

DataWorks: Building Smart Community Capacity

1951818

Betsy DiSalvo, Georgia Institute of Technology

IRG Track 2, 2019

Principal Research Investigators

- Betsy DiSalvo, PI, Georgia Institute of Technology
- Carl DiSalvo, co-PI, Georgia Institute of Technology
- Ellen Zegura, co-PI, Georgia Institute of Technology
- Cedric Stallworth, co-PI, Georgia Institute of Technology
- Ben Shapiro, co-PI, Georgia State University

Georgia Tech Constellations Center for Equity in Computing

Community Partners

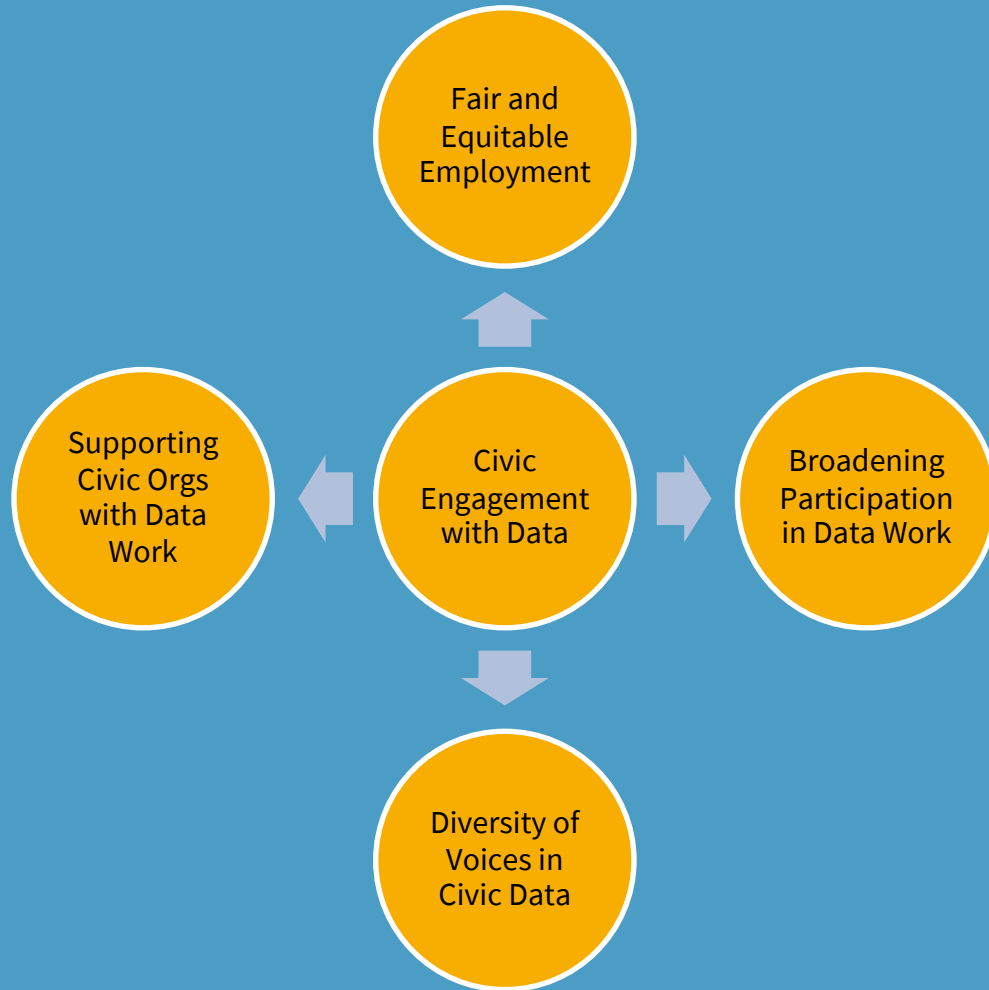
- Atlanta Housing Authority
- Atlanta Regional Commission
- Atlanta Wealth Building Initiative
- Center for Civic Innovation
- GreenLight Foundation
- LifeLine Animal Shelter/ Dekalb Human Society
- Raising Expectations
- CHRIS 180
- PwC

DATA

WORKS

**DATA
FOR ALL.**

Project Overview



DataWorks is a unique data services provider

DataWorks recruits people from economically disadvantaged neighborhoods and underrepresented groups in computing to train and employ them as Data Wranglers.

1. To demonstrate how the context of data work, including the lived experiences of data workers, shapes civic data.
2. To demonstrate new pathways to data literacy, particularly pathways for members of minoritized communities.

Project Overview

Data shapes our cities, but **the people involved in collecting, cleaning, formatting, analyzing and communicating data** to inform civic decisions **are not representative** of the citizens.

DataWorks trains and employs citizens of **Atlanta, Georgia**, from minoritized neighborhoods, to wrangle data for civic partners filling two needs:

1. Improving civic data with a diversity of citizens involved with its preparation
2. Building data literacy and critical data literacy among minoritized communities to:
 - a) improve civic engagement and
 - b) create job opportunities

Research Contributions

1. How can we broaden participation in civic engagement with data work?
2. What tools are needed to scaffold entry-level data work?
3. How can contextual experience be of value to civic data work?
4. How can we use the university as
 - (a) a platform for civic engagement and
 - (b) for equity and inclusion?

Project Update

Clients and Projects

- 14 Clients
- 17 projects (12 completed and 5 in progress)
- Income from Clients: ~\$25,000!

Data Workers

- **4** Part-time workers in Jan 2020
- **3** Part-time workers in April 2020
(1 data worker left for new job)
- **4** Part-time workers in July 2020
- **3** Part-time workers in August 2020
(1 worker let go, could not work from home)
- **5** Full-time workers in April 2021

All workers went **from 20 hours a week to 30 hours a week** in Oct 2020

All workers were hired as **fulltime Georgia Tech employees** as of Jan 2021

Publications

Exploring Approaches to Data Literacy Through a Critical Race Theory Perspective, **ACM CHI 2021 (honorable mention for best paper)**. Britney Johnson, Ben Rydal Shapiro, Betsy DiSalvo, Annabel Rothschild, and Carl DiSalvo

Data collection

- Weekly interviews with data workers
- Weekly interviews with management team
- Interviews with civic clients
- Developing and testing data literacy measures (to be submitted to learning or computing venue in 2021)
- Workshops with data workers
- Developing a Data 101 workshop with civic clients
- Interviews with data workers in civic sector (to be submitted to CSCW in April 2021)
- Data tools catalogue – collecting list from interviews
- +more!

Project Evolution

Civic Engagement and Operations

Requires more than technical skill, it also requires

- Project management
- Self-regulation
- A structured work environment
- Communication

Partners are seeking us out because of data workers personal experience and context:

- DataWorks hired for data collection about issues facing Atlanta in minoritized communities.
- DataWorks hired for annotating ML training data (children's moral narratives and African American Vernacular English projects)

Future Research

- How can we broaden participation in civic engagement with data work?
Findings: Novice data workers' perceptions and vernacular language changes how we might define data science terminology and career paths
- What tools are needed to scaffold entry-level data work?
Findings: Data workers need tools that bridge entry-level and expert tools.
- How can contextual experience be of value to civic data work?
Findings: Data Works is be sought out for expertise in AAVE and civic connections in Atlanta . There is a need in civic sector for this contextual knowledge.
- How can we use the university as a platform for civic engagement and equity & inclusion?

Anticipated outcomes & success measures for next year

Novice Data Workers' Perceptions of Data

How can we measure data literacy and particularly critical data literacy?

How does a change in novice data workers perceptions of data change their civic engagement?

Current: Developing awareness of issues because of data

Current: Understanding how civic issues are shaped by data and data interpretation

Next: Realizing their power with data to shape their community

“Data collection?... that’s the first time I’ve heard that term” – DataWorker, January 2020

“I realized that one thing about data, ... you have to get in good with it, you have to love it for what it is, and you have to be able to love it to organize it” DataWorker, September 2020

Civic Data Literacy for Non-Data Scientist/Analyst

What does it mean to have a distributed, unstandardized workforce completing integral steps of the civic data lifecycle?

How can we support the range of skills that civic data works have ensure quality and accuracy and promote the work experts in civic domains?

Current: Refine definition of data work through literature review and analysis of previously-conducted interviews

Next: develop an in-situ data literacy curriculum for the Data Wranglers