

# Proceedings of the 8th International Conference on Educational Data Mining



26-29 June 2015  
Madrid - Spain

O.C. Santos, J.G. Boticario, C. Romero, M. Pechenizkiy,  
A. Merceron, P. Mitros, J.M. Luna, C. Mihaescu, P. Moreno, A. Hershkovitz,  
S. Ventura, M. Desmarais (Eds)



PEARSON



UNED

Olga Cristina Santos, Jesus Gonzalez Boticario, Cristobal Romero, Mykola Pechenizkiy,  
Agathe Merceron, Piotr Mitros, José María Luna, Cristian Mihaescu, Pablo Moreno,  
Arnon HersHKovitz, Sebastian Ventura, Michel Desmarais

*International Conference on Educational Data Mining (EDM) 2015*

Proceedings of the 8<sup>th</sup> International Conference on Educational Data Mining

Olga Cristina Santos, Jesus Gonzalez Boticario, Cristobal Romero, Mykola Pechenizkiy,  
Agathe Merceron, Piotr Mitros, José María Luna, Cristian Mihaescu, Pablo Moreno,  
Arnon HersHKovitz, Sebastian Ventura, Michel Desmarais (eds.)

Madrid, June 26-29, 2015

ISBN: 978-84-606-9425-0

## Preface

The 8th International Conference on Educational Data Mining (EDM 2015) is held under auspices of the International Educational Data Mining Society at UNED, the National University for Distance Education in Spain. The conference held in Madrid, Spain, July 26-29, 2015, follows the seven previous editions (London 2014, Memphis 2013, Chania 2012, Eindhoven 2011, Pittsburgh 2010, Cordoba 2009 and Montreal 2008).

The EDM conference is a leading international forum for high-quality research that mines large data sets in order to answer educational research questions that shed light on the learning processes. These data sets may come from the traces that students leave when they interact with learning management systems, interactive learning environments, intelligent tutoring systems, educational games or when they participate in a data-rich learning context. The types of data therefore range from raw log files to eye-tracking devices and other sensor data.

This year's conference features three invited talks by Luis von Ahn and Matt Streeter (Duolingo), George Siemens, Ryan Baker and Dragan Gasevic (Athabasca University, Columbia University and University of Edinburgh respectively) and Pekka Räsänen (Niilo Mäki Institute). To facilitate further discussion of the increasingly important research issues, three interactive panels have been organized; on grand challenges in EDM, ethics and privacy considerations in EDM, and practical applications of EDM at scale. This year, together with the Journal of Educational Data Mining (JEDM), we started the JEDM Track with the intention to accommodate researchers who want to contribute a more substantial contribution than space allows in the conference proceedings, and yet to present their work to a live conference audience. The papers submitted to the track followed the regular JEDM peer review process; 4 paper have been accepted to the track and will be presented at the conference. The abstracts of the invited talks, panels and accepted JEDM Track papers can be found in these proceedings.

The main conference calls for papers invited contributions to the Research Track and Industry Track. We received 121 full and 59 short paper submissions, each of which was reviewed by three experts in the field, resulting in 43 full (41 research and 2 industry), and 50 short (46 research and 4 industry) papers accepted for presentation at the conference (some of the full paper submissions have been accepted as short paper). From a separate call for posters we also accepted 39 poster and 3 demo papers. All accepted submissions appear in these proceedings.

The EDM conference traditionally provides opportunities for young researchers, and particularly for PhD students, to present their research ideas and receive feedback from the peers and more senior researchers. This year, the organized Doctoral Consortium will feature 12 presentations.

Besides the main conference program, the participants are program conference also includes 3 workshops (Graph-based Educational Data Mining, SMLIR: Workshop on Tools and Technologies in Statistics, Machine Learning and Information Retrieval for Educational Data Mining, and International Workshop on Affect, Meta-Affect, Data and Learning) and 2 tutorials (Using Natural Language Processing Tools in Educational Data Mining, and Student Modeling Applications, Recent Developments & Toolkits).

We would like to thank UNED for the sponsorship and hosting of EDM'2015. We would like to thank the commercial sponsors (MARi, Pearson and duoLingo), student support sponsors (NSF and Professor Ram Kumar Memorial Foundation) and academic support (UNED). We also want to acknowledge the amazing work of the program committee members and additional reviewers, who with their enthusiastic contributions gave us invaluable support in putting this conference together. Our special thanks to ConferenceNavigator – a social system for conference attendees that provided services for personal scheduling, social linking and personalized recommendations of papers. Last but not least we would like to thank the local organizing team.

June 2015

Cristobal Romero and Mykola Pechenizkiy – Program Chairs

Jesus G. Boticario and Olga C. Santos – Conference Chairs



## Organization

### CONFERENCE CHAIRS

Jesus G. Boticario UNED, Spain  
Olga C. Santos UNED, Spain

### PROGRAM CHAIRS

Cristóbal Romero University of Cordoba, Spain  
Mykola Pechenizkiy Eindhoven University of Technology, the Netherlands

### INDUSTRY TRACK CHAIRS

Agathe Merceron Beuth Hochschule für Technik Berlin, Germany  
Piotr Mitros edX, USA

### JEDM JOURNAL TRACK CHAIR

Michel Desmarais Ecole polytechnique Montreal, Canada

### POSTER CHAIRS

José María Luna University of Cordoba, Spain  
Cristian Mihaescu University of Craiova, Romania

### DEMO CHAIRS

Pablo Moreno e-Learning Group  
Arnon HersHKovitz Tel Aviv University, Israel

### DOCTORAL CONSORTIUM CHAIRS

Sebastián Ventura University of Cordoba, Spain

### WORKSHOP AND TUTORIAL CHAIRS

Katrien Verbert KU Leuven, Belgium  
Kaska Poryaska-Pomsta LKL

### PUBLICITY CHAIRS

Sergio Gutiérrez University of London Birkbeck, UK  
Taylor Martin NSF, USA

### SPONSOR CHAIR

Dragan Gašević University of Edinburgh, UK

### AWARDS CHAIRS

Manolis Mavrikis University of London Birkbeck, UK  
Kalina Yacef University of Sydney, Australia

### ADMINISTRATIVE ORGANIZER

Mar Saneiro UNED, Spain

### LOCAL ARRANGEMENT CHAIR

Sergio Salmeron-Majadas UNED, Spain

### FINANCIAL CHAIR

Pilar Muñoz UNED, Spain

### WEB CHAIR

Emmanuelle Gutiérrez y Restrepo UNED, Spain

## STEERING COMMITTEE/IEDMS BOARD OF DIRECTORS

|                    |   |
|--------------------|---|
| Ryan Baker         | Teachers College, Columbia University           |
| Tiffany Barnes     | University of North Carolina at Charlotte, USA  |
| Michel Desmarais   | Ecole Polytechnique de Montreal, Canada         |
| Sidney D'Mello     | University of Notre Dame                        |
| Neil Heffernan     | Worcester Polytechnic Institute, USA            |
| Agathe Merceron    | Beuth University of Applied Sciences, Germany   |
| Mykola Pechenizkiy | Eindhoven University of Technology, Netherlands |
| John Stamper       | Carnegie Mellon University                      |
| Kalina Yacef       | The University of Sydney, Australia             |

## PROGRAM COMMITTEE RESEARCH TRACK

|                        |   |
|------------------------|---|
| Lalitha Agnihotri      | McGraw Hill Education                             |
| Esma Aimeur            | University of Montreal                            |
| Omar Alzoubi           | School of Computer Science                        |
| Ivon Arroyo            | Worcester Polytechnic Institute                   |
| Mirjam Augstein        | Upper Austria University of Applied Sciences      |
| Ryan Baker             | Teachers College, Columbia University             |
| Tiffany Barnes         | North Carolina State University                   |
| Damian Bebell          | Boston College                                    |
| Behzad Beheshti        | Ecole Polytechnique de Montreal                   |
| Gautam Biswas          | Vanderbilt University                             |
| Mary Jean Blink        | TutorGen, Inc.                                    |
| Jesus Boticario        | UNED  |
| François Bouchet       | LIP6 - Université Pierre et Marie Curie           |
| Kristy Elizabeth Boyer | North Carolina State University                   |
| Javier Bravo Agapito   | Universidad Autonoma de Madrid                    |
| Renza Campagni         | Università degli Studi di Firenze                 |
| Alberto Cano           | University of Cordoba                             |
| John Champaign         | University of Waterloo                            |
| Min Chi                | North Carolina State University                   |
| Germán Cobo            | Universitat Oberta de Catalunya (UOC)             |
| Mihaela Cocea          | School of Computing, University of Portsmouth     |
| Michel Desmarais       | Ecole Polytechnique de Montreal                   |
| Sidney D'Mello         | University of Notre Dame                          |
| Hendrik Drachler       | Open University of the Netherlands                |
| Stephen Fancsali       | Carnegie Learning, Inc.                           |
| Mingyu Feng            | SRI International                                 |
| Davide Fossati         | Carnegie Mellon University in Qatar               |
| Carlos Garcia          | University of Córdoba                             |
| Dragan Gasevic         | University of Edinburgh                           |
| Eva Gibaja             | University of Cordoba                             |
| Daniela Godoy          | ISISTAN Research Institute                        |
| Ilya Goldin            | Pearson   |
| Yue Gong               | Worcester Polytechnic Institute                   |
| José González-Brenes   | Pearson   |
| Joseph Grafsgaard      | North Carolina State University, Computer Science |
| Wilhelmiina Hamalainen | University of Eastern Finland                     |
| Neil Heffernan         | Worcester Polytechnic Institute                   |
| Arnon Hershkovitz      | Tel Aviv University                               |
| Roland Hubscher        | Bentley University                                |
| Sébastien Iksal        | University of Maine                               |
| Patrick Jermann        | EPFL - Center for Digital Education               |
| Mike Joy               | University of Warwick                             |
| Kenneth Koedinger      | Carnegie Mellon University                        |
| Irena Koprinska        | University of Sydney                              |
| Sotiris Kotsiantis     | University of Patras                              |
| Vanda Luengo           | Université Joseph Fourier                         |
| Jose Maria Luna        | University of Córdoba                             |

|                         |  |
|-------------------------|--|
| Maria Luque             | University of Córdoba                            |
| Lina Markauskaite       | University of Sydney                             |
| Manolis Mavrikis        | London Knowledge Lab                             |
| Oleksiy Mazhelis        | University of Jyväskylä                          |
| Riccardo Mazza          | University of Lugano                             |
| Gordon McCalla          | University of Saskatchewan                       |
| Victor Menendez         | Universidad Autónoma de Yucatán                  |
| Agathe Merceron         | Beuth University of Applied Sciences Berlin      |
| Donatella Merlini       | Università di Firenze                            |
| Cristian Mihaescu       | University of Craiova                            |
| Julià Minguillón        | Universitat Oberta de Catalunya                  |
| Piotr Mitros            | EdX  |
| Carlos Monroy           | Rice University                                  |
| Jack Mostow             | Carnegie Mellon University                       |
| Tristan Nixon           | Carnegie Learning Inc.                           |
| Roger Nkambou           | Université du Québec À Montréal                  |
| Juan Luis Olmo          | University of Córdoba                            |
| Andrew Olney            | University of Memphis                            |
| Alvaro Ortigosa         | Universidad Autonoma de Madrid                   |
| Abelardo Pardo          | University of Sydney                             |
| Zach Pardos             | UC Berkeley                                      |
| Radek Pelanek           | Masaryk University                               |
| Kaska Porayska-Pomsta   | Birkbeck College                                 |
| Martina Rau             | University of Wisconsin-Madison                  |
| Steve Ritter            | Carnegie Learning, Inc.                          |
| Robby Robson            | Eduworks Corporation                             |
| José Raúl Romero        | University of Córdoba                            |
| Carolyn Rose            | Carnegie Mellon University                       |
| Olga Santos             | UNED   |
| George Siemens          | SoLAR  |
| Erica Snow              | Arizona State University                         |
| John Stamper            | Carnegie Mellon University                       |
| Jun-Ming Su             | National Chiao Tung University                   |
| Ling Tan                | Australian Council for Educational Research      |
| Sebastián Ventura       | University of Cordoba                            |
| Katrien Verbert         | KU Leuven  |
| Alina Vondaviev         | Educational Testing Service                      |
| Feng-Hsu Wang           | Ming Chuan University                            |
| Stephan Weibelzahl      | Private University of Applied Sciences Göttingen |
| Kalina Yacef            | University of Sydney                             |
| Michael Yudelson        | Carnegie Learning, Inc.                          |
| Amelia Zafra Gómez      | University of Córdoba                            |
| Osmar Zaiane            | University of Alberta                            |
| Alfredo Zapata González | Universidad Autonoma de Yucatan                  |
| Marta Zorrilla          | University of Cantabria                          |

**PROGRAM COMMITTEE INDUSTRY TRACK**

|                   |   |
|-------------------|---|
| Ryan Baker        | Teachers College, Columbia University       |
| Ilya Goldin       | Pearson                                     |
| Michael Yudelson  | Carnegie Learning, Inc.                     |
| George Siemens    | University of Texas - Arlington             |
| Steve Ritter      | Carnegie Learning, Inc.                     |
| Zach Pardos       | UC Berkeley                                 |
| Lalitha Agnihotri | McGraw Hill Education                       |
| Alina Vondaviev   | Educational Testing Service (ETS)           |
| Manolis Mavrikis  | London Knowledge Lab                        |
| Mingyu Feng       | SRI International                           |
| Ling Tan          | Australian Council for Educational Research |
| Robby Robson      | Eduworks Corporation                        |
| Patrick Jermann   | EPFL - Center for Digital Education         |

**ADDITIONAL REVIEWERS**

Aurora Ramirez  
Andreas Formiconi  
Andrew Hicks  
Ani Aghababyan  
April Galyardt  
Aurora Ramirez  
Barbara Furletti  
Beate Grawemeyer  
Behrooz Mostafavi  
Ben Toussaint  
Caitlin Mills  
Chen Lin  
Christopher Maclellan  
David García-Solórzano  
Douglas Selent  
Eliane Stampfer  
Ekaterina Vasilyeva  
Hannah Gogel  
James Segedy

John Kinnebrew  
Josep Grau  
Juan Jesús Alcolea Picazo  
Kelly Rivers  
Korinn Ostrow  
Kristin Stephens-Martinez  
Laura Grassini  
Leonardo Grilli  
Linting Xue  
Maria Cecilia Verri  
Matthew Jacovina  
Michael Sao Pedro  
Mouna Selmi  
Nayyar Zaidi  
Nick Green  
Nigel Bosch  
Paolo Cintia  
Peter Brusilovsky  
Ran Liu

Rinat Rosenberg-Kima  
Sameer Bhatnagar  
Sarah Schultz  
Satabdi Basu  
Sébastien Lallé  
Sergio Salmeron-Majadas  
Seth Adjei  
Shitian Shen  
Sokratis Karkalas  
Srecko Joksimovic  
Steven Tang  
Terry Peckham  
Thomas Price  
Vitomir Kovanovic  
Yan Wang  
Yun Huang  
Zhongxiu Liu



## Sponsors

### Commercial Sponsors

Gold



<https://www.mari.com/>

Gold



<http://researchnetwork.pearson.com/digital-data-analytics-and-adaptive-learning>

Lanyard



<https://www.duolingo.com/>

### Student Support Sponsors



<http://www.nsf.gov>



<http://ramkumarfoundation.org/>

### Academic Sponsors



<http://www.uned.es/>



## Table of Contents

### Invited Talks (abstracts)

- Behind the Scenes of Duolingo** 3  
*Luis Von Ahn, Matt Streeter*
- Personal Knowledge/Learning Graph** 5  
*George Siemens, Ryan Baker, Dragan Gasevic*
- Educational Neuroscience as a Tool to Understand Learning and Learning Disabilities in Mathematics** 7  
*Pekka Räsänen*

### Panels (abstracts)

- The Future of Practical Applications of EDM at Scale (Industry track)** 11  
*Ryan Baker, John Carney, Piotr Mitros, Bror Saxberg (moderator), John Stamper*
- Ethics and Privacy in EDM** 13  
*Dragan Gasevic, Taylor Martin (moderator), Zach Pardos, Mykola Pechenizkiy, John Stamper, Osmar Zaiane*
- Grand Challenges for EDM and Related Research Areas** 15  
*Ryan Baker (moderator), Peter Brusilovsky, Dragan Gasevic, Neil T. Heffernan, Mykola Pechenizkiy, Alyssa Wise*

### JEDM Track journal papers (abstracts)

- Metrics for Evaluation of Student Models** 19  
*Radek Pelánek*
- Multi-Armed Bandits for Intelligent Tutoring Systems** 21  
*Benjamin Clement, Didier Roy, Pierre-Yves Oudeyer, Manuel Lopes*
- Developing Computational Methods to Measure and Track Learners' Spatial Reasoning in an Open-Ended Simulation** 23  
*Aditi Mallavarapu, Leilah Lyons, Tia Shelley, Brian Slattery, Moira Zellner, Emily Minor*
- Move your lamp post: Recent data reflects learner knowledge better than older data** 25  
*April Galyardt, Ilya Goldin*

### Full Papers

- Combining techniques to refine item to skills Q-matrices with a partition tree** 29  
*Michel Desmarais, Peng Xu and Behzad Beheshti*
- On the Performance Characteristics of Latent-Factor and Knowledge Tracing Models** 37  
*Severin Klingler, Tanja Käser, Barbara Solenthaler and Markus Gross*
- Mixture Modeling of Individual Learning Curves** 45  
*Matthew Streeter*
- Accounting for Slipping and Other False Negatives in Logistic Models of Student Learning** 53  
*Christopher Maclellan, Ran Liu and Kenneth Koedinger*
- Learning Environments and Inquiry Behaviors in Science Inquiry Learning: How Their Interplay Affects the Development of Conceptual Understanding in Physics** 61  
*Engin Bumbacher, Shima Salehi, Miriam Wierzchula and Paulo Blikstein*

|   |            |
|---|------------|
| <b>Toward a Real-time (Day) Dreamcatcher: Detecting Mind Wandering Episodes During Online Reading</b>   | <b>69</b>  |
| <i>Caitlin Mills and Sidney D'Mello</i>   |            |
| <b>A Comparison of Face-based and Interaction-based Affect Detectors in Physics Playground</b>  | <b>77</b>  |
| <i>Shiming Kai, Luc Paquette, Ryan S. Baker, Nigel Bosch, Sidney D'mello, Jaclyn Ocumpaugh, Valerie Shute and Matthew Ventura</i>                     |            |
| <b>Exploring Dynamical Assessments of Affect, Behavior, and Cognition and Math State Test Achievement</b>   | <b>85</b>  |
| <i>Maria Ofelia San Pedro, Erica Snow, Ryan Baker, Danielle McNamara and Neil Heffernan</i>   |            |
| <b>Sensor-Free or Sensor-Full: A Comparison of Data Modalities in Multi-Channel Affect Detection</b>  | <b>93</b>  |
| <i>Luc Paquette, Jonathan Rowe, Ryan Baker, Bradford Mott, James Lester, Jeanine Defalco, Keith Brawner, Robert Sottolare and Vasiliki Georgoulas</i> |            |
| <b>Machine Beats Experts: Automatic Discovery of Skill Models for Data-Driven Online Courseware Refinement</b>  | <b>101</b> |
| <i>Noboru Matsuda, Tadanobu Furukawa, Norman Bier and Christos Faloutsos</i>  |            |
| <b>Student Models for Prior Knowledge Estimation</b>  | <b>109</b> |
| <i>Juraj Nižnan, Radek Pelánek and Jirí Rihák</i>   |            |
| <b>Discovering Prerequisite Structure of Skills through Probabilistic Association Rules Mining</b>  | <b>117</b> |
| <i>Yang Chen, Wuillemmin Pierre-Henri and Jean-Marc Labat</i>   |            |
| <b>Choosing to Interact: Exploring the Relationship Between Learner Personality, Attitudes, and Tutorial Dialogue Participation</b>                   | <b>125</b> |
| <i>Aysu Ezen-Can and Kristy Elizabeth Boyer</i>   |            |
| <b>Considering the Influence of Prerequisite Performance on Wheel Spinning</b>  | <b>129</b> |
| <i>Hao Wan and Joseph Beck</i>  |            |
| <b>Comparing Novice and Experienced Students in Virtual Performance Assessments</b>   | <b>136</b> |
| <i>Yang Jiang, Luc Paquette, Ryan Baker and Jody Clarke-Midura</i>  |            |
| <b>The Impact of Incorporating Student Confidence Items into an Intelligent Tutor: A Randomized Controlled Trial</b>                                  | <b>144</b> |
| <i>Charles Lang, Neil Heffernan, Korinn Ostrow and Yutao Wang</i>   |            |
| <b>Analyzing Early At-Risk Factors in Higher Education e-Learning Courses (Industry track)</b>  | <b>150</b> |
| <i>Ryan Baker, David Lindrum, Mary Jane Lindrum and David Perkowski</i>   |            |
| <b>Do Country Stereotypes Exist in Educational Data? A Clustering Approach for Large, Sparse, and Weighted Data</b>                                   | <b>156</b> |
| <i>Mirka Saarela and Tommi Kärkkäinen</i>   |            |
| <b>Student Privacy and Educational Data Mining: Perspectives from Industry (Industry track)</b>   | <b>164</b> |
| <i>Jennifer Sabourin, Lucy Kosturko, Clare Fitzgerald and Scott Mcquiggan</i>   |            |
| <b>Beyond Prediction: Towards Automatic Intervention in MOOC Student Stop-out</b>   | <b>171</b> |
| <i>Jacob Whitehill, Joseph Williams, Glenn Lopez, Cody Coleman and Justin Reich</i>   |            |
| <b>From Predictive Models to Instructional Policies</b>   | <b>179</b> |
| <i>Joseph Rollinson and Emma Brunskill</i>  |            |
| <b>Your Model Is Predictive— but Is It Useful? Theoretical and Empirical Considerations of a New Paradigm for Adaptive Tutoring Evaluation</b>        | <b>187</b> |
| <i>José González-Brenes and Yun Huang</i>   |            |
| <b>Automated Session-Quality Assessment for Human Tutoring Based on Expert Ratings of Tutoring Success</b>  | <b>195</b> |
| <i>Benjamin Nye, Donald Morrison and Borhan Samei</i>   |            |
| <b>A Framework for Multifaceted Evaluation of Student Models</b>  | <b>203</b> |
| <i>Yun Huang, José González-Brenes, Rohit Kumar and Peter Brusilovsky</i>   |            |

|   |            |
|---|------------|
| <b>Predicting Student Performance In a Collaborative Learning Environment</b>   | <b>211</b> |
| <i>Jennifer Olsen, Vincent Alevan and Nikol Rummel</i>  |            |
| <b>Learning Instructor Intervention from MOOC Forums: Early Results and Issues</b>  | <b>218</b> |
| <i>Muthu Kumar Chandrasekaran, Min-Yen Kan, Bernard C.Y.Tan and Kiruthika Ragupathi</i>   |            |
| <b>Investigating How Student's Cognitive Behavior in MOOC Discussion Forum Affect Learning Gains</b>  | <b>226</b> |
| <i>Xu Wang, Diyi Yang, Miaomiao Wen, Kenneth Koedinger and Carolyn Rose</i>   |            |
| <b>Methodological Challenges in the Analysis of MOOC Data for Exploring the Relationship between Discussion Forum Views and Learning Outcomes</b>             | <b>234</b> |
| <i>Yoav Bergner, Deirdre Kerr and David Pritchard</i>   |            |
| <b>Influence Analysis by Heterogeneous Network in MOOC Forums: What can We Discover?</b>  | <b>242</b> |
| <i>Zhuoxuan Jiang, Yan Zhang, Chi Liu and Xiaoming Li</i>   |            |
| <b>Modeling Learners' Social Centrality and Performance through Language and Discourse</b>  | <b>250</b> |
| <i>Nia Dowell, Oleksandra Skrynyk, Srecko Joksimovic, Arthur Graesser, Shane Dawson, Dragan Gašević, Pieter de Vries, Thieme Hennis and Vitomir Kovanovic</i> |            |
| <b>You are your words: Modeling Students' Vocabulary Knowledge with Natural Language Processing Techniques</b>  | <b>258</b> |
| <i>Laura Allen and Danielle McNamara</i>  |            |
| <b>Automatic Identification of Nutritious Contexts for Learning Vocabulary Words</b>  | <b>266</b> |
| <i>Jack Mostow, Donna Gates, Ross Ellison and Rahul Goutam</i>  |            |
| <b>Mining a Written Values Affirmation Intervention to Identify the Unique Linguistic Features of Stigmatized Groups</b>                                      | <b>274</b> |
| <i>Travis Riddle, Sowmya Bhagavatula, Weiwei Guo, Smaranda Muresan, Geoff Cohen, Jonathan Cook and Valerie Purdie-Vaughns</i>                                 |            |
| <b>Automatic Classification of Question &amp; Answer Discourse Segments from Teacher's Speech in Classrooms</b>   | <b>282</b> |
| <i>Nathaniel Blanchard, Sidney D'Mello, Andrew Olney and Martin Nystrand</i>  |            |
| <b>Topic Transition in Educational Videos Using Visually Salient Words</b>  | <b>289</b> |
| <i>Ankit Gandhi, Arijit Biswas and Om Deshmukh</i>  |            |
| <b>YouEDU: Addressing Confusion in MOOC Discussion Forums by Recommending Instructional Video Clips</b>   | <b>297</b> |
| <i>Akshay Agrawal, Jagadish Venkatraman, Shane Leonard and Andreas Paepcke</i>  |            |
| <b>Seeing the Instructor in Two Video Styles: Preferences and Patterns</b>  | <b>305</b> |
| <i>Suma Bhat, Phakpoom Chinprutthiwong and Michelle Perry</i>   |            |
| <b>Using Partial Credit and Response History to Model User Knowledge</b>  | <b>313</b> |
| <i>Eric Van Inwegen, Seth Adjei, Yan Wang and Neil Heffernan</i>  |            |
| <b>Translating Head Motion into Attention - Towards Processing of Student's Body-Language</b>   | <b>320</b> |
| <i>Mirko Raca, Lukasz Kidzinski and Pierre Dillenbourg</i>  |            |
| <b>Using Visual Analytics Tool for Improving Data Comprehension</b>   | <b>327</b> |
| <i>Jan Geryk</i>  |            |
| <b>Data-Driven Problem Profiling</b>  | <b>335</b> |
| <i>Behrooz Mostafavi, Zhongxiu Liu and Tiffany Barnes</i>   |            |
| <b>Interaction Network Estimation: Predicting Problem-Solving Diversity in Interactive Environments</b>   | <b>342</b> |
| <i>Michael Eagle, Andrew Hicks and Tiffany Barnes</i>   |            |
| <b>Why Do the Rich Get Richer? A Structural Equation Model to Test How Spatial Skills Affect Learning with Representations</b>                                | <b>350</b> |
| <i>Martina Rau</i>  |            |

## Short Papers

|   |     |
|---|-----|
| <b>Spectral Bayesian Knowledge Tracing</b>  | 360 |
| <i>Mohammad Hassan Falakmasir, Michael Yudelson, Steve Ritter and Kenneth Koedinger</i>   |     |
| <b>Direct Estimation of the Minimum RSS Value for Training Bayesian Knowledge Tracing Parameters</b>  | 364 |
| <i>Francesc Martori Adrian, Jordi Cuadros and Lucinio González-Sabaté</i>   |     |
| <b>Goodness of Fit of Skills Assessment Approaches: Insights from Patterns of Real vs. Synthetic Data Sets</b>  | 368 |
| <i>Behzad Beheshti and Michel Desmarais</i>   |     |
| <b>A Transfer Learning Approach for Applying Matrix Factorization to Small ITS Datasets</b>   | 372 |
| <i>Lydia Voß, Carlotta Schatten, Claudia Mazziotti and Lars Schmidt-Thieme</i>  |     |
| <b>Towards Understanding How to Leverage Sense-making, Induction/Refinement and Fluency to Improve Robust Learning</b>                                | 376 |
| <i>Shayan Doroudi, Kenneth Holstein, Vincent Alevan and Emma Brunskill</i>  |     |
| <b>Learning Behavior Characterization with Multi-Feature, Hierarchical Activity Sequences</b>   | 380 |
| <i>Cheng Ye, John S. Kinnebrew, James R. Segedy and Gautam Biswas</i>   |     |
| <b>Discrimination-Aware Classifiers for Student Performance Prediction</b>  | 384 |
| <i>Ling Luo, Irena Koprinska and Wei Liu</i>  |     |
| <b>Language to Completion: Success in an Educational Data Mining Massive Open Online Class</b>  | 388 |
| <i>Scott Crossley, Danielle McNamara, Ryan Baker, Yuan Wang, Luc Paquette, Tiffany Barnes and Yoav Bergner</i>  |     |
| <b>A Comparative Study of Regression and Classification Algorithms for Modelling Students' Academic Performance</b>                                   | 392 |
| <i>Pedro Strecht, Luis Cruz, Carlos Soares, João Mendes-Moreira and Rui Abreu</i>   |     |
| <b>Predicting Student Grade based on Free-style Comments using Word2Vec and ANN by Considering Prediction Results Obtained in Consecutive Lessons</b> | 396 |
| <i>Jingyi Luo, Shaymaa E.Sorour, Tsunenori Mine and Goda Kazumasa</i>   |     |
| <b>Learning the Creative Potential of Students by Mining a Word Association Task</b>  | 400 |
| <i>Cristian Olivares-Rodríguez and Mariluz Guenaga</i>  |     |
| <b>Optimizing Partial Credit Algorithms to Predict Student Performance</b>  | 404 |
| <i>Korinn Ostrow, Christopher Donnelly and Neil Heffernan</i>   |     |
| <b>Identifying Styles and Paths toward Success in MOOCs</b>   | 408 |
| <i>Kshitij Sharma, Patrick Jermann and Pierre Dillenbourg</i>   |     |
| <b>Analyzing Student Inquiry Data Using Process Discovery and Sequence Classification</b>   | 412 |
| <i>Bruno Emond and Scott Buffett</i>  |     |
| <b>Desirable Difficulty and Other Predictors of Effective Item Orderings</b>  | 416 |
| <i>Steven Tang, Hannah Gogel, Elizabeth McBride and Zachary Pardos</i>  |     |
| <b>Variations in Learning Rate: Student Clustering Based on Systematic Residual Error Patterns Across Practice Opportunities</b>                      | 420 |
| <i>Ran Liu and Kenneth R Koedinger</i>  |     |
| <b>Evaluating Educational Videos using Bayesian Knowledge Tracing and Big Data</b>  | 424 |
| <i>Zachary Machardy and Zachary Pardos</i>  |     |
| <b>Measuring Problem Solving Skills in Plants vs. Zombies 2</b>   | 428 |
| <i>Valerie Shute, Gregory Moore and Lubin Wang</i>  |     |
| <b>Strategic Game Moves Mediate Implicit Science Learning</b>   | 432 |
| <i>Elizabeth Rowe, Ryan Baker and Jodi Asbell-Clarke</i>  |     |

|   |     |
|---|-----|
| <b>Predicting learning-related emotions from students' textual classroom feedback via Twitter</b><br><i>Nabeela Altrabsheh, Mihaela Cocea and Sanaz Fallahkhair</i>   | 436 |
| <b>Video-Based Affect Detection in Noninteractive Learning Environments</b><br><i>Yuxuan Chen, Nigel Bosch and Sidney D'Mello</i>   | 440 |
| <b>Modeling Classroom Discourse: Do Models of Predicting Dialogic Instruction Properties Generalize across Populations?</b><br><i>Borhan Samei, Andrew Olney, Sean Kelly, Martin Nystrand, Sidney D'Mello, Nathan Blanchard and Arthur C. Graesser</i>                  | 444 |
| <b>Breaking Off Engagement: Readers' Cognitive Decoupling as a Function of Reader and Text Characteristics</b><br><i>Patricia Goedecke, Daqi Dong, Genghu Shi, Shi Feng, Evan Risko, Andrew Olney, Sidney D'Mello and Arthur C. Graesser</i>                            | 448 |
| <b>Semantic Similarity Graphs of Mathematics Word Problems: Can Terminology Detection Help?</b><br><i>Rogers Jeffrey Leo John, Rebecca J. Passonneau and Thomas S. McTavish</i>   | 452 |
| <b>An Analysis of Peer-submitted and Peer-reviewed Answer Rationales in a Web-based Peer Instruction Based Learning Environment</b><br><i>Sameer Bhatnagar, Michel Desmarais, Chris Whittaker, Nathaniel Lasry, Michael Dugdale, Kevin Lenton and Elizabeth Charles</i> | 456 |
| <b>Learning Analytics Platform. Towards an Open Scalable Streaming Solution for Education</b><br><i>Nicholas Lewkow, Neil Zimmerman, Mark Riedesel and Alfred Essa</i>  | 460 |
| <b>Improving Student Performance Using Nudge Analytics (Industry track)</b><br><i>Jacqueline Feild</i>  | 464 |
| <b>Educational Reports That Scale Across Users and Data (Industry track)</b><br><i>Rob Rolleston, Richard Howe and Mary Ann Sprague</i>   | 468 |
| <b>Mining Login Data For Actionable Student Insight (Industry track)</b><br><i>Lalitha Agnihotri, Ani Aghababayan, Shirin Mojarad, Mark Riedesel and Al Essa</i>  | 472 |
| <b>Building Models to Predict Hint-or-Attempt Actions of Students</b><br><i>Francisco Enrique Vicente Castro, Seth Adjei, Tyler Colombo and Neil Heffernan</i>  | 476 |
| <b>Modeling Students' Memory for Application in Adaptive Educational Systems</b><br><i>Radek Pelánek</i>  | 480 |
| <b>Social Facilitation Effects by Pedagogical Conversational Agent: Lexical Network Analysis in an Online Explanation Task</b><br><i>Yugo Hayashi</i>   | 484 |
| <b>Personalized Education; Solving a Group Formation and Scheduling Problem for Educational Content</b><br><i>Sanaz Bahargam, Dora Erdos, Azer Bestavros and Evimaria Terzi</i>   | 488 |
| <b>An Approach of Collaboration Analytics in MOOCs Using Social Network Analysis and Influence Diagram</b><br><i>Antonio R. Anaya, Jesús G. Boticario, Emilio Letón and Félix Hernández-Del-Olmo</i>  | 492 |
| <b>On Convergence of Cognitive and Non-cognitive Behavior in Collaborative Activity</b><br><i>Diego Luna Bazaldua, Saad Khan, Alina von Davier, Jiangang Hao, Lei Liu and Zuowei Wang</i>   | 496 |
| <b>The Impact of Small Learning Group Composition on Drop-Out Rate and Learning Performance in a MOOC</b><br><i>Zhilin Zheng, Tim Vogelsang and Niels Pinkwart</i>  | 500 |
| <b>Exploring Causal Mechanisms in a Randomized Effectiveness Trial of the Cognitive Tutor Algebra I Program</b><br><i>Adam Sales and John Pane</i>  | 504 |

|   |            |
|---|------------|
| <b>Confounding Carelessness? Exploring Causal Relationships Between Carelessness, Affect, Behavior, and Learning in Cognitive Tutor Algebra Using Graphical Causal Models</b> | <b>508</b> |
| <i>Stephen Fancsali</i>   |            |
| <b>Students at Risk: Detection and Remediation</b>  | <b>512</b> |
| <i>Irena Koprinska, Joshua Stretton and Kalina Yacef</i>  |            |
| <b>Intelligent Tutor Recommender System for On-Line Educational Environments</b>  | <b>516</b> |
| <i>Cristian Mihaescu, Paul Stefan Popescu and Costel Ionascu</i>  |            |
| <b>Discovering the Pedagogical Resources that Assist Students to Answer Questions Correctly – A Machine Learning Approach</b>   | <b>520</b> |
| <i>Giora Alexandron, Qian Zhou and David Pritchard</i>  |            |
| <b>Using Topic Segmentation Models for the Automatic Organisation of MOOCs resources</b>  | <b>524</b> |
| <i>Ghada Alharbi and Thomas Hain</i>  |            |
| <b>How High School, College, and Online Students Differentially Engage with an Interactive Digital Textbook</b>   | <b>528</b> |
| <i>Jeremy Warner, John Doorenbos, Bradley Miller and Philip Guo</i>   |            |
| <b>Modeling Exercise Relationships in E-Learning: A Unified Approach</b>  | <b>532</b> |
| <i>Haw-Shiuan Chang, Hwai-Jung Hsu and Kuan-Ta Chen</i>   |            |
| <b>Using Knowledge Components for Collaborative Filtering in Adaptive Tutoring Systems</b>  | <b>536</b> |
| <i>Peter Halkier Nicolajsen and Barbara Plank</i>   |            |
| <b>Exploring the Influence of ICT in online Education Through Data Mining Tools</b>   | <b>540</b> |
| <i>Javier Bravo, Sonia Janeth Romero, María Luna and Sonia Pamplona</i>   |            |
| <b>Understanding Revision Planning in Peer-Reviewed Writing</b>   | <b>544</b> |
| <i>Alok Baikadi, Christian Schunn and Kevin Ashley</i>  |            |
| <b>Convergent Validity of a Student Model: Recent-Performance Factors Analysis</b>  | <b>548</b> |
| <i>Ilya Goldin and April Galyardt</i>   |            |
| <br>  |            |
| <b>Posters and Demos</b>  |            |
| <b>Automatic Grading of Short Answers for MOOC via Semi-supervised Document Clustering</b>  | <b>554</b> |
| <i>Shumin Jing</i>  |            |
| <b>Discovering Students' Navigation Paths in Moodle</b>   | <b>556</b> |
| <i>Alejandro Bogarin, Cristobal Romero and Rebeca Cerezo</i>  |            |
| <b>Teacher-Student Classroom Interactions: A Computational Approach</b>   | <b>558</b> |
| <i>Arnon HersHKovitz, Agathe Merceron and Amran Shamaly</i>   |            |
| <b>Modeling Student Learning: Binary or Continuous Skill?</b>   | <b>560</b> |
| <i>Radek Pelánek</i>  |            |
| <b>An Analysis of Response Times in Adaptive Practice of Geography Facts</b>  | <b>562</b> |
| <i>Jan Papoušek, Radek Pelánek, Jiří Řihák and Vít Stanislav</i>  |            |
| <b>Achievement versus Experience: Predicting Students' Choices during Gameplay</b>  | <b>564</b> |
| <i>Erica Snow, Maria Ofelia San Pedro, Matthew Jacovina, Danielle McNamara and Ryan Baker</i>   |            |
| <b>How to Aggregate Multimodal Features for Perceived Task Difficulty Recognition in Intelligent Tutoring Systems</b>   | <b>566</b> |
| <i>Ruth Janning, Carlotta Schatten and Lars Schmidt-Thieme</i>  |            |
| <b>Teacher and Learner Behaviour in an Online EFL Workbook</b>  | <b>568</b> |
| <i>Krzysztof Jędrzejewski, Mikołaj Bogucki, Mikołaj Olszewski, Jan Zwoliński and Kacper Łodzиковski</i>   |            |
| <b>Skill Assessment Using Behavior Data in Virtual World</b>  | <b>570</b> |
| <i>Ailiya Borjigin, Chunyan Miao, Zhiqi Shen and Zhiwei Zeng</i>  |            |



|  |            |
|--|------------|
| <b>Pacing through MOOCs: Course Design or Teaching Effect?</b>   | <b>572</b> |
| <i>Lorenzo Vigentini and Andrew Claypahn</i>   |            |
| <b>Integrating a Web-based ITS with DM tools for Providing Learning Path Optimization and Visual Analytics</b>                     | <b>574</b> |
| <i>Igor Jugo, Božidar Kovačić and Vanja Slavuj</i>   |            |
| <b>Different Patterns of Students' Interaction with Moodle and Their Relationship with Achievement</b>                             | <b>576</b> |
| <i>Rebeca Cerezo, Miguel Sanchez-Santillan, Jose C Nuñez and M. Puerto Paule</i>   |            |
| <b>Educational Data Mining in an Open-Ended Remote Laboratory on Electric Circuits. Goals and Preliminary Results</b>              | <b>578</b> |
| <i>Jordi Cuadros, Lucinio Gonzalez, Susana Romero, M. Luz Guenaga, Javier Garcia-Zubia and Pablo Orduña</i>                        |            |
| <b>Discovering Process in Curriculum Data to Provide Recommendation</b>  | <b>580</b> |
| <i>Ren Wang and Osmar Zaiane</i>   |            |
| <b>Improving Long-Term Retention Level in an Environment of Personalized Expanding Intervals</b>                                   | <b>582</b> |
| <i>Xiaolu Xiong and Joseph Beck</i>  |            |
| <b>Exploring Problem-Solving Behavior in an Optics Game</b>  | <b>584</b> |
| <i>Michael Eagle, Rebecca Brown, Elizabeth Rowe, Tiffany Barnes, Jodi Asbell-Clarke and Teon Edwards</i>                           |            |
| <b>Simulating Multi-Subject Momentary Time Sampling</b>  | <b>586</b> |
| <i>Luc Paquette, Jaclyn Ocumpaugh and Ryan Baker</i>   |            |
| <b>Analyzing Students' Interaction Based on Their Response to Determine Their Learning Outcomes</b>                                | <b>588</b> |
| <i>Fazel Keshtkar, Jordan Cowart, Ben Kingen and Andrew Crutcher</i>   |            |
| <b>Exploring the Impact of Spacing in Mathematics Learning through Data Mining</b>   | <b>590</b> |
| <i>Richard Tibbles</i>   |            |
| <b>Data-Driven Analyses of Electronic Text Books</b>   | <b>592</b> |
| <i>Ahcène Boubekki, Ulf Kröhne, Frank Goldhammer, Waltraud Schreiber and Ulf Brefeld. Toward</i>                                   |            |
| <b>How to Visualize Success: Presenting Complex Data in a Writing Strategy Tutor</b>   | <b>594</b> |
| <i>Matthew Jacovina, Erica Snow, Laura Allen, Rod Roscoe, Jennifer Weston, Jianmin Dai and Danielle McNamara</i>                   |            |
| <b>Adjusting the weights of assessment elements in the evaluation of Final Year Projects</b>                                       | <b>596</b> |
| <i>Mikel Villamañe, Mikel Larrañaga, Ainhoa Alvarez and Begoña Ferrero</i>   |            |
| <b>Predicting Students' Outcome by Interaction Monitoring</b>  | <b>598</b> |
| <i>Samara Ruiz, Maite Urretavizcaya and Isabel Fernandez-Castro</i>  |            |
| <b>Hierarchical Dialogue Act Classification in Online Tutoring Sessions</b>  | <b>600</b> |
| <i>Borhan Samei, Vasile Rus, Benjamin Nye and Donald M. Morrison</i>   |            |
| <b>Towards Freshmen Performance Prediction</b>   | <b>602</b> |
| <i>Hana Bydžovská</i>  |            |
| <b>Generalising IRT to Discriminate Between Examinees</b>  | <b>604</b> |
| <i>Ahcène Boubekki, Ulf Brefeld and Thomas Delacroix</i>   |            |
| <b>Detection of Learners with a Performance Inconsistent with Their Effort</b>   | <b>606</b> |
| <i>Diego García-Saiz and Marta Zorrilla</i>  |            |
| <b>A Probabilistic Model for Knowledge Component Naming</b>  | <b>608</b> |
| <i>Cyril Goutte, Serge Léger and Guillaume Durand</i>  |            |
| <b>An Improved Data-Driven Hint Selection Algorithm for Probability Tutors</b>   | <b>610</b> |
| <i>Thomas Price, Collin Lynch, Tiffany Barnes and Min Chi</i>  |            |
| <b>Good Communities and Bad Communities: Does Membership Affect Performance?</b>   | <b>612</b> |
| <i>Rebecca Brown, Collin Lynch, Michael Eagle, Jennifer Albert, Tiffany Barnes, Ryan Baker, Yoav Bergner and Danielle McNamara</i> |            |
| <b>A Model for Student Action Prediction in 3D Virtual Environments for Procedural Training</b>                                    | <b>614</b> |
| <i>Diego Riofrío and Jaime Ramírez</i>   |            |
| <b>The Impact of Instructional Intervention and Practice on Help-Seeking Strategies within an ITS</b>                              | <b>616</b> |
| <i>Caitlin Tenison and Christopher Maclellan</i>   |            |

|   |            |
|---|------------|
| <b>Predicting Performance on Dichotomous Questions: Comparing Models for Large-Scale Adaptive Testing</b>                                     | <b>618</b> |
| <i>Jill-Jénn Vie, Fabrice Popineau, Jean-Bastien Grill, Éric Bruillard and Yolaine Bourda</i>   |            |
| <b>The Effect of the Distribution of Predictions of User Models</b>   | <b>620</b> |
| <i>Eric Van Inwegen, Yan Wang, Seth Adjei and Neil Heffernan</i>  |            |
| <b>Predicting Student Aptitude Using Performance History</b>  | <b>622</b> |
| <i>Anthony F. Botelho, Seth A. Adjei, Hao Wan and Neil T. Heffernan</i>   |            |
| <b>Discovering Concept Maps from Textual Sources</b>  | <b>624</b> |
| <i>Jagadeesh Chandra Bose RP, Om Deshmukh and Ravindra Bhavanam</i>   |            |
| <b>Integrating Process and Product Data: The Case of an Automated Writing Evaluation System</b>   | <b>626</b> |
| <i>Chaitanya Ramineni, Tiago Caliço and Chen Li</i>   |            |
| <b>Application of Sentiment and Topic Analysis to Teacher Evaluation Policy in the U.S.</b>   | <b>628</b> |
| <i>Antonio Moretti, Kathy McKnight and Ansaf Salieb-Aouissi</i>   |            |
| <b>Defining Mastery: Knowledge Tracing Versus N- Consecutive Correct Responses</b>  | <b>630</b> |
| <i>Kim Kelly, Yan Wang, Tamisha Thompson and Neil Heffernan</i>   |            |
| <b>A Toolbox for Adaptive Sequence Dissimilarity Measures for Intelligent Tutoring Systems (demo)</b>   | <b>632</b> |
| <i>Benjamin Paaßen, Bassam Mokbel and Barbara Hammer</i>  |            |
| <b>Carnegie Learning’s Cognitive Tutor (demo)</b>   | <b>633</b> |
| <i>Steven Ritter and Stephen Fancsali</i>   |            |
| <b>SAP: Student Attrition Predictor (demo)</b>  | <b>635</b> |
| <i>Devendra Singh Chaplot, Eunhee Rhim and Jihie Kim</i>  |            |
| <br><b>Doctoral Consortium</b>  |            |
| <b>Dynamic User Modeling within a Game-Based ITS</b>  | <b>639</b> |
| <i>Erica Snow</i>   |            |
| <b>Use of Time Information in Models behind Adaptive Practice System for Building Fluency in Mathematics</b>                                  | <b>642</b> |
| <i>Jiří Řihák</i>   |            |
| <b>Integrating Learning Styles into Adaptive e-Learning System</b>  | <b>645</b> |
| <i>Huong May Truong</i>   |            |
| <b>Modeling Speed-Accuracy Tradeoff in Adaptive System for Practicing Estimation</b>  | <b>648</b> |
| <i>Juraj Nižnan</i>   |            |
| <b>Reimagining Khan Analytics for Student Coaches</b>   | <b>651</b> |
| <i>Jim Cunningham</i>   |            |
| <b>Data Analysis Tools and Methods for Improving the Interaction Design in e-Learning</b>   | <b>653</b> |
| <i>Paul Stefan Popescu</i>  |            |
| <b>Assessing the Roles of Student Engagement and Academic Emotions within Middle School Computer-Based Learning in College-Going Pathways</b> | <b>656</b> |
| <i>Maria Ofelia San Pedro</i>   |            |
| <b>Who Do You Think I Am? Modeling Individual Differences for More Adaptive and Effective Instruction</b>                                     | <b>659</b> |
| <i>Laura Allen</i>  |            |
| <b>Developing Self-Regulated Learners Through an Intelligent Tutoring System</b>  | <b>662</b> |
| <i>Kim Kelly</i>  |            |
| <b>Data-driven Hint Generation from Peer Debugging Solutions</b>  | <b>665</b> |
| <i>Zhongxiu Liu</i>   |            |
| <b>Enhancing Student Motivation and Learning Within Adaptive Tutors</b>   | <b>668</b> |
| <i>Korinn Ostrow</i>  |            |
| <b>Estimating the Local Size and Coverage of Interaction Network Regions</b>  | <b>671</b> |
| <i>Michael Eagle</i>  |            |