

Current state of learning analytics and educational data mining

George Siemens

Ryan S.J.d. Baker

August 2013

Poll #1

- How far along is your institution in using LA/EDM at institutional level?
 - We're thinking about it
 - Ad hoc and uncoordinated
 - Committees have been struck, meetings have been held
 - Strategies and policies have been established
 - Active and successful use

Poll #2

- How far along are researchers at your institution in using LA/EDM for personal research questions?
 - Some folks talking about it
 - Some folks starting to use it
 - Individuals using it
 - Active and successful research groups
 - I have no idea

Trends in LA

Understanding the LA field

- Structured mapping of the LA space
(with Simon Buckingham Shum)

Key emerging shifts

- Expansion of areas of inquiry
- Increased technical sophistication (i.e. machine learning)

Topic areas

- Analytics approaches and methods
 - Data integration, ML, DBR, visualization
- Analytics applications and interventions
 - Intervention, learning monitoring, predictive modeling, dual focus: system & learner
- Settings of learning
 - Formal/informal, work, quantified self

Trends in EDM

Types of EDM/LA method (Baker & Siemens, in press)

- Prediction
 - Classification
 - Regression
 - Latent Knowledge Estimation
- Structure Discovery
 - Clustering
 - Factor Analysis
 - Domain Structure Discovery
 - Network Analysis
- Relationship mining
 - Association rule mining
 - Correlation mining
 - Sequential pattern mining
 - Causal data mining
- Distillation of data for human judgment
- Discovery with models



Key Emerging Shifts

- Latent knowledge estimation has become quite successful...
 - Moving from initial success in intelligent tutors for math and other well-defined domains
 - To more ill-defined domains like scientific inquiry skill
 - And to more challenging contexts like MOOCs

Key Emerging Shifts

- Latent knowledge estimation has become quite successful...
 - Moving from predicting current knowledge
 - To predicting retention of knowledge over time and preparation for future learning

Key Emerging Shifts

- Domain structure discovery has become quite successful...
 - Moving from initial success in intelligent tutors for math and other well-defined domains
 - To more ill-defined domains like video games and logical reasoning

Key Emerging Shifts

- Questions are getting bigger
- For example...

Predicting Long-Term Outcomes (San Pedro et al., EDM2013)

- Applied EDM-based detectors of affect, engagement, learning
- To data from 4k New England Middle school students interacting with educational software all year ('04-'06)
 - 2,107,108 actions within software (within 494,150 problems)
- Predict whether student will attend college ~6 years later
 - National Student Clearinghouse data
 - Logistic regression model
- Cross-validated $A'=0.69$

Predicting Long-Term Outcomes (San Pedro et al., EDM2013)

- Applied EDM-based detectors of affect, engagement, learning
- To data from 4k New England Middle school students interacting with educational software all year ('04-'06)
 - 2,107,108 actions within software (within 494,150 problems)
- Predict whether student will attend college ~6 years later
 - National Student Clearinghouse data
 - Logistic regression model
- Cross-validated $A' = 0.69$
 - Cutting class effect size 0.328σ

Predicting Long-Term Outcomes (San Pedro et al., EDM2013)

- Applied EDM-based detectors of affect, engagement, learning
- To data from 4k New England Middle school students interacting with educational software all year ('04-'06)
 - 2,107,108 actions within software (within 494,150 problems)
- Predict whether student will attend college ~6 years later
 - National Student Clearinghouse data
 - Logistic regression model
- Cross-validated $A' = 0.69$
 - Cutting class effect size 0.328σ
 - Gaming the system effect size 0.341σ

Questions?

International Educational Data Mining Society

EDM Society

- First EDM-named workshop held in 2005, at AAAI
- First EDM conference held in 2008 in Montreal
- Journal of EDM founded in 2009
 - First issue, December 2009, has 189 citations as of this writing –15.75 citations per article per year
- First Handbook of EDM published in 2010
- IEDMS founded in 2011

EDM2013

- 109 Submissions
- 25% acceptance rate
- 209 attendees

Society for Learning Analytics Research

Activities

- First conference, Banff 2011
- In co-operation with ACM (2012, 2013)
- SoLAR formed 2012

Learning Analytics Summer Institute (LASI)

Stanford University, July 1-5 2013

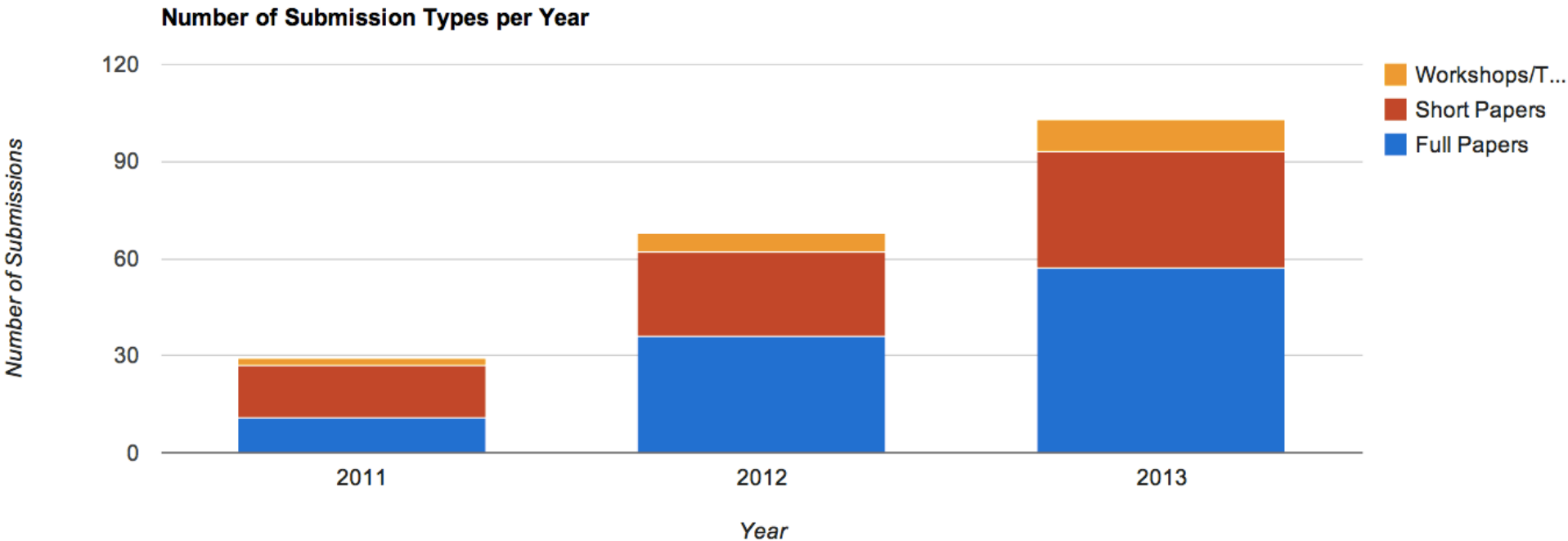


SOLAR
SOCIETY for LEARNING
ANALYTICS RESEARCH

- Distributed doctoral research lab
- Journal of Learning Analytics
- Handbook of Learning Analytics (2014)
- Local events
- Online events

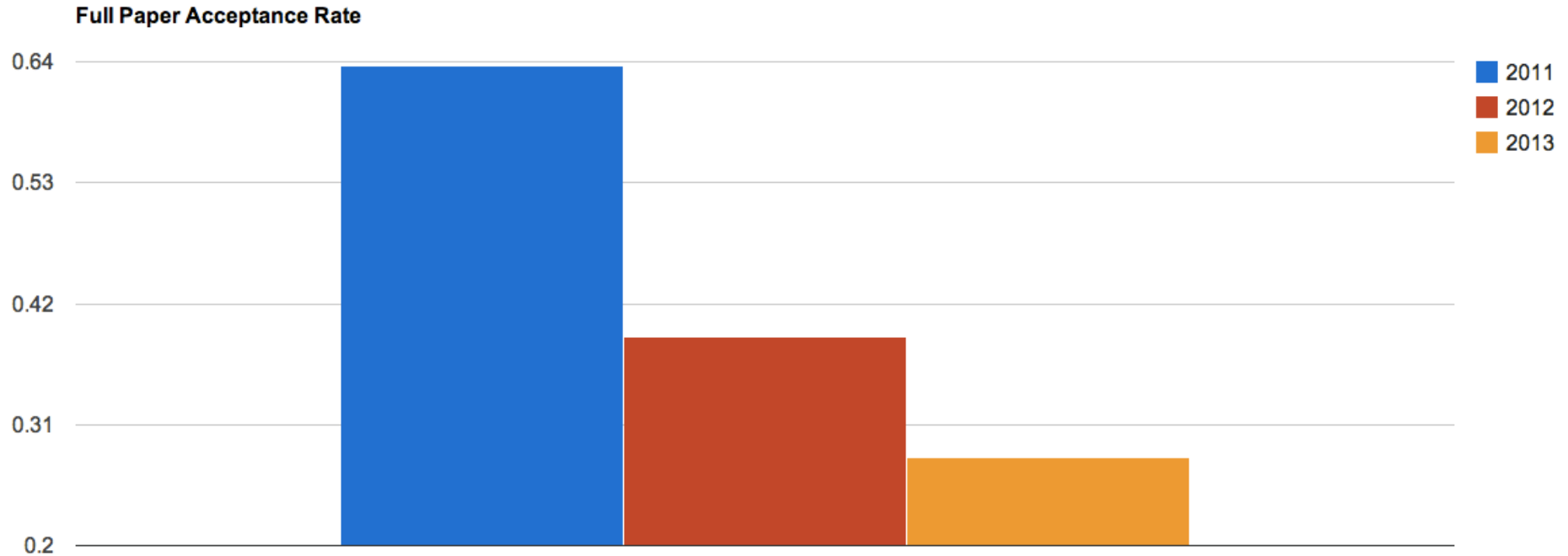
<http://www.solaresearch.org/>

LAK Conference



Courtesy: Xavier Ochoa

LAK Conference



Courtesy: Xavier Ochoa

Getting involved in SoLAR


- SoLAR:
 - Mailing list
 - Membership
 - Conference
 - Online events

<http://www.solaresearch.org/>

Getting involved in EDM Society

- 240 paid members
- You can make it 241!
 - <http://educationaldatamining.org/membership>
- 653 subscribers to edm-announce
 - Free of charge
 - <http://educationaldatamining.org/maillinglists>

Future Directions: EDM & LA



Traces of such activity may be fragmented across multiple logs and may not match analytic needs. As a result, the coherence of distributed interaction and emergent phenomena are analytically cloaked.

Suthers, Rosen 2011



Integrating data across sources and over time

- Unifying data streams to predict long-term engagement, learning, and other key outcomes
- From what's happening now
- Giving actionable information to teachers, or using it in automated interventions
- Predict the future

Integrating data across sources and over time

- Unifying data streams to predict long-term engagement, learning, and other key outcomes
- From what's happening now
- Giving actionable information to teachers, or using it in automated interventions
- Predict the future, change the future

Questions?

“Big Data and Education” on Coursera, Fall 2013



Masters in Human Development with Focus in Learning Analytics

Teachers College, Columbia University

<http://www.columbia.edu/~rsb2162/lak-concentration.html>

EDM2014

- London, UK
- July 4 - July 7

The 4th International Conference on

Learning Analytics and Knowledge

March 24 - 28, 2014

LAK14



Contact us:

Ryan S.J.d. Baker: baker2@exchange.tc.columbia.edu
(@BakerEDMLab)

George Siemens: gsiemens (gmail, twitter)