Board of Regents Advisory Committee on Data Science
Fall 2021

October 5, 2:00 pm – 4:00 pm

Minutes

Present

Member	Institution
Erin Porter	Abraham Baldwin Agricultural College
Faisal Mirza	Atlanta Metropolitan State College
Neal Smith	Augusta University
Kristin Seamon Lilly	Columbus State University
Thomas Gonzalez	Dalton State College
Sebastian Siva	Georgia Gwinnett College
Xin Xu	Georgia Gwinnett College
Melanie Largin	Georgia Highlands College
Joel Sokol	Georgia Tech
Sherrill Hayes	Kennesaw State University
Barry Monk	Middle Georgia State University
Abhinandan Chowdhury	Savannah State University
Shetia Butler Lamar	Savannah State University
Radu P. Mihail	Valdosta State University
Jonathan Watts Hull	System Liaison

Guests:
Barbara Brown, Assistant Vice Chancellor, Transitional and General Education, University System of Georgia
Tristan Denley, Executive Vice Chancellor, Academic Affairs and Chief Academic Officer, University System of Georgia
Gene Pierce, University of Georgia, Georgia Informatics Institute
Art Recesso, Chief Innovation Officer, University System of Georgia

Jonathan Watts Hull opened the meeting at 2:00pm.

BACKGROUND ON A REGENTS ADVISORY COMMITTEE

Barbara Brown provided an overview of the role and function of a Regents Advisory Committee (RAC). She noted that a RAC has two main functions. First when the system needs a discipline-focused voice on a matter, typically a course for approval in the Core, or less frequently for the consideration of a course as a part of the required high school curriculum, it will seek input from the respective discipline RAC. The second, but still central, role is to serve as a point of contact and exchange for ideas and issues in the field among faculty in the System.

Dr. Brown observed that the Data Science community has already been instrumental in the development of the first "templated" course – DATA 1501—which serves as a model for other similarly structured courses. A question was asked about how a templated course differs from other course approvals, specifically MATH 1401. Normally, when any course is added to the curriculum, the institution has to submit a full proposal for approval to the Council on General Education and then is routed to the discipline RAC. Templated courses do not have to satisfy the full requirements so long as they adopt the requirements pre-approved by the RAC, facilitating adoption and approval.

ELECTION OF OFFICERS AND ORGANIZATIONAL DISCUSSION

Sherril Hayes agreed to be nominated to serve as chair of the committee; Shetia Lamar agreed to be nominated to serve as vice chair and recorder. Both candidates were elected unanimously for a term of one year.

A draft set of Bylaws for the Committee will be distributed by email with these minutes. Members are asked to review them and make any suggested changes by October 28. A vote on the recommendations to the bylaws will be conducted the following week, with an electronic vote conducted on Friday, November 5.

DATA 1501 UPDATE

Jonathan Watts Hull (USG) provided a brief background on the DATA 1501 course, Introduction to Data Science. The DATA 1501 course as a commonly numbered, Area D course emerged from the development of the Statistics Pathway with the move of MATH 1401 into Area A2 from Area D, which created a hole, and an opportunity, to build a course to really make a true statistics pathway. The course was built from an already approved course from Coastal Georgia in a collaborative fashion with faculty from a group of interested institutions. Dr. Brown provided the list of institutions where the course has been approved for Area D in the "Related Documents" Section of the Data Science RAC page on the USG website.

Kennesaw State University (Sherrill Hayes) indicated that KSU is using Data 1501 as an extension of the statistics pathway for non-STEM majors. Sixteen undergraduate majors have signed up for the MATH 1401 & DATA 1501 pathway. Initial projections of the enrollment for DATA 1501 was higher than actual enrollment because the courses is part of a natural sequence that students begin with MATH 1401, shifting most enrollment to spring term. KSU has also struggled to find a textbook, and is using a business analytics text that is not a perfect fit. The biggest hurdle thus far is that students are having trouble with the software. The course focuses on using Excel as that is what is most likely to be used by these students later in their careers. Faculty are having to teach about Excel as they are teaching early course concepts. Students are reporting very positive things about the course – saying that is is the best version of a math course they have taken. Structurally, not diving into the technology right away can help students who may not have a background in coding understand the reason why these skills are worth learning because it provides value to these skills.

Georgia Gwinnett College (Sebastian Siva) offers three Data Science related courses MATH 1401, DATA 1501 and MATH 2050 for STEM majors. There has been an evolution of these courses over time. There has been more of a push to move to R and there is no standard path and some stepping on toes with the three courses. The teaching of Python to non-STEM students is something that can be done. The domains that are used to motivate students are usually something students can associate with, such as music, which can make things easier. It was suggested that an introductory data science course is possibly not the place to teach programming. Thus, at GGC, Data 1501 is taught in Excel, which is a valuable skill for freshman to learn.

Paul Radhu Mihail (Valdosta State University) provided background on the Valdosta State implementation. His work started with a topics list and technology set and is built as a blend of python and Excel. VSU also is using Collaboratory on Google to give access to an interactive python environment to allow students to play with tools and libraries. Even though there are no prerequisites for coding, doing so exposes students to basic operations using Python and allows students to understand how programming works without having to teaching programming.

Out of a concern that the need to calibrate to an unknown student population (both in terms of their interests and goals –the course had been advertised to the college of Humanities), VSU only designed the first two weeks of the course to provide an opportunity to recalibrate on the fly. The course is structured to be useful to students in fields like psychology, so ensuring that the material aligns with the interests of the students is also important.

Some things that have been learned this year: students struggled with typing (needed help with tilde, back and forward slashes, etc.), which required some recalibration. The 3-credit course offered M-W; Monday is traditional lecture, and Wednesday is a lab day with student working on the labs independently with faculty support. Spent a bit of time on syntax and basics, but the labs are designed to be something the students can follow without understanding the code. This is an experiment in many ways and students may struggle to connect the labs from one week to the next. Students seem to value the labs as a process.

Erin Porter (Abraham Baldwin Agricultural College) echoed the gaps students had with keyboarding, observing that students were struggling with understanding a range of basic computer skills. Students coming into the courses don't understand things like file structures and have gaps in other essential skills. How to address these aspects of essential skills for students is key to moving forward.

Shetia Butler Lamar (Savannah State) observed a freshman student would be more comfortable learning Excel than Python, but exposure is important. Thus some coding practice, even if they aren't writing or developing the code, can be both important and meaningful. The introductory course can't be highly technical but can give students the exposure they need to understand what the results are when they are presented them in their careers. Savannah is taking an interdisciplinary approach to the DATA 1501 course. What they are trying to do is to make sure students are exposed to general information on Data Science and to exercises that

relate to the various disciplines and explore the way that data science intersects with these fields. The course is still in development at SSU, which plans to first offer it in Fall 2022.

DATA SCIENCE KNOWLEDGE DEMAND ANALYSIS

Introducing the System's interest in the area, **Tristan Denley** (USG) said that the field of Data Science was at a critically important time for institutions, and the System is very supportive of it. Most especially on the collaboration on curriculum and resources to facilitate and accelerate the process of course and program development.

Art Recesso (USG) provided an overview of research on data science demand in the workforce. The USG interviewed employers and data scientists in the field to identify essential skills and key areas within the field. They validated industry needs into a framework and are now working on a career continuum. These are not curricular in nature, but focused on competencies, knowledge and skills. The next steps are to align case studies and resources to support courses.

What was learned in this process is in addition to being a high demand area, the field is also highly differentiated, with the greatest demand for data managers and architects. To meet this demand as a state, there is a need for scalable programs and possibly different models. The current consideration is to establish centralized, system-level assets to address knowledge demand, a sandbox and data range for developing student experience working with real data and solving real problems, and opportunities for collaborative research and development with key sectors. Related to this, faculty at UGA are building some collaborative courses that are focused on machine learning to support skill and knowledge development at the Data scientist level. There is an opportunity to extend this approach to other collaborative support. An opportunity that has emerged from this work is a framework for collaborative assets facilitated by the Georgia Informatics Institute at UGA.

The meeting adjourned at 3:57