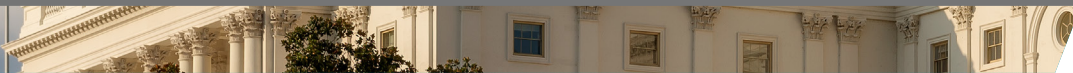# federal CDO council

Data Skills Training Program

CASE STUDIES

This series was developed by the Chief Data Officer Council's **Data Skills & Workforce Development Working Group** to provide support to agencies in implementing the Federal Data Strategy's Agency Action 4 gap-closing strategy training component in FY21.

Contents

Key

To help determine if a case study is relevant to your needs, all cases include a tag from each of the three categories below.

Agency Size

Large Agency   An agency with greater than 10,000 employees

Small Agency   An agency with 10,000 employees or fewer

Program Scale

Pilot   A small-scale evaluation that takes place over a short period of time designed to test your new training approach and identify any deficiencies before substantial resources are committed

Full Program   A full-scale data skills training program should factor in learnings from a pilot, if one was done, for replication on a larger scale, including necessary adjustments to resource allocations and budget estimates for future budget

Training Format

Online Learning   Trainings are offered in an entirely virtual format

In-Person Learning   Trainings are offered in-person only at the sponsoring

Blended Learning   Trainings are offered through a combination of online and in-person learning

CASE STUDY

HHS Data Science CoLab

| Large Agency | Full Program | Blended Learning |

🔒 Agency

The Department of Health & Human Services (HHS) is a federal agency with more than 80,000 employees.

ℹ️ Contact

Bishen Singh, *Senior Advisor, Office of the Assistant Secretary for Health*

Bishen.Singh@hhs.gov | *(301) 337-5785*

👥 Key Stakeholders

» Office of the Chief Technology Officer (CTO) and Office of Business Management and Transformation (OBMT), which launched the Data Science CoLab program, and the Office of the Chief Data Officer (CDO) which now manages the program.

» Biomedical Advanced Research and Development Authority (BARDA) and the Office

of the Inspector General (OIG) which provided classroom space for in-person instruction.

» Educational institution which is contracted to teach the curriculum.

» Human Resources (HR) and Information Technology (IT) offices across the Department which facilitate and support employees'

🎯 Objectives

Design, implement, and sustain a cohort-based data science training program. The primary goals include:

☑ Building a data skills community of practice internally across the agency and providing opportunities for employees to network and enhance their data skills.

☑ Increasing internal capacity for data skills to decrease dependency on contracting.

☑ Generating increased value from existing datasets to drive innovation in untapped areas across the organization.

🧳 Program Overview

The Data Science CoLab provides the opportunity for cohorts of 30-60 participants, including HHS employees as well as contractors, to learn basic and intermediate data skills.

These data skills are interactively applied as participants work with their own program specific datasets and complete a capstone project, such as the development of a tool, model, or system.

The first two cohorts in 2017 and 2018 were eight weeks long and met twice a week, in-person, for eight hours each day. The third cohort in early 2020 blended in-person and virtual learning in response to the COVID-19 pandemic, and the fourth cohort in late 2020 was fully virtual. The late 2020 cohort was adjusted to a two-week basic course for beginners, followed by a six-week intermediate course for graduates of the basic course and those who already have at least basic-level data skills.

> **The Data Science CoLab has increased HHS' capacity for data analysis and has further enabled greater cross-agency collaboration."**

Overall, the Data Science CoLab has increased HHS' capacity for data analysis and has further enabled greater cross-agency collaboration due to the unique combination of capstone projects and structured data skills training, as well as the diversity of the cohorts.

💡 Previous Data Skills Training Approach

This program was the first attempt at a comprehensive, cohort-based data skills training program for the agency. Previously, small-scale one-off training sessions such as day-long seminars and brown bag lunch sessions on basic data skills were occasionally offered.

The program was in part inspired by the HHS Ignite Accelerator, which is an internal innovation start-up program where employees can bring their data challenges and identify solutions, but employees would sometimes lack the necessary skills or capacity to be able to follow through with those solutions.

Key Inputs

» **Funding**: The training program cohorts are only able to take place once funding is available, based on the fiscal year calendar. One cohort of 30 participants costs approximately $75,000 for the eight-week course at a cost of approximately $2,500 per participant, which is covered by the agency.

» **Curriculum**: The agency designs the curriculum for the program, and instruction is provided by a contractor. In recent cohorts, participants from prior cohorts have become involved with helping the new cohorts through the community of practice by mentoring the participants in the current cohort and providing advice and support on

the capstone projects. Participants attend class in the mornings, have lunch together, and then work on their homework and capstone projects in the afternoons in the presence of teaching assistants (TAs).

» **Staffing**: Three employees support the program, including a GS-11 and GS-14 under the Management & Program Analysis job series, and a GS-13 under the IT Management job series. The contractor provides an instructor and teaching assistant for the eight-week program's courses.

» **Accommodations**: Participants in the program receive approval from their supervisors to devote 16 hours per week to the program. The participants resume full-time work after completing the program and return with enhanced data skills and a completed capstone project that can be applied to their assigned duties to increase productivity and/or otherwise improve outcomes.

Key Implementation Activities

1. **Identify a Training Provider**: The Data Science CoLab program and curriculum were conceptualized in late 2016 and early 2017 and a contractor was selected through a proposal process to administer the educational component of the program. This process took roughly 6 months.

2. **Secure Funding**: Initial funding for two cohorts was secured by receiving executive buy-in and the pilot program was announced in early 2017.

3. **Select Participants:** Employees from across the agency were invited through an agency-wide email to apply to the program. Over 100 applications were received, applicants were screened and invited to interview based on

a defined criterion, and 30 interviewees were then selected for participation in the program based on an interview grading rubric which was developed by the implementing team.

4. **Pilot the Program**: The pilot program began in mid-2017 and ran for eight weeks. The curriculum focused on Python programming since it was a widely used and in-demand data science skill. The pilot program was well-received and was authorized to be continued through the ReImagine HHS Data Insights Initiative.

> **" The pilot program was well-received and was authorized to be continued through the ReImagine HHS Data Insights Initiative."**

Important Program Milestones

» The same pilot process was repeated for a full-scale program in 2018 and a new cohort of 30 participants was selected from around 400 applications. The format was the same as the pilot program, except the curriculum for the full-scale focused on R programming instead of Python because R was discovered

to be more commonly used across the agency during the pilot.

» The same process was repeated in early 2020 once additional funding was secured and a new cohort of 30 participants was selected from around 500 applications. The format was the same as the previous cohort, but

in response to the COVID-19 pandemic, this cohort transitioned midway from in-person to virtual learning.

» The program was transitioned from the Office of the CTO and OBMT to the Office of the CDO at the end of FY20 as part of the completion of ReImagine HHS, but the program's staff remained the same.

» A new, fully virtual cohort of 60 participants was selected in mid-2020 from around 800 applications based on available funding and the same selection process, this cohort finished in late 2020. In this cohort, instead of one eight-week course, a two-week basic and six-week intermediate curriculum were offered.

Challenges

» Securing a consistent and recurring source of funding. As a result, new cohort cycles can only take place once funding is secured and the number of interested participants has outpaced the amount of funding available. For example, there was no cohort in 2019 due to a lack of funding.

» Technical considerations such as coordinating IT and access to data and software requires approval, planning, and time. Additionally, logistical considerations such as building access, meeting room

reservations, room setup with the right equipment and technology, etc. can easily be overlooked.

» Fulfilling the demand for data skills training was also a challenge, as there are many more applicants to the program than there are spaces in a cohort. Additionally, the in-person cohorts required in-person participation in the Washington D.C. metro area, while the most recent virtual cohort expanded access to HHS employees located throughout the world.

Lessons Learned

» Collaborate with stakeholders, such as division and executive-level leadership within the agency to enhance awareness and buy-in for the program.

» To foster the connections that build a true "community of practice" beyond the eight-week course, the program can provide additional incentives to collaboration beyond just participation in the program.

» Anticipate the level of demand based on historical data before launching the program to be able to more effectively plan for the number of resources and amount of time required, as well as to bridge potential gaps, such as the limited number of spaces available in the cohort.

Best Practices

» Design and implement an in-person cohort training program and adapt it to accommodate the virtual learning and working environment.

» Partner with an outside training provider to design a curriculum tailored for the agency, while internally building a community of practice through collaboration, teambuilding, and a network of data science training alumni.

» Recruit high potential applicants and select them for interviews based on a defined selection criteria, then evaluate them based on a defined interview grading rubric to select the best candidates for participation in the program.

» Administer a pilot program first and then develop an implementation playbook from the lessons learned for replication as a full-scale program.

» Attain departmental and operational leadership buy-in to ensure sustainability.

» Utilize new insights from performance metrics and survey feedback to continually enhance the program and curriculum.

Future Strategy

» Secure a permanent budget to enable the program to run more consistently and frequently with training offerings advertised well in advance. This will enable more employees to be trained, and employees can more effectively plan for their participation in the program, since they will know when the training opportunities will take place.

» Continue to gather data on the program from participant interviews, feedback, surveys, and performance metrics to continually adjust the program and curriculum. Particularly important is understanding the degree to which participants can apply the data skills they learn in the program to their jobs after completing the program.

» Conduct a full skills assessment regarding data skills in the agency to identify gaps and in-demand data skills and incorporate these results into the program's curriculum.

» Once in-person learning can take place again, consider the initial in-person components with the current virtual components to understand if a blended learning approach may be a viable format for future cohorts and determine which aspects should be in-person or virtual.

" Particularly important is understanding the degree to which participants can apply the data skills they learn in the program to their jobs."

Air Force Public Sector Data Governance Professional Training Certification Pilot

Large Agency **Pilot** **In-person Learning**

Agency

The Department of the Air Force is a federal agency under the Department of Defense (DoD) with approximately 330,000 active-duty employees.

Contact

Eileen Vidrine, *Chief Data Officer, Department of the Air Force*

eileen.vidrine@us.af.mil | *(703) 697-9823*

Key Stakeholders

»   Air Force Chief Data Office which owns and manages the program.

»   Air Force contracting office which manages the training procurement and contracting requirements.

»   Education Provider which is contracted to provide the training.

Objectives

Upskill Air Force data officers to support the emerging need for data governance. The primary goals include:

☑   Improving workforce readiness and data governance standards.

☑   Increasing data literacy and acumen across the agency.

☑   Building a community of practice for data officers to network and gain critical data skills.

Program Overview

The Public Sector Data Governance Professional (PSDGP) Training Certification Pilot program was established to support the Air Force's vision for data management, digital technology, and the future of artificial intelligence within the agency and other branches of the military.

The PSDGP training is a three-day course that prepares chief data officers at the Air Force for data governance certification. This professional certification program utilizes the Institute for Certification for Computing Professionals (ICCP) federal sector data governance certification and a Registered Education Provider from the Project Management Institute (PMI) to provide the training.

The course topics include: Public Sector Data Governance Mission Drivers and Deliverables, Data Governance Roles and Responsibilities, and the Legal and Regulatory Environment in which public sector officials and service providers must operate. The certification requires passing the Public Sector Data Governance Test upon completion of the course, a 90-minute exam administered by the ICCP-approved exam proctors.

Previous Data Skills Training Approach

The PSDGP training builds a foundation for Air Force Chief Data Officers."

Ad hoc training by agencies without an enterprise focus. One of Air Force's previous initiatives, Digital University, was started to enhance the baseline skills across the agency and the PSDGP Training Certification Pilot program was aligned to the standards developed for a professional certification.

The PSDGP training builds a foundation for Air Force Chief Data Officers (CDOs), as previously there was not a centralized skills assessment that employees could take classes for.

Key Inputs

» **Funding:** Funding was provided from the allocated Chief Data Office budget for the fiscal year and covered the cost of training for their employees.

» **Course Material:** The course was developed and delivered by the vendor using instructional systems design (ISD), a systematic process for the assessment and development of training solutions.

Key Implementation Activities

1. **Conduct a Skills Assessment**: In the previous fiscal year, the Air Force Institute of Technology conducted a skills assessment across the Air Force in partnership with a vendor. The assessment results were then used to develop a series of follow-on questions as part of a data maturity assessment to determine the baseline of the

workforce. Leveraging the requirements listed in Section 3520 of the OPEN Government Data Act, 14 functions were established that CDOs must carry out for their agencies including but not limited to transparency, lifecycle data management, data use, data assurance and data infrastructure improvements. The data maturity assessment led to the implementation of the data governance training.

2. **Pilot the Training**: Prior to implementing the training program at the Air Force, the Chief Data Office sent two employees to pilot the course. It was then determined the training was adequate to provide to their workforce.

3. **Secure Funding**: A business case was made through a written statement of objectives to the Air Force contracting office to secure additional funding. Funding was then made available at the time the training program was piloted.

4. **Establish a Contracting Vehicle:** The Air Force CDO established a three-year Blanket Purchase Agreement (BPA) between the agency and the training vendor through support from their contracting office. Establishing a BPA saved time on a repetitive need for the training course, allowing the Chief Data Office to purchase the training on an as needed basis. It also enabled the agency to negotiate a discount for a course with the vendor for a larger number of participants. Other Air Force agencies are approved to utilize the BPA.

> **Establishing a BPA saved time … and enabled the agency to negotiate a discount."**

Important Program Milestones

» **Contract Award**: The Air Force CDO, contracting offices for both the Air Force and the vendor, and the Air Force contract office representative (COR) were the key stakeholders in the creation of the BPA. There was only one vendor providing the desired training course and that vendor was selected as the education provider for the training program.

» **Pre-Planning**: The Chief Data Office worked with the vendor to begin planning for the first course which was originally intended to take place in person but was transitioned to a virtual learning format in response to the COVID-19 crisis. The Air Force COR and the vendor tested and determined an appropriate online tool to host the training that was accessible on both government and personal computers.

» **Registration**: The vendor managed the registration process and ensured participants for the training course had access to the online training tool. At the advice of the vendor as well as budget considerations, the training course accommodated 25 seats.

» **Course Execution:** Post contracting award, the program had the training running within two months. The first course was initiated in June 2020 as a pilot and included members of the Air National Guard and the Navy. Through the existing BPA, the Air Force has provided three large training courses to date.

» **Certification Testing**: Upon immediate completion of the training course, participants took a 90-minute online-proctored certification exam. A total of 80 employees have taken and passed the training, including the two employees who piloted the course in the beginning.

» **Participant Feedback**: While a formal evaluation process is underway, participants of the course have begun providing feedback, sharing insights into areas where the workforce wants to invest their time.

Challenges

» The volume of the course content was challenging. Due to the synchronous learning format, participants were drained from the overload of information provided in the three days of training. Also, the program had to factor working with the different time zones which led to the dense virtual course outline.

» Transitioning from an in-person to a virtual learning format was a challenge. The program had to quickly pivot to a virtual learning format to maintain participant engagement while addressing the agency's response to the COVID-19 crisis.

Lessons Learned

» The program may need to be expanded and broken up into smaller segments over a longer period to cover the extensive amount of information provided in the course.

» The program must be agile to adapt to the virtual learning and working environment. There were collaboration opportunities for engaging participants in an online environment and accessibility challenges for those using government computers.

Best Practices

» Conduct a skills and maturity assessment to baseline your workforce and better understand the training needs for your agency.

» Implement feedback surveys and utilize the insights gleaned to continually adjust and enhance the program and training course. Work with the training vendor to advocate for those adjustments to ensure proper alignment with federal requirements and the agency's employee roles.

.

» Leverage BPAs as a contract vehicle for acquiring training services quickly and on an as needed basis, especially when the demand for training is not certain (i.e., number of training offerings provided in a year).

» Use training opportunities as a platform to build a community of practice. The training program helped to identify and integrate data professionals across and into all facets of the Air Force.

> **Use training opportunities as a platform to build a community of practice."**

Future Strategy

» Continue providing the training once a year at minimum, to provide a baseline for new and upcoming data officers.

» Consider expanding the training and shortening the coursework hours from full days to half days to reduce the overload of information in the current course material and to accommodate participants in different time zones. Participant feedback also indicated wanting more time to study as they felt rushed to take the exam so changing the training format would also solve this issue.

» Develop a formal program evaluation process with help from Air Force CDOs who have taken the course.

» Partner with the Air Force talent management community to continue to build data acumen across the workforce in support of career mobility and flexibility to upskill new talent quickly, adapting this baseline course and sustaining the business of data and digital technology acumen. The Chief Data Office also wants to expand this line of education by helping agency Airmen and Guardians take data courses at the Air Force Academy, the Air Force Institute of Technology, and through Digital University.

" Participant feedback also indicated wanting more time to study as they felt rushed to take the exam."

CDC Data Science Upskilling Program

Small Agency   **Full Program**   **Blended Learning**

Agency

The Centers for Disease Control and Prevention (CDC) is an agency within the Department of Health & Human Services (HHS) and has around 10,000 federal employees.

Contact

Mary Bertulfo, MPH, *Acting Program Director and Program Coordinator*

ktt8@cdc.gov | (404) 747-3983

Key Stakeholders

» Division and agency-level leadership which advocate for the program and help secure funding.

» Subject matter experts (SMEs) within CDC who initially helped design the program and continually provide advice and support.

» Education vendors which provide the learning platforms and data science training.

» Non-profit contractor that supports program design and strategy.

Objectives

Design, implement, and sustain a cohort-based data science upskilling program. The primary goals for the Data Science Upskilling (DSU) program include:

☑ Enhancing the data science capabilities of the CDC workforce in support of its data modernization initiative.

☑ Building a data science community of practice to enable networking, peer-learning, and resource sharing.

📦 Program Overview

The DSU program provides the opportunity for an annual cohort of around 80 learners, including CDC full-time employees (FTEs) and fellows, to learn a wide variety of data science skills.

Teams of around three to five learners on average, who are aligned on specific projects, are nominated to join the cohort, and work together on their capstone project as a part of the cohort. The program uses a blended in-person and virtual learning method and the training courses are curated from a variety of vendors. A community of practice is formed through the cohort as teams collaborate on projects and learn from each other. Learners participate in the program on a part-time basis, around five to ten hours per week, in a bi-weekly format and each cohort lasts around one year.

> **"** The DSU program provides the opportunity for an annual cohort to learn a variety of data science skills."

The pilot cohort ran from November 2019 to July 2020. A second cohort was launched in August 2020 and will complete in July 2021..

💡 Previous Data Skills Training Approach

A landscape analysis and pilot project found very few formal training opportunities around data science at the agency. One-off training from certain vendors were offered, but no comprehensive training program covered the variety of data science literacies. The implementing Division began offering a Public Health Informatics Fellowship Program for several decades prior to the pilot program launch.

🌐 Key Inputs

» **Funding**: Initial seed funding for the pilot was allocated from end of year funds.

» **Data Science Training:** The agency curates training material, certifications, a learning platform, and learning resources, from a variety of vendors based on the needs of the cohort. The agency also facilitates internal guest lectures.

» **Staffing**: Initially two employees supported the pilot program, currently three full time employees and three contractors support the program. These include an acting program director who is also the program coordinator, one program assistant (contractor), two instructional designers (contractors), and two subject matter experts.

» **Accommodations**: Participants in the cohort program are nominated and approved at the Branch, Division, or CIO-level to devote around five to ten hours per week to the program. The participants resume full-time work after completing the program and return with enhanced data science skills and a completed project that supports their duties as assigned.

Key Implementation Activities

1. **Understand Current System**: A landscape analysis of workforce training opportunities was conducted, and data skills training was identified as a gap.

2. **Propose the Program**: A proposal for the data skills training program was written that explained how the program was connected to the CDC public health data modernization initiative and defined the data skills training gap it would address. Initial planning began in June 2019.

3. **Secure Funding:** Proposal for initial funding for a pilot cohort was granted once leadership buy-in was obtained, which resulted in the program receiving end of year funds. The implementation set-up process began in August 2019.

4. **Select Participants**: Employees and fellows from across the agency were nominated as teams aligned on specific projects to participate in the program by branch, division, or CIO-level leadership. The projects were selected based on a scoring rubric which contains specific criteria such as alignment to agency priorities. Generally, when a nominated project is selected, the nominated team members associated with the project nomination are all accepted if they meet the requirement of being an FTE or a fellow.

5. **Pilot the Program**: The pilot program began in November 2019 and concluded in July 2020. The pilot program was well-received, and the program looks to scale-up in the future.

> **The pilot program was well-received and looks to scale-up in the future."**

Important Program Milestones

» Throughout the pilot, the implementing team generated outputs and communicated to leadership about the impacts of the program to attempt to secure additional funding.

» An agency-wide symposium was held at the end of the pilot year in July 2020 where learners presented on their projects to the entire cohort.

» A steering committee including agency stakeholders and subject matter experts was established.

» A similar process was repeated for a full-scale program to run from August 2020 to July 2021 once consistent, yearly funding had been secured.

» A non-profit contractor is currently assessing the program through data analysis and interviews and will deliver a gap analysis report to support the ongoing design and strategy of the program.

Challenges

»   For learners, balancing the learning responsibilities and time commitment while also working at the same time. Learners are well-supported by their supervisors and are working on projects supporting their duties as assigned, but allocating time for learning continues to be a major challenge.

»   Transitioning from a blended in-person and virtual learning environment to a fully virtual learning environment in response to the COVID-19 pandemic. Maintaining peer and team-based learning is more difficult to accomplish while virtual, but project check-ins, practice problem sessions, and use of a videoconferencing communication platform have been helpful.

Lessons Learned

»   Identify potential mentors and establish a mentor network for the learners. While important to build workforce capacity, it is a challenge to not have enough mentors for the number of learners being trained in the program.

»   Collaborate with stakeholders within the agency to enhance awareness and buy-in for the program.

»   Communicate informally with others who are running data skills training programs at different agencies to understand and mitigate challenges, as well as to come up with new program ideas.

> "Collaborate with stakeholders to enhance awareness and buy-in for the program."

»   Anticipate the number of resources required to run the full-scale program when scaling up a pilot program. Be ready for multifaceted requirements and to hire additional resources.

Best Practices

»   Design and implement a blended in-person and virtual cohort training program and adapt it to accommodate the virtual learning and working environment.

»   Partner with multiple vendors to curate a variety of training offerings to enable diverse training opportunities.

»   Develop a nomination and selection process to identify high priority projects and the associated teams for participation in the program.

»   Administer a pilot program first and then use the pilot learnings and outcomes, such as the agency-wide symposium, to establish sustainability of the program through departmental and operational leadership buy-in.

»   Gather performance metrics and conduct interviews to perform a gap-analysis for the program and use these insights to continually enhance the program and curriculum.

Future Strategy

» Consider the program design and strategy when adjusting the curriculum based on the upcoming gap analysis report.

» Currently, the program is cohort-based. In the future, the program should be prepared to offer different tracks in the future to better meet the specific needs of individuals, such as those who require basic data science literacy or those who require complex data science training.

» Expand the program to support skill development in independent learners who want to further their careers in data science but who may not be a part of a team that would be nominated to work on a data science project.

» Leverage the data science community of practice and develop a more formal mentor network to support the growth of learners.

CASE STUDY
Census Bureau Data Science Training Pilot

| Small Agency | Pilot | Blended Learning |

🔒 Agency

The U.S. Census Bureau is a federal agency within the Department of Commerce (DOC) with nearly 5,000 federal employees, excluding field workers.

ℹ️ Contact

Lisa Frid, *Workforce Transformation Program Manager, Human Resources Division*

lisa.frid@census.gov | *(301) 763-5543*

💬 Key Stakeholders

» Human Resources (HR) Division which owns and manages the program.

» Information Technology (IT) Division which facilitates and supports employees' participation in the program as it pertains to technology needs.

» Communications Division which supports internal program marketing and communications.

» Human Capital Council which provides the long-term vision for the program and is a decision-making body of executives from every major area of the Census Bureau, referred to as program areas in this case study.

» Data Science Advisory Council which supports the day-to-day operations and decisions of the program and includes subject matter experts (SMEs) from across the organization with varying levels of data science expertise.

Objectives

A business transformation initiative to provide a coordinated and systematic approach to technical training across the organization. The primary goals for the Data Science Training Pilot include:

- ☑ Upskilling technical employees that use data science to effectively solve problems and find insights as part of their role,

- ☑ Engaging employees about the evolving space of data science through learning and education, and

- ☑ Promoting innovation at the Census Bureau.

Program Overview

The Census Data Science Training Pilot is a competitive cohort-based continuous learning program available to all Census employees, excluding contractors. The program has three main elements:

1. Blended online approach with coursework provided through a vendor, supplemented by live virtual sessions with in-house Census data science experts.

2. Hands-on learning through capstone projects where participants can apply and reinforce data skills to relevant Census challenges benefiting agency missions.

3. Ongoing mentorship from Census data science SMEs who provide specialized advice and feedback as participants undergo training and deliver insights into the actively changing data science landscape within the organization.

The part-time program is six months and the first cohort, commenced in January 2020, required 8 hours a week of the participant's time. In the second and current cohort, that time commitment was increased to 12 hours of each participant's time. Designed intentionally for retention, participants of the program must sign a continuing service agreement of 15 additional months at Census after completing the training program following Office of Personnel Management (OPM) guidelines.

Previous Data Skills Training Approach

Uncoordinated efforts spread across the organization utilizing individual program resources. Training examples included sending staff to outside training or extensive academic programs which became increasingly cost prohibitive. The Data Science Training Pilot is the evolution and expansion of various training approaches at Census.

Key Inputs

» **Funding**: The program received partial funding of $300K from the Office of Management and Budget (OMB) through the Federal Chief Information Officers Council (CIOC) FY18-19 CAP Goal and CXO funding request. This helped acquire needed software and training tools and brought program startup costs down. Since the program is an in-house training offering, it only costs program divisions $400 to send staff to the training. This low-cost training approach is a huge advantage as many program areas are funded through reimbursable funding agreements.

» **Staffing**: In addition to existing contract support, the six-month program is resourced through volunteers from across the organization. There are roughly 40 program volunteers that include the Advisory Council, mentors, capstone project managers and supporters, and IT support. Most volunteers devote 2-3 hours per week while capstone managers spend 5-6 hours per week supporting the HR division.

» **Course Curriculum**: The program worked internally to design a tailored curriculum. Survey and focus group feedback demonstrated that participants returned to their full-time jobs with new and sharpened data skills that were applied to their roles.

> **Survey feedback demonstrated that participants returned to their jobs with new and sharpened skills."**

Key Implementation Activities

1. **Build a Governance Structure**: An executive steering committee was formed with the highest-level executive from each of the Census program areas to provide strategic direction and oversee the training program's progress. At the next level down, an additional advisory council was formed with data science SMEs to help guide the design and implementation of the training program. Much of the pre-work and workforce benchmarking activities were executed by these governance groups.

2. **Secure Funding**: Census can continue funding the training program because of leadership buy-in and OMB's initial sponsorship which put existing infrastructure in place.

3. **Design the Program:**

   a. The curriculum and application process included specific prerequisites for applying to the program such as supervisor support and approval and taking needed refresher courses to ensure all participants have the same foundational knowledge of statistics and data analysis. Participants applied to one of the two learning tracks that will guide them through their training: the data science generalist path, targeted towards employees who have little to no data science training but have math and statistics experience and want to explore the data science field, and the machine learning path, targeted towards those who are familiar with data science at an intermediate level and want to advance their machine learning skills in Python. Capstone projects are selected from new and ongoing efforts at the Census Bureau based on a set of criteria, including impact to agency missions, accessibility of data, and feasibility of completion in the allotted time.

b. The training evaluation included pre- and post-program evaluation to assess the impact of the training. Additional checkpoints were added throughout the duration of the training to assess the usefulness and effectiveness of the program material. Participants cannot graduate from the program until all evaluations are completed. These evaluations provided insights for shaping the future of the program and measuring learning outcomes.

4. **Select Participants**: Employees were invited to apply to the first cohort in November 2019. 125 applications were received and screened by the Data Science Advisory Council, and 51 of those applicants were accepted based on a defined criterion and scoring rubric.

5. **Announce the Program**: Participants were set up for their online learning environments including acquiring licenses for required training tools and the first cohort was announced to the agency in December 2019.

Important Program Milestones

» A needs assessment was conducted in the spring of 2019. The assessment uncovered a skills gap in advanced data science techniques. The insights from the assessment were then used to develop a program plan for the launch of the pilot.

» The capstone project was designed by the advisory councils from September-December 2019, ensuring a plan for secure and ethical use of data. During this time, program staff also searched for volunteers to teach the in-person courses as part of the coursework component of the program.

» A communications campaign was developed and then executed in October and November of 2019. An internal SharePoint page was set up as the central portal for program information.

» The admission period was from November- December 2019. 125 applications were received.

» The program was piloted in January 2020, the coursework ran through March and the capstone projects began in April for the remainder 10 weeks of the program. The training program is currently in its second cohort cycle.

» A cohort graduation celebration was conducted in June 2020. The celebration was reimagined from an in-person to virtual event in response to the COVID-19 crisis. Senior executives and interested parties were invited. Participants presented their capstone projects and were celebrated for their accomplishments.

Challenges

» Gaining access to training software and tools, such as Python and Tableau, required extensive approval and time. In some instances, these modern tools were deployed at the agency for the first time so ensuring the right level of access resulted in different layers of approval.

» Participants lacked enough time to prepare for and understand the capstone project. Participants wanted to know more about the capstone project during the coursework component of the training, including the type of data being used, so they could dedicate time setting up their technical research environments and working teams as needed.

✎ Lessons Learned

» Communicate the program curriculum upfront and the application expectations clearly to yield more qualified participants and lessen the number of applicants.

» Define a clear scoring rubric containing criteria for the desired knowledge and skill level to ensure overqualified candidates do not apply.

» Participant feedback from the first cohort indicated more time was needed to grasp the concepts and techniques taught in the program curriculum.

» The shift to a fully virtual work environment revealed a need for more check-ins between program staff and participants and mentors to drive engagement.

" Participant feedback indicated more time was needed to grasp the concepts and techniques."

♜ Best Practices

» Engage leadership in the initial phases of the program to garner buy-in early on.

» Establish a governance model tailored to the organization's specific needs to ensure appropriate stakeholder engagement.

» Learn from within the organization, including previous efforts and other similar programs to understand what has worked in the past.

» Consider a cohort style training format to offer the opportunity for participants to network and share learnings across different parts of the organization.

» Incorporate informal learning opportunities, giving participants a chance to practice and refine non-technical skills like presentation and communication skills.

» Use a framework or model, such as the Kirkpatrick model, for analyzing and evaluating training and continuous learning program results.

» Get participants to practice the new techniques learned as soon as and as often as they can to promote knowledge retention. Support this through capstone-type projects that tackle real agency challenges.

🗎 Future Strategy

» Acquire additional program resources. For a six-month program and a cohort of roughly 50 participants, five full-time staff is ideal for ongoing development, customer service, and training implementation.

» Pilot a new curriculum with advanced level courses for the advanced learner path.

» Consider different tiers of complexity or "cafeteria style" program offerings based on the needs of employees.

Data Science @NLM Training Program

Small Agency   Full Program   Blended Learning

Agency

The National Library of Medicine (NLM) is one of the 27 Institutes and Centers within the National Institutes of Health (NIH) with approximately 1,700 employees.

Contact

Lisa Federer, *Data Science and Open Science Librarian* | *lisa.federer@nih.gov*

Dianne Babski, *Acting Associate Director of Library Operations* | *babskid@mail.nlm.nih.gov*

Key Stakeholders

»   Leadership including executive sponsors, mid-level managers and supervisors who relay the value of this program to employees and other executive stakeholders.

»   Agency's Information Technology (IT) Office who coordinate with program staff to ensure participants have the appropriate software licenses and tools.

»   Data Science @NLM Training Team with five representatives from various divisions, serving as a planning and steering committee.

Objectives

A key component of NLM's 2017- 2027 Strategic Plan is ensuring a data savvy biomedical workforce and a data ready public. NLM embarked on achieving this goal by introducing an inclusive training program to increase data skills and elevate data literacy across job functions and applicable to a variety of NLM career tracks.

This institute-wide training program allows staff across all NLM divisions to fully leverage the data science ethos at the agency, by giving each employee the opportunity to improve their data literacy and giving them access to hands-on data science training.

🧰 Program Overview

The Data Science @NLM Training Program provides participants with individualized training plans across various skill levels to apply data science skills in their job functions. The program is open to NLM staff and contractors.

Critical to the planning of this program was the process of interviewing 60+ staff members to understand how exactly data science fits across various job functions and organizational priorities. The program utilized introductory sessions to communicate the value of this training and implemented multiple evaluations and assessments to inform the design of the training plans. Each participant received a customized training plan curated to meet their learning goals and skills gaps. The length of each training plan varied depending on the employee's selected goals.

> **Critical to this program was interviewing staff members to understand how data science fits across job functions and priorities."**

💡 Previous Data Skills Training Approach

This is the first institute wide training program implemented at NLM across all divisions.

🌐 Key Inputs

» **Funding**: The Associate Director of Library Operations invested a portion of the division's budget dollars to this program, but there continues to be no dedicated budget. The program received 1.3 million in funding in the first year.

» **Course Catalog**: 200 courses from online platforms and NIH resources created to guide NLM staff and meet ongoing needs of the workforce. Each individual learning plan featured classes from this catalog. This serves as an ongoing resource that can be used in future iterations of this program. While formal evaluations are planned for the future, NLM has begun to receive positive feedback from participants on the training material.

» **Staffing**: Five members from across NLM's divisions dedicated time as subject matter experts, they received contractor support to lead the design and implementation of the program.

> **NLM has begun to receive positive feedback from participants on the training material."**

Key Implementation Activities

1. **Planning Period**: Starting in December 2018, the contractor interviewed 60-70 staff members across NLM to understand the data training needs across the organization. They also analyzed documents and reports from NLM and NIH for additional insights into the skills needed to meet mission goals. These efforts served as a landscape analysis to inform NLM on how data science skills fit into the organization and where the gaps in those skills existed. Next, the contractor developed a training roadmap to help NLM leadership understand what areas were in need of data science training. Lastly, staff interviews helped them develop eight skill development profiles to guide staff on selecting a data science trainee type. The contractor tied the profiles to specific use cases related to data science work at NLM.

2. **I**ntroductory sessions: In January 2019, introductory events began including a kickoff with all staff and an overview session for supervisors. The training kickoff event served as an all hands meeting for staff to share all necessary details on the program. The supervisor overview event created an environment for leadership to understand the importance of the program and learn how they can best support staff coordinating this effort.

These sessions elevated employee understanding of the importance of the data skills training and how it relates to their job duties.

3. **Program Design**: In March 2019, all employees participated in a Data Science Readiness Survey helping create individual training plans and assigning staff a set of classes aligned with their learning goals. Over 100 employees showed interest in a more intensive data science fundamentals course that kicked off a few months later in June, for which 25 seats were available. In April, employees worked with supervisors to finalize individual training plans. In the same month, the data science basics course was provided to all staff and kicked off the training program. In May, everyone began work on their individual training plans. The training plans were curated by analyzing the results from the Data Science Readiness Survey and were aligned with each participant's skills gaps and learning needs.

> **"These sessions elevated employee understanding of the importance of the training and how it relates to their job duties."**

Important Program Milestones

» January 2019 marked the kickoff of the programming and first training event attended by 500+ staff members. This event was an introductory session explaining data science concepts in a simplified format.

» January 2019 also marked the first supervisor briefing, more than 100 supervisors attended multiple program activities.

» March 2019 began the participation in the Data Science Readiness Survey, 700+ participants completed this survey.

» April 2019 participants and supervisors worked together to finalize the individualized training plans.

» April 2019 also marked the start of the first training class, a 101 session on data science terminology and principles attended by 500+ participants.

- » May 2019 participants began work on their individualized training plans.

- » June 2019 marks the beginning of an in-depth data science fundamentals course requested by 25 participants.

- » August 2019 marked the wrapping up of the first-year training cohort through a celebratory event, a data science open house attended by 300+ employees.

Challenges

- » Securing buy-in from staff members across the divisions was a roadblock early on. Many employees viewed this type of training as only appropriate for data scientists, initially failing to understand the importance of basic data literacy in their individual roles. This was addressed by communicating to participants the individual level program benefits and vast training opportunities available for non-data scientists.

- » Participants faced a gap in accessing hands on project opportunities to apply the skills taught by this program. This was addressed by providing an additional mentorship component to the second-year training program cohort. Mentors are matched with learners to help guide them through a capstone project by giving them advice and guidance from practical experience that they can apply to the skills they have acquired.

> **Participants faced a gap in accessing hands on project opportunities to apply the skills taught by this program."**

- » The lack of a dedicated budget and full-time staffing for this program continues to pose a challenge. Currently, five staff members dedicate additional time on top of their full-time duties to run this program.

Lessons Learned

- » **Securing Leadership Buy-In**: Program staff found through post-program evaluation that participants who made the most progress in their individual training plans had better supervisor involvement. The NLM program coordinators found having buy-in from all levels of leadership was crucial.

- » **Messaging**: Marketing the program effectively to staff by breaking down data science terminology through 101 sessions and clear messaging allowed participants from all data skills levels to see value.

- » **Practical Application**: Program staff learned that a capstone project was necessary to include to help participants utilize hands-on application of their training. The capstone project was incorporated into the second-year curriculum.

Best Practices

» Establish a mentorship component for specialized advice and feedback to guide participants through capstone projects and as they undergo their training.

» Employ strategic messaging to impress upon employees the importance of data skills training across all job types and connect the value of the training in advancing specific skill sets.

» Conduct pre-program assessments through staff interviews and analyzing agency documentation, to design a training program that addresses gaps in data science skills and understand the impact training can have across the organization.

» Include individualized training pathways, giving an opportunity for all employees to gain data literacy and data science skills, regardless of job type or years of experience. NLM developed eight skill development profiles to give concrete examples of how participants can fit into and gain value from data skills training.

» Utilize existing courses from online web platforms and multiple agency Learning Management Systems (LMS) to make use of free or already existing resources to generate a course catalog.

Future Strategy

» Hire a full-time staff member to administer and evaluate the program.

» Introduce more peer-to-peer opportunities, allowing participants with similar skill levels or interests to partner on group learning projects.

» Conduct formal program evaluations to understand the full scope of impact on NLM's goals and mission.

CASE STUDY
FDIC Data Literacy Pilot

| Small Agency | Pilot | Blended Learning |

Agency

The Federal Deposit Insurance Corporation (FDIC) is a federal agency with over 5,000 employees.

Contact

Jacques Vilar, *Chief Data Officer*

CVilar@FDIC.gov | ChiefDataOfficer@FDIC.gov

Key Stakeholders

» FDIC's Corporate University which is currently developing the curriculum and providing the in-house training.

» Research and advisory services contractor which provided workforce assessment and program design support.

» Division officers and directors who provided advice and input during the design and planning process.

» Office of the Chief Information Officer (CIO) which provided advice and input during the design and planning process.

Objectives

☑ Building and advancing enterprise-wide data literacy through a blended training approach with e-learning modules, instructor led training and brown bag sessions.

Program Overview

The FDIC Data Literacy Pilot Program is an agency-wide training offering with coursework focused on advancing the data literacy of the FDIC workforce. All employees are encouraged to participate and select from five different data skills training personas with a customized training curriculum to fit individual needs.

FDIC's Corporate University develops the curriculum and provides the in-house training for each of the training personas. The trainings are self-paced and utilize a variety of methods from e-learning modules to instructor-led sessions to elevate FDIC's data literacy. The Pilot Program kicked off in Spring of 2021 and will run until Summer of 2021.

> **The trainings are self-paced and utilize a variety of methods."**

Previous Data Skills Training Approach

Ad-hoc data skills training opportunities within various divisions of the agency, there were no consolidated data literacy programs available to all agency employees.

Key Inputs

» **Funding**: Existing funding from the FDIC CDO Office was utilized, no additional funding was requested.

» **Staffing**: One FDIC employee served as the program manager and handled logistics and coordination with the contractor and Corporate University.

» **Training Curriculum**: The FDIC Corporate University developed a combination of instructor led training and e-learning modules for the pilot program.

» **Outreach and Communication**: Meetings with agency, division, and executive leadership to gather their advice, input and buy-in during the planning and design phase of the program.

Key Implementation Activities

1. **Conduct a Needs Assessment:** FDIC worked with contract support to develop a survey to assess the FDIC workforce's data literacy competencies. The survey was given to all employees, including senior executives like the FDIC Chairman. The survey contained data literacy indicators on a five-level Likert scale and participants self-selected where they fell on the data literacy scale, with five being "most data literate" or "conversational". The survey received a 47 percent response rate, and the results revealed a three out of five stars average on the data literacy scale, indicating that most FDIC employees considered themselves at a "competent" data literacy level.

2. **Develop Training Personas:** The results of the survey helped assess what type of training would advance FDIC's data literacy. With most of FDIC employees indicating they were intermediate to advanced levels in data literacy, the contractor created five personas and sub personas describing the various data roles at FDIC which guided the development of each training curriculum type. The different personas included data and analytics sponsors, data and analytics ambassadors, mission enablers, analytic and data scientists, and data architects. For example, an FDIC executive would be assigned as a data and analytics sponsor with corresponding curriculum to address any specific training gaps or needs.

3. **Design the Curriculum**: Program staff worked with the FDIC Corporate University to design the curriculum for the program. Training plans included a mix of e-learning modules, instructor led training and brown bag sessions that participants complete at their own pace. An 18-month timeline was mapped out to pilot the program and then to re-survey the participants to assess if there were any workforce advancements in data literacy.

Important Program Milestones

» The needs assessment survey was issued to all FDIC employees in March 2020.

» The pilot design concluded in late Fall 2020 and the pilot kicked off in Spring 2021.

» The pilot is projected to end in the Summer of 2021.

Challenges

None currently. The FDIC Data Literacy pilot program is still in the implementation phase.

Lessons Learned

» Adjust expected program goals and design based on learnings from the workforce needs assessment.

» Ensure environmental scan and needs assessment equally assesses data literacy knowledge of employees regardless of level or role.

Best Practices

» Engage senior leadership during the planning and design phase to build awareness of the pilot program and gain executive buy-in.

» Leverage existing agency training resources (e.g., course catalog, educational/training entity) to support program needs.

» Market program efforts early on to data literacy or data skills focused agency sub-groups or employee peer groups to engage employees, get peer input and foster buy-in.

Future Strategy

» Adding the FDIC Enterprise Data Council as an additional stakeholder group to engage and solicit input and advice from.

» Dedicating a program staff member whose job function will be 50 percent focused on the management and coordination of the program.

» Conducting an additional survey to assess the impact of the pilot program on participants.

» Using the insights from the pilot to create a mature training program available to all FDIC employees.

NSF Data Science and Data Analytics Certification Pilot

Small Agency   **Pilot**   **Online Learning**

Agency

The National Science Foundation (NSF) is a federal agency with 2,100 employees including career federal workers, scientists from research institutions, contractors, and staff from the National Science Board (NSB) and the Office of the Inspector General (OIG).

Contact

Adrienne Deitemeyer, *Program Manager*
adeiteme@nsf.gov

Dorothy Aronson, *Chief Data Officer and Federal Sponsor for the NSF Data Science and Data Analytics Certification Pilot Program*
daronson@nsf.gov

Key Stakeholders

» Chief Information Officers (CIO) Council and Office of Management and Budget (OMB) which were the decision-making executive bodies for the pilot program.

» CIO Council's Workforce Committee which provided support and input throughout the pilot program.

» NSF CIO who served as the federal sponsor for the program.

» NSF Program Manager who coordinated between the other stakeholder groups and planned for, developed, and executed the pilot.

» Commercial vendor which provided the online training platform and designed the curriculum.

Objectives

☑ Upskill the federal workforce to use digital tools to organize, analyze and leverage data to derive meaning and solve big data problems
.

📦 Program Overview

The NSF Data Science and Data Analytics Certification Pilot Program was a one-time cohort-based online training taken by ten participants from across various federal agencies, provided by a single vendor platform.

The curriculum focused on using advanced data analytics training techniques to teach learners how to apply data tools to uncover meaning and solve foundational big data issues. This was an initiative led by the CIO Council and NSF was selected to manage the program. Agency CIOs were invited to nominate one to three employees from their organization to enter in the program selection process, only non-2200 series employees were eligible. The training was self-paced, and participants were tasked to complete four to six hours a week of coursework along with their full-time job duties. Participants who successfully completed the pilot program received a nanodegree, a certificate marking their proficiency of the curriculum.

The pilot program was conducted from July 2019 to March 2020 and OMB provided the funding for the program's first year. NSF learned important lessons from the first cohort but found that the vendor-provided online platform did not align with their envisioned program outcomes. This pilot program did not progress past the first-year cohort.

> **"NSF learned important lessons from the first cohort but found that vendor-provided online platform did not align with their envisioned outcomes."**

💡 Previous Data Skills Training Approach

Online data skills training courses were available to employees and various NSF organizations offered individual data skills training programs..

👥 Key Inputs

» **Funding**: OMB provided $15,000 of funding to cover the costs for the first year which required a funds transfer from GSA to NSF, the sponsoring agency.

» **Training Curriculum**: The vendor provided an online training platform through a website that led participants through the coursework. Two training pathways were selected for the curriculum by the CIO Council, the data analyst and data scientist. The data analyst pathway was an intermediate option focusing on organizing data, identifying data findings, and gaining proficiency in Python and SQL programming. There were two participants in this course. The data scientist pathway focused on advanced skills, including using Python and SQL to analyze data from different sources and using machine learning to build predictive models. There were eight participants in this course.

» **Staffing**: One employee served as the program manager whose responsibility was to oversee communication with the participants, manage logistics, and coordinate vendor communication. The program manager held ongoing phone calls with participants

throughout the program to answer questions, garner feedback, and keep them engaged. Vendor representatives were present during the kickoff meeting and communicated with the program manager as the pilot progressed. And the NSF CIO served as the federal sponsor for the program.

» **External Support**: OMB, the CIO Council, and the CIO Council's Workforce Committee provided support and feedback throughout the planning and implementation process. The CIO Council and OMB determined the

need for the pilot and selected the vendor to provide the online tool for the training. The CIO Council's Workforce Committee guided program staff in selecting the training offerings from the vendor and provided support as the pilot progressed. The program manager conducted status meetings with the CIO Council's Workforce Committee to obtain their feedback and input at different points of the program lifecycle.

Key Implementation Activities

1. **Secure the Contract**: In April 2019, NSF worked with the vendor to initiate the contracting process. In May, the contract was finalized, and funds were successfully transferred from GSA to NSF to pay for the first year of the pilot program.

2. **Select the Candidates**: CIOs nominated one to three employees from their agencies to participate in the selection process. Around 60 candidates from 14 agencies were nominated for the pilot.

3. **Testing for Selection**: In June 2019, nominees took an assessment created by the vendor to determine if they fit the criteria for the pilot. It tested for prerequisite knowledge and understanding of data science concepts to be selected, and tested finalists into either the data analyst or data scientist pathway.

4. **Finalize the Cohort**: A cohort of ten participants and five waiting list candidates were finalized in June 2019. The waiting list candidates did not make the official participant list.

5. **Survey the Cohort**: A survey was given in the beginning of the program to understand participants' needs and what they wanted to achieve from the program. A survey was also given during and at the end of the course to gather lessons learned. Results from the first survey found that participants wanted to complete the pilot to either seek data science skills or enhance existing knowledge. After the second and final surveys, participants stated a need for a deeper dive into the artificial intelligence and machine learning coursework and advised future learners to acquire foundational knowledge in database programming to succeed in this type of program. Participants also found significant time needed to be set aside to complete the pilot program along with their full-time duties.

6. **Kick Off the Pilot**: The vendor managed the training through an online platform and assigned each participant a log-in. The course curriculum began in July 2019.

> " **Participants found significant time needed to be set aside to complete the pilot program along with their full-time duties.**"

Important Program Milestones

» The CIO Council worked with agency CIOs to complete the participant selection and nomination process in May 2019.

» A kickoff meeting was held with participants, vendor representatives and program staff to outline expectations and resources for the pilot program.

» Participants began work on their coursework in July 2019, the curriculum was completely virtual and self-paced.

» Program staff had initially planned for all participants to graduate from the program in February 2020, but this was extended to March 2020 to give learners more time to complete their coursework. By March 2020, three out of ten participants completed the pilot program and received a nanodegree.

Challenges

» The process to secure the contract underwent multiple challenges including transferring funds from GSA to NSF and coordinating with a vendor that was not on a GSA Schedule, this led to challenges in establishing the contract and delayed the implementation timeline.

» The pilot program was designed completely by the vendor and not curated to the federal workforce; participants used general data sets in their coursework instead of agency specific information. As a result, participants found it more difficult to apply their learnings directly to real-world scenarios they face in their jobs.

» Coursework was mostly virtual and self-paced; participants lacked a live teaching component where they could seek answers from and engage with a live instructor or subject matter expert (SME).

» Vendor communications lacked consistency and support. The training vendor was not proactive in communicating status updates to the program manager.

» Some participants began to slow in engagement by no longer logging into the vendor platform consistently to complete their coursework. The program manager noticed this around November to December 2019.

Lessons Learned

» Set clear expectations for time, commitment, goals, and outcomes of the training program in the beginning to allow participants to understand upfront the level of effort required from the program.

» Tailor the coursework to align with the federal workforce, using agency specific data sets to

allow participants to visualize how to utilize the learnings within their agency.

» Avoid interagency funds transfers for small investments to ensure the program stays on track and is not faced with bureaucratic delays.

- » Monitor participants virtual training activities to keep them engaged and on track by scheduling check-ins with specific learners who are inconsistent and disengaged.

- » Allow more flexibility for participants either through a shorter program over time or less coursework hours per week to ensure adequate time for completion of the pilot program while working their full-time jobs.

- » Gain buy-in from the participants' supervisors to ensure they are aware of the training and understand the level of effort required by employees to manage their coursework along with their full-time workload.

- » Clearly communicate the type of program support and communication needed from the vendor in the service level agreement to avoid unwanted outcomes.

- » Create more opportunities for peer engagement and networking amongst participants to allow for collaboration as they progress through their coursework.

- » Include a capstone project as a component of the training so all participants can collaborate and solve problems that deal with real-world scenarios in the federal workforce.

> **Create more opportunities for peer engagement and networking amongst participants."**

♜ Best Practices

- » Conduct participant surveys before, during and at the end of the program to capture learning objectives, responses to the coursework and concluding suggestions for future cohorts.

- » Ensure the curriculum and course materials meet the needs of learners and can be applied to their job duties.

- » Foster consistent engagement with executive stakeholder groups to ensure leadership approval is achieved through the multiple phases of the implementation process. This will also build trust for future initiatives requiring executive buy-in.

🗎 Future Strategy

Share learnings from the pilot program with the CIO Council and OMB and apply them to other federal data science training pilots being developed.

Treasury/Internal Revenue Service Data Literacy Program

Large Agency   **Full Program**   **Online Learning**

Agency

The Department of the Treasury (USDT) is a federal agency with more than 100,000 employees; the Internal Revenue Service (IRS) is the largest bureau within USDT, with around 80,000 employees.

Contact

Carmen Andujar, *Director, Human Capital Policy*

lisa.frid@census.gov | (301) 763-5543

Key Stakeholders

» Office of the Chief Human Capital Officer (CHCO) which manages and secures funding for the program.

» Office of the Chief Information Officer (CIO) which manages implementation and integration of the training within the agency's learning management system.

» Leadership, including executive sponsors, which conceptualized and drove implementation of the program.

» Learning Leaders Council, mostly comprised of training officers or those in charge of learning programs within various bureaus.

» Educational technology vendor which provides the learning platform and training content.

Objectives

☑ Design, procure, pilot, and implement an online, module-based data literacy training program. The primary goal of the program is to provide every employee the ability to effectively utilize data by increasing the capacity to digest, interpret, and extract relevant information from data, as well as to increase data literacy across the agency.

🧰 Program Overview

The USDT/IRS Data Literacy Program is an online, module-based program available to all USDT employees. The program is available through the USDT Integrated Talent Management (ITM) system, which is the agency's learning and performance management system, and provides employees with a variety of other learning resources.

The program provides beginner-level training in data literacy through modules provided by a vendor. They include around 25 courses across five modules: Data Visualization, Big Data, Data-Driven Decision Making, Data and Analytics Literacy, and Data and Analytics at Work. The IRS first utilized the training from the vendor in mid-late 2019, which was then replicated across USDT in 2020 after the nine-month pilot.

💡 Previous Data Skills Training Approach

This program was USDT's first attempt at a comprehensive, online module-based data literacy training program available for the entire agency.

Additionally, USDT continues to offer a variety of online data science courses through vendors on an as-needed basis.

🌐 Key Inputs

» **Funding**: Funding decisions were made at the CHCO-level to implement this program as an agency-wide training opportunity. The IRS utilized its own budget for the pilot with the vendor but both USDT and IRS partnered to contribute ongoing funding for the program.

» **Curriculum**: The training modules include Data Visualization, Big Data, Data-Driven Decision Making, Data and Analytics Literacy, and Data and Analytics at Work. The vendor curated previously developed course offerings to fit under each module.

» **Learning Platform**: The USDT ITM includes a large collection of online resources available to all USDT employees and some contractors on a restricted basis.

» **Staffing**: The program is housed in the Office of the CHCO and is supervised by the Director of Human Resources Policy, the Deputy CHCO, and the Learning Leaders Council. The vendor provides three contractors to administer the program, including an account executive, full-time program manager, and a customer success manager.

Key Implementation Activities

1. **Recognition of Agency Training Needs:** Although Treasury had not yet done a comprehensive training needs assessment for the entire agency, it was able to benefit from efforts of one of its subcomponents and recognized an opportunity to cost effectively provide training modules across the agency.

2. **Identify a Training Provider**: The IRS first purchased training modules from the vendor in early 2019 through the vendor's learning platform. The IRS Virtual Learning Management Group is responsible for any type of training within IRS and interfaces with the procurement process. The trainings were well-received within IRS and the vendor had the capability to integrate the training modules within the USDT ITM system.

3. **Establish a Partnership**: The IRS and USDT partnered to run a pilot of the integrated training program primarily through collaboration between the IRS Human Capital Officer and the USDT CHCO. Both agencies committed to funding and supporting the pilot.

4. **Pilot the Program**: The pilot program began in October 2019 and completed in June 2020. The pilot included approximately 500 individuals from various bureaus within USDT. Based on a user survey, more than 90% of the feedback from the pilot was positive and internal evaluations led to the decision to scale-up the program.

5. **Integrate the Program**: After deciding to expand the program to make it available to all USDT employees across the Department, the Office of the CIO and the vendor integrated the training modules within the USDT ITM system. The program became a part of the ITM system in April 2020.

> **More than 90% of the feedback from the pilot was positive and led to the decision to scale-up the program."**

Important Program Milestones

» The vendor previously partnered with the business school of a university to support the development of the training content around data literacy.

» Focus groups, a detailed survey, and an analysis of lessons learned were used to evaluate the pilot and provided insights on whether to invest in scaling-up the program.

» Ongoing performance metrics are being collected to continually assess and manage the program, with a program performance assessment expected in FY21.

Challenges

» There were technical challenges associated with both the IRS and USDT programs. For IRS, the IT and cybersecurity teams worked to overcome challenges in accessing the program through the vendor's platform which sometimes proved difficult due to the IRS firewall. For USDT, there were technical issues associated with integrating the program into the USDT ITM system which took additional time to resolve.

» An ongoing challenge is building awareness and participation in the program, as while all employees can complete the training, not all employees have done so. The training modules were rolled out just before Treasury's attention shifted to responding to the COVID pandemic. Performance metrics will provide insight on the level of participation in the program. Additionally, initial insights suggest employees who have prior data skills may think the program is too basic, while employees who have no prior data skills may not understand the benefits of the training.

Lessons Learned

» An on-line, broadly disseminated, self-paced learning program can be a component of an agency's initial methods of enhancing data skills.

» Acquiring enterprise-wide solutions can be a cost-effective alternative to bureaus or organizations trying to build custom content or micro-managing the curriculum.

» Understand what data skills training activities are taking place at the different bureaus within an agency to get ideas for other areas to scale-up across the agency. Additionally, understand what adjustments should be made to the scaled-up program through a pilot and evaluation process.

» Have a clear understanding of the purpose of the pilot and program goals, and a contingency plan if those goals are not met.

Be prepared for communications to require a significant amount of time."

» Develop a communication strategy to market the program to build awareness and participation but be prepared for communications to require a significant amount of time to navigate the appropriate approval channels.

» Collaborate with relevant stakeholders, such as executive-level and internal council leadership within the agency to enhance awareness and buy-in for the program.

Best Practices

» Design and implement an online, module-based data literacy training program within a bureau, and successfully pilot and implement the program across the entire agency.

» Partner with a vendor to curate an online curriculum tailored for the agency and integrate it within the agency's learning and/or performance management system

to provide the training alongside other important resources for employees.

» Administer a pilot program first and then conduct an evaluation of the pilot to enable replication as a full-scale program by adapting and refining the program.

» Attain departmental and operational leadership buy-in and engage the relevant stakeholders throughout the program's design and implementation to create and sustain the program.

» Utilize new insights from performance metrics and survey feedback to continually enhance the program and curriculum.

Future Strategy

» Conduct a program performance assessment in FY21 to gather data on the level of participation in the program, the level of effectiveness with expected outcomes, and develop strategies to increase awareness and participation. Marketing strategies including emails, bulletins, conversations with managers, presentations, and partnerships with various business units across the agency will continue to enhance employee awareness of the program. Use new Office of the CDO to amplify messaging.

» Continue to gather data on the program from feedback, surveys, and performance metrics to continually adjust the program and curriculum. Particularly important is understanding employee satisfaction with the training offerings and what other topics they would like to see offered.

» In the broader ITM system, there are career development pathways which are called Aspire Journeys and include pathways such as a data analyst to data scientist or business analyst to data analyst. These are comprehensive and are linked to learning resources within USDT. In the future, USDT plans to link data literacy to these pathways through this program.

> " Particularly important is understanding employee satisfaction with the training offerings."

If you have questions or would like more information about the case studies, contact cdocstaff@gsa.gov.

federal CDO council