It’s a pleasure to welcome you to Marist College and our scenic and historic Hudson River Valley. Since Marist was a founding member of the Society for Learning Analytics Research (SoLAR), we’re delighted to welcome our friends and colleagues from around the world to the 5th International Learning Analytics and Knowledge (LAK) conference. Over the past 36 years, I have been privileged to serve as president of Marist, and I have witnessed how the College has grown from a local institution to one with an international reputation for excellence. It is from this perspective that we take great pride in hosting this gathering of the world’s leading researchers and practitioners in the field of learning analytics. As we have prepared for the conference, I have been struck by this year’s theme, “Scaling Up: Big Data to Big Impact,” as it is not only appropriate for the event but for Marist as well. Why? Just as the field of learning analytics is starting to have a big impact on some of higher education’s biggest challenges, the College has also been scaling up our pioneering research in the area of open learning analytics, both across our campus and at campuses throughout the country. Of course, Marist could not be doing this work alone, which is why we are honored to be part of this innovative academic research community and host of this conference. Again, welcome to the 2015 LAK conference and to Marist College!

DENNIS J. MURRAY
PRESIDENT, MARIST COLLEGE

A WELCOME FROM PRESIDENT DENNIS J. MURRAY

LAK 2015 CONFERENCE SPONSORS

Platinum Sponsor

McGraw Hill Education

Gold Sponsors

bluecanary
Oracle
IBM

Silver Sponsors

ani
Blackboard
canvas
Intel Education
Unicon
Welcome to Learning Analytics & Knowledge 2015 (LAK15), the fifth edition of this international conference. This year LAK takes place at Marist College in Poughkeepsie, New York (USA), a city located on the banks of the majestic Hudson River, midway between Albany, the state capital, and New York City. The theme of LAK15, Scaling Up: Big Data to Big Impact, reflects our growing community of researchers, practitioners, and learners and our success in leveraging the power of “big data” to create substantial impact within higher education and learning at increasingly larger scales, while simultaneously reflecting the need for analytics to be effective at the level of individual learning (n of 1). Building on the momentum generated in earlier conferences and in recognition of our growth and of our mission ‘inventing effective means to improve the way students learn’, we have introduced a practitioner track this year. We hope that addressing and discussing the learning from two points of view; that of the researcher/learner and the practitioner, will provide each of us some impulse to think diversely about the many stakeholders in learning analytics and to open new perspectives on the intersection between research and the practice.

LAK15 received record numbers of high quality submissions. The research track received 138 submissions and the practitioner track received 33. We made a decision early in our planning to broaden the program committee to include representation from related fields and to bring in prominent academics that are involved in LA work. As a result, the number of reviews per reviewer remained manageable. Out of the 74 full papers received, it has not been an easy task to select the 20 accepted for presentation, resulting in a 27% acceptance rate, the same rate as last year. The proceedings include a further 39 short papers, 1 panel, 1 demo and 19 posters. The program of the conference is completed by 14 presentations, 2 panels and 9 technology showcases of the practitioner track. The pre-conference program includes a doctoral consortium with 7 students, 6 workshops and 2 tutorials.

We are aware of the need for the LAK conference to continue to improve the quality of reviewer feedback as well as the rigor of the review process. LAK is a broad umbrella that brings together learning scientists and computer scientists. The rich learning that happens during presentations and hallway conversations is a result of diverse viewpoints. For LAK to continue to grow, the review process and diverse representation (researcher, learners, administrators, practitioners) must continue to improve and evolve. Our goal, and that of the Society for Learning Analytics Research (SoLAR) more broadly, is to create a welcoming event where researchers find a fertile space for dialogue, thinking, critique, and advancement of learning in practical and important ways that touch classrooms and policy.

We hope that you will enjoy LAK15 and find a space for your work in our growing community. In a span of five years, we have moved from a small research cluster to a community and organization that has received recognition in government policy, granting agencies, and in both administration and teaching in universities. The LAK conference and SoLAR today reflect a broadening community, one that is poised to play a central role in guiding educational innovation and the use of big data for big impact globally. We feel confident that this event will help continue to define who we are as a community and how our work can impact the important challenges and opportunities facing education today.

Paulo Blikstein
Stanford University

Agathe Merceron
Beuth University of Applied Sciences Berlin

George Siemens
University of Texas, Arlington
**LAK 15 Schedule**

The registration desk will be open 7:30 a.m.–5:00 p.m. March 16–17, and 7:30 a.m.–12:00 noon March 18 just outside the entrance of the Nelly Goletti Theatre located in the Student Center (see map on inside back cover)

### Pre-Conference Workshops, Tutorials and Hackathon at a Glance

<table>
<thead>
<tr>
<th>TIME</th>
<th>MONDAY, MARCH 16</th>
<th>TUESDAY, MARCH 17</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00 a.m.–8:30 a.m.</td>
<td>Continental Breakfast (SC 3104-3105)</td>
<td>Continental Breakfast (SC 3104-3105)</td>
</tr>
<tr>
<td>8:30 a.m.–5:00 p.m.</td>
<td><strong>First Annual Open Learning Analytics Hackathon</strong> <em>(Cabaret)</em></td>
<td><strong>First Annual Open Learning Analytics Hackathon</strong> <em>(Cabaret)</em></td>
</tr>
<tr>
<td></td>
<td>Facilitators: Alan Berg and Sandeep Jayaprakash</td>
<td>Facilitators: Alan Berg and Sandeep Jayaprakash</td>
</tr>
<tr>
<td></td>
<td><strong>VISLA15: VISual Approaches to Learning Analytics (MU 3204)</strong></td>
<td><strong>VISLA15: VISual Approaches to Learning Analytics (MU 3204)</strong></td>
</tr>
<tr>
<td></td>
<td>Authors: Erik Duval, Katrien Verbert, Joris Klerkx, Martin Wolpers, Abelardo Pardo, Sten Govaerts, Denis Gillet, Xavier Ochoa and Denis Parra</td>
<td>Authors: Erik Duval, Katrien Verbert, Joris Klerkx, Martin Wolpers, Abelardo Pardo, Sten Govaerts, Denis Gillet, Xavier Ochoa and Denis Parra</td>
</tr>
<tr>
<td></td>
<td><strong>Doctoral Consortium (MU 3202)</strong></td>
<td><strong>Doctoral Consortium (MU 3202)</strong></td>
</tr>
<tr>
<td></td>
<td>Chairs: Dan Suthers, Simon Buckingham Shum, Katherine Maillet, Stephanie Teasley</td>
<td>Chairs: Dan Suthers, Simon Buckingham Shum, Katherine Maillet, Stephanie Teasley</td>
</tr>
<tr>
<td>8:30 a.m.–12:30 p.m.</td>
<td><strong>It’s About Time: 4th International Workshop on Temporal Analyses of Learning Data</strong> <em>(MU 3204)</em></td>
<td><strong>Topic Modeling for Learning Analytics Researchers</strong> <em>(MU 3203)</em></td>
</tr>
<tr>
<td></td>
<td>Authors: Simon Knight, Alyssa F. Wise, Bodong Chen and Britte Haugan Cheng</td>
<td>Authors: Vitomir Kovanovic, Srećko Joksimovic and Dragan Gasevic</td>
</tr>
<tr>
<td></td>
<td><strong>Ethics and Privacy in Learning Analytics Workshop (MU 3203)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Authors: Hendrik Drachsler, Adam Cooper, Tore Hoel, Rebecca Ferguson, Alan Berg, Maren Scheffl, Gábor Kismihók, Christien Bok and Weiqin Chen Chen</td>
<td></td>
</tr>
<tr>
<td>10:00 a.m.–10:15 a.m.</td>
<td><strong>Break (SC 3104-3105)</strong></td>
<td><strong>Break (SC 3104-3105)</strong></td>
</tr>
<tr>
<td>12:30 p.m.–1:00 p.m.</td>
<td><strong>Box Lunch (Pick up in SC 3104-3105)</strong></td>
<td><strong>Box Lunch (Pick up in SC 3104-3105)</strong></td>
</tr>
<tr>
<td>1:00–5:00 p.m.</td>
<td><strong>The 2nd International Workshop on Open Badges in Education—From Learning Evidence to Learning Analytics (MU 3204)</strong></td>
<td><strong>Hands-on with Bayesian Knowledge Tracing: Assessment and Hypothesis Testing in Digital-Ed (MU 3203)</strong></td>
</tr>
<tr>
<td></td>
<td>Authors: Daniel Hickey, Jelena Jovanovic, Steven Lonn and James Willis</td>
<td>Author: Zach Pardos</td>
</tr>
<tr>
<td></td>
<td><strong>The 3rd LAK Data Challenge (MU 3204)</strong></td>
<td><strong>The 3rd LAK Data Challenge (MU 3204)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>NOTE: Combined with VISLA15 (1-5 p.m.)</strong></td>
<td><strong>NOTE: Combined with VISLA15 (1-5 p.m.)</strong></td>
</tr>
<tr>
<td></td>
<td>Authors: Hendrik Drachsler, Mathieu D’Aquin, Eelco Herder, Stefan Dietze, Davide Taibi, Maren Scheffl and Adam Cooper</td>
<td>Authors: Hendrik Drachsler, Mathieu D’Aquin, Eelco Herder, Stefan Dietze, Davide Taibi, Maren Scheffl and Adam Cooper</td>
</tr>
<tr>
<td>2:30 p.m.–2:45 p.m.</td>
<td><strong>Break (SC 3104-3105)</strong></td>
<td><strong>Break (SC 3104-3105)</strong></td>
</tr>
</tbody>
</table>

See map on inside back cover for room locations
### Main Conference Sessions at a Glance

<table>
<thead>
<tr>
<th>TIME</th>
<th>WEDNESDAY, MARCH 18</th>
<th>THURSDAY, MARCH 19</th>
<th>FRIDAY, MARCH 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:45–8:45 a.m.</td>
<td>Breakfast (Main Dining Hall)</td>
<td>Breakfast (Main Dining Hall)</td>
<td>Breakfast (Main Dining Hall)</td>
</tr>
<tr>
<td>8:45–9:00 a.m.</td>
<td>Opening Remarks (Theatre)</td>
<td>Keynote Session* (Theatre)</td>
<td>Keynote Session* (Theatre)</td>
</tr>
<tr>
<td>9:00–10:00 a.m.</td>
<td>Keynote Session* (Theatre) Chuck Severance The Trials and Tribulations of Predicting the Future for Learning Technology</td>
<td>Keynote Session* (Theatre) Belinda Tynan Learning analytics = (problem+question) X (data+analysis) / (ethics+intervention)</td>
<td>Keynote Session* (Theatre) Danielle McNamara Getting a Big Picture from Big Data: Use Your Words!</td>
</tr>
<tr>
<td>10:00–10:15 a.m.</td>
<td>Break (SC 3102-3105)</td>
<td>Break (SC 3102-3105)</td>
<td>Break (SC 3102-3105)</td>
</tr>
<tr>
<td>10:15 a.m.–12 Noon</td>
<td>Concurrent Sessions 1A: MOOCs – Assessments, Connections and Demographics (Theatre) 1B: Student Engagement and Behaviour (Recital Hall) 1C: Indicators and Tools for Awareness (SC 3101)</td>
<td>Concurrent Sessions 4A: Off-Task Behaviour/ Bayesian Knowledge Tracing (Theatre) 4B: Writing and Discourse Analysis (Recital Hall) 4C: Learning Analytics Tools &amp; Frameworks (SC 3101)</td>
<td>Concurrent Sessions 7A: Multilevel, Multimodal and Network Analysis (Theatre) 7B: Curricula, Network and Discourse Analysis (Recital Hall) 7C: Open Learning Analytics and Tutoring Systems (SC 3101)</td>
</tr>
<tr>
<td>12 Noon–1 p.m.</td>
<td>Lunch (Main Dining Hall)</td>
<td>Lunch &amp; SoLAR Annual General Meeting (Main Dining Hall)</td>
<td>State of the Field Expert Panel &amp; Closing Remarks (Theatre)</td>
</tr>
<tr>
<td>1:00–2:30 p.m.</td>
<td>Concurrent Sessions 2A: Institutional Perspectives (Theatre) 2B: Students At Risk (Recital Hall) 2C: Practice Across Boundaries (SC 3101)</td>
<td>Concurrent Sessions 5A: Learning Strategies and Tools (Theatre) 5B: Text &amp; Discourse Analysis (Recital Hall) 5C: Theoretical Foundations &amp; Frameworks for LA (SC 3101)</td>
<td></td>
</tr>
<tr>
<td>2:30–2:45 p.m.</td>
<td>Break (SC 3102-3105)</td>
<td>Break (SC 3102-3105)</td>
<td>Box Lunch (pick up in main hallway) &amp; Depart Conference</td>
</tr>
<tr>
<td>2:45–4:30 p.m.</td>
<td>Concurrent Sessions 3A: Predicting Achievement (Theatre) 3B: MOOCs – Discussion Forums (Recital Hall) 3C: Student Performance (SC 3101)</td>
<td>Concurrent Sessions 6A: Alternative Methods of Improving Learning (Theatre) 6B: Interventions and Remediation (Recital Hall) 6C: Analyses with LMS Data (SC 3101)</td>
<td></td>
</tr>
<tr>
<td>4:30–4:45 p.m.</td>
<td>Transition Time (Move to Theatre)</td>
<td>Tech Showcase (SC 3102-3105)</td>
<td></td>
</tr>
<tr>
<td>4:45–5:30 p.m.</td>
<td>Firehose (Theatre)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5:30–6:00 p.m.</td>
<td>Posters and Welcome Reception (SC 3102-3105)</td>
<td>Load Buses for FDR (Follow signs)</td>
<td></td>
</tr>
<tr>
<td>6:00–7:00 p.m.</td>
<td>Conference Dinner at the Henry A. Wallace Center at the FDR Presidential Library and Home (Tickets Required, Purchase at Registration Desk)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7:00–9:00 p.m.</td>
<td>Dinner On Your Own (see Information Desk for assistance)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Directly following each keynote there will be a 30 minute discussion with the presenter in MU 3204.
Dr. Charles R. Severance

THE TRIALS AND TRIBULATIONS OF PREDICTING THE FUTURE FOR LEARNING TECHNOLOGY

Sponsored by McGraw-Hill Education

Dr. Charles R. Severance is a Clinical Associate Professor and teaches in the School of Information at the University of Michigan. He also works with the IMS Global Learning Consortium promoting and developing standards for teaching and learning technology. Previously he was the Executive Director of the Sakai Foundation and the Chief Architect of the Sakai Project. Charles teaches two popular MOOCs to students worldwide on the Coursera platform: Internet History, Technology, and Security and Programming for Everybody and is a long-time advocate of open educational resources to empower teachers. Charles is the editor of the Computing Conversations column in IEEE Computer magazine that features a monthly article and video interview of a computing pioneer. Charles is the author of the book, Sakai: Building an Open Source Community that describes the early days of the open source Sakai project. Charles has a B.S., M.S., and Ph.D. in Computer Science from Michigan State University.

Professor Belinda Tynan

LEARNING ANALYTICS = (PROBLEM+QUESTION) X (DATA+ANALYSIS) / (ETHICS+INTERVENTION)

Sponsored by Blue Canary

Professor Belinda Tynan joined the Open University (OU) in April 2013 as Pro Vice-Chancellor Learning and Teaching. She has a doctorate in education in the field of distance education from the University of Western Australia and is an active researcher, invited speaker and is passionate about online, open and distance education. Her research includes numerous refereed publications on the learner experience, faculty development, workload and innovation in the use of technologies. In her current role she has a specific focus in learning and teaching and is chair of several of the University main governance committees associated with policy and innovation for learning and teaching. A range of strategic University projects are lead from the PVC portfolio including Assessment, Tuition, Undergraduate and Postgraduate Frameworks, StageGate and Learning Analytics which is a University wide coordinated approach to supporting learners to reach their study goals. Professor Tynan is also on the Board of FutureLearn and the Steering Committee of the Open data Programme in partnership with the Open Data Institute.

Dr. Danielle S. McNamara

GETTING A BIG PICTURE FROM BIG DATA: USE YOUR WORDS!

Sponsored by Oracle

Danielle S. McNamara is a Professor in the Psychology Department at Arizona State University. She focuses on educational technologies and discovering new methods to improve students’ ability to understand challenging text, learn new information, and convey their thoughts and ideas in writing. Her work integrates various approaches and methodologies including the development of game-based, intelligent tutoring systems (e.g., iSTART, Writing Pal), the development of natural language processing tools (e.g., iSTART, Writing Pal, Coh-Metrix, the Writing Assessment Tool), basic research to better understand cognitive and motivational processes involved in comprehension and writing, and the use of learning analytics across multiple contexts. More information about her research and access to her publications are available at soletlab.com.
observations about approaches to clustering in this context. We also make some
assumptions about how patterns of engagement in these massive learning environments
are influenced by decisions about pedagogy. We previously apply in this case. Instead, we see seven distinct patterns of engagement: Samplers, Strong Starters, Returners, Mid-way Dropouts, Nearly There, Late Completers and Keen Completers. This suggests that patterns of engagement that emerged from MOOCs based on videos and assessments. However, not all platforms take this approach to learning design. Courses on the FutureLearn platform are underpinned by a social-constructivist pedagogy, which includes discussion as an important element. In this paper, we use analyse engagement patterns on four FutureLearn MOOCs and find that only two clusters identified on 451 submissions were graded only by peer grading and automated validation. The results of the experiment suggest that both CTAS and peer grading do not correlate: 0.76). Griffin and Doug Clow
Massive open online courses (MOOCs) are now being used across the world to provide millions of learners with access to education. Many learners complete these courses successfully, or to their own satisfaction, but the high numbers who do not finish remain a subject of concern for platform providers and educators. In 2013, a team from Stanford University analysed engagement patterns on three MOOCs run on the Coursera platform. They found four distinct patterns of engagement that emerged from MOOCs based on videos and assessments. However, not all platforms take this approach to learning design. Courses on the FutureLearn platform are underpinned by a social-constructivist pedagogy, which includes discussion as an important element. In this paper, we use analyse engagement patterns on four FutureLearn MOOCs and find that only two clusters identified previously apply in this case. Instead, we see seven distinct patterns of engagement: Samplers, Strong Starters, Returners, Mid-way Dropouts, Nearly There, Late Completers and Keen Completers. This suggests that patterns of engagement in these massive learning environments are influenced by decisions about pedagogy. We also make some observations about approaches to clustering in this context.

1A. MOOCs—Assessments, Connections and Demographics
Session Chair: Agathe Merceron

Session ID: 1A-01
Best Paper Nominee

ON THE VALIDITY OF PEER GRADING AND A CLOUD TEACHING ASSISTANT SYSTEM (FULL)
Tim Vogelsang and Lara Ruppertz
We introduce a new grading system, the Cloud Teaching Assistant System (CTAS), as an additional element to instructor grading, peer grading and automated validation in massive open online courses (MOOCs). The grading distributions of the different approaches are compared in an experiment consisting of 476 exam participants. 25 submissions were graded by all four methods. 451 submissions were graded only by peer grading and automated validation. The results of the experiment suggest that both CTAS and peer grading do not simulate instructor grading (Pearson’s correlations: 0.36, 0.39). If the CTAS and not the instructor is assumed to deliver accurate grading, peer grading is concluded to be a valid grading method (Pearson’s correlation: 0.76).

Session ID: 1A-02

EXAMINING ENGAGEMENT: ANALYSING LEARNER SUBPOPULATIONS IN MASSIVE OPEN ONLINE COURSES (MOOCs) (FULL)
Rebecca Ferguson and Doug Clow
Massive open online courses (MOOCs) are now being used across the world to provide millions of learners with access to education. Many learners complete these courses successfully, or to their own satisfaction, but the high numbers who do not finish remain a subject of concern for platform providers and educators. In 2013, a team from Stanford University analysed engagement patterns on three MOOCs run on the Coursera platform. They found four distinct patterns of engagement that emerged from MOOCs based on videos and assessments. However, not all platforms take this approach to learning design. Courses on the FutureLearn platform are underpinned by a social-constructivist pedagogy, which includes discussion as an important element. In this paper, we use analyse engagement patterns on four FutureLearn MOOCs and find that only two clusters identified previously apply in this case. Instead, we see seven distinct patterns of engagement: Samplers, Strong Starters, Returners, Mid-way Dropouts, Nearly There, Late Completers and Keen Completers. This suggests that patterns of engagement in these massive learning environments are influenced by decisions about pedagogy. We also make some observations about approaches to clustering in this context.

Session ID: 1A-03

SOCIOECONOMIC STATUS AND MOOC ENROLLMENT: ENRICHING DEMOGRAPHIC INFORMATION WITH EXTERNAL DATASETS (SHORT)
John D. Hansen and Justin Reich
To minimize barriers to entry, massive open online course (MOOC) providers collect minimal demographic information about users. In isolation, this data is insufficient to address important questions about socioeconomic status (SES) and MOOC enrollment and performance. We demonstrate the use of third-party datasets to enrich demographic portraits of MOOC students and answer fundamental questions about SES and MOOC enrollment. We derive demographic information from registrants’ geographic location by matching self-reported mailing addresses with data available from Esri at the census block group level and the American Community Survey at the zip code level. We then use these data to compare neighborhood income and levels of parental education for U.S. registrants in HarvardX courses and the U.S. population as a whole. Overall, HarvardX registrants tend to reside in more affluent neighborhoods. U.S. HarvardX registrants on average live in neighborhoods with median incomes approximately .45 standard deviations higher than the U.S. population. Parental education is also associated with a higher likelihood of MOOC enrollment. For instance, a seventeen year-old whose most educated parent has a bachelor’s degree is more than five times as likely to register as a seventeen year-old whose most educated parent has a high school diploma.

Session ID: 1A-04

HOW DO YOU CONNECT?
ANALYSIS OF SOCIAL CAPITAL ACCUMULATION IN CONNECTIVIST MOOCS (SHORT)
Srećko Joksimović, Nia Dowell, Oleksandra Skrypnyk, Vitomir Kovanović, Dragan Gašević, Shane Dawson and Arthur C. Graesser
Connections established between learners via interactions are seen as fundamental for connectivist pedagogy. Connections can also be viewed as learning outcomes, i.e. learners’ social capital accumulated through distributed learning environments. We applied linear mixed effects modeling to investigate whether the social capital accumulation interpreted through learners’ centrality to course interaction networks, is influenced by the language learners use to express and communicate in two connectivist MOOCs. Interactions were distributed across the three social media, namely Twitter, blog and Facebook. Results showed that learners in a cMOOC connect easier with the individuals who use a more informal, narrative style, but still maintain a deeper cohesive structure to their communication.
TOWARDS BETTER AFFECT DETECTORS: EFFECT OF MISSING SKILLS, CLASS FEATURES AND COMMON WRONG ANSWERS (SHORT)
Yutao Wang, Neil Heffernan and Christina Heffernan

The well-studied Baker et al., affect detectors modeled on boredom, frustration, confusion and engagement concentration with ASSISTments dataset were used to predict state tests scores, college enrollment, and even whether a student majored in a STEM field. In this paper, we present three attempts to improve upon current affect detectors. The first attempt analyzed the effect of missing skill tags in the dataset to the accuracy of the affect detectors. The results show a small improvement after correctly tagging the missing skill values. The second attempt added four features related to student classes for feature selection. The third attempt added two features that described information about student common wrong answers for feature selection. Result showed that two out of the four detectors were improved by adding the new features.

Session ID: 1B-01
HOW SHOULD WE QUANTIFY STUDENT ENGAGEMENT? (PRACTITIONER)
Perry Samson

Student engagement is widely thought to be a key predictor of student motivation and achievement. Engagement has been defined as “both the time and energy students invest in educationally purposeful activities.” Unfortunately this doesn’t identify what specific student actions to include in a quantification of engagement. This interactive presentation invites participants to consider how they would quantify student engagement using technology. The discussion will be informed from lessons learned at the University of Michigan where a rich database of student participation in class has been collected and related to student outcomes.

Session ID: 1B-03
EXPLORING NETWORKS OF PROBLEM-SOLVING INTERACTIONS (FULL)
Michael Eagle, Andrew Hicks, Barry Pedycoreld III and Tiffany Barnes

Intelligent tutoring systems and other computer-aided learning environments produce large amounts of transactional data on student problem-solving behavior. Previous work modeled the student-tutor interaction data as a complex network, and successfully generated automated next-step hints as well as visualizations for educators. In this work we discuss the types of tutoring environments that are best modeled by interaction networks and how the empirical observations of problem-solving result in common network features. We find that interaction networks exhibit the properties of scalefree networks such as vertex degree distributions that follow power law. We compare datasets from two versions of a propositional logic tutor, as well as two different representations of data from an educational programming video game. We find that statistics such as degree assortativity and the scale-free metric allow comparison of the network structures across domains, and provide insight into student problem solving behavior.

Session ID: 1C-02
THE LATUX WORKFLOW: DESIGNING AND DEPLOYING AWARENESS TOOLS IN TECHNOLOGY-ENABLED LEARNING SETTINGS (FULL)
Roberto Martinez-Maldonado, Abelardo Pardo, Negin Mirriaei, Kalina Yacef, Judy Kay and Andrew Clayphan

Designing, deploying and validating learning analytics tools for instructors or students is a challenge requiring techniques and methods from different disciplines, such as software engineering, human-computer interaction, educational design and psychology. Whilst each of these disciplines has consolidated design methodologies, there is a need for more specific methodological frameworks within the cross-disciplinary space defined by learning analytics. In particular there is no systematic workflow for producing learning analytics tools that are both technologically feasible and truly underpin the learning experience. In this paper, we present the LATUX workflow, a five-stage workflow to design, deploy and validate awareness tools in technology-enabled learning environments. LATUX is grounded on a well-established design process for creating, testing and re-designing user interfaces. We extend this process by integrating the pedagogical requirements to
WEDNESDAY, MARCH 18

generate visual analytics to inform instructors' pedagogical decisions or intervention strategies. The workflow is illustrated with a case study in which collaborative activities were deployed in a real classroom.

Session ID: 1C-03

LEARNING ANALYTICS BEYOND THE LMS: THE CONNECTED LEARNING ANALYTICS TOOLKIT (SHORT)
Kirsty Kitto, Sebastian Cross, Zak Waters and Mandy Lupton

We present a Connected Learning Analytics (CLA) toolkit, which enables data to be extracted from social media and imported into a Learning Record Store (LRS), as defined by the new xAPI standard. A number of implementation issues are discussed, and a mapping that will enable the consistent storage and then analysis of xAPI verb/object/activity statements across different social media and online environments is introduced. A set of example learning activities are proposed, each facilitated by the Learning Analytics beyond the LMS that the toolkit enables.

Session ID: 1C-04

DEVELOPING AN EVALUATION FRAMEWORK OF QUALITY INDICATORS FOR LEARNING Analytics (SHORT)
Maren Scheffel, Hendrik Drachsler and Marcus Specht

This paper presents results from the continuous process of developing an evaluation framework of quality indicators for learning analytics (LA). Building on a previous study, a group concept mapping approach that uses multidimensional scaling and hierarchical clustering, the study presented here applies the framework to a collection of LA tools in order to evaluate the framework. Using the quantitative and qualitative results of this study, the first version of the framework was revisited so as to allow work towards an improved version of the evaluation framework of quality indicators for LA.

1:00 p.m.–2:30 p.m. • Nelly Goletti Theatre

2A. Institutional Perspectives

Session Chair: Negin Mirriahi

Session ID: 2A-01

DEVELOPING POLICY FOR THE ETHICAL USE OF LEARNING ANALYTICS AT THE OPEN UNIVERSITY (PRACTITIONER)
Sharon Slade and Avinash Boroowa

Institutions are increasingly collecting, analysing and using student data with the aim of improving student satisfaction and success. The use of learning analytics within the Open University is relatively new and, as such, existing policies relating and referring to potential uses of student data received fresh scrutiny to ensure continued relevance and completeness. In response, the Open University has addressed a range of ethical issues relating to its approach to learning analytics via the implementation of a new policy. This presentation details the process undertaken and summarises the key principles on which the policy is built.

Session ID: 2A-02

RIDING THE TIGER: DEALING WITH COMPLEXITY IN THE IMPLEMENTATION OF INSTITUTIONAL STRATEGY FOR LEARNING ANALYTICS (PRACTITIONER)
Kevin Mayles

Implementing strategy for learning analytics across an institution is a complex task. The Open University UK is undertaking a major change programme to enhance its use of analytics to drive student success. The strategy has been developed around capability in three areas: data availability, creation of insight and the ability to apply analytics in practice to impact the student learning experience. Lessons learned from this case study in institutional strategy formulation and the management of structural, socio-political and emergent complexities during implementation will be highlighted and shared.

Session ID: 2A-03

Best paper Nominee

STUDENT PRIVACY SELF-MANAGEMENT: IMPLICATIONS FOR LEARNING Analytics (FULL)
Paul Prinsloo and Sharon Slade

Optimizing the harvesting and analysis of student data promises to clear the fog surrounding the key drivers of student success and retention, and provide potential for improved student success. At the same time, concerns are increasingly voiced around the extent to which individuals are routinely and progressively tracked as they engage online. The internet, the very thing that promised to open up possibilities and to break down communication barriers, now threatens to narrow it again through the panopticon of mass surveillance. Within higher education, our assumptions and understanding of issues surrounding student attitudes to privacy are influenced both by the apparent ease with which the public appear to share the detail of their lives and our paternalistic institutional cultures. As such, it can be easy to allow our enthusiasm for the possibilities offered by learning analytics to outweigh consideration of issues of privacy. This paper explores issues around consent and the seemingly simple choice to allow students to opt in or opt out of having their data tracked. We consider how 3 providers of massive open online courses (MOOCs) inform users of how their data is used, and discuss how higher education institutions can work toward an approach which engages and more fully informs students of the implications of learning analytics on their personal data.

1:00 p.m.–2:30 p.m. • Recital Hall

2B. Students At Risk

Session Chair: Cassandra Colvin

Session ID: 2B-01

OU ANALYSE: ANALYSING AT-RISK STUDENTS AT THE OPEN UNIVERSITY (PRACTITIONER)
Jakub Kuzilek, Martin Hlosta, Drahomira Herrmannova, Annika Wolff and Zdenek Zdrahal

The OU Analyse project aims at providing early prediction of ‘at-risk’ students based on their demographic data and their interaction with VLE. Four predictive models have been constructed from legacy data using machine learning methods. In Spring 2014 the approach was piloted on two introductory university courses with about 1500 and 3000 students, respectively. Since October 2014 the predictions have been extended to include 10+ courses of different level. For presenting predictions, the OU Analyse dashboard providing course overview and a view of individual students has been implemented. The presentation will include demonstration of the OU Analyse system.
WHO, WHEN, AND WHY: A MACHINE LEARNING APPROACH TO PRIORITIZING STUDENTS AT RISK OF NOT GRADUATING HIGH SCHOOL ON TIME (FULL)
Everaldgo Aguier, Himabindu Lakkaraju, Nasir Bhanpuri, David Miller, Ben Yuhas, Kecia Addison, Shihching Liu, Marilyn Powell and Rayid Ghani

Several hundred thousand students drop out of high school every year in the United States. Interventions can help those who are falling behind in their educational goals, but given limited resources, such programs must focus on the right students, at the right time, and with the right message. In this paper, we describe an incremental approach that can be used to select and prioritize students who may be at risk of not graduating high school on time, and to suggest what may be the predictors of particular students going off-track. These predictions can then be used to inform targeted interventions for these students, hopefully leading to better outcomes.

OPEN-SOURCE ACADEMIC EARLY ALERT AND RISK ASSESSMENT API (PRACTITIONER)
Sanjeev Jayapraaksh, Josh Baron, Gary Gilbert, Eitel Lauria, Erik Moody and James Regan

The presentation sums up the technical decisions and design strategies in building an open source academic early alert system. The system is an automation of the Open Academic Analytics Initiative (OAAI), a multi-year EDUCAUSE research grant focused on impacting the student retention rates. The system uses predictive analytics at its core to identify student population who are potentially at academic risk of not completing the course. The presentation details the data integration and data mining stages to formulate a potent prediction model using big data approaches. It also introduces an open source early alert Application Program Interface (API) developed using a generalized Learning analytics processor framework that can potentially support a wide range of learning analytics solutions.

LEARNING ANALYTICS: EUROPEAN PERSPECTIVES (RESEARCH PANEL)
Rebecca Ferguson, Adam Cooper, Hendrik Drachsler, Gábor Kismihók, Alejandra Martínez Monés, Kairit Tammets and Anne Boyer

Since the emergence of learning analytics in North America, researchers and practitioners have worked to develop an international community. The organization of events such as SoLAR Flares and LASI Locals, as well as the move of LAK in 2013 from North America to Europe, has supported this aim. There are now thriving learning analytics groups in North American, Europe and Australia, with smaller pockets of activity emerging on other continents. Nevertheless, much of the work carried out outside these forums, or published in languages other than English, is still inaccessible to most people in the community. This panel, organized by Europe's Learning Analytics Community Exchange (LACE) project, brings together researchers from five European countries to examine the field from European perspectives. In doing so, it will identify the benefits and challenges associated with sharing and developing practice across national boundaries.

OPENCOURSEWARE OBSERVATORY—DOES THE QUALITY OF OPENCOURSEWARE LIVE UP TO ITS PROMISE? (FULL)
Sahar Vahdati, Christoph Lange and Sören Auer

A vast amount of OpenCourseWare (OCW) is meanwhile being published online to make educational content accessible to larger audiences. The awareness of such courses among users and the popularity of systems providing such courses are increasing. However, from a subjective experience, OCW is frequently cursory, out-dated or non-reusable. In order to obtain a better understanding of the quality of OCW, we assess the quality in terms of fitness for use. Based on three OCW use case scenarios, we define a range of dimensions according to which the quality of courses can be measured. From the definition of each dimension a comprehensive list of quality metrics is derived. In order to obtain a representative overview of the quality of OCW, we performed a quality assessment on a set of 100 randomly selected courses obtained from 20 different OCW repositories. Based on this assessment we identify crucial areas in which OCW needs to improve in order to deliver up to its promises.

A TIME SERIES INTERACTION ANALYSIS METHOD FOR BUILDING PREDICTIVE MODELS OF LEARNERS USING LOG DATA (FULL)
Christopher Brooks, Craig Thompson and Stephanie Teasley

As courses become bigger, move online, and are deployed to the general public at low cost (e.g. through Massive Open Online Courses, MOOCs), new methods of predicting student achievement are needed to support the learning process. This paper presents a novel method for converting educational log data into features suitable for building predictive models of student success. Unlike cognitive modeling or content analysis approaches, these models are built from interactions between learners and resources, an approach that requires no input from instructional or domain experts and can be applied across courses or learning environments.
Session ID: 3A-03
PREDICTING SUCCESS: HOW LEARNERS’ PRIOR KNOWLEDGE, SKILLS AND ACTIVITIES PREDICT MOOC PERFORMANCE (SHORT)
Gregor Kennedy, Carleton Coffrin, Paula de Barba and Linda Corrin

While MOOCs have taken the world by storm, questions remain about their pedagogical value and high rates of attrition. In this paper we argue that MOOCs which have open entry and open curriculum structures, place pressure on learners to not only have the requisite knowledge and skills to complete the course, but also the skills to traverse the course in adaptive ways that lead to success. The empirical study presented in the paper investigated the degree to which students’ prior knowledge and skills, and their engagement with the MOOC as measured through learning analytics, predict end-of-MOOC performance. The findings indicate that prior knowledge is the most significant predictor of MOOC success followed by students’ ability to revise and revisit their previous work.

Session ID: 3A-04
LIKELIHOOD ANALYSIS OF STUDENT ENROLLMENT OUTCOMES USING LEARNING ENVIRONMENT VARIABLES: A CASE STUDY APPROACH (SHORT)
Scott Harrison, Renato Villano, Grace Lynch and George Chen

Tertiary institutions are increasing the emphasis on generating, collecting and analyzing student data as a means of targeting student support services. This study utilizes a data set from a regional Australian university to conduct logistic regression analyzing the student enrollment outcomes. The results indicate that demographic factors have a minor effect while institutional and learning environment variables play a more significant role in determining student enrollment outcomes. Using grade distribution compared to grade point average provides better estimates as to the effect particular grades have on enrollment outcomes. Moreover, the effect of an early alert system on enrollment outcomes shows that early identification has a significant relationship to a student’s choice to stay enrolled versus discontinuing, lapsing or being inactive in their enrollment. These results are vital in the targeting of student support services at the case study institution. The significant results indicate the importance of learning environment variables in understanding student enrollment outcomes at tertiary institutions. This analysis forms part of a much larger research project analyzing student retention at the institution.

2:45 p.m.–4:30 p.m. • Recital Hall
3B. MOOCs—Discussion Forums

Session Chair: Phillip Long

WHAT DO CMOOC PARTICIPANTS TALK ABOUT IN SOCIAL MEDIA? A TOPIC ANALYSIS OF DISCOURSE IN A CMOOC (FULL)
Srecko Joksimovic, Vitomir Kovancovic, Jelena Jovanovic, Amal Zouaq, Dragan Gasevic and Marek Hatala

Creating meaning from a wide variety of available information and being able to choose what to learn are highly relevant skills for learning in a connectivist setting. In this work, various approaches have been utilized to gain insights into learning processes occurring within a network of learners and understand the factors that shape learners’ interests and the topics to which learners devote a significant attention. This study combines different methods to develop a scalable analytic approach for a comprehensive analysis of learners’ discourse in a connectivist massive open online course (cMOOC). By linking techniques for semantic annotation and graph analysis with a qualitative analysis of learner-generated discourse, we examined how social media platforms (blogs, Twitter, and Facebook) and course recommendations influence content creation and topics discussed within a cMOOC. Our findings indicate that learners tend to focus on several prominent topics that emerge very quickly in the course. They maintain that focus, with some exceptions, throughout the course, regardless of readings suggested by the instructor. Moreover, the topics discussed across different social media differ, which can likely be attributed to the affordances of different media. Finally, our results indicate a relatively low level of cohesion in the topics discussed which might be an indicator of a diversity of the conceptual coverage discussed by the course participants.

Session ID: 3B-02
EXAMINING LEARNERS’ COGNITIVE PRESENCE THROUGH LINGUISTIC ANALYSIS IN MASSIVE OPEN ONLINE COURSES (MOOCs) (PRACTITIONER)
Yianna Vovides, Paige Arthur, Rob Pongsajapan and Mindy McWilliams

This study examines a corpus of 21,450 forum discussion posts from 1,937 learners in a Georgetown University edX MOOC on globalization for insight to the cognitive presence of learners and its implications for course performance. By analyzing the use of key terms linked to core course concepts over time as well as estimated level of language abstraction in the discussion forum, we examine the relationship between the results of this analysis, course grade and video content engagement. By combining these varied analytics, we aim to get a better sense of learners’ cognitive presence, course completion and performance.

Session ID: 3B-03
UNSUPERVISED MODELING FOR UNDERSTANDING MOOC DISCUSSION FORUMS: A LEARNING ANALYTICS APPROACH (SHORT)
Aysu Ezen-Can, Kristy Elizabeth Boyer, Shaun Kellogg and Sherry Booth

Massively Open Online Courses (MOOCs) have gained attention recently because of their great potential to reach learners. Substantial empirical study has focused on student persistence and their interactions with the course materials. However, most MOOCs include a rich textual dialogue forum, and these textual interactions are largely unexplored. Automatically understanding the nature of discussion forum posts holds great promise for providing adaptive support to individual students and to collaborative groups. This paper presents a study that applies unsupervised student understanding models originally developed for synchronous tutorial dialogue to MOOC forums. We use a clustering approach to group similar posts, compare the clusters with manual annotations by MOOC researchers, and further investigate clusters qualitatively. This paper constitutes a step toward applying unsupervised models to asynchronous communication, which can enable mass-scale automated discourse analysis and mining to better support students’ learning.

Session ID: 3B-04
CROWD-SOURCED LEARNING IN MOOCs: LEARNING ANALYTICS MEETS MEASUREMENT THEORY (SHORT)
Sandra Milligan

This paper illustrated the promise of the combination of measurement theory and learning analytics for understanding effective MOOC learning. It reports findings from a study of whether and how MOOC log file data can assist in understanding how MOOC participants use (often) messy, chaotic forums to support complex, unpredictable, contingent learning processes. It is argued that descriptions of posting, voting and
viewing behaviours do not in and of themselves provide insights about how learning is generated in MOOC forums. Rather, it is hypothesised that there is a skill involved in using forums to learn; that theory-informed descriptions of this skill illustrate how MOOC participants use forums differently as they progress from novice to expert; that the skill progression can be validated through the use of forum log file data; and that log file data can also be used to assess an individual MOOC participant’s position in relation to this progression — that is, to measure an individual’s skill in learning through forums and similar educational settings. These hypotheses were examined using data drawn from forums in a large MOOC run at the University of Melbourne in 2013.

2:45 p.m.–4:30 p.m. • SC 3101

3C. Student Performance

Session Chair: Eitel Lauria

Session ID: 3C-01

STUDENT-FACING ANALYTICS: DESIGNING AN INTERACTIVE DATA VISUALIZATION TO IMPROVE STUDY HABITS IN AN INTRODUCTORY MATH COURSE (PRACTITIONER)

Bart Pursel, Simon Hooper, Tsan-Kuang Lee, Andrew Fisher, Sherwyn Saul, Benjamin Brautigam, Christopher Millet and Nicholas Warcholak

As an institution, how can we leverage data to better support students throughout their coursework? This presentation outlines the design, development and creation of an interactive data visualization tool designed to help students in a large-enrollment calculus course prepare for assessments. Data were gathered from three different sources in order to understand the variables indicative of success in the course. These variables are then incorporated into the visualization, allowing students to explore their own data points in relation to exam score, and compare their data to students with a similar math aptitude.

Session ID: 3C-02

INVESTIGATING PERFORMANCE OF STUDENTS: A LONGITUDINAL STUDY (SHORT)

Raheela Asif, Agathe Merceron and Mahmood Pathan

In this paper, we investigate how performance of students evolves over the years in their studies. To determine typical progression patterns over the years, students are described by a 4 tuple (e.g. (x1, x2, x3, x4)), these being the clusters’ mean that s/he belongs to in each year of the degree. Two consecutive cohorts have been analyzed using X-means clustering. Interestingly the patterns found in both cohorts show that a substantial number of students stay in the same kind of group during their studies.

Session ID: 3C-03

USING TRANSACTION-LEVEL DATA TO DIAGNOSE KNOWLEDGE GAPS AND MISCONCEPTIONS (SHORT)

Randall Davies, Rob Nyland, John Chapman and Gove Allen

The role of assessment in learning is to evaluate student comprehension and ability. Assessment instruments often function at the task level. What is rarely considered is the process students go through to reach the final solution. This often allows knowledge component gaps and misconceptions to go undetected. This research identified higher levels of knowledge component gaps and misunderstandings when assessing transaction-level knowledge component data than task-level final solution data. Final solution data showed little evidence that students had any misunderstanding or knowledge gaps about the use of absolute references. However, when analyzing these data at the transaction level we found evidence that far more students struggled than the analysis of the final solutions suggested.

Session ID: 3C-04

COLLABORATIVE MULTI-REGRESSION MODELS FOR PREDICTING STUDENTS’ PERFORMANCE IN COURSE ACTIVITIES (SHORT)

Asmaa Elbadrawy, R. Scott Studham and George Karypis

Methods that accurately predict the grade of a student at a given activity or course can identify students that are at risk in failing a course and allow their educational institution to take corrective actions. Though a number of prediction models have been developed, they either estimate a single model for all students based on their past course performance and interactions with learning management systems (LMS), or estimate student-specific models that do not take into account LMS interactions; thus, failing to exploit fine-grain information related to a student’s engagement. In this work we present a class of collaborative multi-regression models that are personalized to each student and also take into account features related to student’s past performance, engagement and course characteristics. These models use all historical information to estimate a small number of regression models shared by all students along with student-specific combination weights. This allows for information sharing and also generating personalized predictions. Our experimental evaluation on a large set of students, courses, and activities shows that these models are capable of improving the performance prediction accuracy by over 20%. In addition, we show that by analyzing the estimated models and the student-specific combination functions we can gain insights on the effectiveness of the educational material that is made available at the courses of different departments.

4:45 p.m.–5:30 p.m. • Nelly Goletti Theatre

Firehose

Sponsored by IBM

Session Chair: Marcelo Worsley

The Firehose Session is a rapid fire introduction to all poster sessions. One author from each accepted poster will have one minute with one slide to provide a brief overview and reason for why you should stop by their presentation/poster. Voting station will be available to vote for “Best Poster”.

5:30 p.m.–7:30p.m. • SC 3102-3105

Poster Session & Welcome Reception

Sponsored by McGraw-Hill Education

Poster: 01

A LEARNING ANALYTICS APPROACH TO CHARACTERIZE AND ANALYZE INQUIRY-BASED PEDAGOGICAL PROCESSES

Carlos Monroy, Virginia Snodgrass Rangel, Elizabeth R. Bell and Reid Whitaker

This paper focuses on the use of learning analytics (LA) for investigating inquiry-based science instruction. To this end, we define several variables that quantify curriculum usage in a more nuanced fashion. We also leverage tools from process mining to examine inquiry-based pedagogical processes. These are initial steps toward measuring and modeling fidelity of implementation of a science curriculum. We use data from one school district’s use of an online science curriculum (N=1,021 teachers and nearly 330,000 page views), along with ProM, an open-source process mining software. We discuss the generalizability of our methods to other pedagogical frameworks as well as their limitations and our lessons learned.
**Poster: 02**

**PREDICTING POST-TRAINING READINESS TO WORK WITH COMPUTERS: THE PREDOMINANCE OF LOG-BASED VARIABLES**

Dalit Mor, Hagar Laks and Arnon Hershkovitz

In today's job market, computer skills are part of the prerequisites for many jobs. In this paper, we report on a study of readiness to work with computers (the dependent variable) among unemployed women (N=54) after participating in a unique training focused on computer skills and empowerment. Associations were explored between this variable and 17 variables from four categories: log-based, computer literacy and experience, job-seeking motivation and practice, and training satisfaction. Only two variables were associated with the dependent variable: Knowledge post-test duration and satisfaction with content. Building a prediction model of the dependent variable, another feature was highlighted: Total number of actions in the course website along the course. Our analyses highlight the predominance of the log-based variables over the variables from the other categories, and we thoroughly discuss this finding.

**Poster: 03**

**INVESTIGATING THE IMPACT OF A NOTIFICATION SYSTEM ON STUDENT BEHAVIORS IN A DISCOURSE-INTENSIVE HYBRID COURSE: A CASE STUDY**

Zhenhua Xu and Alexandra Makos

This study investigated the effects of students’ opting to use notification tools in a collaborative discourse-intensive online graduate course. Social constructivism and self-expectancy theory were applied to frame our understanding of the interactive relationship between the use of the notification tools, student’s online contribution behavior and student’s self-expectancy. Log-data from a 12-week hybrid (online and face-to-face) graduate course at a Canadian faculty of education was analyzed. Findings from the correlation, mediation and ANOVA analyses suggested that activation of the notification tool system positively affected students’ contribution behavior and that the influence of the use of notification tools on student contribution behavior was partially mediated by student’s self-expectancy.

**Poster: 04**

**MINIMUM INFORMATION ENTROPY BASED Q-MATRIX LEARNING IN DINA MODEL**

Shiwei Ye, Yuan Sun, Haobo Wang and Yi Sun

Cognitive diagnosis models (CDMs) are of growing interest in test development and measurement of learners’ performance. The DINA (deterministic input, noisy, and gate) model is one of the most widely used models in CDM. In this paper, we propose a new method and present an alternating recursive algorithm to learn Q-matrix and uncertainty variables, slip and guessing parameters, based on Boolean Matrix Factorization (BMF) and Minimized Information Entropy (MIE) respectively for the DINA model. Simulation results show that our algorithm for Q-matrix learning has fast convergence to the local optimal solutions for Q-matrix and students’ knowledge states A matrix. This is especially important and applicable when the method is extended to big data.

**Poster: 05**

**INTEGRATED REPRESENTATIONS AND SMALL DATA—TOWARDS CONTEXTUALIZED AND EMBEDDED ANALYTICS TOOLS FOR LEARNERS**

Andreas Harrer and Tilman Göhner

We present an approach to support learners by means of visualization and contextualization of learning analytics interventions in the learning process. We follow up on conceptual work of colleagues and derive further design principles oriented towards learners as recipients of LA results. These are shown with implementations in two distinct projects to fulfill learners information in collaborative learning processes.

**Poster: 06**

**FREQUENT SEQUENTIAL INTERACTIONS AS OPPORTUNITIES TO ENGAGE IN TEMPORAL REASONING WITH AN ONLINE GIS**

Raymond Kang, Josh Radinsky and Leilah Lyons

Temporal reasoning (i.e., reasoning about relationships across time) is complex and difficult, particularly when engaged through complex media such as online Geographic Information System (GIS) applications. Partnering with Social Explorer (SE), a Web-based GIS application that allows users to create interactive visualizations of large sociological datasets, we engaged in frequent sequential pattern mining of a database of users’ interactions with SE. The resulting frequent sequences provide initial descriptions of how SE affords opportunities to engage in temporal reasoning.

**Poster: 07**

**THE BRIDGE REPORT: BRINGING LEARNING ANALYTICS TO LOW-INCOME SCHOOLS**

Aaron Hawn

Widespread adoption of learning analytics for risk prediction faces different challenges at low-income secondary schools than at post-secondary institutions, where such methods have been more widely adopted. To leverage the benefits of learning analytics for under-resourced communities, educators must overcome the barriers to adoption faced by local schools: internet access, data integration, data interpretation, and local alignment. We present the case study of an enhanced reporting tool for parents and teachers, the Bridge Report, locally designed to meet the needs of a low-income secondary school in New York City. Parent and Teacher focus groups suggest that addressing local obstacles to learning analytics can create conditions for enthusiastic adoption by parents and teachers.

**Poster: 08**

**IMPROVING UNDERGRADUATE STUDENT ACHIEVEMENT IN LARGE BLENDED COURSES THROUGH DATA-DRIVEN INTERVENTIONS**

Bernie Dodge, John Whitmer and James P. Frazee

This pilot study applied Learning Analytics methods to identify students at-risk of not succeeding in two high enrollment courses with historically low pass rates at San Diego State University: PSY 101 and STAT 119. With input from instructors, targeted interventions were developed and sent to participating students (n=882) suggesting ways to improve their performance. An experimental design was used with half of the students randomly assigned to receive these interventions via email and the other half being analyzed for at-risk triggers but receiving no intervention. Pre-course surveys on student motivation and prior subject matter knowledge were conducted, and students were asked to maintain weekly logs of their activity online and offline connected to the courses. Regression analyses, incorporating feature selection methods to account for student demographic data, were analyzed. Findings from the correlation, mediation and ANOVA analyses suggested that activation of the notification tool system positively affected students’ contribution behavior and that the influence of the use of notification tools on student contribution behavior was partially mediated by student’s self-expectancy.
used to compare the impact of the interventions between the control and experimental groups. Results showed that the interventions were associated with a higher final grade in one course, but only for a particular demographic group.

**Poster: 09**  
**INCREASING THE ACCESSIBILITY OF LEARNING OBJECTS BY AUTOMATIC TAGGING**  
Katiija Niemann  
Data sets coming from the educational domain often suffer from sparsity. Hence, they might comprise potentially useful learning objects that are not findable by the users. In order to address this problem, we present a new way to automatically assign tags and classifications to learning objects offered by educational web portals that is solely based on the objects’ usage.

**Poster: 10**  
**MEASURING STUDENT SUCCESS USING PREDICTIVE ENGINE**  
Shady Shehata and Kimberly E. Arnold  
A basic challenge in delivering global education is improving student success. Institutions of education are increasingly focused on improving graduation and retention rates of their students. In this poster, we describe Student Success System (S3) that can measure student performance starting from the first weeks of the semester and the adoption process for S3 by University of Wisconsin System (UWS).

**Poster: 11**  
**A LEARNING SYSTEM UTILIZING LEARNERS’ ACTIVE TRACING BEHAVIORS**  
Kazushi Maruya, Junji Watanabe, Hiroyuki Takahashi and Shoji Hashiba  
A monitoring system that does not disturb learners’ motivation and attention is important, especially in online learning with massive numbers of participants. We propose a learning system, called the finger trail learning system (FTLS), that can monitor participants’ learning attitude by means of their finger movements. On the display of the FTLS, letters are presented with low contrast in the initial state, and the contrast of the letters changes to high when they are traced by learners. We implemented the FTLS as an iOS application and confirmed that the software can be utilized to monitor learners’ attitudes. In addition, we compared trails of finger movements between participants with high and low performance. The results show that the trail of finger movements recorded by the FTLS can be an index of learners’ attitudes.

**Poster: 12**  
**A CASE STUDY TO TRACK TEACHER GESTURES AND PERFORMANCE IN A VIRTUAL LEARNING ENVIRONMENT**  
Roghayeh Barmaki and Charles E. Hughes  
As part of normal interpersonal communication, people send and receive messages with their body, especially with their hands. Gestures play an important role in teacher-student classroom interactions. In the domain of education, many research projects have focused on the study of such gestures either in real classrooms or in tutorial settings with experienced teachers. Novice teachers especially need to understand the messages they are sending through nonverbal communication as this can have a major effect on their ability to manage behaviors and deliver content. Such learning should optimally occur before experiencing the real classroom. To assist in this process, we have developed a virtual classroom environment—TeachLivE—and used it for teacher practice, reflection and assessment. This paper investigates the way teachers use gestures in the virtual classroom settings of TeachLivE. Biology and algebra teachers were evaluated in our study. Analysis of video recordings from real and virtual environment seems to indicate that algebra teachers gesture significantly more often than biology teachers. These results have implications for providing useful feedback to participant teachers.

**Poster: 13**  
**QUALITATIVELY EXPLORING ELECTRONIC PORTFOLIOS: A TEXT MINING APPROACH TO MEASURING STUDENT EMOTION AS AN EARLY WARNING INDICATOR**  
Frederick Nwanganga, Everaldo Aguiar, G. Alex Ambrose, Victoria Goodrich and Nitesh V. Chawla  
The collection and analysis of student-level data is quickly becoming the norm across school campuses. More and more institutions are starting to use this resource as a window into better understanding the needs of their student population. In previous work, we described the use of electronic portfolio data as a proxy to measuring student engagement, and showed how it can be predictive of student retention. This paper highlights our ongoing efforts to explore and measure the valence of positive and negative emotions in student reflections and how they can serve as an early warning indicator of student disengagement.

**Poster: 14**  
**MEDIA MULTIPLEXITY IN CONNECTIVIST MOOCS**  
Rafa Absar, Anatoliy Gruzd, Caroline Haythornthwaite and Drew Paulin  
In this poster, we present work on exploring use of multiple social media platforms for learning in two connectivist MOOCs (or cMOOCs) to develop and evaluate methods for learning analytics to detect and study collaborative learning processes.

**Poster: 15**  
**USING LEARNING ANALYTICS TO STUDY COGNITIVE DISEQUILIBRIUM IN A COMPLEX LEARNING ENVIRONMENT**  
Marcelo Worsley and Paulo Blikstein  
Cognitive disequilibrium has received significant attention for its role in fostering student learning in intelligent tutoring systems and in complex learning environments. In this paper, we both add to and extend this discussion by analyzing the emergence of four affective states associated with disequilibrium: joy, surprise, neutrality and confusion; in a collaborative hands-on, engineering design task. Specifically, we conduct a comparison between two learning strategies to make salient how the strategies are associated with different affective states. This comparison is grounded in the construction of a probabilistic model of student affective state as defined by the frequency of each state, and the rate of transition between affective states. Through this comparison we confirm prior research that highlights the importance of confusion as a marker of knowledge construction, but put to question the notion that surprise is a significant mediator of cognitive disequilibrium. Overall, we show how modeling learner affect is useful for understanding and improving learning in complex, hands-on learning environments.
Poster: 16

ANALYSIS OF LEARNERS’ STUDY LOGS: MOUSE TRAJECTORIES TO IDENTIFY THE OCCURRENCE OF HESITATION IN SOLVING WORD-REORDERING
Mitsumasa Zushi, Yoshinori Miyazaki, Ken Norizuki
In this paper, we describe a Web application we have been developing in order to help both teachers and learners notice the crucial aspects of solving word-reordering problems (WRPs). Also, we discuss ways to analyze the recorded mouse trajectories, response time, and drag and drop (D&D) logs, because these records are potential indicators of the degree of learners’ understanding.

Poster: 17

HOW DO STUDENTS INTERPRET FEEDBACK DELIVERED VIA DASHBOARDS?
Linda Corrin and Paula de Barba
The large-scale production, collection, aggregation, and processing of information from various learning platforms and online environments have led to ethical and privacy concerns regarding potential harm to individuals and society. In the past, these types of concern have impacted on areas as diverse as computer science, legal studies and surveillance studies. Within a European consortium that brings together the EU project LACE, the SURF SIG Learning Analytics, the Apereo Foundation and the EATEL SIG dataTEL, we aim to understand the issues with greater clarity, and to find ways of overcoming the issues and research challenges related to ethical and privacy aspects of learning analytics practice. This interactive workshop aims to raise awareness of major ethics and privacy issues. It will also be used to develop practical solutions to advance the application of learning analytics technologies.

Poster: 18

LEARNING ANALYTICS IN OZ: WHAT’S HAPPENING NOW, WHAT’S PLANNED, AND WHERE COULD IT (AND SHOULD IT) GO?
Tim Rogers, Cassandra Colvin, Deborah West and Shane Dawson
This poster outlines the process and purpose of two related Australian Office for Learning and Teaching (OLT) commissioned grants to investigate the current usage and future potential of learning analytics in Australian Higher Education, with a view to developing resources to guide Australian universities in their adoption of learning analytics. The commissioned grants run from February 2014 to June 2015. Preliminary results will be available for LAK 15.

Poster: 19

TEXT MINING APPROACH TO AUTOMATE TEAMWORK ASSESSMENT IN GROUP CHATS
Antonette Shibani, Elizabeth Koh and Helen Hong
The increasing use of chat tools for learning and collaboration emphasizes the need for automating assessment. We propose a text mining approach to automate teamwork assessment in chat data. This supervised training approach can be extended to other domains for efficient assessment.

Doctoral Consortium Posters

Poster: 20

MULTIMODAL ASSESSMENT OF TEACHERS IN IMMERSIVE LEARNING ENVIRONMENT
Roghayeh Barmaki
People send and receive messages with their body, especially with their hands in interpersonal communications. Gestures play an important role in teacher-student interactions in classrooms. In the domain of education, many research projects have focused on the study of gestures either in real classrooms or in tutorial settings with practicing teachers. Novice teachers need to hone their required skills, including communication, management and content delivery skills, and prior to entering the classroom. To assist in this process, we have developed a virtual classroom environment—TeachLIVE™—and used it for teacher practice, reflection and assessment. Using semi-automated multimodal annotation procedure, the data collection will be completed. The study tries to perceive the correlations between objective measures of teaching preparedness with nonverbal communication signals. Providing the assessment from the collected data will help teachers in their professional development.

Poster: 21

LEARNING ANALYTICS FOR COMMUNITIES OF INQUIRY
Vitomir Kovanovic
This document outlines the doctoral research project that explores the use of Learning Analytics for improving the quality of learning within communities of inquiry in online and distance education. The primary goal of the proposed research is to improve pedagogical practices in online and distance education by the means of development of different learning analytics models. More precisely, the goal is to provide better understanding and broader adoption of Community of Inquiry (CoI) mode—a major educational theory of online and distance education—through the adoption of different techniques such as text classification, analysis of student social networks, and student clustering based on LMS trace data. The central component of the presented doctoral project is a novel text classification system for (semi)automated classification of student discussion messages in accordance with Community of Inquiry framework.

Poster: 22

EXPLORING AND MEASURING STUDENTS’ SENSE-MAKING PRACTICES AROUND REPRESENTATIONS OF THEIR ACADEMIC INFORMATION
Stephen Aguilar
My dissertation will consist of three studies, the first of which is complete. Study 1 is a quantitative study focused on the relationship between a representation of student achievement and students’ change in mastery orientation over the course of a summer bridge program. Multiple regression analysis indicates that even after controlling relevant variables merely viewing representations of academic information predicts decreases in bridge students’ mastery-goal orientation. Study 2 will use a grounded theory approach to study how students make sense of both “ubiquitous” representations of academic information, and unique representations of academic information typically found in learning analytics interventions. Building and overlapping with study 2, study 3 will use cognitive interviews, exploratory factor analysis, and confirmatory factor analysis to develop a psychometrically valid instrument that will measure students’ proclivity towards adaptive or maladaptive motivational attributions when exposed to various representations of their academic information. All three studies will draw participants from “at-risk” undergraduate student populations.
UNDERSTANDING SCALED ONLINE LEARNING
Oleksandra Skrypnyk
This paper presents an overview of a project that aims to examine shared aspects of learning in MOOCs to inform facilitation of social learning at scale. Observed differences between scaled online learning and conventional online provisions imposed by the course size, the lack of cohort boundaries, tangled diversity of motivations, and consequently, the lack of common terminology, does not allow to directly transfer knowledge base about online learning to learning at scale. Therefore, building upon the theories of adult education, online learning and CSCL, a social network perspective is applied to analyse the transactional structure of patterns underlying peer-to-peer and teacher-to-learner interactions, as well as the content of these interactions in relation to the levels of intersubjective meaning making. Project’s intended outcome is to derive indicators and describe social aspects of scaled online learning.

COMPUTATIONAL ANALYSIS OF PSYCHOSOCIAL AUTHOR CHARACTERISTICS FROM REFLECTIVE TEXT
Andrew Gibson
Reflective writing captures a person’s thoughts and feelings related to a specific context. The text is unstructured and yet often rich with emotion and personal content. When captured over time and analysed, reflections can provide insights on the author, the context of the author, and the underlying narratives. The complex nature of this type of text makes it time consuming for human analysis, creating problems when using reflections in areas like qualitative research or learning analytics. Computational analysis of reflective text would enable researchers to use larger quantities of reflective data, and educators would be able to analyse larger numbers of student reflections. My research is directed at this problem by investigating ways of computationally analysing reflective text in terms of the psychosocial characteristics of the author. In this paper I provide an overview of the research and the potential contributions in the form of a theoretical model and an algorithm.

EARLY ALERT PROGRAMS: A QUANTITATIVE ANALYSIS OF STUDENT RETENTION RATES
Scott Harrison
Early alert programs (EAPs) aim to provide targeted support to students in their learning environments. This study analyses to whom and when targeted support should be offered, as well as why institutions should invest in EAPs. Multinomial and regression models are used to estimate student enrolment outcomes and the timing of departure. They also form the basis for quantifying the economic benefits of EAPs to institutions. Preliminary results analyzing a non-temporal data set indicate that demographic factors play only a minor role in student enrolment outcomes, whilst institution and learning environment variables have a more significant effect. The multinomial model indicates that EAPs have a significant and positive effect in improving the likelihood of student staying enrolled. The complementary stationary multiple regression model estimates when student departure will occur. It shows a significant increase in the number of days students are enrolled when identified by the EAPs.

INTEGRATING LABOUR MARKET INFORMATION IN A STUDENT GOAL SETTING APPLICATION
Vladimer Kobayashi
Students attend formal education in order to acquire skills, competencies, and knowledge that are needed to succeed in the labour market. However, the mismatch between the graduates’ skills and the skills required for entry-level positions continues to beset both employers and applicants. One possible remedy for this problem is to increase the awareness of students about the situation in the labour market. The information from the labour market will guide students in their selection about which learning and employment goals to adapt. This paper discusses our ongoing research to develop software where students can set and follow both learning and employment goals. Moreover, the software will display two dashboards, one for labour market information and the other a learning dashboard that tracks students’ activities, performances, and progress towards goals. This approach combines knowledge from goal-setting theory, labour market analysis, and learning analytics.
Learning analytics (LA) has invested much effort in the investigation of learning technologies such as learning management systems (LMSs), has resulted in large amounts of data about student learning being readily accessible to educational researchers. One common use of this data is to measure time that students have spent on different learning tasks (i.e., time-on-task). Given that LMS systems typically only capture times when students executed various actions, time-on-task measures are estimated based on the recorded trace data. LMS trace data has been extensively used in many studies in the field of learning analytics, yet the problem of time-on-task estimation is rarely described in detail and the consequences that it entails are not fully examined. This paper presents the results of a study that examined the effects of different time-on-task estimation methods on the results of commonly adopted analytical models. The primary goal of this paper is to raise awareness of the issue of accuracy and appropriateness surrounding time-estimation within the broader learning analytics community, and to initiate a debate about the challenges of this process. Furthermore, the paper provides an overview of time-on-task estimation methods in educational and related research fields.

**Session ID**: 4A-03

**Short title**: Penetrating the Black Box of Time-On-Task Estimation (Full)

**Authors**: Vitomir Kovanic, Dragan Gasevic, Shane Dawson, Srecko Joksimovic, Ryan S. Baker and Marek Hatala

All forms of learning take time. There is a large body of research suggesting that the amount of time spent on learning can improve the quality of learning, as represented by academic performance. The widespread adoption of learning technologies such as learning management systems (LMSs), has resulted in large amounts of data about student learning being readily accessible to educational researchers. One common use of this data is to measure time that students have spent on different learning tasks (i.e., time-on-task). Given that LMS systems typically only capture times when students executed various actions, time-on-task measures are estimated based on the recorded trace data. LMS trace data has been extensively used in many studies in the field of learning analytics, yet the problem of time-on-task estimation is rarely described in detail and the consequences that it entails are not fully examined. This paper presents the results of a study that examined the effects of different time-on-task estimation methods on the results of commonly adopted analytical models. The primary goal of this paper is to raise awareness of the issue of accuracy and appropriateness surrounding time-estimation within the broader learning analytics community, and to initiate a debate about the challenges of this process. Furthermore, the paper provides an overview of time-on-task estimation methods in educational and related research fields.

**Session ID**: 4A-04

**Short title**: How does Bayesian Knowledge Tracing Model Emergence of Knowledge About a Mechanical System? (Short)

**Authors**: Hee-Sun Lee, Gey-Hong Gweon, Chad Dorsey, Robert Tinker, William Finzer, Dan Damelin, Nathan Kimball, Amy Pallant and Trudi Lord

An interactive learning task was designed in a game format to help high school students acquire knowledge about a simple mechanical system involving a car moving on a ramp. This ramp game consisted of five challenges that addressed individual knowledge components with increasing difficulty. In order to investigate patterns of knowledge emergence during the ramp game, we applied the Monte Carlo Bayesian Knowledge Tracing (BKT) algorithm to 447 game segments produced by 64 student groups in two physics teachers’ classrooms. Results indicate that, in the ramp game context, (1) the initial knowledge and guessing parameters were significantly highly correlated, (2) the slip parameter was interpretable monotonically, (3) low guessing parameter values were associated with knowledge emergence while high guessing parameter values were associated with knowledge maintenance, and (4) the transition parameter showed the speed of knowledge emergence. By applying the K-means clustering to ramp game segments represented in the three dimensional space defined by guessing, slip, and transition parameters, we identified seven clusters of knowledge emergence. We characterize these clusters and discuss implications for future research as well as for instructional game design.
Session ID: 4B-01

BECOMING STRATEGIC WRITERS: CAPTURING AND DISPLAYING STUDENTS’ STRATEGIES FOR WRITING TO ENHANCE FEEDBACK (PRACTITIONER)

Ge Vue and Tracey Hall

Providing frequent, explicit, and targeted feedback on student writing and strategies for writing is essential for improving writing performance. However, capturing and displaying students use of strategies when writing is difficult. We developed Writer’s Key, a web-based environment that supports middle school students to write and captures their strategies for writing. This allowed us to provide teachers and students with data that they can use to provide targeted feedback and monitor progress throughout the iterative cycle of drafting and revising. We will share examples of how teachers and students used Writer’s Key and visual display of log reports to individualize instruction, graduated scaffolding the development of writing skills, and motivate and sustain student engagement in writing.

Session ID: 4B-02

YOU’VE GOT STYLE: DETECTING WRITING FLEXIBILITY ACROSS TIME (FULL)

Erica L. Snow, Laura K. Allen, Matthew E. Jacovina, Cecile A. Perret and Danielle S. McNamara

Writing researchers have suggested that students who are perceived as strong writers (i.e., those who generate texts that are rated as high quality) demonstrate flexibility in their writing style. While anecdotally this has been a commonly held belief among researchers, scientists, and educators, there is little empirical research to support this claim. This study investigates this hypothesis by examining how students vary in their use of linguistic features across 16 prompt-based essays. Forty-five high school students wrote 16 essays across 8 sessions within an Automated Writing Evaluation (AWE) system. Natural language processing (NLP) techniques and Entropy analyses were used to calculate how rigid or flexible students were in their use of narrative linguistic features over time and how this trait related to individual differences in literacy ability and essay quality. Additional analyses indicated that NLP and Entropy reliably detected narrative flexibility (or rigidity) after session 2 and was related to students’ prior literacy skills. These exploratory methodologies are important for researchers and educators, as they indicate that writing flexibility is indeed a trait of strong writers and can be detected rather quickly using the combination of textual features and dynamic analyses.

Session ID: 4B-04

OPENESSYIST: A SUPPLY AND DEMAND LEARNING ANALYTICS TOOL FOR DRAFTING ACADEMIC ESSAYS (SHORT)

Denise Whitelock, Alison Twiner, John T.E. Richardson, Debora Field and Stephen Pulman

This paper focuses on the use of a natural language analytics engine to provide feedback to students when preparing an essay for summative assessment. OpenEssayist is a real-time learning analytics tool, which operates through the combination of a linguistic analysis engine that processes the text in the essay, and a web application that uses the output of the linguistic analysis engine to generate the feedback. We outline the system itself and present analysis of observed patterns of activity as a cohort of students engaged with the system for their module assignments. We report a significant positive correlation between the number of drafts submitted to the system and the grades awarded for the first assignment. We can also report that this cohort of students gained significantly higher overall grades than the students in the previous cohort, who had no access to OpenEssayist. As a system that is content free, OpenEssayist can be used to support students working in any domain that requires the writing of essays.

Session ID: 4C-01

AN ARCHITECTURE FOR A NEXT GENERATION LEARNING ANALYTICS PLATFORM (PRACTITIONER)

Vineet Sinha and Alfred Essa

Building an Analytics Platform requires balancing multiple competing concerns—from providing near real time data to making available sophisticated calculations. Further the rapidly growing digital learning ecosystem demands for an open analytics platform based on a standard such as IMS Caliper. We discuss some of the core requirements of such a platform, the architecture and main components, as well as some of our experiences.

Session ID: 4C-02

IMS CALIPER ANALYTICS™-STANDARDS-BASED LEARNING MEASUREMENT SERVICE FRAMEWORK TO ENABLE LEARNING DATA INTEROPERABILITY AND ANALYTICS (PRACTITIONER)

Chris Vento and Anthony Whyte

Educational institutions are evolving their use of learning technologies away from the one-application-fits-all “LMS” to a loosely-connected digital learning ecosystem of diverse services that leverage data analytics to drive pedagogical innovation and improve learning outcomes. Yet, an ecosystem rich in services and providers but lacking a common approach to measuring learning activity will find data collection, aggregation and analysis both time-consuming and costly. IMS Global’s Caliper Analytics™ framework addresses this challenge by providing an extensible learning activity information model with controlled vocabularies, together with baseline instrumentation for collecting learning events to transmit to a target record store for subsequent analysis.
DOP8: MERGING BOTH DATA AND ANALYSIS OPERATORS LIFE CYCLES FOR TECHNOLOGY ENHANCED LEARNING (SHORT)
Nadine Mandran, Michael Ortega, Vanda Luengo and Denis Bouhineau
This paper presents DOP8: a Data Mining Iterative Cycle that improves the classical data life cycle. While the latter only combines the data production and data analysis phases, DOP8 also integrates the analysis operators life cycle. In this cycle, data life cycle and operators life cycle processing meet in the data analysis step. This paper also presents a refication of DOP8 in a new computing platform: UnderTracks. The latter provides a flexibility on storing and sharing data, operators and analysis processes. Undertracks is compared with three types of platform 'Storage platform', 'Analysis platform' and 'Storage and Analysis platform'. Several real TEL analysis scenarios are present into the platform, (1) to test Undertracks flexibility on storing data and operators and (2) to test Undertracks flexibility on designing analysis processes.

A HANDWRITING RECOGNITION SYSTEM FOR THE CLASSROOM (SHORT)
Eric Gross, Safwan Wshah, Gary Skinner and Isaiah Simmons
The Xerox Ignite™ Educator Support System (henceforth referred to simply as Ignite™) is a data collection, analysis, and visualization workflow and software solution to assist K-12 educators. To illustrate, suppose a third-grade teacher wants to know how well her class has grasped a lesson on fractions. She would first scan her students’ homework and/or exams into the Ignite system via a range of multifunctional input devices. Xerox Ignite™ reads, interprets, and analyzes the students’ work in minutes. Then the teacher can select how to view the data by choosing from numerous reports. Examples are: an “at a glance” class summary that shows who needs extra help in what areas and who is ready to move on; a “context” report showing how each skill for each student is progressing over time; a grade-level performance report that helps third-grade teachers share best practices and cluster students into learning groups; and a student feedback report that tells each student what he/she needs to improve upon. Ignite™ intent is also to make it easier for districts to administer, score and evaluate content based on academic goals set for schools and students. The scanning and ‘mark lifting’ technology embedded into Ignite™ reduces the time needed to correct papers and frees time for the teacher to apply detailed insights to their day-to-day instruction tasks. Critical to this function is the automated reading of student marks, including handwriting, to enable the digitization of student performance at a detailed level. In this paper we present a system level description of the Ignite™ handwriting recognition module and describe the challenges and opportunities presented in an educational environment.

1:00 p.m.–2:30 p.m. • Nelly Goletti Theatre

5A. Learning Strategies and Tools
Session Chair: Shane Dawson

IDENTIFYING LEARNING STRATEGIES ASSOCIATED WITH ACTIVE USE OF VIDEO ANNOTATION SOFTWARE (SHORT)
Abelardo Pardo, Negin Mirriahi, Shane Dawson, Yu Zhao, An Zhao and Dragan Gasevic
The higher education sector has seen a shift in teaching approaches over the past decade with an increase in the use of video for delivering lecture content as part of a flipped classroom or blended learning model. Advances in video technologies have provided opportunities for students to now annotate videos as a strategy to support their achievement of the intended learning outcomes. However, there are few studies exploring the relationship between video annotations, student approaches to learning, and academic performance. This study seeks to narrow this gap by investigating the impact of students’ use of video annotation software coupled with their approaches to learning and academic performance in the context of a flipped learning environment. Preliminary findings reveal a significant positive relationship between annotating videos and exam results. However, negative effects of surface approaches to learning, cognitive strategy use and test anxiety on midterm grades were also noted. This indicates a need to better promote and scaffold higher order cognitive strategies and deeper learning with the use of video annotation software.

PLANNING FOR SUCCESS: HOW STUDENTS USE A GRADE PREDICTION TOOL TO WIN THEIR CLASSES (SHORT)
Caitlin Holman, Stephen J. Aguilar, Adam Levick, Jeff Stern, Benjamin Plummer and Barry Fishman
Gameful course designs require a significant shift in approach for both students and instructors. Transforming a standard course into a good game involves fundamentally altering how the course functions, most notably by giving students greater control over their work. We have developed an application, GradeCraft, to support this shift in pedagogy. A key feature of the application is the Grade Predictor, where students can explore coursework options and plan pathways to success. We observed students in two gameful courses with differing designs using the Grade Predictor in similar ways: they spent similar amounts of time per session, increased usage when assignments were due and before making significant course decisions, predicted different types of assignments at different rates, and made more predictions in preparation for the end of semester. This study describes how students plan their coursework using the GradeCraft Grade Predictor tool.

A PROCESS MINING APPROACH TO LINKING THE STUDY OF APTITUDE AND EVENT FACETS OF SELF-REGULATED LEARNING (SHORT)
Sanam Shirazi Beheshtia, Dragan Gašević and Marek Hatala
Research on self-regulated learning has taken main two paths: self-regulated learning as aptitudes and more recently, self-regulated learning as events. This paper proposes the use of the Fuzzy miner process mining technique to examine the relationship between students’ self-reported aptitudes (i.e., achievement goal orientation and approaches to learning) and strategies followed in self-regulated learning. A pilot study is conducted to probe the method and the preliminary results are reported.
to credit-bearing programs while maintaining comparable levels of academic achievement.

1:00 p.m.–2:30 p.m. • Recital Hall

5B. Text & Discourse Analysis
Session Chair: Alyssa Wise

Session ID: 5B-01

TOPIC FACET MODELING: SEMANTIC VISUAL ANALYTICS FOR ONLINE DISCUSSION FORUMS (SHORT)
I-Han Hsiao and Piyush Awasthi
In this paper, we propose a novel Topic Facet Model (TFM), a probabilistic topic model that assumes all words in single sentence are generated from one topic facet. The model is applied to automatically extract forum posts semantics for uncovering the content latent structures. We further prototype a visual analytics interface to present online discussion forum semantics. We hypothesize that the semantic modeling through analytics on open online discussion forums can help users examine the post content by viewing the summarized topic facets. Our preliminary results demonstrated that TFM can be a promising method to extract topic specificity from conversational and relatively short texts in online programming discussion forums.

Session ID: 5B-02

EFFECTS OF SEQUENCES OF SOCIALLY REGULATED LEARNING ON GROUP PERFORMANCE (SHORT)
Inge Molenaar and Ming Ming Chiu
Past research shows that regulative activities (metacognitive or relational) can aid learning and that sequences of cognitive, metacognitive and relational activities affect subsequent cognition. Extending this research, this study examines whether sequences of socially regulated learning differ across low, medium or high performing groups. Scaffolded by a computer avatar, 54 primary school students (working in 18 groups of 3) discussed writing a report about a foreign country for 51,338 turns. Statistical discourse analysis (SDA) of these sequences of talk showed that in high performing groups, high cognition was preceded more often by high cognition and less often by denials or low cognition. In medium performing groups, high cognition was preceded more often by high cognition or planning. As these results indicate that different sequences among students' cognitive, metacognitive and relational activities are linked to levels of performance, they can inform a micro-temporal theory of socially shared regulation.

Session ID: 5B-03

DEVELOPING A MULTIPLE-DOCUMENT-PROCESSING PERFORMANCE ASSESSMENT FOR EPISTEMIC LITERACY (SHORT)
Simon Knight and Karen Littleton
The LAKIS theme “shifts the focus from data to impact”, noting the potential for Learning Analytics based on existing technologies to have scalable impact on learning for people of all ages. For such demand and potential in scalability to be met the challenges of addressing higher-order thinking skills should be addressed. This paper discusses one such approach—the creation of an analytic and task model to probe epistemic cognition in complex literacy tasks. The research uses existing technologies in novel ways to build a conceptually grounded model of trace-indicators for epistemic-commitments in information seeking behaviors. We argue that such an evidence centered approach is fundamental to realizing the potential of analytics, which should maintain a strong association with learning theory.
technical systems of learning and the problems of implementation, and has the potential to be emancipatory. Some methodological implications for learning analytics practice are discussed.

2:45 p.m.–4:30 p.m. • Nelly Goletti Theatre

6A. Alternative Methods of Improving Learning
Session Chair: Xavier Ochoa

Best Paper Nominee
CLASSIFYING STUDENT DIALOGUE ACTS WITH MULTIMODAL LEARNING ANALYTICS (FULL)
Aysu Ezen-Can, Joseph F. Grafsaard, James C. Lester and Kirsty Elizabeth Boyer

Supporting learning with rich natural language dialogue has been the focus of increasing attention in recent years. Many adaptive learning environments model students’ natural language input, and there is growing recognition that these systems can be improved by leveraging multimodal cues to understand learners better. This paper investigates multi-modal features related to posture and gesture for the task of classifying students’ dialogue acts within tutorial dialogue. In order to accelerate the modeling process by eliminating the manual annotation bottleneck, a fully unsupervised machine learning approach is utilized for this task. The results indicate that these unsupervised models are significantly improved with the addition of automatically extracted posture and gesture information. Further, even in the absence of any linguistic features, a model that utilizes posture and gesture features alone performed significantly better than a majority class baseline. This work represents a step toward achieving better understanding of student utterances by incorporating multimodal features within adaptive learning environments. Additionally, the technique presented here is scalable to very large student datasets.

ANALYSING REFLECTIVE TEXT FOR LEARNING ANALYTICS: AN APPROACH USING ANOMALY RECONTEXTUALISATION (SHORT)
Andrew Gibson and Kirsty Kitto

Reflective writing is an important learning task to help foster reflective practice, but even when assessed it is rarely analysed or critically reviewed due to its subjective and affective nature. We propose a process for capturing subjective and affective analytics based on the identification and recontextualisation of anomalous features within reflective text. We evaluate 2 human supervised trials of the process, and so demonstrate the potential for an automated Anomaly Recontextualisation process for Learning Analytics.

TOWARDS DATA-DRIVEN MASTERY LEARNING (SHORT)
Behrooz Mostafavi, Michael Eagle and Tiffany Barnes

We have developed a novel data-driven mastery learning system to improve learning in complex procedural problem solving domains. This new system was integrated into an existing logic proof tool, and assigned as homework in a deductive logic course. Student performance and dropout were compared across three systems: The Deep Thought logic tutor, Deep Thought with integrated hints, and Deep Thought with our data-driven mastery learning system. Results show that the data-driven mastery learning system increases mastery of target tutor-actions, improves tutor scores, and lowers the rate of tutor dropout over Deep Thought, with or without provided hints.

2:45 p.m.–4:30 p.m. • Recital Hall

6B. Interventions and Remediation
Session Chair: Erik Moody

Session ID: 6B-01

USING DATA TO ENACT EARLY INTERVENTION STRATEGIES AND DEEPEN INSTRUCTION IN intro PSYCHOLOGY COURSES (PRACTITIONER)
Erik Moody, Larry Rudiger and John Rinderle

This panel will explore both the challenges and opportunities of teaching introductory psychology with predictive learning analytics and courseware based on cognitive science design principles. This session will feature case studies from the University of Vermont and Marist College utilizing newly re-designed courseware and learning analytics dashboard from Acrobatiq, a Carnegie Mellon University start-up and extension of the Open Learning Initiative. The presentations and discussion will immerse participants in an analysis of how precise learning data repositories can enable educators, regulatory bodies, university leadership, employers, and especially students to better understand what’s working and what isn’t within a classroom or program.

REDDUCING SELECTION BIAS IN QUASI-EXPERIMENTAL EDUCATIONAL STUDIES (SHORT)
Christopher Brooks, Jared Tritz, Omar Chavez and Stephanie Teasley

In this paper we examine the issue of selection bias in quasi-experimental (non-randomly controlled) educational studies. We provide background about common sources of selection bias and the issues involved in evaluating the outcomes of quasi-experimental studies. We describe two methods, matched sampling and propensity score matching, that can be used to overcome this bias. Using these methods, we describe their application through one case study that leverages large educational datasets drawn from higher education institutional data warehouses. The contribution of this work is the recommendation of a methodology and case study that educational researchers can use to understand, measure, and reduce selection bias in real-world educational interventions.

DISCOVERING CLUES TO AVOID MIDDLE SCHOOL FAILURE AT EARLY STAGES (SHORT)
Manuel Ángel Jiménez-Gómez, José María Luna, Cristóbal Romero and Sebastián Ventura

The use of data mining techniques in educational domains helps to find new knowledge about how students learn and how to improve the resources management. Using these techniques for predicting school failure is very useful in order to carry out actions to avoid drop out. With this purpose, we try to determine the earliest stage when the quality of the results allows for clarifying the possibility of school failure. We process real information from a Spanish high school by structuring the whole data in incremental datasets, which represent how students’ academic records grow. Our study reveals an early and robust detection of the risky cases of school failure at the end of the first out of four courses.
SESSION ID: 6B-04

AUTOMATED DETECTION OF PROACTIVE REMEDIATION BY TEACHERS IN REASONING MIND CLASSROOMS (SHORT)

Among the most important tasks of the teacher in a classroom using the Reasoning Mind blended learning system is proactive remediation: dynamically planned interventions conducted by the teacher with one or more students. While there are several examples of detectors of student behavior within an online learning environment, most have focused on behaviors occurring fully within the context of the system, and on student behaviors. In contrast, proactive remediation is a teacher-driven activity that occurs outside of the system, and its occurrence is not necessarily related to the student’s current task within the Reasoning Mind system. We present a sensor-free detector of proactive remediation, which is able to distinguish these activities from other behaviors involving idle time, such as on-task conversation related to immediate learning activities and off-task behavior.

2:45 p.m.–4:30 p.m. • SC 3101

6C. Analyses with LMS Data
Session Chair: Tim Rogers

SESSION ID: 6C-01

SCALABILITY AND FLEXIBILITY THROUGH OPEN RESEARCH (PRACTITIONER)
Thieme Hennis and Pieter de Vries

Delft University of Technology (TUD) is the largest university of engineering sciences in the Netherlands. TUD has been an active member of the Open Resources Movement since 2007 and in 2012 it has joined the EdX consortium. Openness in research and education is central in the TUD mission. It was the first edX partner to issue all its MOOCs under a creative commons license. In this presentation, we present our open research approach and toolbox. This toolbox facilitates sharing of data and the process of collaborative research in the context of MOOCs/online education. The presentation explains our rationale for open research, and our experiences with the toolbox, including research highlights.

SESSION ID: 6C-02

FORMATIVE AND SUMMATIVE ANALYSES OF DISCIPLINARY ENGAGEMENT AND LEARNING IN A BIG OPEN ONLINE COURSE (SHORT)
Daniel T. Hickey, Joshua D. Quick and Xinyi Shen

Situative theories of knowing and participatory approaches to learning and assessment were used to offer a big open online course on Educational Assessment using Google CourseBuilder in 2013. The course was started by 160 students and completed by 60, with relatively extensive instructor interaction with individual learners. This yielded much higher levels of engagement and learning than are typical of open or conventional online courses. The course was further refined and offered a second time in 2014, where it was started by 76 students and completed by 22, with a much lower level of support. Comparable levels of engagement and learning were obtained, suggesting that this participatory approach to learning and assessment can indeed be managed with more typical instructor support. Nonetheless, additional automation and streamlining is called for if the model is to eventually be used in massive online courses with thousands of students or as an autonomous self-paced open course.

SESSION ID: 6C-03

“SCALING UP” LEARNING DESIGN: IMPACT OF LEARNING DESIGN ACTIVITIES ON LMS BEHAVIOR AND PERFORMANCE (SHORT)
Bart Rienties, Lisette Toetenel and Annie Bryan

While substantial progress has been made in terms of predictive modeling in the Learning Analytics Knowledge (LAK) community, one element that is often ignored is the role of learning design. Learning design establishes the objectives and pedagogical plans which can be evaluated against the outcomes captured through learning analytics. However, no empirical study is available linking learning designs of a substantial number of courses with usage of Learning Management Systems (LMS) and learning performance. Using cluster and correlation analyses, in this study we compared how 87 modules were designed, and how this impacted (static and dynamic) LMS behavior and learning performance. Our findings indicate that learning design strongly influences how students were engaging online. Finally, learning design activities seem to have an impact on learning performance, in particular when modules rely on assimilative activities. Our findings indicate that learning analytics researchers need to be aware of the impact of learning design on LMS data over time, and subsequent academic performance.

SESSION ID: 6C-04

COMBINING OBSERVATIONAL AND EXPERIENTIAL DATA TO INFORM THE REDESIGN OF LEARNING ACTIVITIES (SHORT)
Abelardo Pardo, Robert A. Ellis and Rafael A. Calvo

A main goal for learning analytics is to inform the design of a learning experience to improve its quality. The increasing presence of solutions based on big data has even questioned the validity of current scientific methods. Is this going to happen in the area of learning analytics? In this paper we postulate that if changes are driven solely by a digital footprint, there is a risk of focusing only on factors that are directly connected to numeric methods. However, if the changes are complemented with an understanding about how students approach their learning, the quality of the evidence used in the redesign is significantly increased. This reasoning is illustrated with a case study in which an initial set of activities for a first year engineering course were shaped based only on the student’s digital footprint. These activities were significantly modified after collecting qualitative data about the students approach to learning. We conclude the paper arguing that the interpretation of the meaning of learning analytics is improved when combined with qualitative data which reveals how and why students engaged with the learning tasks in qualitatively different ways, which together provide a more informed basis for designing learning activities.
Technology Showcase

The Technology Showcase event has been designed to allow practitioners and researchers to demonstrate new and emerging learning analytics technologies that they are currently piloting or deploying in full production. *Light refreshments will be served.*

Tech: 01

READERBENCH: AN INTEGRATED TOOL SUPPORTING BOTH INDIVIDUAL AND COLLABORATIVE LEARNING

Mihai Dascalu, Lucia Larise Stavarche, Stefan Trausan-Matu, Phillipe Dessus, Bianco Maryse and Danielle McNamara

The core of our ReaderBench software framework exposes a unified vision for predicting and assessing comprehension in both individual and collaborative learning scenarios. ReaderBench aims to improve both the quality and the classification of the analyzed documents by using an expanded range of criteria such as: morphology, semantics, discourse analysis with emphasis on polyphony and dialogism, thus providing reliable support for both tutors and students across a range of educational settings. ReaderBench uses a unitary cohesion-based representation of discourse applied into three major directions, all tightly connected by the underlying model and the Natural Language Processing (NLP) computations: reading strategies, textual complexity, and collaboration evaluation in Computer Supported Collaborative Learning (CSCL) conversations.

Tech: 02

OPEN APEREO LEARNING ANALYTICS PLATFORM

Sandeep Jayaprakash, Joshua Baron, Alan Berg, Gary Gilbert and Russ Little

The technical showcase demonstrates ongoing effort in Apereo Foundation to build an open and scalable architecture for the Learning Analytics. The project touches on key architectural components widely applicable in a variety of Learning Analytics project settings. The technical demonstration focusses on showcasing a prototype system containing the various learning analytics components working as a cohesive system. The components include: Learning Data Collection methods and emerging open standards integration, Learning Data Storage Repository, App store model hosting targeted Data Analysis applications, Intervention systems and a multi-purpose widget based Open Dashboard systems delivering template insights from all the other four systems.

Tech: 03

THE PAR FRAMEWORK ECOSYSTEM, ENABLING A CYCLE OF STUDENT OUTCOME IMPROVEMENT

Russ Little

Learn how the non-profit PAR Framework leverages community driven common data definitions, predictive analytics and intervention inventories to drive a continuous improvement cycle of student outcomes improvement. Session will include an overview of the PAR Framework eco-system, review of the common data definitions, demonstrations of the measurement benchmarks, predictive watch lists, and the student intervention inventory toolset. Integration with 3rd party apps will be outlined and demonstrated. Real world experiences will be discussed, as well as a road map for the future of PAR Framework.

Tech: 04

LEARNING TO DRIVE INTERFACES

Ronald Monson

The use of engaging and tactile interfaces are becoming an increasingly popular way for both illustrating pedagogical concepts as well as often changing professional practice. Consequently their design, implementation and optimal usage are becoming increasingly important components in educational and product life-cycles. We showcase here the role that each component played in the creation of a new interface—Hypothesizing—a tool for illustrating the core logic underpinning Hypothesis Testing and how learning analytics can be applied to enhance the final component - its optimal usage.

Tech: 05

OU ANLYSE: ANALYSING AT-RISK STUDENTS AT THE OPEN UNIVERSITY

Jakub Kuzilek, Martin Hlosta, Drahomira Herrmannova, Annika Wolff and Zdenek Zdrahal

See abstract on page 9

Tech: 06

USING DATA TO ENACT EARLY INTERVENTION STRATEGIES AND DEEPEN INSTRUCTION IN INTRO PSYCHOLOGY COURSES

Howard Lurie

See abstract on page 22

Tech: 07

RIDING THE TIGER: DEALING WITH COMPLEXITY IN THE IMPLEMENTATION OF INSTITUTIONAL STRATEGY FOR LEARNING ANALYTICS

Kevin Mayles

See abstract on page 9

Tech: 08

FLEXPATH: A CASE STUDY IN BUILDING COMPETENCY-BASED, DIRECT ASSESSMENT OFFERINGS

Jeff Grann and Deborah Bushway

See abstract on page 19

Tech: 09

STUDENT-FACING ANALYTICS: DESIGNING AN INTERACTIVE DATA VISUALIZATION TO IMPROVE STUDY HABITS IN AN INTRODUCTORY MATH COURSE

Bart Pursel, Simon Hooper, Tsan-Kuang Lee, Andrew Fisher, Sherwyn Saul, Benjamin Brautigam, Christopher Millet and Nicholas Warcholak

See abstract on page 12

Tech: 10

D0P8_CYCLE: MERGING BOTH DATA AND ANALYSIS OPERATORS LIFE CYCLES AND TECHNOLOGY ENHANCED LEARNING

Nadine Mandran, Michael Ortega, Vanda Luengo and Denis Bouhineau

See abstract on page 19
Multilevel, Multimodal and Network Analysis

Session Chair: Xavier Ochoa

Session ID: 7A-01

LEVERAGING MULTIMODAL LEARNING ANALYTICS TO DIFFERENTIATE STUDENT LEARNING STRATEGIES (FULL)
Marcelo Worsley and Paulo Blikstein

Multimodal analysis has had demonstrated effectiveness in studying and modeling several human–human and human–computer interactions. In this paper, we explore the role of multimodal analysis in the service of studying complex learning environments. We compare uni-modal and multimodal; manual and semi-automated methods for examining how students learn in a hands-on, engineering design context. Specifically, we compare human annotations, speech, gesture and electro-dermal activation data from a study (N=20) where student participating in two different experimental conditions. The experimental conditions have already been shown to be associated with differences in learning gains and design quality. Hence, one objective of this paper is to identify the behavioral practices that differed between the two experimental conditions, as this may help us better understand how the learning interventions work. An additional objective is to provide examples of how to conduct learning analytics research in complex environments and compare how the same algorithm, when used with different forms of data can provide complementary results.

Session ID: 7A-02

FROM CONTINGENCIES TO NETWORK-LEVEL PHENOMENA: MULTILEVEL ANALYSIS OF ACTIVITY AND ACTORS IN HETEROGENEOUS NETWORKED LEARNING ENVIRONMENTS (FULL)
Dan Suthers

Learning in social settings is a complex phenomenon that involves multiple processes at individual and collective levels of agency. Thus, a richer understanding of learning in socio-technical networks will be furthered by analytic methods that can move between and coordinate analyses of individual, small group and network level phenomena. This paper outlines Traces, an analytic framework designed to address these and other needs, and gives examples of the framework’s practical utility using data from the Tapped In educator professional network. The Traces framework identifies observable contingencies between events and uses these to build more abstract models of interaction and ties represented as graphs. Applications are illustrated to identification of sessions and key participants in the sessions, relations between sessions as mediated by participants, and longer-term participant roles.

Session ID: 7A-03

UBIQUITOUS LEARNING ANALYTICS IN THE CONTEXT OF REAL-WORLD LANGUAGE LEARNING (SHORT)
Kousuke Mouri, Hiroaki Ogata and Noriko Uosaki

This paper describes a method of the visualization and analysis for mining useful learning logs from numerous learning experiences that learners have accumulated in the real world as the ubiquitous learning logs. Ubiquitous Learning Log (ULL) is defined as a digital record of what learners have learned in the daily life using ubiquitous technologies. It allows learners to log their learning experiences with photos, audios, videos, location, RFID tag and sensor data, and to share and reuse ULL with others. By constructing real-world corpora which comprise of accumulated ULLs with information such as what, when, where, and how learners have learned in the real world and by analyzing them, we can support learners to learn more effectively. The proposed system will predict their future learning opportunities including their learning patterns and trends by analyzing their past ULLs. The prediction is made possible both by network analysis based on ULL information such as learners, knowledge, place and time and by learners’ self-analysis using time-map. By predicting what they tend to learn next in their learning paths, it provides them with more learning opportunities. Accumulated data are so big and the relationships among the data are so complicated that is difficult to grasp how closely the ULLs are related each other. Therefore, this paper proposes a system to help learners to grasp relationships among learners, knowledge, place and time, using network graphs and network analysis.

Session ID: 7A-04

AN EXPLORATORY STUDY USING SOCIAL NETWORK ANALYSIS TO MODEL EYE MOVEMENTS IN MATHEMATICS PROBLEM SOLVING (SHORT)
Mengxiao Zhu and Gary Feng

Eye-tracking is a useful tool to understand students’ cognitive process during problem solving. This paper offers a unique perspective by applying techniques from social network analysis to eye movement patterns in mathematics problem solving. We construct and visualize transition networks using eye-tracking data collected from 37 8th grade students while solving linear function problems. By applying network analysis on the constructed transition networks, we find general transition patterns between areas of interests (AOIs) for all students, and we also compare patterns for high- and low-performing students. Our results show that even though students share general transition patterns during problem solving, high-performing students made more strategic transitions among AOI triples than low-performing students.
been developing a curriculum analysis method that can help to find out differences among curricula. On the basis of “Computing Science Curricula CS2013”, a report released by the ACM and IEEE Computer Society, we applied our method to analyzing 10 computer science (CS) related curricula offered by CS departments of universities in the United States. Using the method enables us to compare courses across universities. Through an analysis of course syllabi distribution, we found that CS2013 uniformly covered a wide area of computer science. Some universities emphasized human factors, while others attached greater importance to theoretical ones. We also found that some CS departments offered not only a CS curriculum but also an electrical engineering one, and those departments showed a tendency to have more “Architecture and Organization (AR)” related curricula. Furthermore, we found that even though “Information Assurance and Security (IAS)” has not yet become a very popular field, some universities are already offering IAS related courses.

Session ID: 7B-02

“TWITTER ARCHEOLOGY” OF LEARNING ANALYTICS AND KNOWLEDGE CONFERENCES (FULL)
Bodong Chen, Xin Chen and Wanli Xing

The goal of the present study was to uncover new insights about the learning analytics community by analyzing Twitter archives from the past four Learning Analytics and Knowledge (LAK) conferences. Through descriptive analysis, interaction network analysis, hashtag analysis, and topic modeling, we found: extended coverage of the community over the years; increasing interactions among its members regardless of peripheral and in-persistent participation; increasingly dense, connected and balanced social networks; and more and more diverse research topics. Detailed inspection of semantic topics uncovered insights complementary to the analysis of LAK publications in previous research.

Session ID: 7B-03

DISCOURSE COHESION: A SIGNATURE OF COLLABORATION (SHORT)
Mihai Dascalu, Stefan Trausan-Matu, Philippe Dessus and Danielle McNamara

As Computer Supported Collaborative Learning (CSCL) becomes increasingly adopted as an alternative to classic educational scenarios, we face an increasing need for automatic tools designed to support tutors in the time consuming process of analyzing conversations and interactions among students. Therefore, building upon a cohesion-based model of the discourse, we have validated ReaderBench, a system capable of evaluating collaboration based on a social knowledge-building perspective. Through the intertwining of different participants’ points of view, collaboration emerges and this process is reflected in the identified cohesive links between different speakers. Overall, the current experiments indicate that textual cohesion successfully detects collaboration between participants as ideas are shared and exchanged within an ongoing conversation. At a macroscopic level, the inter-relationships are circular in the sense that:

a) discourse coherence can be achieved collaboratively, and

b) cohesion, which provides the microstructure level connectedness, is a building block for achieving a truly collaborative discourse.

Session ID: 7B-04

CORRELATIONS BETWEEN AUTOMATED RHETORICAL ANALYSIS AND TUTORS’ GRADES ON STUDENT ESSAYS (SHORT)
Duygu Simsek, Agnès Sandor, Simon Buckingham Shum, Rebecca Ferguson, Anna De Liddo and Denise Whitelock

When assessing student essays, educators look for the students’ ability to present and pursue well-reasoned and strong arguments. Such scholarly argumentation is often articulated by rhetorical metadiscourse. Educators will be necessarily examining metadiscourse in students’ writing as signals of the intellectual moves that make their reasoning visible. Therefore students and educators could benefit from available powerful automated textual analysis that is able to detect rhetorical metadiscourse. However, there is a need to validate such technologies in higher education contexts, since they were originally developed in non-educational applications. This paper describes an evaluation study of a particular language analysis tool, the Xerox Incremental Parser (XIP), on undergraduate social science student essays, using the mark awarded as a measure of the quality of the writing. As part of this exploration, the study presented in this paper seeks to assess the quality of the XIP through correlational studies and multiple regression analysis.
Session ID: 7C-03

IMPROVING STUDENTS’ LONG-TERM RETENTION PERFORMANCE: A STUDY ON PERSONALIZED RETENTION SCHEDULES (SHORT)
Xiaolu Xiong, Yan Wang and Joseph Barbosa Beck

Traditional practices of spacing and expanding retrieval practices have typically fixed their spacing intervals to one or few predefined schedules [5, 7]. Few have explored the advantages of using personalized expanding intervals and scheduling systems to adapt to the knowledge levels and learning patterns of individual students. In this work, we are concerned with estimating the effects of personalized expanding intervals on improving students’ long-term mastery level of skills. We developed a Personalized Adaptive Scheduling System (PASS) in ASSISTments’ retention and relearning workflow. After implementing the PASS, we conducted a study to investigate the impact of personalized scheduling on long-term retention by comparing results from 97 classes in the summer of 2013 and 2014. We observed that students in PASS outperformed students in traditional scheduling systems on long-term retention performance (p = 0.0002), and that in particular, students with medium level of knowledge demonstrated reliable improvement (p = 0.0209) with an effect size of 0.27. In addition, the data we gathered from this study also helped to expose a few issues we have with the new system. These results suggest personalized knowledge retrieval schedules are more effective than fixed schedules and we should continue our future work on examining approaches to optimize PASS.

12:15 p.m.–12:45 p.m. • Nelly Goletti Theatre
State of the Field Expert Panel
Moderator: Josh Baron; Panelists: Simon Buckingham Shum, Australia; Caroline Haythornthwaite, Canada; Dragan Gasevic, UK; Stephanie Teasley, USA

This panel contextualises the findings from LAK15 in 4 different perspectives from Australia, Canada, United Kingdom and United States of America providing a state of the field analysis of learning analytics. They will discuss trends and future directions for the field and ways to get involved.

12:45 p.m. –1:00 p.m. • Nelly Goletti Theatre
Closing Remarks
MARIST COLLEGE is a comprehensive institution grounded in the liberal arts, with its main campus located in the historic Hudson River Valley of New York and a branch campus in the rich cultural center of Florence, Italy. The College is listed as one of the nation’s top institutions for undergraduate education by The Princeton Review, which has featured Marist in its “best colleges” guide for 12 consecutive years. Only about 15 percent of America’s 2,500 primarily undergraduate colleges are featured in the annual guide. For the 8th consecutive year, Kiplinger’s Personal Finance has included Marist on its list of America’s best values in private colleges. U.S. News & World Report ranks two Marist online degree programs—the MBA and the BA/BS in liberal studies—in its list of “Best Online Education Programs.” Marist is distinguished by its longtime partnership with IBM, which supports Marist’s world-class technology platform, and the Marist Institute for Public Opinion, one of the premier independent survey research centers in the nation.

A bequest by one of the 20th century’s leading businessmen and industrialists established the RAYMOND A. RICH INSTITUTE FOR LEADERSHIP DEVELOPMENT to educate individuals in the art of leadership for careers in business, government, and the nonprofit sector. The institute is based at the 60-acre Hudson Valley historic riverfront estate bequeathed by Mr. Rich to Marist College in 2009. For information visit www.marist.edu/richinstitute.

Marist offers a variety of programs at its branch campus in FLORENCE, ITALY, ranging from a structured freshman-year experience and semester- and year-long placements in the junior or senior year to four-year bachelor’s degree programs in seven fields of study and a Master of Arts in museum studies. For information visit www.marist.edu/florence.