LALink: Educational Data Science (Virtual) Demos

Christopher Brooks & George Siemens, University of Texas Arlington
Pete Smith, University of Michigan
Serdar Abaci, Joshua Quick & Katy Börner, IUB

CNS Talk, Social Science Research Commons, Woodburn Hall 200

October 17, 2016
Learning Analytics

**Empowering Teachers:** How to make sense of the activities of thousands of students? How to guide them?

**Empowering Students:** How to navigate learning materials and develop successful learning collaborations across disciplines and time zones?

**Empowering Researchers:** How do people learn? What pedagogy works (in a MOOC) and when?

**Empowering MOOC Platform Designers:** What technology helps and what hurts?
Figure 1: Analysis types vs. user needs.

Educational Data Science: Precision Learning, Teaching, and Leadership
IU Emerging Area of Research Proposal

“We will develop, validate, and optimize models that explain and help predict the impact of different interventions on student success at IU and in life.”
The Team

- **Katy Borner**, Victor H. Yngve Distinguished Prof of Information Science, ILS, SOIC
- **Raymond Burke**, E.W. Kelley Prof of Marketing, KSB
- **Robert Goldstone**, Chancellor's Prof, Psychological & Brain Sciences, COAS
- **Dennis Groth**, Vice Provost for Undergraduate Education
- **Daniel Hickey**, Prof, Learning Sciences Program, SoE
- **Michael Kaganovich**, Prof of Economics, Economics, COAS
- **George Rehrey**, PI Consultant with IU’s CITL; Director of SOTL
- **Anastasia Morrone**, Prof of Educational Psychology, IUPUI School of Education; Associate Vice President for Learning Technologies, OVPIT; Dean of IT at IUPUI
- **Jennifer Meta Robinson**, Prof of Practice, Anthropology, COAS
- **Linda Shepard**, Senior Assistant Vice-Provost for Undergraduate Education; Director of Bloomington Assessment & Research
- **Timothy F. Slaper**, Indiana Business Research Center, IUB
Big Questions

• What would college students, faculty, and other stakeholders do differently if they had easy, first-hand access to the data already created by college life in the information age?
• What wisdom about learning and life could students actualize from pathways visualized through documents, data, code, expertise, laboratory outcomes, class performance, and grades?
• What leverage points for learning could faculty discern and operationalize?
• What interventions should faculty/units/institutions implement for positive gains?
Changes in Higher Education

Today
- Large crowd of students
- Directional signs
- Classroom setting
- Graduation celebration

Tomorrow
- Graduation cap saying "HIRE ME PLEASE"
- Newspaper with "JOBS" magnifying glass
- World map
- Professional career icons
Research Cores

The team will perform cutting-edge, interdisciplinary research in **Educational Data Science (EDS)** at the intersection of four research areas:

- **Cognitive Science > Classroom Experiments** investigates the cognitive and social variables, patterns, and leverage points in learning and teaching.
- **Learning Science > Student Support** investigates the impact of curricular interventions on student success at IU and in life.
- **Decision Science: Economics of Higher Education** investigates the economic value of education across scales—from micro to macro.
  
  **Management/Student Choice Research** investigates the impact of incentives and educational product offerings on short-term and long-term decision making.
- **Data Science > Learning Analytics** performs research on data mining, modelling, and visualization techniques that increase “data (visualization) literacy” and data-driven decision making.
Cyberinfrastructure Core

- Implements novel means to provision sensitive data via secure data enclaves and federated Denodo virtualized databases.
- Develops novel functionality for existing learning management systems (LMS) such as Canvas using LTI and Caliper.
- Uses/extends Tableau to serve actionable dashboards for IU leadership.
Establishing EDS and Ensuring IU Leadership

Capitalizing on existing IU strengths:

• Student Learning Analytics (SLA) Fellows Program
• Scholarship of Teaching and Learning Program
• Learning Technologies, UITS
• Learning Science Research, PBS, COAS
• Cognitive Science Program, IUB
• Learning Sciences Program, School of Education
• Bloomington Assessment and Research (BAR) office
• Indiana Business Research Center, http://ibrc.indiana.edu
• Decision Support Initiative, http://dsi.iu.edu

Proactive collaborations with other institutions:

• Unizin—11-institution digital learning consortium, http://unizin.org
• Bay View Alliance—8-institution Student Learning Analytics (SLA) initiative
The University of Michigan has engaged in a breadth approach to learning analytics, and is involved in scholarly activity in the field, applied technology development, and institutional infrastructure investment. In this talk, the rich ecosystem of educational innovation initiatives will be surveyed, with a particular focus on (a) investments in scholarly learning analytics work, including two $1.25M interdisciplinary learning analytics grants funded in part through the UM data science initiative, (b) activities in the Digital Innovation Greenhouse (DIG), which serves as an on-campus education technology accelerator to address this challenge (http://ai.umich.edu/about-ai/digital-innovation-greenhouse), and (c) institutional investment in Unizin and the development of a learning analytics architecture to enable data-driven rescission making.
Large scale data has resulted in increased interest in learning sciences and related research. Much of this research interest is coming from non-traditional education fields as physicists, biologists, and others begin to analyze the data generated by learners in online and blended environments. A second trend has been to incorporate the practices of business intelligence to improve how universities make decisions about student support, recruitment, and institutional resource allocation. In most universities, the research and the practice of analytics are treated as separate silos. At University of Texas Arlington, we have created an integrated model where our learning analytics research (LINK Research Lab) coordinates extensively with our University Analytics department. This discussion will focus on the components of an integrated research/practice system as well as the challenges and ongoing opportunities.
Learning Analytics Initiatives at Indiana University

Serdar Abaci and Joshua Quick, IUB

We will present the initial developments of learning analytics initiatives at Indiana University at an institutional level. We will also give an overview of our experience in processing, analyzing, visualizing, and interpreting the e-textbook reading behavior data that is available from the Unizin Engage e-text reader and discuss the research implications of studying instructional activity data from digital learning environments such as the Learning Management System (LMS).
Learning Analytics Innovation at the University of Michigan

Creating Scholarly, Practical, and Institutional Impact

Dr. Christopher Brooks
Research Assistant Professor, School of Information
Director of Learning Analytics & Research,
Office of Academic Innovation
University of Michigan
brooksch@umich.edu @cab938
U-M Approach To Learning Analytics

- Has a **scholarly** component...
  - a dozen faculty involved in grant writing, publishing, and graduate student mentorship in learning analytics
- Has a **practical** component...
  - building next generation tools and learning experiences for U-M learners based upon analytics and data
- Is an **institutional** activity...
  - driven at the highest levels (Provost) with service units in deep collaboration

Christopher Brooks, University of Michigan  🔄@cab938  🚮broosch@umich.edu
Scholarly

- The Provost’s Learning Analytics Task Force, Dr. Tim McKay (2012-2015)

- A significant community building activity through the Student Learning and Analytics at Michigan (SLAM) seminar series

- Increased involvement of U-M scholars with SOLAR
  - Dr. Stephanie Teasley, incoming President
  - Learning Analytics Summer Institute (LASI) 2016
Scholarly

- Continued nurturing of local community of scholars through Academic Innovation at Michigan (AIM) Analytics (2014-Present)
- Supporting next-generation learning analytics through the Michigan Institute for Data Science (MIDAS)
  - Two $1.25M interdisciplinary internal projects on learning analytics awarded for AY16
- Several external awards (NSF) to continue applied learning analytics projects (e.g. eCoach $1.9M)
Scholarly

- Broad array of topics and interests:
  - Predictive modeling of student success
  - Nudge interventions
  - Text and educational discourse analysis (META, Mwrite)
  - Massive open online courses
  - Holistic modeling of students, psychological (affective) states
- Inter- and Multidisciplinary approach
  - Faculty from Information, CS, Education, Physics, Chemistry, English...
Practical

- Learning analytics are actively used to support academic innovation at U-M
- One place is through the Digital Innovation Greenhouse (Dr. Tim McKay)
  - Interesting model of institutional funded scholarly work turned institutional asset!
  - Educational technology incubation in higher ed
    - Academic Reporting Toolkit (ART) 2.0
    - Electronic Coaching (ECoach)
    - Advisor-facing Early Warning System (Student Explorer)
Practical

- ART 2.0
- Dr. Gus Evrard
- Dr. Chris Teplov
- Data in agg from student course review

SI 365
Cyberscience: Computational Science and the Rise of the Fourth Paradigm

This course invites students to understand the ways in which data-rich, compute-intensive, collaborative research drives discovery in the natural and social sciences. Through readings, activities, and discussion with invited guests, students will gain an appreciation of the opportunities and challenges posed by the fast-growing, interdisciplinary subject of data science.

Advisory prerequisites:
Computing language course (e.g., EECS 183, 280, ENGR 101), Introductory statistics (e.g., Stats 250)

Credits:
3.00

Evaluation Data
Learn more about this evaluation data.

- 65% of respondents expressed a strong desire to take this course.
- 86% of respondents reported learning a lot from this course.
- 22% of respondents reported the workload as heavy.
Practical

- Simple data about the makeup of students in the course
Practical

- Explorable details about data about students in the course (for faculty)
Student Explorer

- Kris Steinhoff,
  Dr. Steve Lonn,
  Dr. Stephanie Teasley

- At a glance early warning for advisors
Student Explorer

- Behaviors of students in courses
ECoach

- Nudging students based on behavior and profiles
- Dr. Tim McKay
ECoach

SET YOUR GOAL
What grade are you aiming to get on your upcoming exam?

[Graph showing a rating of 80%

How motivated are you to get this grade?

1 2 3 4 5 6
Not at all

How important is it for you to achieve this grade?

1 2 3 4 5 6

Welcome back, Ben
You made it through the first statistics exam.

You scored a **61 out of 75 points** or **81.3%** which corresponds to a letter grade of a B

Here is where your exam grade falls in the class-wide distribution of exam grades.

ECoach

CHOOSE YOUR TOOLS
Which resources will help you prepare for this exam?

Please take a moment to consider the types of questions that will be on the exam.

To achieve the grade that you want, what resources (listed below) do you think will help you prepare for this exam most effectively?

- Practice exam questions (from Canvas)
- Private tutoring
- Office hours with a GSI
- Study group (or discussions with students)
- Textbook readings
- Problem Roulette past exam questions
- Video recorded lectures (on Canvas)
- ECoach
- Asking questions in lecture or lab
- Office hours with a lecturer
- Formula card
- Lab materials (ILPS)
- Past recommended HW problems
- Past required HW problems

Christopher Brooks, University of Michigan  @cab938  brooksch@umich.edu
Institutional

- Creation of a unit office and Vice Provost: Digital Education & Innovation, now the Office of Academic Innovation
- Compliments existing institutional assets
  - Not IT!
  - Not Teaching and Learning!
- Aim is to seed investment in new educational technologies, models, and methodologies
  - Technology incubation, academic innovation experimentation funding...
Institutional Assets

- Extends beyond the immediate U-M environment with Unizin Consortium
- Strong support within IT unit on campus
  - Learning Analytics Architecture (LARC) dataset
  - Helping to develop IMS Caliper specification and implementations
  - New processes for data access (MOU, IRB)
- We take (1) cool, proven novel research and (2) scale it to the institution and then (3) share it in our broader consortium
Takeaway

- Learning Analytics at Michigan is
  - Scholarly
  - Practical
  - Institutional
- A broad array of actors doing many things, this summary just scratches the surface!
Educational Data Science, Many Names

Educational Data Mining

Learning Analytics

“Big Data” in Education

And increasingly…

Machine Learning

Artificial Intelligence
Data-Enriched Educational Products

Online courses that enable the constant logging and tracking of learners through their clickstream data;

E-textbooks that can ‘learn’ from how they are used;

Adaptive learning systems that enable materials to be tailored to each student’s individual needs through automated real-time analysis;
As time goes on…

New forms of data analytics that are able to harvest data from students’ actions, learn from them, and generate predictions of individual students’ probable future performances;

Automated personal tutoring software that monitors students and gives constant real-time support and shapes the pedagogic experience.

—Mayer-Schönberger & Cukier (2014), *Learning with Big Data: The Future of Education*
And emerging today…

New forms of data analytics that are able to harvest data from students’ affective states, social and cognitive engagement;

More recently: machine learning drives AI tools such as chatbots, “smart” discussion fora, automated coaching, etc.

“Smart Campus UTA”
Behind it all…

…are models and “training data” for personal profiles
e-curriculum pathways
models of student activity, engagement, affective states
models for natural-language interaction with learners
What data are feeding our models?

At UTA, primary sources are our Student Information System (SIS) and Learning Management System (LMS).

Additional Campus Systems: Student Affairs, Library, Housing and Food Services

Federation of data from neighboring two-year colleges is/will be taking place.

Expanding Geographical Context: Arlington and the DFW Metroplex as “Smart Cities”
UTA Experience

University Analytics

New University Unit of 25 FTE
Data Scientists for Data Mining, Analytics Across the Campus Academic and Business Enterprise

Learning Innovation and Networked Knowledge (LINK) Lab

Research Facility of 24 including Faculty, Staff, Postdocs, and Graduate-level Researchers
UA Hardware and Toolsets

Civitas Learning
  Multivariate Modeling of Student Persistence, Graduation
  IaaS around Student Data

SAS
  Visual Analytics
  Enterprise Miner
  Prediction Suite
  Viya Machine Learning/Neural Network Modeling

450 Core Server Farm (Planned)
UTA “Big Data Questions”

How will big data and new models provide a more complex understanding of the learner in higher education today?

How can universities use big data to improve student success (retention and successful progress to graduation)?

Can higher education develop new, more multivariate models of student engagement? How might these models drive faculty, staff, and coaches to improve student cognitive and social presence in formal coursework?

How can we better understand learners of diversity and personalize the educational experience for engagement and success?
About Civitas Learning

- Founded in 2011 by Charles Thornburgh and Mark Milliron
- Provides cloud-based predictive analytics applications for administrators, faculty, students, and advisors
- Helps answer the question of what’s working, what’s not working, for which students, at each point in their learning journeys
## LMS Data in UTA/Civitas Model

<table>
<thead>
<tr>
<th>Activity Type</th>
<th>Civ Activity Type Id</th>
<th>Dim Date Id</th>
<th>Dim Enrollment Id</th>
<th>Dim Person Id</th>
<th>Dim Section Id</th>
<th>Event Ts</th>
<th>Raw Activity Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Announcement</td>
<td>42619</td>
<td>6469735</td>
<td>8052836</td>
<td>266752</td>
<td>9/7/2016 10:55:19 AM</td>
<td>COURSE_ACCESS_announcement</td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>42619</td>
<td>6469735</td>
<td>8052836</td>
<td>266752</td>
<td>9/7/2016 10:48:52 AM</td>
<td>COURSE_ACCESS_announcement</td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>42611</td>
<td>6415398</td>
<td>8814225</td>
<td>263394</td>
<td>8/30/2016 12:49:29 PM</td>
<td>COURSE_ACCESS_announcement</td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>42611</td>
<td>6415398</td>
<td>8814225</td>
<td>263394</td>
<td>8/30/2016 12:49:14 PM</td>
<td>COURSE_ACCESS_announcement</td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>42610</td>
<td>6446205</td>
<td>9265419</td>
<td>265888</td>
<td>8/29/2016 11:33:26 PM</td>
<td>COURSE_ACCESS_announcement</td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>42610</td>
<td>6446187</td>
<td>9035214</td>
<td>265888</td>
<td>8/29/2016 3:17:15 PM</td>
<td>COURSE_ACCESS_announcement</td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>42580</td>
<td>6354675</td>
<td>331765</td>
<td>261193</td>
<td>7/30/2016 12:41:45 PM</td>
<td>COURSE_ACCESS_announcement</td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>42580</td>
<td>6354675</td>
<td>331765</td>
<td>261193</td>
<td>7/30/2016 12:41:34 PM</td>
<td>COURSE_ACCESS_announcement</td>
<td></td>
</tr>
</tbody>
</table>
Sample Student Persistence Model Output
Sample Student Persistence Model Output
Learning Analytics Maturity Model

Figure 4: Learning Analytics Sophistication Model

Maturity of Learning Analytics Deployment
What does it mean to be human in a digital age?
Projects - dLRN

$1.6M Bill and Melinda Gates Foundation (PI)
linkresearchlab.org/dlrn
Projects - dLRN Conference

linkresearchlab.org/dlrn2015
Projects - Smart Science Network

$5.2M Bill and Melinda Gates Foundation (Co-PI)
linkresearchlab.org/research
Projects - BCC: Community and Capacity for Educational Discourse Research

$254K NSF (Co-PI)
linkresearchlab.org/research
Projects - BIGDATA: Collaborative Research

$1.6M NSF (Co-PI)
linkresearchlab.org/research
Projects - DALMOOC

linkresearchlab.org/dalmooc
Projects - Emerging Technologies and their Practical Applications in K12 Teaching and Learning MOOC

goo.gl/w9Bkdx
Projects - INTERlab

interlab.me
aWEAR
Expanding data collection to include broadening scope of data collection
Holistic learning
Individual well-being
Preparing learners for the future of work and life
## New Paradigm(s)

<table>
<thead>
<tr>
<th>Data type</th>
<th>Mode of data collection</th>
<th>Assessment genres: Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine assessments</td>
<td>Computer adaptive testing</td>
<td>Select response assessments, quizzes (e.g., reading comprehension, grammar, vocabulary)</td>
</tr>
<tr>
<td></td>
<td>Natural language processing</td>
<td>Automated essay scoring, feedback on language features</td>
</tr>
<tr>
<td></td>
<td>Procedure-defined processes</td>
<td>Games, intelligent tutors</td>
</tr>
<tr>
<td></td>
<td>Argument-defined processes</td>
<td>Rubric-based peer review of writing</td>
</tr>
<tr>
<td></td>
<td>Machine learning processes</td>
<td>Semantic tagging and annotation, text visualizations, accepted textual change suggestions</td>
</tr>
<tr>
<td>Structured, embedded data</td>
<td>Incidental “data exhaust”</td>
<td>Keystroke patterns, edit histories, clickstream and navigation paths, social interaction patterns</td>
</tr>
<tr>
<td>Unstructured, incidental data</td>
<td>Dedicated devices for collecting unstructured data</td>
<td>Video capture, eye trackers, movement detectors</td>
</tr>
</tbody>
</table>

Cope & Kalantzis (2016). “Big Data Comes to School”
# New Paradigm(s)

<table>
<thead>
<tr>
<th>Traditional research model</th>
<th>Emerging research model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher as <em>independent observer</em></td>
<td>Researchers recruit subjects as data collectors, <em>co-researchers</em></td>
</tr>
<tr>
<td>Optimal <em>sample N</em> to produce reliable results</td>
<td>There is no marginal cost for $N = \text{all}$, and data are rich enough to support $N = 1$</td>
</tr>
<tr>
<td>Practical limits to research perspective determined by the <em>scale of data collection</em></td>
<td><em>Multiscalar perspectives</em>, from $N = 1$ to $N = \text{all}$</td>
</tr>
<tr>
<td><em>Fixed time frames</em>, long enough to demonstrate overall effect; longitudinal analyses expensive and thus infrequent</td>
<td><em>Short time frames</em>, feeding small incremental changes back into the learning environment; <em>longitudinal time frames</em> as a consequence of data persistence</td>
</tr>
<tr>
<td><em>Standardization</em> effects (fidelity, average effect)</td>
<td>Tracing <em>heterogeneity</em> in data, e.g., different paths in adaptive learning environments, salient activities of outliers</td>
</tr>
<tr>
<td>Causal effects: <em>overall</em>, for whole populations or population subsets</td>
<td><em>Microgenetic</em> casual analysis, e.g., learning progressions for different students, differential effects traceable in varied learning paths</td>
</tr>
<tr>
<td>Relatively <em>separate quantitative and qualitative research practices; low significance of theory</em> in empirical analyses</td>
<td><em>Integration of quantitative and qualitative analyses; increasing importance of theory</em> in data analyses</td>
</tr>
</tbody>
</table>

Cope & Kalantzis (2016), "Big Data Comes to School"
Automated Content Analysis of Online Discussion Transcripts

Vitomir Kovanovic  
Simon Fraser University  
Vancouver, BC, Canada  
vitomir_kovanovic@sfu.ca

Sreko Joksimovic  
Simon Fraser University  
Vancouver, BC, Canada  
sjoksim@sfu.ca

Dragan Gasevic  
Athabasca University  
Edmonton, AB, Canada  
dgasevic@acm.org

Marek Hatala  
Simon Fraser University  
Vancouver, BC, Canada  
mhatala@sfu.ca

ABSTRACT
In this paper we present the results of an exploratory study that examined the use of text mining and text classification for the automation of the content analysis of discussion transcripts in the context of distance education. We used Community of Inquiry (CoI) framework and focused on the content analysis of the cognitive presence construct given its central position within the CoI model. Our results demonstrate the potentials of proposed approach; The developed classifier achieved 58.4% accuracy and Cohen’s Kappa of 0.41 for the 5-category classification task. In this paper we analyze different classification features and describe the main problems and lessons learned from the development of such a system. Furthermore, we analyzed the use of several novel classification features that are based on the specifics of cognitive presence construct and our results indicate that some of them significantly improve classification accuracy.

1. INTRODUCTION
One of the important aspects of modern distance education is the focus on the social constructivist theories of educational research, and focus mostly on the tacit aspects of the trace and log data. Given the need for the qualitative aspects of the learning products this is not to address this issue, we base our transcript analysis of the well-established Community of Inquiry (CoI) model of education [10, 11] which is used for more than a decade this type of questions.

In this paper we present a results of a study which focuses on automation of the content analysis of discussion transcripts. Community of Inquiry coding technique. We developed a classifier for automatic classification of the discuss transcripts in accordance with the CoI framework, and we discuss in detail the challenges and issues with this type of text classification, most notably the creation of relevant classification features.

2. BACKGROUND WORK
We base our work on the theoretical foundations of the Community of Inquiry framework and previous work done in the field of text classification. In this section we will present an overview of previous studies and methods used in the field of text classification.
Intercultural Frontier

Programa de Capacitación para geração e uso de energias limpas na América Latina.

Este curso, desenvolvido para ONU*SEU em cooperação com o Centro de Pesquisa para Energia, Meio Ambiente e Tecnologia (CEMAET) da Universidade de Salamanca e a Universidade Politécnica de Madrid, aborda as principais táticas para geração de energia e construção de edifícios energéticamente eficientes.

---

International Marketing

本课程通过生动的视频资料和案例分析帮助同学们深刻理解国际市场营销的核心知识与理念。掌握课程知识不仅可以让你能够解释和分析实际生活中的国际营销案例，而且还能够助力你走向未来创业创新之路。
Distributed Research Networks

Stanford
TC, Columbia
U of Edinburgh
Collaboration with UniSA
University of Penn

Knowledge is networked.
Learning is network formation.
Discovering connections and patterns is research.
Making connections is innovation, creativity, and knowledge generation.
Institutional Analytics with the Engage Platform

Serdar Abaci, PhD
Educational Research and Evaluation Specialist

Joshua Quick
Graduate Research Assistant

UITs, Learning Technologies Division
Indiana University, Bloomington
IU eText Program

• Different business model for e-text adoption
  • Instructor adopts an e-textbook
  • E-textbooks are available 1st day of the class
  • On average, 60% savings in textbook cost
  • Students maintain access to e-texts
  • Students access all of their e-texts through single platform (Courseload – Now Unizin Engage)

<table>
<thead>
<tr>
<th>PILOT</th>
<th>FULL IMPLEMENTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009-2011</td>
<td>SPRING 2012</td>
</tr>
</tbody>
</table>
Engage Platform

• Unizin online/mobile content delivery platform
• Enables students and instructors to read and annotate a shared document
• Allows for readily accessible feedback and collaboration
• Platform records these events and occurrences at institutional, course, and section levels
Such debates continue. Yet over and over again we will see that in contemporary science the nature–nurture tension dissolves: Nurture works on what nature endows. Our species is biologically endowed with an enormous capacity to learn and adapt. Moreover, every psychological event (every thought, every emotion) is simultaneously a biological event. Thus, depression can be both a brain disorder and a thought disorder.
2012-2015 Use

Engage Student Users by Year/Term

Year/Term

Count

2012 2013 2014 Spring 2015

Engage Instructor Users by Year/Term

Year/Term

Count

2012 2013 2014 Spring 2015

10/17/2016
# Student Use by Departments and Schools

<table>
<thead>
<tr>
<th>Department/School</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>Spring 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business School</td>
<td>1732</td>
<td>4456</td>
<td>6856</td>
<td>4202</td>
</tr>
<tr>
<td>Economics Department</td>
<td>481</td>
<td>350</td>
<td>667</td>
<td>68</td>
</tr>
<tr>
<td>Journalism Department</td>
<td>174</td>
<td>388</td>
<td>425</td>
<td>0</td>
</tr>
<tr>
<td>School of Public &amp; Environmental Affairs</td>
<td>75</td>
<td>266</td>
<td>222</td>
<td>110</td>
</tr>
<tr>
<td>School of Public Health</td>
<td>893</td>
<td>1661</td>
<td>1479</td>
<td>878</td>
</tr>
<tr>
<td>Statistics Department</td>
<td>0</td>
<td>735</td>
<td>1078</td>
<td>47</td>
</tr>
</tbody>
</table>
Initial Efforts

Source Data
- # number of pages viewed
- Highlights
- Notes
- Questions
- Aggregated by section

Calculations
- % of page viewed → proxy for reading engagement with platform
- Measures of central tendency and deviation were taken.
- Distribution of these section level statistics were taken to identify higher use sections for exploratory analyses.
Existing Engagement Data

Average Percentage Viewed by Section

<table>
<thead>
<tr>
<th>eText Page Percentage Viewed</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>45</td>
<td>50</td>
</tr>
<tr>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>85</td>
<td>90</td>
</tr>
<tr>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>95</td>
<td>100</td>
</tr>
</tbody>
</table>

10/17/2016 Abaci & Quick
Exploratory Analytics

• Analysis is limited to Business School data (fall 2013)
• Instructors use of features is LOW
  • 0 → no use of features
  • 1 → any use of features
• Student page view
  • LOW → % of page view at or below section median
  • HIGH → % of page view above section median
• Students use of features
  • NONE → no use of a particular feature
  • LOW → use of a particular feature at or below section median use
  • HIGH → use of a particular features above section median use
Analytic Efforts

PRIOR.GPA < 3.2

yes

PRIOR.GPA < 3.2

no

READING.CATEGORY = LOW

3.1
n=1053 100%

2.8
n=563 53%

2.6
n=313 30%

3.1
n=250 24%

3.5
n=490 47%
Challenges and Developments

• Low usage of interactive features
• Connecting to real-time Engage data
• Linking Engage data with other institutional data sources (e.g., SIS, LMS)
Future Directions

• Identifying eText usage patterns at departmental levels
• Identifying eText usage patterns at departmental levels across time
• Developing robust predictive models of eText usage and student engagement and achievement
• Text mining (i.e., Instructor and Student Notes)
Questions to Answer

• Does the level of content interaction predict the outcomes (learning, retention)?
• Does first day access to eText predict course outcomes?
Questions?
Thank You

- Serdar Abaci
  - sabaci@iu.edu

- Joshua Quick
  - jdquick@iu.edu