

The Learning Analytics Workgroup (LAW) Report

Building the Field of Learning Analytics for Personalized Learning at Scale

By Roy Pea, D.Phil., Oxon. David Jacks Professor of Education and Learning Sciences at Stanford University.

Many students come to college unprepared, as evidenced by as many as 40% of 1st-year college students being placed in developmental courses, with fewer than 60% of students completing college within 6 years. With technological advances, there are new ways to provide support to students to improve college and career readiness. We fall behind when it comes to providing teachers, administrators, and families with the tools that they need to track progress and ensure all students achieve college and career success.

The Learning Analytics Workgroup (LAW) Project was initiated at the convening of a multisector group by the Bill and Melinda Gates Foundation on August 3, 2011, at the University of Chicago's Computation Institute. At this meeting began the discussion of how to best build capacity in the field for creating an innovative and sustainable ecosystem dedicated to advancing the state of learning data and learning analytics on behalf of all children's college and career readiness. The LAW Project has focused on making progress towards personalized learning at scale by building the field of learning analytics to support and advance initiatives and state-wide technology infrastructure to support new ecosystems of personalized learning and teaching.

A Conceptual Framework for Building the Field of Learning Analytics

In building the field of learning analytics, we are targeting the challenge of **advancing personalized learning at scale** for all learners with varying needs, skill levels, interests, dispositions, and abilities, arguing that continuously capturing, deriving meaning from, and acting on the production of vast volumes of data produced by learners engaged with digital tools is fundamental to personalized learning. Personalized learning provides the following opportunities: (a) improving educational performance, (b) facilitating cost efficiencies through educational productivity and organizational optimization, and (c) accelerating educational innovation.

The exponential growth of education data to be generated by digitally-enhanced learning environments requires education data scientists and people with diverse sense-making talent able to bring these data sets into productive interactive systems so that the various stakeholders can visualize learning at different levels of aggregation and use it to guide their decision-making.

Data science, as a distinct professional specialization, is in its infancy. What we are calling for is an even newer specialization, **Education Data Science**. People with skills in this area currently come from a wide array of academic disciplines that initially did not have to do with "data science," but all of which involved dealing with and managing enormous datasets: business intelligence, oceanography, meteorology, particle physics, bioinformatics, proteomics, nuclear physics, fluid and thermal dynamics, and satellite imagery data. What all of these people have in common today is their lack of affiliation to any school of education or education industry. Building the field of learning analytics will require leveraging the talents, skills, and other resources of (a) the academy, (b) nonprofits, (c) industry, (d) private foundations, and (e) governmental agencies.

Critical Questions for Understanding How to Build the Field of Learning Analytics

First, we consider it vital to **foreground the challenges of educators** in relation to the prospects of personalized learning. Based on insights from close to 800 teachers and administrators that had been interviewed across six states, we identified "opportunity areas" for technology innovation. Second, we recognize that different educational stakeholders will have different **success metrics for learners**. What outcomes should we care about in the development of personalized learning? Among the topics of special attention today are the so-called "noncognitive factors" in learning, such as academic persistence or perseverance and self-regulation. Third, a pre-eminent objective is **creating a model of the learner**. We have identified sources for building a learner model: metrics of student interaction during learning activities, social metrics, data concerning student mindset, past performance, learning media or genre preference, perseverance and persistence, administrative data, demographic information, temporal history, emotional state, and social network. Next, a broad set of topics is encompassed in the question of **how to establish a well-functioning, personalized-learning research infrastructure**. There are needs in the areas of data sharing, analysis and visualization tools, collaboration practices, data-management policies, and Institutional Review Board (IRB) reforms that will enable development of learning analytics as a field and implementation of personalized learning at scale. Finally, the transformations of educational systems into personalized learning, when actualized, will have

important consequences **for the preparation and professional development of teachers and educational leaders** of schools, districts, and states. Data literacy is an important skill for teachers, as making data-enhanced decisions in the classroom will depend upon the ability of a teacher to quickly make sense of data.

Articulating and Prioritizing New Tools, Approaches, Policies, Markets, and Programs

We present three "grand challenges" for research in learning analytics. We see these grand challenges as areas where early success could demonstrate the value of education data sciences. These challenges could be supported by competitions to create predictive learner models that get the greatest percentage of learners to competency in the shortest time at the lowest cost.

- **Grand Challenge 1: Learning progressions and the Common Core State Standards.** How can learning analytics help refine our understanding and practices involving learning progressions in digital learning environments for Common Core State Standards in math and language arts and the Next Generation Science Standards?
- **Grand Challenge 2: Standards-based assessments for digital learning.** How can we systematize the mapping of standards onto a bank of formal and informal assessment questions, with the goal of assessing content mastery and making recommendations for teacher practice in response to evaluation of learner competencies?
- **Grand Challenge 3: Creating multimodal learning analytics.** How can we expand education data to capture contextual features of learning environments that will allow assessment not only to focus on student demonstrations of knowledge on pre-designed assessment tasks but also to capture aspects of learners interacting with each other and their environment.

Determining Resources Needed to Address the Priorities

There is a need to develop training programs to develop capacity for learning analytics in education. First, we recommend a **faculty cross-training** approach. Bringing current education faculty—especially those who study psychometrics and educational measurement—into learning analytics is an important goal. Second, there is a need to develop **postdoctoral cross-training**. Graduates from computer science, data science, learning and educational sciences, computational statistics, computational linguistics, and other areas are all potential fits for learning analytics postdoctoral training. Third, there is a need for **degree and certificate options**. A range of certification options will need to be developed, including full degree programs at a variety of educational levels, certification programs, summer institutes, and courses (both traditional and online, as well as specialized seminars and survey courses). Finally, building the field of learning analytics will require **knowledge networking and online community building**. Recognizing and developing indicators of quality and establishing reputations for courses and programs will help establish a trusting relationship between stakeholders in learning analytics.

The Value Proposition for Different Stakeholders

The learning analytics community needs to step forward with a plan to address the challenges and opportunities discussed in this report. As we make our recommendations, we realize the importance for each stakeholder of communicating the value proposition in relation to problems of practice. **Institutions of higher education** could show leadership in addressing the emerging market demand for education data scientists trained in learning analytics by developing educational programs that contribute to human capacity building in this field. **Foundations and government agencies** need to provide Requests for Proposals for programs of research funding to which researchers, universities, and industry (when appropriate as partners) can respond. **University and nonprofit researchers** need to propose foundational research projects that solve key problems in the fields of learning analytics and education data science. **Industry** needs to offer compelling products and services that meet increasingly varied learner needs. **Educational systems** (states, districts) need to participate in co-design and co-study of the new learning and teaching ecosystems employing cyberinfrastructure to advance goals of college- and career-ready high school students.

Road Map to Implement the Field-Building Strategy and How to Evaluate Progress

To develop a road map for building the field of Learning Analytics, we began by brainstorming four essentials to grow learning analytics as a field: Human Capital, Research, Policy, and Tools. We also considered how we could measure progress in growing the field. Then we determined the necessary actions and identified potential areas already doing some of this work to learn from and organizations to include as partners in this work. In the following pages we provide a short overview of these recommendations.



Human Capital

Phase 1: Year 1



University degree, certificate programs, fellowships

- Network to develop training programs
- Provide dissertation fellowships, pre-doctoral fellowships, PhD student internships
- Provide faculty fellowships



Industry internships

- Internships for EdTech Professionals
- PhD Students



Annual event: LASI

- Organize annual capacity building activity
- Similar to Learning Analytics Summer Institute (LASI2013/LASI2014)



Preparing education researchers

- Develop short focused summer programs
- Education researchers bring data and get support for analysis as they learn



Changing teacher and leader preparation

- Identify a group of quality teachers who use data successfully
- Identify teacher preparation programs
- Create plan for integrating Learning Analytics into programs

Phase 2: Year 2-3



Start-up accelerator center

- Develop a cutting-edge startup accelerator for analytics-driven research
- Determine the best way to train people in the field



Integrate data-based decision-making into educator preparation

- Organize a committee to work with universities on implementation
- Integrate learning analytics in teacher and school leader preparation
- Integrate learning analytics in programs for in-service teachers and leaders

Phase 3: Year 4-5



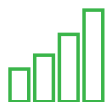
Establishing university education data science programs

- Provide competitive awards for establishing university Education Data Science PhD programs
- Develop programs that encompass departments of statistics, computer science, and education/psychology



Start-up accelerator center: worked examples

- Create a resource for teaching or for newcomers to the field
- Provide a data set and worked examples that highlight the types of questions that can be asked and answered with different analytic techniques



Research

Phase 1: Year 1



Research to prevent reinventing the wheel

- Develop the What Might Work and Why Clearinghouse
- Identify what has been done and what research is out there



Researcher & ed tech startup connector

- Identify a few Ed Tech startups to design competitions around solving real problems in education
- Provide strong financial incentives to produce professional products



Case studies to inform capacity building and policies for learning analytics

- Identify and develop case studies that demonstrate how to build capacity and policies
- Use case studies as tangible models for other districts and states to follow



Multimodal methods of measurement

- Develop multimodal learning analytics techniques
- Researchers use multimodal methods to examine unscripted, complex tasks in more holistic ways



Measuring success

- Determine what mastery or success looks like in structured and less structured learning environments
- Use new measures of success as outcomes



Prototype of personalized learning system Optimization of personalized learning

- Develop a prototype of a personalized learning management recommendation and reporting system
- Selected schools/districts pilot the system and provide feedback
- Determine explicit and/or tacit measures to serve as predictors of success
- Use predictors in future analyses

Phase 2: Year 2-3



Center for learning at scale

- Develop a center that will conduct a longitudinal study that follows a group of eighth or ninth graders to college
- Explore how to best collect longitudinal data that supplements classroom/learning data and take advantage of archival data

Phase 3: Year 4-5



Research to prevent reinventing the wheel

- Continue to identify what has been done and what is out there
- The What Might Work and Why Clearinghouse to provide a continually evolving research-based guide to learning



Personalized learning pilot

- Refine the system based on user feedback and roll it out as a free pilot
- Conduct projects to discover and validate/scale best practice for usage and visualization of data



Researcher & ed tech startup connector

- Continue participation in the Imagine K12 start up incubator program
- Disseminate success stories
- Create social network to link researchers and Ed tech startups



Tools

Phases 1,2 & 3: Years 1-5



Data science resource center

- Develop a center to create and provide a “Data Marketplace” or collection of datasets and streams for data scientists and developers
- Provide “Tools and Services” and help people use those tools to achieve their goals with big data



Competitions

- Incentivize innovation using industry models of competitions
- Identify the top 5-10 problems or grand challenges to be solved
- Competitions held in phases 2 and 3



Policy

Phase 1: Year 1



Templates

- Develop a set of templates for best data practices based on use cases
- Collaborate with the U.S. Department of Health and Human Services as the appropriate governance body for learning data policy issues



K-12 Data sharing and privacy standards

- Work with one or more states and a large district within each state to better understand how data sharing work
- Determine the issues, what is working, and effective methods for ensuring secure access to data

Phases 2 & 3: Years 2-5



Trust frameworks

- Develop and disseminate trust frameworks
- Develop tools, approaches, and practices for data sharing and privacy protection



Governing body

- Continue collaboration with U.S. Department of Health and Human Services to manage templates and create trust frameworks

Milestones for Measuring Progress

As we consider the road map to building the field of learning analytics, we also need to consider what milestones there would be to document that progress is being made along these funding lines. Some examples of milestones for measuring progress are provided in the following table.

Milestone Area	Description of Measurable Progress
Human Capital	<ul style="list-style-type: none">➤ An increase in percent of Carnegie-classified High and Very High Research university programs in Learning Analytics➤ A decrease in the human capital gap as measured by an increase in percent of trained people in the field➤ Improved decision-making on the part of districts, schools, and teachers to select products that are informed by learning analytics and have the greatest potential for improving outcomes for students➤ Improved decision-making for teachers and administrators using data based on new understanding of learning analytics.
Research	<ul style="list-style-type: none">➤ An increase in the percent of learners engaged in personalized learning environments developed with information from the field of learning analytics➤ Publication of case studies that inform capacity building with tangible models for districts to follow➤ Publication of metrics for success and guidance for how to use learning analytics to apply these metrics
Tools	<ul style="list-style-type: none">➤ An increase in the development and use of tools for learning analytics by members of the education community➤ Publicly available toolkit for use by education researchers and districts for learning analytics
Policy	<ul style="list-style-type: none">➤ Improvements in policy related to data privacy and data sharing for education, corporations and universities that support Learning Analytics➤ Collaboration with Health and Human Services (HHS) as the appropriate government body on learner data privacy and learning analytics with trust➤ Publicly available templates for best data practices

Branding and Getting the Word Out

There is a need for the field to identify a set of common messages that can be disseminated through conferences and other events. Funding would be needed to sponsor talks at conferences with a data science area of focus, such as the O'Reilly Strata conference, the O'Reilly Open Source Convention (OSCON), and SXSWedu conference. We recommend a collaboration with Strata and O'Reilly leaders to develop a StrataEdu conference.

Realizing the Promise: It Takes a Community of Funders to Build a Field

We hope to have made the case with the LAW Report for the needs and the transformative potentials of making K-12 personalized learning a reality continuously improved by education data science and learning analytics. We encourage funders to consider an essential area aligned with their mission and interest and one or more activities to financially support in order to aid in this field building effort throughout the three phases of this work over the next five to ten years.

