RESEARCH-ARTICLE

A Window into DataWorks: Developing an Integrated Work-Training Curriculum for Novice Adults

LARA KARKI, Georgia Institute of Technology, Atlanta, GA, United States

DANA PRIEST, Georgia Institute of Technology, Atlanta, GA, United States

JAMES G DUBOSE, Emory University, Atlanta, GA, United States

ZAJERRIA GODFREY

ANNABEL ROTHSCHILD, Georgia Institute of Technology, Atlanta, GA, United States

BEN RYDAL SHAPIRO, Georgia State University, Atlanta, GA, United States

View all

Published: 12 February 2025

Citation in BibTeX format

SIGCSE TS 2025: The 56th ACM Technical Symposium on Computer Science Education
February 26 - March 1, 2025
PA, Pittsburgh, USA

Conference Sponsors:
SIGCSE

A Window into DataWorks: Developing an Integrated Work-Training Curriculum for Novice Adults

Lara Karki
Georgia Institute of Technology
Atlanta, GA, USA
lschenck3@gatech.edu

Dana Priest
DataWorks, Georgia Institute of Technology
Atlanta, GA, USA
dpriest3@gatech.edu

James G. DuBose
Emory University
Atlanta, GA, USA
james.g.dubose@gmail.com

Zajerria Godfrey
Atlanta Public Schools
Atlanta, GA, USA
zajerria.godfrey@atlanta.k12.ga.us

Annabel Rothschild
Georgia Institute of Technology
Atlanta, GA, USA
arothschild@gatech.edu

Ben Rydal Shapiro
Georgia State University
Atlanta, GA, USA
bshapiro@gsu.edu

Betsy DiSalvo
Georgia Institute of Technology
Atlanta, GA, USA
bdisalvo@cc.gatech.edu

Abstract

Computing education is often confined to the context of formal education or after-school programs; however, there is a growing industry built around adult education, including workshops, coding intensives, online learning, and apprenticeship programs. Amidst these efforts, little research has explored the workplace as a site for novice adult learners to develop computing skills. In this experience report, we present an integrated training curriculum for adults at DataWorks, an organization that trains and employs novice adults from groups historically underrepresented in computing who seek to advance their career through on-the-job learning. "Data Fellows" are hired to complete client projects by providing data services for local organizations, nonprofits, and businesses. Training is integrated into employees' weekly responsibilities at DataWorks, and the curriculum consists of four modules: Microsoft Excel, Critical Data Literacy, Python Fundamentals, and Career Development. In this report, we reflect holistically on the evolution of the curriculum over four years. We distill our reflection into insights to inform other integrated training programs that aim to equip novice adults with computing skills in the workplace.

CCS Concepts

• **Social and professional topics** → **Adult education**; **Computing occupations**; **Informal education**; **Model curricula**.

Keywords

Training Curriculum, Novice Adults, Workplace, Data Work, Communities of Practice

SIGCSE TS 2025, February 26–March 1, 2025, Pittsburgh, PA, USA.

ACM ISBN 979-8-4007-0531-1/25/2
https://doi.org/10.1145/3641554.3701826

ACM Reference Format:
Lara Karki, Dana Priest, James G. DuBose, Zajerria Godfrey, Annabel Rothschild, Ben Rydal Shapiro, and Betsy DiSalvo. 2025. A Window into DataWorks: Developing an Integrated Work-Training Curriculum for Novice Adults. In *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE TS 2025), February 26–March 1, 2025, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3641554.3701826

1 Introduction

There is a growing interest in computing education for adults and growing recognition of out-of-school learning pathways in computing [5, 40]. Graphic design work, for example, can be a context in which to learn many foundational computer science (CS) concepts informally through building and modifying graphics software plugins [13, 14]. Scholars similarly find that members of online fandom communities develop computing skills through the collective maintenance of open-source platforms to host fandom stories [15, 16]. Another study demonstrates how administrative professionals using Salesforce CRM learn software development skills and pursue software development careers through professional meetup groups [28]. These studies, the ubiquity of computing work occurring across domains, and the rapidly evolving industry of adult computing education highlight the importance of studying computing education for adults outside of formal CS spaces [2, 15].

The increased demand for adults with data skills[1] has warranted a growth in popularity of data-related coding boot camps, online courses, and apprenticeship programs. However, few adult data skill training programs are accessible to novice working adults. In response, we created DataWorks to explore the potential of integrated workplace training in data skills as an approach to adult computing education and broadening participation in computing. DataWorks is a data service provider housed at Georgia Tech that recruits adults with little to no technical background, employs them full-time for one year, and trains them in data work while they complete real-world projects on datasets submitted by local companies,

[1]Since there are many similarities between data analytics and data science [1], we use the term "data skills" to capture the foundational knowledge of how data drives technology and how to work with and understand clean data in a workplace context.

nonprofits, researchers, and other community organizations. Employees, known as "Data Fellows" complete in situ learning through activities and experiences that allow them to assume central roles through DataWorks.

In this experience report, we present our curriculum and reflect on its evolution over the course of four years with four distinct cohorts and 15 employees. We highlight what worked, what didn't work, and what changes we made to the curriculum and the structure of DataWorks during this period and why. We synthesize our reflection into three overarching insights: (1) Adults' Motivation is Nuanced and Variable, (2) Adaptability is Key to Stability, and (3) Workplace Needs Challenge Traditional CS. These insights can guide other researchers, practitioners, and organizations who aim to develop curricula and computing education experiences for novice adult learners in the workplace. Finally, we make our curriculum design freely available online [7].

2 Related Work

Resources for adult computing education are largely inaccessible to novice adult learners who did not go to college, who need to work to meet living expenses, or who have childcare commitments. Online courses are focused on individual learning and rarely build upon learners' personal interests and cultural backgrounds [23]. They require large time commitments with little learning motivation, lack opportunities for collaborative learning, and do not provide opportunities to exercise new knowledge in an authentic context. In-person programs such as coding boot camps or apprenticeships may offer such opportunities, but they often require learners to demonstrate or obtain significant computing knowledge before they begin the program (e.g., [3, 4, 6]) and tend to benefit learners who already have a college education [29]. Further, the intensity of such initiatives may disrupt aspects of adults' lives [36].

The workplace is one site for learning that is accessible to novice adult learners, but little computing education research or programs focus on novice adults in the workplace. Most studies that consider adults' CS learning at work focus on how experienced engineers continue learning on-the-job. These studies aim to understand just-in-time, incidental, and self-directed learning rather than how to integrate curricula into work practices [25, 27, 30]. There is a need for computing education research to incorporate insights from adult and workplace learning more broadly to understand how learning can be intentionally integrated into—and mediated by—everyday work practices [9, 21, 24].

Data skills are increasingly relevant to work requirements across domains. Accordingly, data-centric curricula are growing in popularity in both formal education and informal settings. Typical CS curricula tends to focus on programming and theoretical CS concepts; in contrast, data-centric curricula closely links computing concepts to their application through data projects [22, 39]. Further, learning sciences and computing education research find data-centric curricula are an effective way to teach broader computational thinking skills and to engage diverse groups of learners [10, 18, 19, 26, 31, 38, 41]. However, few curricular initiatives are focused on adult learners and even fewer consider how to integrate data skills training into the workplace.

Our work at DataWorks is motivated by bringing together the accessibility of the workplace as a site for adult learning with the benefits of data-centric curricula. DataWorks explores how a curriculum in data skills can be integrated into data work (e.g., data cleaning, entry, and annotation) to scaffold apprenticeship learning through legitimate peripheral particiation (LPP) in a workplace community of practice. LPP provides an analytical lens to understand apprenticeship learning in a variety of contexts, describing how newcomers in a community of practice learn through participation in legitimate, low-stakes activities that are authentic to the goals and values of a community of practice [24]. Prior work establishes that adults learn computing skills through LPP [11, 16]. For example, adults learn data skills through LPP in domain-specific communities such as administrative work, local government, and non-profits [33]. Outside of these few studies, however, little computing research considers apprenticeship learning through an LPP lens. Instead, a few studies consider apprenticeship and work-based learning as possible alternatives to degree-level education [32, 35], and others are concerned with how formal apprenticeship programs can provide CS students the professional skills employers want but CS degrees are not delivering [8, 17].

3 DataWorks' Background

Founded in January 2020, DataWorks is a data services provider that operates within Georgia Tech, completing data projects for external and internal clients. With funding from the university, DataWorks recruits employees with little to no technical background from communities historically excluded from computing for full-time, paid positions at Georgia Tech. DataWorks is a model for members of historically excluded communities to learn entry-level data skills through full-time paid work and to sustain a community of practice and research context generating novel approaches to data science work and education. Our curriculum can be adapted and used in other kinds of workplaces to support novice adult learners from diverse backgrounds to engage in and learn about data work [7].

In the first year, DataWorks had four employees working part-time and primarily did pro bono work for non-profit organizations. During this time, we sought to understand the different types of data work that we could do for clients and the training that would be needed to prepare the Data Fellows. Initially, we provided training for Data Fellows as-needed based on client projects. We found there was a need for critical data literacy training to help Data Fellows understand the implications of data and the data life-cycle. Following this, we continued to do ad-hoc training on different systems we were asked to work in such as Microsoft Excel, various CRMs, and homegrown systems.

During year two, we recognized that in-depth training on Excel would provide us with skills that were transferable to most projects. Similarly, as we were asked to help with more data collection, we continually relied on graduate students to help write Python scripts for web scraping. This led us to develop a Python programming curriculum, which taught practical scripting techniques often used in data collection and cleaning. Finally, during our second year, when our first cohort of full time workers were ready to enter the workforce, we recognized the need for career development. The Data Fellows requested training to help them in constructing

their resumes, finding appropriate positions, and developing relevant professional relationships. It was clear that our curriculum must explicitly prepare the Data Fellows to find their next position, namely by addressing the invisible, culturally-specific skills and connections required to procure a job. The first author, a doctoral student returning to school after a career as an informally trained software engineer, developed and continues to facilitate the career development program at DataWorks.

The research team has a close working relationship with the DataWorks employees and meets with them several times per week for topics related to both research and DataWorks' operations. Today, DataWorks operates with seven employees: five Data Fellows, one Training Coordinator, and one Project Manager. Both the Training Coordinator and Project Manager are former Data Fellows, and they are responsible for the day-to-day operations of DataWorks. The Training Coordinator is the second author of this paper. The current DataWorks employees consist of five women and two men. Across all cohorts, Data Fellows are Black Americans and/or people of color, range in age from early 20s to 50s, come from diverse career backgrounds such as construction, makeup artistry, sales, and the service industry, and most pursued educational pathways outside of a traditional bachelor's degree.

Data Fellows are hired for one year, and their time is divided between working on client projects and learning foundational data skills, including data cleaning, formatting, and labeling. They learn through formal training and exposure to a variety of datasets for client projects. After six to eight months of training, the Data Fellows shift to self-directed learning and job search as they find their next role. A minimum of 10 hours/week is dedicated to training-related work at DataWorks. The team meets roughly two to three times per week for training-related sessions. During a typical week, one session may include formal instruction or working through video content, while another will be a remote "Skillshare" meeting where Data Fellows present a project or problem they have been working on during the week related to their training. Many weeks also have dedicated career development sessions, for example, a guest speaker "Career Chat" or a facilitator-supported working session on topics like professional networking or resume writing.

4 Training Curriculum

The four-part training curriculum at DataWorks aims to provide novice data workers with foundational data skills and the professional skills to pursue a career in data. In this section, we provide an overview of the four curriculum modules.

4.1 Microsoft Excel

DataWorks' Excel module prepares learners to use Microsoft Excel to clean, organize, and store data, and to perform basic data analysis. Core elements of the curriculum include Excel's functions and formulas to sort and filter data, summarize data with pivot tables, and visualize data using charts. The training is video-based and initially leveraged video courses from LinkedIn Learning [2]. Recently, the curriculum was refined to cover the same topics using freely available online resources from Data Quest [3]. A key feature of the Excel curriculum is that learners choose a small project and demonstrate a concept from the training during the weekly Skillshare meetings.

Module	Weeks	Objective
Excel	6-8	Clean, process, and analyze data using Microsoft Excel.
Critical Data Literacy	6-8	Describe which factors to consider when leveraging data to support decision-making.
Python Fundamentals	8-10	Organize and manipulate data with Python.
Career Development	Ongoing	Find and apply to jobs aligned with career interests.

Table 1: DataWorks' Training Curriculum and Objectives.

4.2 Critical Data Literacy

The Critical Data Literacy module, developed by members of the research team, is a lecture and discussion based course designed to teach the foundational concepts of data, how it is used to inform decision making in various contexts, the lifecycle of data and documentation, and how to utilize data visualization tools to tell a story with data. The course contains four guided modules:

(1) **How to Navigate a World of Big Data**. Introduces the big picture concepts of data, e.g., data types and how data is used to make evidence-based decisions in variable contexts.
(2) **Data Organized and Stored**. Introduces the idea of telling a story with data, beginning with how data is stored and organized utilizing tables and databases. Investigates how data can be interpreted to further its meaning and impact.
(3) **Documentation and Data Lifecycles**. Addresses the lifecycle of data and each of its stages, emphasizing the importance of data documentation.
(4) **Data for Identifying Patterns to Create Models**. Explains how patterns in data are recognized using statistical techniques to make informed predictions. Learners investigate real-world applications for utilizing models to inform decision-making in a variety of contexts.

Discussions and examples center the importance of clean, accurate, ethical data and the potential ramifications bad data can have on society. Incorporating critical perspectives into data literacy training equips the Data Fellows to voice concerns when data, its projected use, or work processes raise ethical issues [12].

4.3 Python Fundamentals

DataWorks' Python module uses data manipulation as a central theme to teach fundamental Python skills. The course progresses through various topics, including string manipulation, list handling, file reading, conditionals, loops, functions, and external libraries using data that is personally relevant to workers [38]. It concludes with a practical hands-on project. The course content is delivered through verbal lectures and discussions, using Google Colaboratory for practice problems to reinforce learning. During the Python curriculum, Skillshares are devoted to letting Data Fellows' present

[2] https://www.linkedin.com/learning/
[3] https://www.dataquest.io/

solutions to practice problems and clarifying concepts with the Training Coordinator. This course was developed by members of the research team, and refined in collaboration with the Training Coordinator, who authored practice problems, contributed to the revised version of the course, and now instructs the course fully.

4.4 Career Development

The career development module aims to prepare the Data Fellows to set and pursue short-term and long-term career goals, develop professional relationships, perform job search-related activities, and develop required materials for job search. The curriculum—developed by the research team and shaped by the Data Fellows—consists of lectures, discussion-based workshops, scaffolded exercises, and facilitator-supported work time. External volunteers give one-hour informal Zoom "Career Chat"s where they share career stories and answer questions from the Data Fellows. Career development occurs in phases throughout the Data Fellows' one-year term and consists of weekly meetings, 1-1 coaching as needed, and work time to prepare job search materials.

4.5 Weekly Skillshare

Skillshare is a weekly one-hour meeting to reinforce concepts from the Data Fellows' training. During the preceding week, Data Fellows choose a topic from their training and design their own 5-7 minute demonstration that identifies how the topic is relevant to their work. In the meeting, everyone presents their demonstration.

5 Reflection

This report contributes a holistic reflection on the evolution of DataWorks and its training curriculum. Given the complexity, situatedness, and unpredictability of work and workplace learning more generally, we do not attempt to evaluate the curriculum using quantitative measures. Instead, we leverage our intimate knowledge of formative data collected through a variety of avenues, including ethnographic field notes, meeting recordings, workplace artifacts, and interviews with employees at their start and during their 8^{th} month in the program. Collectively, this lends us thousands of hours of ethnographic participant observation and experience in key roles of the DataWorks organization.

5.1 Consistency Matters

In the early days of DataWorks, we didn't have such formalized training. Instead, we supported the Data Fellows' learning on an as-needed basis for whatever project they worked on. Notably, the first cohort of workers were less interested in formal training; they wanted to learn on an as-needed basis while working on a project. However, we quickly realized workers needed to understand the purpose of the data to contribute to project work, and our initial ad-hoc approach was insufficient. Later, we emphasized the role of training when interviewing Data Fellow candidates. It took us a year to identify which fundamentals we need to teach and to hire a cohort that was interested in more structured training.

As DataWorks is housed in a university, a unique aspect of the Data Fellow's work environment is their access to graduate students. When DataWorks first began, those students were in charge of developing and teaching the curriculum. However, busy schedules, combined with students graduating and moving on, led to erratic lesson schedules and subject matter that underwent drastic changes with each subsequent student. This scattered learning schedule made it difficult to install a permanent training curriculum. Over time, we realized the need to commit to more consistency. One our first steps was to promote a Data Fellow to Training Coordinator and give them responsibility to facilitate the curriculum. Having a full-time member of the team take the lead helped stabilize the schedule and enabled DataWorks to be more intentional about what skills the Data Fellows learn. Our persistence and intentional effort to focus on training allowed us to gradually build what has become our current program. As the team began to learn these new skills, we realized the valuable interplay between work on our client projects and training.

5.2 Symbiotic Nature of Work and Training

By working on client projects while completing training, Data Fellows have an authentic context to apply what they learn. The opposite is also true, where project work shapes what content we cover in training, and experiences from work projects motivate the Data Fellows' learning.

One example is a project for *GET Cities* where the Data Fellows applied their Excel and Python skills and in doing so, improved their abilities and increased their interest to incorporate automation in client projects. Due to DataWorks' work with civic and non-profit organizations, most of the previous datasets the team worked on were small; automation was never necessary. This project required workers to manipulate over 705,000 Excel cells, more than three times the data volume the team was used to handling. Realizing that manually working on the project file would take too much valuable time and result in missing the client's deadline, the team brainstormed how to automate the process and work more efficiently. They drew from their Excel training to develop various formulas to convert thousands of cells of data in minutes. Yet, they realized the data had so many varying factors, the formulas were not precise enough to catch all of them. Going back to the drawing board, the team realized they could write a Python script that would recognize the variables Excel could not.

How does this example showcase the symbiosis of work and training? Our work knowledge and efficiency improved from the training and our training was reinforced as we applied it to a real project scenario. While both elements of our program work independently, the two sides feed into each other and produce better results when done in tandem. This project taught the team a valuable lesson in persistence and problem-solving. When one course of action failed, the team iteratively reassessed and challenged their capabilities until they achieved desirable results. Moving forward, this project also gave insight into future training curricula DataWorks can incorporate to further support client work.

5.3 Re-Usable Curriculum

Discovering how effective training and working in tandem can be, we aimed to keep improving and adjusting the curriculum, not only for the internal team, but for the public and other work contexts. As previously mentioned, consistency matters, so streamlining the material and making lessons reusable was the next priority in our

training curriculum building journey. Over the summers of 2023 and 2024, DataWorks hired a teacher to help us. The teacher shared her experiences and the most effective methods for creating learning materials for adults. Following collaborative brainstorming meetings with the entire team, DataWorks and the teacher were able to identify changes to make, decide how to restructure the lesson plans (or modules), and create rubrics to measure the Data Fellows' (or any potential student's) knowledge and success. The lesson plans were put into presentations available for public use in group settings or individually at a self-pace. We also developed learning objectives and measures to assess learning, which we continue to advance to evaluate the effectiveness of the curriculum.

5.4 Broadening Targeted Career Pathways

The career development module of DataWorks' curriculum was initially developed at the request of an early cohort of Data Fellows. When it came time for them to find a job following their term, they requested training and support. The career development module evolved from a series of six semi-formal workshops in professional networking and job search skills to be ongoing throughout the one-year term, including a weekly meeting, guest speakers, and 1-1 coaching with the facilitator. While the initial modules' content remains similar, weekly meetings respond to the Data Fellows' needs during their job search, and the "Career Chats" allow them to develop a professional network through DataWorks.

It was not until the end of the earliest cohort's term that we considered to *what* jobs the Data Fellows would apply. Following training at DataWorks, the Data Fellows are likely overqualified for roles such as data entry clerk, and many were not interested in clerical roles. During the first and second cohorts of career development, the Data Fellows were encouraged to find and apply for entry-level analyst positions such as data analyst or business analyst. Even though the curriculum does not address statistical analysis or visualization in depth, Data Fellows leave with an intimate understanding of data and getting to know a dataset. While we maintain the Data Fellows may be excellent candidates for entry-level analyst roles, we have seen that these roles often require bachelor's or master's degrees or years of related work experience.

Despite several attempts applying to analyst jobs and attempts to build a local data-focused network through "Career Chats", Georgia Tech connections, and attending data-related events, none of the Data Fellows from early cohorts procured a data-specific role following the program. Further, we have found that entry-level candidates seeking a data-focused job must either have a personal referral in their network or complete significant work to craft supplementary job search materials beyond a resume to appeal to employers. These efforts include personal branding and marketing, networking at meetups (which typically occur in evenings), building a portfolio website, and additional study of data analysis skills outside of what the DataWorks training offers. We find that this workload is prohibitive for Data Fellows, many of whom have family commitments, as it requires significant time outside of work hours.

Given the need for Data Fellows to find positions following their one-year term, we pivoted our focus. We broadened our ideas about what careers to consider, and we were more explicit about what additional efforts would be required for different career pathways.

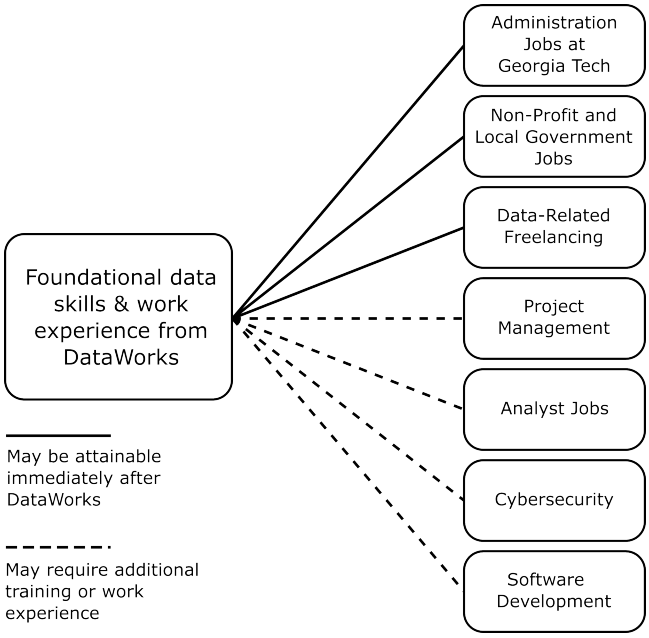

Figure 1: Career options after DataWorks, based on our observations of 15 Data Fellows participating in the program over the course of four years.

Instead of presenting DataWorks as one step in a linear pathway toward a data-specific career, we now present the skills obtained at DataWorks as a starting point for many potential career pathways (Figure 1). We are beginning to broaden this further by introducing freelance work platforms and other entrepreneurial options [20].

5.5 Challenges of Novelty in the Real World

DataWorks started as a research project that aimed to establish a different approach for local data work and developing a local data workforce. We consistently run into the issue of being a unique training program. The data work performed at DataWorks is often outsourced to platforms such as Mechanical Turk, or these tasks are completed by someone in addition to another role. Creating a full-time position in data work that focuses on the data itself rather than gleaning insights from data is novel. Further, performing data work in a team setting where data workers have direct contact with clients is also novel [12]. As an organization, DataWorks brings together the isolated steps of data preparation with model work and analysis, making data work a form of legitimate peripheral participation in data analysis and data science.

While DataWorks' novelty is important and exciting for a research project, when it comes to growing the program or to the Data Fellows marketing their skills, this novelty creates practical challenges. For example, potential employers, clients, and partners misinterpret DataWorks due to its connection with an elite technical university. Also, when communicating the program structure to Data Fellow candidates in interviews, we need to explicitly highlight the nature of the position in contrast to training intensives that are completed *before* getting a job. Now, we are clear about

the structure of the program: it is first and foremost a full-time job with a 1-year term, and we do not place candidates in their next job. We need to directly address what DataWorks is and isn't in all contexts, and we are often limited by language and broader societal data literacy when it comes to clearly communicating what we do.

6 Takeaways and Insights

After experiencing growing pains, we gradually learned what worked and didn't work. Our main takeaways are that consistency, continually improving our curriculum, broadening our ideas about career paths, and facing challenges of novelty all play a role in how the program is shaped today. We found workers benefit from concurrent on-the-job training, via real client projects, and semi-formal training in Excel, Python, critical data literacy, and career development. New skills learned on a weekly basis are transferable to client work, where they are reinforced and applied to projects, making our work more efficient. DataWorks is an alternative route for adults who are less frequently afforded opportunities to enter the tech world and explore different career paths.

In this section, we build on our reflective themes to synthesize key insights about workplace learning programs to guide other practitioners and researchers. Additionally, our curriculum design is freely available online [7]. It can be adapted and used in other kinds of workplaces to support novice adult learners from diverse backgrounds to engage in and learn about data work.

6.1 Adults' Motivation is Nuanced and Variable

DataWorks is a training program that pays Data Fellows to work and learn on the job as opposed to a paid training program or a formal degree. Data Fellows are typically *not* motivated by achieving an elite or cutting-edge tech job or obtaining a formal degree, which is often the motivation for adults in programs like coding boot camps or degree programs. The adults at DataWorks, a diverse group with a variety of career and educational backgrounds, aged from early 20s to 50s, have a variety of motivations. These include financial stability, career transition, or finding alternatives to college education. We find that the Data Fellows are motivated by the opportunity to learn while being paid at a full-time job. They take the program seriously and demonstrate a strong work ethic while maintaining a work-life balance that doesn't impede on family and personal commitments.

6.2 Adaptability is Key to Stability

Being flexible and willing to adapt our process on the fly indirectly led us to stabilizing our program. For example, a research scientist was initially hired to do project management. When she suddenly left the position, we adapted by promoting a Data Fellow into the project management role. This set the tone for then promoting another Data Fellow when that project manager moved into a different job at Georgia Tech, and for us to create the position of Training Coordinator, promoting the second author from her former role as Data Fellow. Today, these two roles are critical to the stability of day-to-day operations at DataWorks, and have the close visibility that allows them to adapt on a micro-scale to maintain a balance between project work and training, which is constantly in flux based on our client project workload. Similarly, our willingness to adapt to the difficulties Data Fellows faced with job search was the catalyst for adding career development to the curriculum. Finally, broadening the roles we considered in job search led us to a future direction for DataWorks outside of computing-specific contexts.

6.3 Workplace Needs Challenge Traditional CS

The constraints of the workplace led us to challenge and re-shape traditional ideas about what is taught in CS and outcomes of CS learning [37]. Our program establishes the need for further research into how to support novice adults learning computing in the workplace [34]. Formal education and resources like coding bootcamps and online courses are inaccessible and uninteresting to many novice adults. At DataWorks, our focus on the needs and strengths of our unique workplace results in a stronger and more accessible program than if we had conformed to traditional ideas of CS curricula and career pathways. In doing so, we developed a novel, stable program and curriculum that allows novice adults to learn data skills while they are paid to work a full-time job.

In pursuit of long-term stability, DataWorks will join forces with the Office of Sponsored Programs at Georgia Tech to adjoin grants administration training to the DataWorks program. Grants administration may not align with traditional ideas about computing jobs, e.g., software engineer or data scientist. However, this partnership will both ensure long-term stability for DataWorks and it will be an opportunity for Data Fellows to combine data skills learned at DataWorks with domain-specific skills leading to gainful employment at universities all over the country, if not world.

6.4 Limitations and Conclusion

While we are thrilled by the evolution of DataWorks, our program has several limitations. The size of our team (seven) limits the scale of work we can accommodate. Smaller organizations with less data and less capacity to leverage their data benefit from our model. DataWorks' curriculum focuses on a broad set of skills relevant to this context. While we believe our curriculum is a foundation for technical work within and outside of computing-specific jobs, these may not be the most valuable skills for other organizations. Moreover, attempting to adopt aspects of DataWorks into a larger organization will require institutional and administrative labor (e.g., recruitment and funding) beyond the scope of this paper. Access to professors and graduate students who are not beholden to profit-driven outcomes is also unique to DataWorks' context in a research institution. Finally, DataWorks is a relatively new program, and our ability to define and evaluate success *while* developing the program has been limited. As DataWorks matures and more Data Fellows find their next roles, we hope to improve our evaluation metrics.

Acknowledgments

This report is supported by the National Science Foundation under Grant No. 1951818, the Constellations Center for Equity and Computing at Georgia Institute of Technology, and the Public Interest Technology University Network Challenge Fund, a fiscally sponsored project of New Venture Fund. We thank Michael Johnson, Idel Martínez, Carl DiSalvo, and Kayla Evans for their support in preparing this experience report.

References

[1] Cheryl L Aasheim, Susan Williams, Paige Rutner, and Adrian Gardiner. 2015. Data Analytics vs. Data Science: A Study of Similarities and Differences in Undergraduate Programs Based on Course Descriptions. 26 (2015).

[2] Gregory D. Abowd. 2012. What next, Ubicomp?: Celebrating an Intellectual Disappearing Act. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*. ACM Press, Pittsburgh, Pennsylvania, 31. https://doi.org/10.1145/2370216.2370222

[3] Ada Developers Academy. 2023. Eligibility Requirements. https://web.archive.org/web/20230930061838/https://adadevelopersacademy.org/eligibility-requirements. Accessed: 2024-07-15.

[4] Fullstack Academy. 2024. Bootcamp Prep Course. https://web.archive.org/web/20240225151813/https://www.fullstackacademy.com/bootcamp-prep. Accessed: 2024-07-15.

[5] Friday Joseph Agbo. 2024. Broadening Participation in Adult Education: A Literature Review of Computer Science Education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. ACM, Portland OR USA, 11–17. https://doi.org/10.1145/3626252.3630797

[6] New Apprenticeship. 2024. Data Analytics Apprenticeship. https://web.archive.org/web/20240229121221/https://newapprenticeship.com/apprenticeships/data-analytics-apprenticeship. Accessed: 2024-07-15.

[7] DataWorks at Georgia Tech. 2023. Data Tools Curriculum. https://dataworkforce.gatech.edu/data-tools. Accessed: 2024-11-11.

[8] Matthew Barr and Syed Waqar Nabi. 2022. The Development of Students' Employability Skills on a Work-Based Software Engineering Degree Programme. In *2022 IEEE Frontiers in Education Conference (FIE)*. 1–9. https://doi.org/10.1109/FIE56618.2022.9962611

[9] Stephen Billett. 2014. Learning in the Circumstances of Practice. *International Journal of Lifelong Education* 33, 5 (Sept. 2014), 674–693. https://doi.org/10.1080/02601370.2014.908425

[10] Tamara L. Clegg, Keaunna Cleveland, Erianne Weight, Daniel Greene, and Niklas Elmqvist. 2023. Data Everyday as Community-Driven Science: Athletes' Critical Data Literacy Practices in Collegiate Sports Contexts. *Journal of Research in Science Teaching* 60, 8 (2023), 1786–1816. https://doi.org/10.1002/tea.21842

[11] Betsy DiSalvo, Mark Guzdial, Amy Bruckman, and Tom McKlin. 2014. Saving Face While Geeking Out: Video Game Testing as a Justification for Learning Computer Science. *Journal of the Learning Sciences* 23, 3 (July 2014), 272–315. https://doi.org/10.1080/10508406.2014.893434

[12] Carl DiSalvo, Annabel Rothschild, Lara L. Schenck, Ben Rydal Shapiro, and Betsy DiSalvo. 2024. When Workers Want to Say No: A View into Critical Consciousness and Workplace Democracy in Data Work. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (April 2024), 1–24. https://doi.org/10.1145/3637433

[13] Brian Dorn and Mark Guzdial. 2006. Graphic Designers Who Program as Informal Computer Science Learners. In *Proceedings of the 2006 International Workshop on Computing Education Research*. ACM Press, Canterbury, United Kingdom, 127. https://doi.org/10.1145/1151588.1151608

[14] Brian Dorn and Mark Guzdial. 2010. Learning on the Job: Characterizing the Programming Knowledge and Learning Strategies of Web Designers. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 703. https://doi.org/10.1145/1753326.1753430

[15] Brianna Dym, Namita Pasupuleti, Cole Rockwood, and Casey Fiesler. 2021. "You Don't Do Your Hobby as a Job": Stereotypes of Computational Labor and Their Implications for CS Education. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. ACM, Virtual Event USA, 823–829. https://doi.org/10.1145/3408877.3432396

[16] Casey Fiesler, Shannon Morrison, R. Benjamin Shapiro, and Amy S. Bruckman. 2017. Growing Their Own: Legitimate Peripheral Participation for Computational Learning in an Online Fandom Community. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, Portland Oregon USA, 1375–1386. https://doi.org/10.1145/2998181.2998210

[17] Wouter Groeneveld, Joost Vennekens, and Kris Aerts. 2022. Identifying Non-Technical Skill Gaps in Software Engineering Education: What Experts Expect But Students Don't Learn. *ACM Transactions on Computing Education* 22, 1 (March 2022), 1–21. https://doi.org/10.1145/3464431

[18] Samantha Hautea, Sayamindu Dasgupta, and Benjamin Mako Hill. 2017. Youth Perspectives on Critical Data Literacies. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 919–930. https://doi.org/10.1145/3025453.3025823

[19] Jessen Havill. 2019. Embracing the Liberal Arts in an Interdisciplinary Data Analytics Program. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*. ACM, Minneapolis MN USA, 9–14. https://doi.org/10.1145/3287324.3287436

[20] Julie Hui, Kentaro Toyama, Joyojeet Pal, and Tawanna Dillahunt. 2018. Making a Living My Way: Necessity-driven Entrepreneurship in Resource-Constrained Communities. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 1–24. https://doi.org/10.1145/3274340

[21] Anne Karhapää, Raija Hämäläinen, and Johanna Pöysä-Tarhonen. 2023. Digital Work Practices That Promote Informal Workplace Learning: Digital Ethnography in a Knowledge Work Context. *Studies in Continuing Education* (Oct. 2023), 1–18. https://doi.org/10.1080/0158037X.2023.2274596

[22] Ayaan M. Kazerouni, Jane Lehr, and Zoë Wood. 2024. Community Action Computing: A Data-centric CS0 Course. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. ACM, Portland OR USA, 646–652. https://doi.org/10.1145/3626252.3630807

[23] Ada S. Kim and Amy J. Ko. 2017. A Pedagogical Analysis of Online Coding Tutorials. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*. ACM, Seattle Washington USA, 321–326. https://doi.org/10.1145/3017680.3017728

[24] Jean Lave and Etienne Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge [England] ; New York.

[25] Hieu-Trung Le and Aditya Johri. 2021. Engineers' Situated Use of Digital Resources to Augment Their Workplace Learning Ecology. In *2021 IEEE Frontiers in Education Conference (FIE)*. 1–8. https://doi.org/10.1109/FIE49875.2021.9637421

[26] Victor R Lee. 2018. Personal analytics explorations to support youth learning. In *Digital technologies and instructional design for personalized learning*. IGI Global, 145–163.

[27] Soila Lemmetty and Kaija Collin. 2020. Self-Directed Learning as a Practice of Workplace Learning: Interpretative Repertoires of Self-Directed Learning in ICT Work. *Vocations and Learning* 13, 1 (April 2020), 47–70. https://doi.org/10.1007/s12186-019-09228-x

[28] Louise Ann Lyon and Chelsea Clayton. 2021. Arising of Informal Women's Learn-to-code Communities: Activity Systems as Incubators. *ACM Transactions on Computing Education* 21, 2 (June 2021), 1–24. https://doi.org/10.1145/3433167

[29] Louise Ann Lyon and Emily Green. 2021. Coding Boot Camps: Enabling Women to Enter Computing Professions. *ACM Transactions on Computing Education* 21, 2 (June 2021), 1–30. https://doi.org/10.1145/3440891

[30] Michael Madaio, Shivani Kapania, Rida Qadri, Ding Wang, Andrew Zaldivar, Remi Denton, and Lauren Wilcox. 2024. Learning about Responsible AI On-The-Job: Learning Pathways, Orientations, and Aspirations. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Rio de Janeiro Brazil, 1544–1558. https://doi.org/10.1145/3630106.3658988

[31] Lionel Mew. 2020. Designing and Implementing an Undergraduate Data Analytics Program for Non-Traditional Students. (2020), 10.

[32] Jack Parkinson and Sebastian Dziallas. 2024. Institutional Perspectives on Formal Work-based Learning Programs in the UK. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. ACM, Portland OR USA, 1028–1034. https://doi.org/10.1145/3626252.3630879

[33] Annabel Rothschild, Amanda Meng, Carl DiSalvo, Britney Johnson, Ben Rydal Shapiro, and Betsy DiSalvo. 2022. Interrogating Data Work as a Community of Practice. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 1–28. https://doi.org/10.1145/3555198

[34] Ben Rydal Shapiro, Amanda Meng, Annabel Rothschild, Sierra Gilliam, Carl DiSalvo, and Betsy DiSalvo. 2023. "Bettering Data": The Role of Everyday Language and Visualization in Critical Novice Data Work. *Educational Technology & Society* 25, 4 (2023), 109–125.

[35] Ella Taylor-Smith, Sally Smith, Khristin Fabian, Tessa Berg, Debbie Meharg, and Alison Varey. 2019. Bridging the Digital Skills Gap: Are Computing Degree Apprenticeships the Answer?. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*. ACM, Aberdeen Scotland Uk, 126–132. https://doi.org/10.1145/3304221.3319744

[36] Kyle Thayer and Amy J. Ko. 2017. Barriers Faced by Coding Bootcamp Students. In *Proceedings of the 2017 ACM Conference on International Computing Education Research*. ACM, Tacoma Washington USA, 245–253. https://doi.org/10.1145/3105726.3106176

[37] Mike Tissenbaum, David Weintrop, Nathan Holbert, and Tamara Clegg. 2021. The Case for Alternative Endpoints in Computing Education. *British Journal of Educational Technology* 52, 3 (May 2021), 1164–1177. https://doi.org/10.1111/bjet.13072

[38] Michelle Hoda Wilkerson and Joseph L. Polman. 2020. Situating Data Science: Exploring How Relationships to Data Shape Learning. *Journal of the Learning Sciences* 29, 1 (Jan. 2020), 1–10. https://doi.org/10.1080/10508406.2019.1705664

[39] Timothy T. Yuen and Kay A. Robbins. 2015. A Qualitative Study of Students' Computational Thinking Skills in a Data-Driven Computing Class. *ACM Transactions on Computing Education* 14, 4 (Feb. 2015), 1–19. https://doi.org/10.1145/2676660

[40] Jia Zhu, Stephanie J. Lunn, and Monique Ross. 2023. Characterizing Women's Alternative Pathways to a Computing Career Using Content Analysis. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*. ACM, Toronto ON Canada, 158–164. https://doi.org/10.1145/3545945.3569798

[41] Abigail Zimmermann-Niefield, R. Benjamin Shapiro, and Shaun Kane. 2019. Sports and Machine Learning: How Young People Can Use Data from Their Own Bodies to Learn about Machine Learning. *XRDS: Crossroads, The ACM Magazine for Students* 25, 4 (July 2019), 44–49. https://doi.org/10.1145/3331071