Meeting #1 Notes

Council For Big Data, Ethics, and Society[1]

November 14, 2014
New York University / 239 Greene Street, 2nd Floor Conference Room

Schedule:

- 9:00-10:00: Breakfast and Check-in
- 10:00-10:30: Introductory Framing and Background on Council
- 10:30-12:15: Existing Ethical Conundrums Break-Outs (moderator: danah)
- 12:15-1:15: Lunch
- 1:15-3:00: Process for Addressing Ethical Issues (moderator: Kate)
- 3:00-3:30: Break
- 3:30-4:30: Empirical Project Brainstorm (moderators: Geof & Jake)
- 4:30-5:00: Closing

Attendees:

Council Members:

- Solon Barocas, Princeton University
- Geoffrey C. Bowker, University of California-Irvine (co-PI)
- danah boyd, Data & Society / Microsoft Research (co-PI)
- Kate Crawford, Microsoft Research / New York University (co-PI)
- Rachelle Hollander, National Academy of Engineering
- Barbara Koenig, University of California at San Francisco
- Arvind Narayanan, Princeton University
- Alondra Nelson, Columbia University
- Helen Nissenbaum, New York University (co-PI)
- Frank Pasquale, University of Maryland
- Seeta Peña Gangadharan, New America / Data & Society
- Latanya Sweeney, Harvard University
- Sharon Traweek, University of California at Los Angeles
- Matt Zook, University of Kentucky

Staff:

- Jake Metcalf, postdoc
- Ellen Mendlow, project management

Council members unable to join us: Alessandro Acquisti, Mark Andrejevic, Ed Felten, Alyssa Goodman, Eric Meslin, Paul Ohm

Introductory Framing and Background on Council:

danah boyd started the meeting by welcoming everyone and launching a round of introductions. She then discussed the broad goal of the Council to step back from the fast-moving public discussions about big data with all of its ethical conundrums and focus for a moment on what ethically-driven "big data" research can and should look like. How do we integrate ethical frameworks into different types of technical research? What does intermediation between academic and industry ethics look like?

In light of the diverse group of people in the Council who are coming from many different communities and research disciplines, boyd asked us to imagine what kind of impact we can have on policies, research agendas, and public conversations. The Council was structured so that members could create a co-constructed agenda and collectively imagine the project.

Conversation lead by boyd touched on some primary areas of activity the Council could pursue.

1) **Influence what the NSF is doing.** The NSF funded the Council through an EAGER grant in order to support the NSF in its efforts to support ethically sound research. The NSF has no specific and direct requests but is looking to the Council for ideas on how to weave ethics through its existing practices. The Council has substantial latitude on this matter, but the project proposal promised a white paper, public commentary, and regular meetings with PIs funded by CISE's BIGDATA program.

   a) **ELSI Model**: One possible outcome is that the NSF decides to model a data ethics program on the NIH's ELSI (Ethical, Legal, and Social Implications) funding for ethics research (broadly conceived). When the National Human Genome Research Institute decided to invest heavily in genomics research, it decided that 3-5% of research budgets should be invested in ethics research in order to sustain societal support for basic research in genomics. ELSI funds supported many soft-money research centers, and is sometimes jokingly called 'the full-employment act for bioethicists.' The Council was supportive of investigating the strengths and weaknesses of the genomics ELSI model and considering proposing something similar for data ethics at the NSF.

   b) **Data Management Plans**: A plausible point of leverage that the NSF could have over the ethical practices of big data research is the mandatory 'data management plans' (DMP) that are included in every project application. Presently, DMPs are not expected to address many ethical concerns, other than the exhortation to make research data available to others. They could possibly address provenance, privacy, repurposing, results sharing, etc., and help propagate new data ethics practices in the field.

2) **Affect research agendas writ large, with and beyond the NSF, in academia and industry.**
   Data ethics is still a relatively nascent field. The Council could have a substantial effect in shaping the priorities of the field.

3) **Speak to a much more public audience.** This past year has seen a major increase in attention to data ethics, especially with recent industry controversies. The Council could build venues for quick responses to problems as they arise.
    a) **International Scope**: Although terms of our NSF grant limits us to US-residing researchers and encourages us to work on topics covered by the NSF BIGDATA program, many in the Council hope to participate in international conversation.

4) **Generate updated ethics codes and policies for the major computing professional societies and data practitioners.** The ethics codes of many major professional computing societies pre-date the Internet Age, let alone the rise of big data practices. The Council could encourage and support revisions to such codes.

5) **Consider how we might perpetuate ourselves.** The Council is currently on a two year schedule via the EAGER grant. It is an open question how/whether the Council should limit itself to the scope and timeline originally proposed or if we should begin seeking opportunities to build out the project beyond the limited scope mapped out in the initial grant.

Existing Ethical Conundrums Break-Outs & Discussion

This session was lead by boyd. The pre-reading <LINK> for this session organized data research ethics controversies into major thematic areas. Her opening comments encouraged the Council to focus on controversies arising from research, not just industry or government practices that may have had the most widespread public attention. The goal of this session was to collaboratively develop a map of the ethical challenges that researchers face when trying to design big data research projects. We looked at existing controversies and the issues that they raise and then we will turn to consider the kinds of questions that researchers should be asking of themselves and of each other. boyd directed attendees to not focus just on external controls (likes IRBs), but to think about how all researchers can and should be engaged in ethical thinking and focus on what should be discussed and debated as opposed to a list of dos and don'ts.

Attendees were assigned to three break-out groups, numbers 2-4. (The original first group was split across the other three groups because of space/attendance dynamics.) Below are some of the major ideas introduced by each group.

Group #2:
Ethical conundrums:
- **Extra visibility** BD brings to certain communities might not be a good thing. The benefits and risks of big data are not spread equitably across society.
- **Traveling algorithms**: As algorithms are adapted to new purposes we should expect to see mission creep with unexpected/unintended consequences.
- **What sort of society** is BD leading us toward? We shouldn't focus too tightly on concerns about ethical practice, we should remain attentive to the big picture.

- **By poking systems** you modify systems. What are the ethical dynamics of making small modifications and tests to data systems, especially compared to the ethics of hypothesis-driven human experimentation that is much more familiar.
- **Displacement of moral authority** onto the algorithm is a major risk. 'The algorithm did it' should not be a plausible response to ethical problems, someone must be accountable and we should help establish models of responsibility.

Actions:
- **Credentialing in professional organizations** could be a plausible option for expanding the scope of data ethics.
- **Avoid ethics as compliance paradigm.** How do the outputs of this group not just become obligatory checkboxes for researchers to pass through?
- **Teach BD practitioners** how to align internal goals of accuracy with broader ethical goals.

Group 3:
Ethical conundrums:
- **Computer science researchers historically see all problems as a matter of breaking through constraints.** However, ethics is often framed as a matter of adding in constraints. If we try to squeeze ethics into accuracy, it might be perceived as a constraint to break.
- **Ethics refers to a multiplicity of things:** compliance, deep/nuanced issues, and practice. Need to persuade practitioners they actually care, and we know they care, about the deeper issues, not just compliance.
- **Tension between individual choice models and collective issues.** Ethics discourse is often framed such that the individual has all of the responsibility and risk. How do we think about collective governance instead?
- **Informed consent and extreme difficulty of disparity** between opacity of systems and the ability of engineers to manipulate systems and repurpose data.

Actions:
- **Adapt the ELSI model from genomics.**
    - **3-5% of basic research funds** dedicated to ethical, legal and social implications research.
    - **Think carefully** about how research is funded and how we might leverage that to get out of compliance culture.
        - ELSI route suggests that empirical and ethical research are separate issues
            - How do we subvert this problem?
            - Nancy Tuana has a pedagogical exercise called 'ethics spotting' that asks participants to find nuances in a seemingly straightforward research project.
            - ELSI project managers did try to fund researcher with *in situ* ethicists with mixed results.
- **Think critically about ELSI model** of funding and how to build ethical cultures of BD practice.
- **Is BD human subjects research?** Big data has brought a range of new communities into contact with computer science, do they think they are doing human subjects research and are they required to use IRBs? Are IRBs a plausible, if problematic, point of leverage?

Group 4:

Ethical conundrums:

- **Reproducibility of Results vs. Ethical Principles**
    - **Data moves too fast** for researchers to reliably reproduce others' results.
    - **Withholding data for privacy reasons** is often at odds with the ethos of openness preferred by data practitioners.
    - **Visibility is not always the best outcome.** Sometimes protecting a community requires obfuscating data. For example, medical and social research into transgender issues is regularly written in Dutch to hide it from US social conservatives.
- **Regulatory Arbitrage.** Data practitioners and companies can try to move between regulatory realms to avoid ethical scrutiny. For example, 23andMe claiming to be an entertainment service, and not medical prediction, to avoid FDA regulations.
- **Impoverished Epistemology of BD.** Big data has received a lot of hype for its potential, but there are some major epistemic shortcomings in this hype.
    - **Causation vs. Correlation vs. Prediction:** Big data lends itself to easy slippage between these different categories of knowledge that should have radically different standards. For example, knowledge gleaned from datasets can only be postdictive over that dataset, not predictive of behavior outside of the dataset. Predictive analytics is really just postdiction, and it is impossible to account for all the future confounders. Similarly, algorithms are superb at finding correlation and rather poor at pointing toward causation.
    - **Hypothesis-free science:** Datasets that are large enough test the edges of hypothesis driven science. Is an algorithm let loose on a large enough dataset equivalent to a hypothesis? Or are algorithms self-fulfilling prophecies?
    - **Predictive leap between individuals and groups.** It is inaccurate to predict future individual behavior based on the behavior of a past dataset.
- **Powerful actors get to create and manipulate data systems.** Data is a social construction because phenomena requires our attention to qualify as a data. Anthropologists have long noted that we typically only end up studying the powerless—our attention is turned toward them as a 'problem.' How do we use data to make the powerful 'a problem' worthy of intensive study?

Actions:

- **Consideration of what the Council should try to accomplish.**
    - Produce principles/code (a la the Belmont Report) that encompass a broad swath of data practices.
    - Arbitration a major conflicts that rise to public attention.
    - Examples of groups trying to play a similar role: Bioethics Commission, WH Privacy research agenda
- **Trustworthiness of Council.** How can the Council establish itself as a trustworthy actor in this field? Whose trust do we need? What is the audience of BDC: the research community, industry, the public and other broad stakeholders?

Ethics codes: history, context and challenges

Kate Crawford lead this portion of the schedule. The pre-reading <LINK> was a historical survey and summary of major professional ethics codes in biomedicine, journalism and computing.

Crawford's introduction to this section raised the question of how the Council might materially engage with professional and research ethics codes about the uses of big data, particularly for NSF projects and academic practices more broadly. Big data practices are a field where there is little by way of explicit ethical frameworks. What can we learn from the history of ethical codes, principles and debates that could inform the council's position here? Importantly, there is some urgency around the demands for more detail about data ethics, especially given controversies around the meeting of academic and industry practices.

Crawford noted that most influential ethics codes are hard-won responses to major disruptions, especially scandals. Such disruptions re-open questions of responsibility and trust, and drive calls for codification of principles, best practices, and what is seen as problematic. The big periods for ethics code include post World War II and the social upheavals of the 1960s-70s. (e.g., Tuskegee, Willowbrook)

There are several principles that can be found at the core of contemporary ethics codes across many domains:
● respect for persons
● balancing of risk to individuals with benefit to society
● careful selection of participants
● independent review of research proposals
● self-regulating communities of professionals
● funding being dependent on adherence to ethical standards

The major ethics codes of computing societies, such as the ACM & IEEE, are over 20 years old, and many argue these professional societies aren't keeping up with quickly shifting norms and technical capacity of the Internet and data-intensive society. What can we, as a council, do to address these gaps, in light of recent scandals?

Crawford noted that ethics codes inevitably raise questions of enforcement. Some professional organizations revoke certifications on the basis of unethical behavior; others that have an academic membership have procedures for revoking the privileges of membership. But there is currently no expectation for people in the technology sector to be part of a single professional community, and big data practices and issues are spread well beyond the tech industry as well.

In the US the biggest instrument available for enforcement is the provisioning of federal research funding, and it has been used mostly widely in biomedical and behavioral research. This depends on meeting Institutional Review Board standards for human subjects research. But as the Facebook emotional contagion study experiment demonstrated there are questions about how IRBs are positioned to manage these issues.

Finally, enforcement may not be as important as it first appears. Many of the purposes of professional ethics codes to affirmatively create a community with common values. By discussing and defining the scope of acceptable activities, establishing expectations and articulating values, a process around ethics defines the boundaries of a community.

Crawford asked attendees to consider some questions in breakout groups:
- Process vs. Principles vs. Code: all or some or one?
- Could a 'data ethics plan' be built into grant applications, similar to the existing requirement of a 'data management plan'? If so, what would it address?
- What broader role should we be taking in academia and/or in industry?
- What can professional organizations do? Is a separate body needed?
- What about enforcement?
- Is there a specific step that the NSF can take? What about academia and industry more broadly?

Breakout group discussions followed these major themes:

How to successfully discuss ethics within and across research communities

Several groups emphasized the need to listen closely to research communities about their ethical practices and dilemmas prior to formulating any code or set of principles. Rather than being an external group trying to control the conversation, we need a lively way to narrate ourselves into the practices of big data and computer science. Ethics needs to be a co-constitutive process between those of us trained in ethical and social dynamics and those of us deeply involved in technical processes. Helping scientists and engineers with specific dilemmas that arise from their lived experience would be a good way to generate collaboration in the long run.

There was discussion of the need to rethink ethics training in science and engineering. Ethics training is often perfunctory and focusses on compliance. We see a need to train scientists and engineers need training in the ability to read ethical context, local for local harms and see nuance. Ethics training done well builds capacity to see problems, not just the ability to follow rules. One group formulated this as the need for data practitioners to 'have a nose for creepy.' This is not contrary to codifying principles, but the process and practice is just as important as the codes that are generated.

One group also noted that while we often frame these issues as social scientists talking to/at scientists and engineers, there is actually a significant need to examine ethical practices in social science research. For example, the Facebook emotional contagion study was a research scandal for both data practitioners and social scientists. And research communities can have wildly different senses of ethical principles and conceptions of harm, even within a discipline.

Codes & Principles & Process

One group felt that principles alone are too vague, and a code without principles is empty and inevitably compliance-oriented. Others felt that we were not well placed to produce codes, but are well placed to promulgate principles and practices. Furthermore, different domains may want different kinds of

engagement from the Council. For example, business interests generally want clear rules to follow, whereas academics want maximal freedom to interpret principles. What would a process be like that would allow us to navigate between these poles? The Council would need to handle trade offs between tackling substantive issues of the day and avoiding the risk of calcifying the current news cycles within generalized principles.

Whatever ethics codes or principles are produced by the Council should attempt to break out of the individualist mode. It must articulate some degree of collective responsibility and offer consideration of who gets to define respect, benefits, harms, etc. There may also be good reason to formulate a consensus process that can deal with power differentials.

Algorithms are a likely focal point for the Council's efforts on this front. Developing principles and practices for the transposition of algorithms across domains was offered as an example. Moving between domains raises issues of generalizability and specificity—can an algorithm built for one domain adequately address another. This raised the classic joke of a physicist attempting to improve milk production on a dairy farm. She tells the farmer, "Let us first assume a spherical cow."

What can we accomplish at NSF?

Several members argued that the Council should make improving data ethics policy and practices at the NSF our top priority. This is where the Council has the most influence, and would in itself be a major achievement.

Many groups touched on data management plans (DMP) as a point of leverage within the NSF. After DMPs were instituted as a required component of all grants with the past decade, scientists came to realize DMPs were more challenging than first imagined. It is easy enough to store data, it is much harder to determine how to share it and how to balance competing ethical concerns about data, such as the conflict between mandates to share with fellow scientists and simultaneously protect the privacy of research subjects. This creates a concrete need in research community for advice from an entity like the Council. Several attendees asked if the Council can we get from the NSF concrete examples of scientists struggling to balance competing criteria in DMPs.

Broader Impacts were also proposed as another point of leverage. The Broader Impacts criterion is ostensibly co-equal with the Intellectual Merit criterion for NSF grants, and is meant to require applicants to identify how the proposed project can positively affect certain social goods. Now that data scientists should at least consider themselves doing human subjects research, how can we expand their sense of responsibility? Presently, successful applicants to the BIGDATA program appear to not take this requirement particularly seriously, with some indicating that general scientific progress is an adequate Broader Impact for the basic science proposed. Other programs would typically have a much higher bar for specificity of Broader Impacts. The appropriateness of requiring basic science projects to articulate their connection to broader social goods is a topic of public policy discussion in recent years and this may be a relevant place for the Council to act.

Several groups noted the need to marry data scientist's pragmatic need for concrete advice on ethically-relevant research issues to the aspirational move of wanting to do good in the world. The Council could frame our work with the NSF as helping to produce better science through better ethics. Along these lines, we would also hope to short circuit the ability to outsource ethics from the empirical project. For example, at Johns Hopkins researchers are able to essentially hand over their DMPs to the university library and have little incentive to engage with the deeper issues behind DMPs. Other universities offer similar services with Broader Impacts.

Enforcement (and other models for addressing ethics)

Also focusing on the NSF, one group proposed a clear actionable plan: persuade the Computer and Information Science and Engineering (CISE) Directorate at the NSF to functionally address big data work as human subjects research and start requiring review by IRBs. This proposal was contentious. What we assume to be human subjects research may not easily fit into big data research processes in many cases, even if there are ever more edge cases.

Among the reasons this would be challenging:
- Data science practices notably slide between algorithm tweaking and hypothesis-testing science. Where exactly on that scale would they fall under IRB oversight?
- There is also an acute sense of competitive disadvantage of falling under human subjects protocols that would provide a disincentive to work on human data.
- Many questions whether university IRBs are the correct model of oversight.
- Data scientists operate on a spectrum ranging from creating algorithms to building products, and only a few points on that spectrum are traditionally addressed by IRBs.
- Latanya Sweeney pointed to her work on re-identification as a model of data research that confounds IRBs. She highlighted major problems with assumptions that had shaped genomic and medical privacy protections required by many IRBs, but it could never have been approved an IRB.

Reasons offered why IRBs might be a reasonable route:
- If a funding/regulatory agency comes out in favor of such oversight, universities will have a lot of incentive to make the right resources available to mitigate much of the concern about competitive disadvantage.
- University IRBs are not the only model. Private/independent IRBs tend to be more collaborative and less compliance-oriented than university IRBs because they have different incentives.
- It is ultimately easier to get an institution to incorporate new standards within existing mechanisms than to create new mechanisms.

Groups also discussed the consequences of institutional non-compliance with ethical standards established by funding agencies. Routine violations result in 'ethical receivership' with substantial outside control and upscaled standards with new restrictions. Historically, the universities getting the most federal funding ($1B+) are found with the most violations.

Other routes for addressing ethics that de-emphasize enforcement were proposed. Several proposals focused on the Council acting as a research ethics consultation service. Some government agencies proactively offer advice on certain financial transactions at the request of the parties—if researchers were uncertain about the advisability of certain practices it would be useful for them to have somewhere to turn other than IRBs.

Empirical Research Ideas

Geof Bowker and Jake Metcalf lead this session. Time was short so not all of the proposed goals were covered. This session was intended to raise the question of how the Council could coordinate empirical research agendas around the many issues discussed in the meeting. Bowker and Metcalf covered what they considered the main thematic areas developed during the meeting and introduced their ideas for empirical projects. There was inadequate time to discuss in depth how Council members might coordinate research projects.

Major thematic areas discussed:
- Algorithmic accountability
- Blurring Research and Practice, Science and Entertainment
- Immutability and Blurring of Records
- Experimentation and A/B Testing
- Networked Consent
- Epistemologies of BD: Correlation, Causation, Prediction and Spherical Cows
- Models of Oversight: IRBs, watchdogging, consultations, arbitration DMPs, Broader Impacts, ELSI, RAC, etc.
- Reproducibility of Results vs. Ethical Principles
- Power and responsibility: individual responsibility, collective culture/governance, addressing inequity and tracking powerful actors
- Repurposing data: opacity of systems and informed consent
- Reproducibility of Results vs. Ethical Principles
- Traveling algorithms

Bowker and Metcalf have proposed studying the ethics of prediction using two case studies. Their major interests are:

- Prediction is performative in the present.
- Do predictive algorithms confirm themselves?
- On the surface, some of these algorithmic interventions appear to be win-win, driven by 'neutral' data.
- What ethical issues are present when prediction shapes our common experience?
- How do we build accountability and transparency as ethical norms around prediction?
- How do algorithms travel? What ethical/social issues are carried across/with predictive practices?

For their case studies, they will be examining 1) a predictive policing firm that uses an adapted earthquake prediction algorithm, and 2) a BIGDATA project that is building a medical-outcomes prediction algorithm using ~80M lifetime medical records. Council members shared suggestions for building these research projects.

Closing Session

Helen Nissenbaum led this session. She noted that the original proposal was from an EAGER grant, but that the Council was encouraged to imagine something larger scale and treat the EAGER funds as a seed. What kind of impact could we achieve with these resources and how might we scale?

The original proposal did not specify the Council's activities in great detail, beyond convening the group present today, hosting meetings with and for BIGDATA PIs, building a public plan to engage with public controversies, and writing white paper(s). Nissenbaum framed some questions that will guide how the Council will develop as an organization going forward.

Should the Council serve as a group that impacts/guides:
- research community
- the public/broad stakeholders
- those specifically impacted by BD research

What are we here to accomplish?
- excavate and identify substantive ethical issues
- vet research practices
- create principles/procedures
- pursue empirical projects together

She invited members to discuss what might be useful given this very interdisciplinary group. Contributions are identified here by initials.

- FP: There are radically new opportunities opened up by big datasets and techniques. Who has access? Are there guidelines available prior to IRB stage? We should work through case studies on this contextually because precedents build up by digesting cases.
- AN: The Council should provide the public service of giving people a sense that things aren't running out of control. This alleviates a potential sense of futility, victimhood to a new technology.
    - HN: Are we catching stuff ahead of time or dealing with post-facto?
    - KC: We could develop early warning-systems by building narratives about problems.
- BK: We should heed warnings about panels of self-appointed experts. Such panels are great for white papers, not so great for voting oneself into self-perpetuation without the support of a broad audience.
- MZ: We need to pay attention to difference between academic, industry, government and public interest structures. Are we able to nimbly provide a reaction to problems?
    - db: Be careful about saying BD work happens in delimited organizations, because the work spills over the boundaries, lots of different people/orgs want advice about how to do research well.

- LS: We don't yet know who we are talking about/to. Someone developing an algorithm will probably just ignore us. We won't have any standing in small companies. Caution is warranted if the role is only in the space of post-facto advising. On the other hand, we have great leverage at NSF/funding agencies who can create ripple effects. Go after low-hanging fruits, like tweaking DMPs, notwithstanding how we might comment on more public issues.
    - RH: We can be optimistic about low-hanging fruits at agencies. But we will not be able to address a sense of things out of control.
- GB: Is there a structural role for something like BDC? By law, IRBs don't address social obligations; they only address potential harm in individual cases. Building a capacity to address the social is a critical need. We need to be able to engage the people writing the code, and we are funded by CISE. It would be a big win to get discussion launched at NSF about changing DMPs.
- LS: We need to be aware of the risk of mutual eye-glazing in conversations with computer scientists.
    - db: How do you structure collaboration? Can't just focus on funding, need to tease out processes of teaching and embedded ethical research. What is the intellectual content that we need to do collectively can we map out how that could be done successfully? Maybe it is not one road map, but a series of trials for different configurations. We will only have succeeded if we can explain that.
- FP: It might not be that hard to come up with consensus principles for a community.
- KC: Identified three outputs for the Council so far that sound promising:
    1) *Focused, pragmatic intervention at NSF*.
    2) *Group workshop for PIs*. Rather than mutual eye-glazing we listen to what raises ethical questions from their research/work.
    3) *White paper* to reflect what we've learned rather than telling them what we think. This would become a foundational document going forward
- AN: Similar promising dynamic for listening in genomics, especially around race. This is a model worth thinking about. Becomes a feather in their cap to do this.
- LS: The technology ethics experience at Carnegie-Mellon has lots of lessons for the Council. Right before 9/11 people were dropping face recognition like a hot rock due to a lot of negative ethical attention on that technology and on Carnegie-Mellon specifically. After 9/11 the money poured back in. So-called 'privacy preserving data mining' was the 'solution' to those ethics complaints which ultimately did nothing to protect with privacy. A similar dynamic can be see with link analysis.
- SB: It is interesting to see how hard it was to solicit meaningful answers when attempting to start a collaboration or make an intervention because of how much practitioners get drilled on proper discourse. You get ready answers rather than productive disclosures.
- db: The conversation that happens at the highest level is pretty irrelevant to the micro-practices of coding where there are lots of tradeoffs. For coders the meta-conversation is irritating, we think meta-first. for us, writing is thinking. For them, writing is a record of thinking.
- AN: There is an analog to the green economy. The processes cost more, the product cost more, etc., but there is an advantage to engage with the politics of it. We need to make our constraints add value by engaging elite data scientists.
- HN: Ethics add value because it engages in the things the scientists already care about.

- GB: We can't hope to parachute in. As a Council, we should take intervention off the table and instead think about engagement.
- KC: There's a different dynamic to being a council inviting researchers to find value in engagement than doing social science research.
- LS: We don't have to make the green economy argument; we only have to make the NSF argument.
- LS: I can't use the term 'data science' anymore because I don't know what it means. Maybe CISE isn't the right body at the NSF, but rather look to social science. For the CS community the goal is to not be disrupted. The goal is to build something that solves the most problems, no one cares about special cases. Addressing a special case or building constraints is considered bad computer science.
- RH: The core of BDC will need to do a lot of work with the BIGDATA PIs to get buy-in. Get cases in advance of meeting. How do you get at ethics in technical projects? Think about risk prediction built into algorithms. Balancing false positives and false negatives.