Envisioning and Realizing a Statewide Data Science Ecosystem

Dr. Karl D. Schubert FIET, University of Arkansas

Dr. Karl D. Schubert is a Professor of Practice and serves as the Associate Director for the Data Science Program at the University of Arkansas College of Engineering, the Sam M. Walton College of Business, and the Fulbright College of Arts & Sciences.

Shantel Romer, University of Arkansas
Stephen R. Addison, IEEE Educational Activities
Tina D Moore
Laura J Berry, North Arkansas College
Jennifer Marie Fowler, Arkansas State University
Lee Shoultz, University of Arkansas
Christine C Davis

Envisioning and Realizing a State-wide Data Science Ecosystem

Abstract

This paper describes the vision, strategy, plan, and realization of a state-wide rigorous data science educational ecosystem. The need for developing data science degree programs and education has been well-established and, in our state, a blue-ribbon panel with industry, academic, and government representatives defined the needs of the state. Additionally, a well-established "think and do tank" published several reports on the importance of data science education and graduates. As we began to develop our programs separately, it occurred to us that we were in a small enough state that, if we chose to do so, we could work together for consistent degree programs. Because we were a small state, we needed to work together to develop quality data science education at all levels. In late 2019, we co-hosted a workshop with representatives from industry, academia, state government, and students at all levels, to explore the potential for developing a state-wide data science educational ecosystem. The response was overwhelmingly "yes." As a result, and with the collaboration of our state's division of higher education, we developed a vision, strategy, and plan to do so with an "opt-in" approach and this paper presents the vision, strategy, plans, results, and experiences and continuous improvement over more than four years of collaboration. These include the development of A.S. degrees in Data Science. B.S. degrees in Data Science, 2+2 programs in data science (A.S. Data Science in a 2-year college and transfer to a 4-year university for the +2 B.S. Data Science with no loss of credits), proposed certificates in data science, a common curriculum state-wide, a high-school data science track based on the common curriculum, and a vision realized of "start anywhere, finish anywhere." Finally, we look to the future in expanding the "opt-in" academic institutions and significantly increasing the number of data science graduates at all levels.

Introduction

The development of a statewide ecosystem in data science builds on earlier statewide initiatives to provide access to education in computer science to all Arkansans. Former Arkansas Governor Asa Hutchinson made headlines in 2015 with the enactment of new legislation requiring all Arkansas public high schools to offer at least one computer science course. At the time, only a few states had implemented such requirements, and Information Technology & Information estimated that computer science courses were offered at only one in ten schools nationwide (Armitage, 2015). The same Arkansas legislation, Act 187, also called for the establishment of a statewide task force to oversee the project and declared the overall lack of graduates with computer science skills to be a public emergency (Arkansas State Legislature, 2015). In late 2020, the Arkansas Board of Education adopted new rules for the 2021-22 school year that emphasized computer science and coding curriculum for kindergarten through eighth grades (Arkansas General Assembly, 2021). This was followed by Act 414, which established completion of a computer science course as a graduation requirement for all Arkansas public high school students.

The incentive for such drastic measures was undoubtedly tied to Walmart's decision in 2012 to locate a $100 million data center in Colorado Springs, CO (Laden 2012). This news, combined with the slow out-migration of operational investments by other key businesses, sparked concerns about the ability to retain companies that serve as pillars of the State's economy. Arkansas has long benefitted from a handful of homegrown Fortune 500 companies like Dillard's, Tyson Foods, and J.B. Hunt Transport Services. Walmart, however, has maintained its status as the world's largest company – and the top spot on the Fortune Global 500 list – for over a decade (Fortune Media (USA) Corporation, 2023). These companies, along with every other major retailer, distributor, and manufacturer, have become increasingly reliant on data analytics and computer programming.

Soon after the passage of Act 187, Governor Asa Hutchinson commissioned a group of public and private sector representatives to guide the State regarding the relevance and need for a workforce skilled in data analytics and computing. The commission members included executives from Walmart, Murphy USA, E-Z Mart, AT&T, Tyson Foods, J.B. Hunt Transport Services, Stephens, Acxiom, and Systematics (the latter two being massive financial analytics companies), as well as leaders from institutions of higher education and relevant state government agencies. The commission published the resulting report in 2017, titled "Recommendations on Advancing the Economic Competitiveness of Data Analytics and Computing in Arkansas." The document, referred to as "The Blue-Ribbon Report," outlined several recommendations to address significant gaps in the projected supply and demand for data-skilled workers (DART Project, 2023). These recommendations included the creation of a nonprofit entity that would be charged with bringing industry needs and the talent pipeline into closer alignment.

The Blue-Ribbon Report was published at a time when the State's Science Advisory Committee (SAC) was deliberating on future priorities for research and development. The SAC is responsible for oversight of the State's Science and Technology Plan, a required component in the submission of proposals to a specific program within the National Science Foundation (NSF) called the Established Program to Stimulate Competitive Research (EPSCoR). Arkansas would be eligible to submit a new proposal under the EPSCoR Track-1 program in 2019. Track-1 projects were five years in duration and consisted of a $20 million award from NSF, plus a $4 million state match, to support statewide research and education aligned with a thematic scientific topic as decided by the SAC. In consideration of the growing momentum around data analytics at the time, the SAC decided that data science and data analytics would serve as the scientific topic for the upcoming EPSCoR Track-1 proposal. Stakeholders from academic institutions around the state collaborated to develop the proposal that would become DART - Data Analytics that are Robust and Trusted.

DART was awarded in 2020 and has since evolved into the largest and most ambitious Track-1 project in Arkansas history. The project's 75+ faculty participants represent thirteen colleges and universities and are loosely organized into six teams. Four of these teams are focused primarily on big data research, one is focused on cyberinfrastructure, and one is focused exclusively on expanding data science education. When the project began, there were no formal baccalaureate

degree offerings in data science. Students who were interested in pursuing data science education were resigned to creative patchworking of courses from computer science, information science, statistics, math, and other programs. Meanwhile, much of the curriculum used in those programs was significantly outdated. Today, undergraduates can choose from three institutions in different regions offering consistent and high-quality bachelor's programs in data science or begin with an affordable associate degree that is fully transferrable to a bachelor's program via innovative 2+2 agreements. Additionally, students enrolled at any public 2-year institution in Arkansas can take the first two courses in the bachelor's program online through the STEM Prep program, which offers virtual synchronous sections of the courses at home tuition prices.

Vision

The vision for a state-wide ecosystem for data science was motivated by taking advantage of the lack of formal data science degree programs at the undergraduate level in the state (though the UCA had data science track in their computer science and mathematics degree programs). This blank slate provided an opportunity for the post-secondary institutions to collaborate to create a high-quality, consistent, data science curriculum throughout the state. We started with four considerations:

1. We are a small enough state that we could all work together if we chose to.
2. We are a small enough state that we cannot afford to not work together.
3. We will broadly collaborate and emphasize transferability for meaningful degree pathways.
4. "The needs of the many, over the needs of the few...or, the one". (Paramount Pictures, 2002)

The proposal was presented at a state-wide workshop in September of 2019, beginning with a proposal to do so, then breakout workshops, and then coming back together to summarize what we had heard from the participants and propose the way forward. The participants at the workshop included faculty from 2-year and 4-year institutions, high-school faculty and students, state administrators from the Division of Education[1] and the Economic Development Commission, the Arkansas Center for Data Sciences, the Arkansas's STEM Coalition, and interested industry representatives from across the state.

The slide in Figure 1, below, provided the rationale for the proposal.

[1] Arkansas Act 910 of 2019 reduced the number of cabinet-level departments from 42 to 15. These consolidations took place over the time that we were planning and developing the statewide educational system for data science. For consistency of use, and to avoid confusion, we have elected to use the current titles of divisions and agencies of state government rather than using the names that were in use at different phases of our effort.

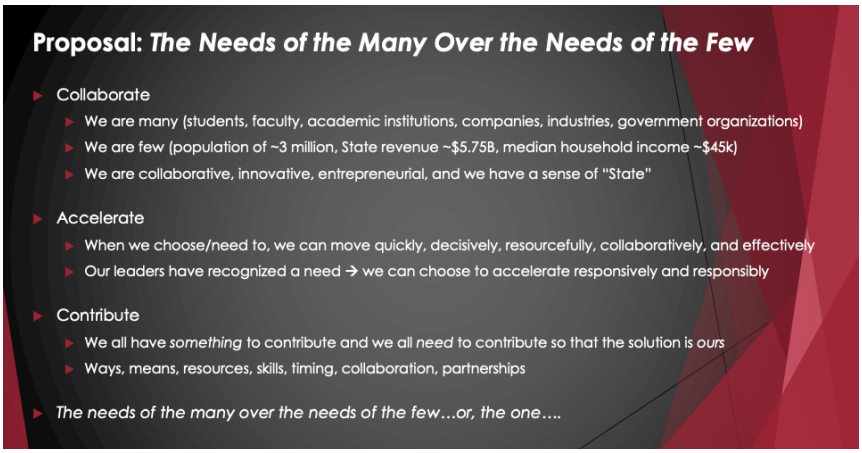

Figure 1. Proposal Rationale

The next figure, Figure 2, below, provides the proposal itself.

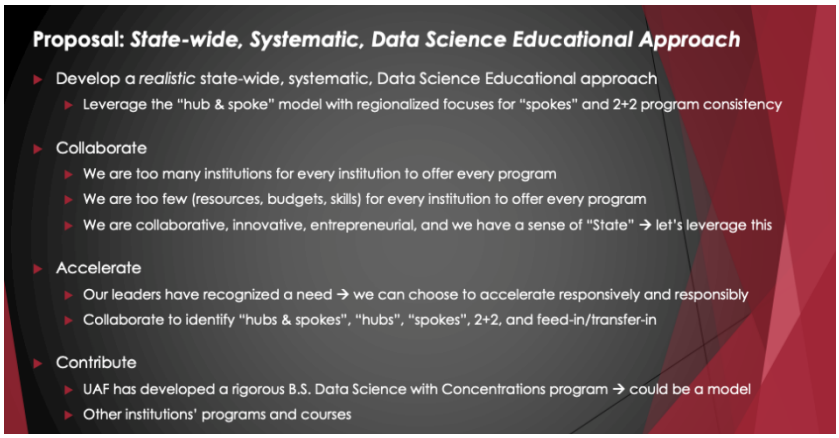

Figure 2. Proposal for a State-wide Data Science Ecosystem

And we began to use the phrase "Start Anywhere, Finish Anywhere" to help people internalize the vision. Over the next year, as we began to build the ecosystem, we formalized our vision and mission as part of an NSF EPSCoR (Established Program to Stimulate Competitive Research) Grant (NSF EPSCoR, 2020):

Vision

> A statewide educational ecosystem where learners receive a designed, consistent, scaffolded education in data science with further educational and/or job opportunities available at appropriate points in their careers.

Mission

> Create a model Data Science and Analytics program for our state's schools to promote problem-based and experiential-based pedagogy in critical thinking and analysis, technology familiarity, and foundation in math and statistics.

We have used our vision and mission and the "start anywhere, finish anywhere" mantra as our "North Star" in developing and expanding our ecosystem. Of course, it is one thing to have a

vision and mission, but there needs to be a strategy in place to realize them. That is our next topic.

Strategy

To develop our planned statewide educational ecosystem, we had to adopt a variety of strategies. We began by identifying the stakeholders. The stakeholders identified are the state's two- and 4-year institutions, the Arkansas Division of Higher Education, the Arkansas Division of Elementary and Secondary Education, the Arkansas Economic Development Commission, the Arkansas STEM Coalition, the Arkansas Center for Data Science, and the companies and corporations around the state that are dependent on data science technologies. We began engaging these stakeholders immediately through a series of informational meetings that were hosted under the auspices of the Arkansas STEM coalition and the Arkansas Center for Data Science. Though all sectors were invited to these meetings, the largest participating groups were faculty and administrators from 4-year institutions with more limited participation from representatives of 2-year institutions. However, this enabled us to reach our initial intended audience – and we were able to explain the steps that were underway to develop data science programs at the University of Arkansas (UofA) and the University of Central Arkansas (UCA). This provided two models for developing programs. At the UofA, the program was developed after cross-college discussions as a standalone program, with faculty being drawn from programs across those colleges. At the UCA, the program was developed by first creating courses in the departments of mathematics and computer science and the college of business, adding faculty as those initial offerings expanded in tracks in degree programs, and reconfiguring those tracks into a standalone degree program. This demonstrated that programs could be developed with differing initial budget commitments and showed that you could begin developing a program without a large initial commitment of resources.

As well as sharing the structure of the programs, we made it clear that we would share our experiences, syllabi, and course development materials to facilitate the development of programs across the state. Thus, in addition to developing a start anywhere, finish anywhere model, we also made clear that we would be operating a "no secrets" model. Our overarching philosophy was to build a statewide, collaborative data science program without competition among campuses and with each campus determining its own role under the hub-and-spoke model described in the next section.

As well as sharing a curriculum, we met with the Arkansas Division of Higher Education to facilitate the approval of new programs. The division agreed to assist programs that followed our framework to get any approvals for new programs.

As well as aiding in the organization of information meetings, the Arkansas Center for Data Sciences aided us in developing links with corporations and aided in providing interested campuses with information about current and expected employer needs for data science professionals. This information was important in the decision-making processes on some campuses and was often the key to generating interest in the first place.

It was important to work with K-12 educators to create pathways into data science. However, it was also clear that our time would be taken up collaborating with partners from higher education. To impact K-12 education we initiated meetings with the Department of Elementary and

Secondary Education and discussed the need to add data science options to the current K-12 offerings in computer science. Fortuitously, K-12 computer science was due for review and Stephen Addison and Karl Schubert were invited to participate in the review panels. This enabled them to ensure that data science was embedded in K-12 curricula and enabled them to aid the development of new strands in Data Science and Computer Engineering.

The development of this strategy is ongoing and ever-broadening as we bring in new partners. We are currently focused on developing additional partnerships with 2-year schools as well as broadening engagements with potential employers.

As the program has grown, we have made a major shift in our strategy for developing data science content at some institutions. Initially, we believed that if we held workshops at which data science faculty explained how they taught their classes and provided all the class materials, the faculty participating in those workshops would be able to return to their campuses to create and offer similar courses. This worked for some participants, but it did not work for others. Since data science can be viewed as a fusion of statistical and probabilistic thinking fused with modern computer technologies and viewed through a problem-solving lens, some participants had difficulty in understanding where to go because they did not have a background in all areas. We have since developed workshops where we explicitly teach the different skills separately and then use them to address data science programs. This should enable participants to adapt some of the methods into their current courses and develop the confidence needed to teach the data science courses in our framework.

Hub-and-Spoke Model

We realized that the "old school" point-to-point education model can only support the transfer of knowledge between two institutions at a time. The point-to-point model is unable to reach the full potential of collaboration between both institutions. A major drawback of the point-to-point model is that it does not scale. The model lacks a global benchmark to govern the transition of knowledge between institutions. We concluded that we needed a new model that was able to overcome some of these drawbacks and would be able to leverage institutional resources and support inter-institutional continuous learning (i.e., "start anywhere, finish anywhere"). Using a hub-and-spoke model would provide us with a Venn Diagram of capabilities and offerings.

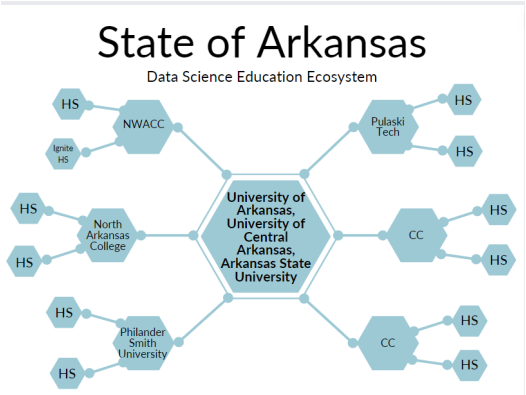

Figure 3. Hub-and-Spoke Model

Hub-and-spoke models were first used as a form of transport topology optimization in which traffic planners organize routes in a series of “spokes” that connect outlying points to a central “hub”. (Wikipedia contributors, 2024). The hub-and-spoke model requires fewer routes. For a network of *n* nodes, only *n*-1 routes are necessary to connect all notes so the upper bound is *n*-1. We saw an opportunity to utilize the “hub-and-spoke" model as an educational model that could overcome the drawbacks discussed earlier. To do so, we defined *distance* as a function of the resources, location, regional and state industry needs, and institution type.

This became the “Architectural Vision for Data Science for Arkansas”. We defined the “hubs” as 4-year institutions that would “opt-in” to implement the standardized 8-semester plan and courses and be destination institutions for 2-year college’ graduates. We adjusted our 8-semester plan working with the “opted-in” 2-year institutions to enable their “2+” student to graduate with an A.S. degree and then transfer to the “+2” 4-year institutions without being disadvantaged compared to the 8-semester plan students. We are also developing certificates collaboratively with the 2-year institutions for consistency across the state. The processes for these will be described later.

The Venn diagram of needs has another advantage: we can have more concentrations in the 4-year institutions across the state, as an ecosystem, and better serve the needs of those areas of the state. We recognized that there are reasons why a common set will be offered by all (for example, Business Data Analytics) and there are also regional differences in industries that inform concentrations for that region. For example, one area of the state may have a heavy agricultural industry suggesting a concentration of Agricultural Data Analytics. Another area may be heavy into manufacturing, and that would suggest an Operations Analytics Concentration, and so on. This regional specialization makes partnerships with industry significantly easier and more likely for guest lectures, student field trips, mid-term team projects, and Practicum/Capstone partnerships.

Once we had the vision, mission, strategy, and model in place, it was time to focus on collaborating to make this happen. We chose to do so through three workshops per year. Fall and Winter Workshops were where we share experiences and progress, and regular status meetings and the Summer Workshop is focused on pedagogy. The workshops and meetings are our next topic.

Education Theme Meetings & Workshops

Each year, the education theme hosts various events to continue participation and engagement across the theme. These events include:

- Annual DART All Hands Meeting Education Theme Breakout Sessions
- Annual DART Retreat Education Theme Breakout Sessions
- Education Theme Statewide Informational Meetings
- Monthly virtual collaboration meetings

As part of the NSF EPSCoR (Established Program to Stimulate Competitive Research) Grant (NSF EPSCoR, 2020), we planned annual teaching workshops to host events at which members can meet to work collaboratively on their theme and across themes. At both the Annual DART All-Hands meeting and the Annual DART retreat, the education theme hosted breakout sessions to collaborate with both the "hub" and "spoke" institutions to discuss their progress on the project, share resources, and create a game plan for moving forward with data science on their campuses. These breakout sessions have served as a great check-in point for all participants throughout the length of the grant.

Education Theme Statewide Informational Meetings

Throughout the life of the grant, the education theme has hosted monthly virtual collaboration meetings and invited all participating institutions across the state to join this meeting each month. During the meeting, participants shared an update on what they have been working on at their respective campuses and asked for guidance on how they can overcome barriers or assist them in moving forward with their curriculum development. The meetings also encouraged the sharing of ideas that may inspire an idea to another institution that may be experiencing similar challenges.

During Spring 2024, we are preparing to host three concentrated 2+ teaching workshops across the state of Arkansas. Each of the three workshops will be hosted in each of the three regions (eastern, central, and northwest) of the state in which our partner institutions reside. The idea of these workshops is to bring the learning environment closest to the 2+ partner institutions' communities so that they can have faculty attend with as few barriers as possible. During these workshops, we will teach the faculty how to use programs such as Excel, Python, R, PowerBI, and Tableau as well as how to conduct case studies.

Recruiting 2+ Colleges (Opt-in)

Our recruiting strategy has evolved over time. Initially, we held informational meetings at a variety of locations across the state, explained what we were trying to accomplish, and invited the participants to join the data science infrastructure as opt-in participants. Many people attended these sessions and indicated that they were interested but, in most cases, they did not follow up to join the program as participants. It was time to consider a different strategy. On reflecting on those who had participated in our informational meetings, we recognized that many of the participants were junior faculty or administrators. However, it was rare that both types of participants attended as representatives of the same institution. The faculty often did not feel able to move forward, and some were suspicious of our intentions – they wondered why we were willing to share everything with them.

To gain more partners in our hub-and-spoke model we quickly recognized that we needed to be talking to both administrators who could facilitate the development of programs and faculty who would develop courses and who could be trained to offer the necessary coursework. Thus, we

began reaching out to Presidents, Provosts, and Deans, summarized our plans, and requested that they identify a team who would participate in our efforts if they chose to “opt-in”.

When we described the program to potential participants, the initial reaction was often that “We are already doing these things, and we just need to show you how our courses fit into the program.” In most cases, the courses addressed a subset of what was offered in our model programs, so it was clear that major changes would be needed to adequately prepare students to complete a B.S. in data science after transfer. We developed an approach in which we analyzed prospective partners' offerings and provided feedback on how well their existing courses fit in the statewide data science ecosystem we are developing.

Thus today, recruiting schools to opt-in to the 2+2 program is a process that involves close collaboration with potential partner institutions. To begin the engagement, the leads of the program reach out to deans who indicate that they are interested in participating, and then a Graduate Research Assistant meets with representatives from these schools. During such sessions, discussions center around the courses already offered by the institution, allowing for a thorough understanding of their existing academic offerings. Our Graduate Research Assistant collaborates with the various academic representatives to design a program that integrates their institution’s offerings into the statewide ecosystem.

As part of this collaborative effort, a preliminary course equivalency assessment is conducted. This involves an examination and comparison of the courses already established at the institutions. This initial evaluation allows us to identify potential areas of alignment and establish the groundwork for the integration of those courses into the program.

Through these engagements, we not only provide valuable insights into the benefits of opting into the program but also actively involve partner institutions in shaping the program to suit their unique strengths and goals. This collaborative approach nurtures a sense of ownership among participating schools and strengthens the foundation for a successful and mutually beneficial partnership in the development of our state-wide data science academic ecosystem.

2-year College Course Equivalencies & 4-semester Plans

Course equivalencies refer to a comparison or mapping of courses offered at different educational institutions, typically to determine whether courses from one institution can be considered equivalent to those at another. This is essential in the 2+2 program as it creates a flexible “start anywhere, finish anywhere” ecosystem, allowing prospective students to commence their data science education at local 2-year colleges. They can seamlessly transfer to any of the 4-year institutions (hubs) in the state that have joined the program to complete a 4-year Bachelor of Science degree in Data Science.

Achieving this vision involves establishing course equivalency cross-checks between the courses offered in the UofA’s Data Science Program and those provided by other participating institutions. This ensures a smooth transition and pathway for students pursuing a comprehensive education in Data Science.

The program for course equivalency involves a thorough comparison between established university courses and those developed at other institutions. To attain approval as equivalent, a minimum of 80% similarity in course syllabi, learning objectives, and lesson plans is required. The initial cross-checks are conducted in an Excel spreadsheet by our Graduate Research Assistant and Undergraduate Research Assistants in the UofA Data Science program, individuals who have completed the courses. This process enables students to provide input on the integration of major topics into the collaborating institution's syllabi. Subsequently, the spreadsheet undergoes review with the Associate Director of the Data Science Program. Once the course equivalencies are certified, they are submitted to an Associate Dean in the College of Engineering. The approval process involves further review by the Provost and Legal to finalize approval of a Memorandum of Agreement (MOA) / Articulation Agreement, outlining the program and its operational mechanics.

For example, outlined below are the 2-year plans of both the UofA (4-year institution) and North Arkansas College (2-year institution). North Arkansas College (NorthArk) is the first approved school for the program and their 2-year plan leads to an associate degree in data science while the UofA is one of the 4-year hubs that students can transfer to upon completion of the associate degree.

In the 2-year plan, students at both institutions are advised to complete a total of sixty-two credits. While the programs share similarities, there are minor differences. UofA students cover Economics, Multivariable Math, and Probability and Statistics, whereas, due to Arkansas Division of Higher Education requirements, NorthArk students focus more on required electives and humanities. Nevertheless, graduates from NorthArk can seamlessly transfer to the UofA to pursue a bachelor's degree in data science within the 4-year time frame as shown below in Table 1. (Schubert, Shoultz, & Romer, 2023)

Table 1. Data Science 4-semester Plan 2+2 Plan for Course Equivalency and Sequencing

UofA				**NorthArk**	
Year 1: Semester 1 (16 credits)				**Year 1: Semester 1 (16 credits)**	
MATH 2254	Calculus I			**ENGL 1013**	English Composition I
GNED NNN4	Gen ED, Science Elective			**CIS 1603 & CIS 2011 OR DVSC 1104**	Introduction to Programming & R for Data Science OR Programming Languages for Data Science
ENGL 1013	Composition I			**DVSC 1003**	Introduction to Data Science
DASC 1003	Intro to Data Science			**ANTH or ECON or GEOG of HIST or PLSC or PSYC or SOC**	Elective
DASC 1104	Programming Languages for Data Science (Python and R)			**ART or DRAM or ENGL or FA or FL or MUS or PHIL**	Elective

UofA			NorthArk	
Year 1: Semester 2 (15 credits)			**Year 1: Semester 2 (16 credits)**	
MATH 2564	Calculus II		**ENGL 1023**	English Composition II
ECON 2143	Gen ED, Basic Economics: Theory and Practice		**ART or DRAM or ENGL or FA or FL or MUS or PHIL**	Elective
ENGL 1033	Gen ED, Technical Composition II		**ANTH or ECON or GEOG of HIST or PLSC or PSYC or SOC**	Elective
DASC 1204	Intro to Object-Oriented Programming for DASC (JAVA)		**MAT 2204**	Analytical Geometry and Calculus I
DASC 1223	Role of Data Science in Today's World		**DVSC 1013**	Intermediate Data Science

UofA			NorthArk	
Year 2: Semester 1 (16 credits)			**Year 2: Semester 1 (17 credits)**	
DASC 2594	Multivariable Math for Data Scientists		**HIST 2003 or HIS 2013 or PLSC 2003**	U.S. History I or U.S. History II or American National Government
INEG 2313	Applied Probability and Statistics for Engineers I		**BIOL or CHEM or GEOL or PHSC or PSYC or SOC**	Elective
DASC 2213	Data Visualization & Communication		**MAT 2304**	Analytical Geometry and Calculus II
DASC 2113	Principles & Techniques of Data Science		**DVSC 2113**	Principles & Techniques of Data Science
DASC 1222	Role of Data Science in Today's World		**DVSC 2213**	Data Visualization & Communication
GNED NNN3	Gen Ed, Fine Arts Elective			

UofA			NorthArk	
Year 2: Semester 2 (15 credits)			**Year 2: Semester 2 (13 credits)**	
SEVI 2053	Business Foundations		**BIOL or CHEM or GEOL or PHSC or PSYC or SOC**	Elective
INEG 2333	Applied Probability & Statistics for Engineers II		**SPCH 2303 or Social Sciences or Humanities Elective**	Public Speaking or Social Sciences or Humanities Elective
DASC 2103	Data Structures & Algorithms		**DVSC 2203**	Data Management & Database

DASC 2203	Data Management & Data Base		**CIS 2203**	Data Structures & Algorithm
DASC 1222	Role of Data Science in Today's World			
Total	62 credits		**Total**	62 credits

This format can vary depending on the 2-year institution. Nonetheless, it leads to the "start anywhere, finish anywhere" ecosystem.

2-year College Pedagogical Workshops

Each summer, we have provided our 2-year college partners with an opportunity to participate in professional development by hosting a pedagogical workshop. These workshops are hosted over 2 days (afternoon, full day, morning). We also try to host the workshops in a location that allows both the participants and the teaching faculty to have a productive learning experience. All meals were also provided with each workshop.

The workshops have provided the "hub" institution faculty with the opportunity to "teach the teachers," meaning that they are teaching the faculty at the 2-year colleges how to teach the courses in the curriculum. This is particularly beneficial to the 2-year colleges because they are obtaining professional development in a specific subject area with minimal financial commitment. Each faculty member teaching the workshops is provided with a stipend for providing their time and expertise. Faculty members conducting the workshops are required to not only attend and produce the workshops but also to provide their teaching materials for all participants. It is only after the faculty member has completed all obligations that they are paid their stipend.

Table 2 below shows our agenda for the 2023 2-year college pedagogical workshop which was hosted at the UCA. We chose this location, as it is centrally located within the state, with the hope of achieving better participation from our partner schools that are located across the state. Hosting it in a central location would allow participants to only drive 3 hours (at the most) instead of 5-6 hours to Northwest Arkansas.

Table 2. Summer Data Science Workshop

Day	Time	Topic
Sunday, July 30	12:00 – 1:00 pm	Check-In- Snacks Provided
	1:00 – 1:15pm	Intros & Workshop Overview
	1:15 – 1:30pm	Data Science Analytics Process
	1:30 – 3:00pm	Introduction to Python in DASC – Part 1
	3:00 – 3:15pm	[Break]
	3:15 – 5:00pm	Introduction to Python in DASC – Part 2
	5:30 -7:00pm	Dinner- Provided

Monday, July 31	8:30-9:00am	Breakfast- Provided
	9:00am – 10:00am	Introduction to Python in DASC – Part 3
	10:00 – 10:15am	[Break]
	10:15am – noon	Introduction to R in DASC – Part 1
	Noon – 1:00pm	Lunch- Provided
	1:00 – 3:00pm	Introduction to R in DASC – Part 2
	3:00 - 3:15pm	[Break]- Snacks Provided
	3:15 - 5:15pm	Data Visualization & Communication
	6:00 – 7:30pm	Dinner- Off Campus @ Mike's Place
Tuesday, August 1	8:30-9:00 am	Breakfast- Provided
	9:00am – 10:00am	Probability & Statistics in Data Science
	10:00 – 10:30am	[Break]- Snacks Provided
	10:30am – 12:30pm	Exploratory Data Analysis (EDA) in Python & R
	12:30 pm	Lunch – Boxed Lunch Provided

Based on our experience with the pedagogical workshops, we have found that both participation and engagement are higher when offering the workshops in person. Through the grant period so far, we have held two workshops in person and one virtually. The virtual workshop had the lowest number of participants and the least amount of verbal engagement in the virtual environment. However, with regards to in-person workshops, timing is key as many faculty are on 9-month contracts and therefore are not obligated to conduct work-related projects outside of the 9 months. Therefore, we found that it was best to survey our anticipated participants as to what dates would work best for them before selecting a date to host the workshop to have the greatest attendance possible.

2-year College Experiences Opting-In

North Arkansas College (NorthArk) is a small 2-year college serving a largely rural six-county area in North Central Arkansas. Although only seventy-five miles from the financial center of the state in Northwest Arkansas, the economic and technological landscape is vastly different. This partnership will not only help supply the statewide need for data scientists but will drive economic growth and upward mobility in NorthArk's region by (1) providing employers with the skilled data science workforce needed to meet current and future needs, (2) prompting expansion and innovation in regional business and industry, and (3) equipping local residents with knowledge and skills for well-paying and rewarding careers in this rapidly growing data science field.

NorthArk had several key components in place to partner with UofA and the EPSCoR team, including a history of transfer articulations with the UofA in Business and Agriculture, long-time participation in UofA's STEM Prep program, existing courses that could be applied to a new data science program, and a commitment to providing advanced STEM career opportunities to rural residents. NorthArk also has well-credentialled and talented faculty in math, programming, and statistics, but these faculty did not have expertise in the multiple flavors of the emerging data science field. In addition to the need for faculty development, this project required (1) new courses, (2) new takes on content in existing courses, (3) adjunct faculty to teach data science courses and/or offload existing courses from current faculty, and (4) an entirely new transfer associate degree.

Establishing equivalent course content and the design of the associate of science degree were the most challenging pieces of this curricular puzzle. Some of the necessary content was entirely new and led to new courses such as *Introduction to Data Science*, while other courses, such as *Data Structures*, were already part of NorthArk's offerings. Some existing courses, such as *Introduction to Programming*, were not a perfect fit because they targeted an audience of programmers rather than an audience who will use programming as a tool to do data science.

Each challenge has had a solution due to the resources provided by the DART grant. For example, detailed course content for all first-year and sophomore-level data science courses was provided to project participants. These materials included syllabi, course outcomes, lesson plans, assignments, and videos. NorthArk and UofA faculty brainstormed to ensure that all necessary content was covered even when the course sequencing was not in perfect alignment. To prepare faculty, the project provides extensive faculty development through summer training workshops, in addition to the previously mentioned curricular materials. NorthArk has already added two adjunct instructors through industry and college connections and is widening its network of employer partners.

NorthArk chose to develop a new associate of science in data science, rather than creating an option in an existing degree, for two reasons: (1) to aid in the recruitment and promotion of both the degree and the partnership and (2) to provide students with a degree title that correctly describes their training; this will benefit graduates who take the associate degree directly into the workforce. Designing the degree was a challenge primarily due to Arkansas requirements that associate degrees designed for transfer must include the 35-credit general education core but must also total exactly sixty credits. NorthArk can, therefore, provide several more hours toward the bachelor's degree in data science than will fit in the state-approved degree. Students who choose to do so may take these courses from NorthArk at the 2-year college's tuition rates even after beginning upper-level courses at UofA.

Possibly the most difficult part of this project was the months-long wait as multiple UofA departments and their legal team evaluated course equivalencies and the articulation agreement. The wait was over – and worth the wait -- when UofA hosted the NorthArk team for a Memo of Agreement (MOA) signing event in November 2023. NorthArk is proud to be the first 2-year college to join with UofA as part of this DART grant. The challenging work by faculty and staff at both institutions will not only provide new and rewarding opportunities for rural residents in NorthArk's service area, but also provide a template for multiple partnerships that will benefit the entire state. The feedback from the educators and institutions has been overwhelmingly positive. For example, in a press release after signing the NorthArk MOA with the UofA

(University of Arkansas, 2023), both the UofA's Provost and NorthArk's administrators reflected on the collaboration:

> "Given that data science is one of the fastest growing career fields in the nation, creating a multidisciplinary data science program was a great way to advance our land-grant mission of service to our students and our state," said U of A Provost Terry Martin. "As we established our own data science program, the idea of partnering with other institutions to further increase access to data science training took hold. We are grateful that North Arkansas College agreed to develop this program with us, as they have created a wonderful model for other two-year institutions to follow."

> "Having this opportunity for our students to have a destination and assurance in place when they graduate Northark is such a positive for our students and our program," said Northark President Rick Massengale. "We have a very aggressive, technology driven curriculum in STEM education and this is another opportunity for our students to start local, and partner with a global player in the discipline of Data Science. We are grateful for this pathway for our students and for the continued partnership we have with U of A to provide opportunity and excellence."

> "Northark is thrilled to be the first two-year college to join with the University of Arkansas as part of the EPSCoR grant," said Laura Berry, interim dean of Health Professions and director of special projects at Northark. "This is truly the most collaborative effort between two institutions that I've experienced, and it will provide a fantastic new career and transfer opportunity for students in Northark's region. My thanks to Dr. Schubert and his team for including us in this partnership!"

> "It was a real honor and pleasure to work with Dean Berry and Northark to develop this MOA," said Karl Schubert, professor of practice and associate director of the U of A's Data Science Program. "It will serve as the blueprint for others to join the state-wide data science ecosystem and potentially for schools in other EPSCoR states."

NorthArk was the first of approximately a dozen 2-year colleges who have "opted-in" to the ecosystem. They are at various stages of their development. The 2-year college that is next to be ready is Northwest Arkansas Community College.

Northwest Arkansas Community College (NWACC), located in Bentonville, Arkansas, is ideally situated for a 2-year data science program to succeed. The college is part of a standard metropolitan statistical area of about 500,000 (Northwest Arkansas Council, 2023) people, which

is home to many industries such as transportation, retail, health care, finance, and food manufacturing. Each of these industries makes heavy use of data analytics companies, and often struggles to find analysts and data scientists. (Souza, 2019) Community college students could contribute significantly to building a strong data science workforce, as "85 percent of community college degree earners stay close to their alma maters, contributing to the workforce, buying homes, raising families and becoming the next generation of community leaders." (Glenn, 2020) NWACC already maintained strong partnerships with programs at the UofA in other related programs, such as business and computer science, with many students transferring each year. The data science initiative in Arkansas, provided a cost-effective means of establishing a data science pathway at the college, making a career in this field accessible to NWACC's students, with the option to transfer to the UofA, or another participating 4-year institution in the state.

The faculty at NWACC have adopted the shared curriculum. An evaluation and alignment process were undertaken, to ensure that the course outcomes, assignments, and technology would be familiar to NWACC students when they transferred to the UofA or another participating 4-year institution. The faculty has attended the professional development opportunities offered by the grant and attended the regional and state-wide meetings to ensure curricular alignment. The meetings also offer networking opportunities with other programs in the state and the opportunity to share assignments, data sources, and other pedagogical innovations.

The lock-step course transfer pathway ensures that NWACC students will transfer seamlessly with junior status if the student follows the pathway and can complete the baccalaureate degree in 4 years. Support for program development has been coordinated through the Arkansas Division of Higher Education, with input and participation in the EPSCoR project.

The DART grant has provided a cost-effective means of developing a new program, or transfer pathway for community college students. NWACC has elected to include its transfer pathway under its more general Associate of Science transfer degree, making use of the same courses that would have been included in a stand-alone A.S. Data Science degree plan. This saves the college program initiation fees and other administrative costs associated with a state program review.

NWACC began offering the DASC 1003 Introduction to Data Science in the Spring of 2024 as a trial course. The MOU with the University of Arkansas, which will allow the college to offer the full program beginning in the Fall of 2024, will be signed in May 2024, so the program is just starting this semester. In the Spring of 2024 two sections of DASC 1003 were offered, one on the main campus in Bentonville, and one at a High School which participates in the Early College Experience (concurrent enrollment) program offered by NWACC. The majority of the students enrolled in the class this semester are at the High School location. Currently, this is the only high school in the ecosystem, though more are expected to follow over the next year or two. As with the relationship of the 2-year college spokes to the 4-year college hubs, the relationship of the high schools for dual credit are geographically-based and in this case, the 2-year college acts as a hub for the high school spoke.

Future

Our efforts to develop a statewide educational infrastructure for Data Science will not end in 2025 when the DART grant ends. The development of Data Science programs at the UofA, and the UCA began before work on the grant proposal began – indeed the existence of these programs aided the development of the grant proposal. We began recruiting partners and sharing our vision for the data science ecosystem long before the grant existed. Workshops and conferences were held under the auspices of the Arkansas Center for Data Sciences and the Arkansas STEM Coalition, and these workshops led to the identification of the initial partners who participated in the first year of grant funding. Continuing the development of the statewide data science infrastructure is not dependent on continued grant funding though continued development will be enhanced, broadened, and potentially accelerated if we are successful in our efforts to secure additional grant funds.

With grant support, we have developed a hub-and-spoke model to enable Arkansans to earn B.S. degrees in Data Science. The hub campuses Arkansas State University (A-State) in Northeast Arkansas, the UofA in Northwest Arkansas, and the UCA in Central Arkansas are well dispersed and serve the majority of Arkansas's population, As our infrastructure continues to develop, we anticipate the development of a hub campus in Southern Arkansas from one of our existing partners. Currently no program in Southern Arkansas has a fully developed B.S. in data science, and it is unlikely that finances will enable the development of a southern hub in the short term. Instead, we will focus on strengthening the ability of programs in Southern Arkansas to offer foundation classes that will enable students to transfer to a hub campus without any loss of credits.

We are currently developing strategies to secure ongoing funding to support the continuing development of the statewide data science ecosystem. We can share course resources independently of grant funding, but securing ongoing funding will be necessary to enable us to continue to offer workshops and to offer summer workshops, and other activities to encourage students to pursue higher education in data science.

Program development also continues at the hub institutions. Two of the hub schools intend to submit their data science programs for ABET (Accreditation Board for Engineering and Technology, Inc.) accreditation when their other programs are undergoing periodic review - this will be in the 2025-26 cycle for UofA and the 2025-26 cycle for UCA. This will then enable these programs to provide additional help to the other partner schools. The first data for learning experiences for each of the courses, and the continuous improvement plans for each of the courses, are being collected in the 2023-24 academic year.

Current, Future, and Lasting Impact for the State

Current state-level education and workforce development efforts are focused on high-skill, high-wage, and high-demand occupations. Given the existing and projected workforce demand for data scientists, the state-wide data science educational ecosystem is vital to meet Arkansas' workforce and economic development needs, with the hub-and-spoke model addressing region-specific needs. (U.S. Bureau of Labor Statistics, 2023) This framework for collaboration and

articulation among 2- and 4-year higher education institutions addresses barriers often associated with developing new programs as well as transfer student success. Furthermore, the ecosystem extends to connect K-12, higher education, and workforce partners and increases access to robust education and training opportunities across the state.

In 2015, Arkansas became the first state in the nation to pass a comprehensive law requiring all public and charter high schools to offer computer science courses. This was followed by 2021 legislation that made a computer science course a graduation requirement, which was amended in 2023 to require one credit in a high school computer science or computer science-related career and technical education for graduation. (Arkansas General Assembly, 2023) Data Science is one of the high school computer science pathways and includes standards for a three-year program of study, closely aligned to introductory-level college courses. Students completing this and other computer science-related pathways and other STEM-focused high school coursework need opportunities to pursue postsecondary data science programs in the state.

High school graduates, as well as non-traditional students, wanting a bachelor's degree often begin at a 2-year institution for a variety of reasons. Forty-one percent of Arkansas' population lives in rural areas and may have greater access to one of the twenty-three public 2-year colleges than a 4-year college due to distance or cost. (ADHE Scholarship Application Management System, 2024) Strong partnerships among community colleges and transfer institutions are key to bachelor's degree attainment for these students. The well-defined course sequences in the articulation agreements along with completion of an associate degree at the 2-year college are supported by the research to significantly increase student success upon transfer. (Monaghan and Attewell, 2015) (Shapiro et al., 2013) (Stern, 2016)

The state-wide data science educational ecosystem was the impetus for the development of the Arkansas Course Transfer System (ACTS) Course, DASC 1003 Introduction to Data Science. Faculty from around the state, including those involved in the ecosystem and those not (yet), came together to develop the course description and expected student learning outcomes that would guarantee transfer to any public 2- or 4-year higher education institution. This, along with the development of an introductory cybersecurity course, was a first step in a broader plan to explore adding such courses as a general education requirement. Additionally, the freshman-level data science course in ACTS will increase opportunities for high school concurrent credit.

The quote attributed to Helen Keller, "Alone we can do so little; together we can do so much," is apropos of what scaling this statewide data science program means to Arkansas. As more campuses "opt-in" to the ecosystem, the cross-pollination of content knowledge and effective pedagogical practices will enrich faculty while the replicable, yet institution-specific articulation agreements will give greater education opportunities to students wanting to pursue the field of data science.

References

ADHE Scholarship Application Management System. (2024, February 4). Institutions in AR.

Retrieved February 4, 2024, from https://sams.adhe.edu/Institutions

Arkansas General Assembly. (2021). Act 414 of 2021: Computer Science Education Advancement Act. Retrieved February 4, 2024, from https://www.arkleg.state.ar.us/Home/FTPDocument?path=%2FACTS%2F2021R%2FPublic%2FACT414.pdf

Arkansas General Assembly. (2023, April 13). Act 654 of 2023: An Act to Regulate Electronic Medical Records. Retrieved February 4, 2024, from https://www.arkleg.state.ar.us/Acts/FTPDocument?path=%2FACTS%2F2023R%2FPublic%2F&file=502.pdf&ddBienniumSession=2023%2F2023R

Arkansas State Legislature. (2015). ACT187. Retrieved February 5, 2024, from https://www.arkleg.state.ar.us/Home/FTPDocument?path=%2FACTS%2F2015%2FPublic%2FACT187.pdf

Armitage, A. (2015, February 26). Arkansas Gov. Directs High Schools to Offer Computer Science Classes. Education Week. Retrieved February 5, 2024, from https://www.edweek.org/technology/arkansas-gov-directs-high-schools-to-offer-computer-science-classes/2015/02

DART Project. (2023, November 13). Recommendations on Advancing the Economic Competitiveness of Data Analytics and Computing in Arkansas. Retrieved February 4, 2024, from https://dartproject.org/recommendations-on-advancing-the-economic-competitiveness-of-data-analytics-and-computing-in-arkansas/

Fortune Media (USA) Corporation. (2023, August 2). Walmart tops the Fortune Global 500 list for 10th consecutive year. PR Newswire. Retrieved February 4, 2024, from https://www.prnewswire.com/news-releases/walmart-tops-the-fortune-global-500-list-for-10th-consecutive-year-301891191.html

Glenn, L. A. (2020, April 20). How are community college students different? Community College Daily. Retrieved February 4, 2024, from https://www.ccdaily.com/2020/04/how-are-community-college-students-different/

Keller, H. (n.d.). Alone we can do so little; together we can do so much. Retrieved February 4, 2024, from American Foundation for the Blind.

Laden, R. (2012, November 23). Wal-Mart's new data center seen as a big plus for city. The Gazette. Retrieved February 4, 2024, from https://gazette.com/news/wal-marts-new-data-center-seen-as-a-big-plus-for-city/article_36ac2f4a-bd18-5258-9230-6d9568c2fefa.html

Monaghan, D. B., & Attewell, P. (2015). The Community College Route to the Bachelor's Degree. Educational Evaluation and Policy Analysis, 37(1), 70-91. Retrieved February 4, 2024, from https://doi.org/10.3102/0162373714521865

Northwest Arkansas Council. (2023, May 23). Northwest Arkansas Now Nation's 100th Largest Metro, Census Estimates. Retrieved February 4, 2024, from https://nwacouncil.org/2023/05/23/northwest-arkansas-now-nations-100th-largest-metro-census-estimates/

Paramount Pictures; directed by Nicholas Meyer. (2002). Star Trek II: The Wrath of Khan. Hollywood: Paramount Pictures

Schubert, K. D., Shoultz, L., & Romer, S. (2023). Expanding, Improving, and Completing a Multi-College Interdisciplinary B.S. Data Science Program with Concentrations. Paper presented at 2023 ASEE Annual Conference & Exposition, Baltimore, MD. Retrieved from https://peer.asee.org/43534

Shapiro, D., Dundar, A., Ziskin, M., Chiang, Y.-C., Chen, J., Harrell, A., & Torres, V. (2013). Baccalaureate attainment: A national view of the postsecondary outcomes of students who transfer from two-year to four-year institutions (Signature Report No. 5). Herndon, VA: National Student Clearinghouse Research Center. Retrieved February 4, 2024, from https://nscresearchcenter.org/signaturereport5

Souza, K. (2019, February 1). Walmart exec says Northwest Arkansas could be data science hub. Talk Business & Politics. Retrieved February 4, 2024, from https://talkbusiness.net/2019/02/walmart-exec-says-northwest-arkansas-could-be-data-science-hub/

Stern, J. M. B. (2016). The effect of articulation agreements on community college transfers and bachelor's degree attainment. Community College Journal of Research and Practice, 40(5), 355-369. Retrieved February 4, 2024, from https://doi.org/10.1080/10668926.2015.1065209

University of Arkansas. (2023, February 1). U of A, North Arkansas College partner to develop Arkansas data science talent pool. *University of Arkansas News*. Retrieved April 1, 2024, from https://news.uark.edu/articles/66826/u-of-a-north-arkansas-college-partner-to-develop-arkansas-data-science-talent-pool

U.S. Bureau of Labor Statistics. (2023, September 6). Data Scientists: Occupational Outlook Handbook. Retrieved February 4, 2024, from https://www.bls.gov/ooh/math/data-scientists.htm

Wikipedia contributors. (2024, February 4). Hub and spoke. In Wikipedia, The Free Encyclopedia. Retrieved February 4, 2024, from https://en.wikipedia.org/wiki/Hub_and_spoke