INL – ML & AI Symposium 2.0 – July 9th, 2020

Idaho National Laboratory

- **Purpose of Meeting:**
  - Introduce ML and AI Current Ideas & Collaborations
  - Provide examples of how ML an AI are being applied across other industries
  - Discuss current ML & AI research and capabilities at INL
  - Discuss planned activities, including engagement opportunities and collaboration opportunities

- **Presentations will include:**
  - Provide overview on Topic Area;
  - Describe the status of industry
  - Identify Issues (if any) and potential impact
  - High level discussion of planned activities and outcomes

NS&T ML-AI

Big Data, Machine Learning, Artificial Intelligence

INL – ML & AI Symposium 2.0 – July 9th, 2020

- **Agenda for Machine Learning and Artificial Intelligence Symposium**

Time	Subject	Speaker
11:00	Welcome, Introductions, and Agenda	Curtis Smith
11:05	Overview of DOE Office of AI and Technology Priorities	Margaret Lentz
11:20	Artificial Intelligence: A NIST strategic priority	Elham Tabassi
11:35	Data Driven Decision Making (3DM)	Thiago Seuaciuc-Osorio
11:45	AI for Materials Science	Lars Kotthoff
11:55	AI for Nuclear Core Design	Koroush Shirvan
12:05	Domain-Enriched Deep Architectures and Applications	Min Xian
12:15	The Future with AI: Sci-Fi or Reality	Milos Manic
12:25	Application of Deep Learning on NPP Related Data	Alper Yilmaz
12:35	Explainable ML for Decision Support Systems	Kasun Amarasinghe
12:45	ML for Risk-Based Decision Making, Command and Control	Dan Cole
12:55	More Letters into the “AI” Acronym	Hany Abdelkhalik
1:05	Wrap Up / Next Steps	Curtis Smith

NS&T ML-AI

Curtis Smith

Organization/Role: INL - Division Director for Nuclear Safety and Regulatory Research

Education/Experience: BS, MS, and PhD in Nuclear Engineering at ISU and MIT, 29.7 years at INL

Current ML/AI work: Leading the Risk-Informed Systems Analysis Pathway for LWRS

Title: My Motivation for AI/ML in Science, Math, and Engineering

Overview: A discussion on how AI/ML has advanced in the science, math, and engineering communities and how these advances may be used with INL applications such as computational risk assessment.

These topics provide an insight into the potential for advanced analysis and operations for complex systems.

NS&T ML-AI

Welcome to the AI/ML Symposium 2.0

Dr. Curtis Smith, Director
Nuclear Safety and Regulatory Research Division
Idaho National Laboratory

www.inl.gov

"And I told him, AI and ML aren't the thing.

They're the thing that gets us to the thing."

(See *Halt and Catch Fire*)

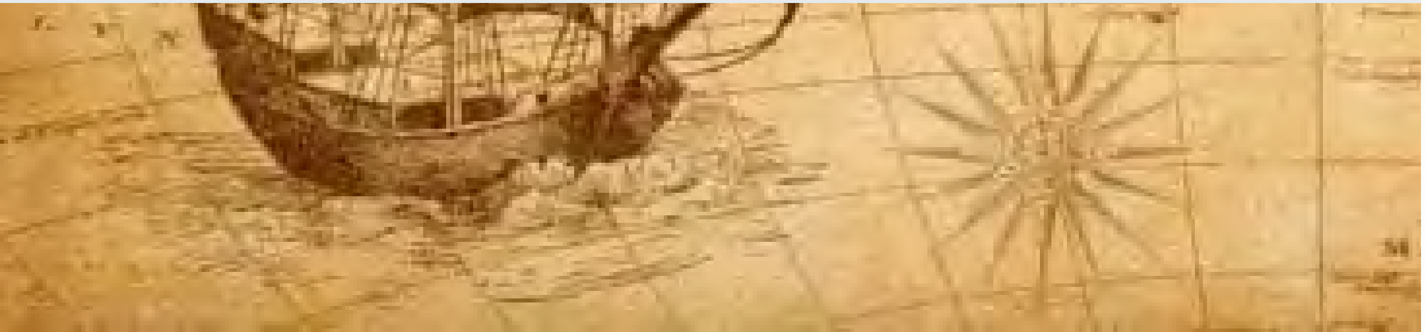
Learning Internal Representations by Error Propagation

DAVID E. RUMELHART, GEOFFREY E. HINTON, and RONALD J. WILLIAMS

THE PROBLEM

We now have a rather good understanding of simple two-layer associative networks in which a set of input patterns arriving at an input layer are mapped directly to a set of output patterns at an output layer. Such networks have no *hidden* units. They involve only *input* and *output* units. In these cases there is no *internal representation*. The coding provided by the external world must suffice. These networks have proved useful in a wide variety of applications (cf. Chapters 2, 17, and 18). Perhaps the essential character of such networks is that they map similar input patterns to similar output patterns. This is what allows these networks to make reasonable generalizations and perform reasonably on patterns that have never before been presented. The similarity of patterns in a PDP system is determined by their overlap. The overlap in such networks is determined outside the learning system itself—by whatever produces the patterns.

The constraint that similar input patterns lead to similar outputs can lead to an inability of the system to learn certain mappings from input to output. Whenever the representation provided by the outside world is such that the similarity structure of the input and output patterns are very different, a network without internal representations (i.e., a network without hidden units) will be unable to perform the necessary mappings. A classic example of this case is the *exclusive-or* (XOR) problem illustrated in Table 1. Here we see that those patterns which overlap least are supposed to generate identical output values. This problem and many others like it cannot be performed by networks without hidden units with which to create



Moving from 1.0 to 2.0

- **Last quarter, INL sponsored a symposium on Artificial Intelligence (AI) and Machine Learning (ML) approaches and activities related to science and engineering**
  - The "1.0 Symposium" focused on internal-to-INL activities and capabilities
  - Eleven speakers discussed a variety of current topics and future applications
  - Over 200 INL staff participated in the symposium
- **For Symposium 2.0 we wanted to have more of an industry vision/overview and platform for university applications and collaborations**
- **The field of AI/ML is evolving, I encourage all to continue our education in these areas, for example**
  - University of Idaho (Min Xian) Deep Learning; Digital Image Processing; Python for ML
  - University of Wyoming (Lars Kotthoff) Advanced Topics in AI
  - North Carolina State University (Xu Wu) Advanced Topics In Nuclear Engineering - Scientific Machine Learning
- **AI/ML will be a key technology moving forward as we continue our R&D**

Curtis.Smith@inl.gov

Thank you and enjoy the symposium!

Margaret Lentz

Organization/Role: Special Advisor to the Artificial Intelligence & Technology Office (AITO) at DOE

Education/Experience: BS’98 Carnegie-Mellon; PhD’02 Purdue University; a scientist with a 20+ yr research history in chemistry, imaging physics, neuroscience, and AI/ML.

Current ML/AI work: Current ML/AI work: Working with DOE’s Programs and national laboratories to coordinate and advance DOE’s strategic goals and priorities in AI.

Title: Overview of DOE’s Artificial Intelligence & Technology Office

Overview: A discussion of DOE’s AITO’s priorities, mission, vision and strategic goals.

www.inl.gov

Idaho National Laboratory

Machine Learning & Artificial Intelligence Symposium
July 9, 2020

Margaret R. Lentz, PhD

Artificial Intelligence & Technology Office, Department of Energy

Artificial Intelligence at DOE

ARTIFICIAL INTELLIGENCE & TECHNOLOGY OFFICE

AITO

UNITED STATES DEPARTMENT OF ENERGY

DEPARTMENT OF ENERGY · UNITED STATES OF AMERICA

Administration’s AI Priorities

AITO was created in support of the Executive Order on Artificial Intelligence and supports the broad goals of accelerating the development of AI capabilities for the Department and maintaining American leadership in AI.

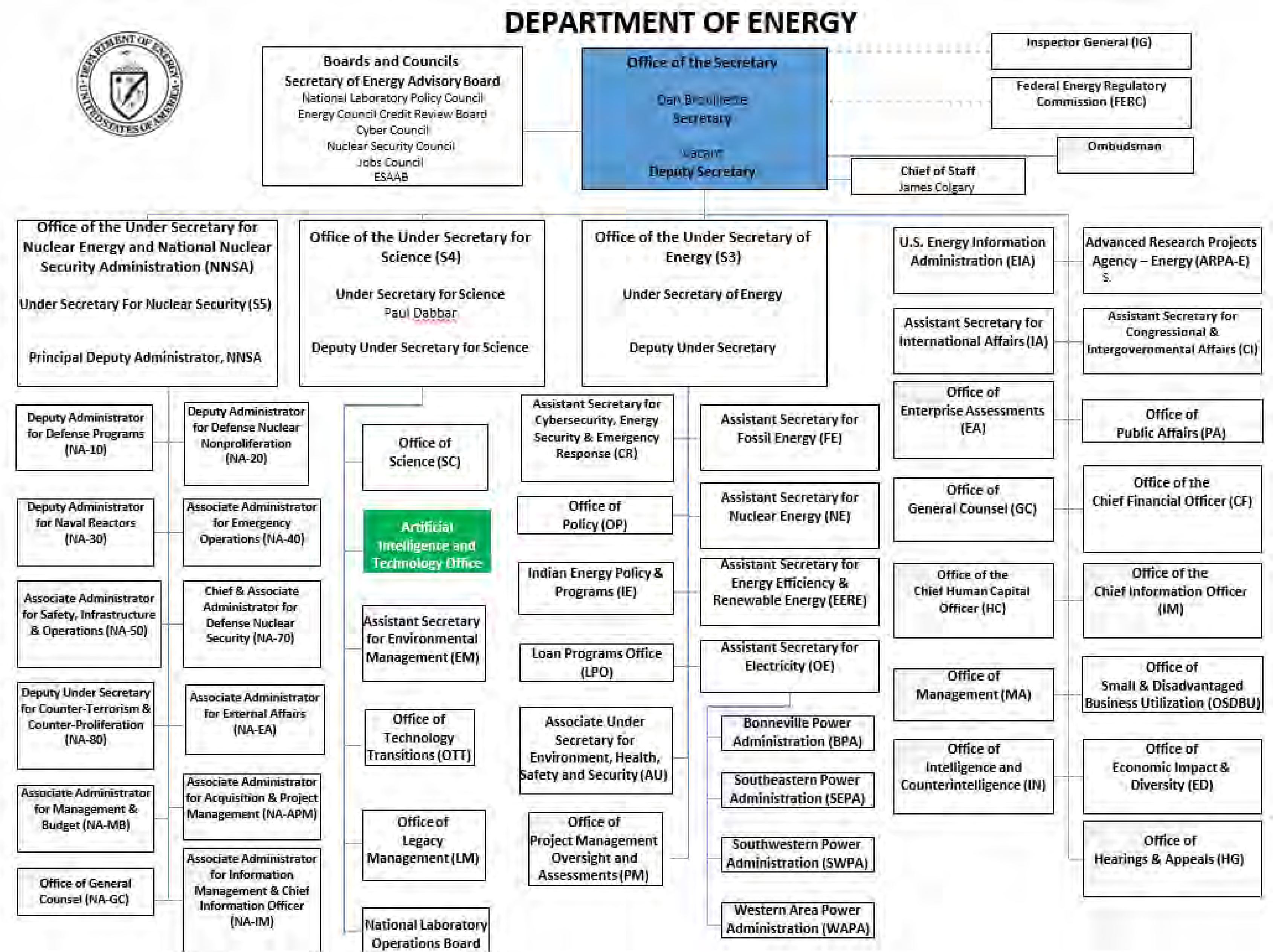

Executive Order on Maintaining American Leadership in Artificial Intelligence

- Using AI for government services
- Removing barriers to AI innovation
- Training the next generation of American workers
- Achieving strategic national security advantage
- Accelerating AI research & development
- Engage international and private sector
- Foster public trust and confidence in AI

AITO Vision & Mission

Vision: Transform the Department of Energy into the United States Government's lead agency in the civilian use of artificial intelligence (AI) by accelerating the research, development , delivery, and application of AI.

Mission: The Artificial Intelligence & Technology Office (AITO), the Department of Energy's center for artificial intelligence, will accelerate the delivery of AI-enabled capabilities, scale the Department-wide development of AI, synchronize AI applications to advance the agency's core missions, and expand public and private sector strategic partnerships, all in support of American AI leadership.



AITO (AI-1) reports to the Under Secretary for Science

Why Does AI Matter for DOE?

AI is a technology that performs tasks which mimic human intelligence

This includes pattern recognition, decision making, visual perception, speech recognition, information processing, behavior adaptation, autonomous control, optimization, etc.

AI is affecting many technologies used DOE wide

Intelligent Sensors, Machine Learning, Data Sciences & Data Analytics, Robotics, Autonomous Systems, Data structure and management, Information and Business Management Systems, Edge Devices/Distributed Systems, Natural Language Processing, Human-machine Interface & Biometrics.

AI has implications for high consequence areas with little room for failure

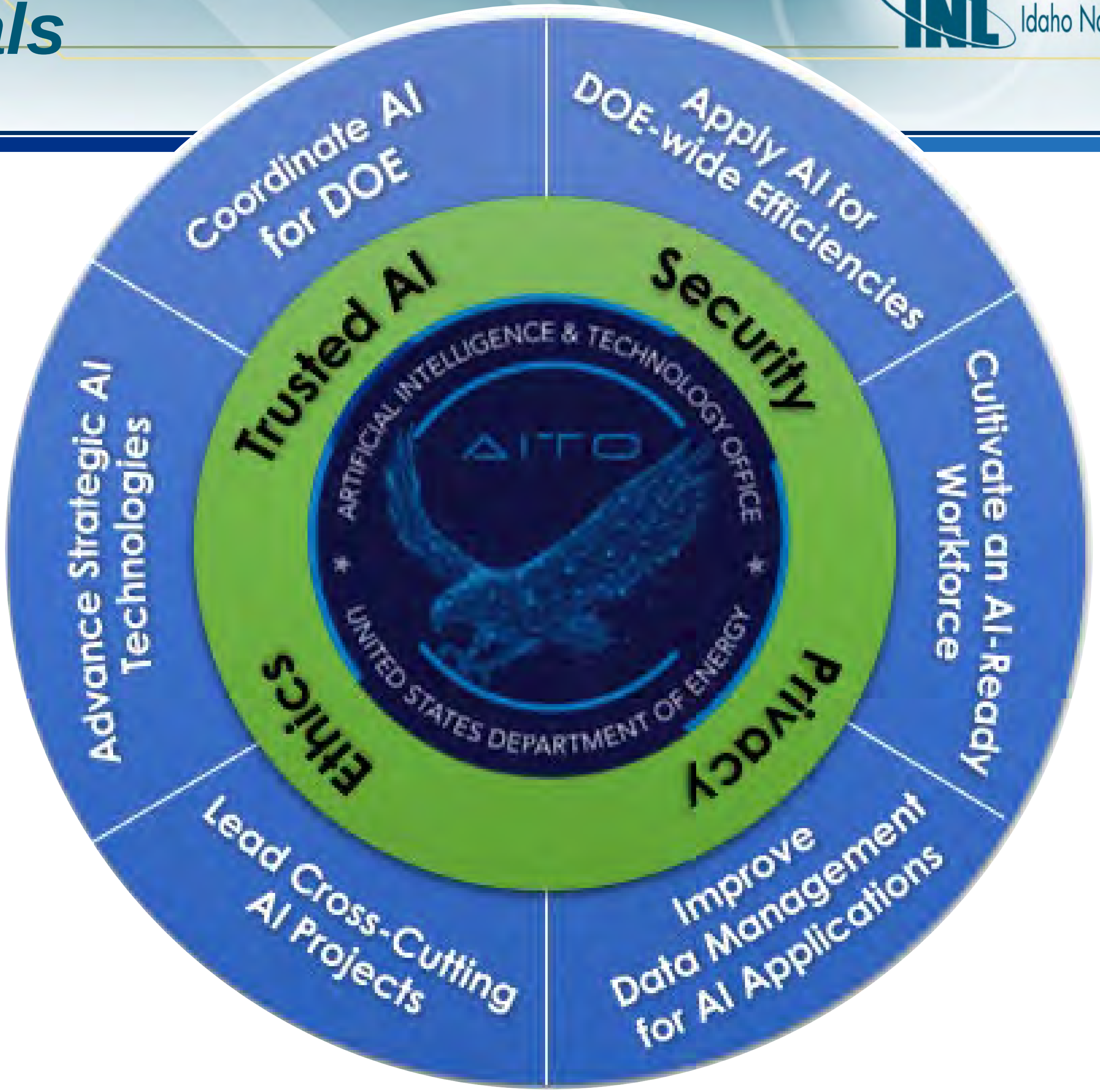
Workforce Development, Energy Security, Cybersecurity, Physical Security, National Security, Economic Security, Science & Highly Hazardous Operations.

There is a need for trustworthy AI that is accurate with high confidence, proven to be unbiased & reliable.

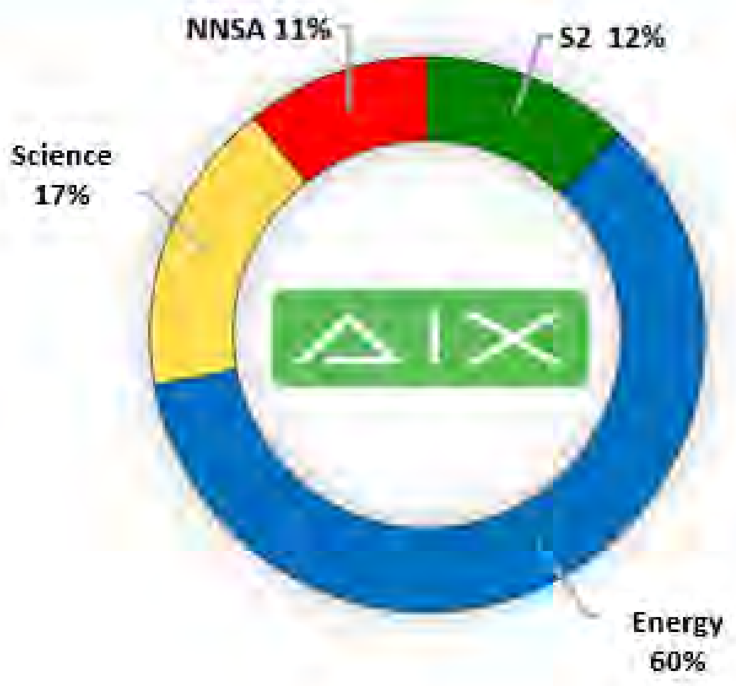
AITO's Strategic Goals

Idaho National Laboratory

To achieve AITO's vision and ensure the creation of trusted AI systems that address core values such as privacy and security, AITO is expected to:



Artificial Intelligence Exchange (AIX)



AIX Content by DOE Organization

AITO created the AIX database to capture DOE AI investments

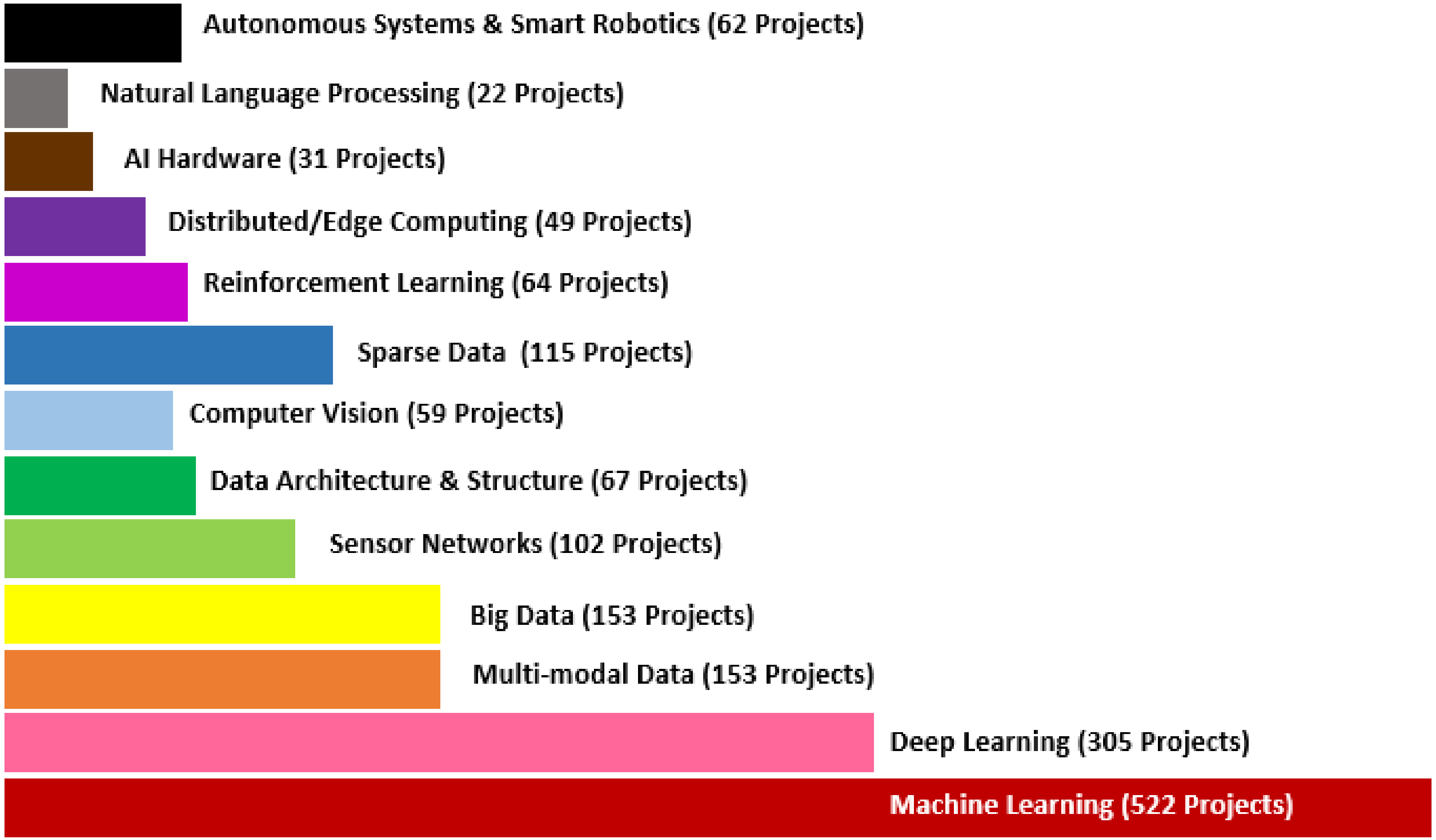
- Allows AITO better coordination across DOE programs, enhances our ability to reduce duplicative efforts, and enables programs to leverage resources for cross-cutting efforts in AI.
- Enables AITO to assess strengths and weaknesses in DOE's AI investments, and identify opportunities in emerging AI technologies.
- Ensures the alignment of projects with the Administration's strategic priorities in AI

AIX By White House Strategic Priorities

Projects in AI Exchange Organized by OSTP/NSTC AI Strategy

Out of 685 Projects Reported.
Projects can be associated with more than one strategy.
Projects reporting is from a data call and are not representative of LDRD or SPP projects within the DOE National Labs.

AIX By Technology Type



Out of 685 projects reported.
Projects can be associated with more than one technology type.
Projects reporting is from a data call and are not representative of LDRD or SPP projects within the DOE National Labs.

AITO Current Activities

- **Working across DOE mission, business and operational functions to identify and surface priority areas for AI**
- **Creating the DOE AI Strategy**
- **Institutionalizing the AI Exchange**
- **Supporting White House and Congressional AI efforts, including:**

  National Strategic Commission on AI (NSCAI), OSTP task force on liberating data for AI, WH strategy on Principles for AI in Government.

- **Exploring AI Leadership Training and AI Workforce Opportunities**

  AITO is engaging federal leads across the department to identify needs and best approach for: hiring, management training, and engaging the future workforce.

- **Workshops & Summits**

For more information on AITO or AI at DOE, please feel free to contact me at margaret.lentz@hq.doe.gov.

AITO

U.S. DEPARTMENT OF ENERGY | Artificial Intelligence and Technology Office

Elham Tabassi

Organization/Role: NIST - Chief of Staff in the Information Technology Laboratory (ITL) at NIST

Education/Experience: Undergraduate degree from Sharif University of Technology, and a Master of Science from Santa Clara University / As a scientist she has been working on various computer vision research projects with applications in biometrics evaluation and standards since 1999

Current ML/AI work: Leads ITL's fundamental and applied research in computer science and engineering, mathematics, and statistics that cultivates trust in information technology and metrology

Title: Artificial Intelligence: A NIST strategic priority

Overview: A discussion of NIST activities, priorities and strategic goals.

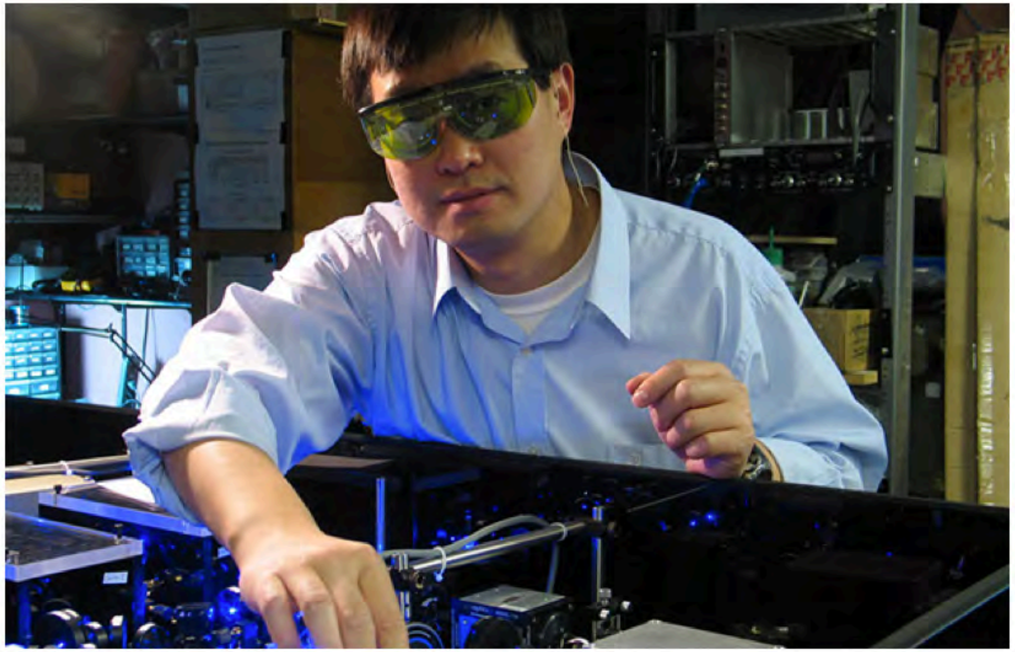
NS&T ML-AI

Artificial Intelligence: A NIST strategic priority

National Institute of Standards and Technology
U.S. Department of Commerce

Information Technology Laboratory

Cultivating Trust in IT and Metrology



NIST National Institute of Standards and Technology
U.S. Department of Commerce

ITL

From innovation to adoption

Fundamental Research

Applied Research

Standards + Best Practice Guides

Adoption

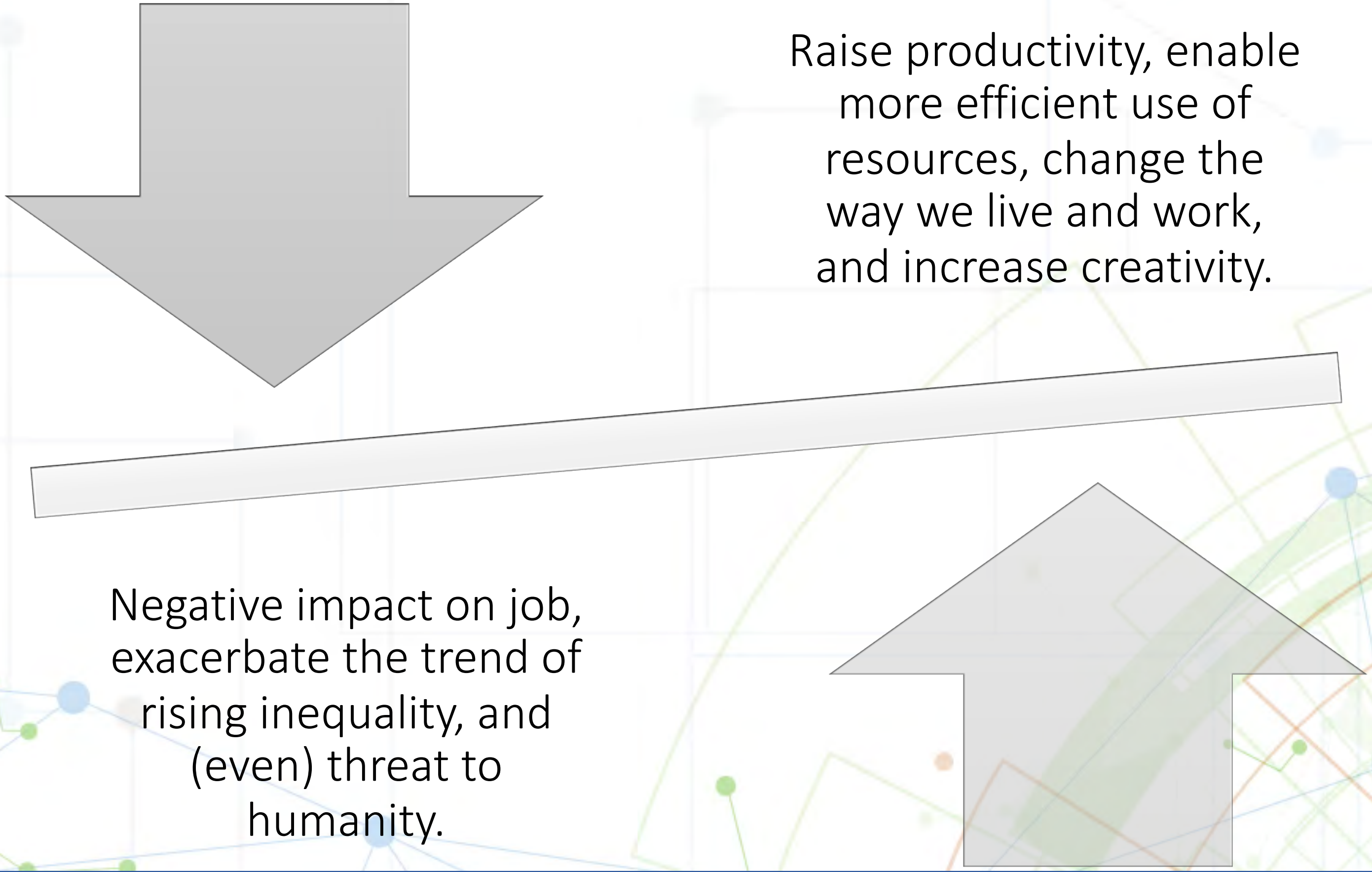
Image Credit: wsj.com

National Institute of Standards and Technology
U.S. Department of Commerce

Trustworthy AI

4

Major advances in artificial intelligence



Technical requirements for trustworthy AI

accurate

secure

robust

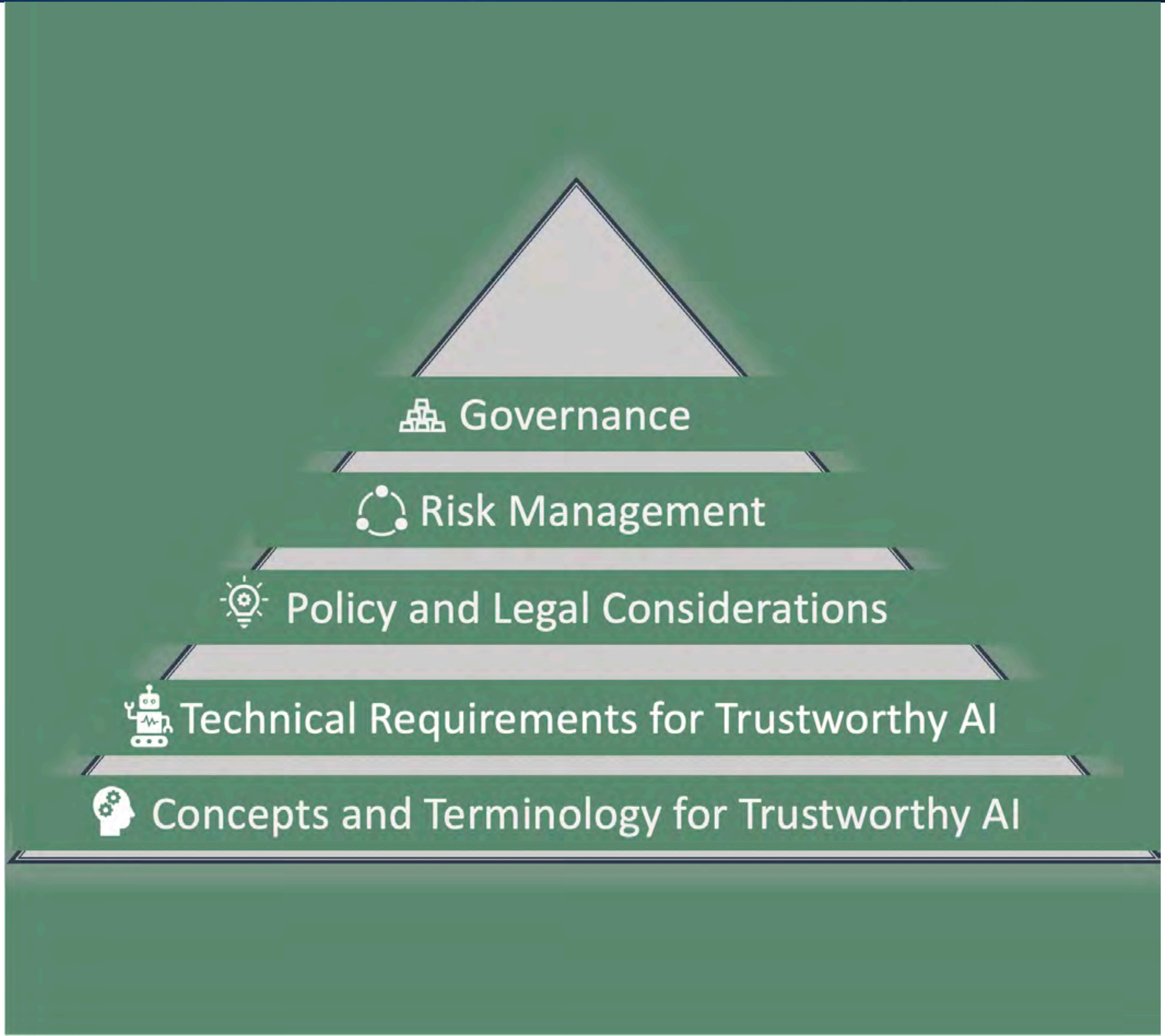
explainable

objective

reliable

and more …

Foundational research for trustworthy AI



Pillars of NIST AI Program

Foundational Research

establish the needed scientific foundation for design, development, and assessment of trustworthy AI

Use-inspired Research

advance AI as a tool to accelerate scientific discoveries and technological innovations

Evaluation

benchmarks to understand the theoretical capabilities and limitations of AI

Standards

tools and guidelines for vocabulary, data, metrics and testbeds for AI

Policy and Engagement

forums and research to engage scientists, engineers, psychologists, and lawyers on issues of trustworthiness

AI happenings in Summer 2020

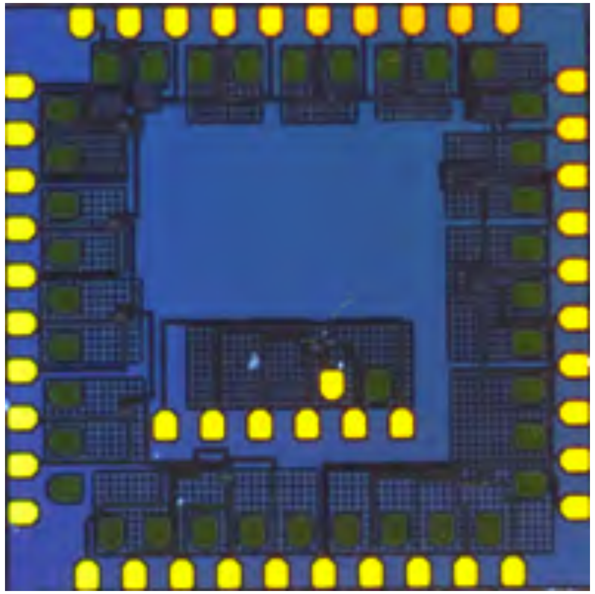
Develop a shared understanding of what constitutes trustworthy AI (e.g., accuracy, security, explainability, reliability, free from bias) and establish the needed scientific foundation for design, development, and assessment of trustworthy AI.

Kickoff webinar on August 6, 2020.

Bias in AI: workshop on August 18, 2020.

Secure AI: Terminology and Taxonomy; 2nd draft for public comment.

Novel computational paradigms for AI



Schneider, M.L., Donnelly, C.A., Haygood, I.W. et al. Synaptic weighting in single flux quantum neuromorphic computing. Sci Rep 10, 934 (2020)

- Establishing metrics and benchmarks for AI hardware.
- Foundational analysis of the computational capacity of a physical system.
- Analysis and development of algorithms for spike-based computation.

National Institute of Standards and Technology
U.S. Department of Commerce

Federal Engagement in Artificial Intelligence Standards

11

EO on Maintaining American Leadership in AI

NIST

EXECUTIVE ORDERS

Executive Order on Maintaining American Leadership in Artificial Intelligence

INFRASTRUCTURE & TECHNOLOGY | Issued on: February 11, 2019

★ ★ ★

www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/

Within 180 days…

Secretary of Commerce, through Director of NIST, shall issue a plan for Federal engagement in the development of technical standards and related tools in support of reliable, robust, and trustworthy systems that use AI technologies.

By the Numbers

97
RFI RESPONSES

>400
WORKSHOP ATTENDEES

6
BREAKOUT SESSIONS

43
PUBLIC COMMENTS

2
DOCUMENTS

10
AUGUST

13

Recommended Actions

Coordination

Bolster AI standards-related knowledge, leadership, and coordination among Federal agencies to maximize effectiveness and efficiency.

Research

Promote focused research to accelerate broader exploration and understanding of how aspects of trustworthiness can be practically incorporated within standards.

Partnership

Support and expand public-private partnerships to develop and use AI standards and related tools to advance trustworthy AI.

Engagement

Strategically engage with international parties to advance AI standards for U.S. economic and national security needs.

Coordination Activities

15

Interagency coordination and leadership

AI Select Committee
Chaired by OSTP, NSF, DARPA

Office of Science and Technology Policy

NSTC

MLAI Subcommittee
Chaired by NIST, OSTP, DoE

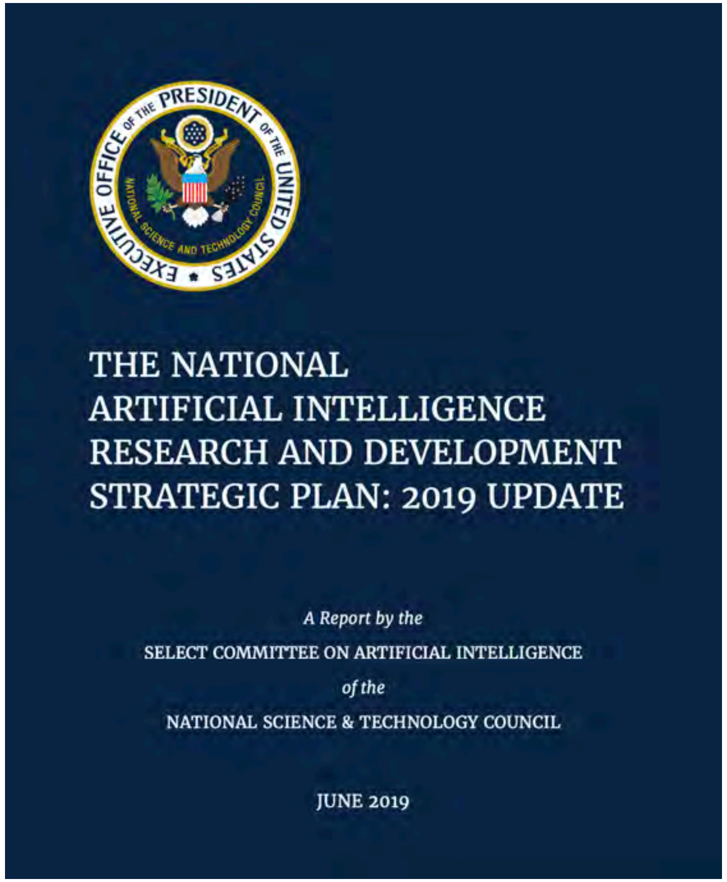
USG AI Standards Coordinator

Networking and Information Technology R+D (NITRD)

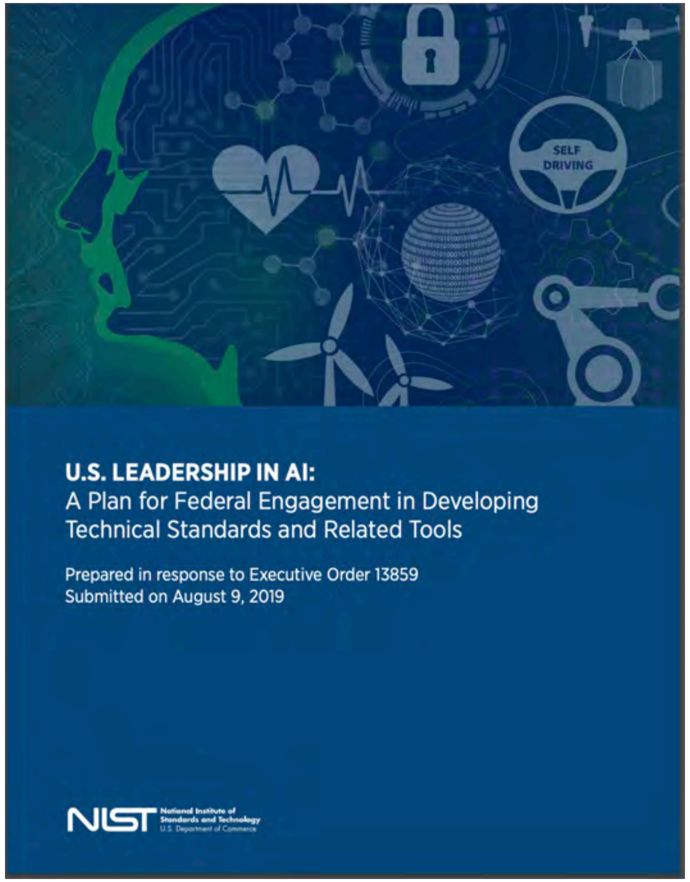
AI Interagency Working Group

National Security Commission on AI
Chief Technical Advisor

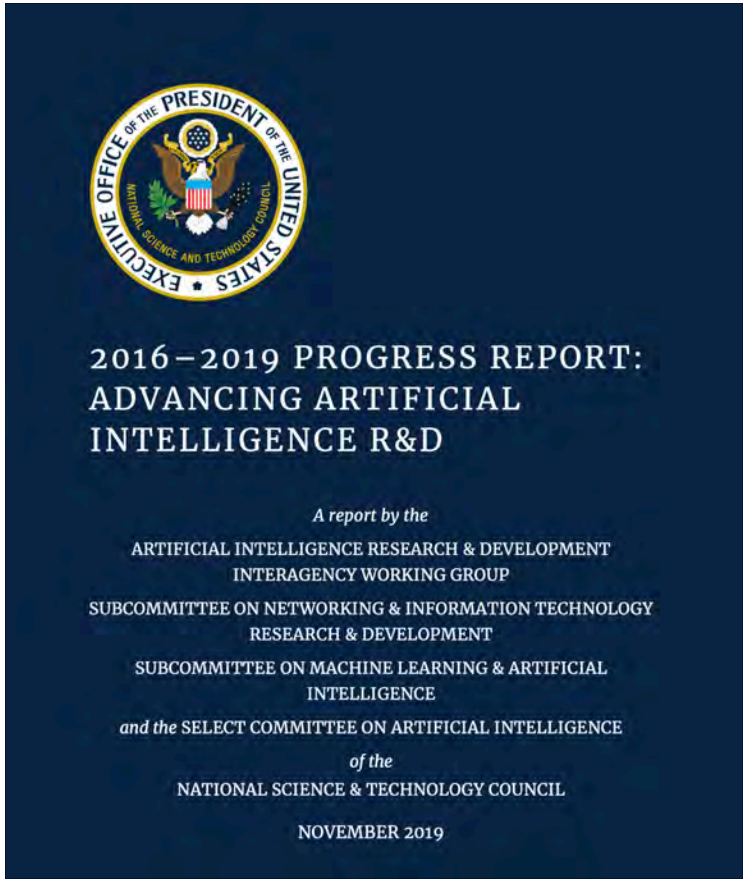
Policy documents in 2019



www.nitrd.gov/news/AI-Progress-Report-2016-2019.aspx



www.nist.gov/sites/default/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf



www.nitrd.gov/pubs/AI-Research-and-Development-Progress-Report-2016-2019.pdf

QUESTIONS?

Elham Tabassi
tabassi@nist.gov

Thiago Seuaciuc-Osorio

Organization/Role: EPRI - Senior Technical Leader in the Nuclear Nondestructive Evaluation (NDE) group at EPRI

Education/Experience: BS in Physics, MS in Mechanical Engineering, currently pursuing MS in Computer Science for Data Science. Nearly 10 years at EPRI.

Current ML/AI work: Leading ML/AI projects related to NDE in Nuclear and helping coordinate ML/AI efforts in the Nuclear Sector at EPRI

Title: Data Driven Decision Making (3DM)

Overview: An overview of how EPRI is seeing and approaching AI in the Nuclear Sector, with some examples of current work in the area.

ELECTRIC POWER RESEARCH INSTITUTE

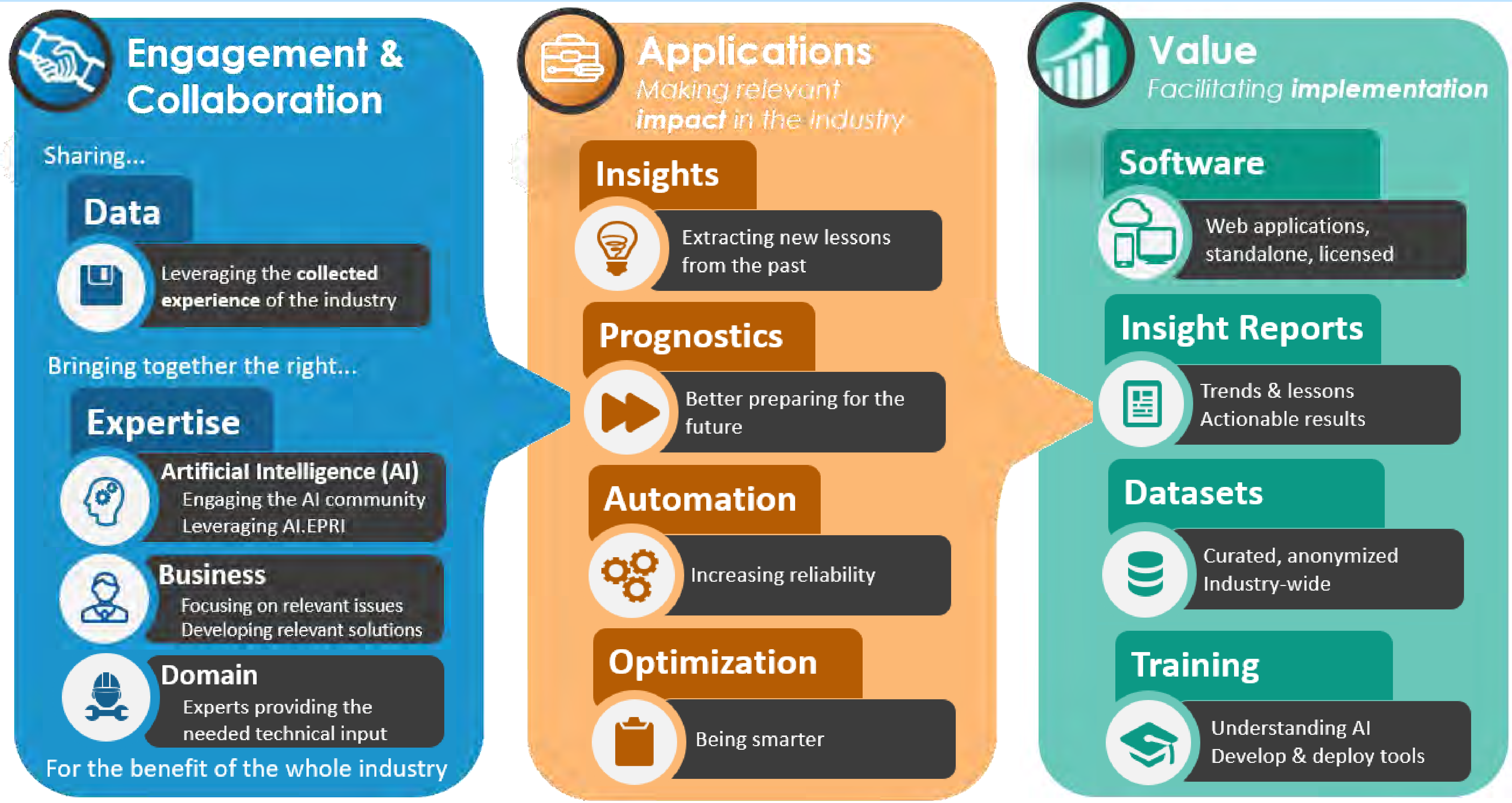
Data Driven Decision Making (3DM)

Rob Austin
raustin@epri.com

Thiago Seuaciuc-Osorio
tosorio@epri.com

July 9, 2020

Enabling data-driven decision making through the collaborative application of data science technologies



Application Examples

Insights

- Event Management Response Tool (EMRT)
- Mining work order database

Automation

- Automating analysis of NDE data
- Automatic diagnostics and processing of CAP data

Prognostics

- Using CHECWORKS database to improve FAC wear rate predictions
- Plant historian data analysis

Optimization

- Decision logic for source term reduction
- Use of AI to enhance inventory management

Together…Shaping the Future of Electricity

Lars Kotthoff

Organization/Role: University of Wyoming - Assistant Professor of Computer Science, Director of Artificially Intelligent Manufacturing (AIM) Center

Education/Experience: PhD in CS/AI (St Andrews), postdoctoral appointments at University College Cork and University of British Columbia, faculty since 2017

Current ML/AI work: Developing and applying techniques from AI and ML to problems in materials science

Title: AI for Materials Science

Overview: Provide information on work in applying AI to problems in materials design and will briefly touch on the AI background.

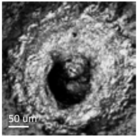
AI for Materials Science

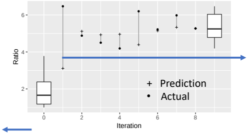
Lars Kotthoff
University of Wyoming
Artificially Intelligent Manufacturing Center
larsko@uwyo.edu

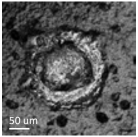
INL ML/AI Symposium, 09 July 2020

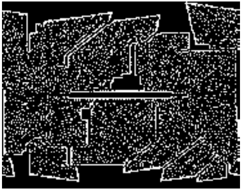
Overview

- Application of AI and ML techniques to Materials Science
- Bayesian Optimization to optimize materials design and production



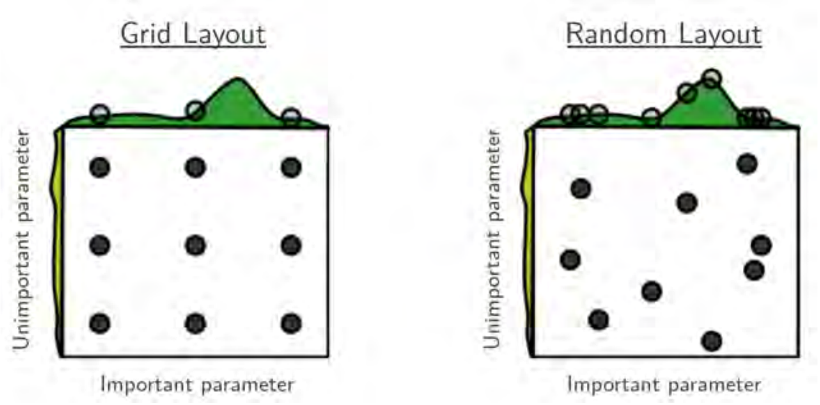






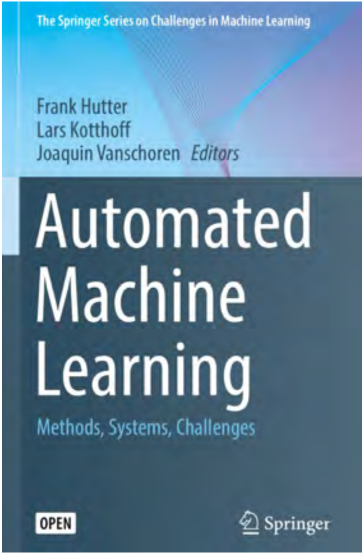
Background – AI-Automated Processes

- Automated tuning of black-box processes
- Mature techniques used in many areas of AI and elsewhere

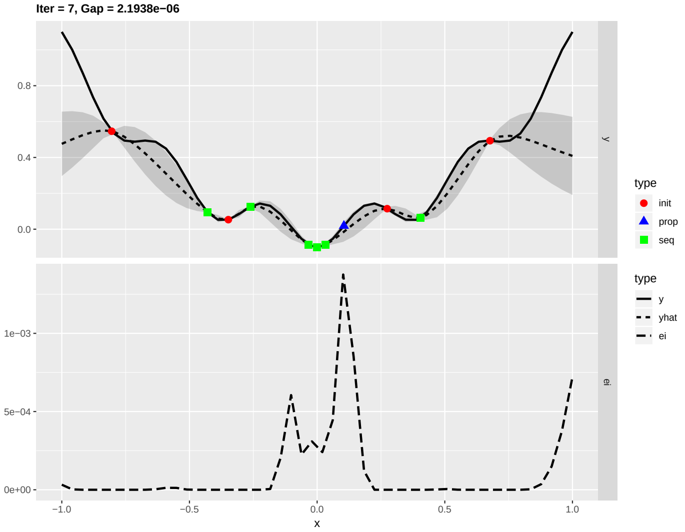


Bergstra, James, and Yoshua Bengio. "Random Search for Hyper-Parameter Optimization." J. Mach. Learn. Res. 13, no. 1 (February 2012): 281–305.

Automated Machine Learning



Bayesian Optimization with Surrogate Models



Bischl, Bernd, Jakob Richter, Jakob Bossek, Daniel Horn, Janek Thomas, and Michel Lang. "MlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions," March 9, 2017. http://arxiv.org/abs/1703.03373.

Challenges and Opportunities

- Other applications
- Inform understanding of optimized process by what surrogate model has learned
- Multi-scale and multi-fidelity measurements and simulations
- Multi-objective optimization

Koroush Shirvan

Group: Massachusetts Institute of Technology

Education/Experience: BS’08 UF, SM’10 PhD’12 MIT Research/Principal Scientist ’12-17, Assistant Prof. (current) all in nuclear power engineering

Current ML/AI work: Reinforcement learning for core design, physics-informed ML for nuclear safety

Title: AI for Nuclear Core Design

Overview: Turn nuclear core reload design tactics into game-play strategy and apply reinforcement learning to achieve more optimized loading patterns

Machine Learning & Artificial Intelligence Symposium
July 9, 2020

www.inl.gov

Idaho National Laboratory

Koroush Shirvan
Massachusetts Institute of Technology
AI for Nuclear Core Design

Massachusetts Institute of Technology

NSE
Nuclear Science and Engineering
science : systems : society

Core Design Today

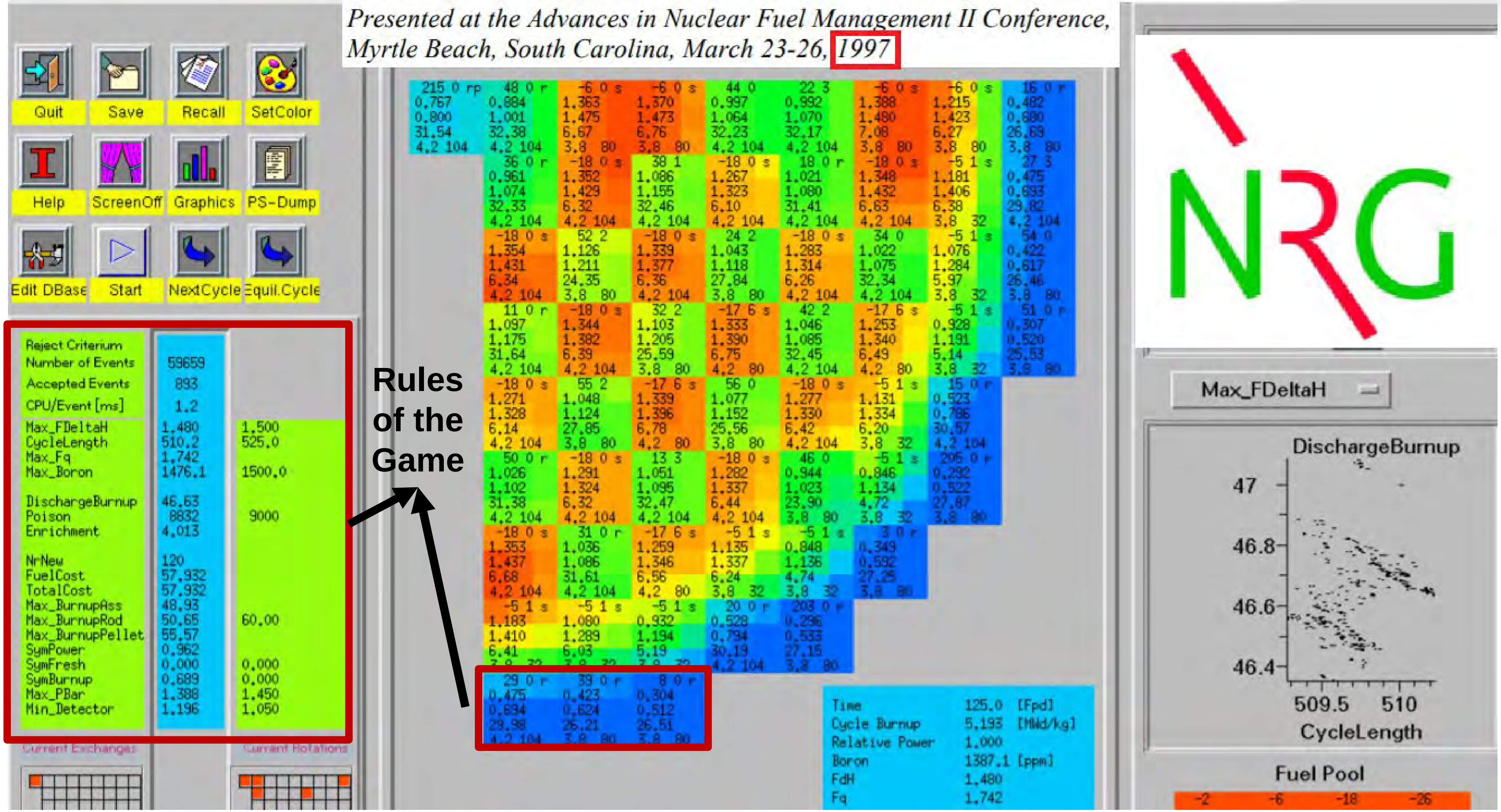
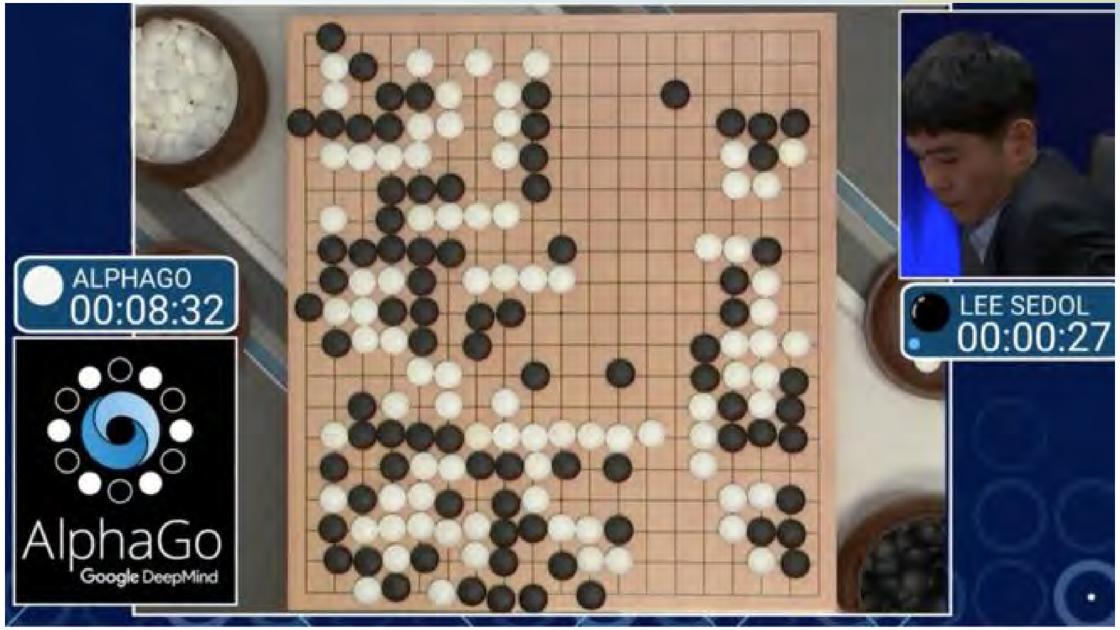


Figure adapted From: https://www.nrg.eu/fileadmin/nrg/Afbeeldingen/producten/5._Asset_Optimalisatie/rosamb.pdf



https://www.bbc.com/news/technology-35785875

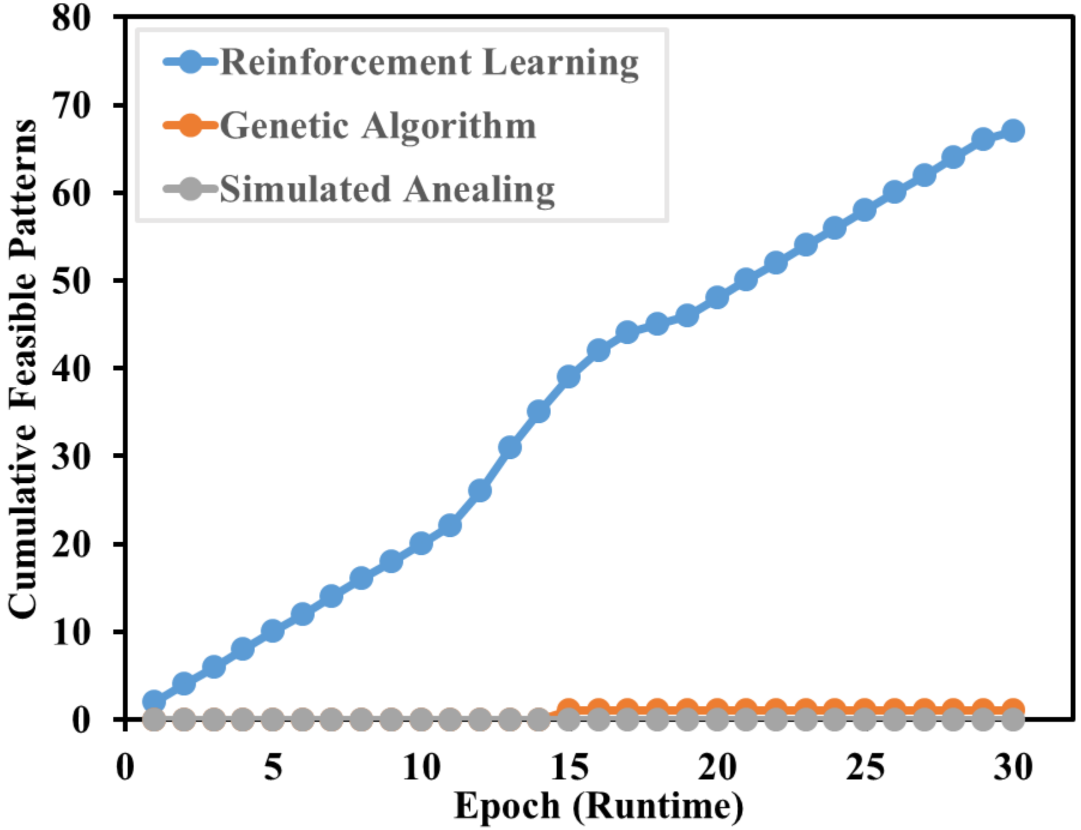
2016: Google DeepMind's ***AlphaGo*** *defeats Go Champion Lee Sedol which AI researcher thought was not possible to achieve in next* ***20 years***

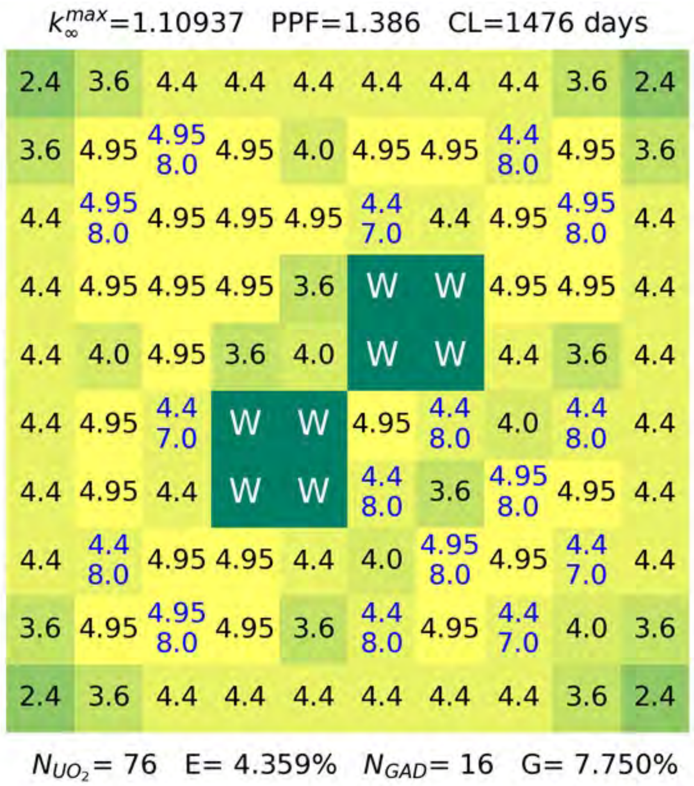
MIT will reshape itself to shape the future, investing $1 billion to address the rapid evolution of computing and AI — and its global effects. At the heart of this effort: a $350 million gift to found the MIT Stephen A. Schwarzman College of Computing.

Photo: Christopher Harting

Can Reinforcement Learning trump Stochastic algorithms?

To be Submitted: Radaideh M., et al., 2020



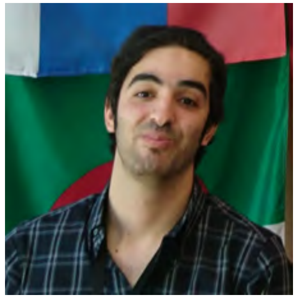


- Value of AI vs. Stochastic Algorithms for ~10^{30} combinatorial problem with licensed methodology:
  - >1000x faster and more efficient exploratory features for the specific problem setup
  - Incorporation of physics-based game tactics was key to the success of AI

Looking Ahead

- **This project: deliver the first software package to Exelon for testing in September 2020**
- **Safety vs. non-Safety Application of ML/AI:**
  - **"When used as a surrogate for a detailed model, the impact on risk needs' to be shown to be insignificant with respect to the results being used to support the decision." NUREG-1855 (2017)**
- **We need more design optimization studies along with high fidelity tool development**
- **Final remarks on value of ML/AI to nuclear energy:**
  - **In my view, the largest value proposition is attracting the best talent to the nuclear energy area**

AI Core Design Team (Students and Postdoc)


Majdi Radaideh


Jane Reed


Paul Seurin


Haijai Wang


Isaac Wolverton

Exelon

Idaho National Laboratory

Questions?

Clean. Reliable. Nuclear.

Koroush Shirvan, Ph.D.

Massachusetts Institute of Technology

AI for Nuclear Core Design

kshirvan@mit.edu

Min Xian

Organization/Role: University of Idaho - Assistant professor at the University of Idaho

Education/Experience: Ph.D. in Computer Science at Utah State University, and 3 years at U of I

Current ML/AI work: Focusing on developing robust and efficient deep learning architectures

Title: Domain-Enriched Deep Architectures and Applications

Overview: Discuss problems in purely data-driven models and exciting recent progress in domain-enriched deep learning and provide insight into the future research directions in deep learning.

Machine Learning & Artificial Intelligence Symposium
July 9, 2020

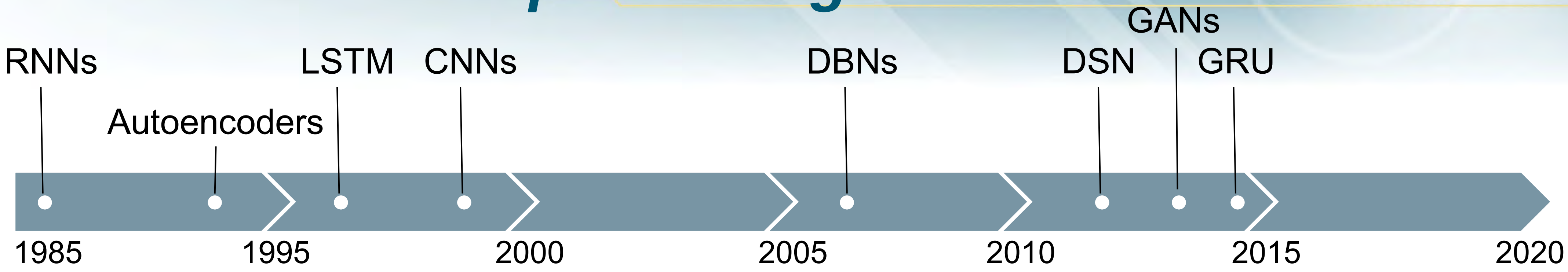
www.inl.gov

Idaho National Laboratory

Min Xian
University of Idaho

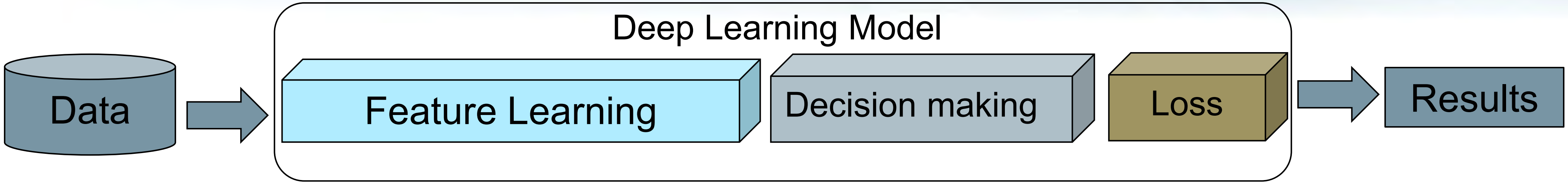
Domain-Enriched Deep Architectures and Applications

Architectures in Deep Learning



Architecture	Application
RNNs	Time series data modeling, speech recognition, handwriting recognition,
Autoencoders	Data Anomaly detection, dimensionality reduction, information retrieval
LSTM/GRU networks	Time series data modeling, natural language text compression, handwriting recognition, speech recognition, gesture recognition, image captioning
CNNs	Image recognition, video analysis, natural language processing
DBNs	Image recognition, information retrieval, natural language understanding, failure prediction
DSNs	Information retrieval, continuous speech recognition
GANs	Image synthesis, image translation, video games, Speech2Face

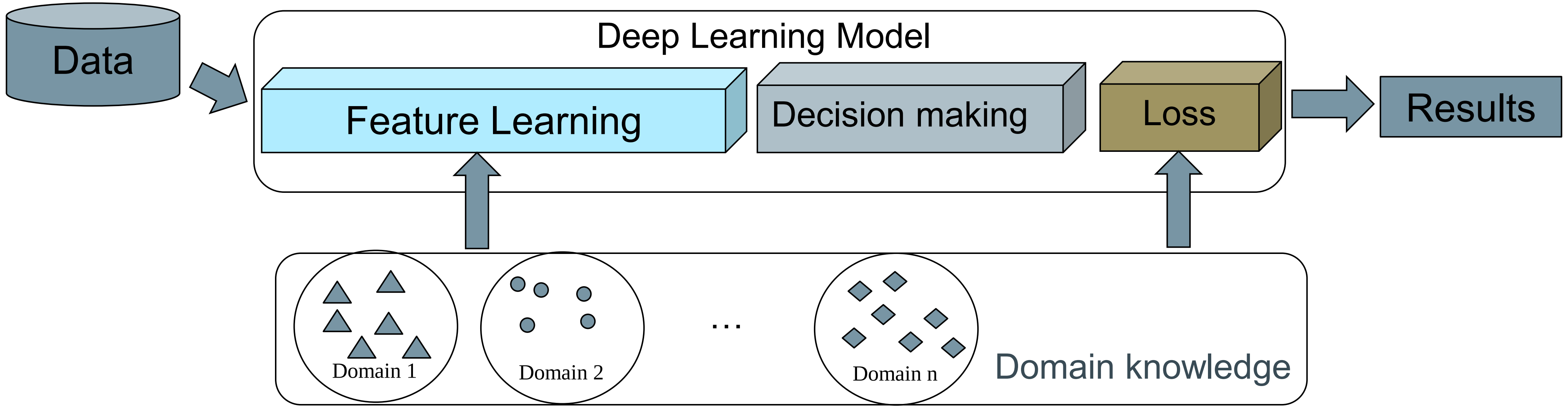
Why Domain-Enriched Deep Architectures?



One of the major contributions of deep learning is the automated feature learning processing.

However, compelling open challenges remain:

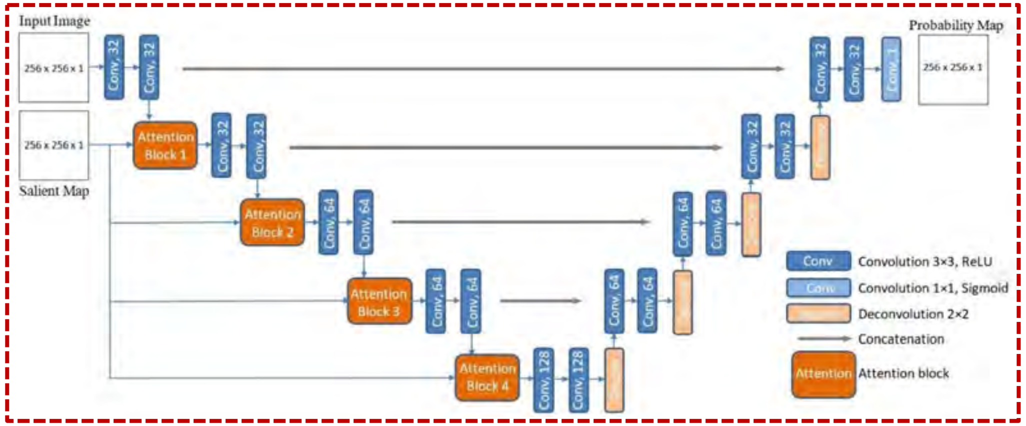
- The performance of a purely data-driven approach heavily depends on ***the quantity and quality of the training data***.
- It is not straightforward to interpret.



Applications

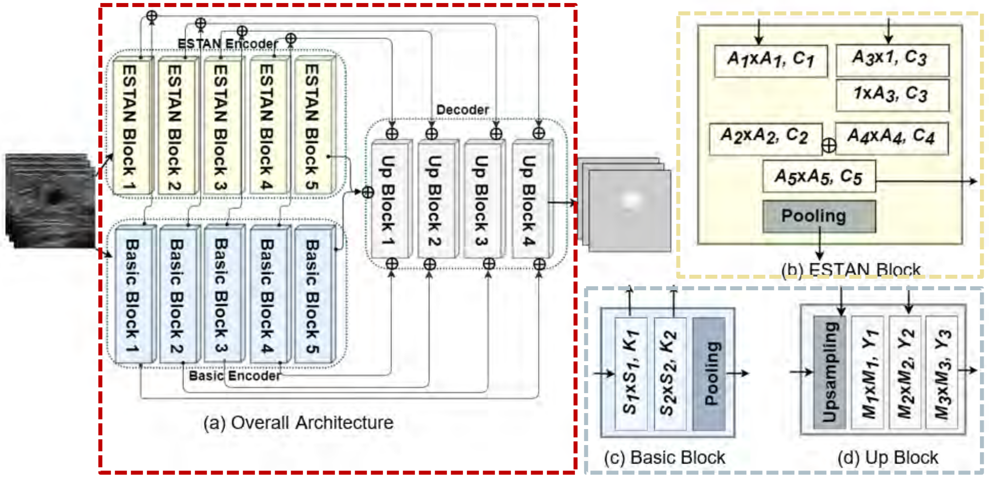
Attention-enriched architecture

A. Vakanski and M. Xian, 2020



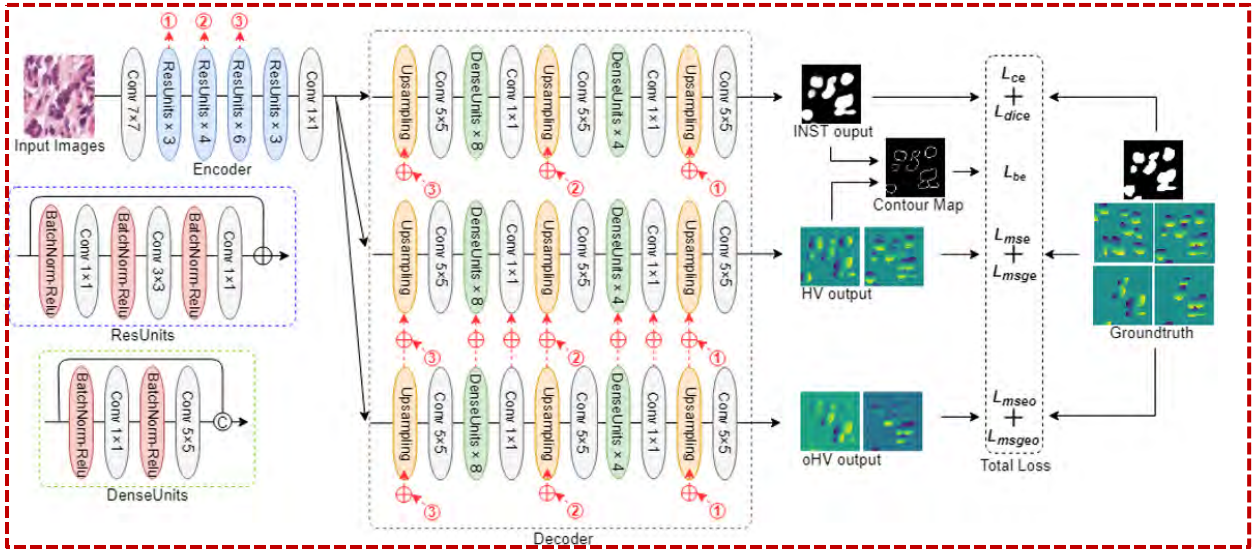
STAN for small object detection

B. Shareef and M. Xian, 2020



Bending Loss-regularized multi-task learning

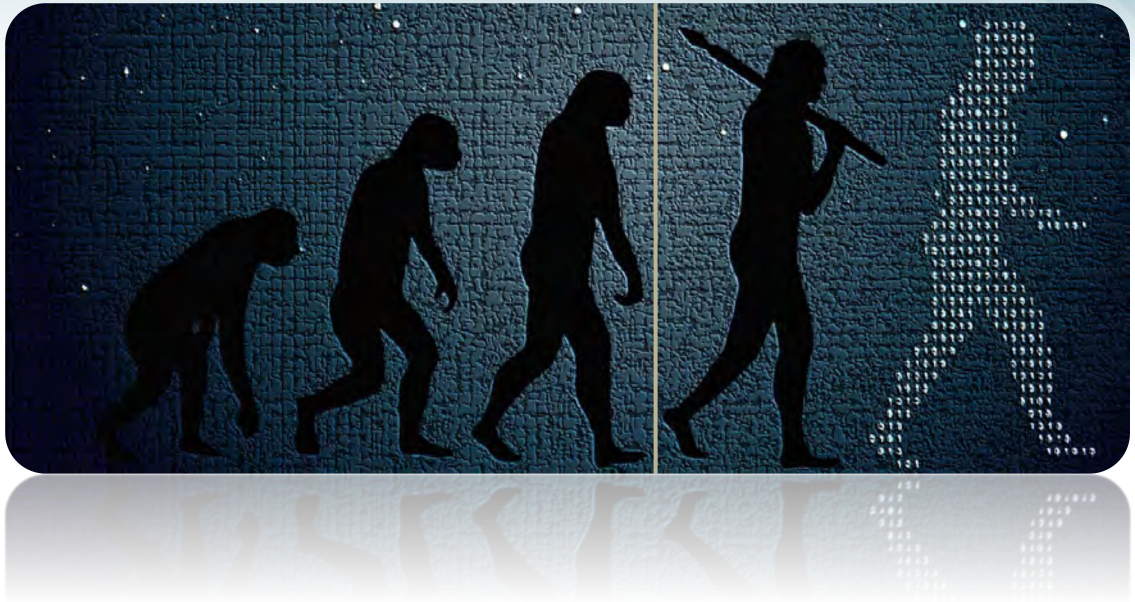
H. Wang, and M. Xian, 2020



Other applications:

Data reconstruction	**Cybersecurity**
Medical image analysis	Face recognition
Self-driving cars	Machine translation
Anomaly detection	Games
	…

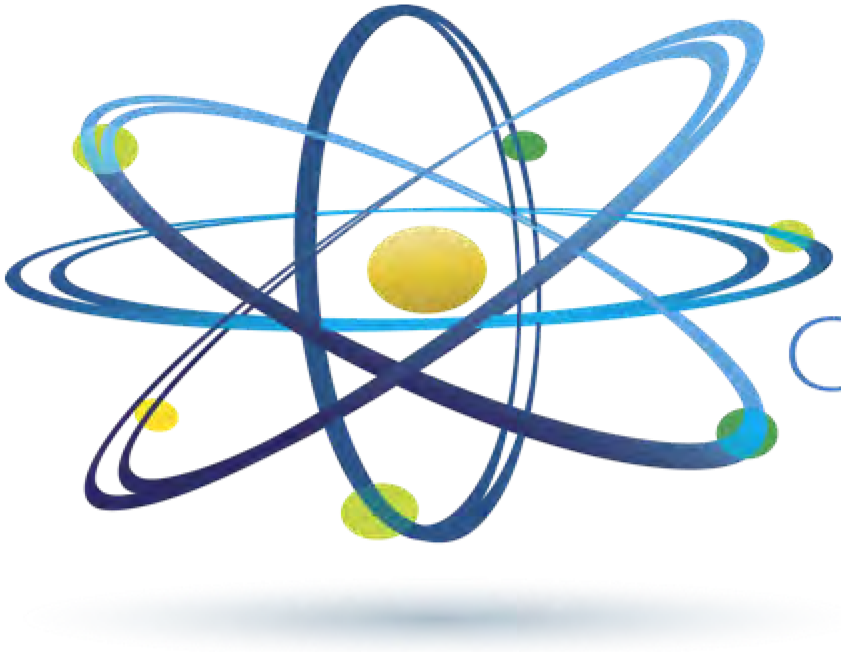
Future Deep Learning and AI



- **What we have**
  - Automounous cars
  - Accurate face recognition
  - Computer-aided medical image analysis
  - Useful Chatbots
  - Acceptable language translation
  - Numerous applications in energy, finance manufacturing, biology games, …

- **What we will have**
  - **Automated deep learning.** In the future, model architectures will be learned, rather than handcrafted by engineer-artisans
  - **Modular subroutine reuse.** Not only leverage previously learned features (submodel weights), but also model architectures and training procedures
  - **Artificial General Intelligence**. A machine has the capacity to understand or learn any intellectual task that a human being can.

Questions?



Clean. **Reliable. Nuclear.**

MIN XIAN, Ph.D.
Assistant Professor

College of Engineering
Department of Computer Science
University of Idaho

mxian@uidaho.edu | 208-757-5425

Milos Manic

Organization/Role: Virginia Common Wealth University - Prof., Virginia Commonwealth University, Dir., VCU Cybersecurity Center, JA, INL

Education/Experience: Ph.D. degree in Computer Science, University of Idaho. Over 40 research grants completed in the area of machine learning in cyber security, critical infrastructure protection, energy security, and resilient intelligent control (DOE, NSF, industry).

Current ML/AI work: Trustworthy AI, Explainable, Reliable, Secure, Fair, Unbiased AI

Title: The future with AI: Sci-Fi or Reality

Overview: Discuss ethics of AI and provide insight into the latest trends in deep and adversarial learning, trustworthy and explainable intelligence, and present the challenges and directions in which AI/ML techniques are developing.

The future with AI: Sci-Fi or Reality

Explainable, Trustworthy, Reliable, and Secure

Machine Learning & Artificial Intelligence Symposium
July 9, 2020

Milos Manic, PhD
Professor, Virginia Commonwealth University,
Affiliate, Idaho National Laboratory

VIRGINIA COMMONWEALTH UNIVERSITY
Modern Heuristics Research Group
http://mhrg.vcu.edu

Research Overview
AI/ML in Resilience and Security

VIRGINIA COMMONWEALTH UNIVERSITY MHRG Modern Heuristics Research Group

Milos Manic, Ph.D., Director, VCU Cyber Security Center, NSA CAE-CD

Director, MHRG Group, Virginia Commonwealth University, Richmond, VA, Joint Appt, INL

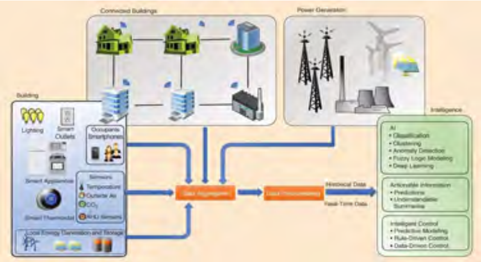
Smart, Sustainable, Resilient, and Secure Cities

We view community resilience as **data** and **human knowledge** driven goals

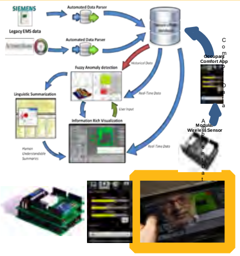
100RC (100 Resilient Cities) – capacity to survive, adapt, and grow, regardless of stressors (aging infrastructure, food, energy & water security, cyber attacks, etc.)

Role of **Machine Learning** and **AI**:
connected, **smart**, **efficient** modern municipalities and critical infrastructures.

Holistic view: **understanding** sub-systems, **interdependencies**, and the **risks** cities face.



AI fuel efficient transportation



TEMST – Targeted Energy Management Toolset for Building Managers

Visual Data Mining
VCU Virtual Reality Lab

Cybersecurity and Resilience

Dr. Milos Manic's research areas: Data Analytics, Machine Learning (ML) and Artificial Intelligence (AI) approaches applied to resilience and security of critical infrastructures.

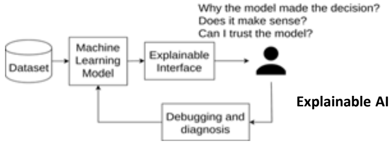
ML driven cybersecurity: anomaly detection, holistic CPS cyberhealth and state awareness, intelligent controls, software vulnerability identification.

Explainable AI (XAI): trust in AI systems; transparency of complex AI models (e.g. deep learning).

Adversarial machine learning: exploits/strengthening AI algorithms and data.

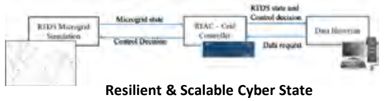
Embedding domain-knowledge in ML: combining physics based (diff. eqs) and data driven (AI/ML) modeling of complex control systems.

Autonomic Intelligent Cyber Sensor (AICS) AI cybersecurity for industrial control systems



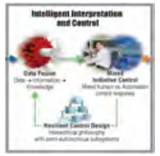
Explainable AI

VTR



Resilient & Scalable Cyber State Awareness through Big Data

Vulnerabilities in Power Grids



Resilient Control Design

- Over 40 invited talks on ML in critical infrastructures, big data integration, nuclear security and energy resiliency, and intelligent human-machine interfaces
- AICS, 2018 R&D 100 Award, one of top 100 science and technology worldwide innovations in 2018.
- Over 200 peer-reviewed publications, 12 book chapters, 33 journals, 8 IEEE Transactions Editorials
- Over 40 research efforts as PI/CO-PI
- Founding Chair of IEEE IES Technical Committee on Resilience and Security in Industry
- 10 advisee awards for outstanding researcher/dissertation
- 15 best-paper and presentation awards at IEEE conferences
- IEEE IES Officer, General Chair of IEEE IECON 2018 http://www.iecon2018.org/, IEEE HSI 2019 http://hsi2019.welcometohsi.org

http://www.people.vcu.edu/~mmanic

mmanic@vcu.edu

What **IS** AI…?

Artificial = Made by humans; Created, produced - rather than natural.

Defining **Intelligence** – much harder!

DARPA DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

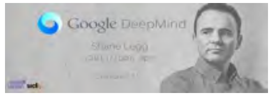
- The capacity to **acquire** and **apply knowledge**.
- The **ability to learn** or understand things or to deal with new or trying situations: the skilled use of reason.
- Terminology..
  - AI, ML, CI, Deep learning

AI our attempt to build models of ourselves?

AI today…

- *"data driven"*
- *takes many forms*

Why NOT use AI today?



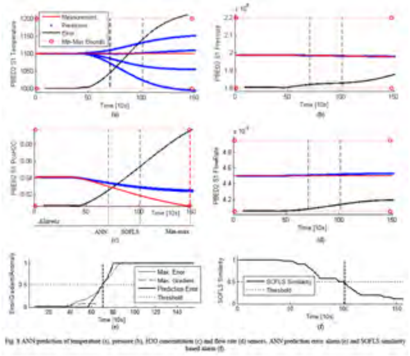
The difficult questions…

- How do you…
  - *…replicate something we do not understand?*
  - *Sentience…emotion, love, dream, conciseness, fear, anger, memory (ours is subjective, fallible)*
  - *Trust and trustworthy, how to quantify?*
  - *Regulate? Public and gov policies*
  - *Autonomous vehicles and intelligence - ethical, moral questions questions*

"AI would be the biggest event in human history. Unfortunately, it might also be the last" Elon Musk (Tesla)

"If a super-intelligent machine decided to get rid of us, I think it would do so pretty efficiently" Shane Legg, DeepMind co-founder

Resilient Anomaly Detection System



Detection 3 times faster !!!

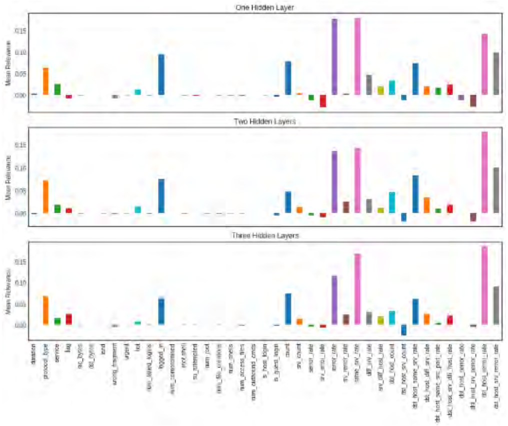
HYTEST

Resilience to transient faults, early warnings

AI Powered Bio-fuel Generation

Increased reliability from 63% to 96%

Anomaly Detection



Different DNN Models

Similar Classification Accuracies

Input Feature contributions **ARE DIFFERENT**

Optimal PMU placement in the grid

AI in cyber manufacturing

AI –> up to 70% cost reduction

For successful adoption of AI…

- Trustworthy AI, transparency, explainability

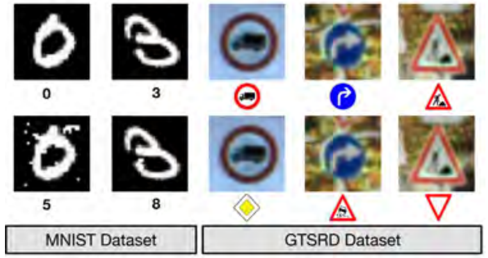
So… AI resiliency, accurate modeling, fairly doable…
*but….**is performance enough?***

Accuracy scores ?

- *Any model will do*
- *Each one uses a different set of features, different learning!*

=> Trustworthy and Explainable AI…

=> Transparency in safety-critical domains



A Holy Grail of Machine Learning - *Generalization*

On previously unseen scenarios

vs.

Adversarial learning

“breaking” generalization

TrustFactory.ai

The difficult questions…

- *Can we develop generalized explaining methods?*
- *How do we measure explainability?*
- *What is a sufficient level of explainability?*
- *Explainability is application/user dependent*

Other…

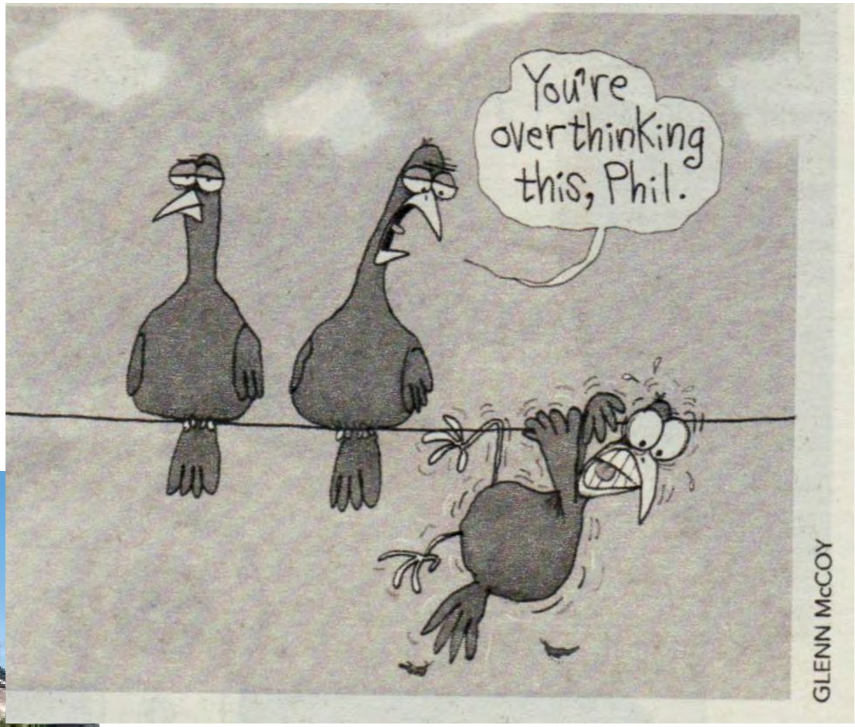
- *Explainable AI*
- *Fair and Unbiased AI*
- *Privacy-Preserving AI*
- *Reliable/Verifiable AI*

Crosscutting Areas

- *Education & WFD*
- *Policy, Governance, Ethics,*

Thank you ☺

Prof. Milos Manic	*mmanic@vcu.edu* http://www.people.vcu.edu/~mmanic



"Simplicity is the ultimate sophistication."
~ Leonardo da Vinci

Alper Yilmaz

Organization/Role: Ohio State University - Professor, The Ohio State University

Education/Experience: BS, MS, and PhD in Computer Science and Engineering, 16 years at OSU

Current ML/AI work: Nuclear Plant Data Analysis, Autonomous Driving, GPS denied geo-localization, sensor fusion, Computer Vision

Title: Application of Deep Learning on NPP Related Data

Overview: Discuss ongoing projects, with brief description on how data should be utilized and what type of results can be obtained. The two projects are on 1) images 2) non-image data.

Applications of Deep Learning on NPP Data

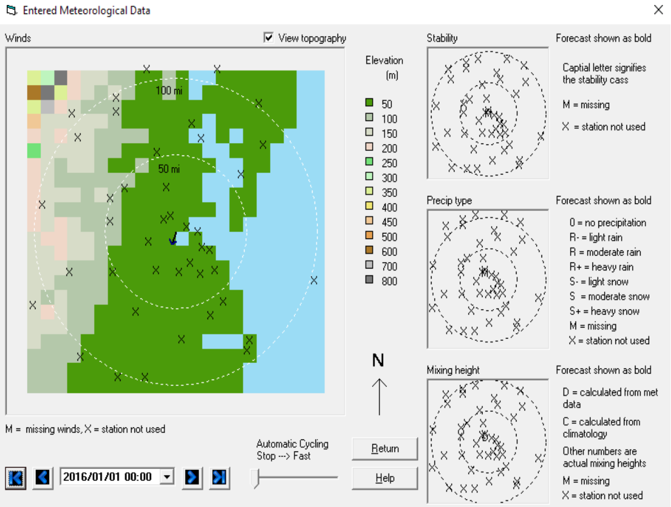
Alper Yilmaz, PhD
Professor, Civil Environmental and Geodetic Engineering
Professor, Computer Science and Engineering (by courtesy)
The Ohio State University

yilmaz.15@osu.edu
@osupcv
http://pcvlab.engineering.osu.edu

Types of Data

- Visual
  - Images sequences from room mounted cameras
  - Images from Augmented Reality mounts

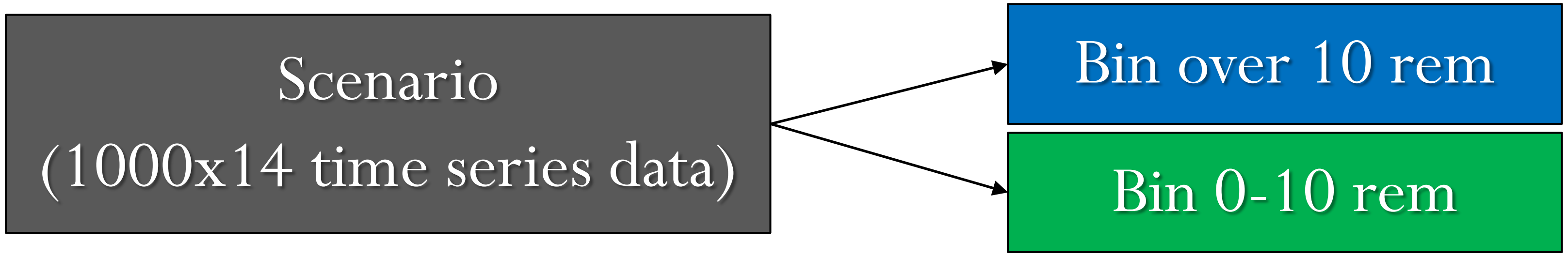
- Non-visual
  - Component states
  - Work orders
  - Images of monitors



Off-Site Emergency Call

- Training data generation

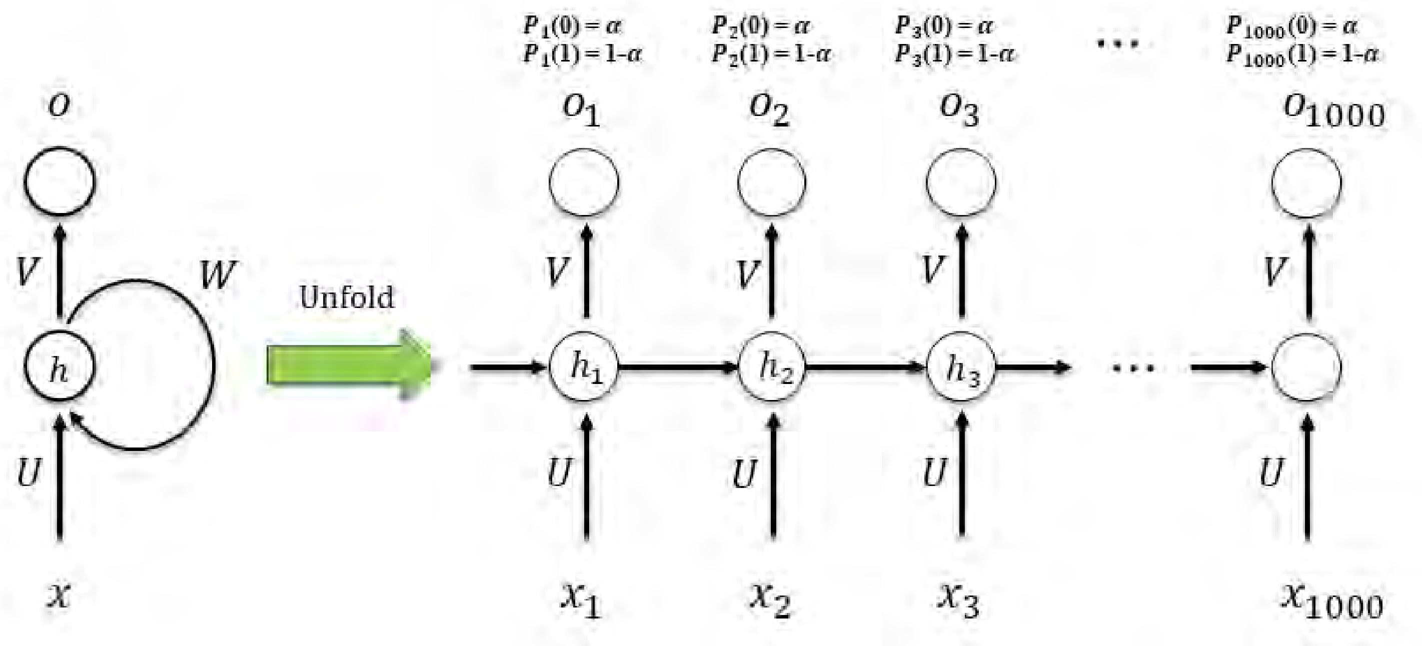
  - Dynamic event tree (DET)
  - Label the pathway evolution following an initiating event
  - Estimate radioactive release based on weather conditions
- Predict Radioactive Material Release for Each Scenario



Transient Data Models

- Long short-term memory (LSTM) for Classification

- Why LSTM
  - Short memory feedback
  - Variable-length I/O
  - Remembers past
  - Flexible structure



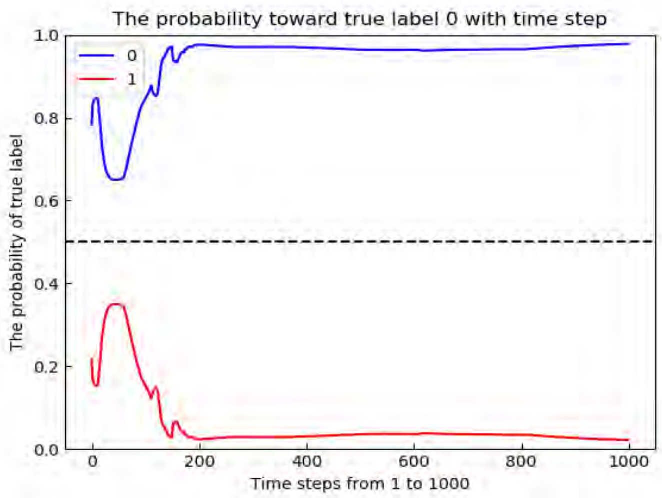
$$\boldsymbol{h}_t = f(W\mathbf{x}_t + V\boldsymbol{h}_{t-1} + \boldsymbol{b}_h)$$
$$\boldsymbol{o}_t = g\left(U\boldsymbol{h}_t + \boldsymbol{b}_y\right)$$

Results

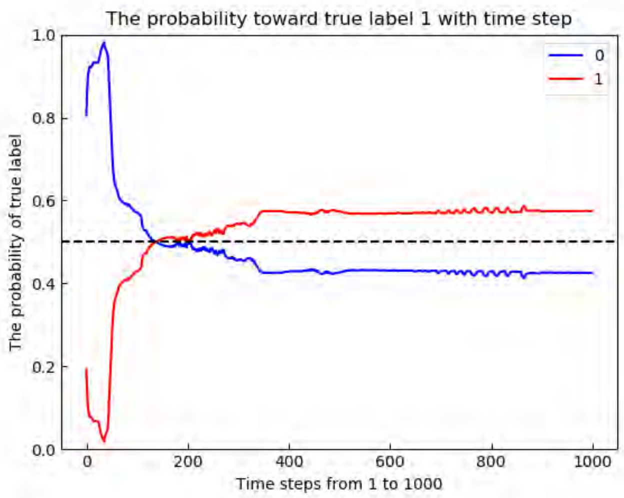
TABLE I: Case Study Data Statistics					
Group	Class	Train	Validation	Test	Total
1	0	1625	407	508	2540
	1	74	18	24	116
2	0	101	26	2413	2540
	1	83	21	12	116

TABLE II: Accuracy of each experiment		
Group	Validation Accuracy	Testing Accuracy
1	0.9976	0.99
2	0.9761	0.99

Probability of label 0 ($> 10rem$)



Probability to label 0 ($\leq 10rem$)



Context-Aware Safety Information Display

- Recognize physical workspaces with maintenance processes

- Real-time overlay of safety information displayed via AR goggles
- Assist field workers in
  - Assessing workspace risks,
  - Locating task-relevant objects,
  - Carrying out the tasks in the correct order

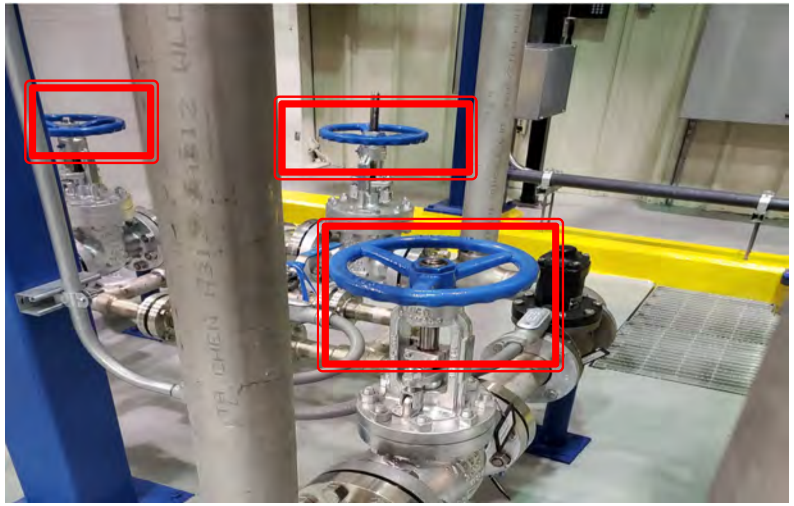
Application Scenarios

Nuclear power plant field sample

Nuclear power plant main control room

Data Engineering & Algorithmic Flow

- Limited training data
- Different class, similar visual content
- Few-Shot / One-Shot learning
- Location as meta data



Contact information

Alper Yilmaz

yilmaz.15@osu.edu

https://pcvlab.engineering.osu.edu

Kasun Amarasinghe

Organization/Role: Carnegie-Mellon University - Postdoctoral Researcher in Machine Learning and Public Policy

Education/Experience: BS (2011) University of Peradeniya, Sri Lanka, and Ph.D. (2019) in Computer Science Virginia Commonwealth University, VA

Current ML/AI work: Conducting research on using Machine Learning for public policy with a focus on ML transparency and fairness for ensuring equitable policy outcomes.

Title: Explainable Machine Learning for Decision Support Systems

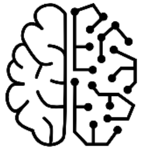
Overview: Discuss on how important explainability is for real-world ML applications, an example framework for explainability, a brief account of existing research, and my view of the future for the field of explainable ML.

Explainable Machine Learning for Decision Support Systems

Kasun Amarasinghe, Ph.D.

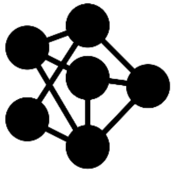
Machine Learning Dept. & Heinz School of Public Policy
Carnegie Mellon University

Black-box vs Explainable Machine Learning



We are incorporating ML in high-stakes settings

- Detecting threats to our critical infrastructure
- Public resource allocation
- Bank loans,
- Incarceration decisions



- Complex data needs complex models!
- But we end up with unintelligible models

What do they learn?

What drives their decisions?


BBC Sign in News More
NEWS
Home Video World US & Canada UK Business Tech Science More
Business Market Data Global Trade Companies More

Apple's 'sexist' credit card investigated by US regulator

11 November 2019

Share

Steve Wozniak

Replying to @dhh

The same thing happened to us. We have no separate bank accounts or credit cards or assets of any kind. We both have the same high limits on our cards, including our AmEx Centurion card. But 10x on the Apple Card.

6:58 AM - Nov 10, 2019

See Steve Wozniak's other Tweets

@STEVEWOZ

Report

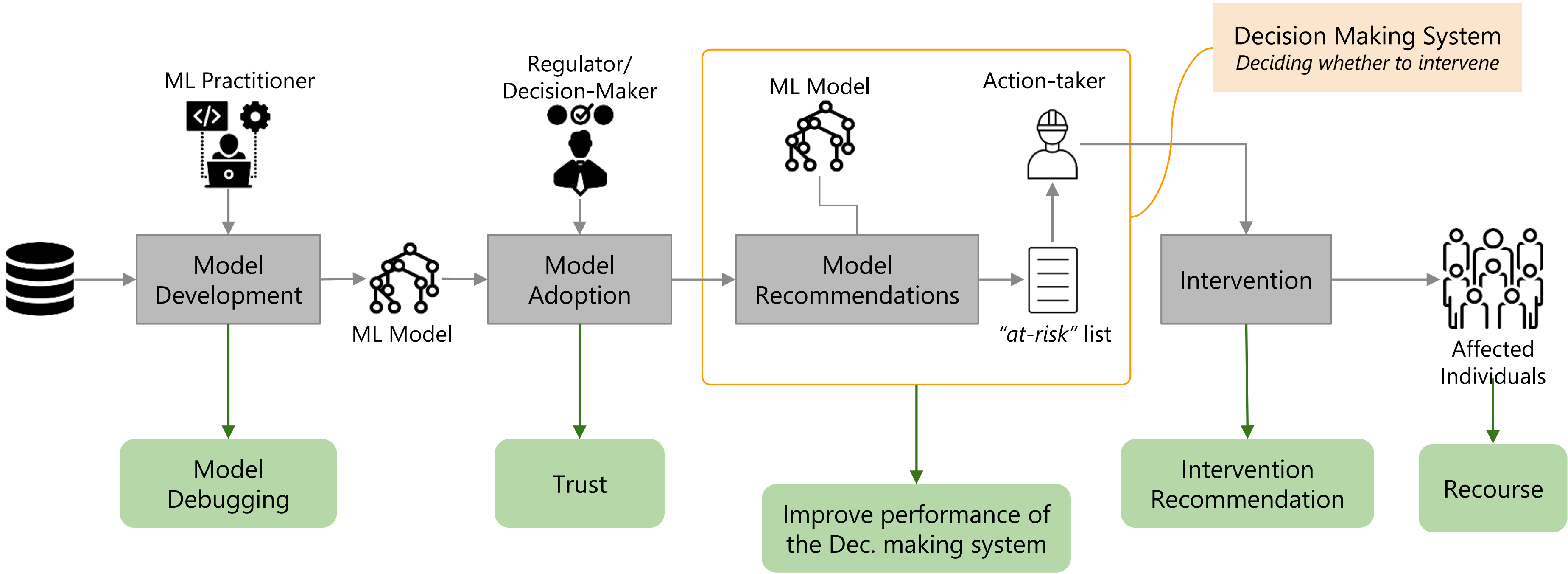

Forbes

Don't Trust Artificial Intelligence? Time To Open The AI 'Black Box'

We need to open the black box

Explainable ML

Why is Explainability so crucial?



Explainable ML framework for Public Policy Applications[1]

Different human "actors" interact with the ML model at different stages and levels

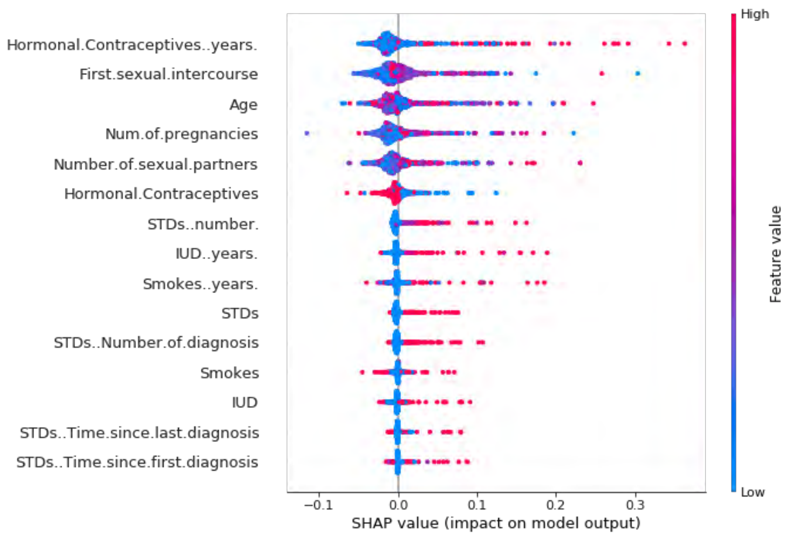
[1] "Making explainable machine learning work for public policy", K. Amarasinghe, K. Rodolfa, R. Ghani

What progress have we made?

- **Two main approaches** to explainable ML:
  - o **Post-hoc explainability** methods for black-box models,
    - LIME, SHAP, LRP, Anchors
    - Most popular out of the two
    - Model-agnostic, and model-specific methods
  - o Developing **inherently explainable models**
    - RiskSLIM, GA2M, MAPLE

- Existing methods:
  - o **Feature attribution** has been the chosen method
  - o **Local and global** explainability

- But, **testing** highly **reliant on synthetic data**, with **“synthetic” users** (AMT)



Feature Attribution Explanations[1]

Theory has not met practice!

[1]https://christophm.github.io/interpretable-ml-book/shap.html

There's a long way to go…

- **Bridge the bap between theory and practice:**
  - Explainability is a domain specific notion
  - Move beyond the buzz-word
  - **Domain specific research** is needed to address the nuances

- **Develop and evaluate existing methods partnering with end-users**:
  - Evaluate **real-world utility** with real users
  - Evaluate the ability to **improve the system outcomes**

- **Move beyond simple feature attribution to generate "complete" explanations:**
  - Explanations with more context than a simple feature importance.

- **Develop methods to generate "useful" explanations by tightly coupling the development process with the end users**
  - Tailor the explanations to include information that is useful to the end-user to accomplish the task

Questions?

Kasun Amarasinghe, Ph.D.
Postdoctoral Researcher
Machine Learning Dept. & Heinz College of Public Policy
Carnegie Mellon University
amarasinghek@cmu.edu

Dan Cole

Organization/Role: University of Pittsburgh - Associate Professor of Mechanical Engineering and Materials Science at University of Pittsburgh

Education/Experience: BS, MS, and PhD in Mechanical Engineering at Virginia Tech, 14 years at Pitt

Current ML/AI work:

- Advanced Online Monitoring and Diagnostic Technologies for Nuclear Plant Management, Operation, and Maintenance
- Data, Modeling, and Forecasting for Nuclear Plant Systems
- Process Anomaly Detection of a Nuclear Power Plant

Title: Machine Learning for Risk-Based Decision Making, Command and Control

Overview: A Discussion of how AI/ML can be used to improve decision making in nuclear power plants, and what needs to be achieved to bridge between advanced simulation to real-time implementation to realize improved command and control.

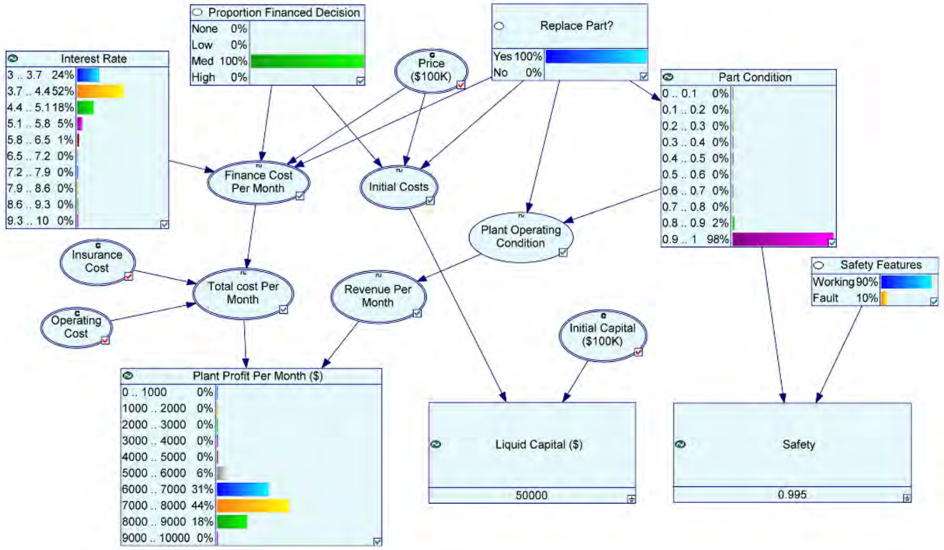
Machine learning for risk-based decision making, command and control

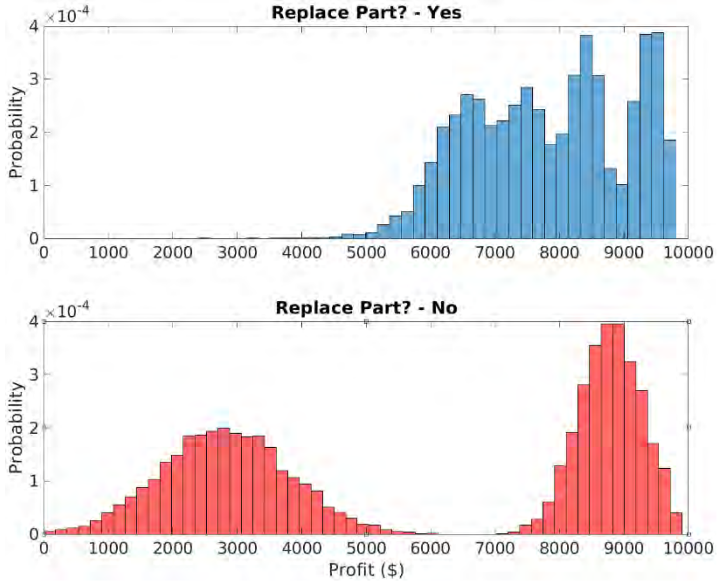
Daniel G. Cole

Mechanical Engineering and Materials Science
Swanson School of Engineering
University of Pittsburgh

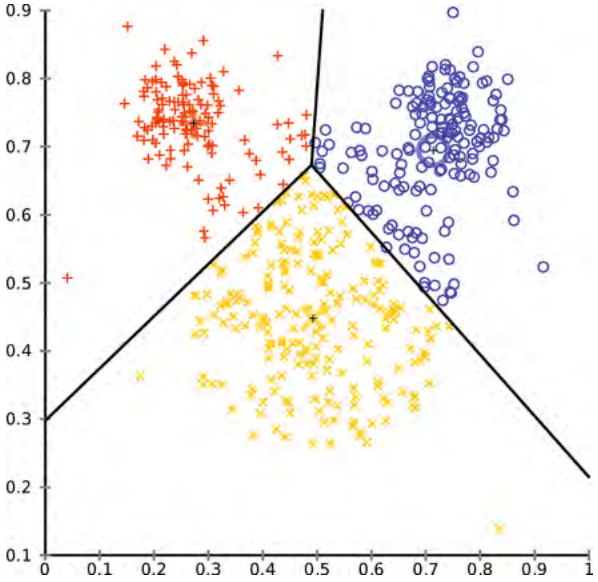
Machine Learning and Artificial Intelligence Symposium

9 July 2020

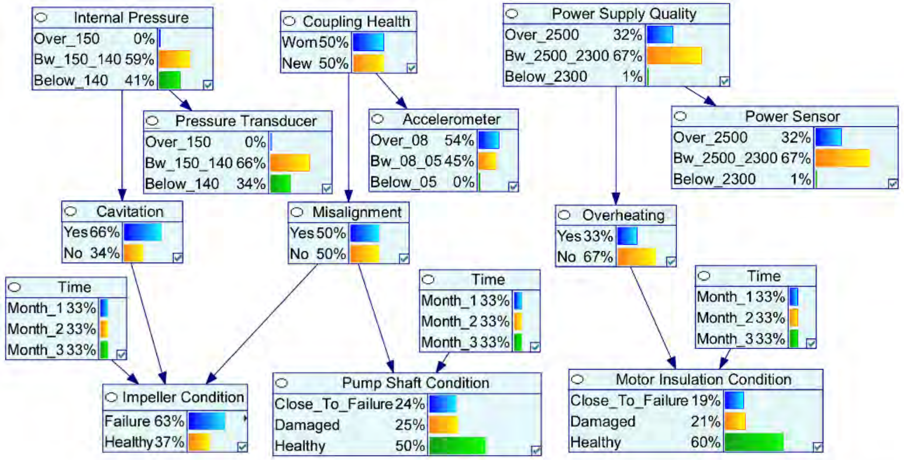




Our work combines data + simulations to determine faults, forecast health, and make decisions



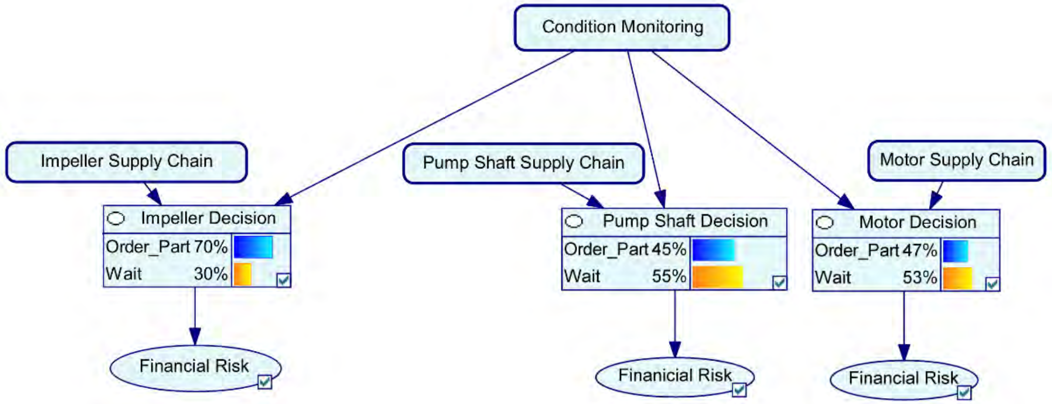
Classifier
Anomaly detection



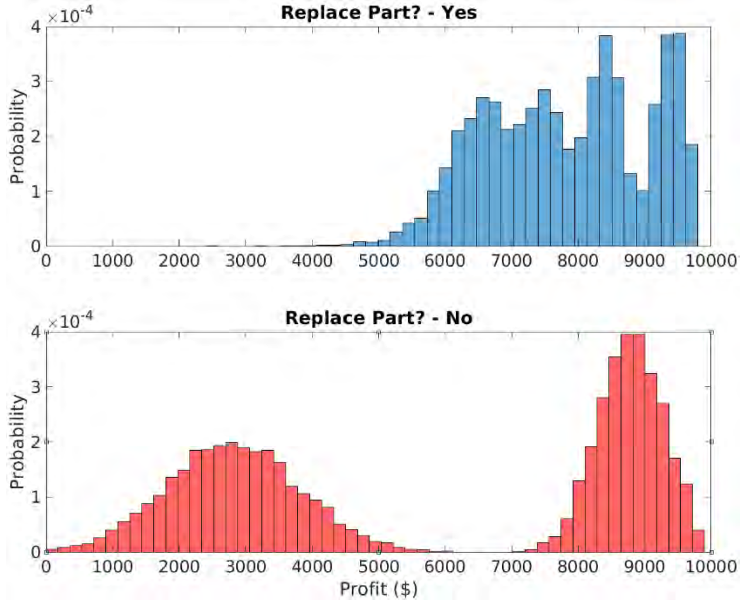
Bayesian networks
Condition monitoring

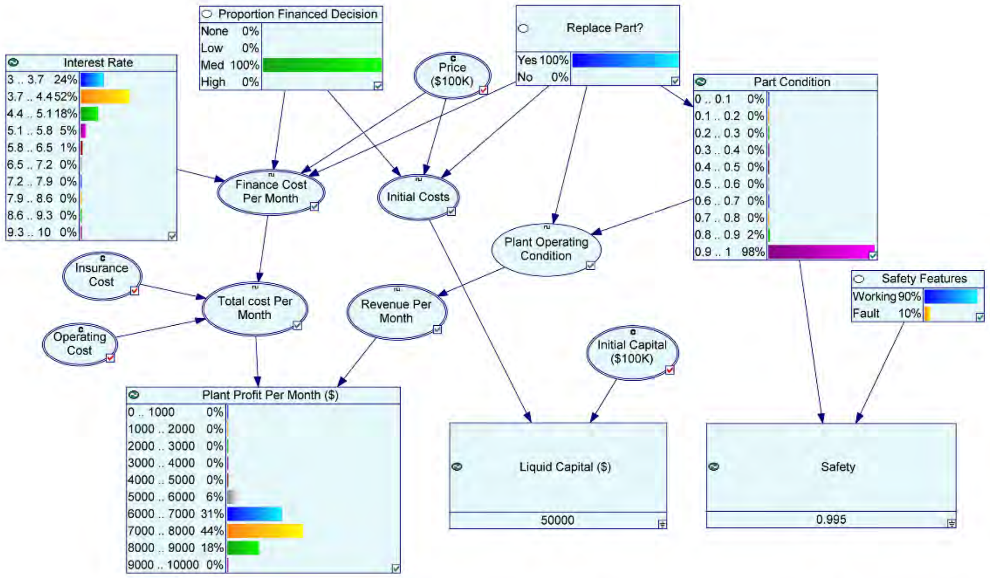
https://en.wikipedia.org/wiki/K-means_clustering

We are integrating health monitoring, supply chain risk, and financial risk for better O&M and asset management



The health of a part can combined with
resource availability make a risk-informed decision about replacing a component.





AI + ML enable us to achieve improved real-time, risk-based command and control

Input

Physical plant (high-burden system)

Simulated plant (costly calculations)

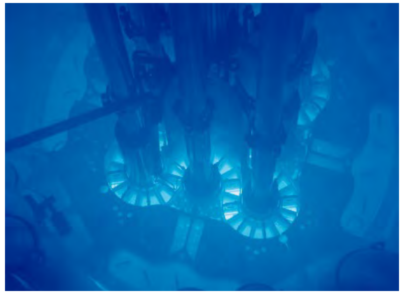
Learning

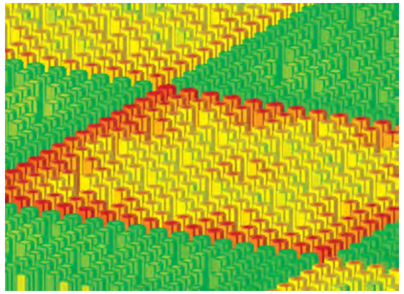
Surrogate model (cheap calculations)

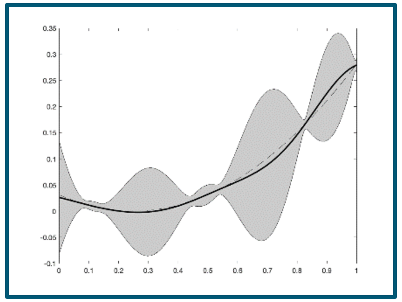
Output

Output matches physical plant

Output matches simulated plant

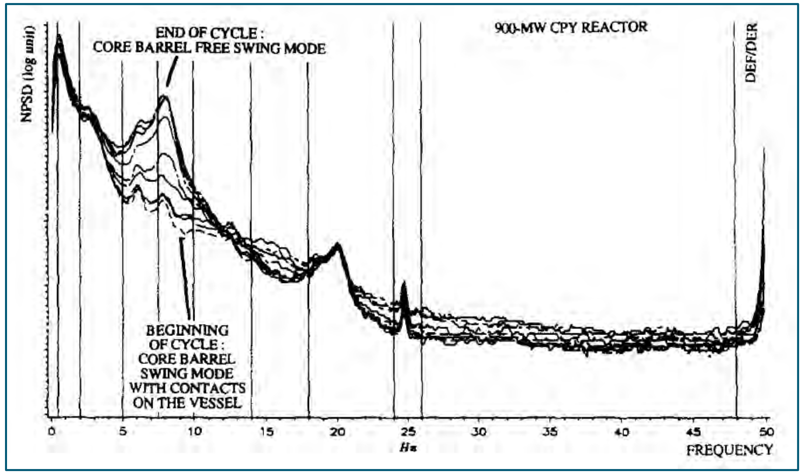




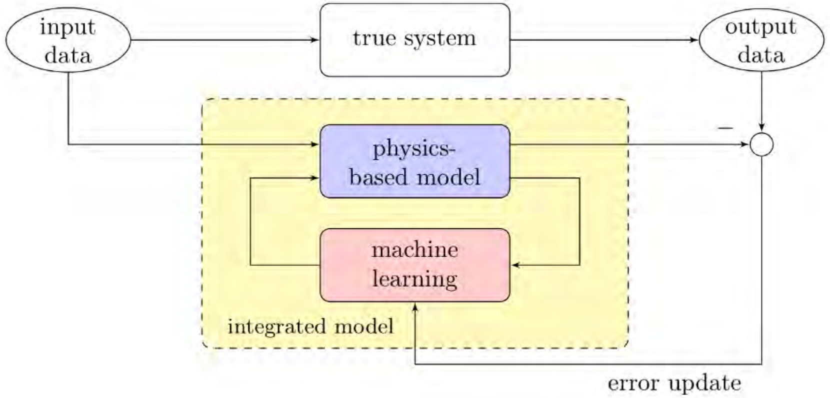


https://commons.wikimedia.org/wiki/File:VERA_reactor_core.jpg

We need methods to handle imbalanced data sets and and to integrate physical models with data-driven ones

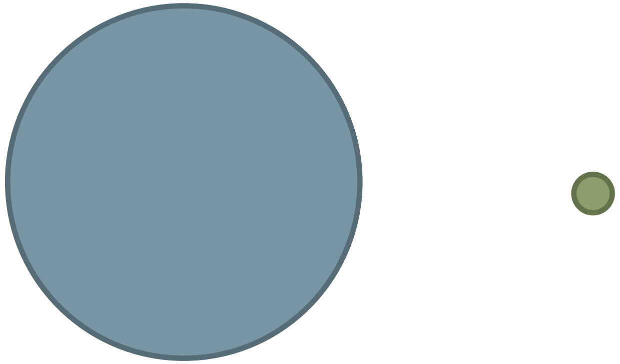


Source: Trenty, Prog. Nuc. Energy, 29(3/4), 347—56, 2995



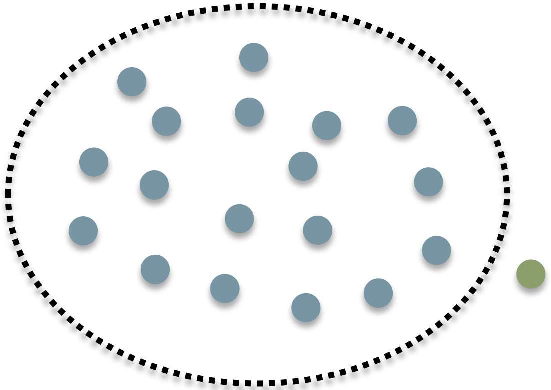
Hybrid models
Physics-based + Machine learning

Data Sources



Normal data

Abnormal data



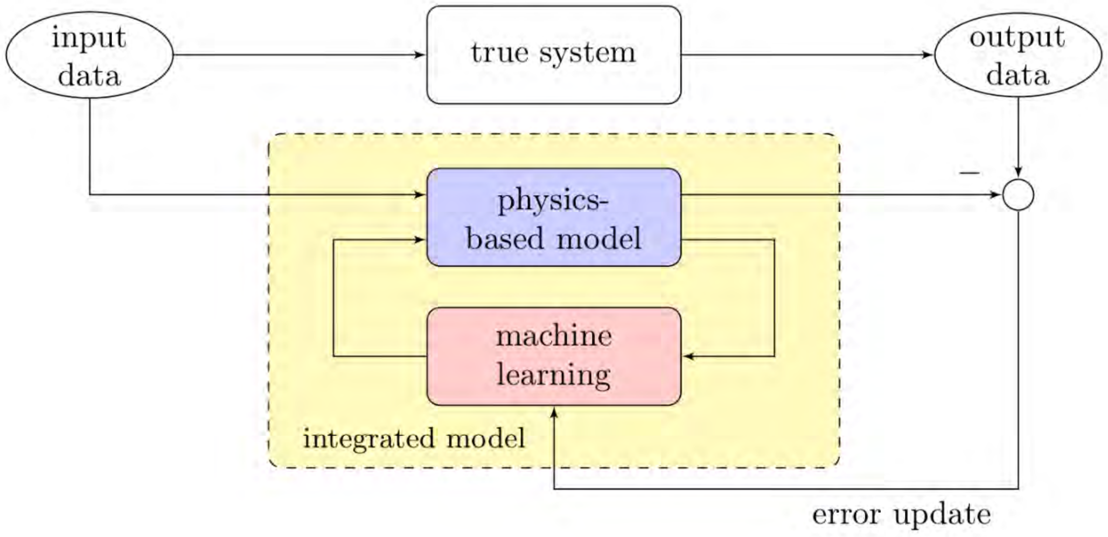
Detecting Anomalies
Classifiers

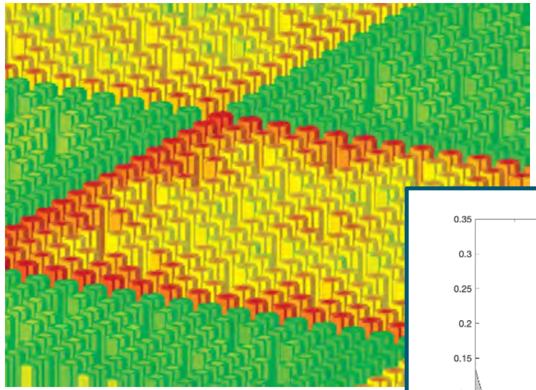
AI & machine learning are necessary to enable real-time, risk-based decision making, command and control

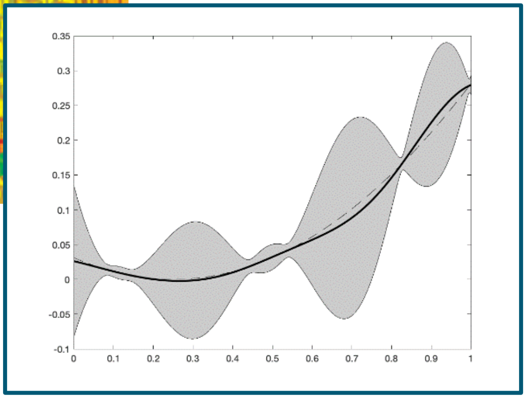
Daniel G. Cole

dgcole@pitt.edu
412-624-3069

Questions?







https://commons.wikimedia.org/wiki/File:VERA_reactor_core.jpg

Hany Abdel-Khalik

Organization/Role: University of Pittsburgh - Associate Professor, School of Nuclear Engineering, Purdue University

Education/Experience: BS, MS, PhD all in Nuclear Engineering. PhD 2004 from North Carolina State University. Worked at AREVA-NP, Lynchburg, NCSU, and Purdue (past 6 years).

Current ML/AI work: Methods development to support nuclear systems performance, safety, and security. I am a computational reactor physicist with R&D interests in data analytics as a basis for improving systems performance, safety, and security

Title: More Letters into the “AI” Acronym

Overview: An overview of new capabilities needed for the ever-increasing role of AI that will focus on “Active AI” designed to influence system operation in order to better-learn its behavior and subsequently better-optimize its operation.

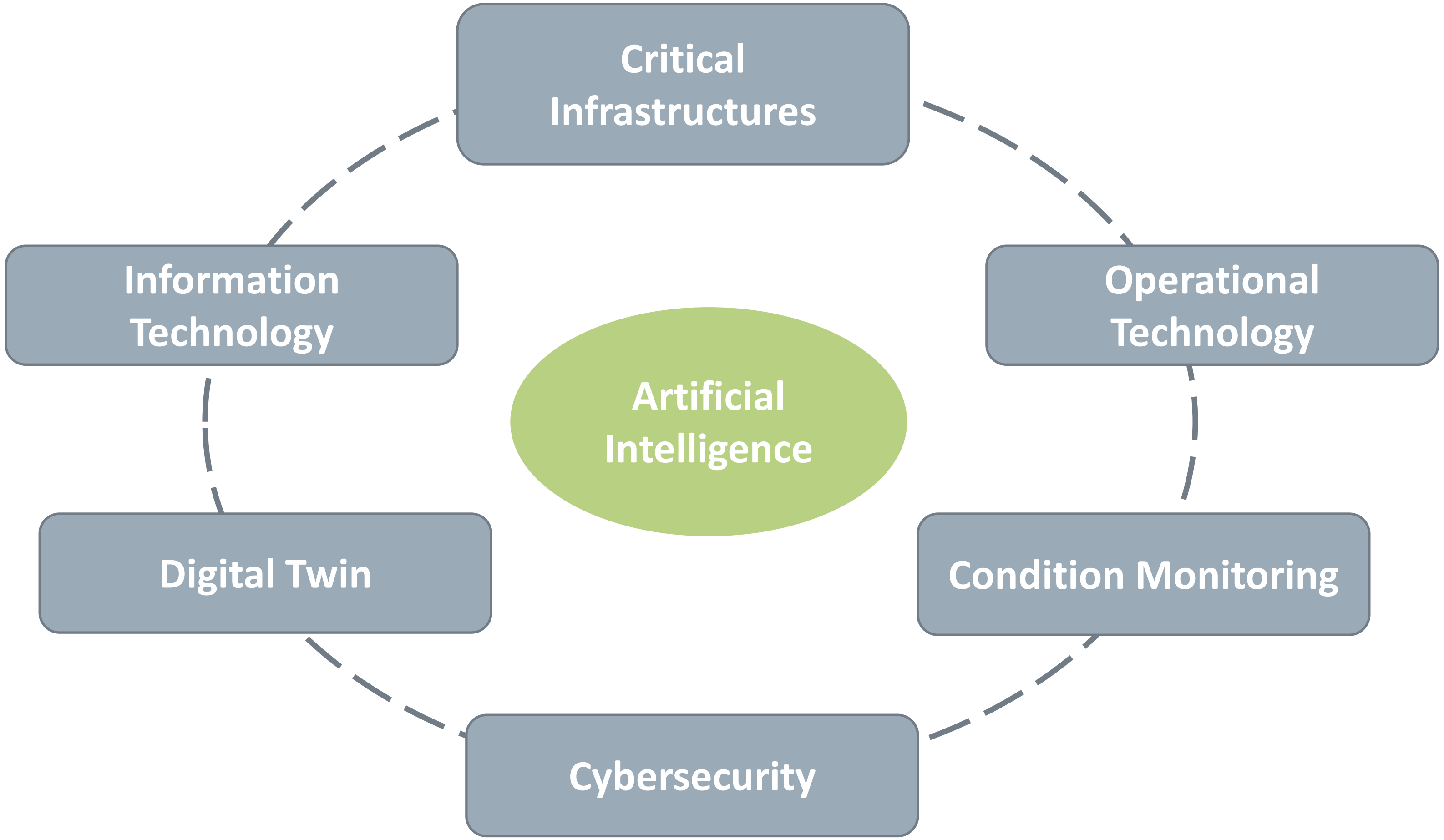
More Letters into the "AI" Acronym

"A": Active, Adversarial,
"C": Covert, Causal, etc.

Hany Abdel-Khalik, Purdue University

INL AI/ML Symposium, July 9th, 2020

Computerized Decision Making Capability @ Center of 21st Science and Engineering Challenges

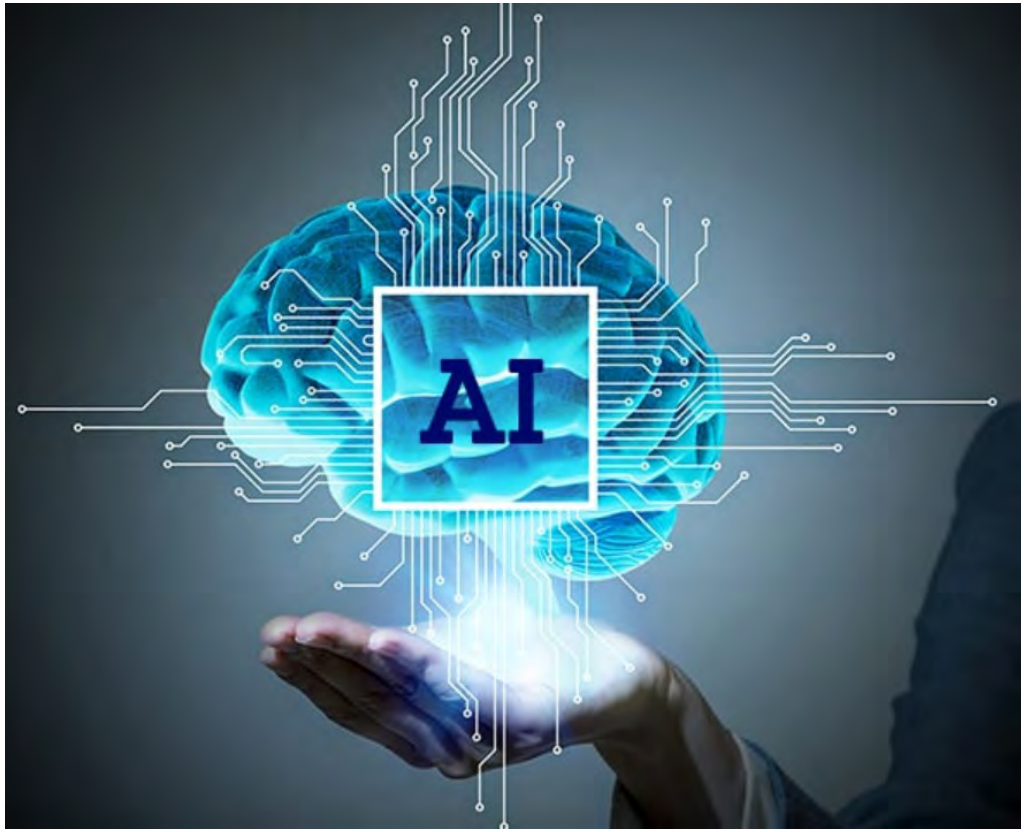


Artificial Intelligence

› AI premised to emulate HI (Human Intelligence)

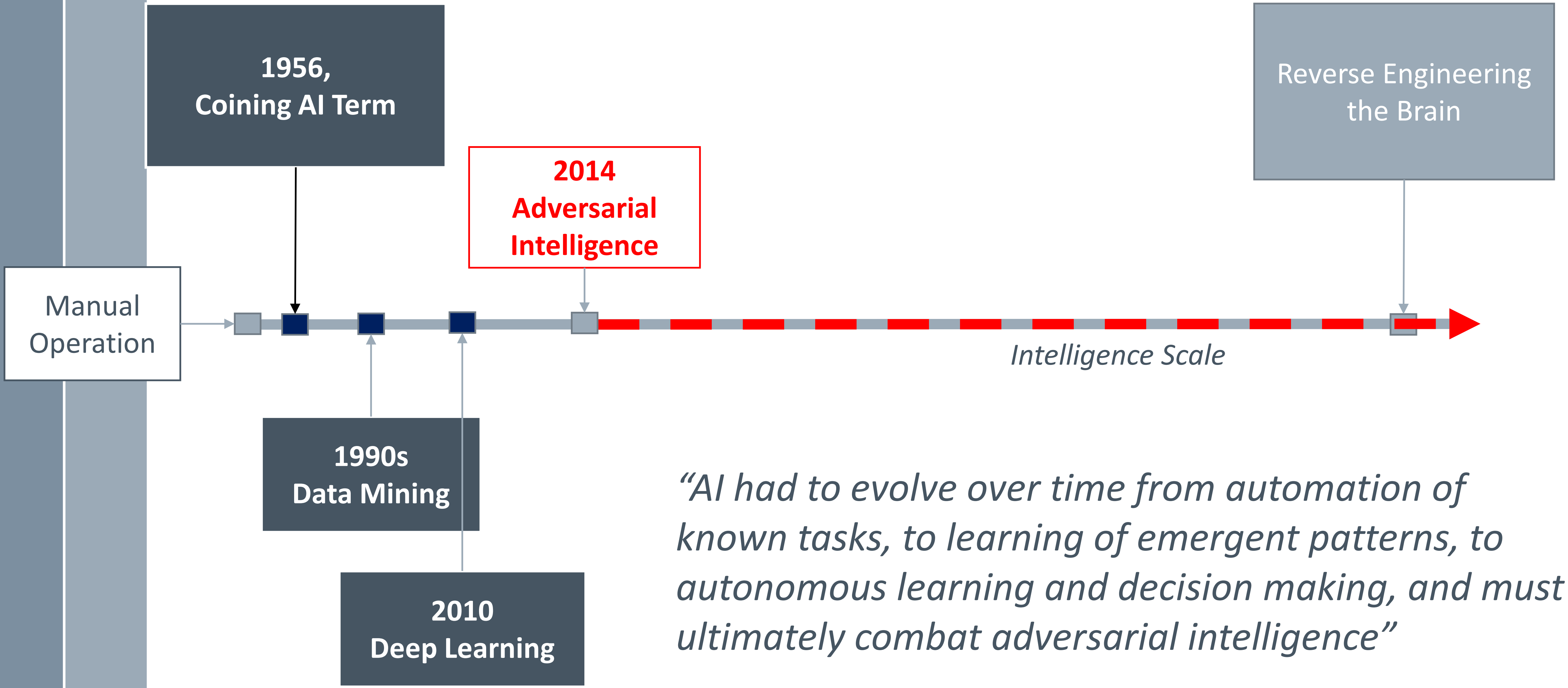
› Past decade has witnessed a huge comeback for AI in almost all sectors of science and engineering, due to:
  – Massive data
  – Advanced learning algorithms
  – Powerful computers

› Premised to render optimum, safe, secure operation for complex systems



https://www.usmsystems.com/top-45-artificial-intelligence-companies/

Towards Automating Intelligence



"AI had to evolve over time from automation of known tasks, to learning of emergent patterns, to autonomous learning and decision making, and must ultimately combat adversarial intelligence"

Current AI Trends

R&D mainstream is focusing heavily on **Predictive AI**, and more recently on **Explainable AI** (XAI) and **Adversarial AI** (AAI), and less on **Causal AI** – with ***passive*** application

- ❑ PAI: Discovering Association Rules
- ❑ XAI: Identifying Key Contributors to PAI
- ❑ AAI: Misleading PAI
- ❑ Causal AI: Distinguishing Cause from Effect

My Current R&D Focus

- Performance:
  - How to optimize process control under ***<u>uncharacterized</u>*** sources of uncertainties?
  - How to support training of computationally-intensive exercises, e.g., optimization, uncertainty analysis, etc.?
- Safety:
  - With huge data collected from operating reactors, how to derive ***<u>defendable</u>*** basis for inference?
  - For **<u>FOAKs</u>**, how to derive “defendable” low-uncertainty estimates of key performance parameters?
- Security:
  - For well-understood industrial processes, how to design ***<u>covert</u>*** defenses using ***<u>adversarial</u>*** AI?

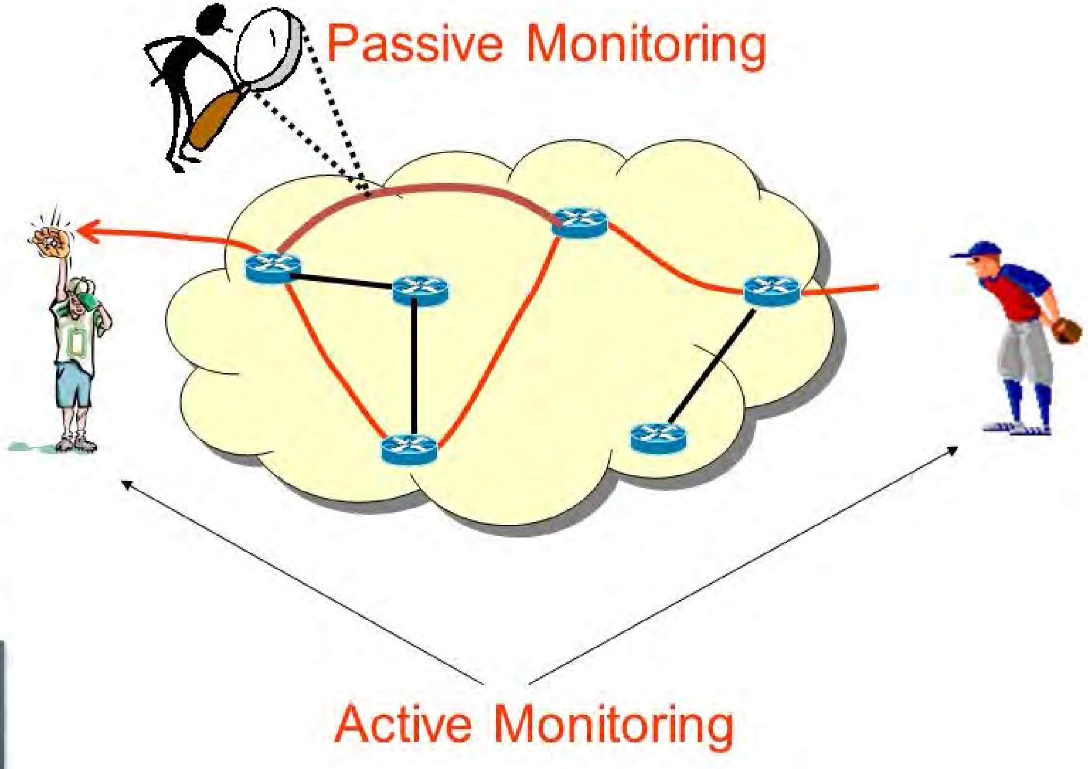
Passive vs. Active AI

PASSIVE

Monitoring without Interfering

ACTIVE

Interfering for better Monitoring



https://slideplayer.com/slide/7395148/

Active vs. Passive AI

To find out what happens to a system when you interfere with it,
you **have to interfere** with it
(not just **passively** observe it).

George Box,
“Use and Abuse of Regression,” Technometrics, Nov. 1966

Example for Active AI Project

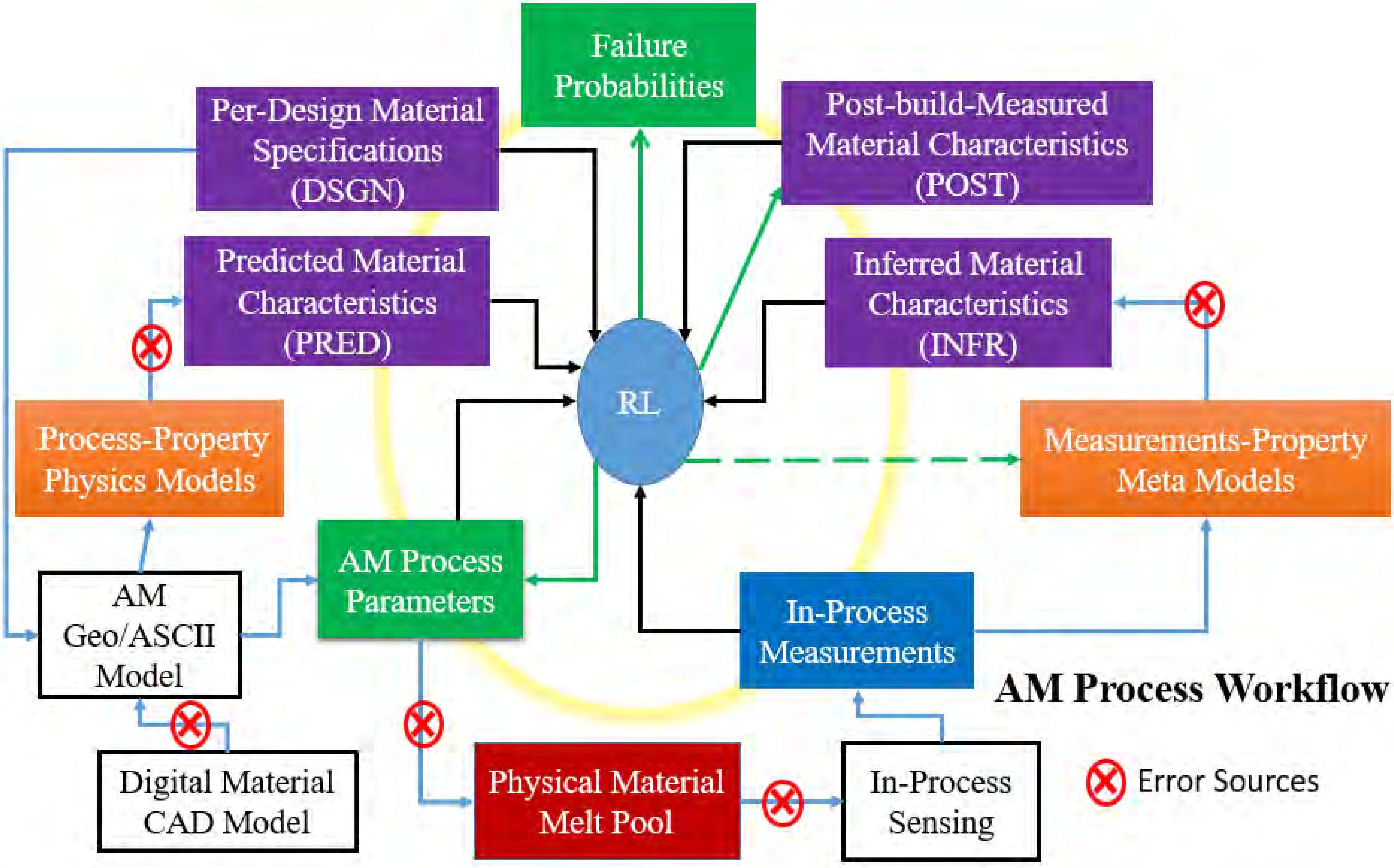
Reinforcement Learning Validation Framework for
Quality Assurance of AI-guided
Additive Manufacturing
Digital Platforms

DOE-NEUP 2020-2022, *in collaboration with*
John Sutherland and Xinghang Zhang (Purdue Univ.), and
Sherri Buchanan and Vincent Paquit (ORNL’s TCR team)

Reinforcement Learning

$$V(s) = \max_a \left(r(s,a) + \gamma \sum_{s'} p(s' \mid s,a) V(s') \right)$$

- Emulates human-like reward system to optimize actions
- Abstracts any system to live in multiple states, with actions transitioning system between states.
- Value function serves as importance measure for states
- Requires data-rich environment for training
- Works best for well-understood systems with no surprises
- Works well with model-based and pure data-driven settings.


Failure Probabilities
Per-Design Material Specifications (DSGN)
Post-build-Measured Material Characteristics (POST)
Predicted Material Characteristics (PRED)
Inferred Material Characteristics (INFR)
RL
Process-Property Physics Models
Measurements-Property Meta Models
AM Process Parameters
AM Geo/ASCII Model
In-Process Measurements
AM Process Workflow
Digital Material CAD Model
Physical Material Melt Pool
In-Process Sensing
Error Sources

END OF PRESENTATION

THANK YOU

For questions, please contact me at:

abdelkhalik@purdue.edu

Wrap Up / Next Steps

Clean. Reliable. Nuclear.

Curtis Smith, Ph.D.
Idaho National Laboratory
NS&T Division Director
Curtis.Smith@inl.gov