The National Academies of
SCIENCES • ENGINEERING • MEDICINE

AGENDA

Challenges in Machine Generation of Analytic Products from Multi-source Data: A Workshop

August 9-10, 2017
The Keck Center of the National Academies of Science, Engineering, and Medicine
500 5th Street, NW – Keck 103
Washington, D.C. 20001

OBJECTIVES

Research Challenges
1. Machine-based methods for generating analytic products
2. Machine-based methods for automating the evaluation of analytic products

Research Questions:
1. What are the technical objectives and metrics needed for success?
2. What are the primary issues?
3. What are the current and "next level" key performance metrics?
4. What is the "level after next" of expected research and development performance?
5. What is the research knowledge base?
6. How can the Government best prepare the scientific workforce to enhance discovery in this area?
7. **What are the requisite enabling technologies?**

DAY 1: WEDNESDAY, AUGUST 9, 2017

7:30 A.M. Registration and Breakfast (on your own)

SESSION 1 — Plenary

8:00 A.M. **Sponsor Remarks and Expectations of the Workshop**
- Dr. David M. Isaacson, ODNI

8:15 A.M. **Generation of Capability Technology Matrix**
- Dr. Rama Chellappa, UMCP, Planning Committee Chair
- Dr. George Coyle, RSO, AFSB/ICSB

8:30 A.M. **Progress in Machine Learning**
- Dr. Tom Dietterich, Oregon State University

9:05 A.M. **Industry Perspective**
- Dr. Josyula R Rao, Watson IBM Fellow

9:45 A.M. **Operational Perspective – Project MAVEN**
- Dr. Travis W Axtell, OSD OUSD (I)

10:25 A.M. BREAK

SESSION 2 — Machine Learning from Image/Video/Map data

10:45 A.M. **Learning from Overhead Imagery**
- Dr. Joe Mundy, Vision Systems, Inc.

11:20 A.M. **Deep learning for Learning from Images and Videos: Is It Real?**
- Dr. Rama Chellappa, UMCP

11:55 A.M. **Learning about Human Activities from Images and Videos**
- Dr. Anthony Hoogs, Kitware, Inc.

12:30 P.M. LUNCH

SESSION 3 — Machine Learning from Natural Languages (ML-NLP)

1:15 P.M. **Machine Learning from Text: Applications**
- Dr. Kathy McKeown, Columbia University

1:50 P.M. **Deep Learning for NLP**
- Dr. Dragomir Radev, Yale University

2:25 P.M. **Machine Learning from Conversational Speech**
- Dr. Amanda Stent, Bloomberg

3:00 P.M. BREAK

SESSION 4 — Learning from Multi-Source Data

3:15 P.M. **Situational Awareness from Multiple Unstructured Sources**
- Dr. Boyan Onyshkevych, DARPA

3:50 P.M. **Discussion on Preparing the Capability Matrix**
- Compile enabling technologies from 1st Day

5:30 P.M ADJOURN

DAY 2: THURSDAY, AUGUST 10, 2017

7:30 A.M. Breakfast in Cafeteria (on your own)

8:00 A.M. **Sponsor Remarks**

- Dr. David Honey, Director of Science & Technology, ODNI

SESSION 5— Learning from Noisy, Adversarial Inputs

8:15 A.M. **Harnessing Machine Learning for Global Discovery at Scale**

- Dr. Mikel Rodriguez, MITRE

SESSION 6 — Learning from Social Media

8:50 A.M. **Presentation**

- Dr. Rob Fergus, NYU

9:25 A.M. **What can we learn from Social Media posts?**

- Dr. Benjamin Van Durme, Johns Hopkins University

10:00 A.M. BREAK

SESSION 7 — Humans and Machines Working Together with Big Data

10:15 A.M. **Sensemaking Systems and Models**

- Dr. Peter Pirolli, Institute for Human and Machine Cognition

10:50 A.M. **Crowd Sourcing for Natural Language Processing**

- Dr. Chris Callison-Burch, University of Pennsylvania

SESSION 8 — Use of Machine Learning for Privacy Ethics

11:25 A.M. **Toward Socio-Cultural Machine Learning**

- Dr. Mark Riedl, George Institute of Technology

12:00 P.M. LUNCH

1:00 P.M. **SESSION 9 - Panel on Evaluation of machine-generated products**

- Dr. Anthony Hoogs, Kitware
- Dr. Jason Duncan, MITRE
- Mr. Jonathan Fiscus, NIST
- Dr. Rob Fergus, NYU

2:00 P.M. BREAK

SESSION 10 — Capability Technology Matrix Panel: NSF, DoD, NIST, DoE

2:10 P.M. **Machine Learning for Energy Applications**
- Dr. Devanand Shenoy, DOE

2:30 P.M. **Using Metrology to Improve Access to 'Unstructured' Data**
- Dr. Ellen Voorhees, NIST

2:50 P.M. **Challenge Problems for Multi-Source Insights**
- Dr. Travis W Axtell, OSD OUSD (I)

3:10 P.M. **An Overview of NSF Research in Data Analytics**
- Mr. James Donlon, NSF

3:30 P.M. **Discussion on Preparing the Capability Matrix**
- Compile enabling technologies from 2ND Day
- Complete Matrix

5:00 P.M. ADJOURN