

*From Rasters to Rows:*  
New Methods and  
Applications in Automated  
Data Extraction

AUGUST 15-16, 2024



# Agenda

Thursday, August 15

**12:00 noon–1:00 p.m.**      **Lunch and Registration**

**1:00 p.m.–1:10 p.m.**      **Welcome**

Daniel Moulton, CREED, Federal Reserve Bank of Philadelphia

**1:10 p.m.–1:20 p.m.**      **Opening Remarks**

Patrick T. Harker, President and CEO, Federal Reserve Bank of Philadelphia

**1:20 p.m.–2:20 p.m.**      **Session 1: The Redlining Lab**

Wenfei Xu, Cornell University

Katherine Thomas, New York University

**2:20 p.m.–3:05 p.m.**      **Session 2: Racial Covenants in Philadelphia**

Robyn Smith, CREED, Federal Reserve Bank of Philadelphia

**3:05 p.m.–3:25 p.m.**      **Break**

**3:25 p.m.–4:25 p.m.**      **Session 3: Ghost Neighborhoods: Digitizing the Columbus Sanborn Maps**

Harvey Miller, The Ohio State University

**4:25 p.m.–5:10 p.m.**      **Session 4: The Historical Housing Prices Project**

Allison Shertzer, Federal Reserve Bank of Philadelphia

**5:10 P.M.**                      **Reception**

# Agenda

Friday, August 16

- 8:00 a.m.–9:00 a.m.**      **Breakfast and Registration**
- 9:00 a.m.–9:25 a.m.**      **Session 5: A Compound AI System for Cheaper and Efficient Table Digitization at Scale**
- Vitaly Meursault, Federal Reserve Bank of Philadelphia
- 9:25 a.m.–10:25 a.m.**      **Session 6: The Decennial Census Digitization and Linkage Project**
- Katie R. Genadek, University of Colorado Boulder and U.S. Census Bureau
- 10:25 a.m.–10:35 a.m.**      **Break**
- 10:35 a.m.–11:25 a.m.**      **Session 7: Automatic Digitization of the Census of Housing Block Statistics**
- Jeffrey Lin, Federal Reserve Bank of Philadelphia  
Daniel Moulton, CREED, Federal Reserve Bank of Philadelphia
- 11:25 a.m.–12:25 p.m.**      **Session 8: The Costs of Housing Regulation: Evidence from Generative Regulatory Measurement**
- Arpit Gupta, New York University
- 12:25 p.m.–12:30 p.m.**      **Closing Remarks**

# Biographies

## Katie R. Genadek

***Director, Decennial Census Digitization and Linkage (DCDL) Project***

**Economist, U.S. Census Bureau**

Katie R. Genadek is the director of the Decennial Census Digitization and Linkage (DCDL) project and an economist working at the U.S. Census Bureau. She is also a faculty associate at the University of Colorado Boulder. She previously worked at the University of Minnesota, where she managed the IPUMS-USA data project. She is a demographer and economist, and her research is focused on the relationship between policy, work, family, and time use.

Genadek has a Ph.D. in applied economics from the University of Minnesota, and an M.S. in applied economics and a B.S. in economics from Montana State University.

## Arpit Gupta

***Assistant Professor of Finance***

**New York University Stern School of Business**

Arpit Gupta joined the New York University Stern School of Business as an assistant professor of finance in September 2016.

Gupta's research interests focus on using large data sets to understand default dynamics in household finance, real estate, and corporate finance. His recent papers examine the role of foreclosure contagion in mortgage markets and estimate the impact of adverse health events on foreclosures and bankruptcies. He received the 2016 Top Finance Graduate Award from Copenhagen Business School.

Gupta has a Ph.D. in finance and economics from Columbia Business School and a B.S. in mathematics and economics from the University of Chicago.

## Patrick T. Harker

***President and CEO***

**Federal Reserve Bank of Philadelphia**

Patrick T. Harker took office on July 1, 2015, as the 11th president and chief executive officer of the Federal Reserve Bank of Philadelphia. In this role, Harker participates on the Federal Open Market Committee, which formulates the nation's monetary policy.

Before taking office at the Philadelphia Fed, Harker was the 26th president of the University of Delaware. He was also a professor of business administration at the university's Alfred Lerner College of Business and Economics and a professor of civil and environmental engineering at the College of Engineering.

# Biographies

Before joining the University of Delaware in 2007, Harker was dean and Reliance Professor of Management and Private Enterprise at the Wharton School of the University of Pennsylvania. Before being appointed dean in 2000, he was the Wharton School's interim dean and deputy dean as well as the chair of its operations and information management department. In 1991, he was the youngest faculty member in Wharton's history to be awarded an endowed professorship as UPS Transportation Professor of the Private Sector. He has published/edited nine books and more than 100 professional articles. From 1996 to 1999, he was editor-in-chief of the journal *Operations Research*.

In 2012, Harker was named a fellow of the Institute for Operations Research and the Management Sciences (INFORMS) and a charter fellow of the National Academy of Inventors. He was also named a White House fellow by President George H. W. Bush in 1991 and was a special assistant to FBI Director William S. Sessions from 1991 to 1992.

Harker is a member of the Select Operating Committee of Select Greater Philadelphia. He previously was on the boards of Catholic Relief Services, Pepco Holdings, Inc., and Huntsman Corporation and was a founding member of the board of advisors for Decision Lens, Inc. He was also a nonbanking Class B director of the Philadelphia Fed from 2012 to 2015.

Harker has a Ph.D. in civil and urban engineering, an M.A. in economics, and an M.S.E. and B.S.E. in civil engineering, all from the University of Pennsylvania.

## Jeffrey Lin

***Vice President and Economist***  
**Federal Reserve Bank of Philadelphia**

Jeffrey Lin is vice president and economist at the Federal Reserve Bank of Philadelphia. He leads the regional and microeconomics section and the team responsible for producing the *Manufacturing Business Outlook Survey* and other economic indicators. He is codirector of the Philadelphia Federal Statistical Research Data Center. His research interests are urban and regional economics and economic growth. In recent research, he has studied factors shaping long-run metropolitan and neighborhood dynamics and simultaneous invention by multiple independent inventors.

His articles have appeared in leading economics journals, including the *Quarterly Journal of Economics*, *Review of Economic Studies*, *American Economic Review*, and *Review of Economics and Statistics*. His 2012 article with Hoyt Bleakley, "Portage and Path Dependence," received an IPUMS Research Award.

Lin has a Ph.D. in economics from the University of California, San Diego, and a B.A. in economics and mathematical methods in the social sciences from Northwestern University.

# Biographies

## Vitaly Meursault

***Machine Learning Economist***  
**Federal Reserve Bank of Philadelphia**

Vitaly Meursault is a machine learning economist at the Federal Reserve Bank of Philadelphia. He uses machine learning to study questions in economics and finance. His current work focuses on extracting economically meaningful signals from language, combining his economics training with his interest in history. One goal of his research is to develop applications of natural language processing for regulatory purposes because regulators often make decisions based on large amounts of text.

Meursault has a Ph.D. in financial economics from Carnegie Mellon University, a M.A. in financial economics from the European University at St. Petersburg (Russia), and a B.A. in history from St. Petersburg State University.

## Harvey Miller

***Professor of Geography and Director, Center for Urban and Regional Analysis***  
**The Ohio State University**

Harvey Miller is the Bob and Mary Reusche Chair in Geographic Information Science, professor of geography, and director of the Center for Urban and Regional Analysis at the Ohio State University. He is also a courtesy professor of city and regional planning at the Ohio State University and a faculty affiliate of the university's Translational Data Analytics Institute, Sustainability Institute, and Institute for Population Research, in addition to being an affiliate of the Center for Injury Research and Policy at Nationwide Children's Hospital in Columbus, OH.

Miller's research focuses on sustainable, resilient, and equitable mobility and communities. He develops and applies spatial and moving objects analytics to questions and problems involving active and public transportation, multimodal transportation systems, traffic safety, and public health. Miller is a pioneer of analytical time geography, an extension of Torsten Hägerstrand's theory of the necessary space-time conditions for human activities. He also develops concepts and theories that reframe spatial analysis for a data-rich era.

Miller has a Ph.D. in geography from the Ohio State University, and an M.A. and a B.A. in geography from Kent State University.

## Daniel Moulton

***Senior Manager of Data Science and Engineering***  
**Federal Reserve Bank of Philadelphia**

Daniel Moulton is senior manager of data science and engineering in the Consumer Finance Institute (CFI) at the Federal Reserve Bank of Philadelphia. He is responsible for building and managing research support staff, maximizing the efficiency of programming and computing infrastructure, building experience in cloud computing

# Biographies

and machine learning, and developing and disseminating data sets that contribute to many areas of research. He cofounded and helps lead the Center for the REstoration of Economic Data (CREED). During the pandemic, Moulton brought novel, near-real-time data into the Bank and led work to connect this information to policy. His research focuses on the use of machine learning (ML) for data extraction, ML and the credit markets, and on the long-run effects of discrimination in credit markets and housing.

Before joining the Philadelphia Fed, Moulton built the data science team at HotelTonight and subsequently led the hotels-focused data science team at Airbnb. He spent four years at Palantir Technologies in a hybrid data science/data engineering role and five years at the hedge fund Bridgewater Associates doing macroeconomic research.

Moulton has an A.B. in mathematics from Princeton University.

## Allison Shertzer

***Economic Advisor and Economist***  
**Federal Reserve Bank of Philadelphia**

Allison Shertzer is an economic advisor and economist in the Research Department at the Philadelphia Fed. Shertzer's research focuses on the economic development of American cities and the connection between a city's history and present-day economic opportunity.

Her work explores how past decisions about land use and infrastructure have shaped cities today. Shertzer is interested in housing affordability and migration, particularly how to make economic opportunity more widely available.

Shertzer is passionate about delving into archives and visualizing the history of cities through map digitization, a tool that can help everyone — from students to journalists to policymakers — better understand the development of U.S. cities. In the past, she used this type of historical digitization to study redlining, segregation, and zoning.

Before joining the Bank in 2023, Shertzer was an associate professor of economics at the University of Pittsburgh, having started there as an assistant professor of economics in 2011. Shertzer was previously a faculty research fellow with the National Bureau of Economic Research from 2012 to 2019 and then became a research associate.

Shertzer has a Ph.D. and an M.A. in economics from UCLA, and a B.S. in industrial engineering and a B.A. in mathematics from Arizona State University.

## Robyn Smith

***Data Scientist***  
**Federal Reserve Bank of Philadelphia**

Robyn Smith is a data scientist in the Center for the Restoration of Economic Data (CREED) at the Federal Reserve Bank of Philadelphia. Smith joined the Philadelphia Fed in 2023. Her research interests focus on the ways that laws, policies, and systemic bias influence residential composition and wealth accumulation in Philadelphia and other major cities. She applies her analytical expertise to synthesize data and build models.

# Biographies

Smith developed her expertise in image processing, temporal analysis, and model building and fitting using both new and archival imaging and spectroscopic data from various ground- and space-based telescopes. Her thesis work was aimed at understanding the complex regions immediately surrounding accreting supermassive black holes at the center of other galaxies. Her work in the field of astronomy also taught her the value of archival data and the importance of its public availability.

Smith has a Ph.D. in astronomy from the University of Maryland, College Park, and a B.S. in physics from Drexel University.

## Katherine Thomas

*Ph.D. Student*  
**New York University**

Katherine Thomas is a Ph.D. student in sociology at New York University. Her work uses quantitative and computational methods to address the causes and consequences of spatial inequality. Two of her current projects address the relationship between environmental risk and residential mobility, and the impacts of historical mortgage lending on contemporary racial inequality and segregation. Before graduate school, Thomas was a research analyst at the Urban Institute.

Thomas has a B.A. in statistics and sociology from Rice University.

## Wenfei Xu

*Assistant Professor, Cornell University*  
**Director, Urban Data Research Lab**

Wenfei Xu is an assistant professor in urban planning in the Department of City and Regional Planning at the College of Architecture, Art, and Planning at Cornell University and director of the Urban Data Research Lab. Xu's main research traces the legacies of historical structural disinvestment and asks how they have shaped present-day inequality and processes of neighborhood change.

Xu has a Ph.D. in urban planning from Columbia University, a dual master's degree in architecture and urban planning from MIT, and a B.A. in economics from the University of Chicago.

---

### ORGANIZING COMMITTEE

Jeffrey Lin

Daniel Moulton

Larry Santucci

Allison Shertzer

Bryan Stuart