# 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 education 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 predesigned 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. **Education 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 table we provide a short overview of these recommendations.

**Table 1. Road Map to Implement Build the Field of Learning Analytics**

<table>
<tr><th></th><th>Human Capital</th><th>Research</th><th>Policy</th><th>Tools</th></tr>
<tr><td rowspan="6">Phase One</td><td rowspan="2">University degree and certificate programs and fellowships</td><td>Research to prevent reinventing the wheel</td><td rowspan="3">Templates for best data practices</td><td rowspan="9">Phases 1-3:<br><br>DATA SCIENCE RESOURCE CENTER to provide a data marketplace of tools and services<br><br>Competitions to incentivize innovation</td></tr>
<tr><td>Prototype of personalized learning system</td></tr>
<tr><td>Changing teacher and leader preparation to incorporate data-based decision-making</td><td>Establish an EdTech Startup Connector to connect researchers and EdTech startups</td></tr>
<tr><td>Annual capacity building event (e.g., LASI)</td><td>Case studies to support capacity building and policies for learning analytics</td><td rowspan="3">Development of K-12 data sharing and privacy standards</td></tr>
<tr><td>Industry internships</td><td>Measuring success</td></tr>
<tr><td>Preparing education researchers</td><td>Multimodal methods of measurement</td></tr>
<tr><td rowspan="2">Phase Two</td><td>START-UP ACCELERATOR CENTER to develop cutting-edge start-up accelerator for analytics-driven research</td><td rowspan="2">CENTER FOR LEARNING AT SCALE to focus on understanding personalized, contextualized learning at scale using analytics</td><td rowspan="3">Phases 2 & 3:<br><br>Development of trust frameworks</td></tr>
<tr><td>Integrate data-based decision-making into educator preparation</td></tr>
<tr><td rowspan="3">Phase Three</td><td>Establish university data science programs</td><td>The What Might Work and Why Clearinghouse guide to research on learning</td></tr>
<tr><td rowspan="2">Worked Examples Resource created by the Start-up Accelerator Center</td><td>Roll out personalized learning systems free-of-charge as a pilot</td></tr>
<tr><td>EdTech Start-up Connector creates Social Network</td></tr>
</table>