AI & Data
Strategic Vision Team Update

Heather Evans
NIST Program Coordination Office

The team

Heather Evans (PCO)

Kirk Dohne (EL)

Mike Garris (ITL)

Will Guthrie (ITL)

Bob Hanisch (MML)

Craig Schlenoff (EL)

Jim Warren (MML)

Adam Wunderlich (CTL)

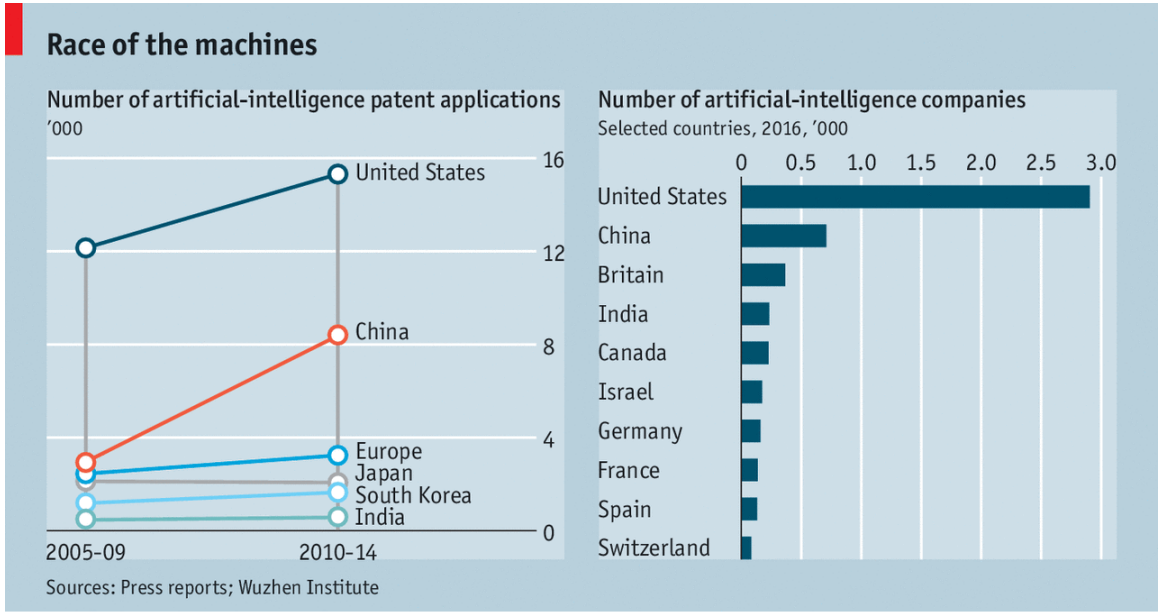
Brian Zimmerman (PML)

Why now?

"The current wave of progress and enthusiasm for AI began around 2010, driven by three factors that built upon each other: the **availability of big data** from sources including e-commerce, businesses, social media, science, and government; which provided raw material for **dramatically improved machine learning** approaches and algorithms; which in turn relied on the capabilities of **more powerful computers**."

White House Report,

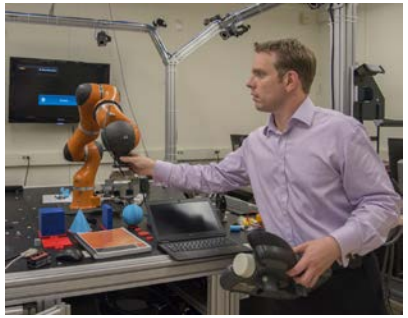
Preparing for the Future of AI



AI&D Overview & Rationale

- AI and Data are essential to the future of research and scientific discovery, industrial competitiveness, and social progress
- NIST has expertise and resources in place that serve as a strong foundation, but are subcritical
- Two roles: external and internal

MDCS
User Community



Vision and Strategies

NIST's investigation and deployment of data and AI technologies builds confidence and trust that drives new measurement research outcomes and an expanded commercial marketplace.

Three key strategies:

1. Expertise
2. Data and Analysis
3. Infrastructure (Networking, Data Storage, and Computational)

Expertise

NIST must grow its expertise in **data engineering** (e.g., data and network infrastructure, database engineering), **data science** (e.g., analytics, informatics, management), **AI/machine learning (ML) methods**, and matching of the appropriate analysis methods to different research domains.

Specifically:

- Increase our in-house expertise with additional specialists
- Partner with academic, government, and industry leaders to rapidly gain knowledge in these areas, for example postdoctoral researchers and exchange programs
- Increase cross-OU collaboration and centralized resources to rapidly build scale and success stories

Data and Analysis

A culture change within NIST is essential to make data usable across space and time, from acquisition to storage, access, and use of data (observed or simulated). NIST staff should be able to routinely draw on AI/ML tools to gain insight into their research.

Specifically:

- Train staff and provide them with the tools to practice high quality data management, for example electronic laboratory notebooks, metadata capture, databases, etc.
- Make tools available for both *specific* types of analysis (e.g. clustering or classification) and the *general*, undirected study of data (e.g. “modern” exploratory data analysis).
- Develop validation tools and datasets for assessing AI systems and to further research in high impact areas, such as advanced communications, health care, and manufacturing

Networking, Data Storage, and Computational Infrastructure

NIST needs a robust infrastructure to support faster data transfer, centralized storage, and shared computational resources for the application of modern tools for data analysis and AI algorithms.

- Network infrastructure does not support the amount of data that needs to be transferred and stored
- Data processing is computationally intensive and not widely accessible to all NIST programs
- Future AI will have computation demands more than an order of magnitude greater than any current in-house capacity

What does success look like?

- A truly interdisciplinary cross-laboratory data sharing culture at NIST
- It is easy to collect, store, and analyze complex datasets
- NIST researchers routinely draw on AI/ML tools to gain insight into their research data and code implementations
- NIST's programs are providing validation tools and datasets for the most in-demand ML and AI application areas
- NIST is seen by industry and academia as a leader providing widely accepted tools/technologies for data and AI

Key Questions

- How can NIST ensure access to talent in this competitive field?
- Which partners are essential for NIST's success in realizing this vision?
- Which research areas are the highest impact for application of ML/AI?
- What best practices in culture change should NIST implement to cultivate data management?
- How do we best integrate advanced data analysis (ML, AI) techniques in our measurement science?