ADAPD:

Advanced Data Analytics for Proliferation Detection

Lawrence Livermore National Laboratory | Los Alamos National Laboratory EST. 1943 | Sandia National Laboratories | Oak Ridge National Laboratory | Pacific Northwest National Laboratory

DSI Workshop – August 8, 2018

Eddy Banks

LLNL – Jim Brase
LLNL – Eddy Banks
LANL – Jim Smith
SNL – Danny Rintoul
ORNL – Phillip Bingham
PNNL – Mark Greaves

LLNL-PRES-754931

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory | NNSA National Nuclear Security Administration

ADAPD – Goal: Early Detection of Low-Profile Proliferation

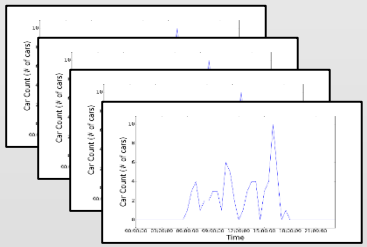
ADAPD will build science-based capabilities to *predict* and *detect* proliferation observables and *characterize* associated activities.

Data Science to the rescue:

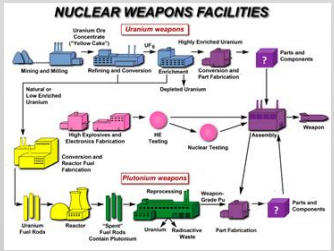
- Emerging analytic methods that integrate physics-based simulation models with experimental data have demonstrated significant increases in prediction performance for complex physical modeling applications.
- ADAPD will draw upon and extend these emerging methods to integrate physics and subject-matter-expert models with DOE/NNSA testbed data to improve prediction of proliferation activities.
- These methods will provide new capabilities for detection of low-profile proliferation observables in noisy, limited data environments with quantified uncertainties.

Advances in data science enable new approaches to low-profile proliferation detection

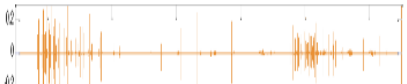
Knowledge, models, testbeds and activities



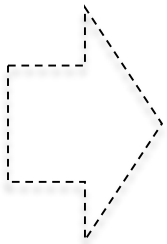
Testbed ground truth data

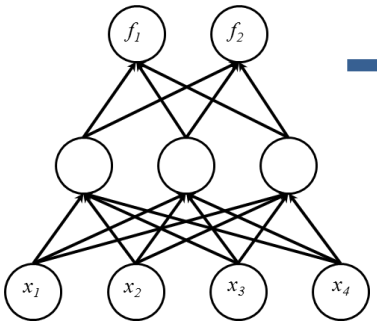


Testbed process simulations



"Learn the physics"





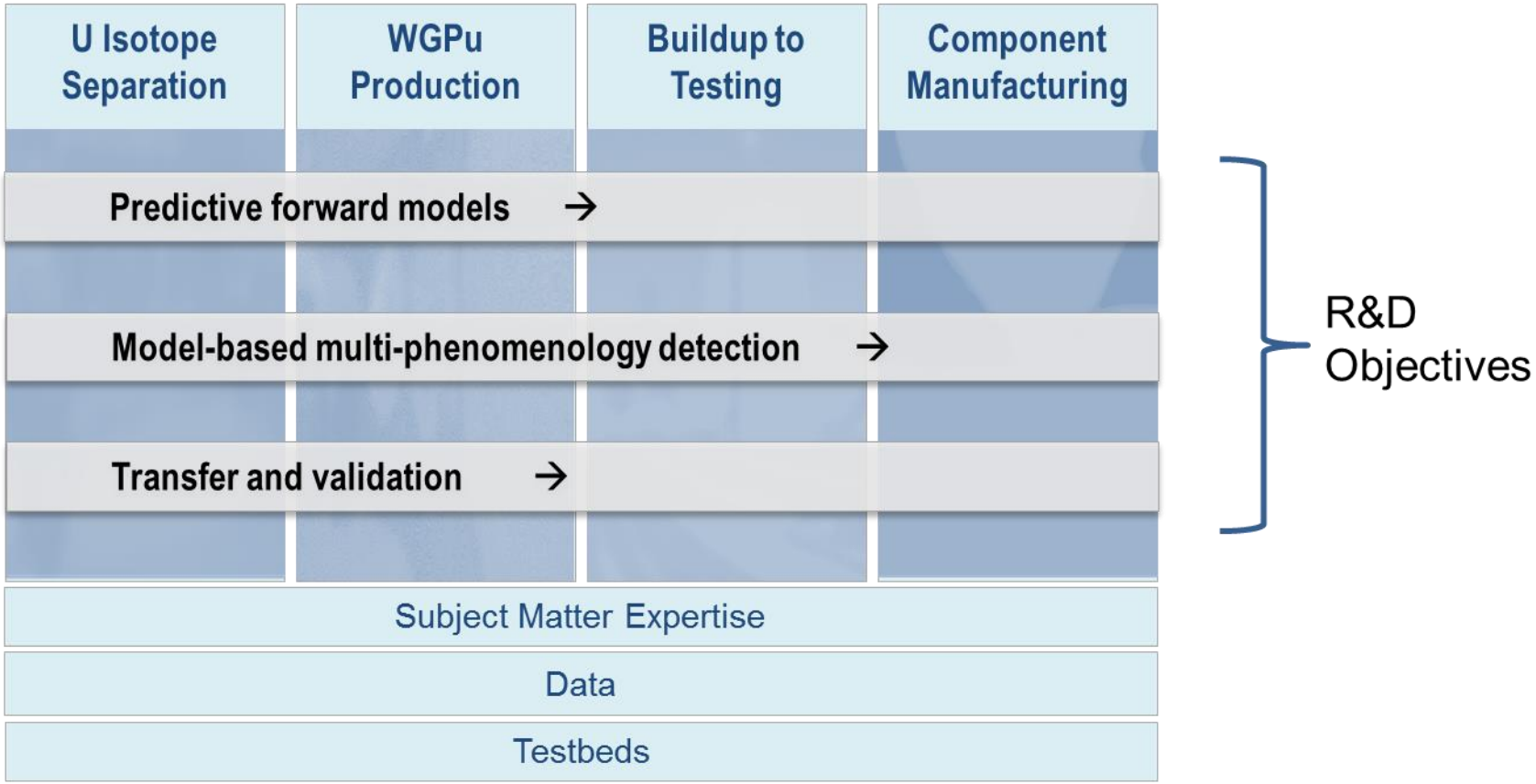
Train with combined facility data and SME-built simulations

Re-train the top layers using limited observable data

Prediction of observables in the new environment

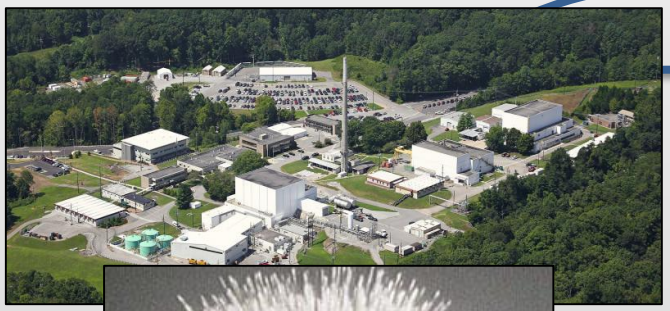
Query the model with new process state and environmental inputs

ADAPD – R&D Objectives



Approach: Predictive models of proliferation observables driving model-based detection

DOE/NNSA testbeds



Objective 1 – Combine SME-built with data-driven models for predicting multiple observables

Objective 2 – Multi-phenomenology detection for complex/ varied environments

Objective 3 – Model validation and transfer to new Low-Profile environments

Predictive models integrated into multisource analytics are essential to limited data applications

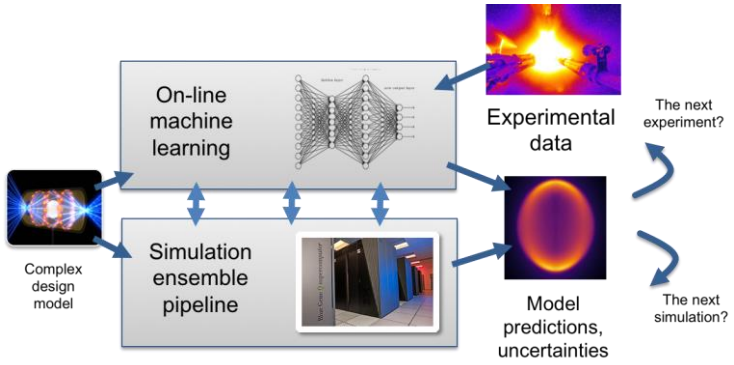
- SPE
- REDC/HIFR
- …

- Acoustic, seismic
- Optical
- Facility power
- Materials procurement and shipping patterns
- Patterns of life
- …

- Integrating direct and indirect data in the context of a predictive model will enable detection in the noisy, limited data environments and extend capabilities for quantitative estimates of rates and uncertainties

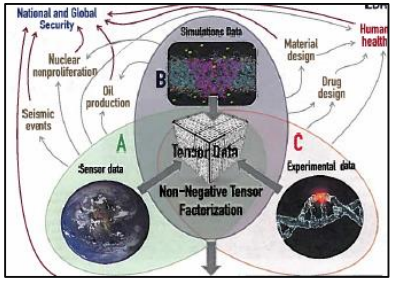
To detect Low-Profile Proliferation, ADAPD will address key data science gaps

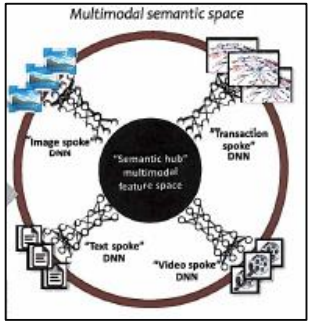
Predictive models integrating physics and data-driven approaches



Collaboration with DP efforts of LLNL and LANL

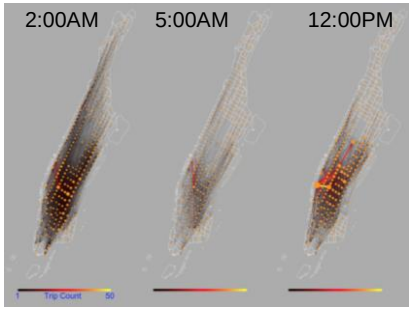
Multi-source data integration





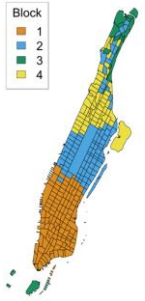
- Deep multimodal learning methods
- Sparse and noisy data – tensor models
- SME model-based transfer learning

Scalable network and graph analytics



Aggregated trip counts between Manhattan census tracts

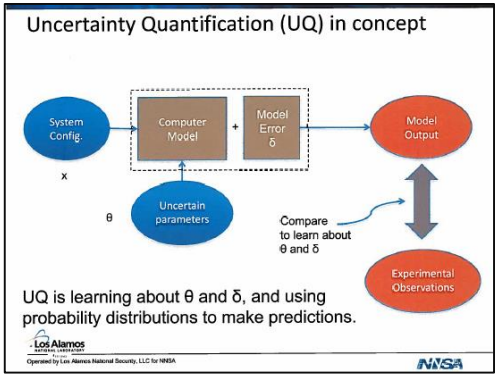
November 2 2013 (marathon day)



Partition based on temporal activity

- Pattern of life analytics
- Temporal activity learning
- Rare event and anomaly learning
- Learning from sparse data

Uncertainty quantification in data-driven systems



- Models learned from data
- High dimensional wave forms, spectra, images
- Accountability from inference back to data

ADAPD will be driven by a set of science questions

Objective 1: Predictive modeling –Can we predict the broad range of observables associated with Low-Profile Proliferation?

- Can we combine physics, SME, and data-driven models to accurately predict a broad range of observables?
- What observations and experiments on existing testbeds are needed to develop and validate predictive models?

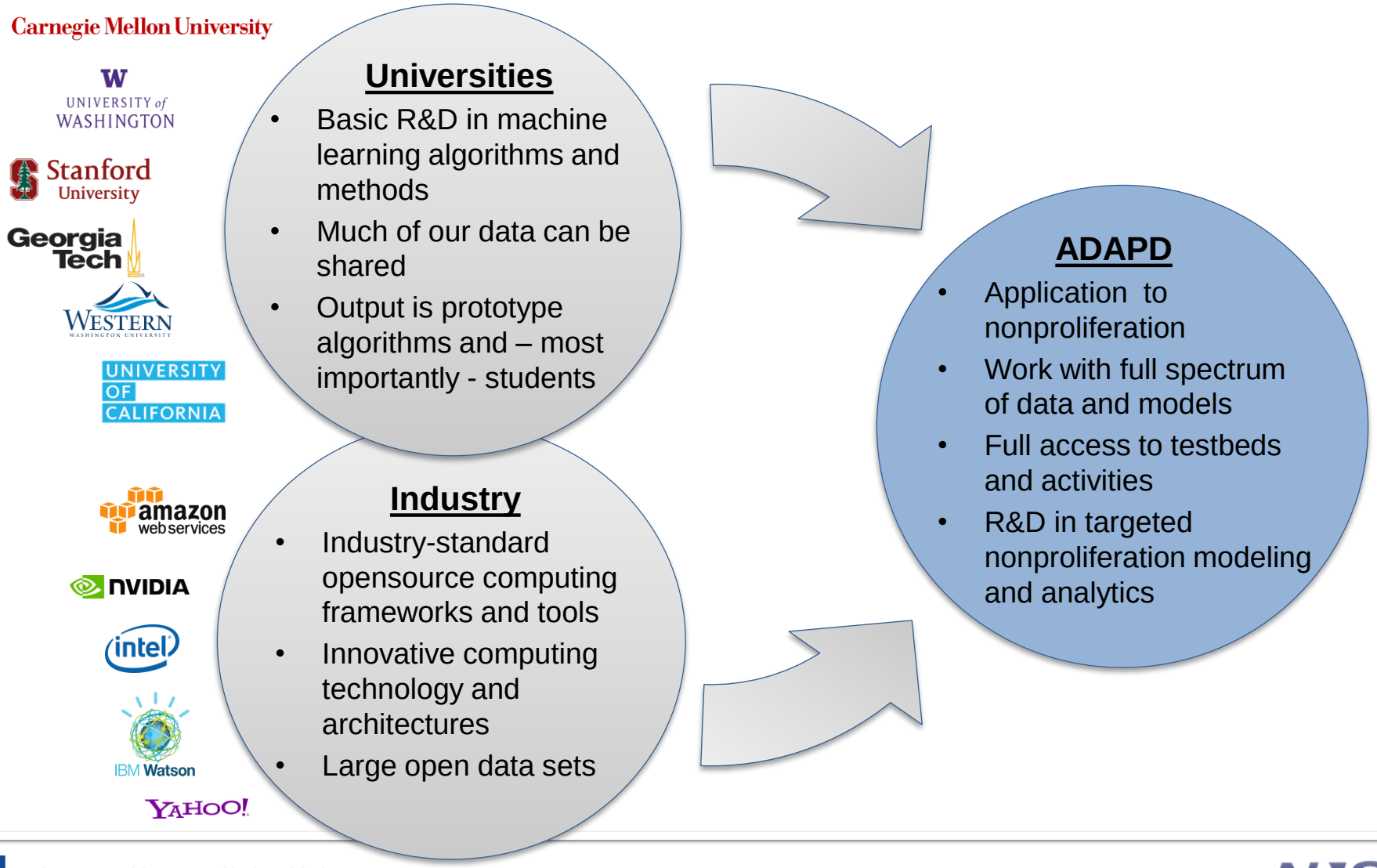
Objective 2: Multi-phenomenology detection and characterization – Can we use multiple observables to detect and characterize proliferation activities?

- Can we use predictive models to improve performance with noisy and sparse data?
- Can we quantify uncertainty in detection results that are based on multiple uncertain observables?

Objective 3: Validation and extension to new environments – Can we transfer a learned model to new geographical locations, different infrastructures, ….

- What tests should be performed on existing testbeds to best validate and calibrate models?
- Can we combine models for proliferation activities involving multiple facilities and sites?

Industry and university partnerships are an important component of the ADAPD plan



The ADAPD R&D Roadmap focuses on building cross-cutting and sustainable capabilities

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8
Predictive modeling	• Predictive model development – baseline scenarios		• Predictive model development –expanded observables – both direct and indirect		• Full UQ integration with integrated models		• Apply & evaluate performance for partially unknown proliferation scenarios	
Multi-phenomenology detection		• Detection and characterization development – baseline methods		• Detection and characterization development – inverse inference methods		• Uncertainty quantification for inverse inference models		• Detection & characterization performance enhancement in high-noise, sparse and missing data environments
Transfer to new environments		• Develop synthetic data approaches to enhanced model generalization		• Test and evaluate transferability of learned models to new data sets on the same testbed		• Test and evaluate transferability of learned models to partially unknown or limited data environments		• Test and evaluate transferability of learned models to new testbeds & environments
Testbed and Experiment Partnerships		• Collaborative testbed exp 1 training set development • Collaborative experiment 2		• Indirect data planning and collection campaigns • Evaluate testbeds and experiments for transfer experiments			• Validate predictive models against extended range of testbed experiments	
Data Management and curation	• Assess existing data sets • Catalog and curate existing data sets		• Catalog and curate new data sets		• Develop uniform cataloging, curation, and control across DNN			

ADAPD – Questions?