. For example, data used by the UMES action analytics tool consisted of records from the student information system such as name and contact information, admissions data, student financial information including financial aid and account balances, course management, dining services, public safety, and attendance information. The data was visualized for users in a tabular format, with one column of color visualizations to augment the spreadsheet. The visualizations used three symbols based loosely on driving signage metaphors. Student data was compared to institutional performance targets for admission, retention and progress toward graduation, as defined by the university. The comparative algorithm for the data was not exposed to the academic support staff using the system.

UMES set the student progress goals and retention attributes: neither of these metrics were shared in the case study. Over the four years of the study, UMES used the analytic tool to increase average incoming SAT scores, decrease the freshman class size by 34%, increase transfer student enrollment by 150%, increase third and fourth year retention rates, maintain enrollment growth and “the school's commitment to first-generation college students” (Forsythe et al., 2012). Use of the analytics tool was reported as having been successful because the institutional goals were met, even in light of the university's strategic decision to limit freshman enrollment to reduce the number of remedial courses, adjunct faculty, tutorial and support staff required to support at-risk students. The limitation of their freshman enrollment resulted in a dramatic increase in SAT scores of accepted students, which was based on the use of the academic analytics. The increase in transfer student enrollments was used to help balance class loads. The success of the academic analytics tool ensured its continued use - but at what cost to learners, particularly at risk learners? Proponents of increasing educational opportunities for disadvantaged individuals may not likely see the same results in such a positive light; neither would the perspective students with lower SAT scores who were denied enrollment so the university could meet its fiscal goals.

It is mentioned in the results of the case study that UMES implemented additional dashboards to monitor the integrity of specific integral data records “to eliminate holes in its data and give end users confidence in the accuracy of institutional data” (Forsythe et al., 2012) . Matters of trust, particularly in response to a targeted question, do not lend themselves to quantitative review. A qualitative review might have revealed not only what caused these data holes, but also unveiled why users wouldn’t trust the data from the original sources. Though authors Forsythe et al. describe the dashboard as encouraging “campus-wide participation in the retention effort” (Forsythe et al., 2012), it is not clear how the dashboards impacted the motivation of the individuals using them. In fact, there is no evidence in either case study that the users of the tool or those impacted by it were ever consulted on its design, use or effectiveness.

Bowie State University

BSU Bowie State University’s (BSU) academic analytics tool was used to identify individuals needing support, after a slight decline in the school's second-year retention rate. This rate was identified as the single dependent variable for this case study. Like UMES BSU used “students’ significant variables” to track their progress toward graduation and to provide alerts to resource staff to prevent attrition, but this is where the similarities end.

The University’s Student Success Monitoring System (SSMS) aggregates data on: demographics, the learners' socioeconomic profile, academic program choices, attendance, course rosters, community group membership, formative and summative assessments. Alerts may be automatically generated by the system because of low grades or attendance problems, or generated manually. The SSMS conceptual model is based on

Swail’s eclectic model of retention (Swail, 2004). There are many data models for student retention; each of them identifies different key variables and models collected data in different ways. For example, Tinto’s Model (Tinto, 1975) utilizes two core variables, goal commitment and institutional commitment, that are conditioned by factors like academic and social integration, familial and institutional environments. This dominant sociological theory of retention has persisted over 35 years. The model put forth by Draper (Draper, 2008) focuses on a balance between intrinsic and extrinsic motivation. An eclectic model uses a set of variables involved in the learning process, but allows the researcher to weight the impact of the variables on the target population. The data collected is from student support, including the advising program, placement testing, counseling, athletics, band and other related activities; financial aid including socioeconomic data, need estimate's and the students’ ability to pay; recruitment and admission information including demographics, the students’ high school GPA and SAT scores; registration and advising information such as the students’ major and minor, general education requirements, academic calendar, and add/drop/withdraw information; academic services such as testing services, learning communities, and tutoring; and curriculum and instruction information such as the course and section rosters, the class schedule, class attendance and activity, formative evaluations, and summative assessments.

BSU’s academic analytics were designed for a wider set of users including students and faculty, in addition to academic support staff. Designed with an interface that is more accessible to non-expert users such as students and faculty, the software presents a different interface depending on the user's role. Tailored for each role’s specific use, the student interface reflects their personal profile, courses, academic services and programs available, email and calendar for scheduling appointments. In the tutorial for students, the system is introduced as the students' academic networking tool, part of their success network (*Student Starfish Tutorial*, 2011).

Unlike UMES, participation with the Student Success Monitoring System was voluntary. This presented a challenge that was met in part by recruitment efforts focused on 22 high enrollment courses in the diverse disciplines of Math, English and the Natural Sciences. When piloted in the spring of 2011, the program involved 40 faculty, 9 advisors, 19 support groups and approximately 1,500 students (Chacon, Spicer, & Valbuena, 2012). In the second deployment of the SSMS, user created profiles nearly doubled (Chacon et al., 2012). Though it was too soon to have empirical results on real usage, the increased participation in the second pilot deployment was reported as promising.

The analytics tools used by the University of Maryland and Bowie State made little use of visualizations to communicate data to students, which is common with academic analytics tools. The queries run by universities generally measure key performance indicators from the top down, supporting strategic planning, resource allocation, and administrative functions. Further, data experts, who have little need for elaborate visualizations to explain the data or to entice them to explore it, since they already know what they are looking for, normally run these queries. Though the institutional data collected for LATs visualization is similar, wide variations in adopted retention models mean that there are no "best practices" however this soon may change. Open data initiatives such as the one put forth by Predictive Analytics Reporting

Framework (WCET Cooperative for Educational Technologies, 2012 may allow data models that can be leveraged by multiple university systems. So far its member institutions have aggregated more than 640,000 student records that will be used to create data models and investigate questions surrounding student progress, completion, and retention (WCET Cooperative for Educational Technologies, 2012).

Focused less on individual learners, action analytics tools such as the ones used at UMES and BSU focus more on aggregations of learners, instructional design, curriculum development, and course content management at the institutional, program and department levels (Norris, Baer, Leonard, Pugliese, & Lefrere, 2008). In this way, action analytics provide institutions broad, longitudinal metrics on retention, performance and graduation rates.

Ross, Morrison and Lowther (Ross, Morrison, & Lowther, 2010) argue that research studies on "cutting-edge" technologies focus on proving effectiveness, failing to address more important, contextual issues. They argue that relevant, quality educational technology research must do more than simply present findings on how well a technology application worked, but should also be able to interpret *why* the technology worked for a particular user group within a particular context (Ross et al., 2010). The two preceding case studies represent what seems to be an underlying objectivist philosophy in academic analytics. An objectivist view of knowledge recognizes knowledge as having meaning independent of the individual, such that it can be transmitted between individuals (Reigeluth, 1983). This philosophy may not be representative of the teaching philosophies being promoted inside the classrooms these analytic tools are meant to support creating a disconnect between the learning environment being observed and the one being lived.

Theories of Learning with Learning Analytics Tools

While socially driven approaches began to emerge in academic analytics tools around 2003, specific underlying pedagogical theories are not often seen in the literature (Ferguson, 2012). The LATs reviewed for this study – CourseVis (Mazza & Dimitrova, 2007), GISMO (Mazza & Botturi, 2007), Uatu (McNely, Gestwicki, Hill, Parli-Horne, & Johnson, 2012) , LOCO:Analyst (Jovanović et al., 2007), Gephi (Gottardo & Vida Noronha, 2012), gStudy (Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007), Signals (Campbell, 2007), and Social Networks Adapting Pedagogical Practice (SNAPP) (Bakharia & Dawson, 2011) – reflect a range of learning theories and approaches, each will be reviewed in turn.

CourseVis

The CourseVis (Mazza & Dimitrova, 2007) learning analytics tool was developed in 2004 to help online educators visualize learner performance data gathered from the WebCT LMS (Mazza & Dimitrova, 2007) to be able to quickly assess the actions of learners and their own teaching effectiveness. Qualitative methods were prevalent in the design and implementation of this tool. First, qualitative data from 98 purposely sampled respondents (Patton, 2002) helped researchers determine the design requirements. Notably, the educators surveyed listed email as the main tool used to communicate with and engage students (85%), followed by discussion forums (80%) and chat (56%) (Patton, 2002). The surveys were also used to capture the educators' thoughts on the LMS, student behaviors that should be tracked, feedback and assessment. This information was used to construct the tool's model of how students are engaged socially, cognitively, and behaviorally in online learning (Mazza & Dimitrova, 2003). Follow up qualitative interviews (DiCicco Bloom & Crabtree, 2006) were conducted to "uncover aspects that had not been captured by the questionnaire" (Mazza & Dimitrova, 2004). When the tool was constructed, a focus group was used to evaluate the representations used within each of the tool's visualizations. Later a controlled experiment was conducted to collect quantitative data on the efficiency of the tool. This was immediately followed by a semi-structured interview with the same participants to gather data on the finished tool's effectiveness, efficiency and usefulness (Mazza, 2006).

The CourseVis tool used trace data to visualize learners' social, cognitive, and behavioral data, then presented the relationships between the three *en masse* to educators (Mazza & Dimitrova, 2004). CourseVis (Mazza & Dimitrova, 2007) utilized a number of comparatively complex visualizations to help instructors form mental models of what was happening in their courses. The ability of visualizations to help individuals form mental models and thus a better understanding of the data presented is an information visualization theory advanced by Spence (Patton, 2002). Visualizations are commonly used to present, confirm or explore data. CourseVis uses scatterplots and matrixes along with color, proximal placement, rotation and perspective projection (Tufte, 1990) to present multidimensional data for exploration (Mazza & Dimitrova, 2004).

Mazza and a new researcher partner, Luca Botturi, learned a valuable related lesson with their subsequent learning analytics tool, GISMO.

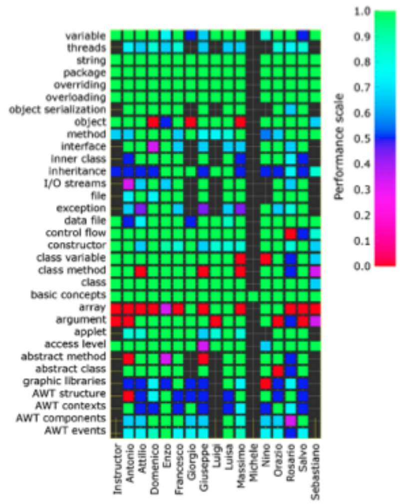

Figure 2. Cognitive matrix for visualizing student performance on quizzes related to domain concepts (Mazza & Dimitrova, 2007)

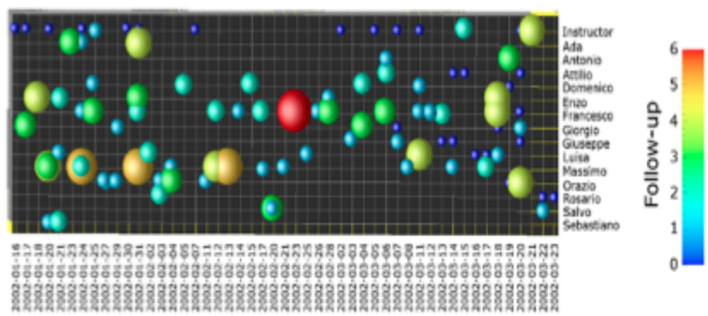

Figure 3. Discussion plot visualization of discussion threads focusing on the students who have initiated the threads (Mazza & Dimitrova, 2007)

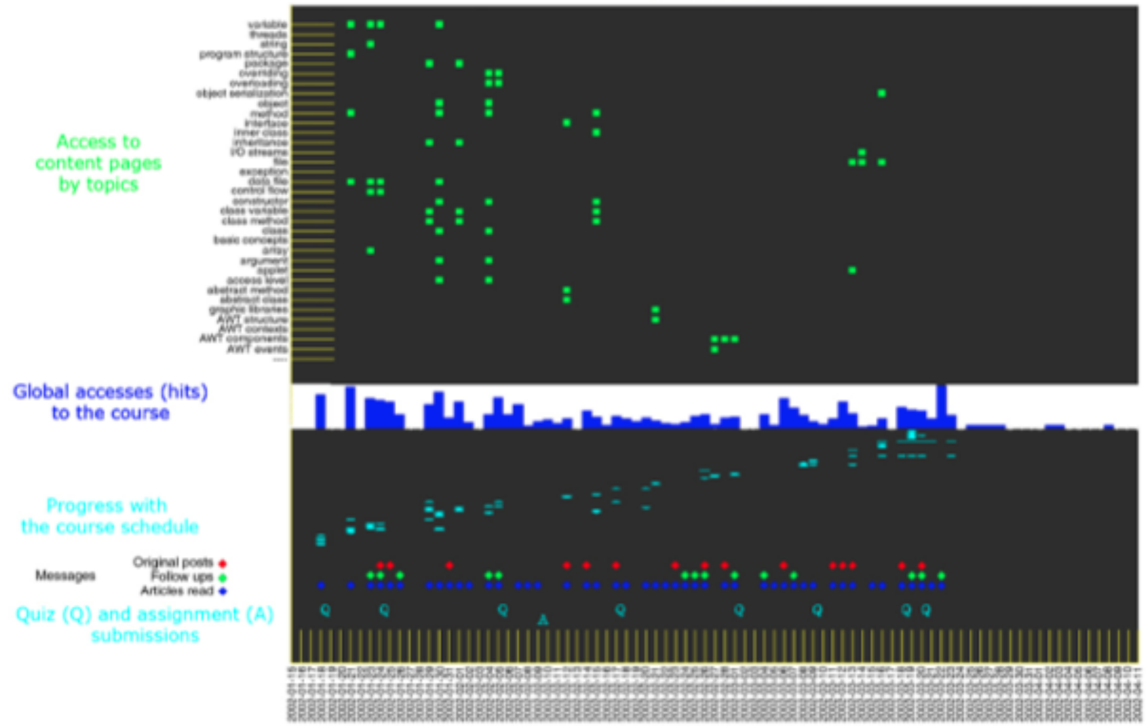

Figure 4. Comprehensive time series representation of the students' behaviour in the student behaviour graph including content accessed and learning progress (Mazza & Dimitrova, 2007)

GISMO

GISMO was built for the Moodle LMS to support a project-based Instructional Design course (Mazza & Botturi, 2007). A free, open source tool, GISMO utilized the data models and visualizations created for CourseVis, with surprising results. Researchers Mazza and Botturi applied a successful learning analytics tool on a similar LMS with the same users, to a new learning context. GISMO could not be utilized as intended, due to the misinterpretation of the data it visualized for learners. While GISMO efficiently monitored what was happening in the course, it could not be used for project-based evaluations due to the misinterpretation of the data visualizations.

In this case, a valuable lesson for all researchers was found in the researchers' reflections. Though Luo noted that this data could be a source of valuable information for educators (Luo, 2011), they are also important for the designers of educational technologies. LATs can be used to infer a good deal of information including learners' knowledge during the course of learning (Corbett & Anderson, 1994; Pavlik, Cen, & Koedinger, 2009), their metacognitive processes including self-efficacy (Litman & Forbes-Riley, 2006; McQuiggan, Mott, & Lester, 2007), their confusion (D'Mello, Craig, Witherspoon, Mcdaniel, & Graesser, 2008), help avoidance (Aleven, Mclaren, Roll, & Koedinger, 2006), un-scaffolded self-explanation (Ryan S Baker, Corbett, Roll, & Koedinger, 2008c; Shih, Koedinger, & Scheines, 2011)), and their level of engagement or undesirable learning behaviors (Baker, Corbett, & Koedinger, 2006; RSJ Baker & de Carvalho, 2008; Baker, Corbett, & Aleven, 2008a; Baker, 2007; Baker, Corbett, & Aleven, 2008b; Cetintas, Si, Xin, Hord, & Zhang, 2009; Walonoski & Heffernan, 2006).

Even so, the interpretation of the data collected by learning analytics tools is not straightforward; this in compounded by the different ways different LMS categorize and count the data they collect. For example, Moodle forum posts are embedded in nested threads. Though a learner may read several posts by different individuals to the same thread, their actions are counted as a single read post count by the LMS. Behaviors can also be interpreted many different ways. Different students have different learning methods. While one learner may be able to read something once and understand it, another may need to read it several times or consult various texts to understand the subject matter. It may be extended then that learning behaviors are difficult to quantify in LMS. Data directly collected from these systems must be uniform, but learning behaviors are not. Incremental assessments are often used in LMS to demonstrate acquired knowledge, but the use of automated or peer quizzes and surveys do not lend instructors the same sort of insight into project-based learning. Though technologically more advanced than CourseVis – GISMO's Application Programming Interface (API) was designed to be portable, allowing it to be used in LMS other than Moodle – GISMO failed to effectively visualize the group dynamics of the Instructional Design course projects. As exemplified by the GISMO case study, LATs are subject to misuse, misrepresentation or misunderstanding – particularly when more complex visualizations come into play. Ironically, transparency in learning analytics – used in many institutions to promote transparency – is an ongoing issue.

Uatu

The Uatu LAT (McNely, Gestwicki, Hill, Parli-Horne, & Johnson, 2012) was created to give university students formative feedback on their collaborative writing projects. Students' writing contributions were visualized in real time, along with the edit history of the collaboratively written documents they contributed to. The tool was model after the industrial practices common to knowledge work, including Agile development in project-based collaboration, regular peer-to-peer interaction and feedback to improve final deliverables. This study is unique in that it is one of the few learning analytics tool studies that used qualitative saturation to evaluate the tool. The Uatu tool (McNely, Gestwicki, Hill, Parli-Horne, & Johnson, 2012) visualizes the collaborative writing activities of university students, specifically novice computer programmers. Using Dourish's theories of embodied interaction (Dourish, 2001; Jovanović, Gašević, Brooks, Devedžić, & Hatala, 2007), the authors' approach to the interactive experience of computation is unique in that it recognizes how social phenomena unfold in real time, in realistic settings. Embodiment is at the centre of phenomenology (Creswell, 2012), an area of qualitative research that rejects the separation of knowledge and experience.

Six participants participated in the study. All were undergraduate students at a midsized public research university in the Midwestern United States. To determine participants in a realistic setting, the authors conducted a qualitative case study conducted with ethnographic methods that included: 20 different classroom observations, 14 observations of student writing in collaboration accompanied by talk aloud protocols, 70 photographs, 19 participant produced artifacts, usability observations of pair and group programming and presentations, followed by stimulated recall interviews, and 24 semi structured interviews with participants spread evenly over 15 weeks (McNely et al., 2012). The systematic nature of their study increased the reliability of their data and deepens their understanding of collaborative work. A clear application of Uatu is in online education, where visualizations of ongoing writing activity may help instructors provide more productive formative feedback and assessment, helping students learn as they work.

LOCO:Analyst

The LOCO:Analyst tool was built with the intention to raise educators' awareness in online learning environments (Jovanović et al., 2007), using an ontology previously developed by the researchers. The Learning Object Context Ontologies (LOCO) framework used semantic web technologies and was compliant with IMS Content Packing specifications. Described as a generic ontological framework, it integrates several learning-related ontologies including the Learning Context ontology, Domain Model Ontology and User Model Ontology (Jeremić et al., 2011).This enabled teachers to use the same tool to evaluate learners' activities that they did to set up their e-learning courses. Further, the researchers intended to formalize extracted learning patterns according to IMS Learning Design specifications, so that they would be reusable (Jovanović et al., 2007) embedded within iHelp (Brooks et al., 2006), a LMS that is open source standard complaint.

Two versions of the LOCO-Analyst tool were developed, the first in 2006 and the second in 2009 (Jovanović et al., 2007). Each time the qualitative studies first gathered initial feedback on the tools' proposed functionality and features, while follow-up studies allowed users to use the tool for some time before providing additional feedback. As with the GISMO tool, LOCO-Analyst was qualitatively evaluated using questionnaires and group interviews with education facilitators, namely instructors, teaching assistants and research students/practitioners (Jovanović et al., 2007). Jovanović et al. found that while all the user groups found the tool effective, the opinions of teaching assistants and instructors varied greatly on the perceived usefulness of feedback on individual students' interactions. The researchers attributed this difference to the instructors' "higher ability to identify useful patterns and draw relevant inferences from the presented feedback" (Jovanović et al., 2007). The researchers recognized the need to verify users' needs with actual users, rather than experts, in addition to the obvious differences between each of these user groups. Including representatives from each group within the case studies was an excellent example of establishing cultural validity.

Social Networks Adapting Pedagogical Practice (SNAPP) tool

In LMS interactions, dialog and collaboration are factors used to determine the nature and quality of learning. Social network analyses (SNA) illustrate the social patterns of learners' interactions, including their peer resource networks. Like game theory, these analyses are often used to model community dynamics (Chris Teplovs, Fujita, & Vatrapu, 2012). SNA is an effective way to model class interactions as a whole, helping educators to quickly identify outliers, learners whose participation patterns are different from those of the group, or highly active, well connected learners. Maladaptive behaviours that can be logging by LMS such as poor attendance, missed assignments and poor discussion participation are easily identified in visual representations of SNA. Reffay and Chanier's social network analysis model is based on cohesion, projecting the assumption that there is a dominant learning pattern that successful (Reffay & Chanier, 2003). Reffay and Chanier argue SNA tools may be more effective than content analyses.

Based on social constructivist theory, the Social Networks Adapting Pedagogical Practice (SNAPP) tool was developed to provide educators the ability to dynamically visualize the evolution of learner relationships within the LMS (Bakharia & Dawson, 2011). SNAPP's underlying foundation is the correlation between academic success and the learner's connections with their peers. It's creators, researchers Bakharia and Dawson, relate the learners' connectedness with their level of engagement. Its visualizations allow instructors to quickly analyze learner interactions visually, rather than interpreting these interactions through the review of discussion threads. Upon a glance dense, reciprocal or transitive relationships could be identified. Ego networks (see Figure 4 for an example) are composed of several learners with strong ties. Bakharia and Dawson note that individuals within ego networks tend to share common interests and attributes, but do not state the theory behind this assumption (2011).

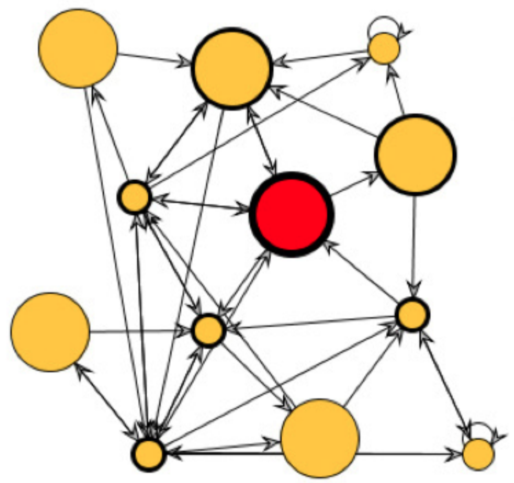

Figure 5. Ego network sociograms rendered with the Java Universal Network Framework (Jung) Library

In the second iteration of the SNAPP, an annotation feature was added so that educators could mark key events along that transpired in the course. The ability to add personalized notes let educators indicate the things important to them, such as instructional interventions, within a particular timeline.

In addition to its stated theoretical approach, SNAPP is unique in that it is LMS agnostic. Further, the tool is cross browser compatible and cross-platform compatible, including both open source and commercial LMS's, because the developers used client-side browser development techniques. Currently, there is no specific platform to target for development of learning analytics tools which would integrate with all LMS's simultaneously. As a result, many learning analytics tools were designed for a single platform. For example, as previously discussed CourseVis was designed for WebCT while GISMO was designed for Moodle. Each LMS has its own individual Application Programming Interface (API) and programming language. Although there are Learning

Tools Interoperability (LTI), Sharable Content Object Reference Model (SCORM), and Experience (Tin Can) API standards, LMS's do not collect uniform data or use a uniform API. SNAPP uses a bookmarklet. A bookmarklet is a small computer application, stored ing the URL of a bookmark in a web browser that extends the functionality of the browser. SNAPP's bookmarklet allows the tool to work in multiple browsers, extracting form data for multiple LMS and embedding the socio-grams visualizations within forums. To install the bookmarklet, a user just had to drag a link to their browser toolbar, or add it to a favorites list.

Gephi

Researchers Gottardo and Noronha wrote an SQL program to extract the data from a Moodle course in a format compatible with Gephi, an open-source interactive SNA visualization and exploration program. To produce visualizations of dynamic, complex systems and hierarchical graphs, Gephi needs data arranged into two GEXF file formats. One file must contain information about the learners in vertices or nodes. The other file contains information about the interactions between the participants, called the edges, arranged by the communication source, target and type. The provision of files organized in a different way from other LMS tools is a common data organization requirement for SNA tools.

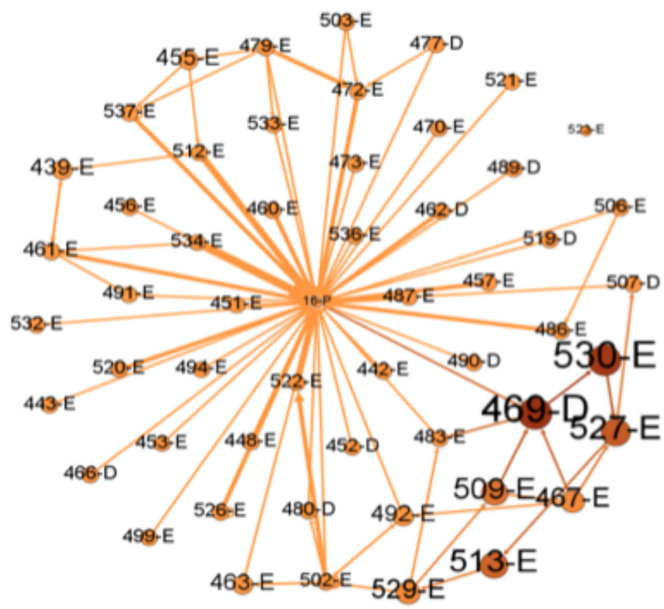

Figure 6. A Fruchterman-Reingold visualization illustrating discussion forum interactions demonstrating the centrality and overall connectedness of this learner group (Gottardo & Vida Noronha, 2012)

Oshima et al. argue that “existing social network models are unable to examine how community knowledge advances through learners' collaboration” (2012). Oshima et al.’s social network model is based on the words learners use in their discourse with their SNA tool, the Knowledge Building Discourse Explorer where the visualization uses words as nodes and the resultant node groupings are representative of the learning community’s knowledge (Oshima et al., 2012).

Personal Learning Environments Networks and Knowledge (PLENK)

With each fundamental shift in the needs of society, educational systems have changed to reflect the occupational needs of the economy. Following the Industrial Revolution, schools adopted a regimented model focused on rote memorization and routine to meet the labor needs of factories. In today's Information Age, information literacy – an intellectual framework for "understanding, finding, evaluating, and using information" – is gaining a foothold in classrooms (Libraries, 2013). In particular, information literate individuals – those who know how to access information effectively and efficiently – are needed to fill the dearth of employees knowledgeable in science, technology, education and mathematics. Learning analytics tools offer a way of presenting knowledge that builds learners' fluency with information technologies, an important competency in today's workplace.

MOOCs are social environments, where learners embark on a process of learning with massive numbers of their peers. The Personal Learning Environments Networks and Knowledge (PLENK) (Downes, Siemens, Cormier, & Kop, 2010) course offered in the fall of 2010 registered 1,641 learners. The course was a joint venture between the National Research Council of Canada's (NRC) Institute for Information Technology, Learning and collaborative Technologies Group's PLE Project, The Technology Enhanced Knowledge Research Institute (TEKRI) at Athabasca University, and the University of Prince Edward Island. Rather than prepared content, this connectivist course was distributed across the web. The ability to connect to each other outside of the LMS was quite popular, with learners quickly adopting communication on blogs and Twitter.

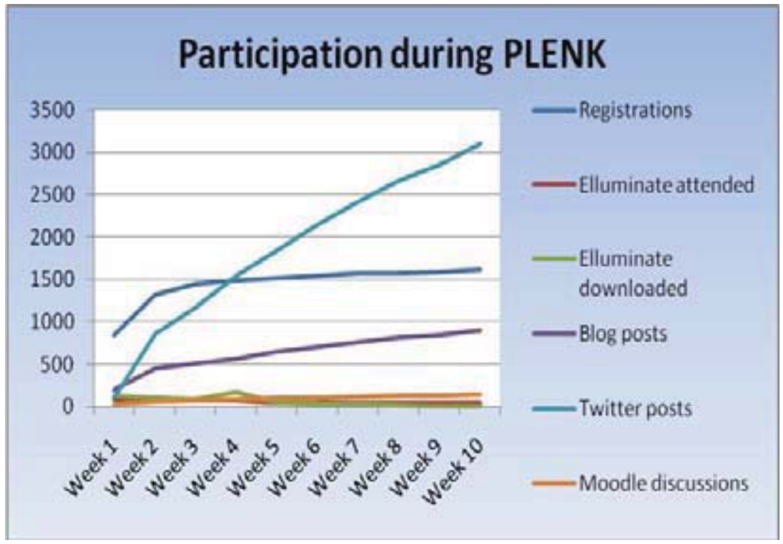

Figure 7. Participation during PLENK (Stephen Downes, et al., 2010)

Surprisingly, only 40-60 learners were active learning content producers; the remaining 1,580 individuals weren't visibly active. This is one example of the high attrition rates in MOOCs.

As noted by one of the learners in PLENK, "if nobody is an active producer, it limits the resources that all participants can use to develop their ideas, to discuss, think, and be inspired by in their learning" (Downes et al., 2010). Connectivist environments rely on participants for the collaborative construction of knowledge. In these learning environments it is normal for a learner to never interact with the educator. As of yet, the strategies used in MOOCs to mitigate this include peer grading and rating systems that pass on the expertise of assessment to people who are neither subject matter experts in instruction or the learning content. Learning analytics tools are a natural completion to the provision of feedback because they can provide a personalized experience without increasing the instructional load of the educator.

Signals

The fundamental difference between learning analytics and educational data mining is that unlike educational data mining, learning analytics tools are used to seek to understand relationships within the context of whole systems to support decision making by students and teachers. Educational data mining, on the other hand, seeks to develop methods to automate discovery that leverages human judgment at the organizational or institutional level (Siemens & Long, 2011). EDM reduces learning into components in order to analyze them individually and then discover relationships between them, focusing on the generalizability of the models created from them. In this regard, Learning analytics tools can benefit from the use of EDM tools and methods.

Signals (Campbell, 2007) is based on Purdue's premise that academic success is defined as a function of aptitude as measured by standardized test scores and similar information, and effort, as measured by participation within the LMS. Signals combines predictive modeling with data mining to communicate real-time, frequent interventional feedback directly to instructors and students (Arnold, 2010; Arnold & Pistilli, 2012; Barneveld et al., 2012). Signal's proprietary prediction algorithm consists of four components: performance, measured by percentage of points earned in course to date; effort, as defined by interaction with the Blackboard LMS as compared to students' peers; prior academic history, including academic preparation, high school GPA, and standardized test scores; and, student characteristics, such as residency, age, or credits attempted. Signals also collects data on student effort and engagement as exemplified by usage, assessment, assignment and calendar information from the LMS (University, Purdue, 2011a). The resulting feedback was visualized using a 3-color stop sign metaphor.

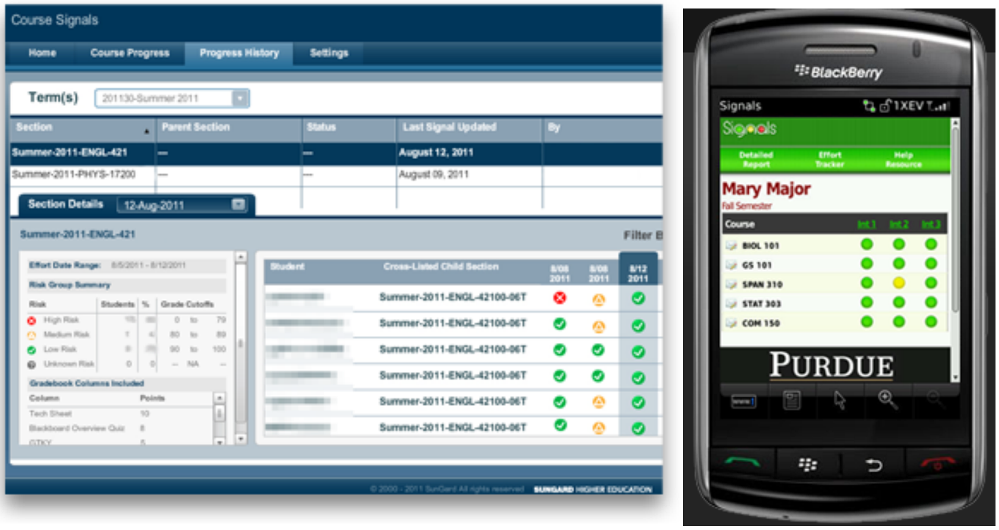

Figure 8. Purdue Signals tool desktop (Purdue University, 2001b) and mobile versions (Purdue University, 2011a)

Piloted in 2007, automated in 2009 and commercialized in 2010, Signals goes even further than the previous LATs reviewed because it directly engages learners in an effort to increase their graduation rates. Originally designed for educators to give relevant feedback to students (Arnold & Pistilli, 2012), the system is based on Tinto's Model (Tinto, 1975) of student retention to increase student success and to help students become academically integrated. The interaction model for this learning analytics tool is unique because it takes place between educators and students, rather than at the institutional level, or directly with academic support teams as seen with some of the earlier case studies and LAT examples.

Signals doesn't support a specific approach to learning but is an excellent example of 'nudge analytics' because of its use of predictive models to incite at-risk learners to action (Carmean & Mizzi, 2010). Nudges are subtle interactions that can influence people's actions without infringing on their freedom of choice (Carmean & Mizzi, 2010). These sorts of interventions persuade users to act, similar to persuasive technologies that prompt behavioural changes like healthy eating or smoking cessation (Consolvo, Markle, Patrick, & Chanasyk, 2009).

Three years of comparative usage data was used to indicate that use of the Signals tool resulted in higher grades (Arnold & Pistilli, 2012); earlier and more frequent help seeking (Pistilli, Arnold, & Bethune, 2012). User surveys gathered from more than 1,500 students across five semesters were used to confirm the tool’s successful adoption (Arnold & Pistilli, 2012). Arnold and Pistilli balanced both positive and negatives comments by instructors and students in their report (2012). Negative results can challenge researchers’ preconceived notions and help them to understand the complexity of the phenomena being studied. In the case of Signals, the negative feedback from some faculty members identified rich areas of inquiry. The concerns, including new students’ possible dependency on the tool and an apparent lack of best practices for using this class of tools (Arnold & Pistilli, 2012), echo ongoing debates in the learning analytics community about their usage and impact.

The success of Signals surely influenced the design of commercial analytics applications developed, specifically the Blackboard LMS in 2012 (Blackboard Inc., 2012). The analytics tool suite provided in the current Blackboard LMS offers both course-specific and institution-wide data and is currently being deployed in field trials with select four-year colleges and universities and two-year community colleges. Results of this pilot program are forthcoming.

gStudy

Bandura, Zimmerman and Sadler's theories of Self Regulated Learning contribute to the perspectives seen in contemporary research and practice in various instructional settings, particularly online learning environments. Bandura describes self-regulation as the way an individual may influence their external environment through self-observation, self-judgment and self-reaction (Bandura, 1991). It is from this definition that Zimmerman conceived the theory of self-regulated learning (Zimmerman, 1989) as a process that "occurs largely from the influence of students' self-generated thoughts, feelings, strategies, and behaviors, which are oriented toward the attainment of goals" (Zimmerman, 1989). Students who self-regulate their learning monitor, evaluate, and adjust their behavior, cognition, and motivation as necessary for effective learning and the successful completion of academic tasks (Zimmerman, 1989). gStudy is an example of an educational technology designed to support academic self-regulation. gStudy and programs similar to it, represents the future of learning analytics because they put the power of all that has been learned about analytics in academic settings completely in the hands of learners. Further, these tools support the motivational constructs that help learners help themselves.

In a recent gStudy study, Hadwin et. al. (Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007) examined the self-regulated study activities of leaners using trace data in an online environment, matching self-reports with behavioural data. Their exploratory case study of eight students was performed to create self-regulated learning profiles based on leaners' self-identified study types, juxtaposed to their trace data from using the gStudy instrument. This is key because in practice, learners' metacognitive knowledge may or may not be accurate, making the monitoring and evaluation of one's metacognition invaluable to the educational process. Metacognitive knowledge is difficult to measure due to the nature of self-reporting bias, however metacognitive skills are measurable in that goal setting and attainment are observable.

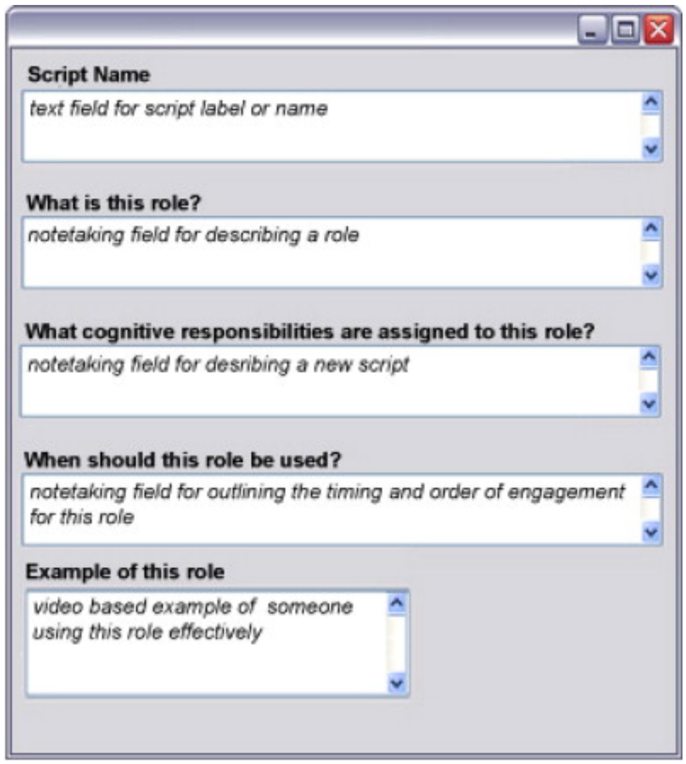

Figure 9. gStudy (Morris et al., 2010)

The gStudy visualizations consisted of prompts to guide their learning throughout a suite of tools. Data collected from log files provided information about the frequency, patterns, and duration of the students' actual study activities (Hadwin et al., 2007). The log files provided explanatory data for the qualitative information and vice versa. The resultant information on actual practice contrasted greatly with students' perceptions of their self-regulatory study habits, which was as expected (Hadwin et al., 2007).

This sort of study may help to explain log data collected on motivational constructs that impact learning, adding value to the trace data collected. This also may belie the future of learning analytics. Improved understanding of what behavioral/temporal progressions correspond to these constructs will help to ensure that they are visualized correctly with learning analytics tools.

Chapter 3: Content Analysis

As discovered during the research for this project, directly contacting the top 3 MOOC providers resulted in no information or data on the methods these providers use to collect data about their learners. Even edX, the only non-for profit, would not share the information even though the platform will be publicly offered in approximately 6 weeks. An email from an edX representative stated that "[A]t this time, in depth information on our LMS and analytics is not publicly available… keep up-to-date by "liking" edX online on Facebook, circling us on Google+ or following us on Twitter" (Miratrix, 2013). While following their social media channels helps their popularity, it does little for the research at hand. The only publicly available information on what data is collected about learners directly from edX, Coursera, and Udacity comes from reviewing their usage policies. The lack of publically shared information on the data collected by the top three MOOC providers completely changed the procedure for this project. Therefore, a qualitative document analysis of the available alternate sources of information were reviewed which included the 'terms of use and privacy statements' of the three MOOC providers and current LATs utilized in higher education (as reviewed in chapter 2).

Data Sources

EdX Terms of Use

EdX *may* collect information about individuals' performance and learning patterns by tracking information such as the users' IP address, operating system, browser software, and pages visited by the learner (edX Privacy Policy, 2013). For research in cognitive science and education, edX may use learners' trace data to recommend specific study partners, mentees or and mentors based on an individuals' interests or educational goals (edX Privacy Policy, 2013). Discussion posts may be used in subsequent offerings of a course the learner enrolled in "within the context of the forums, the courseware or otherwise" (edX Privacy Policy, 2013). Learners' trace data may be used to "present different users with different versions of course materials and software" to personalize the learning experience by adapting to the learner's ability and learning style (edX Privacy Policy, 2013).

Coursera Terms of Use

The Coursera terms of use say that data that *may* be collected from learners is: the percentage of assignments completed, forum points calculated by a summary of foreign participation and communicated with a numerical value, information provided by the learner in their profile and their career services settings ("Coursera Terms of Use," 2012). While they note learners' age, gender, and email address will not be shared, there are no definitive answers of what *will* be shared.

Udacity Terms of Use

Udacity's privacy policy states that they track, collect and aggregate learner information that is not personal, including LMS pages visited and when, hyperlinks and URLs accessed when linked to Udacity's LMS. According to the website, "[C]ollecting such information may involve logging the IP address, operating system and browser software used by each user of the Website" ("Udacity Privacy Policy ," 2013). The site also states that they may use cookies or web beacons to: "determine and identify repeat visitors, the type of content and sites to which a user of our Website links, the length of time each user spends at any particular area of our Website, and the specific functionalities that users choose to use" ("Udacity Privacy Policy ," 2013). It goes on to say that this non-personal information is collected to personalize the learner's experience and better serve them. Later in the policy it says that this information is used in aggregate to provide “higher quality, more useful services by performing statistical analyses of the collective characteristics and behavior of our users, and by measuring demographics and interests regarding specific areas of our Website” (“Udacity Privacy Policy ,” 2013). Udacity goes a step further in collecting leaner data to improve their experience in the LMS with a Feedback Program. The program is managed by a user research team that conducts user research studies including interviews, site visits, and online surveys.

To design learning analytics visualizations for use by learners, this project initially sought to explore the data collected by the three prominent MOOC providers – Cousera, Udacity and edX. Once it was discovered that this data is not publicly exposed or accessible by any public means, an alternative route was taken to obtain information about the types of data collected by these providers, why and when it is collected, who uses it and how. To determine what data to collect, one must first understand the learning phenomena that should be tracked to benefit learners in online learning environments such as MOOCs. It is important to understand the data and data models that have been used thus far in body of learning analytics tools, and to see if visualizations of that data make sense for use by learners to guide their learning.

Current Learning Analytics Tools

A representative sample of LATs were analyzed based on the following criteria: motivating factors for their adoption, the type of analytics employed, users, their goals, the topic or area of interest the analytics tools was used to visualize, the data collected, the interpretation of that data, the visualization of the data, the tools' stated or inferred theoretical approach, and the limitations of the tool. The information in Tables 1 and 2 aggregates the evaluation performed on these LATs.

The case studies from the University of Maryland Eastern Shore and Bowie State University's Student Success Monitoring System (SSMS) represent the drivers most universities experience that lead them to implement action analytics. Though they share a similar learner base and are in the same university system, their users, tools, visualizations, and goals differ. The visualization used by SSMS is unique in that it recognizes the needs of information novices. The University of Maryland Eastern Shore's use of action analytics is prototypical of action analytics performed by educational entities, right down to the use of a business analytics tool to provide administrative information in an educational context.

All of the learning analytics tools reinforce one of the two prevalent metaphors of learning – knowledge acquisition and knowledge through participation. CourseVis, Gismo, LOCO:Analyst, SNAPP and Signals were designed for use in the classroom by educators while Uatu was created for use by learners during the course of learning. Though the alerts from the Signals tool are sent directly to learners, they must be further explained by instructors for the learners to be able to act on the data. CourseVis and Gismo use visualization theories to help users form mental models of what transpires in an online classroom for educators. As online educators do not have the benefit of face to face communication, the visualizations are designed to help them better understand the data presented to them, based on Spence's (Patton, 2002) information visualization theory. CourseVis was successful but Gismo was not with similar users, data and visualizations because the learning contexts differed. While Gismo efficiently monitored what was happening in the course, it could not be used for project-based evaluations due to the misinterpretation of the data visualizations.

Tool	Motivation	Type of analytics	User	Goal	Topic or object of interest
Microsoft Performance Point and Microsoft SQL Server	3 year decline in applicant SAT scores, retention & graduation rates	Action	University of Maryland Eastern Shore (UMES) academic support staff	Administrative functions such as balancing enrollment, scheduling and hiring	Daily tracking of student progress as compared
Student Success Monitoring System (SSMS)	Decline in 2nd year retention rate	Action	Bowie State University Academic support staff, students and faculty	Prediction to identify individuals needing support	Tracking student progress, provide alerts
CourseVis	Dynamic-ally visualize learner performance	Learning	Online educators	Quickly assess the actions of learners, teaching effectiveness	See what is going on in distance classroom in real time
GISMO	Dynamically visualize learner performance	Learning	Online educators	Quickly assess the actions of learners, teaching effectiveness	Insight into group dynamics during project-based learning
Uatu	Dynamically visualize learner performance	Learning	Educators	Dynamic visualization of writing activity of novice computer programmers	Give formative feedback on their collaborative writing projects
LOCO: Analyst	Dynamically visualize learner performance	Learning	Educators	Quickly assess the actions of learners, teaching effectiveness	Increase speed and ease of monitoring performance
Social Networks Adapting Pedagogical Practice (SNAPP)	Dynamically visualize learner relationships	Learning	Educators	Model class interactions to quickly identify learners participation patterns that diverge from the group	Replace interaction discussion thread monitoring with visualization review
Knowledge Building Discourse Explorer (KBDeX)	Dynamically visualize learner performance	Learning	Educators	Reveal points pivotal for individual and groups in social knowledge advancement	Integrates social network analysis with discourse analysis to analyze learning
gStudy	Examine self-	Learning	Educators	Create self-regulated	Self-regulated

	regulated study activities		and learners	learning profiles, and explore potential of supporting self and co-regulation in collaborative engagement	learning progression and supportive computer-based learning environments
Signals	Help educators identify learners needing help	Learning and Nudge	Educators and learners	Quickly help learners in danger of failing	Prediction and remediation of low performance
Tool	Motivation	Type of analytics	User	Goal	Topic or object of interest

Table 1. Aggregation of LATs' motivations, type of analytics, users, goals and topics of interest

Though it would have benefited from further empirical study with online learners rather than co-located ones, Uatu is representative of learning tools that support collaborative work during the learning process, an interaction pattern common in MOOCs. Likewise, the tools employing social learning inform the design of visualization for use in social environments like MOOCs. The Knowledge Building Discourse Explorer (KBDeX) and SNAPP are exemplary tools that attempt to support different modes of learning within the same tool. LOCO:Analyst is the only tool that supports several learning ontologies – the researchers' Learning Object Context Ontologies (LOCO) framework, and several other learning-related ontologies including the Learning Context ontology, Domain Model Ontology and User Model Ontology.

Tool	Data	Interpretation of trace data	Visualization	Shared with learner	Pedago-gical Approach	Limits
Microsoft Perform-ance Point and Micro-soft SQL Server	Demographics, socioeconomic profile, admissions data, standardized scores, LMS course assessments, dining services, public safety, and attendance	Filtered with proprietary metrics targeting behaviors thought to lead to attrition	Tabular dashboard, with one column of color visualization s based on loosely on driving signage metaphors	No	Inferred objecti-vist philo-sophy	Not shared with learners, little use of visualiz-ations

Student Success Monitor-ing System (SSMS)	Demographics, high school GPA and SAT scores, student support and counseling, admissions data, socioeconomic profile, LMS course assessments, standardized scores, attendance, community and recreational group membership	Swail's eclectic model of retention (Swail, 2004) Proprietary metrics targeting behaviors thought to lead to attrition	LMS admini-strative interface for non-expert users, including dashboard	Yes	Inferred objecti-vist philo-sophy	Little use of visualiz-ations
CourseVis	Performance and discussions from LMS	Behavioral, cognitive and social engagement	Scatterplots, matrixes, along with color multidimensi onal data for exploration	No	Construct ivist philoso-phy inferred	Learners could not benefit from same informa-tion
GISMO	Performance and discussions from LMS	Behavioral, cognitive and social engagement	Scatterplots, matrixes, along with color multidimensi onal data for exploration	No	Construct ivist philoso-phy inferred	Visualiza-tions misinter-preted, so could not be used for project-based courses
Uatu	Individual and group writing contributions, edit history	Dourish's theories of embodied interaction (Dourish, 2001; Jovanović, Gašević, Brooks, Devedžić, & Hatala, 2007)	Individual and group writing contribu-tions, edit history	Yes	Construct ivist philoso-phy inferred	Tested with co-located learners, but tool better suited to online learning
LOCO: Analyst	Discussion threads, learner name and person communication is directed to	Social learning theory that dialog and collaboration determine nature and quality of learning	Scatterplots, word networks, bar charts	No	Integrates several learning-related ontolo-gies	Informa-tion not shared with learners
Social Networks Adapting Pedagogic al Practice (SNAPP)	Discussion threads	Correlates academic success with peers connections, and	Ego networks	No	Social learning perspecti ve that dialog and	Supports only one learning mode

		connectedne with engagement			collaborat ion determine nature and quality of learning.	
Knowledg e Building Discourse Explorer (KBDeX)	Conversation data filtered by: merging singular and plural, merging words with same meaning, merging conjugated verb forms, removing non-academic conversations	SNA visualization to examine how learners regulate learning, built into discourse analysis	Fruchterman -Reingold type visualization s, networks of words		Social learning	Visualiza-tions may be difficult for novices
gStudy	Frequency, patterns, and duration of the students' actual study activities	Academic success defined as function of aptitude measured by standardized scores and effort (participation)	Template guides of strategic learning and, highlighting and prompts embedded in range of tools	Yes	Self-regulated learning	Little use of summa-tive visualiza-tions
Signals	Performance measured by percentage of points earned in course to date; effort defined by interaction in LMS compared to peers; prior academic history, demographics, assessments	Predictive modeling of data leading to attrition using Tinto's Model (Tinto, 1975) of student retention	3-color stop sign type metaphor	Yes	Construc-tivist pedagogy inferred	Not enough informa-tion provided to students to direct next steps
Tool	Data	Interpretation of educational trace data	Visualization	Shared with learner	Stated or Inferred Pedago-gical Approach	Limita-tions

Table 2. Aggregation of LATs' data collected, interpretation of the data, visualizations, pedagogical approach, limitations and if the LAT was shared with the learner

Though data collection in education is not new, the shear amount of data potentially collected in MOOC environments may offer a way to solve many issues around education in the long term. Presently distance learners, particularly those in MOOCs, need tools to mitigate these challenges.

Chapter 4: Proposed LAT Design

Today today's nontraditional learner juggles many responsibilities and is often operating with an attentional deficit. There are often many things competing for their attention, even while they are studying. A visual representation of their learning process will help to give the learner an overview of their goals, tasks, resource usage, connections and achievements thus far. Similar to the use of academic analytics, the learner may use these visualizations to monitor their learning progress on a daily basis. This project distills the following activities as the most relevant to aid learners in their academic self-regulation in MOOC learning environments:

- Monitoring progress toward course-defined goals with assessment results, interaction patterns with peers, instructor and course resources
- Identifying patterns that impact decision-making
- Presenting information in an aesthetically pleasing way, inviting users to explore their data
- Dialectically highlighting differences between historically successful and unsuccessful learning trajectories
- Allowing the user to set personal learning goals that are tracked by the LAT

Humans acquire visual information in bursts (Ware, 2012). To initially capture a learner’s attention, it is critical that important information is made visually distinct (Duval, 2011). Once recognized, this information is then processed in a slow, linear fashion, but first the image must garner the learners’ visual attention. Learners not only need to be able to see their data, but rather than see in the literal sense, they need to be able to make sense of the images before them. This LAT uses Klein et al.’s definition of sensemaking as a “motivated, continuous effort to understand connections… in order to anticipate their trajectories and act effectively” (Klein, Moon, & Hoffman, 2006). Our connections, including social and technical systems, help us understand how people make sense of what they experience and even who they are. The learner must learn to appropriately evaluate this information from a sometimes-overwhelming flow of information, and then determine what knowledge is valuable enough to commit to long-term memory.

Data proposed for visualization

This LAT will utilize data automatically collected by the LMS and manual information logged by the learner to answer questions such as “if I keep performing at the same level, what final (summative) grade will I likely achieve in this course?” and “who in this course is interested in the things I am interested in?” The learner may obtain more detailed information by clicking on any of the objects within the LAT space.

A key design component of this tool lies in the idea that for learners to be able to improve, they must develop the capacity to monitor the quality of their learning strategies and the work they produce without the necessity of input from an instructor. The student must be able to set goals, self assess, and seek help as necessary. The learner must be able to identify areas of weakness from the visualization; as such graph reading is necessary. Learners are exposed to technology and data analysis, by familiarizing themselves with the data they know best – their own. LATs can help learners cultivate an academic mindset while learning the content of their chosen field.

The LAT is based on connectivist pedagogy (Siemens, 2005); the dominant visual theme is connectivity – the connections between people, learning spaces and activities. It also builds on George Seimens' idea of transparent learning (*TEDxNYED - George Siemens*, 2010). To paraphrase, Seimens believes that when an individual learns transparently, i.e. letting others see how they are learning, the individual acts as both learner and teacher within the learning space. Though Seimens was specifically addressing a teaching model based on networked resources, we extend this model to learning analytics tools visualizations that utilize the performance of others to guide individual learners through the course experience, orienting them within past and present learning communities.

Glyphs, graphical object designed to convey multiple data values (Ware, 2012), are used throughout the design to make it more aesthetically inviting to information novices. The data in the LAT visualization is presented in layers; with the learner's own data on the top layer. The background layers are composed of aggregations of data from current classmates and alumni of the course. This portion of the LAT relies on the accuracy of knowledge tracing and performance predictions.

Performance prediction and knowledge tracing

Knowledge tracing is a technique commonly used in learning analytics and educational data mining to infer learners' knowledge during the course of learning. This sort of performance prediction has historically been a major feature of learning analytics tools, used to identify students in need of help or in danger of dropping out. This is th kind of formative information a learner would get from an educator. Since learners have less access to educators in MOOC environments, LATs may be used as a source of formative feedback. Framed in Bandura's Social Cognitive theory (Bandura, 1977), this tool will enable self-evaluation through intrinsic reinforcement; a form of reinforcement that is potentially more influential that external reinforcement from others.

Performance prediction techniques include the use of decision trees, neural networks, k-nearest neighbor associations, classification and clustering, regression, and artificial intelligence. The use of each of these techniques is impacted by the pedagogical basis of the course, as well as the data capable of being collected from the LMS. For this study, we intend to use the k-nearest neighbor algorithm to make grade assessment predictions. The k-nearest neighbor algorithm classifies each record in a dataset, based on a combination of the records most similar to it and historic datasets.

Proposed visualization relationships

Learners may benefit from educational technologies that give them the ability to understand and reflect on their evolution as life-long learners. To make learners aware of themselves as evolving beings, the proposed visualizations will offer learners a time series overview of what transpires in the classroom. A time series is a set of data observed over time that is often used to make forecasts, like daily rainfall or stock market returns. At a glance, the learner will be able to view and interact with: upcoming and past assignments, formative and summative assessments, time spent on task, time spent in the LMS, time spent in forums and chats, manually collected information on the learners' self-regulatory behaviors and anonymized, aggregated information from previous learners. Visualizations to this information can help learners detect pedagogically important patterns in their assessments and classroom behaviours, alerting individual learners to maladaptive behaviours, and recommend next steps.

The first interaction that the learner will have with the LAT is to select the grade range that they hope to achieve and the amount of time that they believe will be spent studying each week to achieve their goal. This goal will be a starting point of analysis for the learner, allowing them to compare their predictions to the actual amount of time they spend in class and their resultant grades. The learner will be able to easily discern whether or not they are achieving their goals based on where their assessments' position within the screen.

Visualizations will be used to explore the relationships between all of the assorted learning activities by learning objectives, along with their related assessments, documents and discussions. The visualizations will relate learning objectives to learning activities, performance data with behavioural data and performance data with goals. Data such as the amount of time the learner has spent in the classroom is based on the daily data automatically collected by the LMS. The learner can review the time spent in class on a daily basis, as well as their historical average. If desired, the learner can manually indicate how much of this time was spent on an individual task, such as an individual or group assignment. Learners would use the visualizations to answer specific questions or to explore, gaining insight from the relationships visualized. An example would be a visualization that compares the average amount of time spent within the LMS by learners who achieved the summative grade average desired by the learner. This is a way to suggest what behaviors, in this case time on task, are predicted to achieve a certain goal.

Connections

A visualization of their connections would help learners by allowing them to analyze past discussions for learning content, or by highlighting current and potential collaborations that may help to orient the learner within the online class' community. Using the Fruchterman-Reingold visualization technique (Fruchterman & Reingold, 1991), learners can see where they fit within a community, along with the relative closeness of their discourse. The higher the betweenness centrality of the nodes, the higher the relationship is between the learner and the learning resource or peer the connected node represents. The visual representation of human connections allows learners the ability to quickly discern where they are in the class, by orienting them within the sphere. Unconnected nodes represent the possibility space of all of the people possible to interact with within the course.

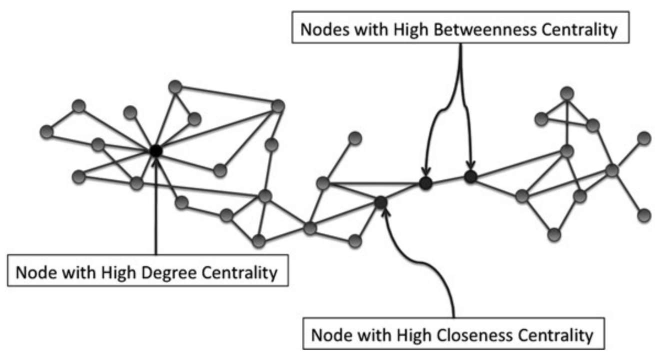

Figure 10. Node Relationship (Oshima et al., 2012).

Topics, trending & keywords

Tools that help the learner monitor what is going on in the classroom will help them find what they need quicker and orient them within the space. If a learner was having trouble with a particular learning objective, they could look it up by topic, with keywords or by searching the trending topics to locate related discussions. Similar to an RSS feed, trending topics could be used to know what discussions are going on. The data for the trending topics relies on the semantic assessment of both the LMS and any connected learning areas such as Facebook or the classes' hash tag on Twitter. This requires a connection to the API of various social media outlets, as well as semantic logs from the LMS. The learner will be able to search for course related keywords in conversations among classmates, representing a new way to incite classroom communication.

Applying a Vygotskian approach like the one taken by Teplovs et al. (Chris Teplovs et al., 2012), a semantic recommender in this LAT would takes the perspective that optimal learning occurs during interactions between individuals who are neither too similar or dissimilar, based on the semantics their written communication. The LAT could collect additional information from the learner to recommend conversations for the learner to join.

Based on the SNA work of Oshima et al. (J. Oshima et al., 2012), topic visualizations would use the same visualization technique, but rather than learners' names, the nodes will be course-related keywords from the learners' discourse. Used by educators to locate outliers, social network analysis shows learners their interaction patterns within a network of their peers and can be used to foster collaboration. A combination of social and semantic analyses can help learners identify the group discussions deemed most valuable to their learning.

Goal setting & monitoring

The addition of tagging functionality for learners' to annotate important events in their learning is inspired by Gismo's tagging feature. We hypothesize that reflection on the actual activity of learning, rather than the subject matter at hand, increases the value of the LAT for learners. An example use case would be marking trends with important auxiliary information. Perhaps a learner's grades dropped because their children were ill or there was a death in the family. Conversely, the learner may want to mark changes in their study habits to determine if these actions have a direct impact on their assessments.

In practice, the goal monitoring would aggregate periodic prompts asking the user to assess their performance on their chosen goal, or to rate their progress over time. The learner could be prompted to enter information on their experience such as the amount of time spent on tasks, and their perceived levels of self-efficacy. The results of this reflection will also be aggregated for the learners' later reflection. If the goals of classmates are shared, learners may be inspired by the goal choices of their peers.

Personalized experience tracing

The exacting accuracy possible through the use of log files is sometimes in direct opposition to learners' memory of what transpired. This is often more about how they felt over the course of achieving a goal, their self-efficacy or their interactions with peers, rather than the actual mechanics of achievement. Though not temporally accurate, the learners' affect and self-efficacy during the process of learning are important to log and present to learners for reflection.

There is a good deal of activity module data collected on assignments, chats, quizzes, wikis and discussion forums in LMS. A Cloze question is one of the types of questions readily available in most LMS. It is a question for which a learner must fill in a piece of removed text. This could be useful for collecting information on the learners' affect and self-efficacy. This type of question could be used to answer questions about learners' experience such as "how did you feel when you began this assignment," "what skills did you master through this course" or "what did you learner about yourself from taking this class?" The learning analytics tool could incorporate a question bank similar to Moodle, so that questions could be reused and combined in any combination. An example of a combination question would be to use how many times a learner attempted a question – information already collected in LMS – against their perceived level of self-efficacy, presented as a Cloze question. Learners could further personalize their activity data by adding tags to different areas to denote their moods, content they want to explore further, etc. The information would have to be collected manually, so learners must be encouraged to use this feature.

Chapter 5: Conclusion and Future Directions

Learners have the greatest ability to implement the changes necessary to improve their learning; access to their data will only increase this ability. Analytics tools have filtered down from the institutional level into individual classrooms. It naturally follows that the next user group to benefit directly from LATs should be learners. Interaction with a LAT may be used to explicitly teach metacognition and self-efficacy (Hannafin & Land, 1997; J. R. Hill & Hannafin, 1997), which may be particularly helpful for non-traditional learners and learners with underdeveloped self-regulatory skills.

Future work will first entail the design of visualizations to support the data identified in this project to help learners in MOOC environments. These visualization prototypes will be empirically validated with learners and visualization experts to explore the degrees of understanding and awareness that learners may gain from the described visual representations of learning performance data, and how learners may use the data displayed to regulate their learning. It must be recognized that increased awareness, even with an abundance of data, does not necessary lead to change. The learner may not know how to address the issue, or know what changes to implement. In a small classroom this might prompt the learner to consult the instructor to be told what their next steps should be. In a MOOC environment, this may not be possible due to the sheer number of learners who want or need feedback. The use of recommender systems similar to those used by Amazon and Netflix to offer the best item or experience for customers, may be the best answer to this. The automated provision of advice or recommendations would help to address this issue, but only if it provided the level of personalization learners need to move forward with the advice.

One of the major limitations to the success of this tool is learners' numeric literacy, along with their ability to understand the abstractions used in the visualizations. In a recent survey among learning analytics experts, only 21% of the 111 respondents felt "that learners would possess the required competences to interpret LA results themselves and determine appropriate actions/interventions from it" (Drachsler & Greller, 2012). A thorough tutorial may help learners' better understand the LAT, but their use of the tool must be empirically studied in situ.

Denzin and Lincoln (Denzin & Lincoln, 2005) recommend multiple methods of inquiry. Using a mixed methodology framework, future empirical study will use behavioural theory, visualization and human computer interaction techniques to provide a mixed methods approach to constructing understanding around the attitudes and activities of learners in online educational environments, providing a realistic view of the needs and practices of these users.

An important component of the value of a visualization lies in its long-term repeated use (Thomas & Cook, 2006). Allowing the user to explore the data gives them a chance to learn the system in a relaxed environment. After the users familiar with the system, the researcher can learn about the users analysis process and findings during natural tasks.

The goal of visualization is to prompt higher order cognitive thought, aiding the learner to gain insight or knowledge based on the unique presentation of or interactions with visual representations of data. Optimally, the user will gain new insight in addition to learning explicit information about the presented data set. Kielman and Ribarsky propose a learning-based evaluation method to measure insight by measuring how users would reapply the knowledge they've gained to a different problem or visual analytics system (Chang, Ziemkiewicz, Pyzh, Kielman, & Ribarsky, 2010). This methodology is unique in that it is not user or domain specific, making it replicable and generalizable. That said, this also presents a weakness in that domain specific information would be lost if this was the sole evaluation method used.

Insight is very subjective – it could be a way to solve a problem or simply a new perspective on the presented data set. It differs among individuals and from study to study, so it is not easily defined across the body of knowledge of information visualization. Evaluating for insight presents challenges due to the varying and sometimes lengthy amount of time educators and learners may need to become familiar with the system, before any insights have been achieved. Often times, information visualization systems are evaluated on usability or lower-level tasks completion activities that are easily defined and replicated for this reason.

Subsequent iterations of the tool will be studied with both learners and education experts to validate its usefulness and contribute to the establishment of benchmark tasks for future study in learning analytics. Findings from future research will contribute to theory and practice in the areas of learning analytics, addressing the visualization of student data for learning at the post-secondary level, also addressing contemporary concerns regarding student retention.

Conclusions

MOOCs have enjoyed one of the fastest adoption rates ever seen in the history of higher education (*NMC Horizon Report, 2013 Higher Education Edition*, 2013). As this explosion of interest in online learning increases the number and variety of learners in these classrooms, it also increases the need for educational technologies that support them, easily scaling up to provide a personalized learning experience to each learner.

Novel learning analytics tools have the potential to personalize assessment feedback for learners in ways never experienced, by using the students' own data to guide their learning. Learning management systems collect this data, but as this data is filtered, analyzed, and visualized, it undergoes syntactical and semantic translations at each stage. To build effective LATs for learners in MOOCs, researchers and developers must be allowed access to the data and how it is analyzed within the LMS. The number of tools that used the Mulce dataset belies the need for more public datasets on MOOC learning environments, along with the need for standardized data formats and extraction tools.

Recently developed tools like GVIS, an integrating infrastructure for adaptively mashing up user data from different sources, may make data extraction from multiple sources more efficient; this should positively impact the availability of this data for use in LATs. If learners could port their data from all of the LMS where they have taken classes, they could create personal learning portfolios to track, reflect on, predict and strategize their learning behaviors - much in the same way that educational institutions mobilize learner data now. To effectively utilize the data in this way, LATs can be used to present the data to in such a way that learners may understand their personal skill competencies, peer learning relationships, and processes of learning.

It is established that progression of learner data is valuable to educational institutions, corporations, instructional designers, instructors and most importantly learners. In innovative learning environments where cutting edge engagement research is being applied within the classroom, the resultant data and methods of sharing that data with learners follows an antiquated model. Udacity selects, trains and films the professors who teach its courses to standardize the quality of the course delivery and content. The same care should be taken to train learners how to get the most out of their learning experience.

Based on connectivist pedagogy and transparent learning, the LAT reflects new learning analytics visualization methods for use by learners. Using the log data exposed by open source LMS such as Moodle, the proposed LAT enables self-evaluation through intrinsic reinforcement, which according to Bandura's social cognitive theory of self-regulation (Bandura, 1991), is potentially more influential that external reinforcement. These logs — observable representations of cognitive, metacognitive and motivational events — are key to more fully modeling the learning processes of non-traditional learners in MOOC environments and supporting their learning with educational technologies designed for their use.

The email received by the representative from edX raises a point important to this research – two of the three most prominent MOOCs are start-up companies who have yet to solidify a sustainable business model. To be sustainable, MOOCs need to draw as many learners as possible, keep them engaged and turn them into repeat customers – one class at a time. This is a lofty endeavor, one that all institutions of higher education seek to achieve and one that may depend on the very data we seek for use in the proposed LAT tool. Another promising revenue stream for MOOCs could come through accreditation tests, or the finder fees employers pay for any hires made through the MOOCs' career services (Ripley, 2012). As learner trace data is shared with more and more entities, it is increasing important that the algorithms, methods of analysis and visualization be made transparent.

Where MOOCs will fit into the overall learning landscape remains to be seen, but one thing is clear – their influence, particularly the technology behind them is a rich opportunity for educational researchers. Open source edX tools will be made available to developers in June 2013, but it is not yet known if the data collection or the algorithms used to display this information will be shared (Agarwal, 2013). Shared standards and competency assessments may mean that educational data could be uniformly collected from kindergarten to college, to career and beyond. This has the potential to provide a more complete lifelong learning portfolio for learners. Data security and sharing issues aside, this presents a rich opportunity for research into the myriad of ways that learning is achieved online and the educational technologies designed to support this process.

Bibliography

Academic Analytics: A New Tool for a New Era. *EDUCAUSE REVIEW*, *42*(4), 40–50. Retrieved from http://www.educause.edu/ero/article/academic-analytics-new-tool-new-era

Anderson, T., & Dron, J. (2011). Three Generations of Distance Education Pedagogy. *The International Review of Research in Open and Distance Learning*, *12.3*, 1–18.

Agarwal, A. (2013, April 4). Stanford and edX Collaborate on OpenSource edX Platform. *edX Blog*. Retrieved April 4, 2013, from http://blog.edx.org/

Aleven, V., Mclaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. *International Journal of …*.

Arnold, K. E. (2010). Signals: Applying Academic Analytics. *EDUCAUSE REVIEW*, 1–10. Retrieved from http://www.educause.edu/ero/article/signals-applying-academic-analytics.

Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using Learning Analytics to Increase Student Success. *LAK '12: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 1–4.

Bache, K., & Lichman, M. (2013). UCI Machine Learning Repository: Data Sets. *University of California, School of Information and Computer Science*. Irvine, CA. Retrieved April 2013, from http://archive.ics.uci.edu/ml.

Baker, RSJ d, Corbett, A. T., & Koedinger, K. R. (2006). Adapting to when students game an intelligent tutoring system. *Intelligent Tutoring …*.

Baker, RSJ, & de Carvalho, A. (2008). Labeling student behavior faster and more precisely with text replays. *… of the 1st International Conference on …*.

Baker, RSJ, Corbett, A. T., & Aleven, V. (2008a). Improving contextual models of guessing and slipping with a truncated training set. *Human-Computer Interaction ….*

Baker, R.S.J.D., Duval, E., Stamper, J., Wiley, D., & Buckingham Shum, S. (2012). Panel: Educational Data Mining meets Learning Analytics. *LAK '12: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 1–3.

Baker, Ryan S J d. (2007). Modeling and understanding students' off-task behavior in intelligent tutoring systems. Presented at the CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM Request Permissions. doi:10.1145/1240624.1240785

Baker, Ryan S, Corbett, A. T., & Aleven, V. (2008b). More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing. Presented at the ITS '08: Proceedings of the 9th international conference on Intelligent Tutoring Systems, Springer-Verlag.

Baker, Ryan S, Corbett, A. T., Roll, I., & Koedinger, K. R. (2008c). Developing a generalizable detector of when students game the system. *User modeling and user-adapted interaction*, *18*(3).

Bakharia, A., & Dawson, S. (2011). SNAPP: a bird's-eye view of temporal participant interaction, 168–173.

Bandura, A. (1977). Self-Efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, *84*, 191–215.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*, 248–287.

Barneveld, A. V., Arnold, K. E., & Campbell, J. P. (2012). Analytics in Higher Education: Establishing a Common Language. *EDUCAUSE Learning Initiative*, 1–11. Retrieved from http://www.google.ca/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CFUQFjAB&url=http%3A%2F%2Fnet.educause.edu%2Fir%2Flibrary%2Fpdf%2FELI3026.pdf&ei=8WBkUeWIA8TEigKyo4DoCg&usg=AFQjCNEn0clkvpn7kb8p15JDofueFM7Gag&sig2=0CPID4ER5EvRj2WcZDY-xA&bvm=bv.44990110,d.cGE

Betbeder, M.-L., Ciekanski, M., & Chanier, T. (2010, September 5). Comparing learner interactions and media use in two groups working in parallel within an audio-graphic environment. *Mulce.org*. Retrieved April 2013, from http://mulce-doc.univ-bpclermont.fr/spip.php?article55

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. doi:10.1080/0969595980050102

Blackboard Inc. (2012, December 11). Blackboard Opens Field Trial for Learning Analytics Solution. *PR Newswire*. Retrieved November 2012, from http://www.prnewswire.com/news-releases/blackboard-opens-field-trial-for-learning-analytics-solution-137020838.html

Boekaerts, M., & Corno, L. (2005). Self-Regulation in the Classroom: A Perspective on Assessment and Intervention. *Applied Psychology: An International Review*, *54*, 199–231.
Campbell, J. P. (2007).

Brooks, C., Hansen, C., & Greer, J. (2006). Social awareness in the iHelp courses learning content management system.

Cadwallard, C. (2012, November 10). UDACITY: COURSERA: edX: Goodbye: to all this? Equities.com. Retrieved April 2013, from http://www.equities.com/news/headline-story?dt=2012-11-10&val=696948&cat=finance

Carmean, C., & Mizzi, P. (2010). The Case for Nudge Analytics. *EDUCAUSE REVIEW ONLINE*, 1–8. Retrieved from http://www.educause.edu/ero/article/case-nudge-analytics

Cetintas, S., Si, L., Xin, Y. P., Hord, C., & Zhang, D. (2009). Learning to identify students' off-task behavior in intelligent tutoring systems. Presented at the Proceedings of the 14th ….

Chacon, F. J., Spicer, D. Z., & Valbuena, A. (2012, April). Analytics in support of student retention and success. *EDUCAUSE REVIEW*. Retrieved March 2013, from http://www.educause.edu/library/resources/analytics-support-student-retention-and-success

Chang, R., Ziemkiewicz, C., Pyzh, R., Kielman, J., & Ribarsky, W. (2010). Learning-based evaluation of visual analytic systems, 29–34.

Ciekanski, M., & Chanier, T. (2008). Developing online multimodal verbal communication to enhance the writing process in an audio-graphic conferencing environment. *ReCALL*, *20*(2).

Consolvo, S., Markle, K., Patrick, K., & Chanasyk, K. (2009). Designing for persuasion: mobile services for health behavior change. *Persuasive '09: Proceedings of the 4th International Conference on Persuasive Technology*.

Cooper, S., & Sahami, M. (2013). Reflections on Stanford's MOOCs. *Communications of the ACM*.

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, *4*(4), 253–278.

Coursera Terms of Use. (2012). Coursera Terms of Use. *Coursera*. Retrieved March 2013, from https://www.coursera.org/about/terms

Creswell, J. W. (2012). *Qualitative Inquiry and Research Design*. SAGE Publications, Incorporated.

Datasets for Data Mining, Analytics and Knowledge Discovery. (2013). Datasets for Data Mining, Analytics and Knowledge Discovery. *KDnuggets*. Retrieved March 2013, from http://kdnuggets.com/datasets

Denzin, N. K., & Lincoln, Y. S. (2005). *The Sage handbook of qualitative research*.

DiCicco Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, *40*(4), 314–321. doi:10.1111/j.1365-2929.2006.02418.x

Dourish, P. (2001). *Where the action is: the foundations of embodied interaction*. The MIT Press.

Downes, S., Siemens, G., Cormier, D., & Kop, R. (2010). Personal Learning Environments Network and Knowledge (PLENK) 2010. *connect.downes.ca*. Retrieved April 2013, from http://connect.downes.ca

Drachsler, H., & Greller, W. (2012). The pulse of learning analytics understandings and expectations from the stakeholders. Presented at the LAK '12: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM Request Permissions. doi:10.1145/2330601.2330634

Draper, S. (2008). Tinto's Model of Student Retention. *www.psy.gla.ac.uk*. Retrieved March 2013, from http://www.psy.gla.ac.uk/~steve/localed/tinto.html

Duval, E. (2011). Attention please!: learning analytics for visualization and recommendation, 9–17.

D'Mello, S. K., Craig, S. D., Witherspoon, A., Mcdaniel, B., & Graesser, A. (2008). Automatic detection of learner's affect from conversational cues. *User modeling and user-adapted interaction*, *18*(1-2), 45–80.

EdX Privacy Policy. (2012, September 26). edX Privacy Policy. *edX*. Retrieved April 2013, from https://www.edx.org/privacy

Ferguson, R. (2012). The State of Learning Analytics in 2012: A Review and Future Challenges. *Knowledge Media Institute, The Open University*, 1–18. Retrieved from http://kmi.open.ac.uk/publications/techreport/kmi-12-01

Forsythe, R. G. J., Chacon, F. J., Spicer, D. Z., & Valbuena, A. (2012, August 31). Two Case Studies of Learner Analytics in the University System of Maryland. Retrieved October 2012.

Friedman, T. L. (2013, January 26). Revolution Hits the Universities. *NYTimes.com*. Retrieved April 2013, from http://www.nytimes.com/2013/01/27-/opinion/sunday/friedman-revolution-hits-the-universities.html?pagewanted=1&_r=0

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164. doi:10.1002/spe.4380211102

Greller, W. (2012). Open Learning Analytics. Reflections on the Knowledge Society (pp. 1–3). Reflections on the Knowledge Society. Retrieved from http://www.greller.eu/wordpress/?p=2573

Gottardo, E., & Vida Noronha, R. (2012). Social Networks Applied to Distance Education Courses: Analysis of Interaction in Discussion Forums. *Web Media 2012*, 1–4.

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, *2*, 107–124. doi:10.1007/s11409-007-9016-7

Hannafin, M. J., & Land, S. M. (1997). The foundations and assumptions of technology-enhanced student-centered learning environments. *Instructional Science*, *25*, 167–202.

Hill, J. R., & Hannafin, M. J. (1997). Cognitive Strategies and Learning from the World Wide Web. *ducational Technology Research and Development*, *45*, 37–64.

Horn, L. J., & Carroll, C. D. (1996). Nontraditional Undergraduates: Trends in Enrollment from 1986 to 1992 and Persistence and Attainment Among 1989-90 Beginning Postsecondary Students (NCES 97-578). *U.S. Department of Education Office of Educational Research and Improvement*, 1–72.

Howison, J., Conklin, M., & Crowston, K. (2006). FLOSSmole: Collaborative collection and analysis of free/libre/open source project data. *International Journal of Information Technology and Web Engineering*. Retrieved April 25, 2013, from http://flossmole.org/

Isenberg, P., Tang, A., & Carpendale, S. (2008). An exploratory study of visual information analysis, 1217–1226.

Jeremić, Z., Milikić, N., Jovanović, J., Radulović, F., Brković, M., & Devedžić, V. (2011). OP4L: Online Presence Enabled Personal Learning Environment, 417–420.

Jovanović, J., Gašević, D., Brooks, C., Devedžić, V., & Hatala, M. (2007). A Quantitative Evaluation of LOCO-Analyst: A tool for Raising Educators' Awareness in Online Learning Environments. In *Creating New Learning Experiences on a Global Scale* (Vol. 4753, pp. 112–126). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/978-3-540-75195-3_9

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *Intelligent Systems, IEEE*, *21*(4), 70–73.

Koedinger, K. R., Baker, R. S. J. D., Cunningham, K., & Skogsholm, A. (2013). A Data Repository for the EDM community: The PSLC Data Shop, 1–21.

Libraries, A. O. C. A. R. (2013, January 30). Information Literacy Competency Standards for Higher Education. *www.ala.org*. Retrieved January 2013, from http://www.ala.org/acrl/standards/informationliteracycompetency

Litman, D. J., & Forbes-Riley, K. (2006). Recognizing student emotions and attitudes on the basis of utterances in spoken tutoring dialogues with both human and computer tutors. *Speech Communication*.

Luo, H. (2011). Qualitative Research on Educational Technology: Philosophies, Methods and Challenges. *International Journal of Education*, *3*(2). doi:10.5296/ije.v3i2.857

MacWhinney, B. (2004). TalkBank: Building an open unified multimodal database of communicative interaction. *Department of Psychology*, 185.

Mazza, R. (2006). Evaluating information visualization applications with focus groups: the CourseVis experience, 1–6.

Mazza, R., & Botturi, L. (2007). Monitoring an Online Course with the GISMO Tool: A Case Study. *Journal of Interactive Learning Research*, *18*, 251–265.

Mazza, R., & Dimitrova, V. (2003). Informing the design of a course data visualisator: an empirical study. *5th International Conference on New ….*

Mazza, R., & Dimitrova, V. (2004). CourseVis: Externalising Student Information to Facilitate Instructors in Distance Learning. *Proceedings of the International conference in Artificial Intelligence in Education*, 1–8.

Mazza, R., & Dimitrova, V. (2007). CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses. *International Journal of Human-Computer Studies*, *65*(2), 125–139. doi:10.1016/j.ijhcs.2006.08.008

McNely, B. J., Gestwicki, P., Hill, J. H., Parli-Horne, P., & Johnson, E. (2012). Learning analytics for collaborative writing: a prototype and case study. Presented at the LAK '12: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM Request Permissions. doi:10.1145/2330601.2330654

McQuiggan, S. W., Mott, B. W., & Lester, J. C. (2007). Modeling self-efficacy in intelligent tutoring systems: An inductive approach. *User modeling and user-adapted interaction*, *18*(1-2), 81–123. doi:10.1007/s11257-007-9040-y

Microsoft PerformancePoint. (2013). Microsoft PerformancePoint. *Microsoft Office*. Microsoft. Retrieved from http://msdn.microsoft.com/en-us/office/bb660518.aspx

Microsoft SQL Server. (2013). Microsoft SQL Server. Microsoft. Retrieved from http://www.microsoft.com/en-us/sqlserver/default.aspx

Miratrix, M. (2013, April 18). LMS data collected. *Gmail*. Retrieved April 19, 2013.

MIT OpenCourseWare: Our History. (2013, April 29). MIT OpenCourseWare: Our History. *MIT OpenCourseWare*. Retrieved April 2013, from http://ocw.mit.edu/about/our-history

Morris, R., Hadwin, A. F., Gress, C. L. Z., Miller, M., Fior, M., Church, H., & Winne, P. H. (2010). Designing roles, scripts, and prompts to support CSCL in gStudy. Computers in Human Behavior, 26(5), 815–824. doi:10.1016/j.chb.2008.12.001

Mulce Repository Documentation. (2012). Mulce repository Documentation. *Mulce.org*. Clermont-Ferrand:France: Mulce (MULtimodal Corpus Exchange); Universite Blaise Pascal. Retrieved April 20, 2013, from http://mulce-doc.univ-bpclermont.fr/*NMC Horizon Report, 2013 Higher Education Edition*. (2013).

National Center for Educational Statistics. (2002). The Condition of Education: Special Analysis 2002: Nontraditional Graduates. *National Center for Educational Statistics*, 1–26. Retrieved from http://www.nces.ed.gov/programs/coe/2002/analyses/nontraditional/sa05.asp

NMC Horizon Report, 2013 Higher Education Edition (pp. 1–44). The New Media Consortium.

Norris, D. M., Baer, L. L., Leonard, J., Pugliese, L., & Lefrere, P. (2008, January 18). Framing Action Analytics and Putting Them to Work. *EDUCAUSE REVIEW*.

Retrieved October 2012, from http://www.educause.edu/ero/article/framing-action-analytics-and-putting-them-work

Open Learning Initiative. (2013, April 25). Open Learning Initiative. *Open Learning Initiative*. Retrieved March 2013, from http://oli.cmu.edu/

Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse. *Educational Technology Research and Development*, *60*(5), 903–921. doi:10.1007/s11423-012-9265-2.pdf

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed., pp. 1–598). Thousand Oaks, CA: Sage Publications.

Pavlik, P. I., Cen, H., & Koedinger, K. R. (2009). Learning factors transfer analysis: Using learning curve analysis to automatically generate domain models, 121–130.

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33–40.

Pistilli, M. D., Arnold, K. E., & Bethune, M. (2012). Signals: Using Academic Analytics to Promote Student Success. *EDUCAUSE REVIEW ONLINE*, 1–8. Retrieved from http://www.educause.edu/ero/article/signals-using-academic-analytics-promote-student-success

Radford, A. W. (2011). Learning at a Distance: Undergraduate Enrollment in Distance Education Courses and Degree Programs (NCES 2012-154). (T. Weko, Ed.)*National Center for Education Statistics*, 1–22.

Reffay, C., & Chanier, T. (2003). How social network analysis can help to measure cohesion in collaborative distance-learning, 343–352.

Reigeluth, C. M. (1983). *Instructional-design theories and models* (Vol. 3, pp. 1–1). Hillsdale, N.J. : Lawrence Erlbaum Associates Publishers.

Ripley, A. (2012, October 13). College Is Dead. Long Live College! *Time Magazine*. Retrieved March 2013, from http://nation.time.com/2012/10/18/college-is-dead-long-live-college/print/

Ross, S. M., Morrison, G. R., & Lowther, D. L. (2010). Educational technology research past and present: Balancing rigor and relevance to impact school learning. *Contemporary Educational Technology*, *1*(1), 17–35.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144.

Science, G. N. L. O., & Technology (Eds.). (2009). DataCite Repository. German National Library of Science and Technology. Retrieved from http://www.datacite.org/

Seiffert, D. (2012, October 12). University of Texas joins edX. Boston Business Journal. Retrieved March 2013, from http://www.bizjournals.com/boston/blog/mass-high-tech/2012/10/university-of-texas-joins-edx.html

Shih, B., Koedinger, K. R., & Scheines, R. (2011). A response time model for bottom-out hints as worked examples. *Handbook of Educational Data Mining*, 201–212.

Siemens, G. (2005). Connectivism: A learning theory for the digital age.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE REVIEW*, *46*(5), 1–6.

SIG dataTEL: Data-driven Research and Learning Analytics. (2007). SIG dataTEL: Data-driven Research and Learning Analytics. *EATEL*. Retrieved April 2013, from http://ea-tel.eu/sig-datatel

Stengel, R. (2012, October 13). Reinventing College. *Time Magazine*. Retrieved March 2013, from http://nation.time.com/2012/10/18/reinventing-college/

Student Starfish Tutorial. (2011). *Student Starfish Tutorial* (pp. 1–2). Bowie, MD: Bowie State University. Retrieved from http://www.bowiestate.edu/about/cabinet/AcademicAffairs/computing_support/sst/

Subotzky, G., & Prinsloo, P. (2011). Turning the tide: a socio-critical model and framework for improving student success in open distance learning at the University of South Africa. *Distance Education*, *32*(2), 177–193. doi:10.1080/01587919.2011.584846

Swail, W. S. (2004). The art of student retention: A handbook for practitioners and administrators, *21*, 2004.

TEDxNYED - George Siemens. (2010). *TEDxNYED - George Siemens* (pp. 1–3). New York: YouTube. Retrieved from http://www.youtube.com/watch?v=4BH-uLO6ovI

Teplovs, Chris, Fujita, N., & Vatrapu, R. (2012). Generating Predictive Models of Learner Community Dynamics. *1st international conference on Learning Analytics & Knowledge*, 147–152. doi:10.1145/2090116.2090139

Teplovs, Christopher. (2010, October 13). Knowledge Space Visualizer. Chris Teplovs. Retrieved from http://chris.ikit.org/ksv/

The Dataverse Network. (1997). The Dataverse Network. *President & Fellows Harvard University*. Retrieved April 25, 2013, from http://thedata.org/

Thomas, J. J., & Cook, K. A. (2006). Illuminating the Path: Research and Development Agenda for Visual Analytics. (T.-M. Rhyne, Ed.)*IEEE Computer Graphics and Applications*, 10–13.

Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, *45*(1), 89–125. doi:10.2307/1170024

Tufte, E. (1990). Envisioning information. *Envisioning information*.

Udacity Privacy Policy. (2013). Udacity Privacy Policy. *Udacity*. Retrieved March 2013, from https://www.udacity.com/legal/privacy#sec3

University, Purdue. (2011a). Signals at Purdue University. *Purdue University Information Technology*. Retrieved February 2013, from http://www.itap.purdue.edu/studio/signals/

University, Purdue. (2011b). Course Signals - Stoplights for Students. *Purdue University Information Technology*. Retrieved November 2012, from http://www.itap.purdue.edu/learning/tools/signals/

U. S. Bureau of Labor Statistics. (2012, June 22). *American Time Use Survey Summary*. Retrieved from http://www.bls.gov/news.release/atus.nr0.htm/

U.S. Department of Labor. (2013). Minimum Wage. *United States Department of Labor*. Retrieved March 2013, from http://www.dol.gov/dol/topic/wages/minimumwage.htm

Vardi, M. Y. (2012). Will MOOCs destroy academia? *Communications of the ACM*, *55*(11), 5–5. doi:10.1145/2366316.2366317

Vetter, A., & Chanier, T. (2006). Supporting oral production for professional purposes in synchronous communication with heterogenous learners. *ReCALL*, *18*(1).

Visual Analytics Benchmark Repository. (2013, April 25). Visual Analytics Benchmark Repository. *SEMVAST Project*. Retrieved March 2013, from http://hcil.cs.umd.edu/localphp/hcil/vast/archive/viewbm.php

Visual analytics for everyone. (2003). Visual analytics for everyone. *Tableau*. Retrieved January 2012, from http://www.tableausoftware.com

Walonoski, J. A., & Heffernan, N. T. (2006). Detection and analysis of off-task gaming behavior in intelligent tutoring systems. Presented at the ITS'06: Proceedings of the 8th international conference on Intelligent Tutoring Systems, Springer-Verlag.

Ware, C. (2012). *Information Visualization: Perception for Design* (3rd ed., pp. 1–536). Morgan Kaufmann.

WCET Cooperative for Educational Technologies (May 2012). Predictive Analytics Reporting (PAR) Framework. *Predictive Analytics Reporting (PAR) Framework*. Retrieved March 2013, from http://wcet.wiche.edu/advance/par-framework

Webley, K. (2012a, September 20). Rethink College: 3 Takeaways from the TIME Summit on Higher Education. *Time Magazine*. Retrieved March 2013, from http://nation.time.com/2012/10/19/rethink-college-3-takeaways-from-the-time-summit-on-higher-education/

Webley, K. (2012b, October 19). Can an Online Degree Really Help You Get a Job? *Time Magazine*. Retrieved March 2013, from http://nation.time.com/2012/10/18/can-an-online-degree-really-help-you-get-a-job/

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning in Metacognition in educational theory and practice. (D. J. Hacker, J. Dunlosky, & A. C. Graesser, Eds.)*The educational psychology series*, 277–304.

Woolf, B. P. (2010). A roadmap for education technology.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329–339.

Zimmerman, B. J. (2000). *Attaining self-regulation: A social cognitive perspective*. (M. Boekaerts, P. R. Pintrich, & M. Zeidner, Eds.) (pp. 13–39). San Diego, CA: Academic Press.