

Developing New Data Pathways in Community Colleges

Joyce Malyn-Smith
Education Development Center, Waltham MA, USA
jmalynsmith@edc.org

Sarah MacGillivray
Education Development Center, Waltham MA, USA
smacgillivray@edc.org

Michael Harris
Bunker Hill Community College, Boston, MA, USA
mdharris@bhcc.edu

James Polzin
Normandale Community College, Bloomington, MN, USA
james.polzin@normandale.edu

Abstract: As more industries come to rely on data collection and analysis, the demand for skilled data workers is growing at a pace that outstrips the capacity of colleges to develop the programs needed to produce qualified employees. Technician education programs and community colleges are expressing growing interest in developing data programs. What is missing are strategies and supports that can enable colleges to rapidly respond to these opportunities and scale-up their efforts to train the next generation of data workers in a sustained and timely manner. This panel will present EDC's Data Pathway Development Process designed to support faculty in the creation and launch of new data programs at two-year institutions that are aligned to local employment opportunities.

The Critical Need for Middle-Skilled Data Practitioners

Data has become ubiquitous in today's technology-intensive world, driving innovations in business and industry, and collected in nearly every facet of our personal lives. This has created a growing demand both for citizens enabled with data literacy and workers at all levels skilled in supporting data-intensive processes. Industry reports point to a need for employees across all sectors who can capture, manage, and interpret data to do their jobs effectively (Columbus, 2017; Sasso, 2018). The World Economic Forum report, *Data Science in the New Economy: A new race for talent in the Fourth Industrial Revolution*, found that between 2013–2018 demand for data skills rose 59% in Information and Communications Technologies industries, 50% within Media and Entertainment, 69% in Professional Services (69%), and by 88% in Financial Services industries (2019). With demand rising, jobs in data science and analytics have become highly competitive, and companies have reported difficulties in locating enough qualified employees to fill their open data positions (Bloomberg Next, 2018). Projections indicate that this gap will continue to grow (Manyika et al., 2017).

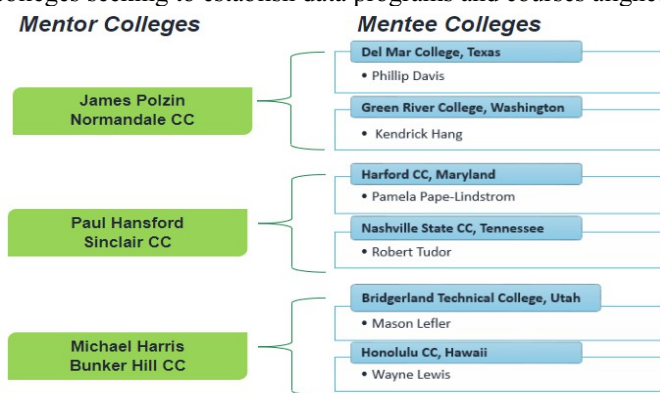
For example, the 2017 publication *The Quant Crunch: How the Demand for Data Science Skills is Disrupting the Job Market*, Burning Glass, IBM and the Business-Higher Education Forum (BHEF) reported that: 1) By 2020, the number of jobs for all US data professionals will increase by 364,000 openings to 2,720,000 according to IBM. 2) Annual demand for the fast-growing new roles of data scientist, data developers, and data engineers will reach nearly 700,000 openings by 2020. 3) Data Science and Analytics (DSA) jobs remain open an average of 45 days, five days longer than the market average. And 4) Data-centric jobs are well paid with jobs requiring machine learning skills paying an average of \$114,000. Advertised data scientist jobs pay an average of \$105,000 and advertised data engineering jobs pay an average of \$117,000 (Markow et al., 2017). The key to meeting this demand is to develop training programs for middle-skilled data workers. In a report on middle-skill careers, Burning Glass found that jobs requiring data skills were growing faster, paying better, and offer more opportunities

for career advancement (Bradley et al., 2017). Middle skill jobs typically require less than a four-year degree, pay livable wages, and are precisely the occupations targeted by community and technical college programming. Many of these institutions have begun the process of developing data programs, but are lacking strategies and supports that will enable them to rapidly respond to these opportunities, and to scale-up their efforts to train the next generation of data workers in a sustained and timely manner.

Building Data Pathways at Community Colleges

Recognized by the National Science Foundation (NSF) as part of their 2016 *Ten Big Ideas* “Harnessing the Data Revolution,” acknowledges that data science has become a topic worthy of investment and a principal focus of research, education, and workforce development programs (NSF 2019). NSF’s Advanced Technological Education (ATE) Program is an example of this investment, with a focus on workforce preparation for high-tech fields within two-year colleges and technical education institutions. Education Development Center’s *Oceans of Data Institute*, alongside our partners at Bunker Hill, Normandale, Sinclair, and Johnson County Community Colleges, have explored how best to define and develop pathways of the middle-skilled data practitioner through two ATE-funded grants. The overarching goal of these projects is to prepare students in technician education programs for employment in data intensive careers, by supporting community colleges as they develop new data courses and programs that are aligned to local industry needs and expectations for middle skilled data workers.

In our first project, *Creating Pathways for Big Data Careers* (NSF DUE #1501927), our team developed the *Profile of the Data Practitioner*. The profile documented the skills, knowledge, and abilities that are needed by the middle-skilled data worker—defining what the data practitioner needs to be able to know and do. The profile provides a framework for administrators, faculty, and local employers to discuss and prioritize the skills data workers will need for work in their community. Our four college partners used this profile, along with tools such as the *Gap Analysis* and *Curriculum Analysis Matrix*, to develop new data programs and courses in their institutions. The collective experiences of developing these programs enables our partners to provide expert guidance to faculty in other colleges seeking to establish data programs and courses aligned to their own local industry needs.



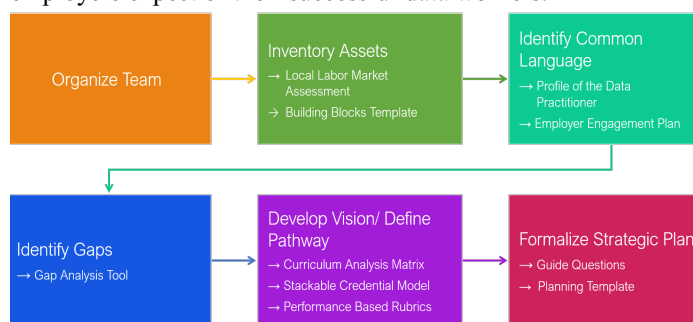
Our next ATE project, *Mentoring New Data Pathways* (NSF DUE #1902568), has taken the *Profile of the Data Practitioner* and associated tools and built upon them, refining them in order to guide new colleges through the pathway development process. This project features our partners Hansford, Harris and Polzin as they mentor six new community colleges ready to develop data pathways at their institutions (Fig. 1). As the mentors and mentees work together, they are joined by an online community of practice of 27 additional colleges, all of whom join in monthly discussions on building data programming at community colleges.

The Data Pathway Development Process

The Data Pathway Development Process is designed to guide colleges through a 6-step progression (Fig. 2); (1) organizing a development team, (2) inventorying of existing assets, (3) identifying a common language between employers and colleges, (4) identifying gaps between industry needs and college offerings, (5) developing curriculum, exploring stackable credentials and rubrics, and (6) formalizing a strategic plan. Along each step, tools

to guide the process are provided, designed in partnership with data science experts and community college faculty. These tools include:

- *Local Labor Assessment*: A process for identifying occupational trends and emerging skill sets in local industry.
- *Building Blocks Template*: A means of capturing an inventory of the existing courses, programs, certifications, degrees, and supportive activities that are in place and can be used to develop a data program.
- *Data Practitioner Profile*: A nationally validated document describing what middle skilled data workers know and are able to do; used by school administrators, faculty, and employers to discuss and identify data worker skills needed in local workplaces.
- *Employer Engagement Plan*: A template that presents an example of a detailed, step-by-step process for identifying, recruiting, and engaging business partners in support of a data pathway.
- *Gap Analysis*: A tool to facilitate a discussion between school administrators, faculty, and employers about the “gap” between the data skills addressed by existing school courses/ programs and the data skills employers expect of their successful data workers.



- *Curriculum Analysis Matrix*: A tool that enables faculty to 1) analyze the extent to which their curriculum includes data skills prioritized by local employers; and 2) plan where to include specific data skills in their courses/programs to ensure that graduates leave school having developed data skills prioritized by local employers.
- *Stackable Credential Model*: A generic representation of a way that colleges can sequence data analytic credentials to enable individuals to build their qualifications for higher skill and higher paying jobs.
- *Performance Based Rubrics*: A collection of contextualized statements that illustrate what data worker skills look like when performed at different degrees of proficiency expected by today’s employers.
- *Guide Questions and Planning Template*: A tool that can be used to record a school’s strategy for developing a data program that specifies anticipated actions, responsibilities, and deadlines.

Data Programs Developed Using this Process

Community College partners have engaged deeply with these tools; first to develop their own data programs and then in providing feedback and refinements in preparation for leading their mentees through the process.

Mini Case Study: Normandale Community College

Normandale introduced an A.S. in Data Science degree in fall, 2017 with seven specialized courses plus general education requirements; and in fall, 2018 began offering three certificate programs that focus on the Data Practitioner (MSBDES): Certificate in Analytical Statistics, Certificate in Computational Analytics, and the Certificate in Data Visualization and Presentation. Normandale’s certificate in Data Visualization is part of a strategy aimed at augmenting the school’s A.S. Degree in Data Analytics. The certificate will create a “package of courses” for data workers who want to develop skills in that area. Other certificate programs are also being considered to combine various skillsets into smaller, more manageable collections which can then be used individually, or stacked together in pursuit of the A.S. degree. School leaders continue to examine opportunities for coordinating with additional bachelor’s degree programs and master’s programs. The aim is to establish and

streamline multiple routes to additional future education in the area of data analytics and data science that can build off of any certificates and the A.S. degree that Normandale is currently providing.

Normandale continued work on certificate programs, expanding the skill sets of working adults to develop and upgrade their data skills now required by industries local to the college. Normandale fine-tuned existing curriculum expanding transfer options to several additional 4-year universities in the area, positioning Normandale as a convenient, low-cost starting point for students pursuing a bachelor's degree, as well as those returning to school. Normandale's data programs have a small but steady enrollment that they are working to increase through collaborative efforts with the school's Marketing Department. Their goal is to reach the public and through ties to local high schools, raising awareness of the available undergraduate and 2-year degree programs, and the potential job opportunities in this growing field.

Other Examples:

At Bunker Hill Community College (MA), leaders created a new data science associates degree. The degree contains two stackable certificates: the Data Management and Data Analytics certificates. Bunker Hill is in the process of changing their Data Management Certificate by removing CIT-234 (Advanced Excel) and replacing it with CIT-130 (Data Visualization with Tableau). Bunker Hill will be graduating its first student from the Data analytics AS degree in fall 2020. Enrollment in the Intro to R course, along with the Data Management Certificate, has been steady while overall Bunker Hill has seen a significant drop in enrollment over the past few years. Enrollment has grown into double digits for the Data Analytics Certificate. A new course on Hadoop was designed based upon previously created labs. Hadoop will now be offered as an introductory course, outside of the AS degree in analytics program. This will allow school leaders to develop a future certificate on cloud systems and infrastructure.

Sinclair Community College developed a two-prong approach to addressing Big Data skills, one in Data Analytics and one in Business Analytics. Sinclair began their Associate of Applied Science in Data Analytics degree in fall 2018 with ten courses that complemented an existing ten-course certificate in Data Analytics. Further, in January of 2019, they began an Associate of Applied Science in Business Analytics degree with 20 courses. They have consulted with and developed partnerships with 14 businesses. Sinclair continues to offer new courses adding CIS 2265 Data Visualization (Power BI & Tableau) and CIS 2266- Data Analysis with Scripting Languages. Sinclair modified their business analytics program (BIS.S.AAS) adding four elective paths effective fall 2020. They launched their Data Fundamentals short-term technical certificate (DF.S.STC) in the fall of 2019 as part of the CS/IT department, one year after the CIS AA degree, as a stacked credential; and are prepared to launch a new short-term technical certificate in Data and Information Management (DIM.S.STC) in fall of 2020. Enrollment in Sinclair's new courses, Data Visualization (CIS 2265) and Data Analysis with Scripting Languages (CIS 2266) and pre-requires for BIS 1600 are all trending up in course enrollment.

Johnson Community College introduced a Data Science Certificate in fall of 2017 with nine courses, eight of which are newly developed data courses. They have consulted with and developed partnerships with 13 businesses, all of whom provide work and career opportunity for big data program graduates.

Our panel presentation will enable participants to learn about our Data Pathway Development Process through the experiences of our community college partners. We will engage attendees in a rich discussion of the need for middle-skilled data workers, will share the experiences of each college and their successful data programs, and will encourage participants to share their own experiences in developing data courses and programming.

Developing a Data Program?

A primary goal of our ATE-funded projects is to share the tools and process with as many colleges as possible. Access to the development process tools and materials is provided through the Oceans of Data website [<http://oceansofdata.org/our-workforceprep-work>]. Colleges interested in developing data programming are invited to continue the conversation by joining our online community of practice. For more information about this ongoing work, contact the project team.

References

- Bloomberg Next (2018) *Building Tomorrow's Talent: Collaboration Can Close Emerging Skills Gap* [White paper]. Retrieved October 4, 2018 from Bloomberg BNA: <https://forms.workday.com/en-us/whitepapers/bloomberg-building-tomorrows-talent/form.html>
- Bradley, B., Restuccia, D., Rudnicki, C., Bittle, S. (2017). *The Digital Edge: Middle-skill workers and careers*. Retrieved October 3, 2020, from Burning Glass Technologies: <https://www.burning-glass.com/research-project/digital-skills-gap/>.
- Columbus, L. (2017, May 14). *IBM Predicts Demand for Data Scientists Will Soar 28% by 2020*. Retrieved on October 4, 2019 from Forbes: <https://www.forbes.com/sites/louiscolumbus/2017/05/13/ibm-predicts-demand-for-data-scientists-will-soar-28-by-2020/#53b5838a7e3b>
- Malyn-Smith, J., Ippolito, J. (2017). *Profile of the Data Practitioner*. Retrieved October 16, 2019 from Oceans of Data Institute: <http://oceansofdata.org/sites/oceansofdata.org/files/FINAL%20Data%20Practitioner%20Profile-September%202016.pdf>
- Manyika, J., Lund, S., Chui, M., Bughin, J., Woetzel, J., Batra, P., Ko, R., and Sanghvi, S. (2017). *Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation*. Retrieved October 16, 2020 from McKinsey Global Institute: <https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages>
- Markow, W., Bragnaza, S., Taska, B., Miller, S., & Hughes, D. (2017). *The Quant Crunch: How the demand for data science skills is disrupting the job market*. Retrieved October 12, 2020 from Burning Glass Technologies: https://www.burning-glass.com/wp-content/uploads/The_Quant_Crunch.pdf
- National Science Foundation (2019). *Harnessing the Data Revolution (HDR) at NSF*. Retrieved October 12, 2020 from the National Science Foundation: <https://www.nsf.gov/cise/harnessingdata/>
- Sasso, M. (2018). *This Is America's Hottest Job*. Retrieved October 15, 2020 from Bloomberg News: <https://www.bloomberg.com/news/articles/2018-05-18/-sexiest-job-ignites-talent-wars-as-demand-for-data-geeks-soars>.
- World Economic Forum (2019). *Data Science in the New Economy: A new race for talent in the Fourth Industrial Revolution*. Retrieved October 16, 2020 from the World Economic Forum: http://www3.weforum.org/docs/WEF_Data_Science_In_the_New_Economy.pdf.

Acknowledgements

These projects are funded by the National Science Foundation, NSF DUE grant numbers 1501927 & 1902568. Any opinions, findings, and conclusions or recommendations expressed in these materials are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.