

Agenda for the NIST Data Science Symposium 2014

National Institute of Standards and Technology
100 Bureau Drive, Gaithersburg, MD 20899
Building 101, Red Auditorium

Day 1 - March 4, 2014

Arrival

8:00	8:45	Check-in	(Administration Building 101)
		Poster-setup	(Hall of Flags)

Welcome & Opening Remarks

		John Garofolo - Senior Advisor for Program Development, IAD (NIST)
8:45	9:00	James St. Pierre - Deputy Director of the Information Technology Laboratory (NIST)

The NIST Data Science Program and Symposium Objectives

9:00	9:15	Ashit Talukder - Chief, Information Access Division (NIST) <i>Overview of the NIST Data Science Program</i>
9:15	9:30	Craig Greenberg - Project Leader, Multimodal Information Group (NIST) <i>Overview of the NIST Data Science Evaluation and Metrology Plans</i>
9:30	9:40	Q&A Discussion

Challenges and Gaps in Data Science Technologies

9:40	9:45	Session Introduction
9:45	10:00	Doug Cutting - Chief Architect (Cloudera) <i>Topic: Scalable Computing</i>
10:00	10:15	Haesun Park - Professor & Executive Director, Center for Data Analytics (Georgia Tech) <i>Data and Visual Analytics for Large Scale High Dimensional Data</i>
10:15	10:30	Eric Frost - Director, Viz Center (San Diego State) <i>Topic: Large-Scale Complex Visualisations</i>
10:30	10:45	Pak Chung Wong - Chief Scientist (Pacific Northwest National Laboratory) <i>Topic: Challenges in Very Large-Scale Visualization and Human-Computer Interaction</i>
10:45	11:00	Q&A Discussion

11:00	11:25	Break and Networking
-------	-------	----------------------

Data and Use Cases for Data Science Research

11:25	11:30	Session Introduction
11:30	11:45	Jeanne Holm - Evangelist (Data.gov) <i>Datasets that Enable Rigorous Data Science Research</i>
11:45	12:00	Stanley Ahalt - Director, Renaissance Computing Institute (University of North Carolina) <i>Research Data Sets and Infrastructure: The Challenge of Taming Research Data for Better Science</i>
12:00	12:15	Volker Markl - Professor and Chair of The Database Systems & Information Management Group (TU Berlin) <i>Data Science Research from the Database Management Systems Perspective</i>
12:15	12:30	Chesley Richards - Deputy Director for Public Health Scientific Services (CDC) <i>Topic: Health Data, Challenges and Opportunities</i>
12:30	12:45	Q&A Discussion

12:45 **John Garofolo** - Working Lunch Logistics

12:45 2:00 *Lunch: NIST cafeteria accepts cash or credit*

Lecture Room B: Focus Group on HCI and Data Science

Lecture Room E: Focus Group on Big Data Analytics for Smart Manufacturing Systems

Portrait Room: Focus Group for NIST Data Science Program Plans

Data Science Poster Session

2:00	3:25	Poster Presentations - Hall of Flags
------	------	--------------------------------------

Data Science Benchmarking and Measurement

3:25	3:30	Session Introduction
3:30	3:45	Milind Bhandarkar - Chief Scientist (Pivotal Software Inc.) <i>Deep Analytics Pipeline: A Benchmark Proposal</i>
3:45	4:00	Shahram Ghandeharizadeh - Associate Professor & Director Database Laboratory (University of Southern California) <i>A Benchmark for Interactive Social Networking Actions</i>
4:00	4:15	Shashi Shekhar - Professor (University of Minnesota) <i>Topic: Spatial Data Mining and Spatial Data Benchmarking</i>
4:15	4:30	Thomas Karl - Director, National Climatic Data Center (NOAA) <i>Climate Archives in NOAA: Challenges and Opportunities</i>
4:30	4:45	Michael Hurley - Technical Staff (MIT Lincoln Laboratory) <i>Information Theoretic Evaluation of Data Processing Systems</i>
4:45	5:00	Q&A Discussion

5:00 **John Garofolo** - Brief wrap-up and adjourn

Agenda for the NIST Data Science Symposium 2014

National Institute of Standards and Technology
100 Bureau Drive, Gaithersburg, MD 20899
Building 101, Red Auditorium

Day 2 - March 5, 2014

Arrival

8:00	9:00	Check-in	(Administration Building 101)
------	------	----------	-------------------------------

Keynote Address

		Charles Romine - Director of the Information Technology Laboratory (NIST) <i>Welcome and Introductions</i>
9:00	9:30	Dr. Francine Berman - Chair, Research Data Alliance / U.S. Edward P. Hamilton Distinguished Professor in Computer Science Director, Center for a Digital Society, Rensselaer Polytechnic Institute
9:30	9:45	Q&A Discussion

Breakout Sessions

9:45		Mark Przybocki - Group Leader, Multimodal Information Group (NIST) <i>Breakout Instructions</i>
10:00	11:30	Breakouts <i>Data Science Benchmarking & Performance Measurement</i> Craig Greenberg (Chair) Portrait Room <i>Datasets and Use Cases for Data Science Research</i> Ashit Talukder (Chair) Green Auditorium <i>Challenges and Gaps in Data Science Technologies</i> John Garofolo (Chair) Red Auditorium

11:30	1:00	<i>Lunch and Focused Topic Working Lunches</i> Lecture Room E: Focus Group on Automated Metadata and Ontologies for Heterogeneous Data
-------	------	--

Government Data Science R&D Panel

1:00	2:45	Christopher White - Program Manager, XDATA Program (DARPA) Jill Crisman - Program Manager (IARPA) Xiaoming Huo - Program Director (NSF) Lucy Nowell - Program Manager, Advanced Scientific Computing Research (DOE) Mike Huerta - Associate Director, US National Library of Medicine (NIH) Stephen Dennis - Innovation Director (HSARPA) Kenneth Rice - Chief, ISR Integration Division (NGA)
------	------	---

2:45	3:00	Q&A Discussion
------	------	----------------

3:00	3:15	<i>Break</i>
------	------	--------------

Breakout Reports

3:15	3:25	Data Science Benchmarking & Performance Measurement
3:25	3:35	Datasets and Use Cases for Data Science Research
3:35	3:45	Challenges and Gaps in Data Science Technologies
3:45	3:55	Poster Session
3:55	4:05	Focus Group: Automated Metadata and Ontologies for Heterogeneous Data (Anne Plant)
4:05	4:15	Focus Group: HCI and Data Science (Sharon Laskowski)
4:15	4:25	Focus Group: Big Data Analytics for Smart Manufacturing Systems (Sudarsan Rachuri)

4:25	4:45	Discussion & Next Steps; Facilitator: Ashit Talukder
------	------	---