Meeting Synopsis:

1. Call to Order
2. Approval of the Minutes from May 7, 2014
3. Announcements
4. Data Science Project
5. Sharing RCR for Collaborative Research (Use of Subaccounts)
6. Adjourn

1) Call to Order

The meeting was called to order by Chair Miller at 9:00 a.m.

2) Approval of the Minutes from May 7, 2014

The minutes from May 7, 2014 were not approved due to lack of quorum.

3) Announcements

Miller announced that he is rotating off a chair of FCR. Rosenfeld will chair the council next year.

4) Data Science Project

Ed Lazowska presented the UW data science project at UW. Lazowska argued that data science has led to exponential improvements in technology and, such as:

- A proliferation of sensors
- Ever more powerful models producing data that must be analyzed
- The creation of almost all information in digital form
- Dramatic cost reductions in storage
- Dramatic increases in network bandwidth
- Dramatic cost reductions and scalability improvements in computation
- Dramatic algorithmic breakthroughs in areas such as machine learning that might lead to a revolution in discovery.

Lazowska explained four paradigms of data science:

1. Empirical + experimental
2. Theoretical
3. Computational
4. Data-Intensive
Lazowska explained the potential importance of the ability to extract knowledge from large, heterogeneous, noisy datasets.

Lazowska discussed the goals of the UW eScience Institute and how it would ensure UW is a leader in advancing data science techniques and technologies. The History of the eScience Institute was presented.

The goal of data science at UW is to do breakthrough science in scientific theme and data science methodology areas, enable breakthrough science through new tools and methods, and establish a “virtuous cycle”:

- **Scientific theme areas**
  - Biological sciences
  - Environmental sciences
  - Physical sciences
  - Social sciences
- **Data science methodologies**
  - Machine learning
  - Data management
  - Data visualization/usability
  - Statistics
  - Sensors
  - Programming environments
  - Scalable hardware & software systems
- **Bridges that connects discoveries**
  - Career paths and alternative metrics
  - Education and training
  - Software tools, environments and support
  - Reproducibility and open science
  - Working spaces and culture
  - Ethnography and evaluation

**Key activities of the eScience Institute**

**Promote interdisciplinary careers**

- Interdisciplinary graduate students
  - New, interdisciplinary “Data Science” Ph.D. tracks and program
- Interdisciplinary postdocs (“Data Science Fellows”)
  - Dual-mentored postdocs with interests in both methods and a domain science
- Interdisciplinary research scientists (“Data Scientists”)
  - Work across disciplines to solve people’s data science challenges
- Interdisciplinary faculty
  - Supported with special hiring and funding initiatives
- “Senior Research Fellows”
  - Short-term and long-term visitors
• A diverse faculty steering committee

Re-establish the “watercooler”

• An open shared R&D space where researchers from across the campus will come to collaborate
• A resident data science team
  o Permanent staff of ~5 Data Scientists – applied research and development
  o ~15-20 Data Science Fellows (research scientists, visitors, postdocs, students)
  o Entrepreneurial mentorship
• Modes of engagement
  o Drop-in open workspace
  o Studio “Office Hours”
  o Incubation Program
  o Plus seminars, sponsored lunches, workshops, bootcamps, joint proposals

Create scalable impact through a Data Science Incubation Program

• Scale and concentrate efforts
  o Move from “accidental” encounters to engineered partnerships
  o Identify emerging opportunities around campus
  o Provide a shared environment where researchers can learn from an in-house team, external mentors, and each other
• A startup environment!
  o “Seed grant” program
    ▪ Lightweight – 1-page proposals
  o Significant potential for technology spinout – new markets for existing technology and new technology for existing markets

Foster interdisciplinary education

• Ultimate goal: A new Ph.D. program with interdisciplinary cohorts!
  o Initial goal: Big Data Ph.D. tracks in multiple departments
  o Education highlights: Data science courses, co-advising, and internships
• End-to-end research agenda
  o Big Data management, analytics, modeling, and collaboration
• Cyberinfrastructure development
  o Big Data analysis service
• Coursera MOOCs
• Traditional courses
• UW Educational Outreach
• Workshops and bootcamps

Advance the state-of-the-art in data science

• In methodology areas
  o Data management, machine learning, visualization, statistics, etc.
• In domain sciences
Astronomy, oceanography, earth and space sciences, biology, neuroscience, genome sciences, sociology, etc.

- In best practices
  - Reproducibility
  - Open science
  - Ethnography

Example projects include:

- AstroDB: Cosmology at Scale (Andrew Connolly – Astronomy)
- Role of Microbes in Marine Ecosystems (Ginger Armbrust – Oceanography)
- Devices + Neuroscience + Data Science

Lazowska envisioned additional funding from foundations to increase the number of data scientists and post docs are the project would become a perpetual motion machine which brings in more grants and researchers. A comment was raised expressing excitement for the project as managing large data sets can be difficult.

Aaragon discussed her personal experiences as a data scientist at UW.

5) Sharing RCR for Collaborative Research (Use of Subaccounts)

Lidstrom reported on her recent efforts to discuss concerns surrounding faculty collaboration on research projects. A committee working on this issue went to the ABB steering committee requesting a strong supporting statement from the deans to endorse this proposal. While it is easy to collaborate within one’s own academic unit, it becomes much more difficult when collaborating across campus. Two common issues arise from collaboration:

1. Sharing of resources
2. Sharing of credit

The committee has already developed a guidance document for large collaborative efforts but Lidstrom has developed a document to expand to all types of collaborations. The basic principle is that resources should be shared between units. Additionally, the guidelines are to allow for flexibility when unique situations arise. Lidstrom explained that there are three categories of resource sharing:

- Subaccounts. While this is a clean way to share resources it does create administrative burdens.
- In-kind support
- Look for long-term balance of shared resources between units. This is common with the School of Public Health and the School of Medicine which have collaborated together for years. In looking back at the resources shared the two schools had practically balanced out.

Members discussed the hiring of new faculty on collaborative projects. Lidstrom stressed the need to push for this document to deans. Discussion moved to shared credit and the impact on promotion and tenure committees. Lidstrom explained that those committees should endorse this as a principle as well. Credit will be reflected into EGC1 which will list all other investigators on a grant. A question was raised about NIH funding and the complications calculating indirect costs. For example, a PI may have an
indirect cost rate of 77% while co-investigators have 53%. Lidstrom explained that the School of Medicine has a policy in which the indirect costs follows the indirect cost rate of the PI.

Concern was raised to have strongly-worded guidelines that clarify what is expected from departments. A comment was raised stressing concern about the large power differential between PIs and co-investigators. Lidstrom reemphasized the importance of having a good endorsement from deans. In order to vet this proposal Lidstrom will be meeting with several stakeholder groups to build support including FCR, council of associate deans, and other faculty groups.

Discussion moved to collaborative teaching and funding for interdisciplinary groups. One example is a possible minor in food studies between different units. While there are program overlaps between existing programs it could create a turf war. One could see funding mechanisms and incentives to allow for this but each dean makes their own decisions about what is important for their department.

A comment was raised about UW graduate students who attend UW because of the institution’s emphasis on collaboration between departments. Faculty have a tendency to fund their own PhD students which reduces the incentive for interdisciplinary students. Lidstrom emphasized the importance of developing a strong business plan before implementing new interdisciplinary programs. A question was raised asking if units could coordinate joint recruitment of new faculty and graduate students. Lidstrom suggested inviting Dave Eaton from the Graduate School to discuss this issue. Lidstrom commented that the 80/20 funding model needs to be revisited which inhibits collaborative work between units.

Lidstrom emphasize the importance of the council putting pressure on her committee to address concerns surrounding collaborative research. Lidstrom hopes to implement her proposal later in Fall Quarter 2014.

6) Adjourn

The meeting was adjourned by Chair Miller at 10:30 a.m.