NATIONAL ACADEMIES
Sciences
Engineering
Medicine

AI Day for Federal Statistics: CNSTAT Public Event
May 2, 2024
National Academy of Sciences Building
2101 Constitution Avenue NW, Washington, DC

PROGRAM

Location: Fred Kavli Auditorium

10:00am–10:10am	**Introductory Remarks** Katharine Abraham, University of Maryland, College Park & CNSTAT Chair
10:10am–10:30am	**Federal Statistics at the Forefront of Artificial Intelligence** Karin Orvis, Chief Statistician of the United States, Office of Management and Budget
10:30am-12:00pm	**Panel on Cross-Cutting Issues Related to AI Adoption in the Federal Statistics System** *(speakers will have up to 20 minutes for their remarks)* **Statistics at the AI Leadership Table** Chris Marcum, Office of Management and Budget **Confidentiality Protection in AI** Xiao-Li Meng, Harvard University **Moving AI Tools into Government** Nadia Fawaz, White House Presidential Innovation Fellows **Return on Investment (ROI) on AI** Kent Cunningham, Microsoft Federal
12:00pm–1:30pm	**Lunch break**
1:30pm–2:30pm	**The Future of AI in Federal Statistics** Nancy Potok, NAPx Consulting LLC
2:30pm–2:45pm	**Break (move to breakout rooms)**

CNSTAT
Serving the Statistical Community since 1972

FCSM
Federal Committee on
Statistical Methodology

NISS | National Institute of
Statistical Sciences

Location: Various Rooms

2:45pm–4:00pm   **Topical Breakout Sessions** *(all sessions concurrent, speakers will have up to 20 minutes for their remarks followed by audience Q&A and discussion)*

1) **Use of AI in the Survey Process** *(Room 120)*
   - Gizem Korkmaz, Westat
   - Michael Link, Ipsos
   Moderator: Minsun Riddles, Westat

2) **Record Linkage and AI** *(East Court)*
   - Carla Medalia, Census Bureau
   - Cordell Golden, National Center for Health Statistics
   Moderator: Keenan Dworak-Fisher, Bureau of Labor Statistics

3) **Building the Future: Leveraging the National AI Research Resource and National Secure Data Service for AI Capabilities** *(Lecture Room)*
   - Katerina Antypas, National Science Foundation Office of Advanced Cyberinfrastructure
   - Lisa Mirel, National Center for Science and Engineering Statistics
   Moderator: Travis Hoppe, Office of Science and Technology Policy

4) **Statistical Core of AI** *(West Court)*
   - Tian Zheng, Columbia University
   - Hongtu Zhu, University of North Carolina at Chapel Hill
   Moderator: Rebecca Nugent, Carnegie Mellon University

5) **Current Scope of AI Implementations** *(Auditorium)*
   - Ken Haase, Census Bureau
   - David Oh, Bureau of Labor Statistics
   Moderator: Bob Siviniski, Office of Management and Budget

6) **Quantifying Uncertainty Due to AI** *(Room 125)*
   - Lucas Janson, Harvard University
   - David Matteson, National Institute of Statistical Sciences and Cornell University
   Moderator: Nicholas Rios, George Mason University

7) **Ethics in AI** *(Members Room)*
   - Rochelle Tractenberg, Georgetown University
   - Michael B. Hawes, Census Bureau
   Moderator: Mark Prell, Economic Research Service

8) **AI and Blended Data** *(Board Room)*
   - H. V. Jagadish, University of Michigan
   - Elizabeth Stuart, Johns Hopkins University
   Moderator: Brian Habing, National Institute of Statistical Sciences and University of South Carolina

Location: West Court and Great Hall

4:00pm–5:00pm   **Poster Session and Reception**

Location	Poster Title	Presenter
P1	Setting the Foundation for AI Usage in Regional Statistics	**Nicholas Wetzler** Bureau of Economic Analysis
P2	Offense Text Autocoder	**Suzanne Strong** Bureau of Justice Statistics and RTI
P3	Clustering Federal Register Comments for Efficient, Manual Review	**Brandon Kopp** Bureau of Labor Statistics
P4	Machine Learning Classification of Product Categories in Scanner Data used CPI	**Brendan Williams** Bureau of Labor Statistics
P5	AI Institute for Societal Decision Making	**Rebecca Nugent** Carnegie Melon University
P6	Using Artificial Intelligence Tools to Automate Manual Data Processing and Generate Alternative Data Sources	**Haley Hunter-Zinck** U.S. Census Bureau
P7	Automated Change Detection for MAF/TIGER	**Elvis Martinez** U.S. Census Bureau
P8	Classifying Consumer Ultra-Processed Food Intake Using NLP and Convoluted Neural Networks	**Mariah Ehmke** Economic Research Service
P9	Capture-Recapture in the Age of AI	**Habtamu K. Benecha** National Agricultural Statistics Service
P10	Using Machine Learning Algorithms to Identify Farms on the 2022 Census of Agriculture	**Gavin Corral** National Agricultural Statistics Service
P11	A Natural Language Processing Examination of Stress and Financial Wellbeing in Consumer Complaints	**Rebecca Martin** Consumer Financial Protection Bureau
P12	Expanding Text Analysis Use at EIA	**Mark Schipper** Energy Information Administration
P13	Use of AI and Impact on Employees: Data From the 2022 Annual Business Survey	**Alexander Rhodes** National Center for Science and Engineering Statistics
P14	Classification of Open-Ended comments From Surveys Based on XGBoost and Natural Language	**Lu Chen** National Institute of Statistical Sciences
P15	Automating Declassification and FOIA Requests: When AI Leads to Increased Government Transparency	**Sam Stehle** State Department
P16	Imputing Measures of Diet Quality using Circana Scanner Data	**Alexander Stevens** Economic Research Service

Location	Poster Title	Presenter
P17*	Measuring and Analyzing Community- and Officer-Initiated Activity	**Jacob Cramer** Bureau of Justice Statistics and RTI
P18*	Unsupervised LLM Topic Modeling for Enhanced Federal Survey Interpretation	**Irina Belyaeva** U.S. Census Bureau
P19	Tracking Firm Use of AI in Real Time: A Snapshot from the Business Trends and Outlook Survey	**Cory Breaux** U.S. Census Bureau
P20	Autocoder for OCWC and OEWS	**Sertan Akinci** Bureau of Labor Statistics
P21	CE Diary Autocoder	**Melissa Pollock** Bureau of Labor Statistics
P22	Explainable Artificial Intelligence for Bias Identification and Mitigation in Demographic Models	**Atul Rawal** U.S. Census Bureau
P23	A Semi-Supervised Active Learning Approach for Block-Status Classification.	**Atul Rawal** U.S. Census Bureau
P24	Corn Yield Prediction Using Remote Sensing Observations and Multi-Source Unsupervised Domain Adaptation	**Zhengwei Yang** National Agricultural Statistics Service
P25	Using Neural Networks to Assess Uncertainty for the Imputation of an Area Survey	**Luca Satore** National Agricultural Statistics Service
P26	Using Data for Training and Validating ML Prediction Models and Explainable AI	**Orlando Davy** National Center for Health Statistics
P27	Using AI and ML to Identify Non-Name Text in Name Fields that Should be Removed Prior to Linkage	**Frances McCarty** National Center for Health Statistics
P28	Implementing Retrieval-Augmented Generation with Survey Question Evaluation Reports	**Priyam Patel** National Center for Health Statistics
P29	AI-Powered Synthetic Data Generation: Algorithms, Artifacts, Analyses	**Gary Howarth** National Institute of Standards and Technology
P30	Mixed Modeling Approach for Characterizing the Genetic Effects in a Longitudinal Phenotype	**Pei Zhang** National Institutes of Health, Graduate Partnership Program
P31	AI in Psychometrics: Estimating Factor Loadings with ChatGPT	**J. Peter Leeds** Office of Personnel Management
P32	A Comparative Analysis between AI and Human Coding in Survey Research	**Isabela Bertolini Coelho** University of Maryland-JPSM