bigdata@CSAIL
MIT BIG DATA INITIATIVE

MIT BIG DATA INITIATIVE at CSAIL Annual Member Meeting

November 12-13, 2014

LOCATION: MIT Media Lab
Bldg E14 6th Floor event space, multi-purpose room
MIT Cambridge, Massachusetts

AGENDA
DAY 1: Wednesday Nov 12 2014
8:30am Registration and Breakfast

9:00am Introduction, MIT Big Data Initiative at CSAIL
Sam Madden and Elizabeth Bruce, MIT Big Data Initiative at CSAIL co-Directors

9:15-9:45 **KEYNOTE: "BIG DATA AND SYSTEMIC RISK"**
MUNTHER DAHLEH, Acting Director of Engineering Systems Division and Professor EECS, MIT

9:45-10:45am **BIG DATA APPLICATIONS**

"Making Roads Safer by Making Drivers Better"
Prof. Hari Balakrishnan, Professor of Computer Science, MIT CSAIL and CTO, Cambridge Mobile Telematics

"Bayesian Regression and Bitcoin"
Prof. Devavrat Shah, Associate Professor, MIT CSAIL, Laboratory for Information and Decision Systems

"Places Database: Large Scale Visual Scene Recognition"
Prof. Antonio Torralba, Associate Professor, MIT CSAIL, Graphics and Vision

10:45-11:15am BREAK

11:15-11:45am **KEYNOTE: "DEEP LEARNING: OVERVIEW & TRENDS"**
ANDREW NG, Chief Scientist of Baidu and Associate Professor of Computers Science at Stanford University

11:45-12:30pm **BIG DATA PREDICTIVE ANALYTICS & MACHINE LEARNING**

"Towards a Recommender System That Suggests Models and Parameters for Data"
Kalyan Veeramachaneni, Principal Research Scientist, MIT CSAIL

"Unfolding Psychological State: Mortality Modelling in Intensive Care Units"
Marzyeh Ghassemi, Ph.D Student, MIT CSAIL, Clinical Decision Making Group

"Methods and Models for Interpretable Classification"
Berk Ustun, Ph.D. Student in EECS, MIT, Prediction Analysis Lab

12:30-1:30 LUNCH

1:30-2pm **"STATISTICAL LEARNING AND BIG DATA"**
Prof. Josh Tenanbaum, Computational Cognitive Science, MIT

What are the limits of statistical learning in the context of big data? What is state of the art in "deep learning" today and how do we assess its value and its limitations?

2:00-3:30pm **BIG DATA AND PRIVACY**
Session Chairs: Elizabeth Bruce (MIT), Mona Vernon (Thomson Reuters) and Danny Weitzner (MIT) -- Working Group Chairs

This Big Data Privacy Session will review recent work by the Big Data Privacy Working Group in developing use case scenarios that illustrate the unique challenges of managing privacy when it comes to big data and mapping emerging technology approaches. Panelists will discuss specific scenarios and key challenges, as well as promising technology solutions.

Introduction and Big Data Privacy Working Group - Elizabeth Bruce, MIT

Big Data Privacy & Policy - Danny Weitzner, MIT

Una-May O'Reilly, Principal Research Scientist, MIT CSAIL
ALFA Research Group

Cameron Kerry, Visiting Scholar, MIT
MIT Media Lab

Lalana Kagal, Principal Research Scientist, MIT CSAIL
Decentralized Information Group (DIG)

Salil Vadhan, Vicky Joseph Professor, CS + Applied Mathematics,
SEAS, Harvard

Panel Discussion moderated by Danny Weitzner, Director of MIT CSAIL Decentralized Information Group

3:30-5:30pm **BIG DATA COOL TOOLS POSTERS, DEMOS AND RECEPTION**
LOCATION: MIT Media Lab
Bldg E14 6th Floor, Multi-purpose room + Winter Garden room, MIT, Cambridge, Massachusetts

5:30pm Meeting adjourns for day

6:00pm Dinner at local restaurants (members only) – confirm RSVP at registration desk

DAY 2: Thursday Nov 13
LOCATION: MIT Media Lab
Bldg E14 6th Floor, Multi-purpose room, MIT, Cambridge, Massachusetts

8:00-9:00am Big Data Advisory Board Meeting (advisory board members only)
Location: Media Lab (E14-493)

8:30-9:30am Breakfast / Networking

9:30-10:30am INDUSTRY PANEL: "BIG DATA IN PRACTICE: WHAT'S HARD? "

The purpose of this session is to hear from industry members on what is really challenging when it comes to piloting a big data project. What are you experiencing as the limits of the systems and tools being used today – from integration and cleaning to visualization and analytics. Sam Madden will serve as a moderator and facilitate dialog to provide discussion around potential solutions and approaches from a research perspective.

John Cardante, Big Data and Analytics Technologist, EMC
Surajit Chaundhari, Distinguished Scientist, Microsoft Research
Anu Krishnan, Manager of IT Innovation Centre for Excellence, Shell

Moderated Discussion – Sam Madden (MIT) with Mike Stonebraker (MIT) and Matei Zaharia (MIT)

10:30-11:30am SYSTEMS AND TOOLS FOR BIG DATA

"MIT Research in Data Integration"
Prof. Mike Stonebraker, Adjunct Professor, MIT CSAIL

"Designing a Composable Toolbox for Data Analysis in Spark"
Prof. Matei Zaharia, Assistant Professor, MIT CSAIL

"The MIT Datahub Project"
Prof. Sam Madden, Professor MIT CSAIL

11:30-12pm "CONNECTION SCIENCE"
Prof. Alex (Sandy) Pentland, Toshiba Professor of Media Art and Sciences

Big data allows us to begin creating accurate mathematical models of human behavior. This new capability allows us to design transportation, energy, health, finance, and organizational systems to be more productive and innovative than ever before.

12:00-12:45pm LUNCH

1:00-3:00pm BIG DATA + FINANCE

Session Chairs: Prof. John Guttag and Prof. Andrew Lo

The Big Data Finance session will highlight examples of current research directions through an interactive panel discussion with MIT faculty. Audience participants are encouraged to share their thoughts about the applications of Big Data in the financial services sector in areas such as: consumer finance, credit and privacy; policy and compliance; and emerging digital currencies.

Welcome: Daniela Rus, Professor of EECS and Director of the Computer Science and Artificial Intelligence Lab (CSAIL)

Introduction: Andrew Lo, Charles E. and Susan T. Harris Professor of Finance at the MIT Sloan School of Management, Director of MIT's Laboratory for Financial Engineering, and CSAIL PI

Panel:

Nickolai Zeldovich, Associate Professor, MIT CSAIL

Andrew Lo, Professor of Finance, MIT Sloan School of Management

Vinod Vaikuntanathan, Assistant Professor, CSAIL

Antoinette Schoar, Professor of Finance, MIT Sloan School of Management

Roberto Rigobon, Professor of Applied Economics, MIT Sloan School of Management

Moderated Discussion: Andrew Lo

3:00-3:30pm **REFRESHMENTS & NETWORKING**

3:30pm Meeting Adjourns