



March 2018

Roundtable
HIGHLIGHTS

Roundtable on Data Science Postsecondary Education

Meeting #5 - December 8, 2017

The fifth Roundtable on Data Science Postsecondary Education was held on December 8, 2017, at the Keck Center of the National Academies of Sciences, Engineering, and Medicine in Washington, D.C. Stakeholders from data science education programs, government agencies, professional societies, foundations, and industry convened to discuss the integration of ethics and privacy concerns into data science education. This Roundtable Highlights summarizes the presentations and discussions that took place during the meeting. The opinions presented are those of the individual participants and do not necessarily reflect the views of the National Academies or the sponsors. Watch meeting videos or download presentations at nas.edu/DSERT.

Welcoming Roundtable participants, co-chair Eric Kolaczyk, Boston University, noted that there are inherent ethical and privacy implications in the choices data scientists make while framing, obtaining, cleaning, manipulating, and interpreting data. He highlighted the value of integrating this context of data science practice into data science education, and he hoped that the conversations at this gathering of the Roundtable would contribute to a more principled awareness of the ethics of data science.

TEACHING ALGORITHMIC ACCOUNTABILITY IN DATA SCIENCE EDUCATION

Cathy O'Neil, mathbabe.org

O'Neil began her presentation to the Roundtable by suggesting that data science ethics be reconceptualized as “algorithmic accountability.” She noted that although countless organizations use algorithms to score individuals (e.g., to estimate their propensity toward some desirable or undesirable behavior), their processes are not always scientific or ethical, and privacy and accountability may not be at the forefront of their concerns. What is most unfair, O'Neil described, is that recipients of such scores have no means to understand them, and there is often no mechanism in place to appeal decisions made as a result of these scoring systems. While these potentially destructive scoring algorithms rise to “secret laws,” in O'Neil's point of view, she said that many companies have yet to find evidence that they are effective in reflecting the true likelihood of what they purport to score.

An algorithm, according to O’Neil, makes predictions based on historical patterns. Although the definitions in an algorithm used to score individuals are crucial, these definitions are often determined secretly by those in power. Concerns also arise about the understanding of false positives and false negatives generated by the algorithm—balancing failures is just as important as having an accurate algorithm, O’Neil explained. She emphasized that it is already technically challenging to understand how and why various algorithms fail in different ways; it becomes even more difficult to hold algorithms accountable when they are optimized to a secret definition of success.

O’Neil provided three examples in which unaccountable, discriminatory algorithms are used in society:

1. *Teacher assessment based on students’ test scores.* Such a scoring system relies on bad proxies (i.e., test scores), bad statistics (i.e., low correlations), and unconstitutional practices.
2. *Job application filters such as mental health assessments and gender.* Such a scoring system is discriminatory, difficult to measure, and even more challenging to fix.
3. *Police dispatch to neighborhoods with high arrest data or arrest of low-level criminals to prevent violent crime in the future.* This unscientific system uses biased data and bad proxies (i.e., crime data are not the same as arrest data).

O’Neil commented that because lawyers and policy makers often do not have the appropriate levels of technical expertise, it is unreasonable to expect the legal system to keep pace with advances in data science. She encouraged academicians to address this issue of accountability in data science classrooms. She advocated for exposing future data scientists to these problems and teaching them to see themselves as accountable for ethically responsible products. She also suggested that, instead of only critiquing existing algorithms, data scientists who build algorithms could help policy makers by producing white papers geared toward non-experts and could involve lawyers in the development of ethical guidelines for algorithms.

O’Neil hopes that university data science institutes will also play a larger role in the devel-

opment of accountable algorithms. She noted the value of having a Hippocratic Oath for data science and encouraged data scientists to focus on their roles as translators of ethics instead of arbiters of truth. In response to a question from Solon Barocas, Cornell University, she suggested that data scientists reject jobs with organizations that do not build ethical (and legal) algorithms. For organizations utilizing algorithms for decision making, she suggested a scaffolding of monitors to ensure that algorithms are fair and legal and that data are clean. In response to a question from Patrick Perry, New York University, she elaborated that such monitors are valuable because they provide a continuous version of scientific algorithmic testing. She acknowledged that external data would be needed for validation throughout such testing.

Victoria Stodden, University of Illinois, Urbana-Champaign, asked O’Neil how she would teach these concepts at the graduate level. O’Neil responded that it is useful if every question to be addressed by an algorithm corresponds to a randomized experiment and if extreme mathematical cases are introduced. Aaron Roth, University of Pennsylvania, noted the bias that exists in data, even when humans make decisions, and wondered how machine learning is distinct from human decision making in terms of fairness. O’Neil highlighted the misconception that machine learning removes bias and encouraged humans to make their values explicit in the development of algorithms. Charles Isbell, Georgia Institute of Technology, asked how far the legal framework could be extended in algorithm development, and O’Neil responded that algorithms are already subject to the law; the questions that remain are whether these laws are enforced and when regulators will have the appropriate tools to measure legality. In response to a question from Perry about algorithmic definitions of success, O’Neil suggested having stakeholders complete an ethical matrix of their concerns about an algorithm. Such a matrix reveals that fairness is always a balancing act when trying to optimize with so many constraints. Perry countered that it seems implausible to determine the cost of making the wrong decision, but O’Neil reiterated that while considering the ethical implications is difficult, it is essential. Barocas identified this as another example in which challenges related to fairness still exist even when the data are reliable.

UNCOVERING THE SUBSTANCE OF A DATA SCIENCE ETHICS EDUCATION

Solon Barocas, Cornell University

Barocas focused on the content of data science ethics education as opposed to the structure through which it is delivered (e.g., stand-alone courses versus integration throughout an entire course of study). He began by extending standard concepts of professional responsibility common to many fields—to do work that is valid, reliable, and transparent—to data science practice and education. Similarly, common professional virtues to strive to instill within future data scientists include skepticism about how models will perform, humility regarding the limits of the models that one develops, honesty to avoid misleading users, and vigilance to ensure that models work well after deployment. Standard ethical dilemmas such as the trolley problem can motivate students to question and develop their own moral agency and moral intuition. These generic approaches to professional ethics do provide value in the context of data science education, particularly in helping students to connect concepts of validity and reliability to questions of fairness and bias in algorithms with relative ease. However, Barocas commented that these approaches are not specific to data science and thus may be inadequate for data science ethics education.

Barocas remarked on the growing interest in the field of “data ethics,” noting that it is unclear what this field entails. Standard approaches underscore privacy (i.e., adherence to the [Fair Information Practice Principles](#) and use of anonymization to safeguard personal information); however, clearly new ethical issues are arising in data science that fall outside of this narrow purview. The past few years have seen increased interest in adapting research ethics principles (i.e., autonomy, beneficence, and justice), which are historically designed to protect research participants, to the use of data analytic tools in companies. This is not a surprising approach, explained Barocas; however, research ethics still does not encompass the breadth and complexity of the ethical and normative questions future data scientists will face.

Barocas described a new upper-level undergraduate elective at Cornell University—[INFO 4270: Ethics and Policy in Data Science](#)—targeted toward aspiring data scientists from the disciplines of information science, computer science,

and quantitative social science. He mentioned that much of the syllabus grew from the annual “[Fairness, Accountability, and Transparency in Machine Learning Workshop](#),” which seeks to build a technical community interested in deeper normative questions in data science work. While interest in this conference and its subject has grown rapidly, Barocas worried that some researchers still mistakenly think that formalizing decision making through algorithms ensures fairness or prevents bias. In part, this is based on experience with human decision makers who do exhibit bias, which can be ameliorated through more formal decision processes (e.g., actuarial scoring tools). He emphasized that using machine learning does not ensure fairness and that misuse of data science can foster inequality in and prevent opportunity for segments of the population.

“Ethics and Policy in Data Science” challenges students to explore familiar technical problems—for example, detecting unobserved differences in model performance, coping with observed differences in model performance, and understanding the causes of differences in predicted outcomes—with greater ethical specificity. Focusing on an example of model validation, Barocas said that data scientists must make normative decisions during validation (e.g., to validate with respect to accuracy of predictions, with respect to differences in error rates, or with respect to differences in outcomes across subpopulations) and that data science ethics education can engage students in deliberation about the ethical implications associated with their modeling decisions. Regarding differences in outcomes, Barocas suggested that data scientists consider the historical events that shape algorithmic outputs about an individual (e.g., whether that person’s family has a history of interaction with the criminal justice system) and to perhaps consider algorithmically aided decision making as a way to remedy past injustices.

“Ethics and Policy in Data Science” consists of 12 broad modules: cultivating a critical disposition in students toward data science and their own work; understanding bias in humans, algorithms, and data; case studies and opportunities in algorithmic auditing; formalizing fairness and trade-offs between different measures of fairness; individual agency and individualized assessment and the ethical dimensions of modeling individuals based on factors over which they have no control or based on their characteristics in reference

to larger populations; moving from allocative to representational harms; transparency, interpretability, and explainability of algorithms and models from the perspective of policy makers or tool users; privacy protections and loss of privacy from precise, automated inference; price discrimination in marketing and insurance models; broader questions about algorithms in the public and their impact on democracy; and the ethics of autonomous experimentation by algorithms deployed in the real world. The final module in the course is about refusal and rejection, where data science students and practitioners explicitly choose not to pursue specific projects because they are ethically or practically objectionable. Barocas closed his presentation by appealing to senior data scientists to lead by example in refusing ethically questionable projects, which in turn will provide an example and protection for more junior researchers, practitioners, and students wishing to reject a project.

RECOGNIZING AND ANALYZING FALSE CLAIMS FROM BIG DATA

Jevin West, University of Washington

West opened his presentation by noting that while many students excel in the execution of mechanics, they often lack the skills both to engage with ethical considerations for data analysis and to understand basic experimental design. In his classroom, West reveals to students, who may not appreciate the limits of technology, that machines make mistakes and harbor bias similar to humans. Instead of offering only a brief unit of study on ethics, he integrates these conversations throughout his curriculum. He encouraged faculty to adopt the Humanities' approach to textual analysis, as future data scientists need to develop critical thinking skills to interrogate and interpret data.

West commented that society is drowning in false information, especially with the rise of charts and quantification in the news. In an effort to teach students to recognize and analyze false claims and to be able to communicate this information to broad audiences, West and his colleague Carl Bergstrom developed the course “**Calling Bullshit in the Age of Big Data**” at the University of Washington. The course includes topics in the following areas: false information and misrepresentation, causality, statistical traps and trickery, data visualization, big data manipulation, publication bias, predatory publishing

and scientific misconduct, fake news and other shams, and refutation of falsehoods. Given that the course emphasizes data reasoning, West dedicates much instructional time to causation and refutation. Campuses across the country and abroad have adopted the course, and West and his colleagues also engage local middle and high school students in similar instructional sessions.

West next demonstrated a contrast between “old school bullshit”—empty phrases and circular reasoning are readily detected and disproved—and “new school bullshit”—scientific language and visualizations are presumed to be fact. While he acknowledged that the notion of the “black box” can be daunting to students, they can recognize misrepresentations by looking carefully at the data that is input into the algorithm as well as the output and the interpretation of an algorithm.

West suggested using real-world examples to create engaging classroom exercises that challenge students to identify instances in which an argument's methods or assumptions lead to absurd conclusions or causations. He shared a series of tips for “calling bullshit”: (1) Think about claims that seem too good to be true; (2) Beware of confirmation bias; (3) Recognize multiple working hypotheses; (4) Evaluate orders of magnitude; and (5) Be wary of unfair comparisons. West concluded by emphasizing the value of improving ethical data science education models at the secondary and postsecondary levels and engaging students and the broader public in data reasoning.

OPEN DISCUSSION

Bias and False Information

Jeffrey Ullman, Stanford University, described the “fake news” discussed in West's presentation as an intractable problem and asked for ideas to formally identify it. West admitted that, right now, it is impossible—it is important to arm machine learning consumers with the right skills and hope that artificial intelligence will catch up eventually. He explained that, unfortunately, for every algorithm created to identify fake news, there is another one designed to create fake news. Bill Howe, University of Washington, commended O'Neil's and Barocas's attention to validity but expressed concern that people may be under the false impression that simply building the perfect model solves all problems. He emphasized that the issues are far more complex. O'Neil said

that the other, worse extreme is when people assume that nothing can be trusted and lose faith in technology entirely. She advocated emphasizing the science in data science by testing frameworks around algorithms so that they can be trusted. Barocas shared O’Neil’s concerns but added that the foundation for seemingly objective work is actually subjective (i.e., nothing can be learned without some amount of bias). Howe also noted that data scientists have choices and power before training a model, and he emphasized the value of teaching students about these crucial data management steps. Barocas agreed and suggested starting courses with the question, “What is data?” Stodden observed that the topic of bias was central to all three presenters’ talks. Because bias is defined narrowly in entry-level statistics courses, and that definition may not translate well in larger discussions, she suggested that the data science community think about how to teach what bias is as well as how to think about data science more broadly.

Preparation for Faculty and Students

Michael Fountane, graduate of Harvard University, asked the presenters about senior-level faculty responsibilities in teaching future data practitioners. Barocas noted that, generally speaking, professors want to produce students who will do high-quality work. He added that competitive marketplaces should then reward those who become practitioners and avoid making statistical errors. O’Neil commented that, especially in financial trading, there is a strong incentive to be accurate so as to maximize profit but there are not nearly as many stakeholders as there are in data science spheres. The realm of data science is much more complicated because these many stakeholders have differing definitions of success, and their values have to be balanced against one another. Many people, she explained, either misunderstand this complexity or choose not to think about it. West reiterated that ethics instruction (i.e., a new way of thinking about and communicating the social elements of data) has to carry through all components of a data science education.

David Culler, University of California, Berkeley, wondered how to educate students to exercise good judgment. West noted that his course incorporates case studies and project-based work in which students are set up to fail; they quickly learn about the value of good judgment in such scenarios. O’Neil said that students can

be taught to practice good judgment through exercises in which they work on one algorithm with multiple choices. Barocas discussed the importance of providing students with messy data so as to better prepare them for real-world experiences. Al Hero, University of Michigan, cautioned that although the concept of “calling bullshit” can energize students, it risks showing students that finger-pointing is always justified. Instead, Hero suggested teaching students to ask what evidence would be needed to make a true claim. He described this as a more constructive way to teach about the inadequacy of selected data and to increase appreciation for negative results, since this is how the scientific enterprise is motivated to continue its work. West noted that selection bias and reproducibility are topics of his course lectures, as are the civic and political implications, and he added that the field of data science could also learn from approaches used in applied psychology. Moses Namara, Clemson University, asked how to motivate people to scrutinize data, and West responded that students are both idealists and natural contrarians. He said that it is important for students to understand the consequences of misusing data, but he cautioned against letting students believe that no truth exists anywhere. Nicholas Horton, Amherst College, commented that there is a clear need for a variety of approaches to and a spiraling curriculum for educating future data scientists. He emphasized addressing key concepts early and often in courses and encouraged the building of critical thinking skills at different levels. He urged faculty to identify learning outcomes related to data integration and data fusion and suggested enhanced faculty training. He described “data literacy for all” as a way for people to better understand the world around them without fear.

MATHEMATICAL APPROACHES TO PRIVACY AND FAIRNESS

Aaron Roth, University of Pennsylvania

Roth presented two important social issues in technology: privacy and fairness. He emphasized the value of approaching these complicated issues from formal, mathematical perspectives. He noted that mathematical approaches to privacy already exist, but fairness is an emerging area of study with a recent explosion of research. A standard definition of and a quantitative approach to fairness would be useful, according to Roth, but both privacy and fairness require understanding trade-offs through formal reasoning.

NAVIGATING HISTORY, PRIVILEGE, AND POWER IN INFORMATION AND DATA SCIENCE

Anna Lauren Hoffmann, University of Washington

Roth described privacy as the promise of freedom from harm. Privacy has been a public concern for decades; despite the use of de-identification techniques, people can still be connected to their data. He acknowledged that privacy is more complicated than hiding personally identifiable information or releasing only aggregate statistics. It is impossible for data analysts not to know anything more about a subject after analyzing her data when auxiliary information is present. Roth pointed out that if this instance is treated as a privacy violation, it becomes virtually impossible to do scientific research, as auxiliary information often reveals information data scientists want to learn. He alluded to [an article published in 2006](#) that discussed the notion of differential privacy—a data set (in which each piece of data belongs to an individual) is inputted into a randomized algorithm, and even if the data are changed for one individual from the data set, the behavior of the algorithm should not change substantially.

Roth noted that many statistical problems can be solved privately with convex optimization, deep learning, spectral analysis, and synthetic data generation, for example. He emphasized that trade-offs will always exist (e.g., accuracy, sample sizes, and privacy level)—mathematical thinking simply allows one to better understand those trade-offs. He noted that although there is still much work to be done translating theory into practice, organizations such as the U.S. Census Bureau already rely on differential privacy.

Roth next turned to a discussion of fairness, using a case study of COMPAS—the recidivism risk prediction software. Investigative journalists at [ProPublica](#) described the tool as unfair and biased against black people, owing to differences in false positive and false negative rates between black people and white people. COMPAS analysts responded that they used a different metric for fairness. While Roth explained that both analyses offer reasonable definitions of fairness, no classification tool can simultaneously satisfy both conditions and equalize false negative rates if the base rates in the two populations differ. Equalizing false positive rates across subpopulations is only one measure of fairness, and it is unclear whether this is the appropriate metric. Roth concluded that the benefit of formalizing such fairness measures is that it allows better management of trade-offs, improved algorithmic design, and scientific progress toward more informed policy making.

Hoffmann encouraged the teaching of ethics in applied contexts—better decisions can be made about issues with moral impact if a combination of disciplinary, theoretical, and activist knowledge is considered. It is important for data science students to realize that different problem solving goals require unique considerations. She emphasized that historical and contextual information are essential in ethical decision making.

Hoffmann observed that data ethics is the intersection of moral, methodological, and practical concerns—data scientists need appropriate tools to balance these three areas. She emphasized the value of confronting these issues with disciplinary diversity, utilizing people with varied skill sets to solve complex problems. In her courses, Hoffmann approaches ethical considerations through a study of context, relevant history, key concepts, and the data life cycle. She suggested that ethical issues arise not only in algorithms and analysis, but also in data collection, and thus should be a part of the entire research life cycle. She noted the importance of teaching students to think about how platform design affects data as well as how to think critically and holistically about data and the problems they can solve.

Hoffmann emphasized that, like any tool, data has affordances. Ultimately, data allow one to count, organize, and make decisions. She emphasized that these processes are not wholly new—there is a canon of historical examples that expose the importance of research ethics. Discussion of Nazi experimentation and the Tuskegee studies, for example, help contemporary students understand ethical issues and determine how to apply these lessons to current case studies. She offered multiple examples, including the Henderson Roll—an illegal census in the 19th Century of Native Americans—as evidence of historical precedent about the vulnerability of certain populations in the face of data-driven systems. She emphasized that such injustices could be perpetuated when people voluntarily provide records to the government (e.g., as discussions about Deferred Action for Childhood Arrivals continue in the United States) and reiterated the importance of thinking about uncovering issues

in and using history to solve current problems in new ways.

OPEN DISCUSSION

Teaching Differential Privacy

John Abowd, U.S. Census Bureau, suggested that faculty focus on teaching the ratio of differentially private variance to the regular variance in their courses—the trade-off is clear and the privacy costs are revealed in that instance. Howe mentioned the tension between reductionism and interpretability. For example, he wondered how many people (especially lawyers responsible for decision making) can reliably understand and interpret differential privacy. Roth acknowledged that it will never be possible to write a mathematical constraint about privacy upon which everyone will agree, but formalization helps to reveal incompatible components. He noted that although one may not know a parameter for differential privacy, a quantitative discussion about privacy levels is possible and useful. Hoffmann added that reflective conversations about privacy and debates about trade-offs are more valuable than a focus on finding the “right answer.” Hero noted that differential privacy and its measures place the analyst in the role of determining acceptable levels of privacy. But, he hypothesized, in the future, when individuals can select their own trade-offs, privacy may become a valuable, tradable asset. Roth clarified that the analyst does not set the privacy level and added that differential privacy is only a metric. He mentioned that there have not been many successful markets for private data in big data applications thus far because they are not very useful and are easily replaceable. He noted that people would need to alter the way they think about privacy before data markets would change.

Balancing Trade-offs

Perry noted that privacy, fairness, and accuracy are all trade-offs that are at odds with one another. He wondered if one should either place different weights on each factor and optimize for the objective function or explore the frontier first and then assign weights. He asked if the latter approach is dangerous because it allows a decision to be made *after* seeing the trade-offs—in other words, sacrificing privacy and fairness for accuracy. Roth explained that it is the responsibility of the technologists to identify trade-offs and of the society to balance competing needs. It is only possible to make fully informed decisions

after understanding the trade-offs, according to Roth. Hoffmann refuted this notion: by setting up values as trade-offs, one has already surrendered to certain inequalities. Roth responded by saying that while it is tempting to suggest that because the Constitution guarantees fairness it cannot be a trade-off, that is not the reality in which we live. Trade-offs have to be discussed in every case. Hoffmann acknowledged that discussions about trade-offs are useful unless a concession has already been made earlier in the process and a different set of trade-offs needs to be debated. Roth agreed that it is always reasonable for researchers to step back and evaluate what is most important.

In light of this discussion about fairness, Isbell commented that bias can be built into the data itself. Roth acknowledged that both data and algorithms can be problematic in terms of bias. He explained that problems with data are hard to measure, and, even if those problems were eliminated, fairness would still be an issue. He encouraged investments to be made in the study of both data and algorithms. To simplify the problem, he suggested first thinking about the data and algorithms in isolation. In response to a follow-up question from Isbell, Roth noted that it should be possible to formalize the problem in data collection and reiterated that fairness is only just beginning to be understood. Ullman asked about assumptions about right and wrong that faculty are making in their courses, as well as about how far rights to privacy extend, and wondered whether the conversation should focus only on the implementation of technologies. Roth reiterated that the technologist’s role is to help discover and delineate trade-offs, not to make decisions about policy or morality. He added that it is possible to write definitions with parameters on which trade-offs will occur. Barocas added that the 4th Amendment determines certain rights and noted that it is the responsibility of faculty to show students that long-standing issues are not new simply owing to the onset of data or technology. Hoffmann emphasized that when people are being harmed by data and software, science must progress to make changes.

Navigating Social and Technical Concerns

Barocas asked Roth and Hoffmann to comment on one another’s talks. Roth acknowledged that he has not yet overcome the obstacles of language differences across disciplines. He explained that even though he and his students

are often operating with toy models for which the complicated problems of the world have been abstracted away, the complex problem of fairness persists. Roth continued that because fairness is complicated, it is critical to understand first how social and technical issues work in isolation and then how they work together. Hoffmann said that her work provides the broader social and political motivation for Roth's research. She reiterated that his and others' work have a larger framework; working ahistorically will only further fragment problems. Referencing Stodden's earlier observation about bias, Hoffmann recognized that communities contextualize bias differently—social theory and historical casework can orient people toward a positive vision about a socially acceptable definition. Mark Krzysko, Department of Defense, mentioned that his team regularly confronts many of the issues discussed in Hoffmann's presentation. He added that access and dissemination are also concerns for the Department of Defense and that it is important for future employees to understand and to engage in constructive dialogues about both data and institutional values.

Educating Students

Kolaczyk asked how Hoffmann and Roth raise issues of ethics and fairness in the classroom. Hoffmann requires written assignments including memos, opinion pieces, blogs and, during the data collection stage, reflective exercises. Roth responded that because he teaches mathematics to Ph.D. candidates, his focus is on equipping students with the skills to push research topics forward. He encourages students to look at popular media to explore real-world questions, but he does not teach interdisciplinary courses or content. Abowd asked how Roth would incorporate the notion of "privacy as a public good" into discussions about system design; Roth recognized the value in thinking about privacy in terms of economic quantities to be analyzed, and he supported further collaboration between scientists and economists.

SMALL GROUP DISCUSSIONS AND CONCLUDING CONVERSATIONS

Roundtable members and audience participants formed subgroups to discuss one or more of the following questions: (1) How could ethics be integrated into the data science curriculum? (2) What mechanisms can educators use to help students navigate between the informal (e.g.,

fairness) and formal (e.g., algorithmic accuracy) terminology of data science? (3) How might educators teach students about the ethics of data science without radicalizing or paralyzing them with skepticism?

Isbell shared considerations raised by his group about integrating ethics into a data science curriculum. As most of the Roundtable's discussion focused on real-world consequences of actions, he questioned whether "ethics" needed to be integrated at all. While it is possible to explore real-world consequences through the lens of ethics, there are other ways to do so, he continued. He noted that it is unrealistic to expect faculty with varied levels of expertise to integrate ethics units into their courses, and it would be equally unsuccessful to create an ethics course out of context. He instead suggested asking faculty to focus on motivating the problems they assign to students in each and every class meeting and then working toward a study of real-world consequences. This approach may be both easier to integrate across the curriculum and more interesting for students. An audience participant suggested developing a required ethics course, separate from core requirements and including guest lecturers from other departments, as a way to motivate students to think about the consequences of working with data. This participant also suggested adding a question about real-world consequences to every student project.

On behalf of his group, Fountane suggested incorporating ethics into curricula with the implementation of an orientation course on reasoning at a high meta-cognitive level (i.e., how to write/model ethical standards to peers). He described this as a practical way for students to engage in active reflection, which is a more socially valuable skill than calculation. Faculty may be more likely to buy in to this approach if ethics played a larger role in the professional data science discipline. For example, both the Association for Computing Machinery and the American Statistical Association already have guidelines for the inclusion of ethics in professional practice. An audience participant added that **Bloom's Taxonomy** could be used to structure and integrate conversations on data ethics in either of the classroom models shared by Isbell or Fountane. She explained that Bloom's offers a way for undergraduates to develop evaluation and critical reasoning skills gradually through paced activities.

Hoffmann shared her group's discussion, noting that educators do not yet have a good understanding of how much content from other courses might be useful in the development of ethics curricula. Referencing the National Academy of Engineering report [Infusing Ethics into the Development of Engineers](#), she wondered which models already exist and how much data science educators should create anew. She suggested that industry create incentives for academia to better understand and to better provide data science education. She also added that there are many structural barriers to developing new curricula in higher education—administrators and students alike often do not realize which skills industry values. An audience participant expressed concern about the lack of diversity in STEM Ph.D. programs and noted that it is crucial to discuss diversity in any conversation about ethics. She explained that students are, by nature, curious and can be motivated to study data science when courses are student- and project-centered. She emphasized that embedding ethical conversations and real-world problems into courses can also serve as a mechanism to improve diversity.

Summarizing his group's discussion, Perry noted that although many faculty may want to teach ethics, they may not know the best approach. For this reason, he cautioned against forcing faculty to teach ethics. He observed the need for faculty scripts for smoother incorporation of ethics without disrupting course material and for case studies relevant to material being taught. And so that students do not become paralyzed with skepticism, he explained that it is important that these case studies show students possible solutions to problems. Since undergraduate students are often interested in exploring problems and non-technical material, it may seem more feasible to incorporate ethics at this level than at the master's level, in which students are focused on building technical skills that can be applied in the workforce. Perhaps a way to introduce ethics at the master's level is to talk about the mathematics of differential privacy. Most importantly, Perry explained, it may be detrimental for students to believe that mathematics solves all problems. It is crucial that students are involved in hands-on exercises that show the consequences of fairness. For example, faculty could ask students to build a model to predict something that is relevant to them and randomly assign covariates. Such an exercise offers personal incentives (beyond

the moral imperative) for students. Howe suggested replacing data sets in the classroom (e.g., use Compass instead of Iris) to start a conversation about fairness. He also wondered if there is a way to obtain better curated data sets that could be crafted for teaching purposes, although he recognized that students are rarely excited about "fake" data. He referenced an effective and engaging assignment from danah boyd—the only correct way to complete it was to refuse to complete it on ethical grounds.

Kolaczyk, speaking on behalf of his group, suggested that the best practices used by the bioinformatics community for including ethics in the curriculum could be leveraged, if they are computationally motivated. He added that there are examples from the social sciences for integrating both the quantitative and the qualitative (e.g., survey and sampling taught together). He also noted that the integration of ethics depends on the context—for example, Boston University's statistics practicum includes situational role-play, which may not be as effective in a theoretical course. Hero encouraged the inclusion of statistical consulting in engineering projects as a way for students to learn to interact with both people and data and become more aware of the consequences of their actions. Such an experience is personalized, offering a well-designed teachable moment. Constantine Gatsonis, Brown University, referenced a course he is designing: Case Studies in Health Data Science. The course invites speakers from Brown's School of Public Health to present real data sets and case studies. He will provide a template for privacy and ethical considerations around which speakers will organize their presentations. He hopes that this will be an effective means to generate useful discussion with the students. Stodden highlighted a forthcoming publication in *Communications of the ACM* that emerged from a working group of the Advisory Committee of the Computer and Information Science and Engineering directorate of the National Science Foundation. The document illustrates the life cycle of data science (i.e., acquire, clean, use, reuse, publish, preserve, and destroy), with ethical questions to be addressed from technical and mathematical perspectives at each stage, making it possible to start to frame concretely how ethics fits into data science.

ABOUT THE ROUNDTABLE: The Roundtable on Data Science Postsecondary Education is supported by the Gordon and Betty Moore Foundation, the National Institutes of Health Big Data to Knowledge program, the National Academy of Sciences W. K. Kellogg Foundation Fund, the Association for Computing Machinery, and the American Statistical Association. Within the National Academies, this roundtable is organized by the Committee on Applied and Theoretical Statistics in conjunction with the Board on Mathematical Sciences and Analytics, the Computer Science and Telecommunications Board, and the Board on Science Education. Roundtable meetings take place approximately four times per year. Please address any questions or comments to Ben Wender at bwender@nas.edu.

DISCLAIMER: This meeting recap was prepared by the National Academies of Sciences, Engineering, and Medicine as an informal record of issues that were discussed during the Roundtable on Data Science Postsecondary Education at its fifth meeting on December 8, 2017. Any views expressed in this publication are those of the participants and do not necessarily reflect the views of the sponsors or the National Academies.

ROUNDTABLE MEMBERS PRESENT: Eric Kolaczyk, Boston University, Co-Chair; John Abowd, U.S. Census Bureau; Constantine Gatsonis, Brown University; Alfred Hero, University of Michigan; Nicholas Horton, Amherst College; Bill Howe, University of Washington; Charles Isbell, Georgia Institute of Technology; Mark Krzysko, U.S. Department of Defense; Patrick Perry, New York University; Victoria Stodden, University of Illinois, Urbana-Champaign; and Jeffrey Ullman, Stanford University.

GUESTS PRESENT: Aubra Anthony, USAID; Anna Arnando; Solon Barocas, Cornell University; David Culler, University of California, Berkeley; Michael Fountane, alumnus of Harvard University; Simson Garfinkel, U.S. Census Bureau; Anna Lauren Hoffmann, University of Washington; Brian Kotz, Montgomery College; Aaron Margolis, Federal Bureau of Investigation; Moses Namara, Clemson University; Kyle Novak, American Association for the Advancement of Science; Cathy O’Neil, mathbabe.org; Hridayesh Rajan, Iowa State University; Aaron Roth, University of Pennsylvania; Mary Rudis, Bates College; Dhruv Sharma, Federal Deposit Insurance Corporation; Rochelle Tractenberg, Georgetown University; and Jevin West, University of Washington.

STAFF PRESENT: Linda Casola (via webcast), Janki Patel, Michelle Schwalbe, and Ben Wender.

Division on Engineering and Physical Sciences

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people’s lives worldwide.

www.national-academies.org