Personalizing Education at Scale

Data Science Challenges for Education in the 21st Century
Tim McKay
University Hall was constructed in 1871. This building featured an auditorium seating 3000. This at a time when total enrollment at the University was 1200 students. No small plans were made.
Birth of the industrial university

- In 1900, enrollment had tripled, to 3482, and the industrial era had begun
- In 1950, enrollment expanded by an additional factor of ten, to 43,683
- Michigan became the model of a modern public research university
- Indoor graduation had become impossible... 1949 commencement on Ferry Field
Commencement, June 11, 1949 at Ferry Field
The 20th Century began with an industrial revolution. Public higher education exploded in scale and bureaucratized, adopting standardized tests, measuring outcomes in credit hours, GPAs, majors, and minors.

Since the 1970’s U-M has worked hard to more deeply personalize education, with richer advising, freshman seminars, CSP, learning communities, and more.

Doing this at scale is difficult and costly. So what has been done reaches most students very thinly...
The 21st Century began with an information revolution. Public higher education has been slow to respond. But change has begun: practically all information is online, classes are flipping, many educational activities are digitally mediated.

The real revolution will come from personalization.
Personalizing education at scale

• We must be able to attend to every student:
  – As a person, with evolving background, interests, goals, identity, concerns, purpose, affect, well-being
  – As a student: we need to see what they do, assess what they know, represent their skills

• We must measure and report what matters: the elements of a liberal education
  – Intellectual breadth, disciplinary depth, range of experience, sustained engagement with desirable difficulties, networks of social and professional connection

• We must be able to act at scale:
  – Explore and understand, attend to everyone in real time, deliver actionable information to students, faculty, and staff
Today’s MIDAS Seminar

• Personalized education as holistic Data Science
  – **Ethics**: what are we doing and why
  – **Measurement**: data collection and management
  – **Analysis**: modeling, extraction of meaning, exploration of how background and structures affect outcomes
  – **Action**: decision making, storytelling, creating the motivation for change

• Our goal: Real personalized education, right here at Michigan, in five years (not twenty!)
Ethics:
What we’re doing and why
Ethical challenges

• What principles should govern the collection and use of data about individuals?
• Norms of consent, privacy, autonomy, responsibility?
• How does application of these principles change for students in our communities?

• Asilomar Convention:
  – 2014: Learning Research in Higher Education
  – Goal: to generate principles for ethical conduct of learning analytics to guide the creation of local policy

Second Asilomar meeting planned for June 2016: Learner Data and Records in the Digital Era
Six Asilomar principles

1. **Respect for the rights and dignity of learners:** transparency, protection of privacy
2. **Beneficence:** maximize benefits, minimize harm
3. **Justice:** benefit all, reduce inequalities
4. **Openness:** learning and research are public goods
5. **The humanity of learning:** insight, judgment, & discretion are essential, keep learning humane
6. **Continuous consideration:** ongoing, inclusive discussion of changing ethical circumstances
Applying these principles: e.g. “predictive modeling”

- Many early learning analytics applications use past performance of students to construct “predictive models”
- These models are really just reports of what’s happened in the past: they predict the future only if nothing changes

- How to act when past students haven’t achieved their goals?
  - “Drown the bunnies”
  - Respond with new, personalized systems of support, test and refine them

- We learn from the past in order to change the future
Two tenets for research

1. **Advance the science of learning for the improvement of higher education:** The science of learning can improve higher education and should proceed through open, participatory, and transparent processes of data collection and analysis that provide empirical evidence for knowledge claims.

2. **Share:** Maximizing the benefits of learning research requires the sharing of data, discovery, and technology among a community of researchers and educational organizations committed, and accountable to, principles of ethical inquiry held in common.
Academics and reputation

• Students/universities: an unusual relationship
  – Students want official validation of success
  – To increase value, they give up some control over this reputational record
    • Many rules about creation of the “permanent record”
    • Degree receipt (or not!) openly acknowledged
    • Transcript access, but not content, controlled by student
    • No ‘right to be forgotten’

The agreement between students and the institutions they attend presents an interesting, not yet fully explored topic for scholars of privacy and reputation.
Measurement:
Data collection and management
What do we measure?

• What we measure now:
  – Admissions information
  – Course taking & grades
  – Degrees & honors

• What we’re starting to record: explosive growth!
  – Process of learning: clickstreams, discussions, video, course structures
  – Products of learning: MC, forum posts, essays

• What we want to have:
  Detailed, evolving portraits of every student's background, interests, goals, and accomplishments

• These portraits should be used to offer admission, monitor progress, decide on graduation, and represent success
Cleaning & aggregation

• Many data sources
  – UM Data Warehouse: *hundreds* of organically evolving tables
  – LMS: CTools -> Canvas: course structures, student work
  – Digitally mediated education: chat rooms, Google docs, Piazza, online homework, Problem Roulette...

• Gathering, digesting, and merging presents many challenges
  – Using identity while preserving privacy
  – Handling very diverse student portraits
  – Digesting multimodal data streams, extracting meaning
  – Handling incompletely defined data in real time
Just for student records, there are 157 pages of data description...
A partial solution LArc: Learning Analytics Data Architecture

A ‘public release’ model for cleaning research data – compare to data releases from open science projects like the Sloan Digital Sky Survey or the GAIA space mission.
Measuring learning in classes

- Grades: performance measures of unrecorded tasks, meant to estimate unknown outcomes, quantified on ill-defined scales
- We should be measuring learning – increases in well defined knowledge and skills – and focusing on individual growth over time

- **Direct**: pre and post testing aligned with learning goals. Good for foundational courses?
- **Indirect**: DS methods for extracting meaning from all student work
  - Simple: IRT, topic modeling and beyond
  - Complex: NLP, categorization by comparison
What we measure today

• A credit-hour/degree requirement economy
  – Credit-hour designations only loosely comparable
  – Categorical degree requirements can be met in highly various ways

• Performance measured only by grades, aggregated into GPA
  – Grades awarded vary 25% by field & course level
  – Students taking classes vary dramatically
Measuring what matters

- Liberal education is more than a list of classes and grades
  - Intellectual breadth
  - Disciplinary depth
  - Range of experience
  - Engagement & effort
  - Social & professional networks
- Important outcomes are long term – we need to see beyond campus
- A multidimensional portrait of student progress
- Multiple forms of commitment and success encouraged and recognized
- Authentic goals reinforced and key outcomes noted
- Redefining student success
How might we quantify intellectual breadth?

• One example: explore each student’s network of connection – Kar Epker senior thesis
  • Course co-enrollment: well measured, large bipartite network
    – Students connected by courses
    – Courses connected by students
  • Future expansion to social networks?

• Simple measures:
  – How many connections?
  – What kinds of students?
  – How to weight, classify?

• Diversity of connection: compare interactions to random graph models
  – In degree: similar majors
  – Out degree: different

• Exposes isolation of majors, allows comparison of individuals within a major
(b) RPD between actual and expected weighted interactions of students segmented by major.

<table>
<thead>
<tr>
<th>Major</th>
<th>RPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music Theory and Composition</td>
<td>1.99</td>
</tr>
<tr>
<td>Physical Education Teaching</td>
<td>1.99</td>
</tr>
<tr>
<td>Dance, General</td>
<td>1.99</td>
</tr>
<tr>
<td>Technical Theatre/Theatre Design</td>
<td>1.99</td>
</tr>
<tr>
<td>Musical Theatre</td>
<td>1.99</td>
</tr>
<tr>
<td>Athletic Training/Trainer</td>
<td>1.98</td>
</tr>
<tr>
<td>Music History, Literature</td>
<td>1.98</td>
</tr>
<tr>
<td>Dental Hygiene/Hygienist</td>
<td>1.98</td>
</tr>
<tr>
<td>Geological/Geophysical Engineer</td>
<td>1.98</td>
</tr>
<tr>
<td>Jazz/Jazz Studies</td>
<td>1.98</td>
</tr>
<tr>
<td>Music Technology</td>
<td>1.98</td>
</tr>
<tr>
<td>Music, Other</td>
<td>1.97</td>
</tr>
<tr>
<td>Geological and Earth Sciences</td>
<td>1.97</td>
</tr>
<tr>
<td>Latin Language and Literature</td>
<td>1.97</td>
</tr>
<tr>
<td>Music Teacher Education</td>
<td>1.97</td>
</tr>
<tr>
<td>Sociology</td>
<td>1.57</td>
</tr>
<tr>
<td>Spanish Language and Literature</td>
<td>1.57</td>
</tr>
<tr>
<td>Multi-/Interdisciplinary Studies</td>
<td>1.56</td>
</tr>
<tr>
<td>Entrepreneurial and Small Business</td>
<td>1.56</td>
</tr>
<tr>
<td>General Studies</td>
<td>1.50</td>
</tr>
<tr>
<td>English Language and Literature</td>
<td>1.49</td>
</tr>
<tr>
<td>History, General</td>
<td>1.47</td>
</tr>
<tr>
<td>Engineering, General</td>
<td>1.44</td>
</tr>
<tr>
<td>Neuroscience</td>
<td>1.39</td>
</tr>
<tr>
<td>Biology/Biological Sciences, General</td>
<td>1.38</td>
</tr>
<tr>
<td>Physiological Psychology/Psychology</td>
<td>1.32</td>
</tr>
<tr>
<td>Economics, General</td>
<td>1.31</td>
</tr>
<tr>
<td>Political Science and Government</td>
<td>1.28</td>
</tr>
<tr>
<td>Experimental Psychology</td>
<td>1.22</td>
</tr>
<tr>
<td>Engineering, Other</td>
<td>1.17</td>
</tr>
</tbody>
</table>

Connected in major

Musical Theater

Joint English/MT

Philosophy

Philosophy Premed

General Studies

Transfer GS
Addressing the long term

• To fully understand our impact, we should know
  – More about where students come from
  – More about what they do while with us
  – More about what they do after they leave us

• Exploring life-long impact of education: ISR, IRIS, connections to employment & beyond
  – Research questions, to take place in enclave
The long term

What happens after college...

Analysis:
Modeling, extraction of meaning, exploration of how background and structures affect outcomes
Three examples:

1. Do living-learning programs work?
2. Are placement exams used well?
3. Are our classrooms equitable?
#1: Do living-learning programs work?

Example: Health Sciences Scholars Program at Michigan

Live-in learning community
Longitudinal data collected from 2004-2010
One of several at Michigan, thousands of students

RQ: How does participation in the learning community impact student outcomes?

Our interest was looking especially at two risk groups:
• First in family to attend university
• Ethnic minorities

Background Research: Literature generally filled with “feel good” stories about learning communities, we wanted more quantified results.

Photo Credit: UM HSSP Program

Brooks, Chavez, Tritz, and Teasley, Learning Analytics & Knowledge, 2015

AAC&U Diversity, Helen Morgan, Jennifer Matlby, Christopher Brooks, March 2015, California
Our Method: Matched Samples

Step 1: Identify potential sources of bias in your population and determine how to measure them, such as:

- Achievement level of learners from student information systems
- Interest in the treatment through:
  - Applications to participate
  - Surveys on student outcome interests (e.g. CIRP)
  - Sign-ons to tech-based treatments
- Demographics and other proxies for latent characteristics (e.g. first in family to attend school)

Step 2: Find the best matches for your sample in the general population
- Linear assignment problem (min-cost, max-flow), Hungarian method

Data elements matched on:
- ACT score (or converted SAT)
- Academic school enrolled in
- At risk support program enrollment (2)
- Year enrolled
- Honors enrollment status
- Credit hours achieved
- Citizenship
- Ethnic group (self-reported)
- Sex
- Previous research program experience
- Family income (bands)
- First in family to go to college
- Self-identification of being interested in pre-health programs

Brooks, Chavez, Tritz, and Teasley, Learning Analytics & Knowledge, 2015
### Example results

All students in LC:

No statistical significance 😊

Underrepresented minorities:

<table>
<thead>
<tr>
<th></th>
<th>LC (n=127)</th>
<th>Matches (n=127)</th>
<th>Paired t-test (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergrad in Science</td>
<td>67</td>
<td>47</td>
<td>0.006</td>
</tr>
<tr>
<td>MSc Completed</td>
<td>15</td>
<td>5</td>
<td>0.024</td>
</tr>
</tbody>
</table>

First in Family:

<table>
<thead>
<tr>
<th></th>
<th>LC (n=144)</th>
<th>Matches (n=144)</th>
<th>Paired t-test (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergrad in Science</td>
<td>86</td>
<td>65</td>
<td>0.009</td>
</tr>
<tr>
<td>Grad. Degree in Health Field</td>
<td>12</td>
<td>4</td>
<td>0.032</td>
</tr>
<tr>
<td>Science PhD/MD/DO/DDS</td>
<td>12</td>
<td>4</td>
<td>0.032</td>
</tr>
</tbody>
</table>

HSSP significantly increased the likelihood of BS and advanced degrees for underrepresented and first-generation students.
Traditional 2:2 Introductory Chemistry Curriculum Model:

- 2 Semesters General Chemistry
- 2 Semesters Organic Chemistry

Michigan 1:2:1 Introductory Chemistry Curriculum Model:

- AP scores of 3, 4, or 5
- Placement exams
  - above 70th percentile chemistry and 30th percentile math
  - below 70th percentile chemistry and/or below 30th percentile math

- Macroscopic Investigations and Reaction Principles (general chemistry) with Laboratory
  - Chemistry 130
- Structure and Reactivity I and II (organic chemistry) with Laboratory
  - Chemistry 210 & 215
- Physical Chemistry Principles and Applications
  - Chemistry 230
Regression Discontinuity

#3: Are our courses equitable? Gender and STEM performance

A Longitudinal Study of Engineering Student Performance and Retention. III. Gender Differences in Student Performance and Attitudes

Richard M. Felder
Department of Chemical Engineering
Gary N. Felder
Department of Chemical Engineering
Meredith Mauney
Department of Statistics
Charles E. Hamrin, Jr.
Department of Chemical Engineering

The More Things Change, the More They Stay the Same? Prior Achievement Fails to Explain Gender Inequality in Entry Into STEM College Majors Over Time
Catherine Riegel-Crumb, Barbara King, Eric Grodsky and Chandra Muller
Am Educ Res J 2012 49: 1048 originally published online 24 February 2012
DOI: 10.3102/0002831211435229
The online version of this article can be found at: http://aer.sagepub.com/content/49/6/1048

The Determinants of Success in University Introductory Economics Courses
Gordon Anderson, Dwayne Benjamin, and

Factors that affect the physical science career interest of female students: Testing five common hypotheses
Zahra Hazari,1,2,3,* Geoff Potvin,2,3 Robynne M. Lock,3 Florin Lung,4 Gerhard Sonnert,5 and Philip M. Sadler5

1Department of Teaching and Learning, Florida International University, Miami, Florida, USA
2Department of Physics, Florida International University, Miami, Florida, USA
3Department of Engineering and Science Education, Clemson University, Clemson, South Carolina, USA
4Department of Physics and Astronomy, Mississippi State University, Mississippi State, Mississippi, USA
5Science Education Department, Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA
(Received 8 August 2012; revised manuscript received 28 June 2013; published 22 October 2013)
Gendered performance differences: different outcomes for students with the same background.

Number male = 14199
Number female = 5810

\[ <\text{GPA} - \text{Grade}> \text{ Male } = 0.32 \]
\[ <\text{GPA} - \text{Grade}> \text{ Female } = 0.59 \]
GPD = 0.27
Data from 2000 – 2012 for all ‘giant’ classes, with average enrollments over 400

- Females favored
- Males favored
- Grade bonus
- Grade penalty

Large courses using timed exams for the majority of their grading.
Data from 2000 – 2012 for all large introductory STEM lecture and lab courses

Females favored

These performance differences remain when we account for all measures of background & preparation. The same patterns are observed on five CIC campuses.

Lab courses

Lecture courses

Unexplained performance differences like this are signs of classroom inequity. We must act to address them.
Action:
Decision making, story telling, motivating change
Putting data to work

• The next frontier: use technology to put data in people’s hands. Doing this, we support decision making, trigger personal connections, motivate action, and guide behavior change.

• **DIG**: the UM Digital Innovation Greenhouse has been established to take good ideas developed on campus from innovation to infrastructure, personalizing education at scale.
DIG was born to solve a recurring problem

• Faculty innovators create IT tools which make education > personal, engaged, and life-long.
• Research teams test them, demonstrating effectiveness: they’re ready to spread!
• Then they hit the entrepreneurial “valley of death” between innovation and infrastructure
• These innovations need a nurturing place to mature and spread, both on-campus and off
We DIG within DEI

Learning Experiences should be Personalized, Engaged, and Lifelong

DEI partnerships are Individualized, Creative, and Collaborative

Make an impact, Give Today

It takes a community.

Your Email

Join Newsletter
How we DIG

The University Community

Innovators & pioneering adopters

Communities of practice: faculty, students, staff

DIG team of Developers, U/X Designers, Behavioral Scientists

University IT: support at scale

Startups

UNIZIN
Since last May, DIG is real: a place, a team of innovators
Students are our best creative engine: Fellows, Design Jams & Hackathons!
DIG projects:
A rapidly growing portfolio
Providing information to individuals

1. Learning about classes and more: ART 2.0
2. Supporting advising: Student Explorer
ART 2.0 – information to all

Eventually, course cards will be joined by reports on courses of study (majors and minors) and people (students, faculty)....
Student Explorer: supporting advisors
Acting directly

ECoach: computer tailored electronic coaching for equity and student success
E²Coach: computer tailored communication for student support and help with behavior change

Welcome back, Zoe.

You made it through the first statistics exam.

You scored a 63 out of 75 points or 84.0%, which corresponds to a letter grade of a B+.

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

What can you do now?

Keep up the good work!
ECoach: tailored communication

• Built on CHCR digital health coaching heritage
• Use rich real-time info about students to tailor feedback, advice, encouragement
• Tailoring on both what to say and also how to say it: testimonials from peers, behavior change experts

• Used since 2012 by 10,000+ students
  – usage creates clear impacts on performance
• Preparing to reach every student in fall, designing interventions to change the future for students!
• Key tool for humane personalization: speak and connect
Data and the Future of Education

• We will have tools which expose information enabling everyone on campus to **learn from the experience** of all
• We will be able to explore and represent what students and faculty do in richer, multidimensional ways, **encouraging better experiences**

• **UM is a giant laboratory for higher education!**
• Teaching and learning will be evidence-based, consistently assessed, and continually refined
• We hope Michigan will become a center for using data to understand what higher education does for students and the nation.

**This transformation will be driven by drawing inference from data in smart ways, and owning the responsibility to act on it**
The sheer amount of data that we generate in the course of our lives is growing exponentially as technology plays a larger and larger part in what we do every day. Nowhere is this fact more important than how information technology has become part of the basic infrastructure for education – in formal and informal settings, playing a role whether we are face-to-face with others or interacting solely online. The “data exhaust” that is generated by the systems used today to support how we teach and how we learn provides an unprecedented opportunity to better understand and support learning, and to question the impact of these technologies on individuals, institutions and culture. This effort, however, is necessarily interdisciplinary and requires the use of a diverse methodological toolset.