

CaRDS -Meeting

Minutes

September 15, 2022, 2:00 pm

Present: Brian Bennett (Computer Science), David Currie (Information Technology Services), Nick Hagemeyer (ORSPA), Frank Hagelberg (Physics and Astronomy), Dana Harrison (Management and Marketing), Jeff Knisley (Math and Statistics), Robert Price (Math and Statistics).

We had a conversation with Nick Hagemeyer, Vice Provost for Research and Chief Research Officer, centered on the new master's degree in applied data science (MSADS) and high performance computing (hpc) at ETSU.

Main conversation items:

- Possible interactions between ORSPA and the MSADS, related to identifying funding sources:

Future growth of the program may require adding graduate assistantships, administrative support, and computing resources. This could be achieved through (a) external grants or (b) support by local corporations.

- (a) Proposals addressed to major funding organizations, such as NSF or NIH, fare better if submitted by collaborations, rather than individuals. ORSPA could be instrumental in connecting the ETSU data science initiative with on-campus institutes or programs that might benefit from the data science expertise associated with the MSADS. In this context, it may be worthwhile to explore opportunities for interacting with the Center for Rural Health Research or the Addiction Science Center, both in the College of Public Health. The planned research partnership between Ballard Health and ETSU might open further options for cooperating, specifically submitting joint proposals. Nick mentioned that he attends the periodic meetings between Ballard and ETSU representatives. He'll inquire about potential roles of MSADS students and faculty in a future research collaboration between Ballard and ETSU.
- (b) ORSPA might assist the MSADS in initializing contact with local companies that could be interested in supporting our program, e.g. donating graduate assistantships. Local business may also be encouraged to fund employees who wish to enroll in our program.

- Advancing and promoting the MSADS program:
 - Aligning the MSADS core curriculum more closely with the demands of the focus areas. Specifically, this will include exposure to the data management and visualization tools *Tableau* and *Power-BI* for students who select business as focus area. Tableau is used for data transformation, visual representation, and generating reports for real-time analysis. Power BI, developed by Microsoft, provides cloud-based BI (business intelligence) services and is, like Tableau, routinely used by corporations. Fluency in handling these instruments will be a plus for all MSADS students who seek careers in industry.

- Increasing program visibility *intra* and *extra muros*:
The planned 3+2 option for undergraduate mathematics or computer science majors should be advertised in local high schools through electronic (email) announcements and personal visits.
ORSPA might assist CaRDS in alerting ETSU departments and centers, as well as corporations and administrative or service agencies off-campus, to the data science expertise at ETSU, and specifically the MSADS.

- *e-day* in the academic year 22/23

We agreed to start preparations for our annual *e-day*, highlighting computation and data science in and around ETSU with a program featuring external speakers and ETSU student presentations, in this fall semester. Nick is open to the idea of integrating *e-day* into the 22/23 *ETSU Trailblazer Summit* this academic year, showcasing ETSU Research and Innovation.

- CaRDS seminar 22/23

Jeff presented ideas for future realizations of the CaRDS seminar that might heighten the visibility of data science at ETSU and contribute to advertizing our MSADS. In summary, Jeff envisions combining seminar presentations as we know them – data science practitioners give talks for ETSU students and faculty about topics of their expertise – with a public outreach component. More specifically, three types of novel seminar events are proposed:

- (1) Talks on topics of DS/ML/AI, delivered by speakers who are interested in and capable of addressing general audiences. A speaker would deliver a traditional seminar talk, followed by a popular talk directed at the community. Ideally, there would be some coverage of the latter event by the local press.
- (2) *Lightning talks*: seminar sessions divided into two parts – in the first, several faculty members (maybe four or five) give short presentations, in the second, the speakers engage the audience in a conversation about the presented materials.
- (3) *Tell-me-more talks*: the community would be encouraged to ask questions about DL/ML/AI. We'd design a letterbox for that purpose and identify local experts maximally suited to address the questions. Questions and responses would be presented at special events that would also allow for spontaneous questions from the floor. This outreach activity might address the great curiosity of the public about matters of DS/ML/AI and yield, at the same time, a good opportunity to dispel numerous – and persistent – myths that cloud the general perception of the field.

- Current and future access of ETSU faculty members and graduate students to high performance computing (hpc) resources

At this juncture, ETSU researchers in need of hpc equipment benefit from ETSU's partnership with Amazon Web Services (AWS). AWS provides on-demand cloud computing platforms to individuals, companies, and administrative units. In particular, the Amazon Elastic Compute Cloud (EC2) that allows users to run their own computer applications on virtual computers, is open to ETSU researchers on a pay-as-you-go basis. A modest

amount of computing funds is free of charge. Recent extensions of the AWS offer include Appstream, a cloud service that enables users to run software not available to them on their own computing devices.

Until recently, ETSU had a partnership with XSEDE (*eXtreme Science and Engineering Discovery Environment*), an international network that encompasses shared computing resources, data, and related technical know-how. *XSEDE pilot allocations* provided ETSU researchers with up to 50,000 core hours at national hpc centers. This initial step could be followed by a start-up allocation of up to 200 000 core hours. The XSEDE project has ended on August 31 2022 – it will be interesting to explore the successor organization ACCESS (AdvancedCyberinfrastructure Coordination Ecosystem: Services & Support).

The ETSU Research Development Committee (RDC) has supported AWS-based computing. Nick mentioned a possible restructuring of RDC – about five priority research areas would be specified. These would be eligible for preferential funding. Among these areas might be hpc based research.

The meeting adjourned at 3.10 pm.