

Data-Intensive Research in Education: Ongoing Work in DR K-12

Chris Dede

Harvard University

Chris_Dede@harvard.edu

Elizabeth Burrows

AAAS Fellow at NSF

eburrows@nsf.gov

Approximate session agenda

- 9:30 – 9:45 Setting the stage
- 9:45 – 10:00 Attendee introductions
- 10:00 – 10:30 Small groups
- Brainstorm potential collaborations
 - Begin “Making an Impact”
- 10:30 – 11:00 Synthesis discussion
- Informal dialog about group progress
 - Take-aways
 - Input for NSF

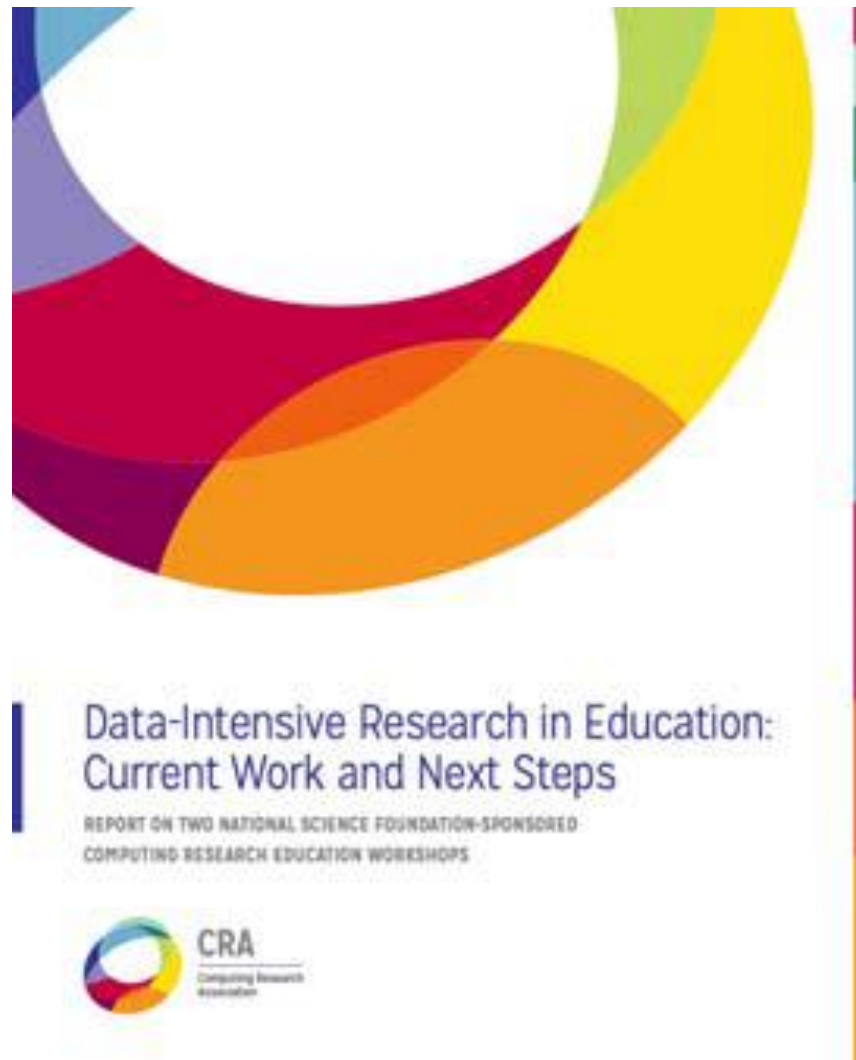
Chris's Current Role in Data-Intensive Research in Education

- Confront “big data” issues in my design-based research in ecosystems science education
- Organized a two workshop sequence on data-intensive research for NSF and the field: insights from relatively mature data-intensive research initiatives in the sciences and engineering were applied to nascent data-intensive research efforts in education

Liz's Current Role in Data-Intensive Research in Education

- As an AAAS fellow in the “big data” track, contributing to government-wide efforts in data science, such as helping to draft the NITRD Big Data Strategic Plan and co-organizing the monthly NSF Big Data Brown Bag
- Aided Chris et al. in organization and reporting of the two workshops on data-intensive research in education

<http://cra.org/cra-releases-report-on-data-intensive-research-in-education/>



Definitions

- Big Data is characterized by the ways in which it allows researchers to do things not possible before (i.e., Big data enables the discovery of new information, facts, relationships, indicators, and pointers that could not have been realized previously).
- Data-intensive research involves data resources that are beyond the storage requirements, computational intensiveness, or complexity that is currently typical of the research field.
- Data science is the large-scale capture of data and the transformation of those data into insights and recommendations in support of decisions.



Illustrative Types of Big Data in Education

- *Micro-behavioral data about students' activities in learning ecosystem science*
- *Micro-behavioral data about diagnostic performance assessments formative for learning and instruction*
- *Micro-descriptive data about activities in MOOCs*
- *Macro- and meso-level data about attributes and outcomes for teachers and schools*
- *Macro-behavioral data related to students' dropping out or staying in college*

Tools, Infrastructures, Repositories;

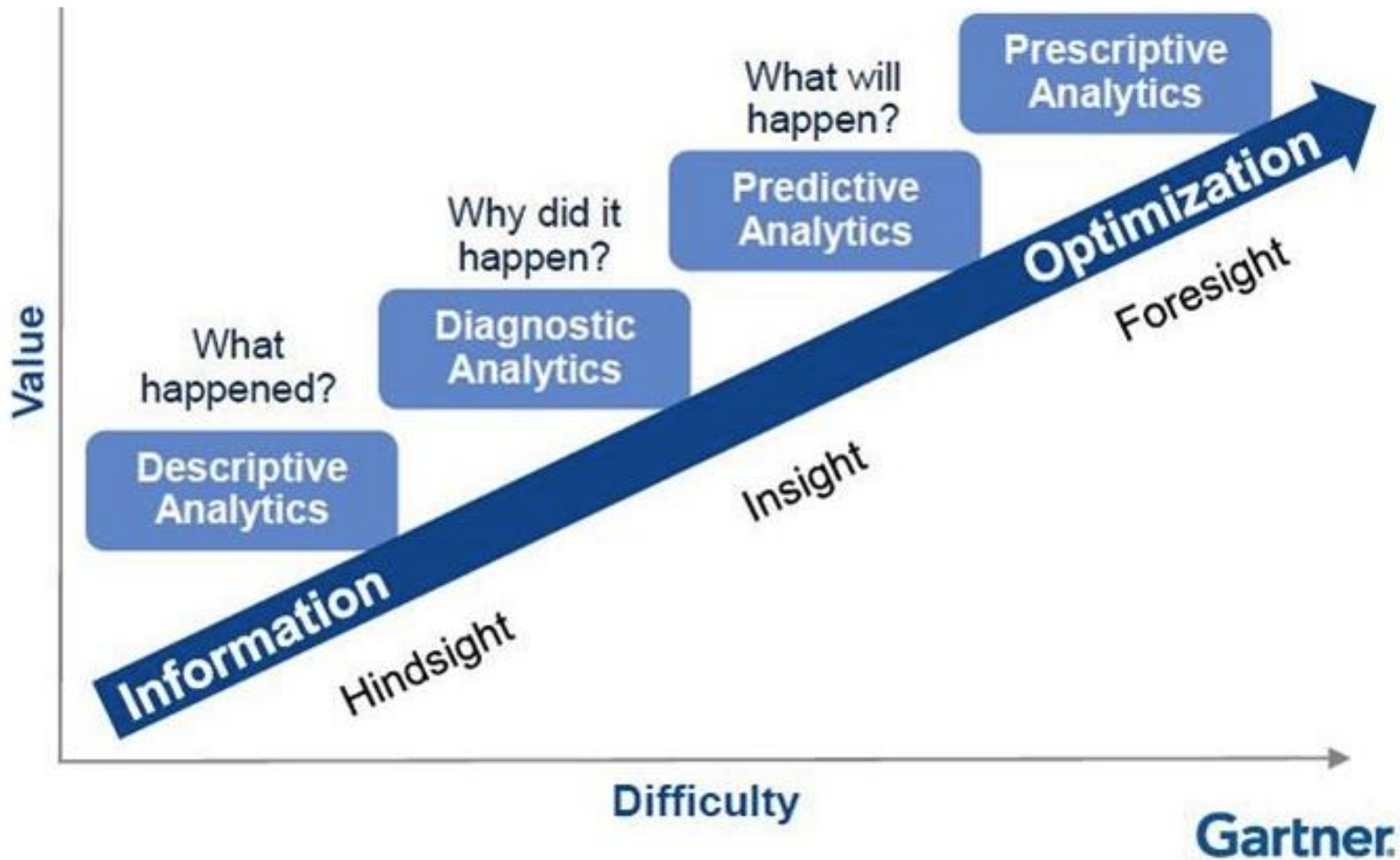
Privacy, Security, Safety;

Models from the Sciences and Engineering

From Description to Prescription

- Determine students' probabilities of failure (*predictions*)
- Determine which students respond to which interventions (*uplift modeling*)
- Determine which interventions are most effective (*explanatory modeling*)
- Allocate resources accordingly (*cost benefit analysis*)

From Hindsight to Foresight



Questions for Field

- To what types of behavioral data could we now apply these methods?
 - *Micro-level* data (e.g., each student's second-by-second behaviors as they learn)
 - *Meso-level* data (e.g., teachers' patterns in instruction; students' patterns in retention)
 - *Macro-level* data (e.g., aggregated student outcomes for accountability purposes) *Gummer's work with EdWise*
- What are the barriers to collecting, storing, sharing and analyzing these data?
- How can we build human and organizational capacity to use evidence-based findings effectively?

Tools for Transformational Insights



Key Next Steps from Report

- Mobilize Communities around Opportunities based on New Forms of Evidence
- Develop New Forms of Educational Assessment
- Develop New Types of Analytic Methods
- Build Human Capacity to Do and to Understand Data Science
- Develop Advances in Privacy, Security, and Ethics
- Infuse Evidence-based Decision-Making throughout Organizations and Systems

Potential Small Groups

- *Micro-behavioral data about students' activities in learning*
Micro-behavioral data about diagnostic performance
assessments formative for learning and instruction
- *Micro-descriptive data about activities in MOOCs*
- *Macro- and meso-level data about attributes and outcomes*
for teachers and schools
- *Macro-behavioral data related to students' dropping out or*
staying in school or college

Tools, Infrastructures, Repositories;
Privacy, Security, Safety;

Illustrative Questions for Small Groups

- Are there ways related projects might work together on joint Big Data initiatives?
- Are there questions NSF could answer on funding opportunities in this area?
- Are there ways NSF could aid in capacity-building for DR K-12 PIs on data-intensive research in education?

Synthesis discussion

- What progress was made in your small groups?
- Any advice or insights that would apply to other groups?
- Any challenges moving forward?
- Input to communicate to NSF?
- Steps you'd like to take on the report recommendations?