

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Roundtable on Data Science Postsecondary Education
Meeting #6 Agenda—March 23, 2018

Improving Reproducibility by Teaching Data Science as a Scientific Process

Hotel Shattuck Plaza
2086 Allston Way, Crystal Ballroom Section 2
Berkeley, CA 94704

Join the Webcast:

http://sites.nationalacademies.org/DEPS/BMSA/DEPS_184420

MEETING TOPICS:

Discuss how data science education can help understanding and improvement of the reproducibility of scientific research

Learn about current practices in training for reproducible data science in academia, industry, and government organizations

Upcoming Dates:

June 13, 2018: Roundtable Meeting #7—Data Science in 2-Year Colleges, Washington, DC (tentative)

September 17, 2018: Roundtable Meeting #8—Improving Female and Underrepresented Minority Participation in Data Science, Atlanta, GA (tentative)

December 10, 2018: Roundtable Meeting #9—Balancing Education in Employable Skills and General Knowledge, Washington, DC (tentative)

Committee on Applied and Theoretical Statistics
Board on Mathematical Sciences and Analytics
Computer Science and Telecommunications Board
Board on Science Education

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Friday, March 23

Improving Reproducibility by Teaching Data Science as a Scientific Process

- 9:00 AM¹ Welcome, new members, and introduction to the day
Eric Kolaczyk, Boston University
Kathy McKeown, Columbia University
- 9:15 AM Data Science as a Science: Methods and Tools at the Intersection of Data Science and Reproducibility
Victoria Stodden, University of Illinois, Urbana-Champaign
- 9:45 AM Teaching Reproducible Data Science: Lessons Learned from a Course at Berkeley
Fernando Perez, University of California, Berkeley
- 10:15 AM Break
- 10:35 AM Reproducible Machine Learning—The Team Data Science Process
Buck Woody, Microsoft Research
- 11:05 AM Group discussion of morning presentations
- 11:45 AM Lunch
- 12:45 PM Training as a Pathway to Improve Reproducibility
Tracy Teal, Data Carpentry
- 1:15 PM Rigor, Reproducibility, and Transparency Training in Biomedical Research
Alison Gammie, National Institute of General Medical Sciences
- 1:45 PM Buried in Data, Starving for Information: How Measurement Noise is Blocking Scientific Progress
Timothy Gardner, Riffyn
- 2:15 PM Break
- 2:30 PM Group discussion of afternoon presentations
- 3:00 PM Begin breakout group discussions
END WEBCAST
- 3:40 PM Report back of breakout group discussions and closing
- 4:05 PM Adjourn meeting

¹ All times Pacific time zone