VIRGINIA TECH™

Hume Center for National Security and Technology

Problematic AI Panel:

What is AI, how is it used and how can it produce erroneous results?

Laura Freeman, Ph.D.
Deputy Director, National Security Institute
Research Associate Professor, Statistics
Assistant Dean for Research, College of Science

nationalsecurity.vt.edu

[Incomplete] Artificial Intelligence Taxonomy

VIRGINIA TECH

Reasoning

Programmed

- Planning/Logic/Expert Systems
- Optimization
- Agent-Based Systems

Learning

Supervised

- Bayesian Networks
- Regression
- Neural Networks

Reinforcement

- Genetic Algorithms
- Markov Decision Processes

Unsupervised

- Clustering
- Principal Component Analysis

Deep Learning

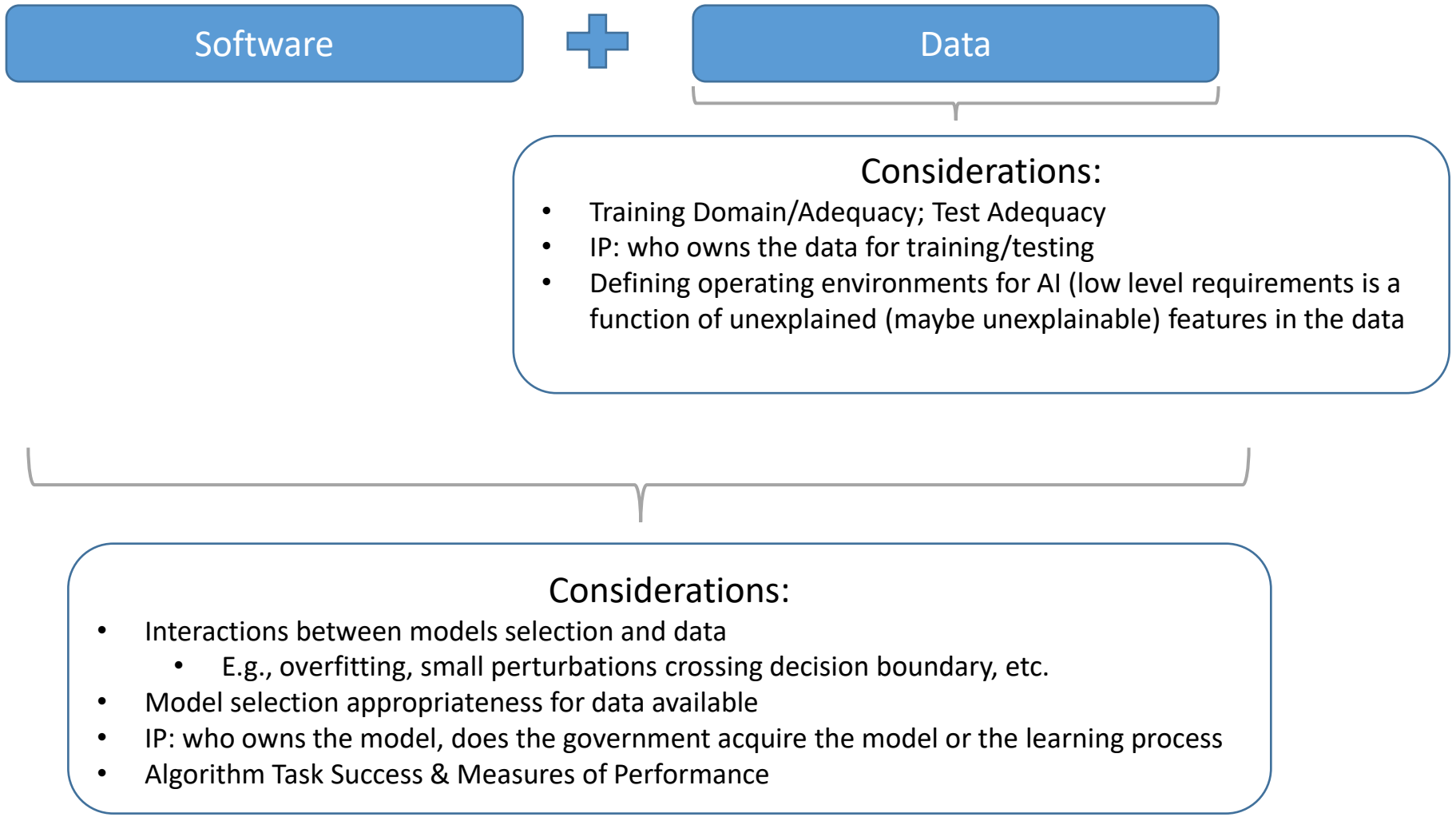
- Convolutional Neural Networks
- Recurrent Neural Networks
- Deep Q Learning
- Autoencoders
- Generative Adversarial Nets

Select National Priorities in AI Research

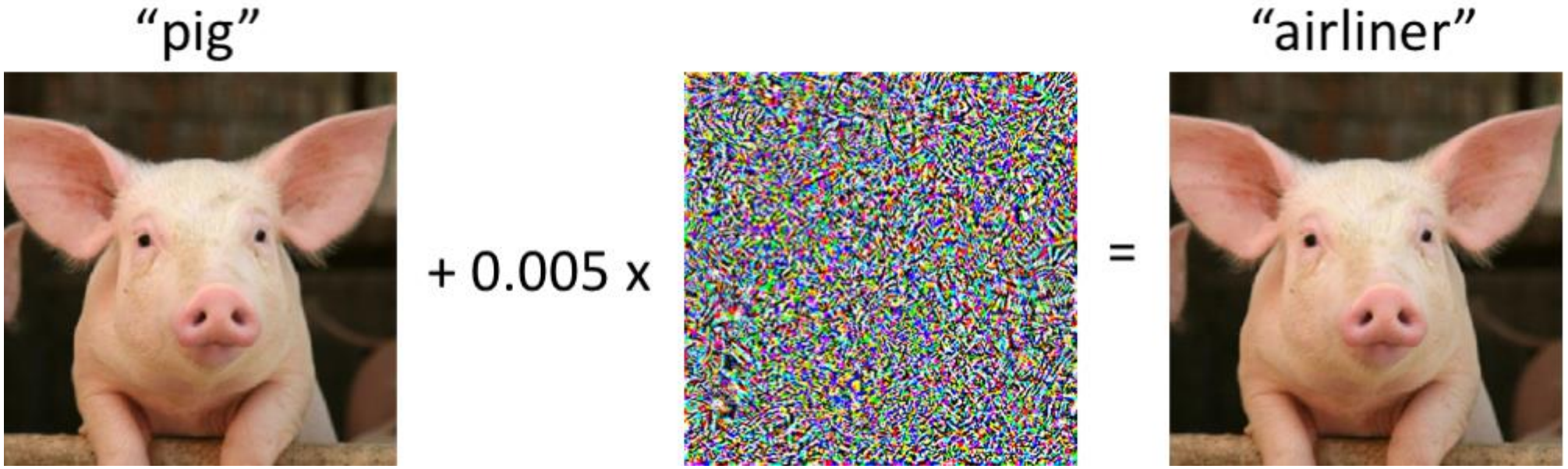
	Office of Scientific and Technical Information - Department of Energy	Department of Defense	National Institute of Standards and Technology
Fundamental / Foundational	Accelerating research insights	"adopt AI principles that reflect our nation's values of a free and open society"	Theme: Measure and enhance the trustworthiness of AI Systems
	Domain Aware	1) Responsible	Secure
		2) Equitable	Objective/Accurate
	Interpretability	3) Traceable	Explainable
	Robustness	4) Reliable	Reliable/Robust
		5) Governable	
Applied / Capability	Scientific data analysis	Key Missions – Situational Awareness, Safety, etc.	Revolutionizing metrology at NIST from experiment design to research results
	Enhance modeling and simulation		
	Management & control of complex systems	Complex System Acquisition	

Current Wave of AI – Software + Data

- Statistical Learning Based Machine Learning and Artificial Intelligence

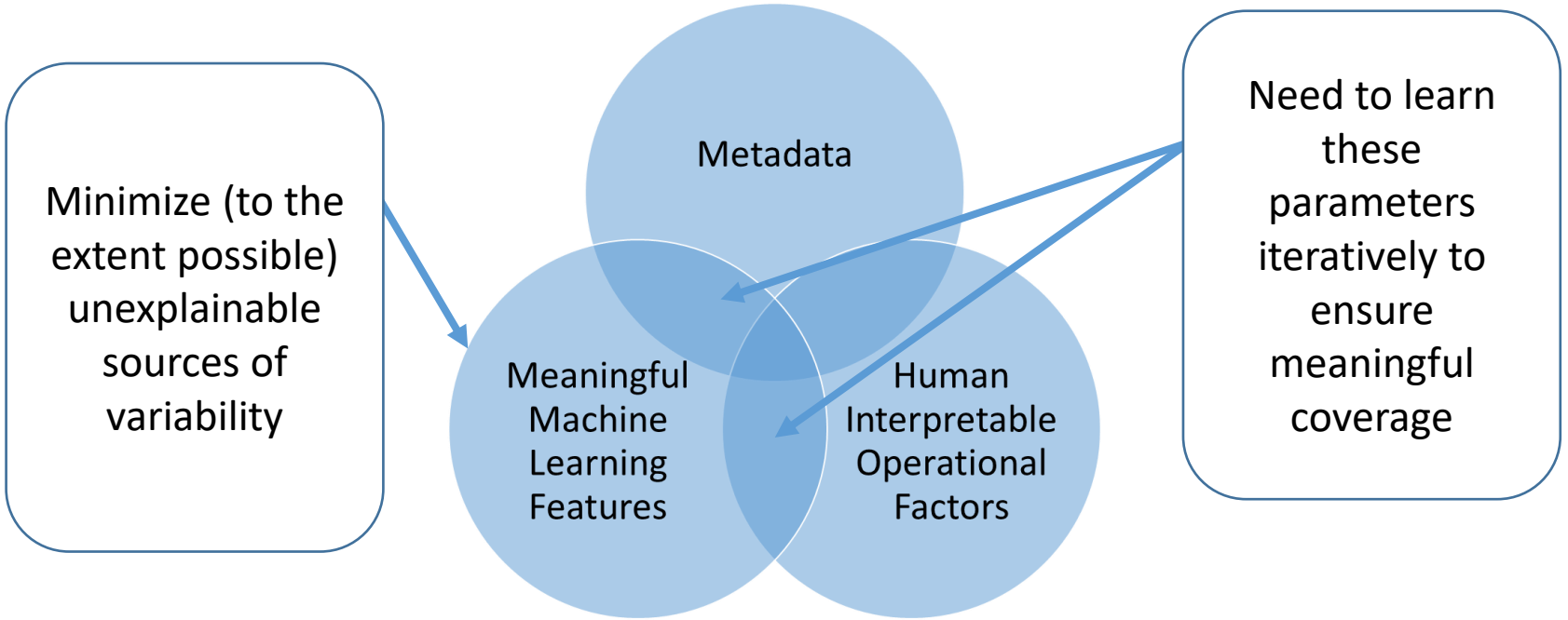


What I am not going to focus on!



Challenge 1: Operating Domains & Explainable AI

- Data defines the operating envelope for an AI-enabled systems
- Humans have to make decisions on when and where to deploy AI



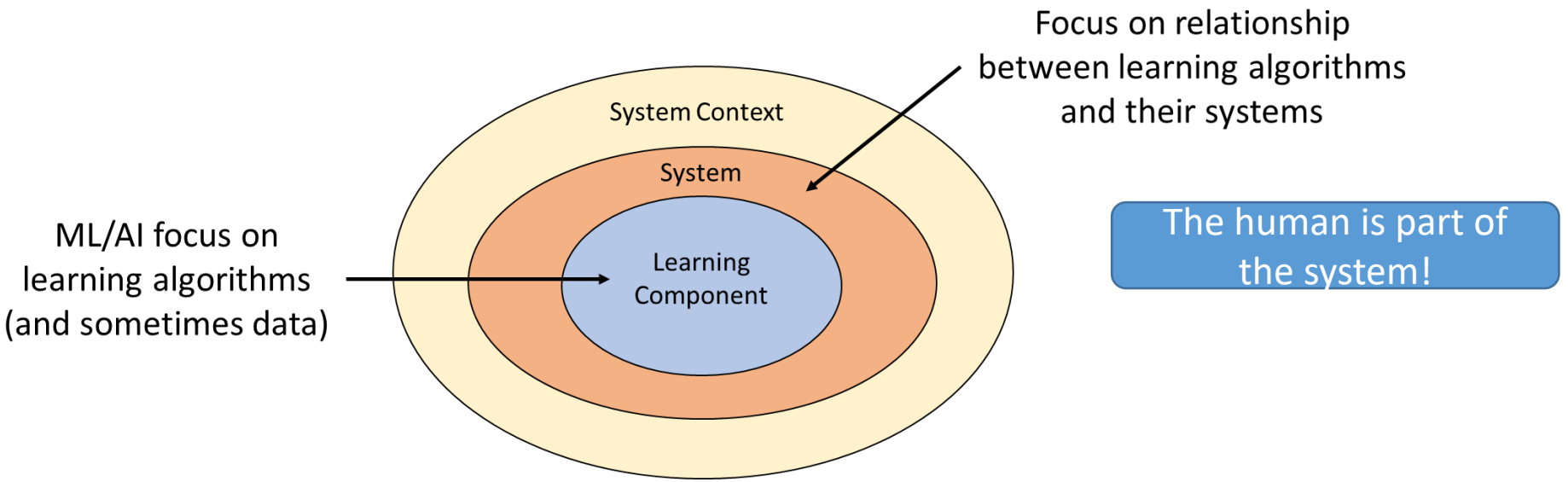
Erroneous results occur when there is a mismatch between what the algorithm is learning and what the human thinks the algorithm is learning.

Challenge 2: Human-System-AI Interactions

Observation

Artificial Intelligence (AI) and engineered systems are **coupled**

Reliability, V&V, prognostics, etc. for AI
cannot be divorced from system (including the human)



Cody, T. “Mesarovician Abstract Learning Systems”. International Conference on Artificial General Intelligence. Springer (2021). *In publishing.*