HPC+ BD + ML
-> Discovery and Innovation

Jim Kurose
Assistant Director, NSF
Directorate for Computer & Information Science & Engineering

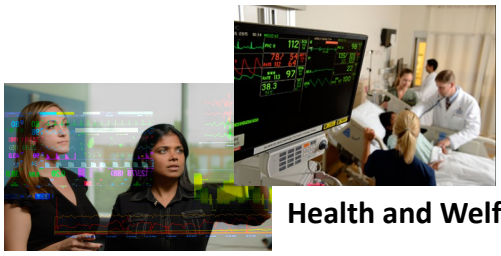
Workshop on the Convergence of High Performance Computing, Big Data, and Machine Learning

October 29-30, 2018

Why are we here?

HPC+ BD + ML ->

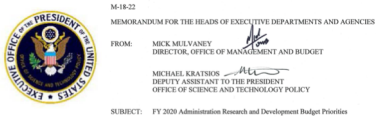
Science discovery



Health and Welfare

National Defense, National Security

Aligned with Administration Priorities



M-18-22

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: MICK MULVANEY
DIRECTOR, OFFICE OF MANAGEMENT AND BUDGET

MICHAEL KRATSIOS
DEPUTY ASSISTANT TO THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY

SUBJECT: FY 2020 Administration Research and Development Budget Priorities

FY 2020 R&D Budget Priorities Memo

"Agencies should invest in fundamental and applied AI research, including machine learning, autonomous systems, and applications at the human-technology frontier. … Agencies should prioritize investment in research and infrastructure to maintain U.S. leadership in strategic computing, from edge devices to high-performance computing, … use of embedded sensors, data analytics, and machine learning "

National Security Strategy

"prioritize emerging technologies critical to economic growth and security, such as data science, encryption, autonomous technologies,… advanced computing technologies, and artificial intelligence. "

National Defense Strategy

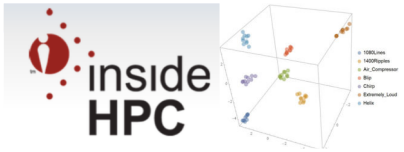
"New technologies include advanced computing, "big data" analytics, artificial intelligence, autonomy, robotics, .."

HPC + BD + ML

ORNL Launches Summit Supercomputer

New 200–Petaflops System Debuts as America's Top Supercomputer for Science

"In addition to scientific modeling and simulation, Summit offers unparalleled opportunities for the integration of AI and scientific discovery, enabling researchers to apply techniques like machine learning and deep learning to problems in human health, high-energy physics, materials discovery …"



Using Ai to detect Gravitational Waves with the Blue Waters Supercomputer

April 14, 2018 by staff Leave a Comment

ARTIFICIAL INTELLIGENCE AND SUPERCOMPUTERS TO HELP ALLEVIATE URBAN TRAFFIC PROBLEMS

Researchers from UT, TACC and City of Austin develop AI tools to automatically analyze road behavior and create searchable databases

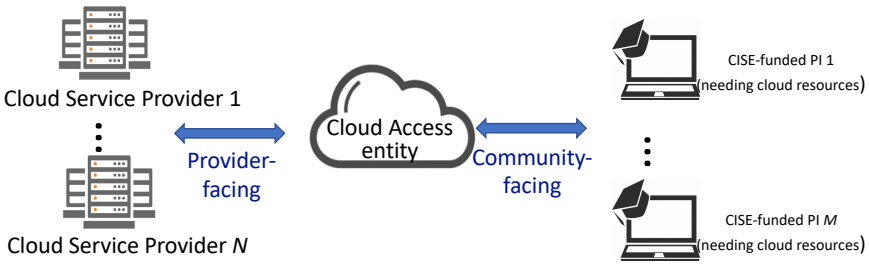
HPC + BD + ML

Enabling Access to Cloud Computing Resources for CISE Research and Education (Cloud Access)

PROGRAM SOLICITATION
NSF 19-510

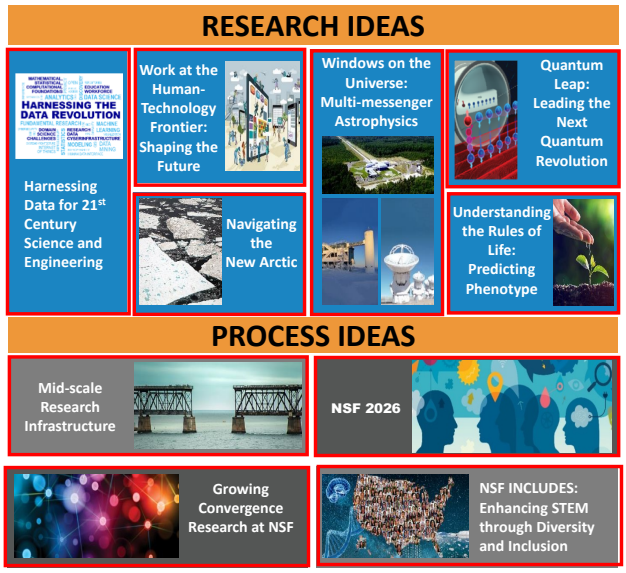
National Science Foundation

Directorate for Computer & Information Science & Engineering
Division of Computing and Communication Foundations
Division of Information & Intelligent Systems
Division of Computer and Network Systems
Office of Advanced Cyberinfrastructure



… building on BIGDATA NSF/AWS/Google/IBM/Microsfot cloud collaboration

HPC + BD + ML



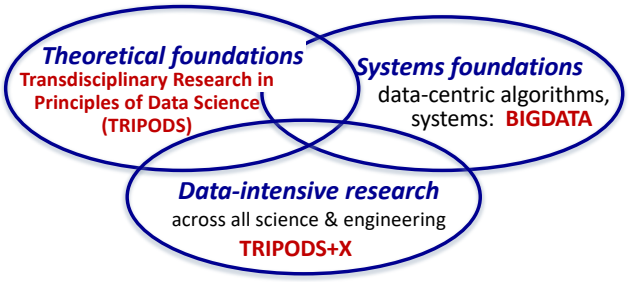
" … bold questions that will drive NSF's long-term research agenda -- questions that will ensure future generations continue to reap the benefits of fundamental S&E research. "

"AI is the universal connector that interweaves all of our Big Ideas; data science is changing the very nature of scientific inquiry, and AI's use of data has the potential to revolutionize everything we do in science."

F. Cordova, Director, NSF, Sept. 2017

HPC + BD + ML : NSF Harnessing The Data Revolution

Research across all NSF Directorates



Educational pathways

Innovations grounded in an education-research-based framework

- **NASEM: undergraduate data science**
- **NSF Research Traineeship**
- **NSF Graduate Research Fellowship Program**
- **Data Science Corps**

Advanced cyberinfrastructure

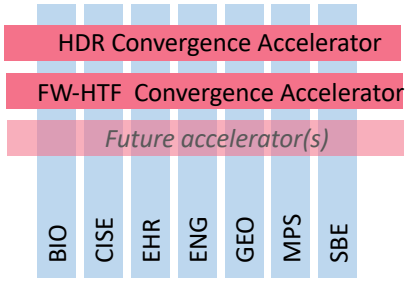
Accelerating data-intensive research.

Cyberinfrastructure for Sustained Scientific Innovation (CSSI);
Scalable data-driven Cyberinfrastructure Dear Colleague Letter (DCL);
Midscale infrastructure (RFI and DCL)

HPC + BD + ML : NSF Harnessing The Data Revolution

HDR Convergence Accelerator

- Translational, use-inspired *convergence* research in HDR
- Projects with clear goals, milestones, directed deliverables (e.g., 6-months)
- More intentional, directed management
- Mission-driven evaluation, rather than peer review
- Partnerships: co-funding, co-design, creation
- FY 2019 launch



HPC + BD + ML : Open Knowledge Network

Goal: design, develop, prototype, and demonstrate an *open knowledge network* – an open semantic information infrastructure based on open standards for creating and maintaining a knowledge graph to enable discovery of non-trivial knowledge from multiple disparate knowledge sources, covering thousands of topic areas, especially scientific information

NITRD Workshops on an Open Knowledge Network: https://www.nitrd.gov/nitrdgroups/index.php?title=Open_Knowledge_Network

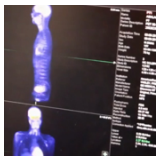
HPC + BD + ML : Artificial Intelligence & Data

Data (big, and small) often sensed in real-time

IoT

Scientific Instruments

Health



Training, baseline data

- Autonomy
- Human-AI interaction
- AI Infrastructure
- Modeling
- Machine Learning Components
- Massive Data Management
- Sensing / Data Acquisition

HPC + BD + ML : Artificial Intelligence & Data

Distributed Data, Privacy, Security, Data-Sharing

Network, distributed data

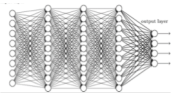
Data discovery, access

Security, Privacy

- Autonomy
- Human-AI interaction
- AI Infrastructure
- Modeling
- Machine Learning Components
- Massive Data Management
- Sensing / Data Acquisition

HPC + BD + ML : Artificial Intelligence & ML

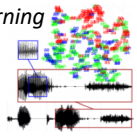
Deep learning, classification



Reinforcement learning

Transfer learning

Representation learning



Semi-supervised learning

Hypothesis generation

Safety, predictability

Fairness, Accountability Transparency, Causality

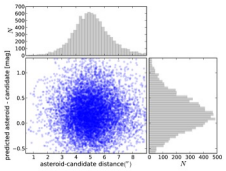
- Autonomy
- Human-AI interaction
- AI Infrastructure
- Modeling
- Machine Learning Components
- Massive Data Management
- Sensing / Data Acquisition

HPC + BD + ML : Artificial Intelligence & Modeling

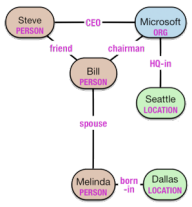
Machine learning, with domain-specific models

Human emotional state

Domain models: classifying "interesting" astronomical events

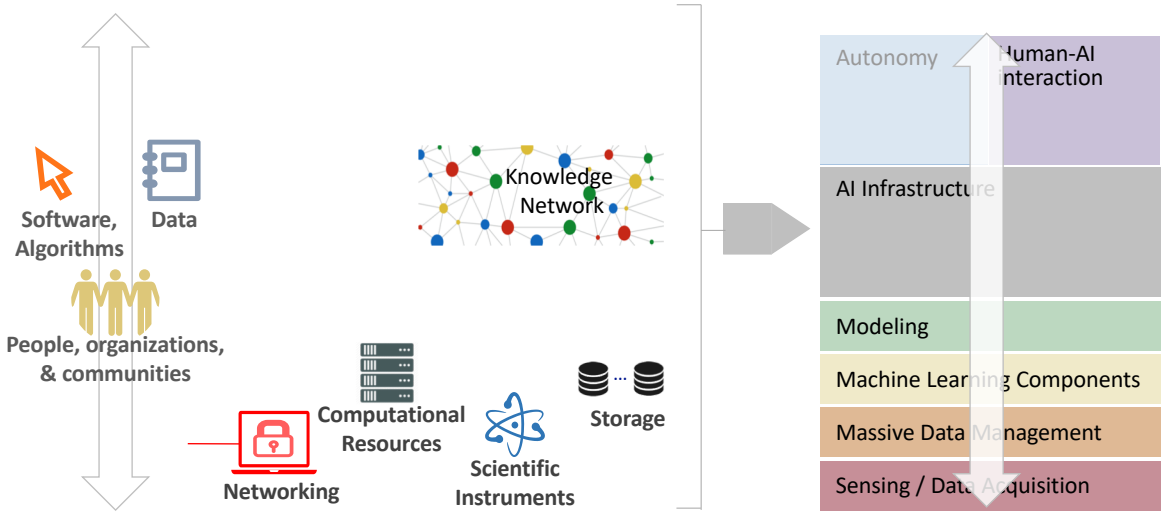


Knowledge Representation

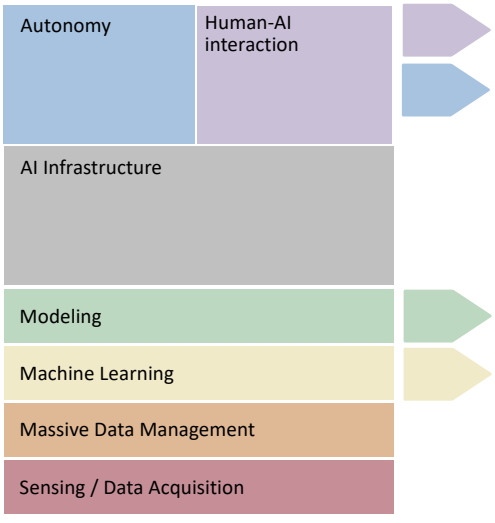


- Autonomy
- Human-AI interaction
- AI Infrastructure
- Modeling
- Machine Learning Components
- Massive Data Management
- Sensing / Data Acquisition

HPC + BD + ML : Artificial Intelligence & Infrastructure



HPC + BD + ML : Artificial Intelligence @ NSF/CISE



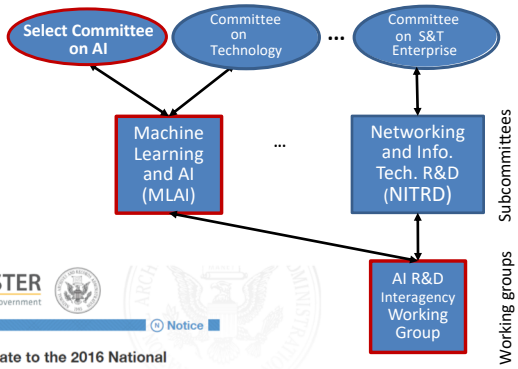
- CISE core research programs:
  - Cyber-human Systems
  - Robust Intelligence
  - Information Integration and Informatics
- Cross-directorate programs:
  - BIGDATA
  - NRI-2.0: Ubiquitous Collaborative Robots
  - Smart and Autonomous Systems
  - Smart & Connected Communities
  - Smart and Connected Health
  - Computational Neuroscience
- CISE Expeditions in Computing
- *AI+X:* ML as a new horizontal
- CISE/IIS budget: $210M

Cross-Government AI Activities

Office of Science & Technology Policy (OSTP)

Lynne Parker
Assistant Director for AI

National Science and Technology Council (NSTC)



Cross-Government HPC+BD+ML: Where?

NITRD

High End Computing IWG

Big Data IWG

AI IWG

HPC+ BD + ML
-> Discovery and Innovation

Cross-Government HPC+BD+ML: *Here!*

- Suggestions for paths forward: *actionable* items
  - technical directions
  - achieving integration (HPC+BD+ML)
  - partnerships (federal, industry, academia)
- Believe it or not: high-quality workshop reports *do matter!*
  - For impact to last beyond this meeting
  - Community voice into agency priorities, activities

High End Computing IWG | Big Data IWG | AI IWG

NITRD Staff: Wendy Wigen, Ji Lee, Faisal D'Sousa

NITRD IWG leadership: Jeff Alstott, Chaitan Baru, Rance Cleaveland, Susan Gregurik, Henry Kautz, Sandy Landsberg, Barry Schneider

"Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program."

The Networking and Information Technology Research and Development (NITRD) Program

Mailing Address: NCO/NITRD, 2415 Eisenhower Avenue, Alexandria, VA 22314

Physical Address: 490 L'Enfant Plaza SW, Suite 8001, Washington, DC 20024, USA Tel: 202-459-9674, Fax: 202-459-9673, Email: nco@nitrd.gov, Website: https://www.nitrd.gov