# DATA INSTITUTE CONFERENCE SCHEDULE
San Francisco | March 12-14, 2023 | #DSCO23

## SUNDAY, MARCH 12

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00am-9:00am</td>
<td>Coffee and Registration</td>
<td>5th Floor</td>
</tr>
<tr>
<td>9:00am-11:30am</td>
<td>Short Course Session I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BAYESIAN METHODS IN A/B TESTING</td>
<td>Nathaniel Stevens, University of Waterloo, Rm 450</td>
</tr>
<tr>
<td></td>
<td>INTRODUCTION TO DEEP LEARNING WITH PYTORCH</td>
<td>Carlos Garcia, University of San Francisco, Rm 452</td>
</tr>
<tr>
<td>12:30pm-3:00pm</td>
<td>Short Course Session II</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ADAPTING PRETRAINED MODELS FOR DOCUMENT CLASSIFICATION</td>
<td>David Guy Brizan, University of San Francisco, Rm 450</td>
</tr>
<tr>
<td></td>
<td>INTRODUCTION TO GRAPH NEURAL NETWORKS AND THEIR GENERALIZATIONS</td>
<td>Mustafa Hajij, University of San Francisco, Rm 452</td>
</tr>
<tr>
<td>9:00am-3:00pm</td>
<td>Datathon</td>
<td>Rm 527</td>
</tr>
<tr>
<td></td>
<td>Hosted by Amazon Web Services &amp; Amazon SageMaker Studio Lab</td>
<td></td>
</tr>
<tr>
<td>3:30pm</td>
<td>Opening Remarks</td>
<td>Rm 150</td>
</tr>
<tr>
<td></td>
<td>Henry Humadi and Shan Wang, Conference Co-Chairs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jeff Hamrick, Senior Director, Data Institute</td>
<td></td>
</tr>
<tr>
<td>4:00pm-5:00pm</td>
<td>Plenary</td>
<td>Rm 150</td>
</tr>
<tr>
<td></td>
<td>Anim Anandkumar, Bren Professor, Caltech &amp; Senior Director ML Research, Nvidia</td>
<td></td>
</tr>
<tr>
<td>5:00pm-6:00pm</td>
<td>Opening Reception</td>
<td>5th Floor</td>
</tr>
<tr>
<td></td>
<td>Light Refreshments and Hors d’oeuvres</td>
<td></td>
</tr>
</tbody>
</table>

## MONDAY, MARCH 13

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00am-9:00am</td>
<td>Coffee and Registration</td>
<td>5th Floor</td>
</tr>
<tr>
<td>9:00am-11:00am</td>
<td>Concurrent Sessions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PRACTICAL ISSUES AND ADVANCES IN A/B TESTING</td>
<td>Nathaniel Stevens, University of Waterloo, Rm 527</td>
</tr>
<tr>
<td></td>
<td>Nicholas Larsen, North Carolina State University</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HODOR: A Two-Stage Hold-Out Design for Online Controlled Experiments on Networks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jiannan Lu, Apple</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Towards Trustworthy, Efficient and Private A/B Tests: Learnings and Insights</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lo-Hua Yuan, Airbnb</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How Airbnb Learns From Product Releases Without A/B Testing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cindy Zhang, Pinterest</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Long-term Holdouts as Ground Truth Measurements</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DATA SCIENCE IN SPORTS</td>
<td>Steve Devlin, University of San Francisco, Rm 529</td>
</tr>
<tr>
<td></td>
<td>Justin Jacobs, Squared2020, NBA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A Brief Tour into Computer Vision in Sports</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jake Toffler, New York Mets</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Model Benchmarking in a Competitive Industry</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thomas Treloar, Hillsdale College</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Global and Iterative Ranking Models Based on Network Diffusion</td>
<td></td>
</tr>
<tr>
<td></td>
<td>David Uminsky, University of Chicago</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NBA Lineup Attribution: A Fourier Transform Approach</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TOPICS IN MACHINE LEARNING</td>
<td>Jeff Hamrick, University of San Francisco, Rm 150</td>
</tr>
<tr>
<td></td>
<td>Oliver Zeigermann, Open Knowledge</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Resilient Machine Learning</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phil Mui, Salesforce</td>
<td></td>
</tr>
<tr>
<td></td>
<td>An Embedding Analysis on the Impact of Diversity Awareness on Diversity of Ideas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yuchi Cho, University of Southern California</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Flexible and Robust Real-Time Intrusion Detection Systems to Network Dynamics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Connor Gibbs, Colorado State University</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ECoHeN: A Hypothesis Testing Framework to Extract Communities from Heterogeneous Networks</td>
<td></td>
</tr>
</tbody>
</table>

**UNIVERSITY OF SAN FRANCISCO**
MONDAY, MARCH 13
11:15am-12:15pm: Mentor-Mentee Lunch, 5th Floor

12:30pm-1:30pm: Plenary, Rm 150
Dr. Patil, former U.S. Chief Data Scientist

1:30pm-2:00pm: Coffee Break
2:00pm-4:00pm: Concurrent Sessions

DATA PRIVACY AND SECURITY RISKS EXPOSED BY AI, Rm 527
Chair: Hyrum Anderson, Robust Intelligence
Daniel Zielocki, Salesforce
The importance of AI to Business, Marketing, and Digital Products

Navaneet Jain, Salesforce
Privacy by Design: Building Privacy into your Data Systems from the Ground Up

Hyrum Anderson, Robust Intelligence
Data and Model Supply Chain Risk in the AI Development Lifecycle

Erwin Quiroga, International Computer Science Institute (ICSI) and Ruhr University Bochum
Dox and Dots of Machine Learning in Computer Security

MACHINE LEARNING APPLICATIONS IN INDUSTRY, Rm 529
Co-Chairs: Jennifer Zhu and Yipeng Sagar
Ravipati, Amazon Web Services
Sean McCurdy, Pinterest
Intro to Pinterest Ads and ML Systems

Saman Sarraf, Johnson & Johnson
Computer Vision Applications Across Industries: Wins and Challenges

Ruijuan Zhang, Airbnb
Machine Learning Driven User Growth and Lifecycle Engagement Marketing

April Liu, Intuit
AI Application in Fighting Synthetic Account Fraud

STATISTICAL THEORY AND APPLICATIONS, Rm 150
Chair: Daniel Jerison
Moinak Bhaduri, Bentley University
Distribution-free, Online Change Detection in Multidimensional Point Processes Through Repeated Testing

Wiranthu Bandara Herath, Drake University
Dimension Reduction for Vector Autoregressive Models

Mustafa Huij, University of San Francisco
What is Topological Deep Learning?

Greeshma Balabhadra, Stony Brook University
High-Frequency Risk Estimators Using Change Point Detection Methods

TUESDAY, MARCH 14
8:00am-9:00am: Coffee and Registration, 5th Floor
9:00am-11:00am: Concurrent Sessions

AI FOR PRECISION MEDICINE AND LEARNING HEALTHCARE SYSTEMS, Rm 527
Chair: William Boz, University of San Francisco
Conrad Yu, Trident.ai
AI in Medicine: What Are We Really Trying to Do?

Vincent Liu, Kaiser Permanente Research
Augmented Intelligence: The Future of AI in Healthcare

Sophia Wang, Stanford University
Envisioning the Future: Artificial Intelligence for Preserving Eyewitness

William Boz, University of San Francisco
AI Meets Cor Pulmonics: Computational Biomedical Knowledge for Integrated, Precision Medicine

PRACTICAL ISSUES AND ADVANCES IN A/B TESTING (II), Rm 529
Chair: Nathaniel Stevens, University of Waterloo
Wenjing Zheng, Netflix
Double-Robust Causal Effect Generalization and Transportation under Covariate Shifts

Nick Ross, University of Chicago
Hidden Integration Costs in Online Controlled Experimentation Platforms

Nathaniel Stevens, University of Waterloo
General-Additive Network Effect Models: A Framework for the Design and Analysis of Experiments on Networks

Steve Howard, The Voleon Group
Augmented Inverse Propensity Weighting for Randomized Experiments

MACHINE LEARNING IN INDUSTRY, Rm 150
Session Chair: Daniel O’Connor, University of San Francisco
Dili Ezeme, Ab-InBev
New Assortment Recommendation Approach for Differentiated Products

Rushi Manglik, University of San Francisco
Fine-Tuning Layout Pager Deep Learning Model for Document Image Analysis and Table Extraction

Diane Woodbridge, University of San Francisco
Bundle Recommender for Complementary Menus

Sanjiv Das, Santa Clara University
Wealth Management Benchmarks

Matthew Dixon, Illinois Institute of Technology
Optimizing Treatment Selection in Colitis Disease Using Patient-Specific Features: An Individual Participant Data Meta-Analysis of Fifteen Randomized Controlled Trials

IN MEDICINE, Rm 150
Chair: Shan Wang
Zezi Zhang, University of California, Irvine
Parameter Inference in Diffusion-Reaction Models of Glioblastoma Using Physics-Informed Neural Networks

Sourendipo Ghosh Dutrill, University of Minnesota Twin Cities
Correcting for Spatial Correlation in Functional Connectivity from task-based fMRI

Vignesh Ravindranath, University of California, San Francisco
Optimizing Treatment Selection in Colitis Disease Using Patient-Specific Features: An Individual Participant Data Meta-Analysis of Fifteen Randomized Controlled Trials

IN A/B TESTING (II)
PRACTICAL ISSUES AND ADVANCES
11:30am-12:20pm: Lunch Break

IN MARKETING
Chair: Chris Brooks
IN INDUSTRY
Chair: Sean McCurdy

12:30pm-1:30pm: Plenary
Eric Befeler, CEO Vibrant Data Labs

1:30pm-2:00pm: Coffee Break
2:00pm-4:00pm: Concurrent Sessions

PRIVACY AND PERSONALIZATION IN MARKETING, Rm 527
Chair: Daniel Zielocki, Salesforce
Carson Carter, Twitch (Amazon)
Putting Personalization to the Test: A Causal Inference Approach

Sherry Guo, Salesforce
From 0 to 1: Leveraging Data Science Algorithms for Personalization

Chad Kimmer, Meta Reality Labs
Navigating the Intersection of Personalization and Privacy: A Marketer’s Perspective

Maura Tuohy, Darcx
The Fine Line: Balancing Personalization and Legal Compliance

REINFORCEMENT LEARNING FOR WEALTH MANAGEMENT, Rm 529
Chair: Matthew Dixon, Illinois Institute of Technology
Kiran Chaitlin, Money Lion
Wealth Management Benchmarks

Sanjiv Das, Santa Clara University
Optimizing The Probabilities Of Obtaining Single And Multiple Financial Goals

Matthew Dixon, Illinois Institute of Technology
Time Consistent Risk-Aware Q-learning of Optimal Consumption under Epstein-Zin Preferences

MACHINE LEARNING AND STATISTICS IN MEDICINE, Rm 150
Chair: Shan Wang

IN A/B TESTING (II)
PRACTICAL ISSUES AND ADVANCES
3:00pm-4:00pm: Concurrent Sessions

4:30pm-5:30pm: Ethical Perspectives of AI-Generated Art, Rm 150
Quinn Keck, Disney Streaming
Chris Brooks, University of San Francisco
Paul Kim, University of California, Santa Barbara

5:30pm-6:30pm: Closing Reception, 5th Floor
Light Refreshments and Hors d'oeuvres

Thank You — Sponsors & Members

Craig Newmark Philanthropies

IN COOPERATION WITH:

SIAH

ASAI

Pendulum
ABOUT THE DATA INSTITUTE

MISSION

• Foster new partnerships between industry and academia to tackle industrial data science problems
  • Build an inclusive community of data scientists
  • Support data science research
• Create innovative curriculum to support the training of the next generation of ethical data scientists
• Partner with nonprofit and civic organizations to seek data-driven solutions to address pressing social, economic and environmental challenges

Founded in 2016 at the University of San Francisco, the Data Institute serves as the umbrella organization for data science education, activities and programming at USF. The Data Institute houses interdisciplinary data science research, the Bachelors and Masters in Data Science degree programs, and continuing education certificates. Through our strong industrial-academic partnerships, we provide unparalleled experiential learning opportunities for our students.

Affiliated Data Institute faculty, researchers and postdoctoral fellows explore the latest theoretical advances and applications in data science. Key research initiatives focus on medical applications of data science, the environment, and ethics.

The Data Institute corporate, nonprofit and civic organizations membership provides engagement opportunities such as faculty consulting, practicum placements, talent acquisition and on-site training. Through our innovative certificate programs, members can provide employee development and education. Membership can also include access to our industrial-academic postdoctoral fellows program.

ABOUT THE UNIVERSITY OF SAN FRANCISCO

The University of San Francisco, the city’s first university, was established by the Jesuits in October 1855. Jesuit tradition defines USF’s approach to learning and our commitment to welcoming students of every faith and no faith. Our vision and mission are the foundations of our university, and reflect the shared views of our institution.

The core mission of the university is to promote learning in the Jesuit Catholic tradition. The university offers undergraduate, graduate, and professional students the knowledge and skills needed to succeed as persons and professionals, and the values and sensitivity necessary to be people for others.

The university distinguishes itself as a diverse, socially responsible learning community of high quality scholarship and academic rigor sustained by a faith that does justice. The university draws from the cultural, intellectual, and economic resources of the San Francisco Bay Area and its location on the Pacific Rim to enrich and strengthen its educational programs.