From Theory to Practice: Leading the Way with Learning Data Principles

Jenn Stringer, M.L.I.S.
UC Berkeley, CNI Spring 2017

IMS Global: **Samantha Birk,** IMS Global, **John Fritz**, University of Maryland, Baltimore County, **Oliver Heyer**, University of California at Berkeley, **Virginia Lacefield**, University of Kentucky, **Virginia Lacefield**, University of Kentucky, **Adam Recktenwald**, University of Kentucky, **Marianne Schroeder**, University of British Columbia

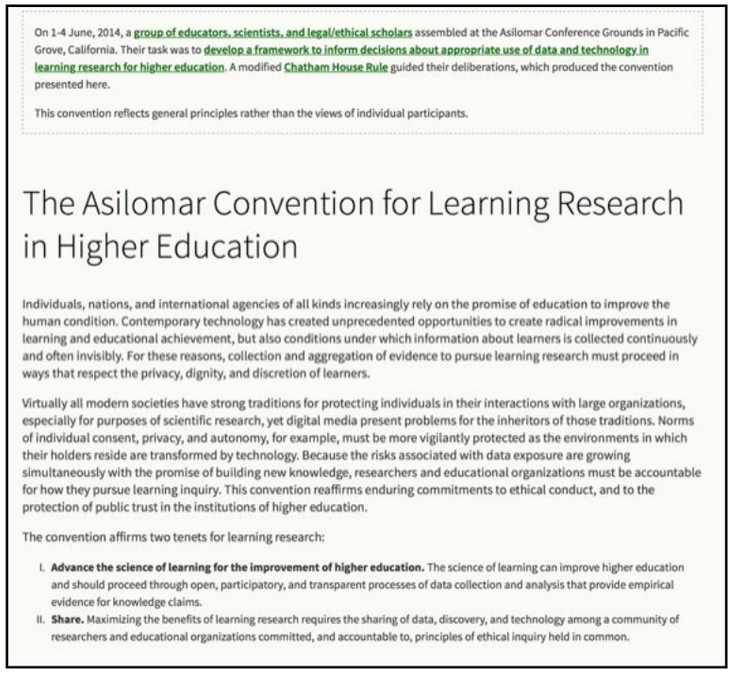
***University of California*: Mary Ellen Kreher,** University of California Office of the President, **Jim Phillips** University of California at Santa Cruz, **James Williamson**, University of California at Los Angeles

Goals of the presentation

1. Set the context for why these are important and why now
2. Give an overview of the landscape of learning data and analytics
3. Present the IMS principles
4. Present UC Principles and Practices
5. Get feedback in particular:
    a. ownership
    b. efficacy

Background and Other Efforts

- Asilomar Convention for Learning Research in Higher Education - http://asilomar-highered.info/
- Asilomar II: Student Data and Records in the Digital Era - https://sites.stanford.edu/asilomar/
- Responsible Use of Student Data in Higher Education - http://ru.stanford.edu/ (Stanford CAROL & Ithaka S+R_
- DELICATE Framework - http://www.laceproject.eu/blog/ethics-privacy-in-learning-analytics-a-delicate-issue/ (Learning Analytics Community Exchange)



On 1-4 June, 2014, a group of educators, scientists, and legal/ethical scholars assembled at the Asilomar Conference Grounds in Pacific Grove, California. Their task was to develop a framework to inform decisions about appropriate use of data and technology in learning research for higher education. A modified Chatham House Rule guided their deliberations, which produced the convention presented here.

This convention reflects general principles rather than the views of individual participants.

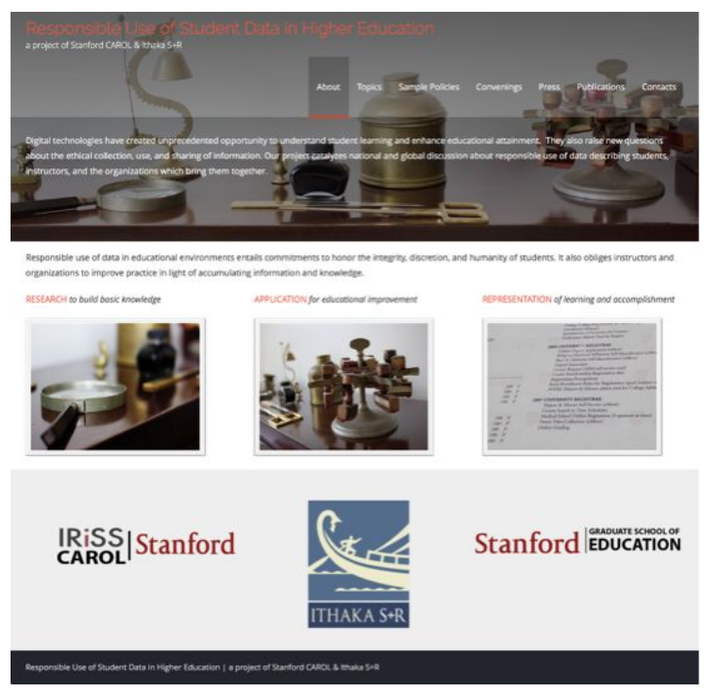
The Asilomar Convention for Learning Research in Higher Education

Individuals, nations, and international agencies of all kinds increasingly rely on the promise of education to improve the human condition. Contemporary technology has created unprecedented opportunities to create radical improvements in learning and educational achievement, but also conditions under which information about learners is collected continuously and often invisibly. For these reasons, collection and aggregation of evidence to pursue learning research must proceed in ways that respect the privacy, dignity, and discretion of learners.

Virtually all modern societies have strong traditions for protecting individuals in their interactions with large organizations, especially for purposes of scientific research, yet digital media present problems for the inheritors of those traditions. Norms of individual consent, privacy, and autonomy, for example, must be more vigilantly protected as the environments in which their holders reside are transformed by technology. Because the risks associated with data exposure are growing simultaneously with the promise of building new knowledge, researchers and educational organizations must be accountable for how they pursue learning inquiry. This convention reaffirms enduring commitments to ethical conduct, and to the protection of public trust in the institutions of higher education.

The convention affirms two tenets for learning research:

I. **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.
II. **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.



Definition: Learning Analytics

"The selection, capture and processing of data that will be helpful for students and instructors at the course or individual level." [1]

"Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs."[2]

SOLAR
SOCIETY for LEARNING ANALYTICS RESEARCH

LAK'11

LAK 17

The 7th International

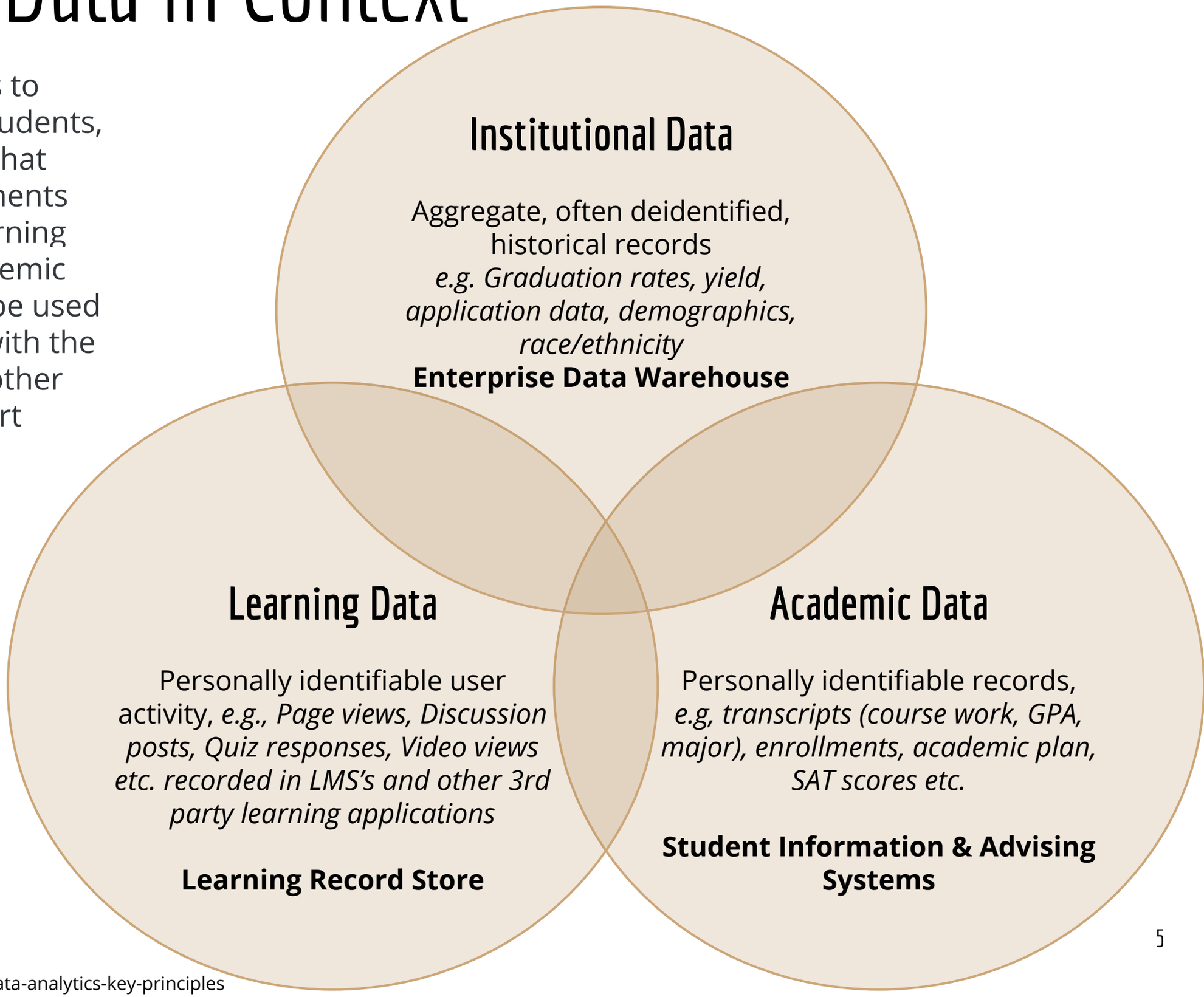
Learning Analytics & Knowledge Conference

Simon Fraser University, Vancouver, BC, Canada
March 13 - 17, 2017 **#LAK17**

[1] Learning Analytics: The Definitions, the Processes, and the Potential, Tanya Elias, 2011
[2] 1st International Conference of Learning Analytics & Knowledge, Banff, Alberta 2011

Learning Data in Context

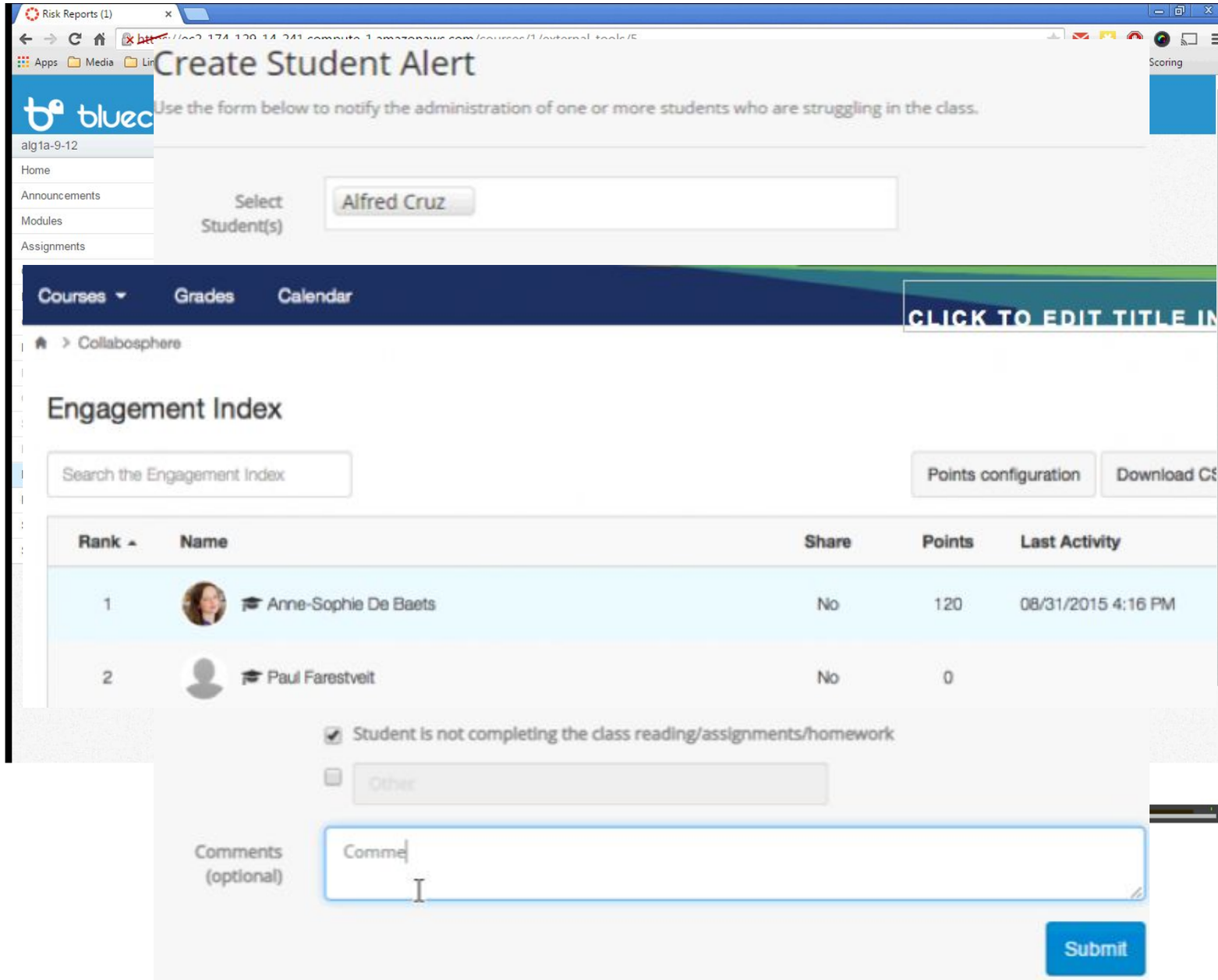
"Learning data refers to data generated by students, faculty, and/or staff that relates to and documents the teaching and learning experience and academic achievement. It can be used alone or combined with the student record and other data points to support student success and research." [3]



[3] https://www.imsglobal.org/learning-data-analytics-key-principles

Why do we care?

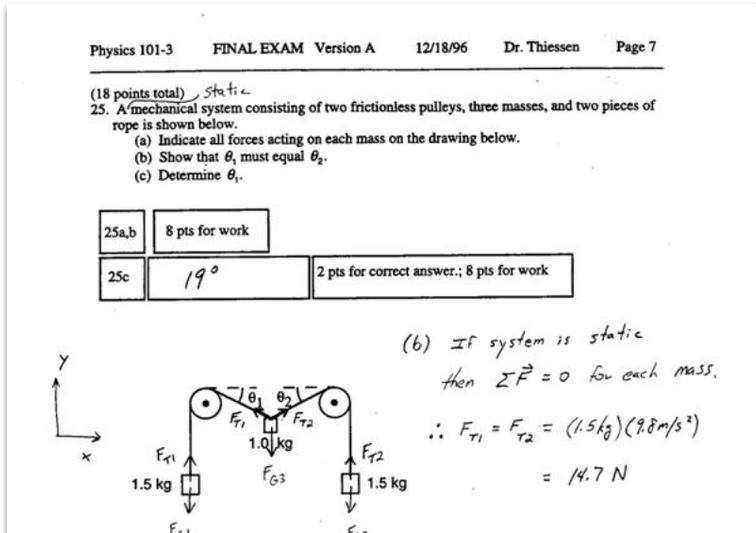
- increase ability to make institutional decisions
- impact student outcomes
- empower students to make changes to their behavior that positively affects their learning
- enable faculty to support students and make changes to their courses based on data
- support faculty teaching and pedagogy
- support educational research



Someone is collecting A LOT of data!

- What: “ Clickstream Data” “Logfile Data”
- Who: Vendors, the institution, libraries, publishers, other third-parties

gradescope



Physics 101-3 FINAL EXAM Version A 12/18/96 Dr. Thiessen Page 7

(18 points total) static

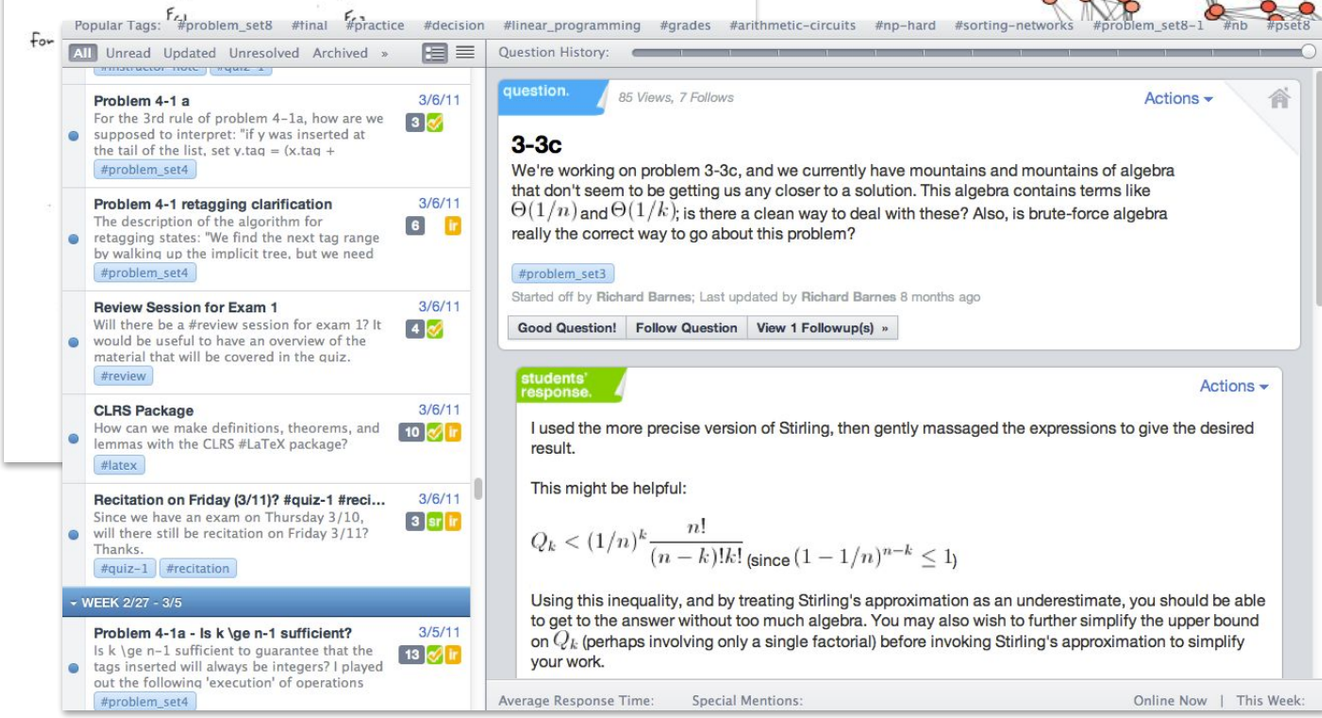
25. A mechanical system consisting of two frictionless pulleys, three masses, and two pieces of rope is shown below.
    (a) Indicate all forces acting on each mass on the drawing below.
    (b) Show that θ_1 must equal θ_2.
    (c) Determine θ_1.

25a,b	8 pts for work	
25c	$19°$	2 pts for correct answer.; 8 pts for work

(b) If system is static then $\Sigma \vec{F} = 0$ for each mass.

$\therefore F_{T1} = F_{T2} = (1.5\,kg)(9.8\,m/s^2)$

$= 14.7\ N$



Popular Tags: #problem_set8 #final #practice #decision #linear_programming #grades #arithmetic-circuits #np-hard #sorting-networks #problem_set8-1 #nb #pset8

All Unread Updated Unresolved Archived »

Problem 4-1 a 3/6/11
For the 3rd rule of problem 4-1a, how are we supposed to interpret: "if y was inserted at the tail of the list, set y.tag = (x.tag + #problem_set4

Problem 4-1 retagging clarification 3/6/11
The description of the algorithm for retagging states: "We find the next tag range by walking up the implicit tree, but we need #problem_set4

Review Session for Exam 1 3/6/11
Will there be a #review session for exam 1? It would be useful to have an overview of the material that will be covered in the quiz. #review

CLRS Package 3/6/11
How can we make definitions, theorems, and lemmas with the CLRS #LaTeX package? #latex

Recitation on Friday (3/11)? #quiz-1 #reci... 3/6/11
Since we have an exam on Thursday 3/10, will there still be recitation on Friday 3/11? Thanks. #quiz-1 #recitation

WEEK 2/27 - 3/5

Problem 4-1a - Is k \ge n-1 sufficient? 3/5/11
Is k \ge n-1 sufficient to guarantee that the tags inserted will always be integers? I played out the following 'execution' of operations #problem_set4

Question History:

question. 85 Views, 7 Follows Actions

3-3c

We're working on problem 3-3c, and we currently have mountains and mountains of algebra that don't seem to be getting us any closer to a solution. This algebra contains terms like $\Theta(1/n)$ and $\Theta(1/k)$; is there a clean way to deal with these? Also, is brute-force algebra really the correct way to go about this problem?

#problem_set3

Started off by Richard Barnes; Last updated by Richard Barnes 8 months ago

Good Question! Follow Question View 1 Followup(s) »

students' response. Actions

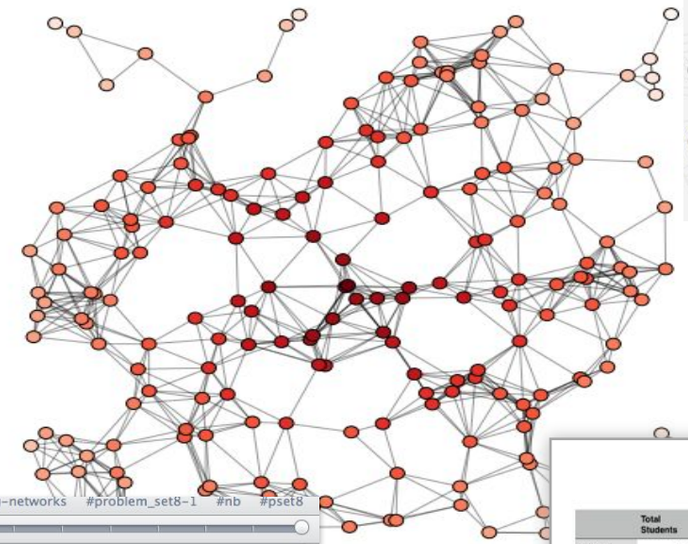
I used the more precise version of Stirling, then gently massaged the expressions to give the desired result.

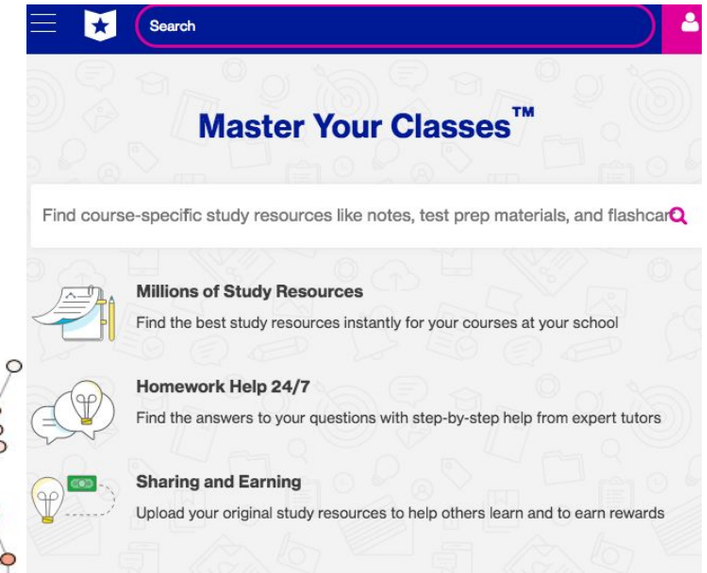
This might be helpful:

$$Q_k < (1/n)^k \frac{n!}{(n-k)!k!} \text{ (since } (1-1/n)^{n-k} \leq 1)$$

Using this inequality, and by treating Stirling's approximation as an underestimate, you should be able to get to the answer without too much algebra. You may also wish to further simplify the upper bound on Q_k (perhaps involving only a single factorial) before invoking Stirling's approximation to simplify your work.

Average Response Time: Special Mentions: Online Now | This Week:





Search

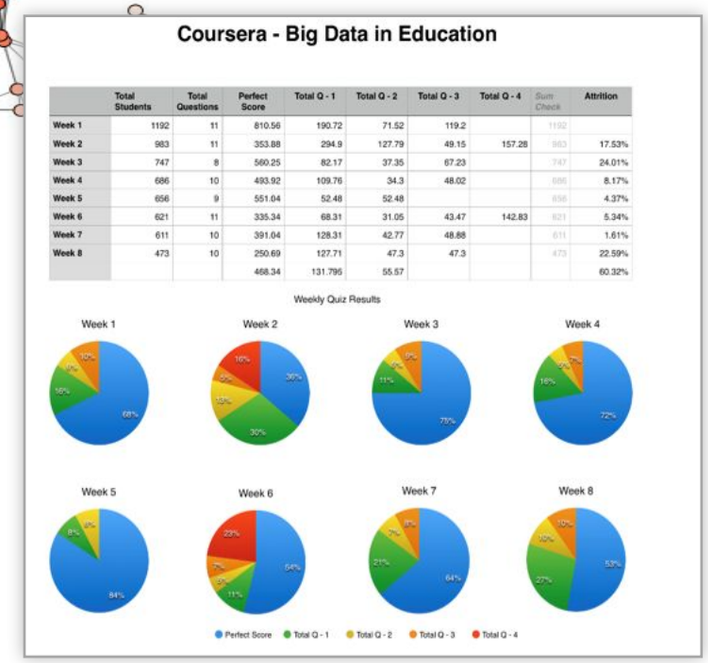
Master Your Classes™

Find course-specific study resources like notes, test prep materials, and flashcar

Millions of Study Resources
Find the best study resources instantly for your courses at your school

Homework Help 24/7
Find the answers to your questions with step-by-step help from expert tutors

Sharing and Earning
Upload your original study resources to help others learn and to earn rewards



Coursera - Big Data in Education

	Total Students	Total Questions	Perfect Score	Total Q - 1	Total Q - 2	Total Q - 3	Total Q - 4	Sum Check	Attrition
Week 1	1192	11	810.56	190.72	71.52	119.2		1192	
Week 2	983	11	353.88	294.9	127.79	49.15	157.28	983	17.53%
Week 3	747	8	560.25	82.17	37.35	67.23		747	24.01%
Week 4	686	10	493.92	109.76	34.3	48.02		686	8.17%
Week 5	656	9	551.04	52.48	52.48			656	4.37%
Week 6	621	11	335.34	68.31	31.05	43.47	142.83	621	5.34%
Week 7	611	10	391.04	128.31	42.77	48.88		611	1.61%
Week 8	473	10	250.69	127.71	47.3	47.3		473	22.59%
			468.34	131.795	55.57				60.32%

Weekly Quiz Results

Week 1 Week 2 Week 3 Week 4 Week 5 Week 6 Week 7 Week 8

Perfect Score Total Q - 1 Total Q - 2 Total Q - 3 Total Q - 4

piazza

vt Home

VoiceThread fills the social presence gap found in online learning interactions

Learning Data

- **“Old Days”**
  - local hosting meant local logs
  - ad-hoc reporting mainly for systems issues
- **“Cloud SaaS”**
  - logs not local and not accessible
  - vendors use to improve systems and troubleshoot issues
- **Contracts**
  - if we have them -- not always specific about ownership of this kind of data
  - they do have security and privacy language
- **No Contract?**
  - goodluck! - at the whim of the third party provider no security/privacy guarantees

Learning Data

- Just the **data** please
  - even if we have “ownership” we need “access”
  - logfile ‘dumps” for historic reporting (nightly, monthly, etc.)
  - realtime for early warning systems, advising, student alerts, etc.

- **Standards**
  - **LTI** - enables interoperability
  - **Caliper and xAPI** - defines learner activity to enable analysis across systems

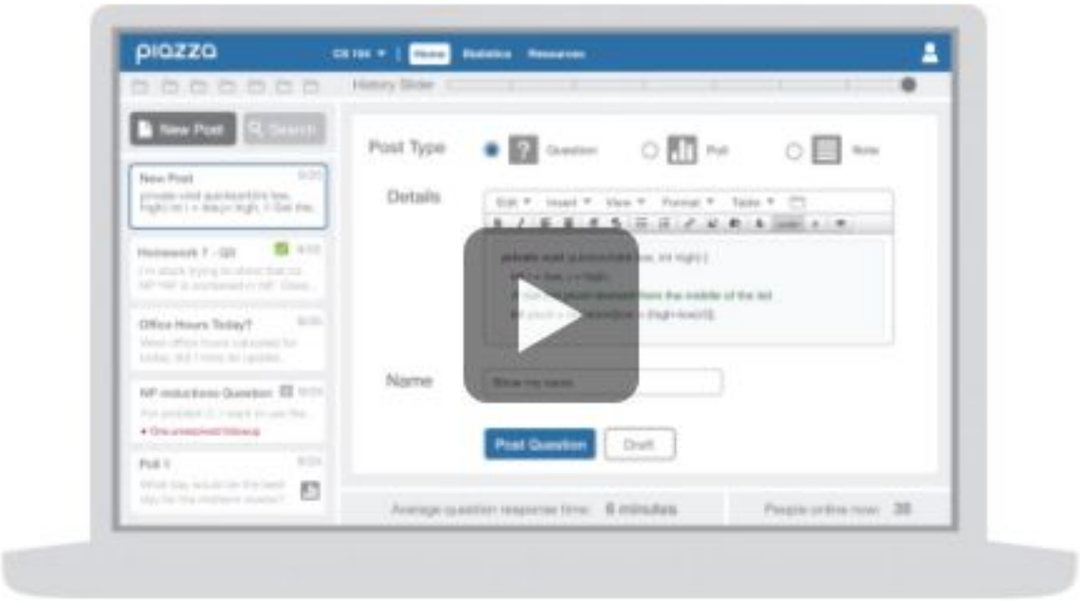
Case Studies

- **With a contract**
  - LMS - wanted to charge users extra for our data
- **Without a contract**
  - a free platform for instructors to efficiently manage class Q&A. Students can post questions and collaborate to edit responses to these questions. Instructors can also answer questions, endorse student answers, and edit or delete any posted content.
  - currently collects a TON of this data and the institution can't get to it
  - revenue model - match students to potential employers
    - the employers pay a fee

piazza Product In Professors' Words Support About Us Piazza Careers Product Sign Up Login

The incredibly easy, completely free Q&A platform

Save time and help students learn using the power of community



- Wiki style format enables collaboration in a single space
- Features LaTeX editor, highlighted syntax and code blocking
- Questions and posts needing immediate action are highlighted
- Instructors endorse answers to keep the class on track
- Anonymous posting encourages every student to participate
- Highly customizable online polls
- Integrates with every major LMS

Students Get Started | Professors and TAs Get Started | View a Real Class

Learn more about how Piazza complies with FERPA

Over 50,000 professors in 1,500 schools and 90 countries have chosen Piazza

Click on an instructor to see why

“If it is free, you are not the customer. You are the product.”[5]

[5]http://blogs.harvard.edu/futureoftheinternet/2012/03/21/meme-patrol-when-something-online-is-free-youre-not-the-customer-youre-the-product/

We're at 1,000 schools in 68 countries

Click a school to see classes and professor stories

Stanford · Cal · MIT Massachusetts Institute of Technology · Princeton University · Penn · Carnegie Mellon

Georgia Tech · Illinois · Cornell University · Michigan · University of Waterloo · The University of Texas at Austin

Berkeley Classes Using Piazza

473 Instructors and 14797 Students at Berkeley use Piazza in 1285 Classes

CHEM 135: Chemical Biology
Ming Hammond
100% of students participated
view class report

COGSCI 131: Computational Mo...
Kevin Canini, Zeyu Li
The average response time was **11 minutes**
view class report

CS 61C: Great Ideas in Computer...
Michael Franklin, Dan Garcia
There were **16429 contributions** in total
view class report

EE 122: Introduction to Commun...
Yahel Ben-David, Scott Shenker
There were **3658 contributions** in total
view class report

ARCH 205 & 249: Studio One
Nicholas de Monchaux
100% of students participated
view class report

CS 61A: The Structure and Interp...
Richard Lan, Tom Magrino
There were **16030 contributions** in total
view class report

Finish setting up your Piazza account:

Account Information (required)

Is this your preferred email address: **akstring@ucsc.edu**

☐ No, use another email

Contact us at team@piazza.com with any questions.

Full Name

Choose Password

Confirm Password

Academic Information (required)

What degree are you currently pursuing?

Graduate Program: Select current program...

Major: Enter current major...

Anticipated Completion: Month | Year

☐ I have two majors

☐ I'm not pursuing a degree

This information will be used for collaborative features on Piazza. We will never share your information without your permission.

Where Piazza is an aid in your class, Piazza Careers is an aid in your career. Get discovered by companies instantly.

☑ Career opportunities are relevant to me

☐ I've read and accept the terms of service

Continue to Piazza

A little better……

piazza

Join 1.5M students on Piazza Careers

Learn about Piazza Careers

Piazza Careers put me on Yelp's radar, something that I was not able to accomplish through applying directly to the company online.

K. W
UMD '14 Grad

Through Piazza Careers, students have been hired by:

- Airbnb
- Quora
- Pinterest
- Facebook
- Google
- Microsoft
- Palantir
- Dropbox
- Morgan Stanley
- Citadel
- Two Sigma
- Verizon

+174 more

- I'm open to hearing from and connecting with companies and alumni (employers can view your Careers profile)
- I don't need any help getting the most fulfilling and rewarding career opportunities at this time

Control what information is shared

Continue

Piazza Careers

Opt in

Save Careers Preferences

Over 30,000 professors have chosen Piazza

Click on an instructor to see why

Jennifer Schwartz
Stanford
Chemistry

"Many of the students who ask questions on Piazza just wouldn't get the opportunity to ask them otherwise."

Ron Lee
UC Berkeley
Economics

"I think it impoverishes the learning environment if students don't talk to other students. Piazza makes it very easy to get a discussion going."

Slobodan Simic
San Jose State
Mathematics

"Nowadays, colleagues tell me that their students always ask, 'Are we going to use Piazza?' It's sort of an expected thing now."

Paul Hegarty
Stanford
Computer Science

"In the last few years, Piazza has replaced my entire website."

We're at 1,000 schools in 68 countries

Click a school to see classes and professor stories

Stanford | Cal | MIT Massachusetts Institute of Technology | Princeton University | Penn | Carnegie Mellon

Georgia Tech | Illinois | Cornell University | Michigan | University of Waterloo | The University of Texas at Austin

Find your school

A little better……

Over 50,000 professors in 1,500 schools and 90 countries have chosen Piazza

Click on an instructor to see why

Jennifer Schwartz
Chemistry

"Many of the students who ask questions on Piazza just wouldn't get the opportunity to ask them otherwise."

Ron Lee
Economics

"I think it impoverishes the learning environment if students don't talk to other students. Piazza makes it very easy to get a discussion going."

Slobodan Simic
Mathematics

"Nowadays, colleagues tell me that their students always ask, 'Are we going to use Piazza?' It's sort of an expected thing now."

Paul Hegarty
Computer Science

"In the last few years, Piazza has replaced my entire website."

Libraries and Learning Data/Analytics

"Though few academic libraries are encountering it just yet, it is only a matter of time before higher education institutions integrate learning analytics at every level of the organization."[6]

"Learning analytics initiatives pose a myriad of ethical questions. For example, are institutions who possess learning data required to act on it? Might learning data be used to "profile" students?"[7]

"At institutions that have committed to a learning analytics future, librarians can also ask questions to clarify the library's role as well as advocate for library inclusion in learning analytics processes."[7]

[6] Bell, Steven. Keeping up with … Learning Analytics. ACRL Blog. http://www.ala.org/acrl/publications/keeping_up_with/learning_analytics
[7] Oakleaf, Megan. "Getting Ready and Getting Started: Academic Librarian Involvement in Institutional Learning Analytics Initiatives." *Journal of Academic Librarianship*. 42(4). 2016.

IMS Global Learning Data & Analytics Key Principles

https://www.imsglobal.org/learning-data-analytics-key-principles

1. **Ownership**: Faculty, staff, and students generate and own their learning data. As governed by institutional policies, individuals, being owners of the data they generate, have the right to access, port, and control the disposition of their data stored by the institution, its service providers, and their affiliated partners.
2. **Stewardship**: As stewards of learning data, institutions should have a data governance plan and governance policies that protect the data and the interests of its owners. These should transcend, but encompass, existing protocols, such as IRB.
3. **Governance**: Learning data use and retention will be governed by institutional policies, and faculty and students retain the right of data access and retrieval.
4. **Access**: Learning data, whether generated locally or in a vendor-supplied system, is strategic to an institution's business and mission and must be available to the institution.
5. **Interoperability**: The collection, use, and access to learning data requires institutional and supplier collaboration, which is dependent upon interoperability standards, protocols, data formats, and content to achieve institutions goals.
6. **Efficacy**: Learning data collection, use, and computational transformation is aimed at student and instructor success and instructional concerns through prescriptive, descriptive, or predictive methodologies.
7. **Security & Privacy**: Individuals' security and privacy relating to collecting, using, and algorithmically transforming learning data is fundamental and must not be treated as optional. It must also be balanced with the effective use of the data.
8. **Transparency**: Individuals have the right to understand the specific reasons, methods, and purposes for which their learning data is collected, used, and transformed. This includes any learning data being shared with third--party service providers and other institutional affiliates or partners. Individuals also have the right to know how their data is transformed and/or used thru processes such as summative or algorithmic modifications, particular outputs, and visualizations.

IMS Global Learning Data & Analytics Key Principles

https://www.imsglobal.org/learning-data-analytics-key-principles

1. **Ownership**: Faculty, staff, and students generate and own their learning data. As governed by institutional policies, individuals, being owners of the data they generate, have the right to access, port, and control the disposition of their data stored by the institution, its service providers, and their affiliated partners.
2. **Stewardship**: As stewards of learning data, institutions should have a data governance plan and governance policies that protect the data and the interests of its owners. These should transcend, but encompass, existing protocols, such as IRB.
3. **Governance**: Learning data use and retention will be governed by institutional policies, and faculty and students retain the right of data access and retrieval.
4. **Access**: Learning data, whether generated locally or in a vendor-supplied system, is strategic to an institution's business and mission and must be available to the institution.
5. **Interoperability**: The collection, use, and access to learning data requires institutional and supplier collaboration, which is dependent upon interoperability standards, protocols, data formats, and content to achieve institutions goals.
6. **Efficacy**: Learning data collection, use, and computational transformation is aimed at student and instructor success and instructional concerns through prescriptive, descriptive, or predictive methodologies.
7. **Security & Privacy**: Individuals' security and privacy relating to collecting, using, and algorithmically transforming learning data is fundamental and must not be treated as optional. It must also be balanced with the effective use of the data.
8. **Transparency**: Individuals have the right to understand the specific reasons, methods, and purposes for which their learning data is collected, used, and transformed. This includes any learning data being shared with third--party service providers and other institutional affiliates or partners. Individuals also have the right to know how their data is transformed and/or used thru processes such as summative or algorithmic modifications, particular outputs, and visualizations.

IMS Global Learning Data & Analytics Key Principles

https://www.imsglobal.org/learning-data-analytics-key-principles

1. **Ownership**: Faculty, staff, and students generate and own their learning data. As governed by institutional policies, individuals, being owners of the data they generate, have the right to access, port, and control the disposition of their data stored by the institution, its service providers, and their affiliated partners.
2. **Stewardship**: As stewards of learning data, institutions should have a data governance plan and governance policies that protect the data and the interests of its owners. These should transcend, but encompass, existing protocols, such as IRB.
3. **Governance**: Learning data use and retention will be governed by institutional policies, and faculty and students retain the right of data access and retrieval.
4. **Access**: Learning data, whether generated locally or in a vendor-supplied system, is strategic to an institution's business and mission and must be available to the institution.
5. **Interoperability**: The collection, use, and access to learning data requires institutional and supplier collaboration, which is dependent upon interoperability standards, protocols, data formats, and content to achieve institutions goals.
6. **Efficacy**: Learning data collection, use, and computational transformation is aimed at student and instructor success and instructional concerns through prescriptive, descriptive, or predictive methodologies.
7. **Security & Privacy**: Individuals' security and privacy relating to collecting, using, and algorithmically transforming learning data is fundamental and must not be treated as optional. It must also be balanced with the effective use of the data.
8. **Transparency**: Individuals have the right to understand the specific reasons, methods, and purposes for which their learning data is collected, used, and transformed. This includes any learning data being shared with third--party service providers and other institutional affiliates or partners. Individuals also have the right to know how their data is transformed and/or used thru processes such as summative or algorithmic modifications, particular outputs, and visualizations.

The University of California should have a say in how suppliers collect, use, and manage our data.

University of California: Learning Data Privacy Principles (DRAFT)

1. **Ownership**: The University of California (UC), its faculty, and students retain ownership of the data and subsequent computational transformations of the data they produce. Individual data owners have the right to determine how their data will be used. The UC acts as stewards of data on behalf of its faculty and students.
2. **Ethical Use**: Learning data collection, use, and computational transformation are governed by pedagogical and instructional concerns, with an aim toward student success through prescriptive, descriptive, or predictive methodologies. As with grades and other sensitive data, uses of learning analytics should be pursued on a “need to know” basis.
3. **Transparency**: Data owners have a right to understand the specific methods and purposes for which their data are collected, used and transformed, including what data are being transmitted to third-party service providers (and their affiliated partners) and the details of how algorithms are applied that shape summaries, particularly outputs and visualizations.
4. **Freedom of Expression**: Faculty and students retain the right to communicate and engage with each other in the learning process without the concern that their data will be mined for unintended or unknown purposes.
5. **Protection**: Stewards, on behalf of data owners, will ensure learning data are secure and protected in alignment with all federal, state, and university regulations regarding secure disposition.
6. **Access and Control**: Data owners have the right to access their data. Given that faculty and students own their learning data and share in its disposition, access to and ultimate authority and control of the data rests with the faculty and student owners, and the data stewards acting on their behalf. Data retention access and control practices will be governed under UC policies and supplier contractual agreements.

University of California:Learning Data Recommended Practices (DRAFT)

1. **Ownership**: Service providers will recognize learning data ownership and access as a right of the faculty and students.
2. **Usage Right:** Through a user's profile setting, service providers will enable users to control the use of their intellectual property. Thus, it will be the user's choice to grant terms such as, "a royalty-free, transferable, perpetual, irrevocable, non-exclusive, worldwide license to reproduce, modify, publish, publicly display, make derivative works."
3. **Opt-in:** Other than those data elements distinctly required for instruction, where appropriate, students will have a choice about the use of learning data collected by faculty and service providers in an "opt in" rather than "opt out" approach.
4. **Interoperable Data:** Service providers will provide learning data to the institution in recognized standard interoperability format(s) to minimize integration costs, support cross-platform and cross-application uses, and promote institutional and academic analysis and research.
5. **Data without Fees:** Service providers will not charge the faculty, students, or other university learning data stewards for the right of access, including the delivery of these data to the University.
6. **Transparency:** Service providers will inform the UC about the learning data they collect and how these data will be used, which in the course of an academic term shall be based on pedagogical concerns and curricular improvement.
7. **Service Provider Security:** All service provider platforms on which student learning data are stored will conform with UC and state mandated security procedures governing the reporting of unexpected incidents and corrections that may occur.
8. **Campus Security:** UC learning data stewards will ensure that all faculty and student data are stored securely in conformance with University data security policy. Learning data stewards will report any learning data security incidents as appropriate to faculty and students, and will provide information about their remedy.

Questions for discussion

- Ownership - students? instructors? institution? vendor? co-creation?

Questions for discussion

- Efficacy - students? instructors? institution? vendor? co-creation?

Discussion