LIDS

LABORATORY FOR INFORMATION & DECISION SYSTEMS

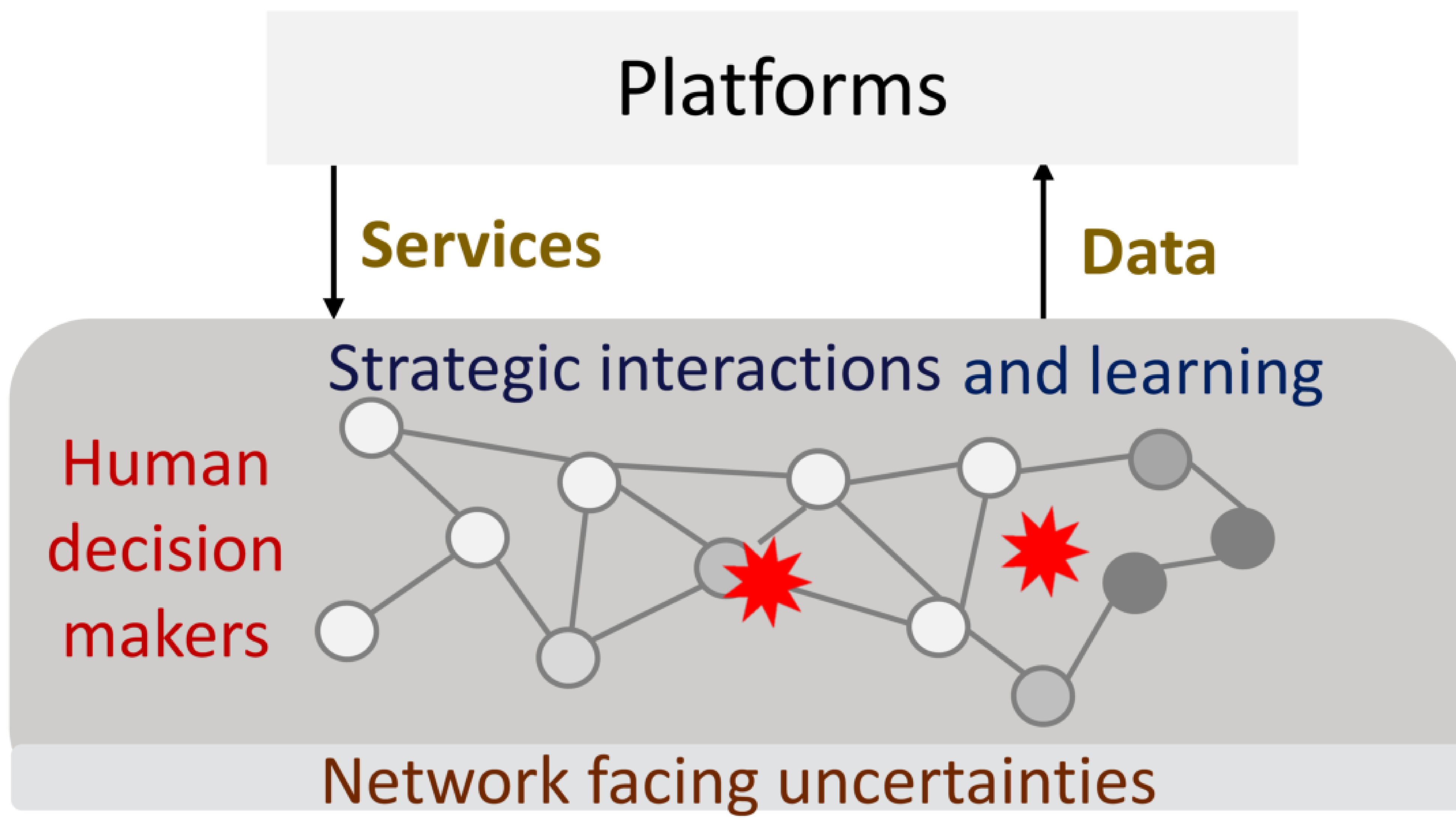
# LIDS Research Groups

## 2024-2025

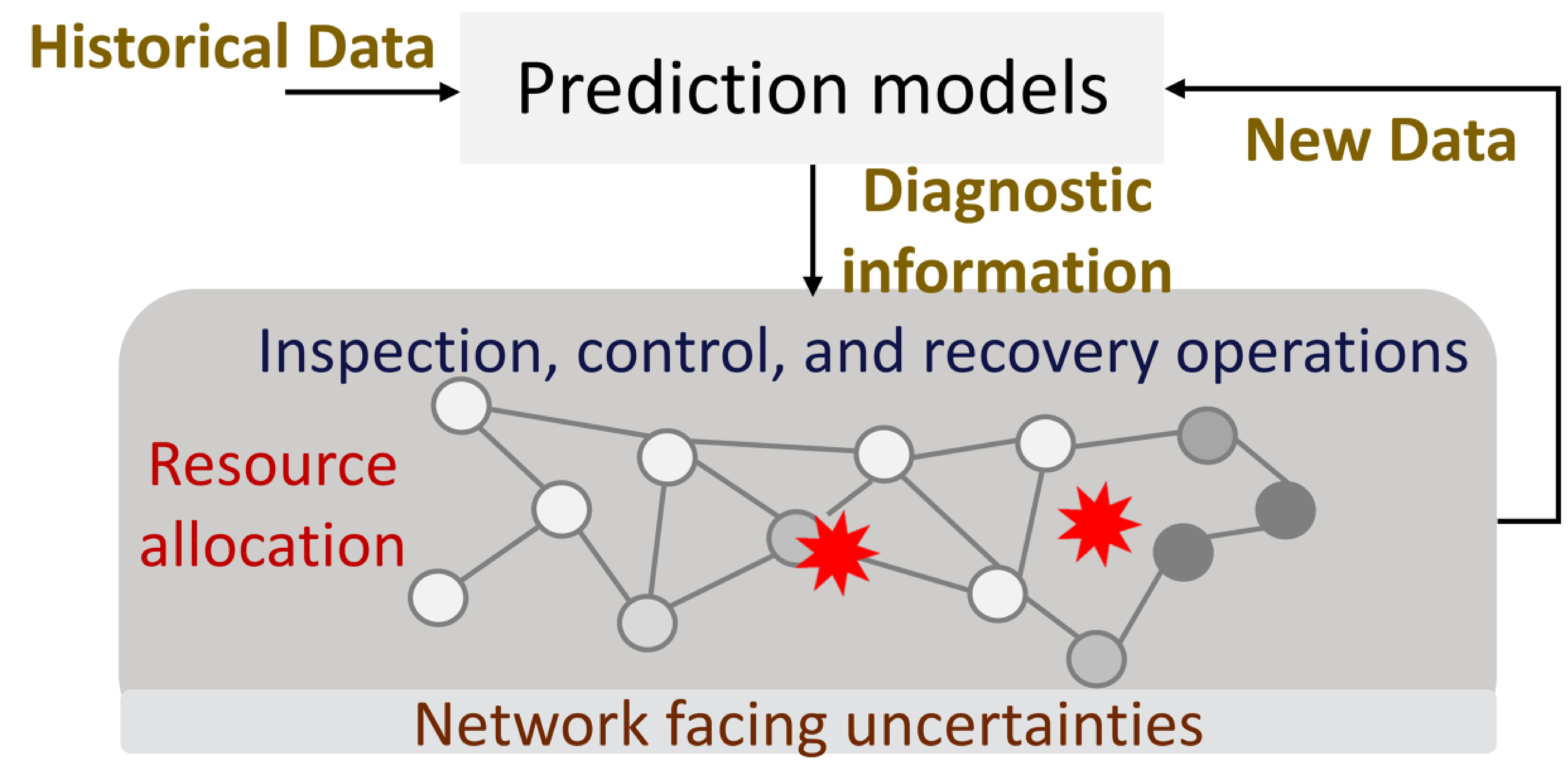
LIDS
RESEARCH GROUPS
2024-2025

# Saurabh Amin

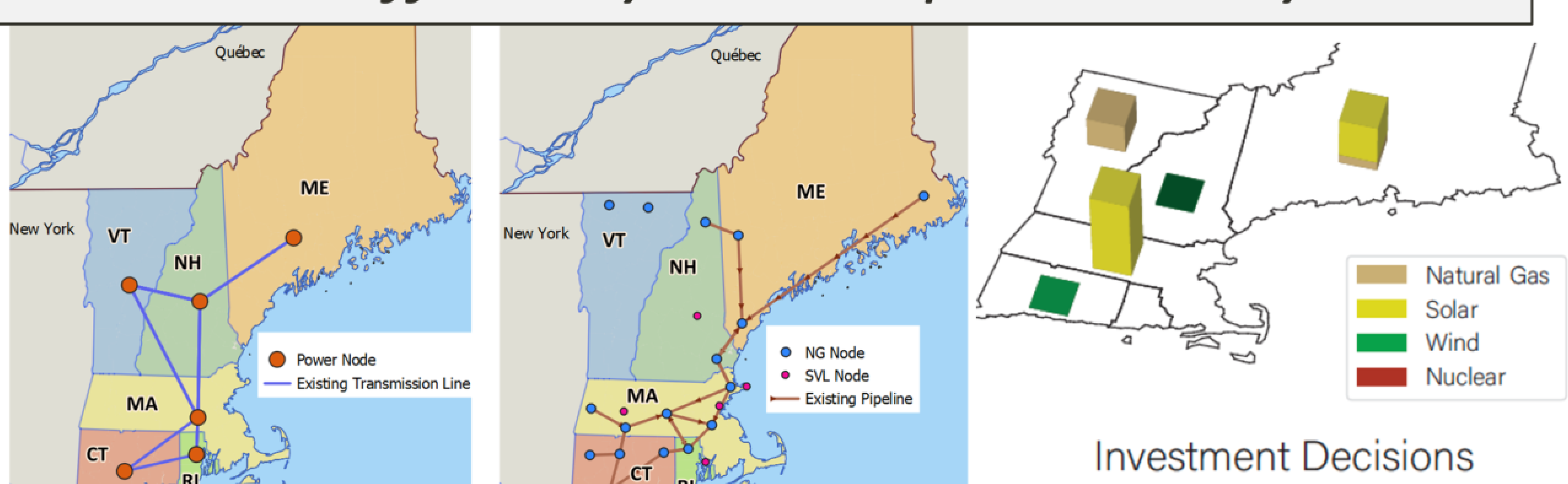
- **Problems:** Energy and Transportation Infrastructure Systems: Climate resilience, Disaster response and recovery, Sustainability and decarbonization, Cyber-physical security
- **Tools:** Stochastic control and optimization, Game theory, Data-driven optimization in networks
- **Solutions:** Control and monitoring strategies, Information design, Incentive mechanisms




*Strategic learning, incentive design, and operational efficiency in transportation systems*




*Predictive and prescriptive analytics for resilience to extreme events and security failures*



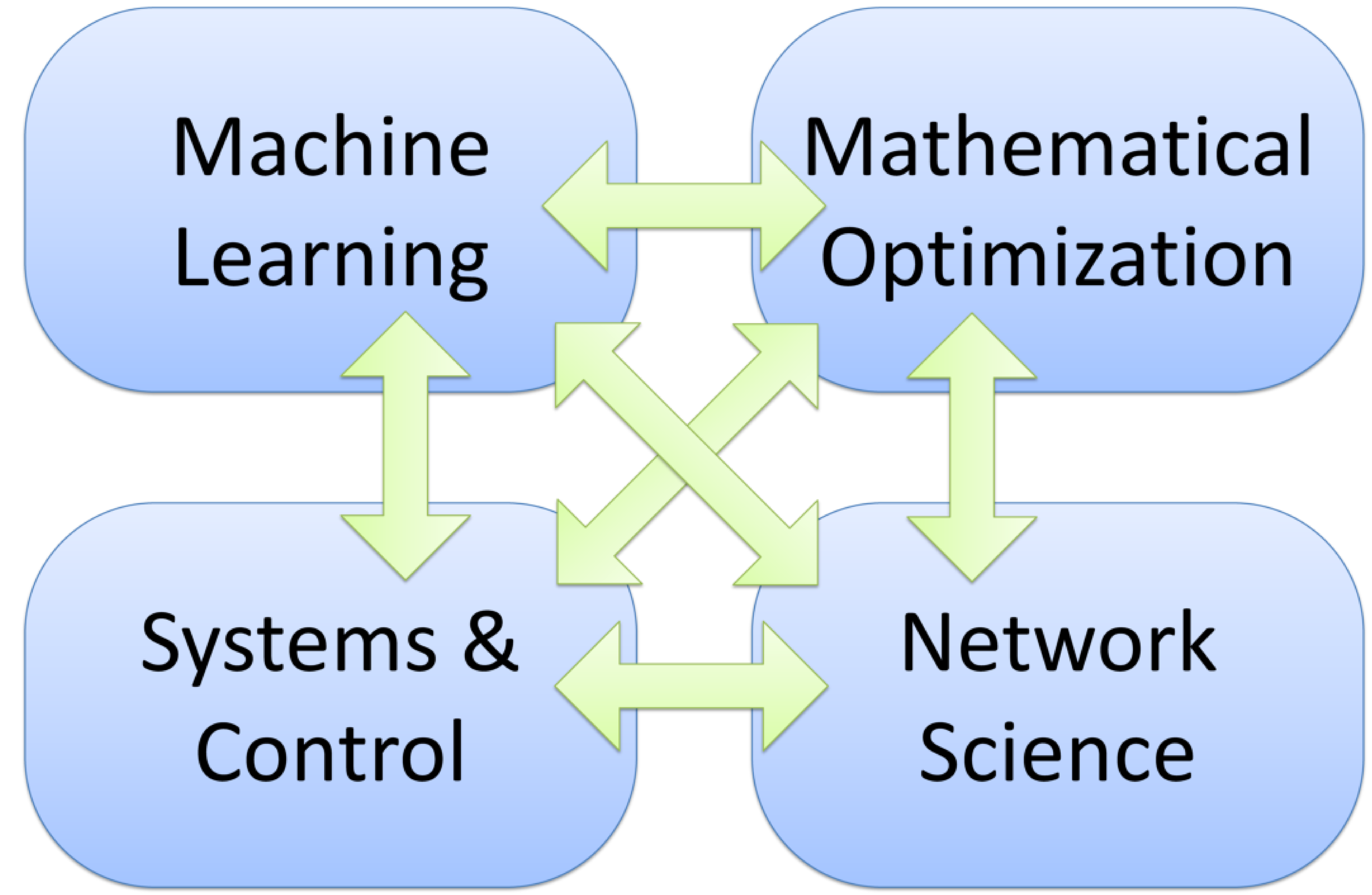

*Data-driven optimization for decarbonized, resilience & equitable energy systems planning*


*Intervention design and network monitoring for sustainability of natural-resource supply chains*

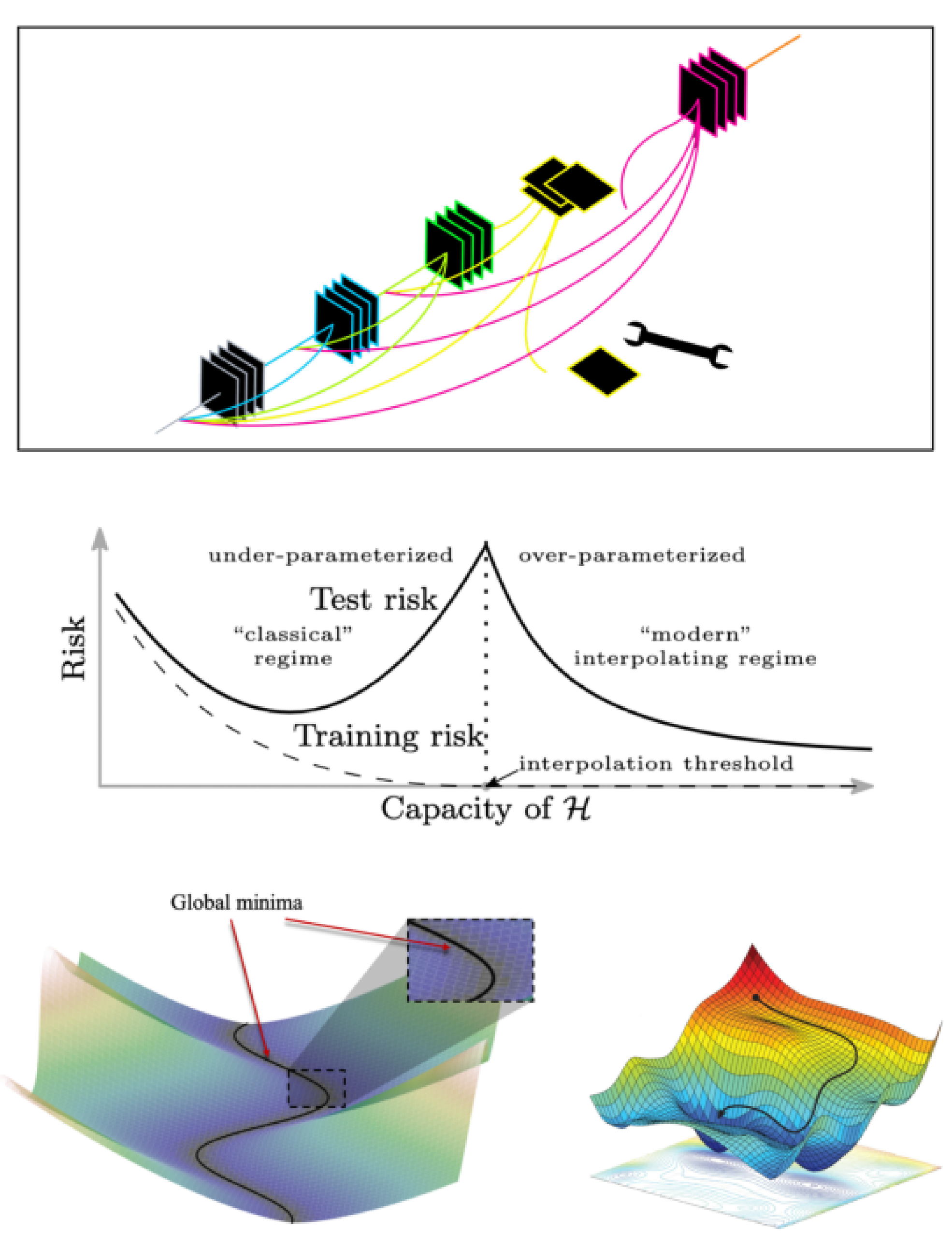
LIDS
RESEARCH GROUPS
2024-2025

# Navid Azizan

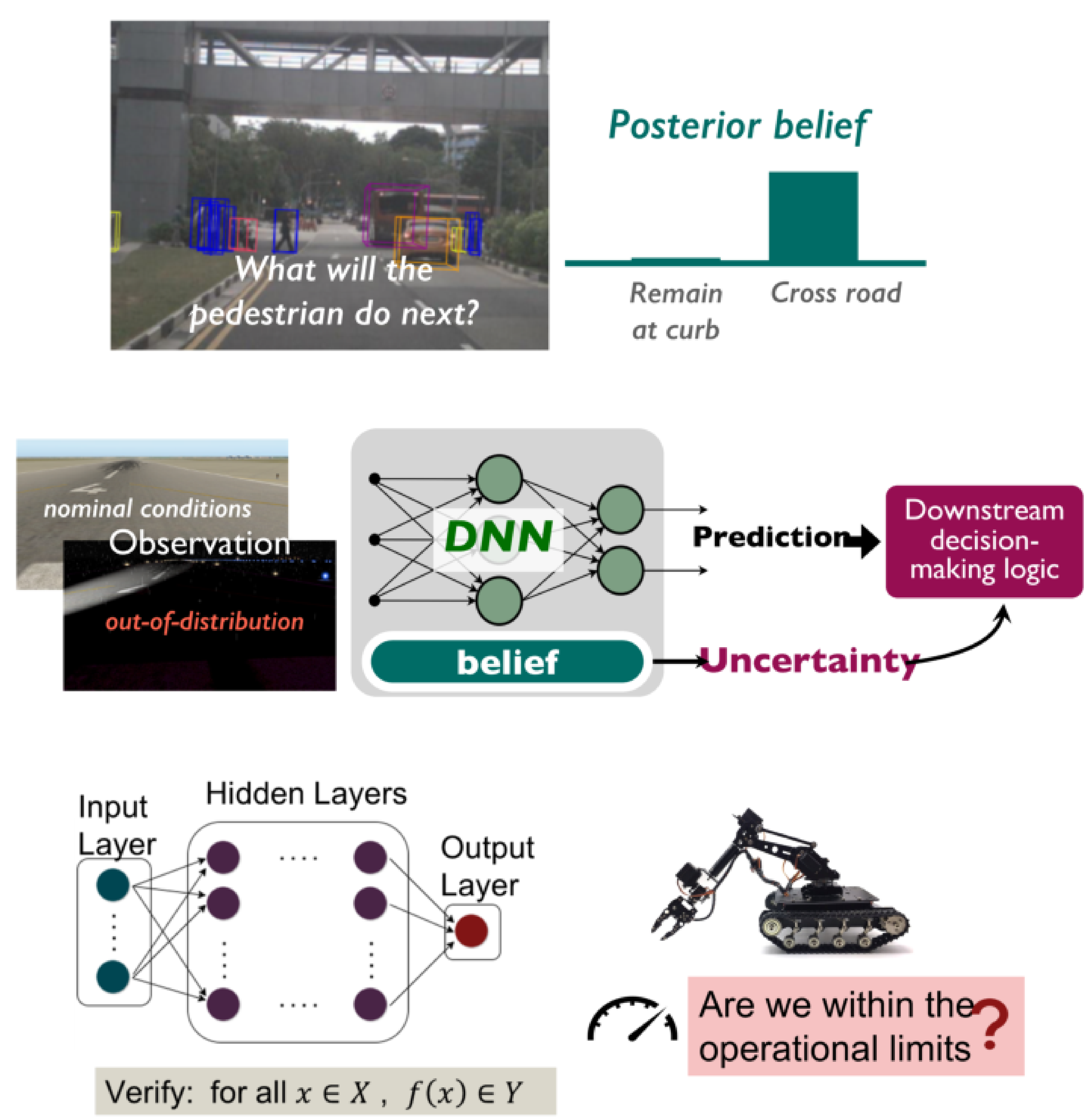
azizan.mit.edu

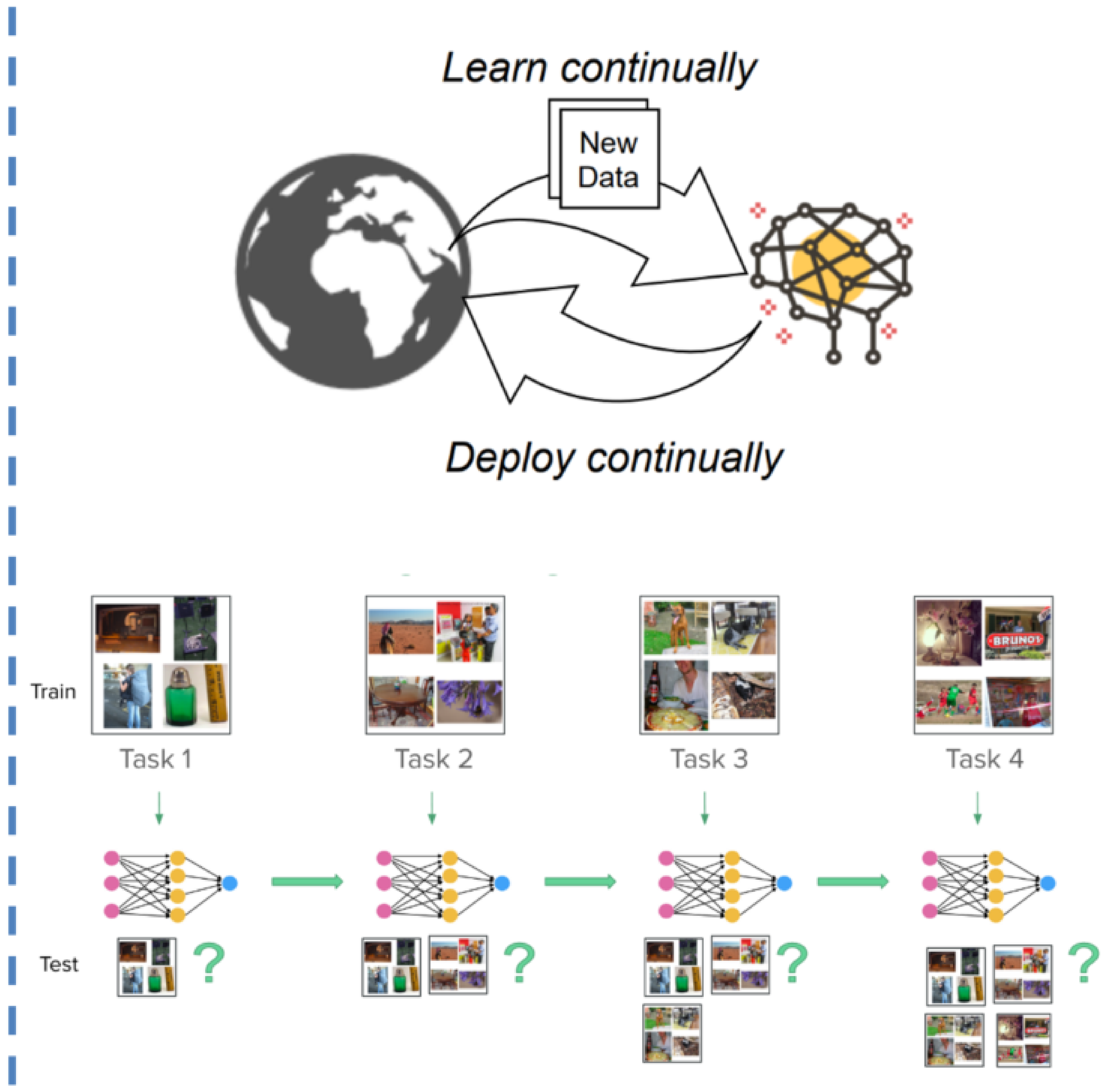

**Research goal**: To develop theoretical foundations and practical methodologies for realizing large-scale intelligent systems that can learn and operate safely, autonomously, and efficiently.



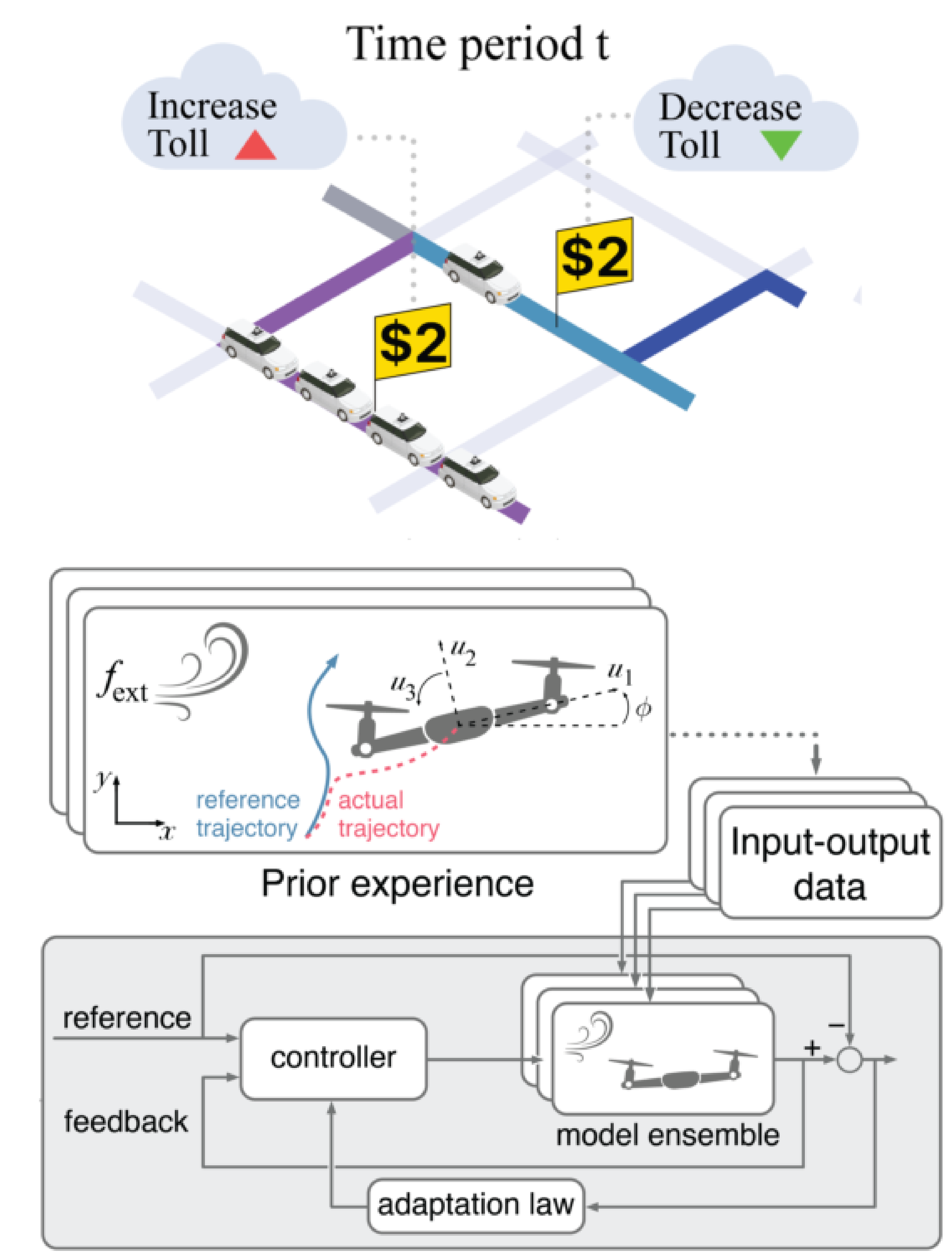

Foundations of robust deep learning




Safety assurances for machine learning




Continual & lifelong learning

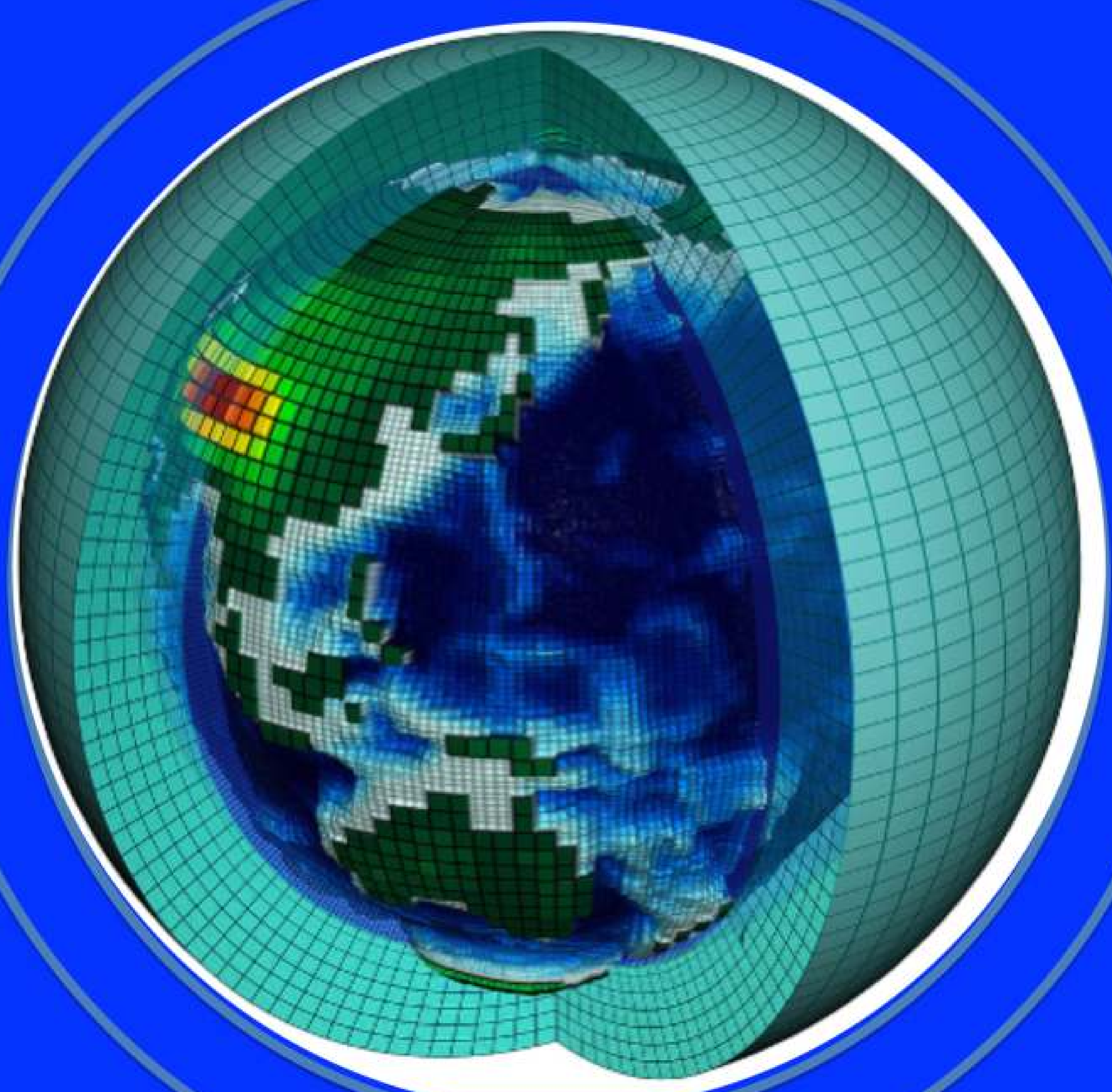



Learning for control & decision making

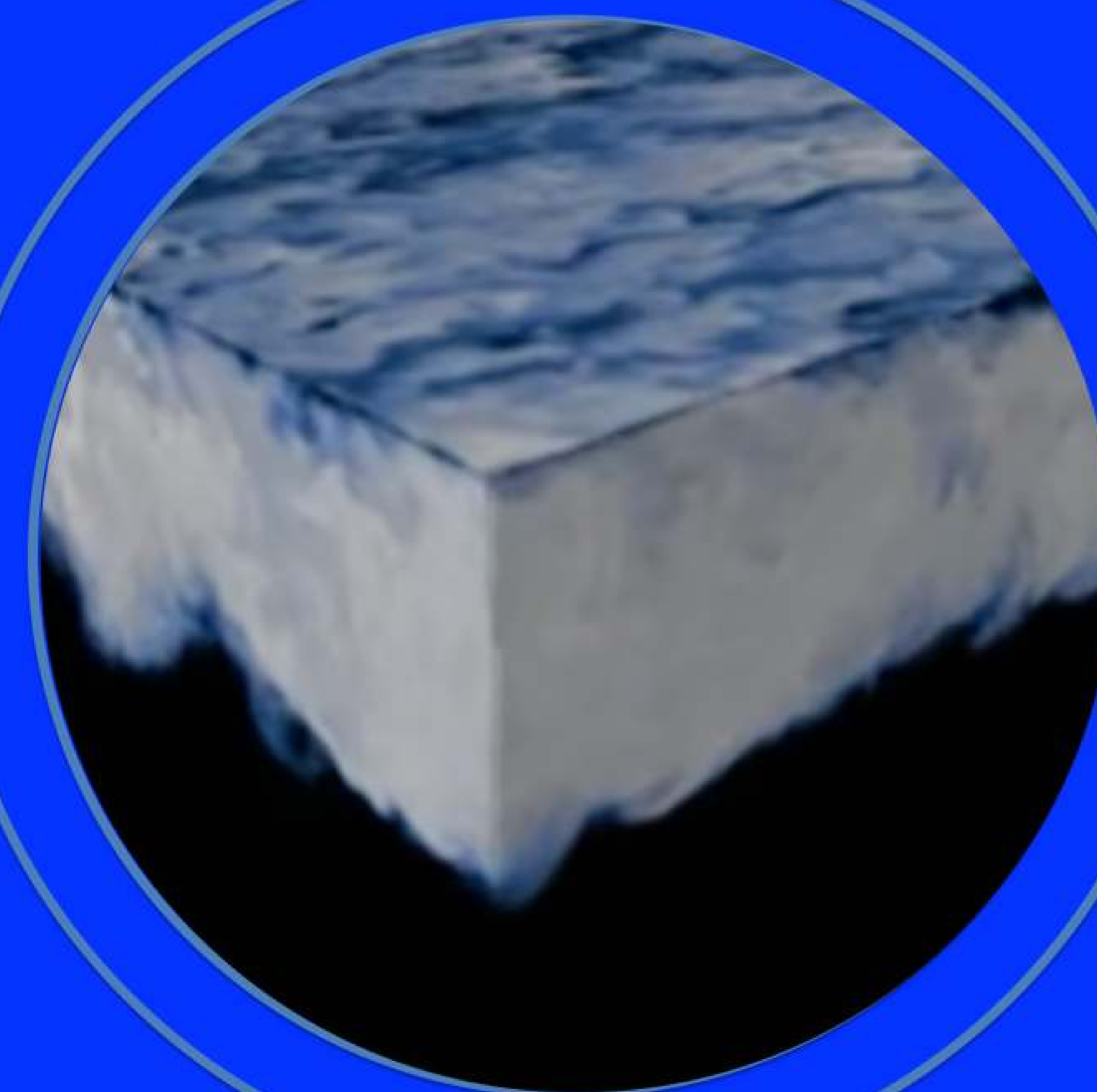
LIDS
RESEARCH GROUPS
2024-2025

# Abigail Bodner, EAPS & EECS

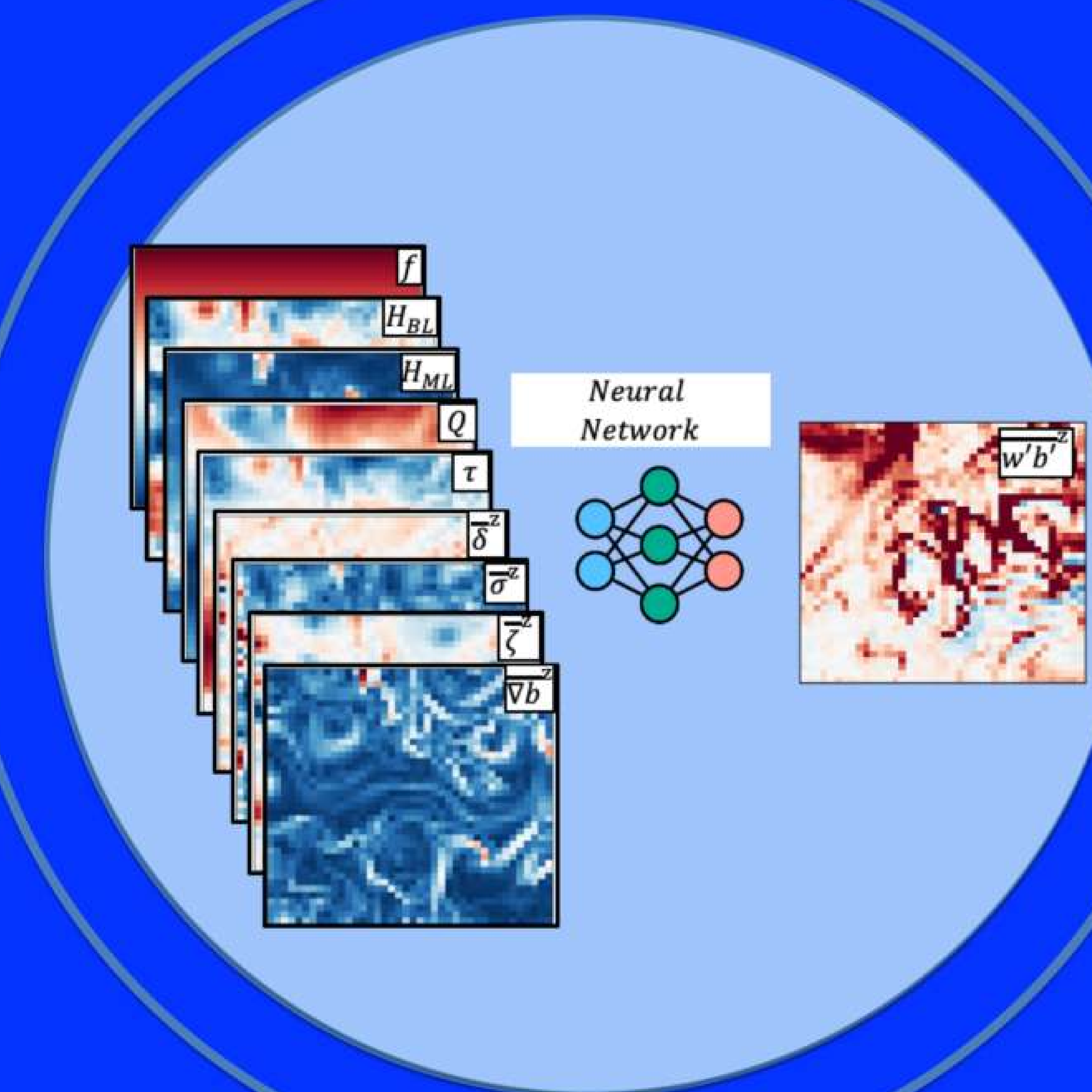
Investigating turbulence in the upper ocean using a combination of theory, observations, climate models, and machine learning.



Climate modeling



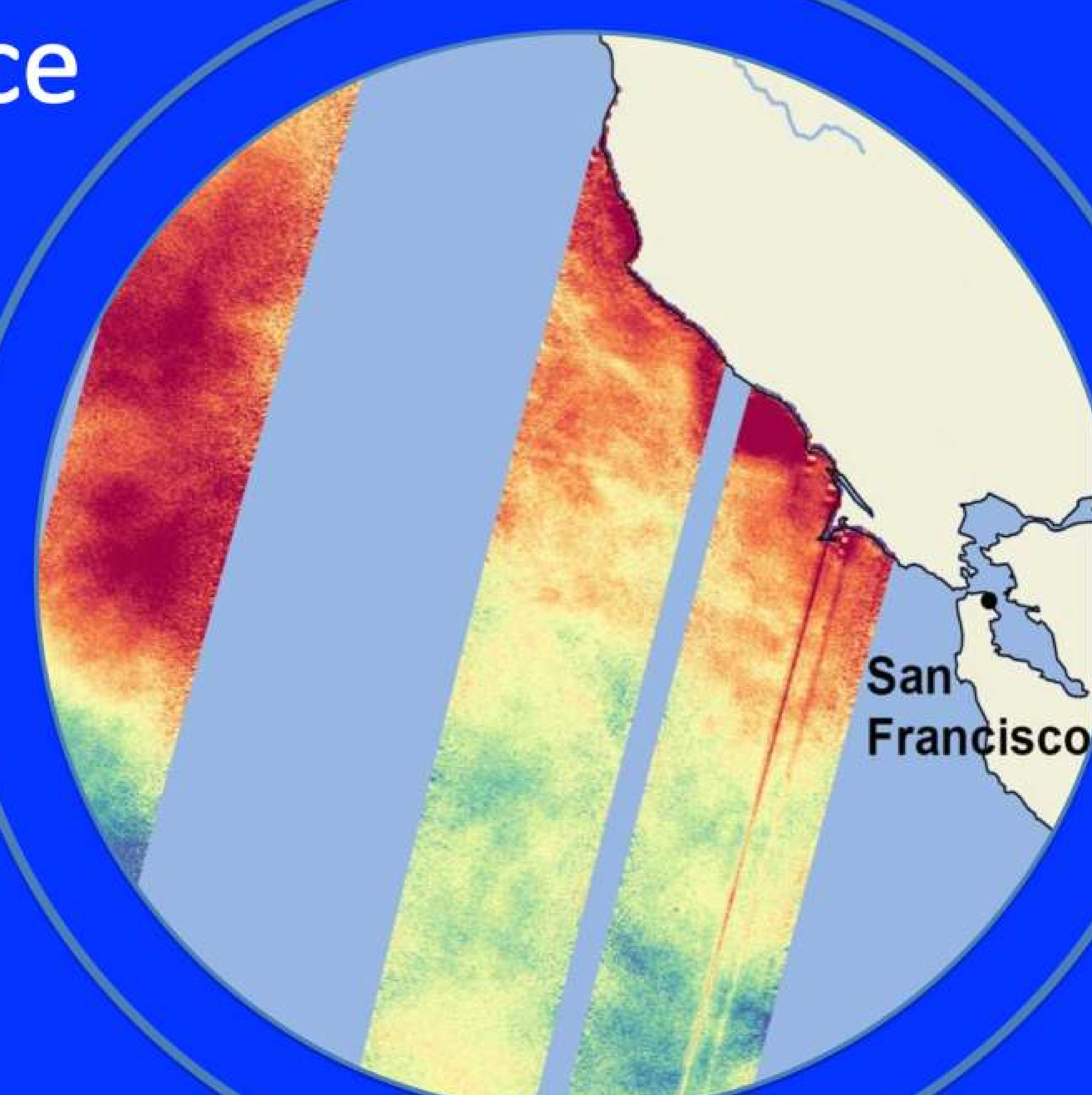
Turbulence




Machine learning

$$\underbrace{\frac{DE}{Dt}}_{\text{change in TKE}} = \underbrace{-\overline{\mathbf{u}'w'}\cdot\frac{\partial\overline{\mathbf{u}}}{\partial z}}_{\text{shear production}} + \underbrace{\overline{w'b'}}_{\text{buoyancy production}} \underbrace{-\frac{\partial}{\partial z}\left(\overline{w'E}+\frac{1}{\rho}\overline{w'p'}\right)}_{\text{transport of TKE}} - \underbrace{\epsilon}_{\textit{Dissipation}}$$

Geophysical fluid dynamics




Ocean satellite observations

LIDS
RESEARCH GROUPS
2024-2025

# Energy Analytics Group

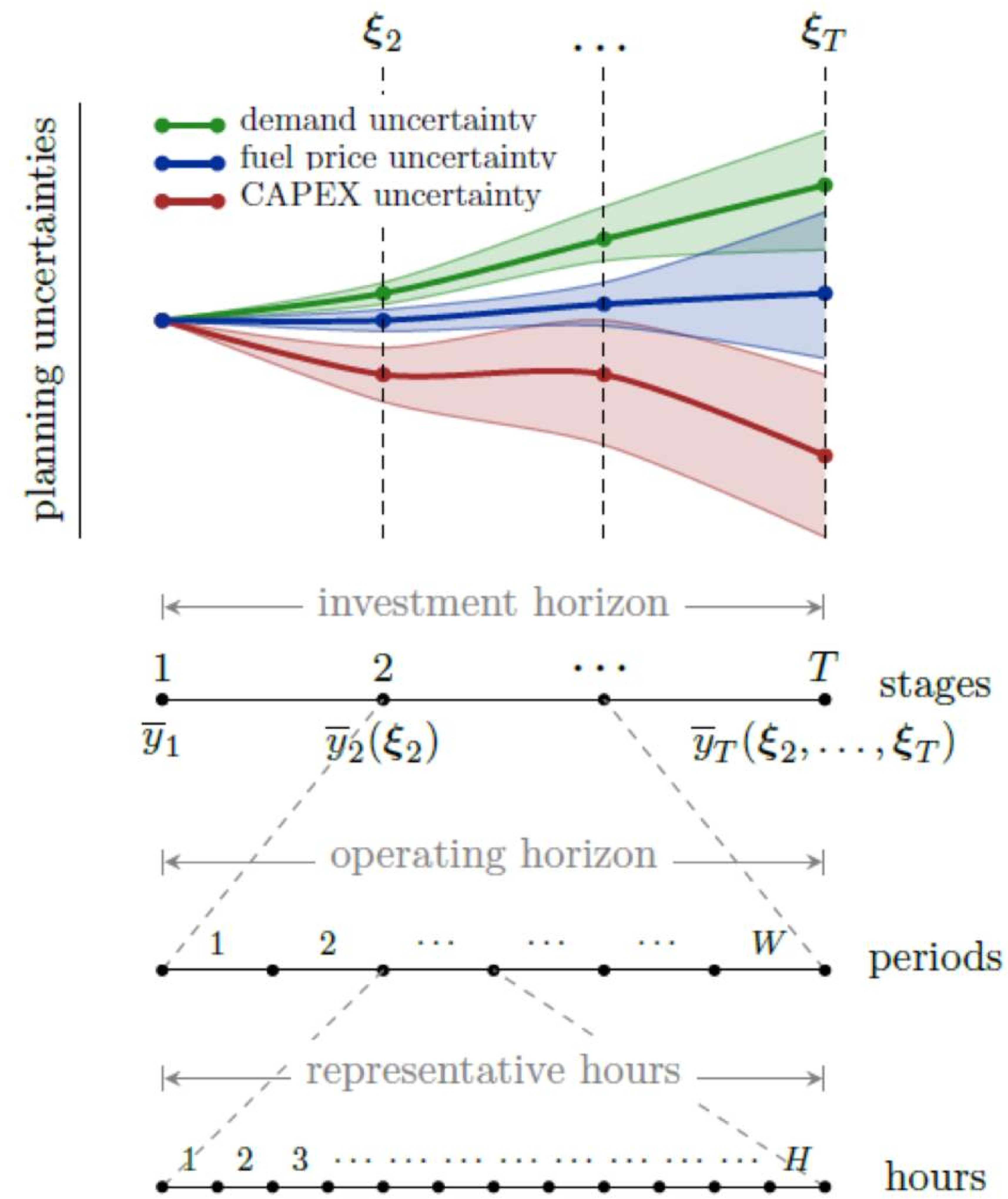
## *Audun Botterud*


## Research Goals

- Develop novel analytics for the planning and operation of clean energy systems of the future
- Improve methods for large-scale grid integration of renewable energy
- Advance the design and operation of low-carbon electricity markets

## Toolbox

- Energy Economics
- Optimization and Simulation
- Machine Learning
- Power Systems Engineering
- Decision Theory
- Game Theory



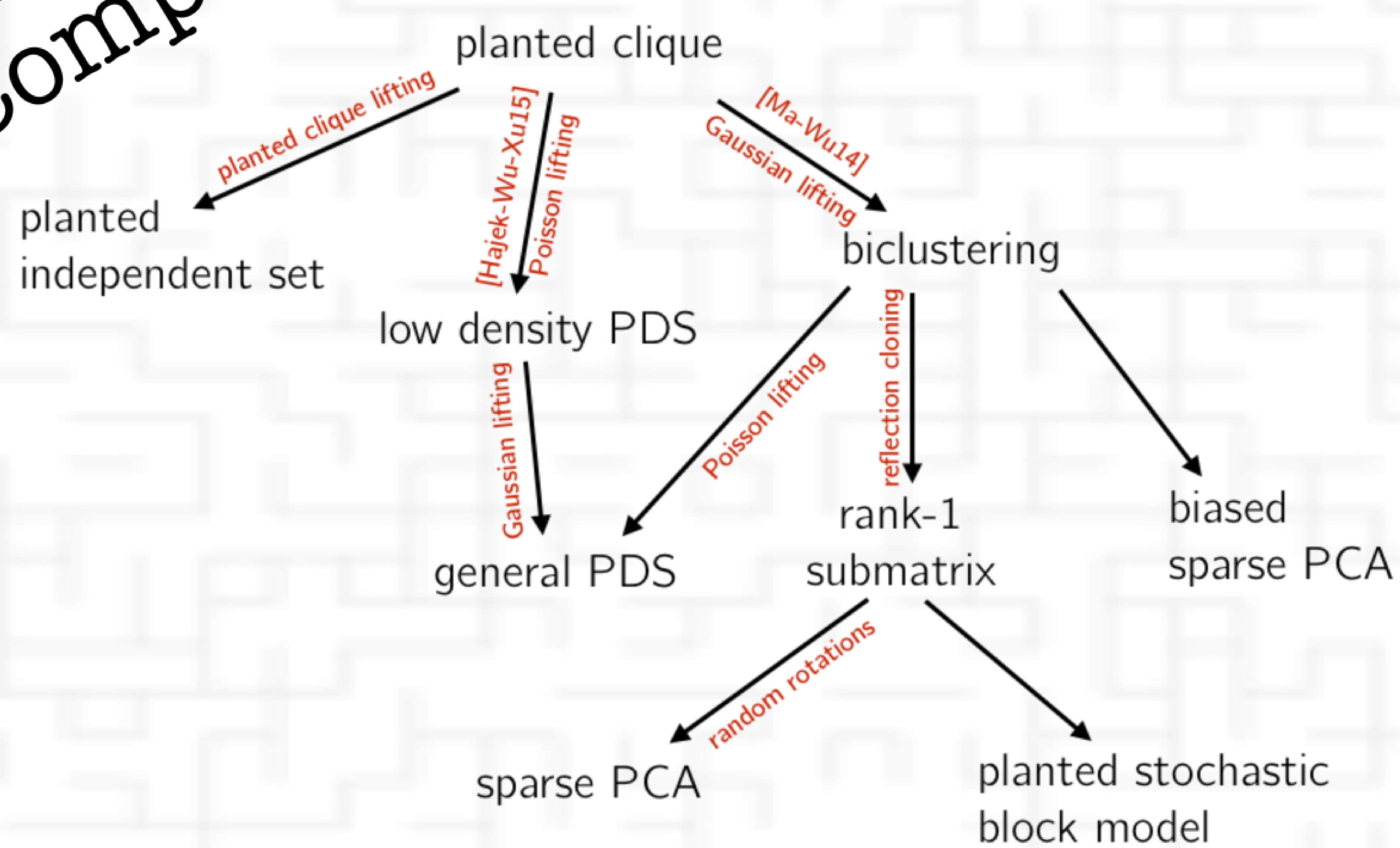

## Current research

- Electric transportation
- Electrification of heating
- Low-carbon buildings
- Zero-caron energy systems and markets
- Clean energy technologies
- Electricity market design
- Grid integration of renewable energy
- Energy storage optimization/control

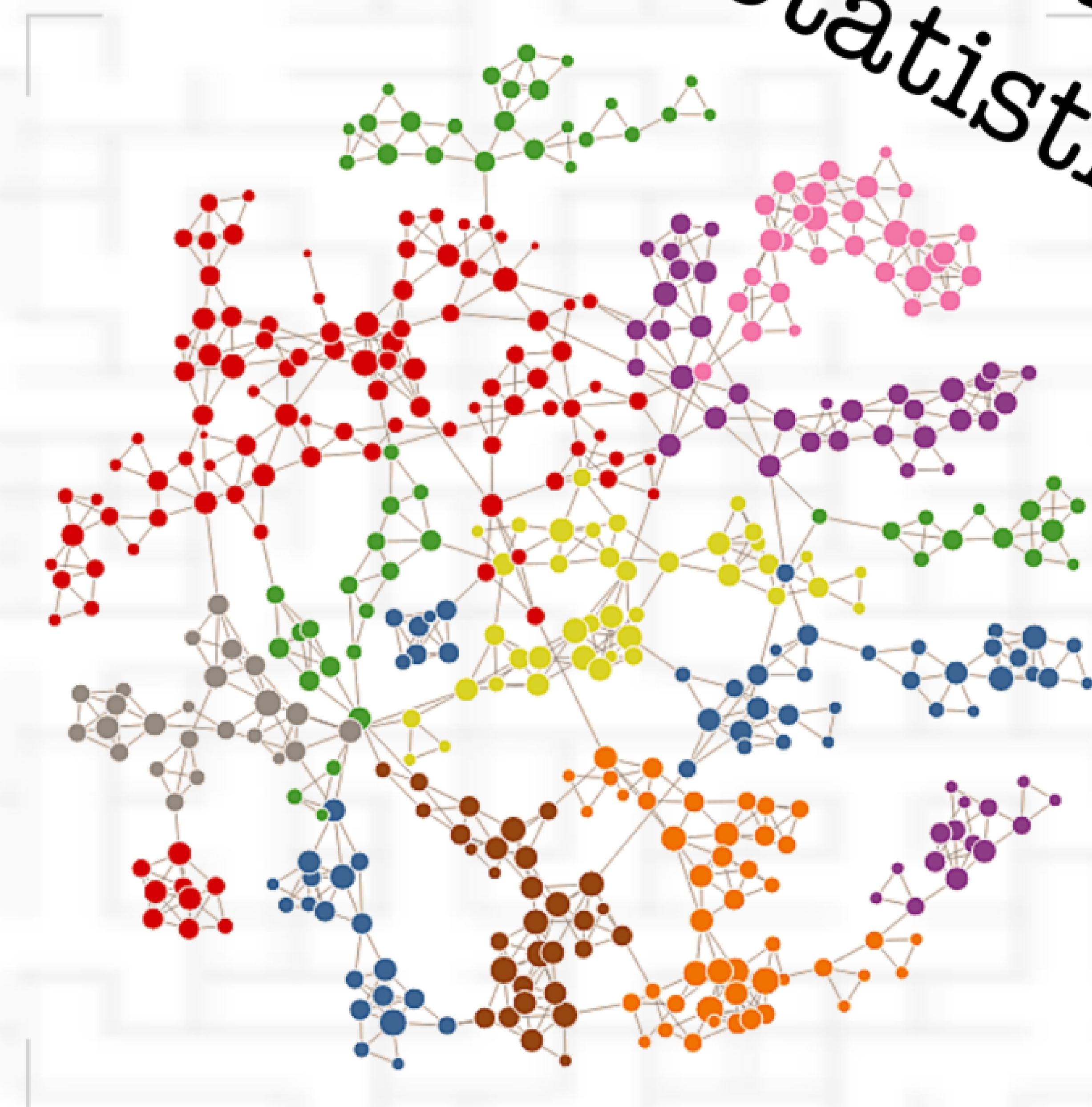
LIDS
RESEARCH GROUPS
2024-2025

# Guy Bresler & Yury Polyanskiy

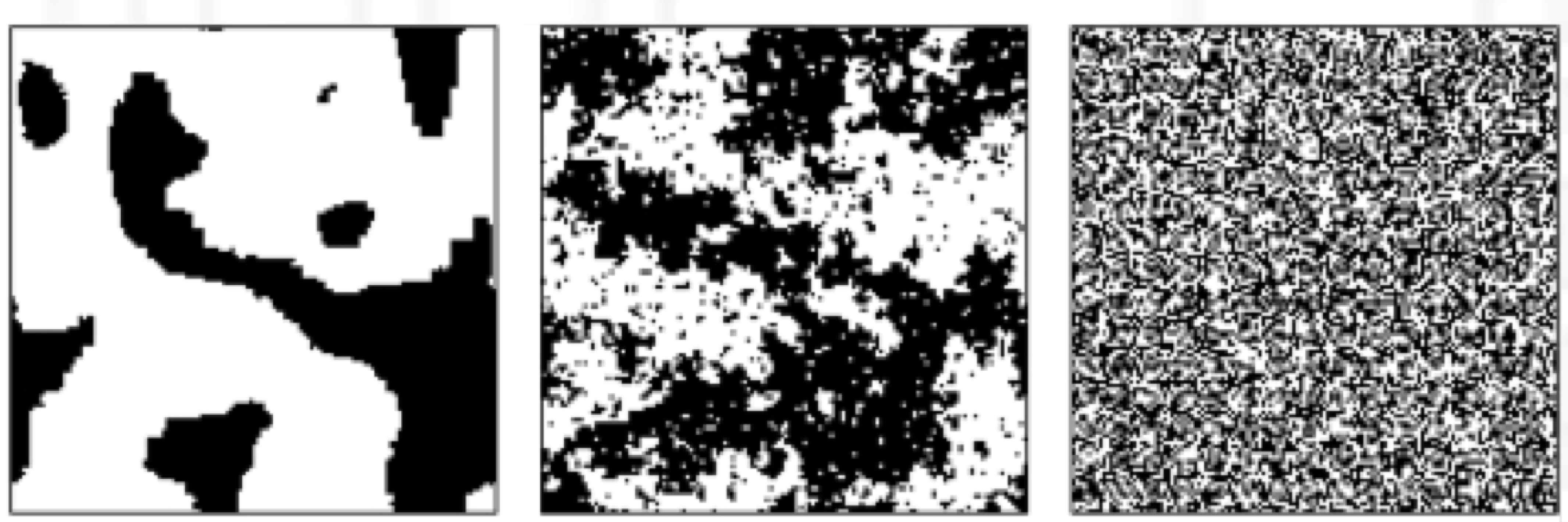
## Average-Case Comp. Complexity



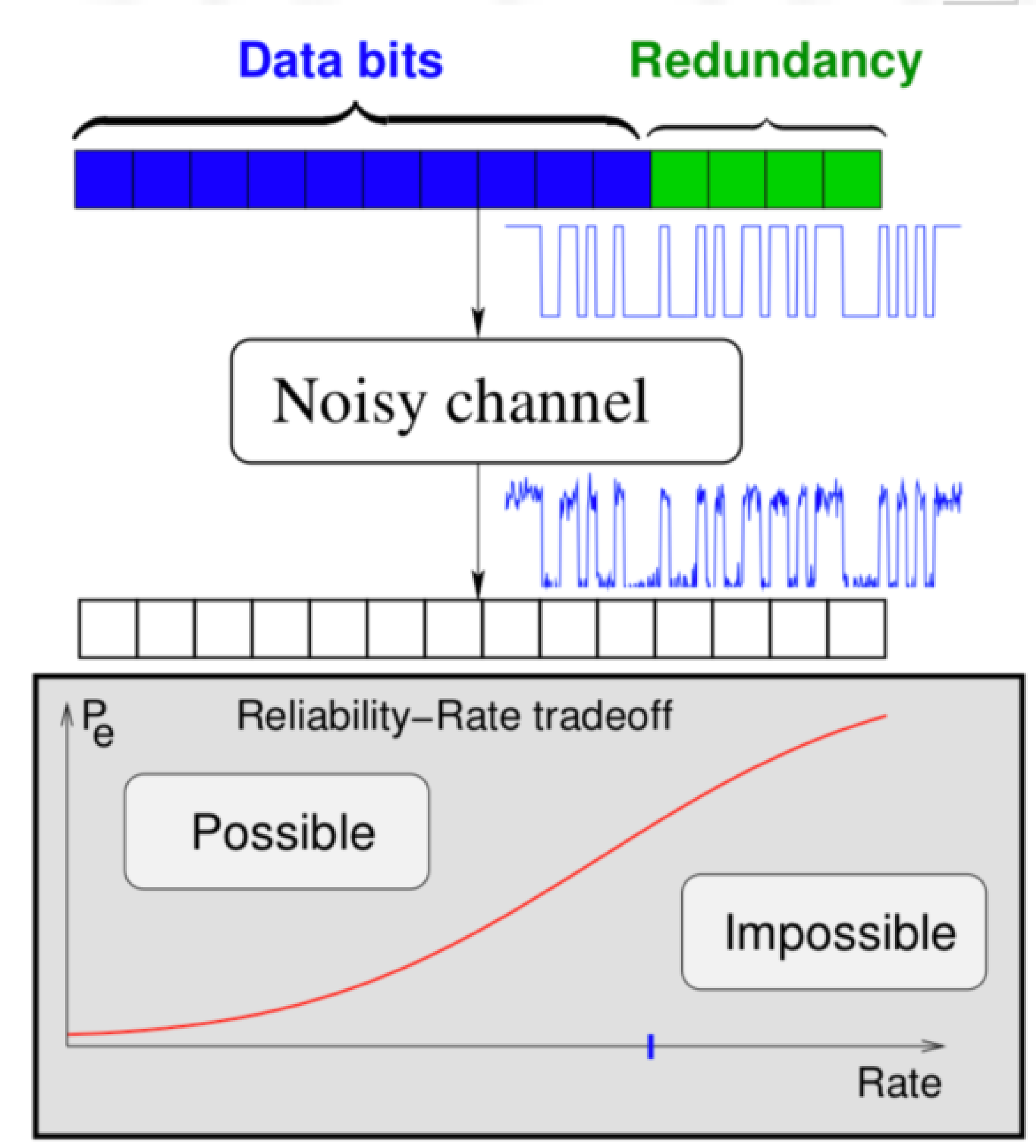

## Information Theory & Statistical Inference



## Probability



## Digital Communication



# Tamara Broderick

tamarabroderick.com

- **My group asks not just what we know, but also: how well do we know it?**
- We study uncertainty and robustness -- and design fast, easy-to-use, and provably accurate decision-making tools.

**Some examples**:

- How many new variants can scientists to expect to find when sequencing new genomes? We provide a state-of-the-art estimate, calibrated uncertainties, and an optimal tradeoff (under a fixed budget) of quantity (# individuals) and quality (sequencing depth).
- Consider an existing famous microcredit data analysis with >16,500 data points. Our work shows that if you drop one data point, the sign of the result changes, and if you drop 15 data points, you can get a significant result of the opposite sign. In general, we provide a tool (and supporting theory) to very quickly discover: if you drop a very small fraction of your data, how much can your substantive conclusions change?
- We develop a method to enable individuals with severe motor impairments (cerebral palsy, locked-in syndrome) to type, draw, game, and generally use computers. We adapt to individual users and limited motor control using statistical inference.

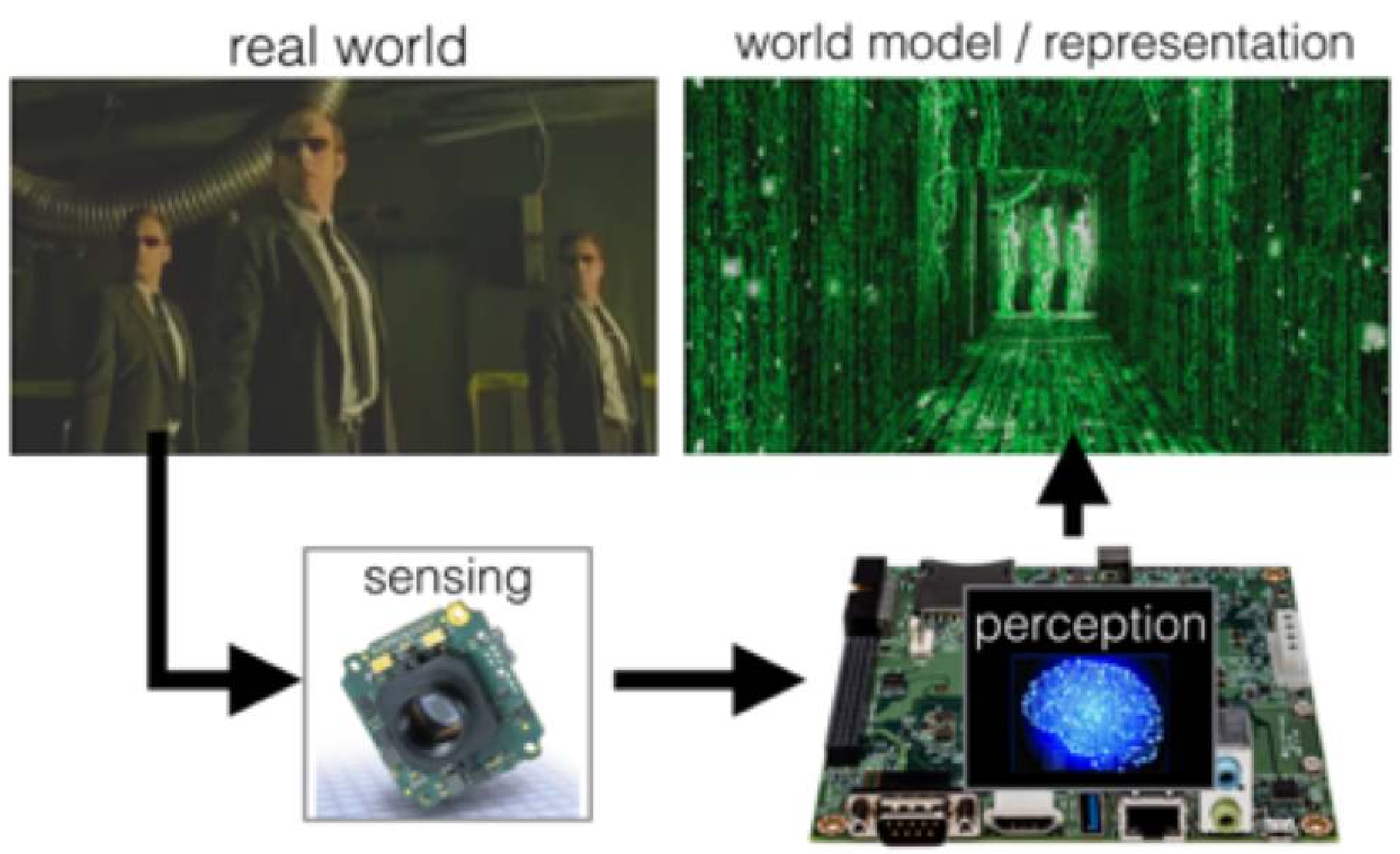
We **collaborate** with: economists, biologists, materials scientists, HCI specialists, and more. Our **methodology and theory** draw on measure-theoretic probability, stochastic process theory, real analysis, optimization, statistical mechanics, and a lot of other fun math.

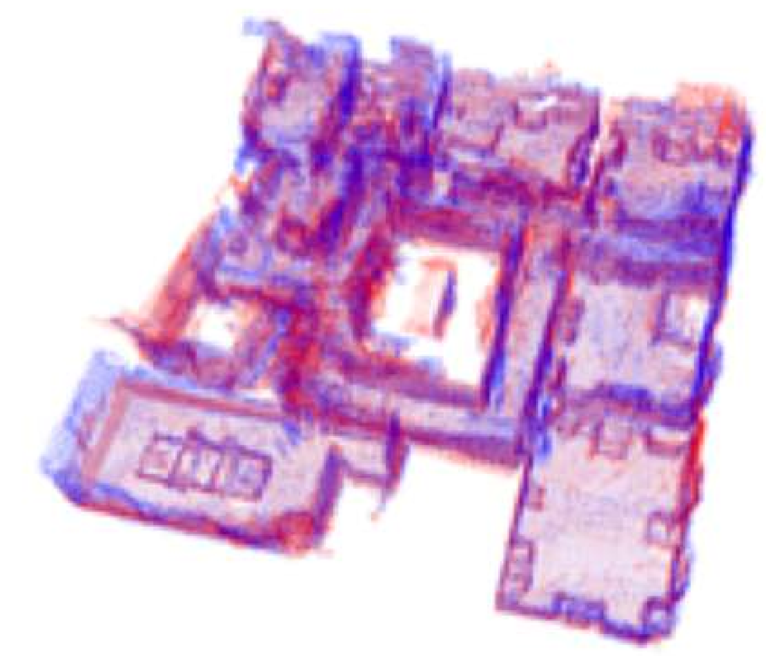
LIDS
RESEARCH GROUPS
2024-2025

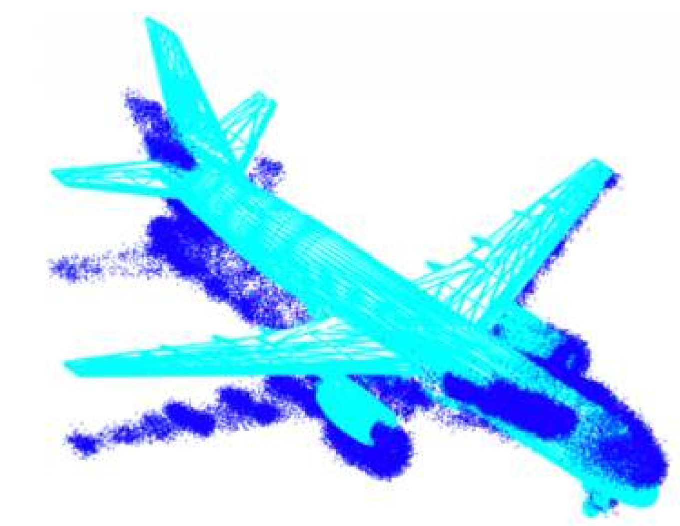
# Luca Carlone

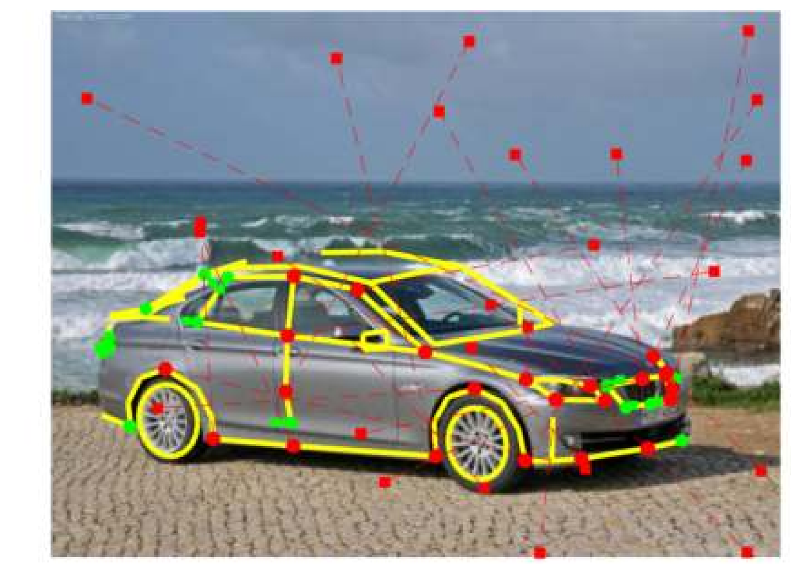
## Sensing Perception Autonomy and Robot Kinetics

SPARK Lab





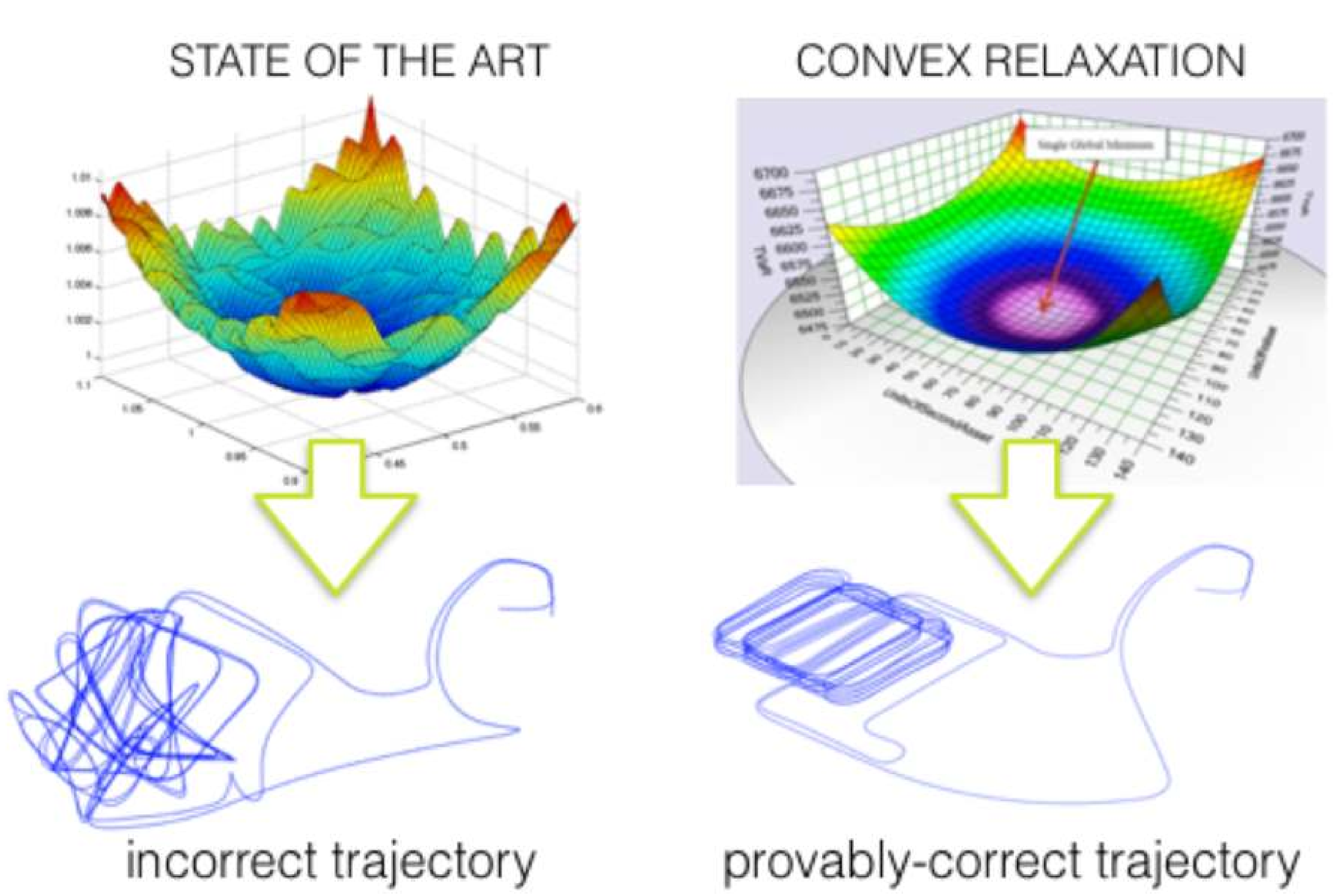


**Goal:** to develop theoretical understanding and practical algorithms to bridge the gap between human and computational (robot) perception

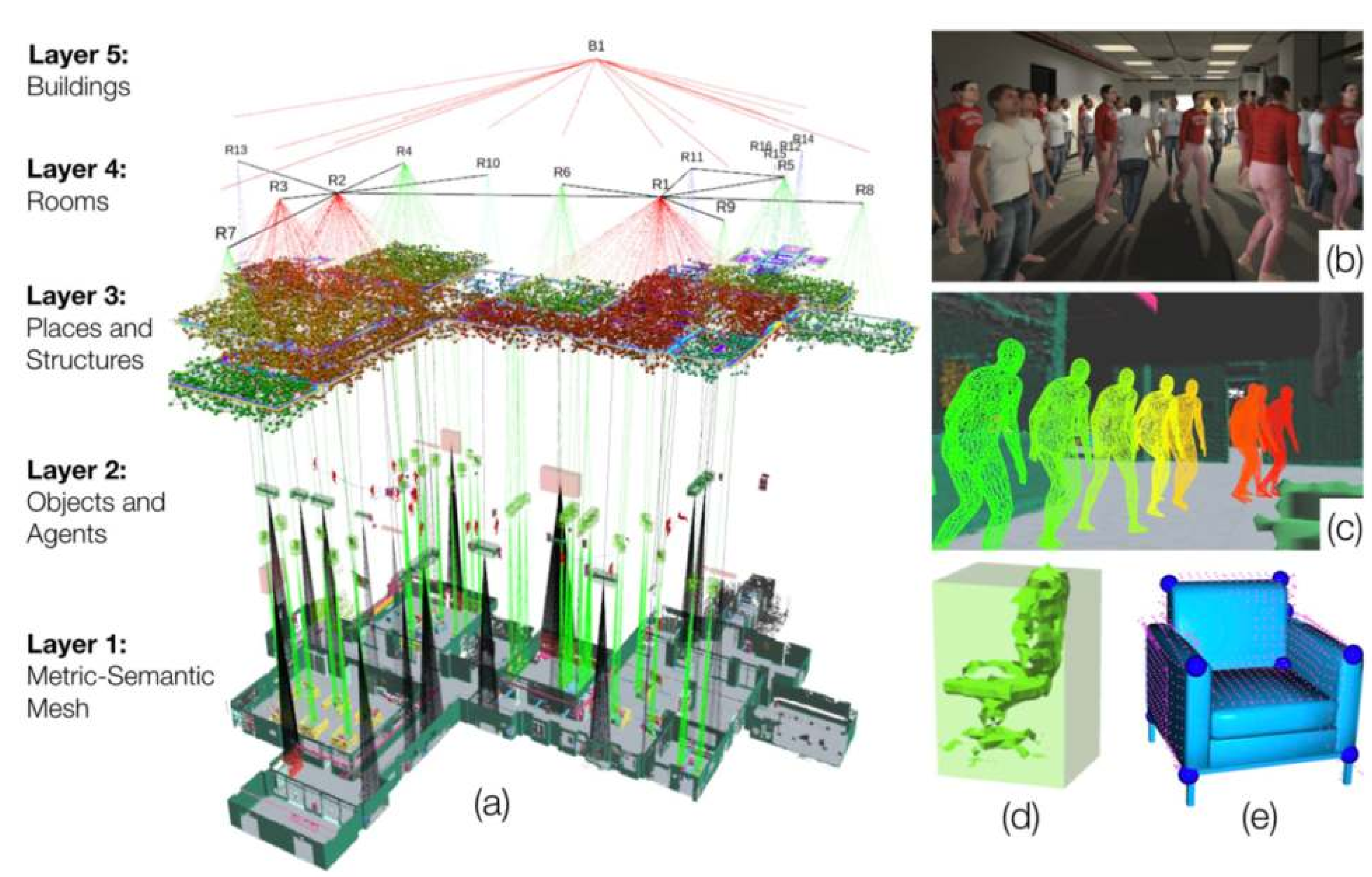
**Technical tools:**

- (non-convex, distributed) optimization
- nonlinear estimation & probabilistic inference
- geometry, graph theory
- control theory, machine learning

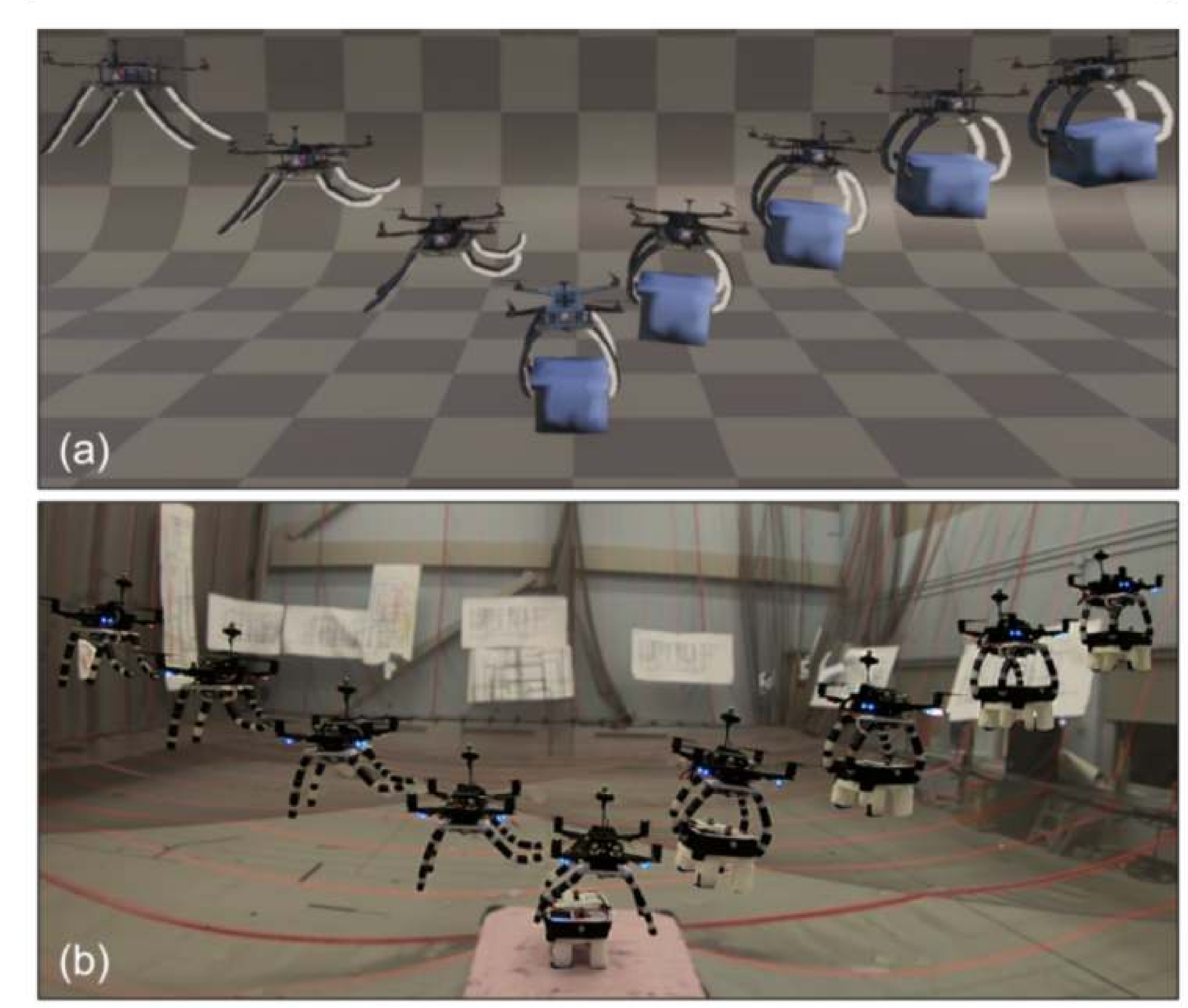
### Certifiable Robustness




### Real-time High-level Understanding




### Robot Co-Design



LIDS
RESEARCH GROUPS
2024-2025

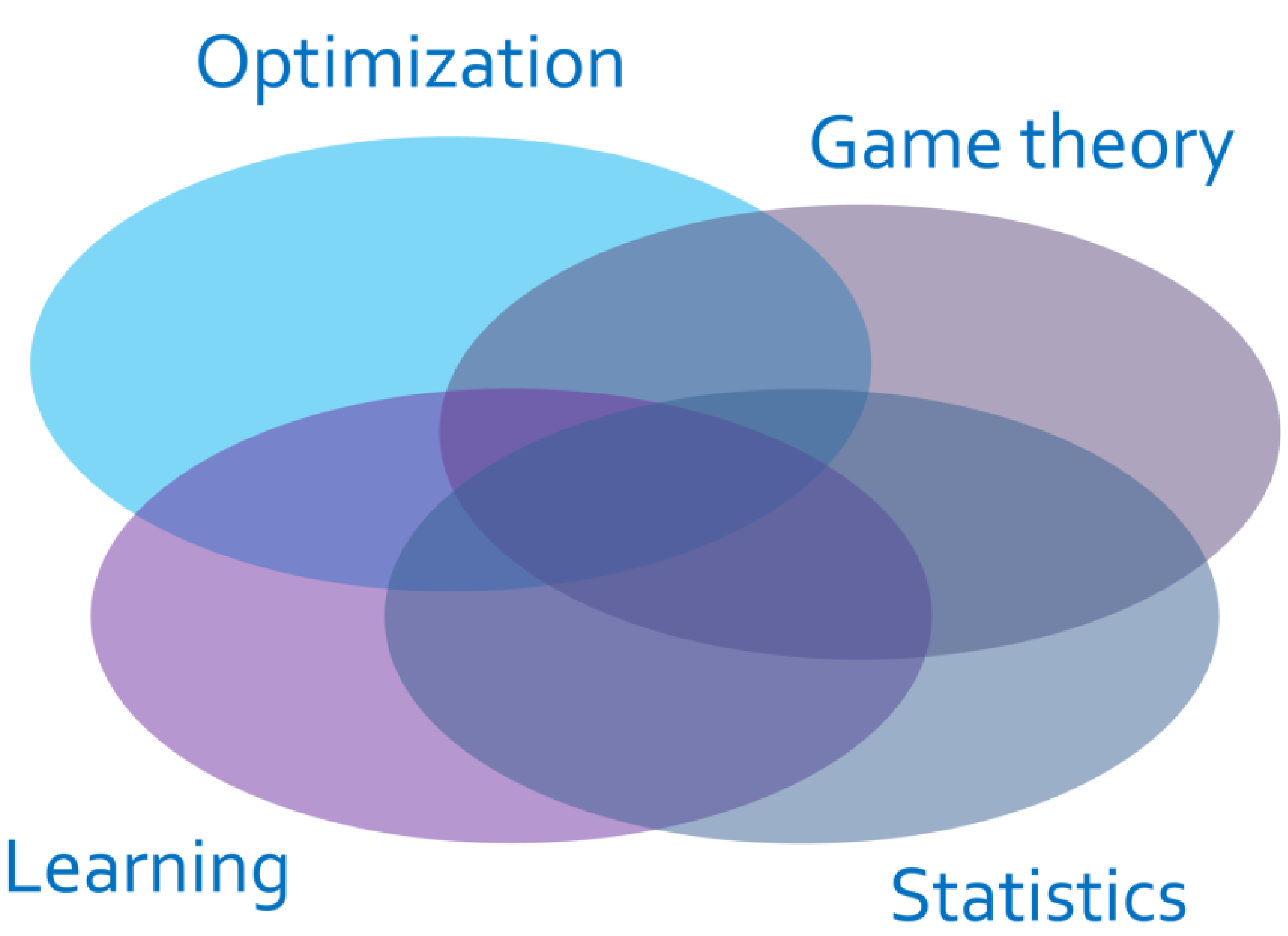
# Munther Dahleh

| | MOTIVATIONS | RESEARCH QUESTIONS | FINDINGS |
|---|---|---|---|
| **ENERGY**<br>Markets for Energy Storage | The power sector is in the cusp of a revolution due to increasing **renewable energy penetration** and **transportation electrification**, necessitating a complete rethinking of electricity market design. | The future grid will consist of millions of EVs that can double as energy storage resources. How to design efficient and incentive compatible mechanisms for EVs to sell battery service to the ISO? | Without carefully-designed incentive structures, integrating EVs for energy storage is counterproductive to energy-efficiency of the grid. Mechanisms to solve this issue have been devised. |
| **FINANCE**<br>Systemic Risk in Financial Networks | **Financial networks** are sensitive to **shocks**. Interbank lending networks are crucial to **allocating liquidity. Default cascades** can be detrimental to national economies. | How many **independent shocks** can hit the network?<br>How does **interbank lending network** function in times of market stress? | The effect of the network structure on the default rate and systemic loss.<br>A method for real-time, automatic, interpretable risk assessment. |
| **DIGITAL FARMING**<br>Information Elicitation and ML to find creditworthy borrowers | Community members know which of their neighbors are likely to repay a loan. We can elicit this information using clever incentives and algorithms. | How to develop incentives which are maximized with truthful reporting? How to create robustness to collusion? How to improve with online learning? | We designed truncated decision scoring rules which incentivize truthfulness in most cases. We will learn more from a pending deployment in Uganda. |
| Reinforcement Learning for Customized Farming | Large-scale data-driven farming is hard as each farm is **a time-variant system**, and **observations are sparse** with respect to interventions and farms. | How do we learn **near-optimal customized policies** for a large number of farms while achieving **provably good performance**? | Given a set of policies, we can learn in finite time and perform almost as well as the best policy considered. |
| **COVID-19**<br>Testing As Control | Attempts to control the spread of the COVID-19 epidemic focus on social distancing, but **testing and contact tracing** should also be considered. | How do we **model testing**/contact tracing/network structure? What are the **qualitative relationships** between testing and disease spread? | Dynamics are **independent of network structure** and simple relationships/formulas determine how testing and disease spread interact. |
| Impacts of Interventions | Interventions for COVID-19 affect people differently due to **variations in age, health conditions, socioeconomic status**, and many other factors. | What are the **impacts** of various interventions on different **communities**? Are there any **trade-offs**? Which interventions are **effective**? | There is a trade-off between saving lives from the pandemic and from recession. The disadvantaged community tends to suffer significantly more than others. |
| **NETWORKS**<br>A Marketplace for Data | Data, an increasingly vital asset, needs to be valued in a systematic way.<br>**Data markets** must consider interactions between and among data buyers, sellers, and intermediaries. | **Robust real-time matching** mechanism to buy and sell training data for Machine Learning tasks?<br>How to allocate and price data sets to buyers in **competition** with each other? | Mathematical model and real time algorithms for a two-sided data market.<br>Welfare and revenue-maximizing mechanisms for selling data to data buyers with negative externalities. |

# Gabriele Farina

## How can machines reason about *strategic* behavior?

**Example**: How do you teach a machine to compute the correct amount of bluffing when playing a poker hand?




What is the objective function?

Can you learn from repeated simulations?

How do you scale these kinds of computations to real-world settings?

1. **Reason about imperfect-information**
   *Not just as an obstacle to sidestep, but also as a strategic opportunity*
2. **What is optimal game-theoretic behavior?**
   *What is the optimization problem we should be solving?*
3. **How do we compute/learn optimal strategic behavior?**
   *For example, when is optimization tractable from a computational and statistical complexity point of view?*
4. **How do you handle interacting with humans?**
   *In cooperative settings?*
   *In adversarial settings?*

LIDS
RESEARCH GROUPS
2024-2025

# REALM: REliable Autonomous systems Lab at MIT

Chuchu Fan

## Sensing and Perception

camera, LIDAR, GPS, computer vision, machine learning, data

Our method can find the most robust design parameter for full-stack autonomy

## Decision and Planning

Decision-making, navigation, path planning

Our algorithm can plan motions automatically from temporal logic specs

## Control and Act

physics, computer, code, engine, actuator

We learn certifiably safe control policies for large-scale autonomy

LIDS
RESEARCH GROUPS
2024-2025

# Aerospace Controls Laboratory
# Jon How

**Multiagent** | **Planning** | **Learning** | **Perception**

**Tools:**
Control Theory
Optimization Theory
Algorithms
Graph Theory
Machine Learning

**Venues:**
AI/ML: NIPS, ICML, ...
Control: CDC, ACC, ...
Robotics: RSS, ICRA, ...

**Projects:**

- ❑ Cooperative Multiagent Reinforcement Learning
- ❑ Lifelong Learning for Distributed Intelligence
- ❑ Pedestrian Motion Prediction
- ❑ Decentralized Dynamic Task Allocation
- ❑ Multiagent Search & Rescue in Forests
- ❑ Resource-aware Spatial Perception
- ❑ Active Perception for Threat Identification

**http://acl.mit.edu**

# EESG @ MIT
## Marija Ilic

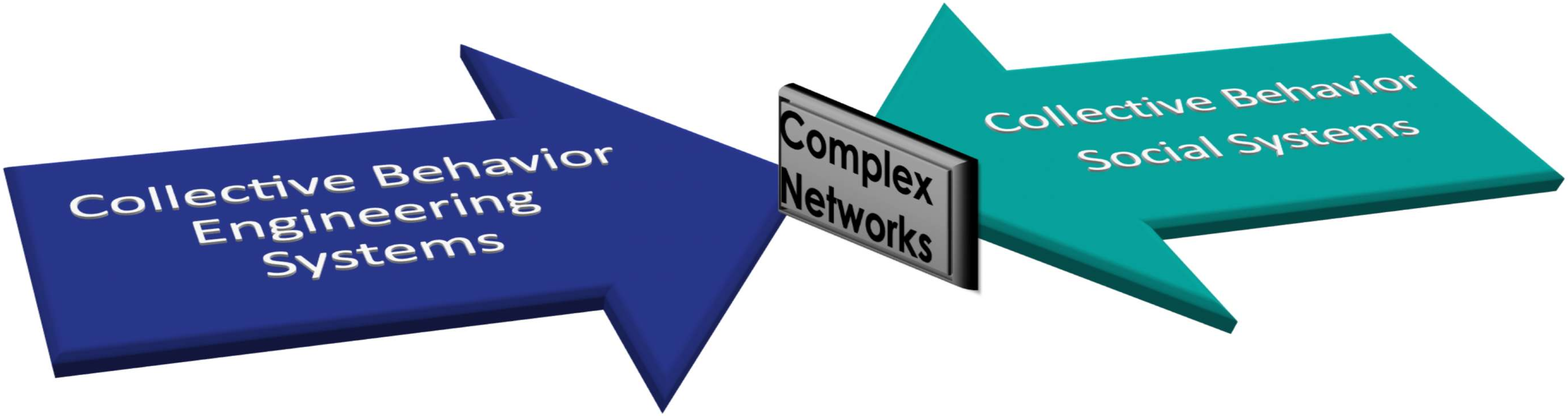
# Electric Energy Systems Group

The Electric Energy Systems Group (EESG) focuses on research related to modeling, control, and communications design of our rapidly changing electric energy system. Our mission is to enable reliable, resilient, sustainable, and cost-effective electric energy service at scale.

We are currently part of MIT's Laboratory for Information and Decision Systems (LIDS). Previously, our group was based at Carnegie Mellon University

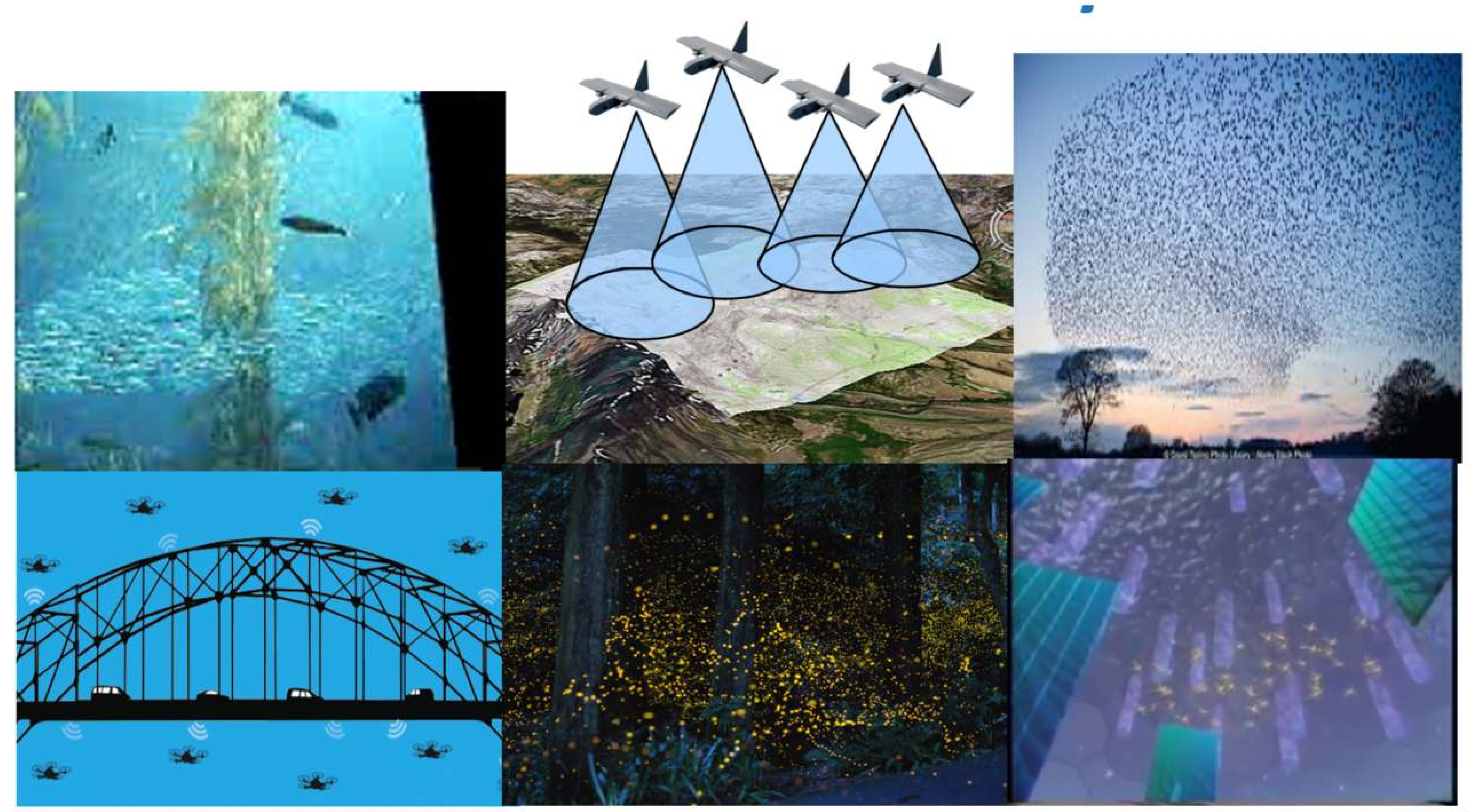
**LIDS**
**RESEARCH GROUPS**
**2024-2025**

# Ali Jadbabaie



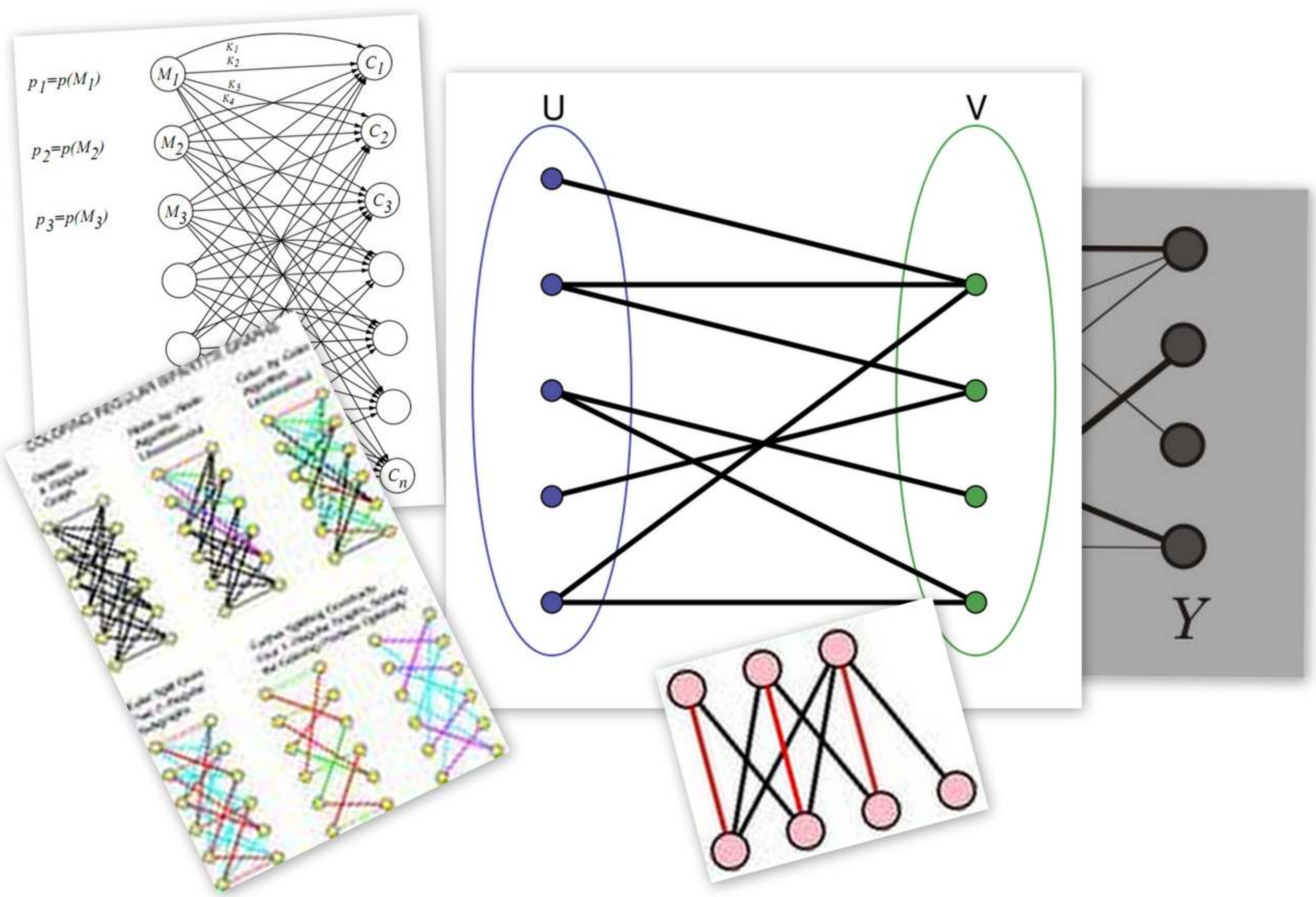

How do people make decisions in groups?
How do misinformation and rumors spread?
What is the role of network structure?

Collective behavior, phenomena, action
Towards a unified theory of group decision-making
News consumption and news sharing



LIDS
RESEARCH GROUPS
2024-2025

# Online Learning and Optimization Group
# Patrick Jaillet



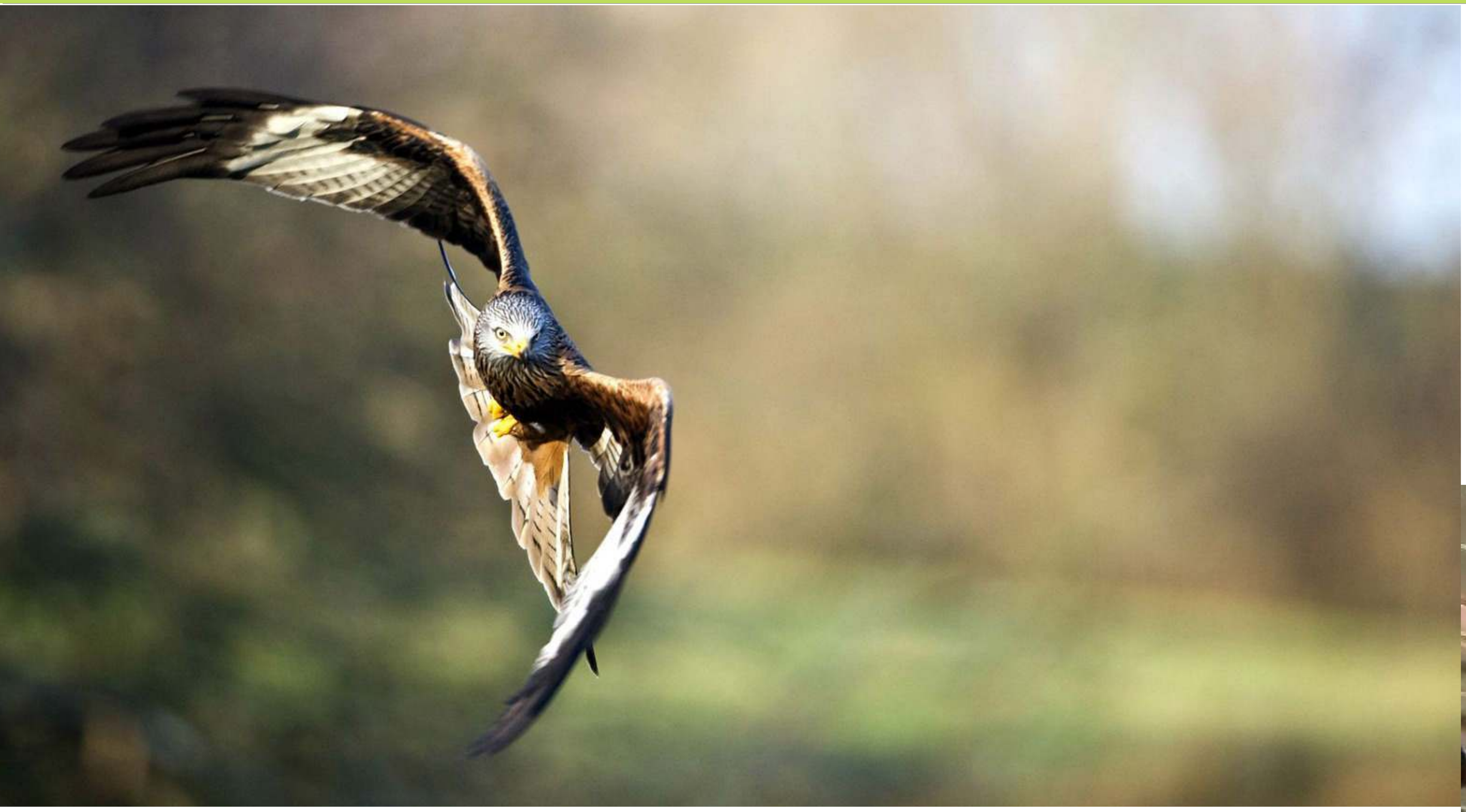

research: online optimization and learning problems; applied probability

focus: theory, models, algorithms

applications, contexts:

- -routing, mobility, spatial explorations
- -internet, dynamic resource allocations
- -cyberinfrastructure security
- -sharing economy
- -networks

# Sertac Karaman



How fast can birds fly through forests? How quickly can robots navigate in cluttered environments? We analyze the performance limits for robotic vehicles operating in cluttered environments.

As driverless cars edge closer to becoming a reality, we ask the question: Can autonomous cars substantially improve performance in traffic intersections? How about all-autonomous transportation networks?

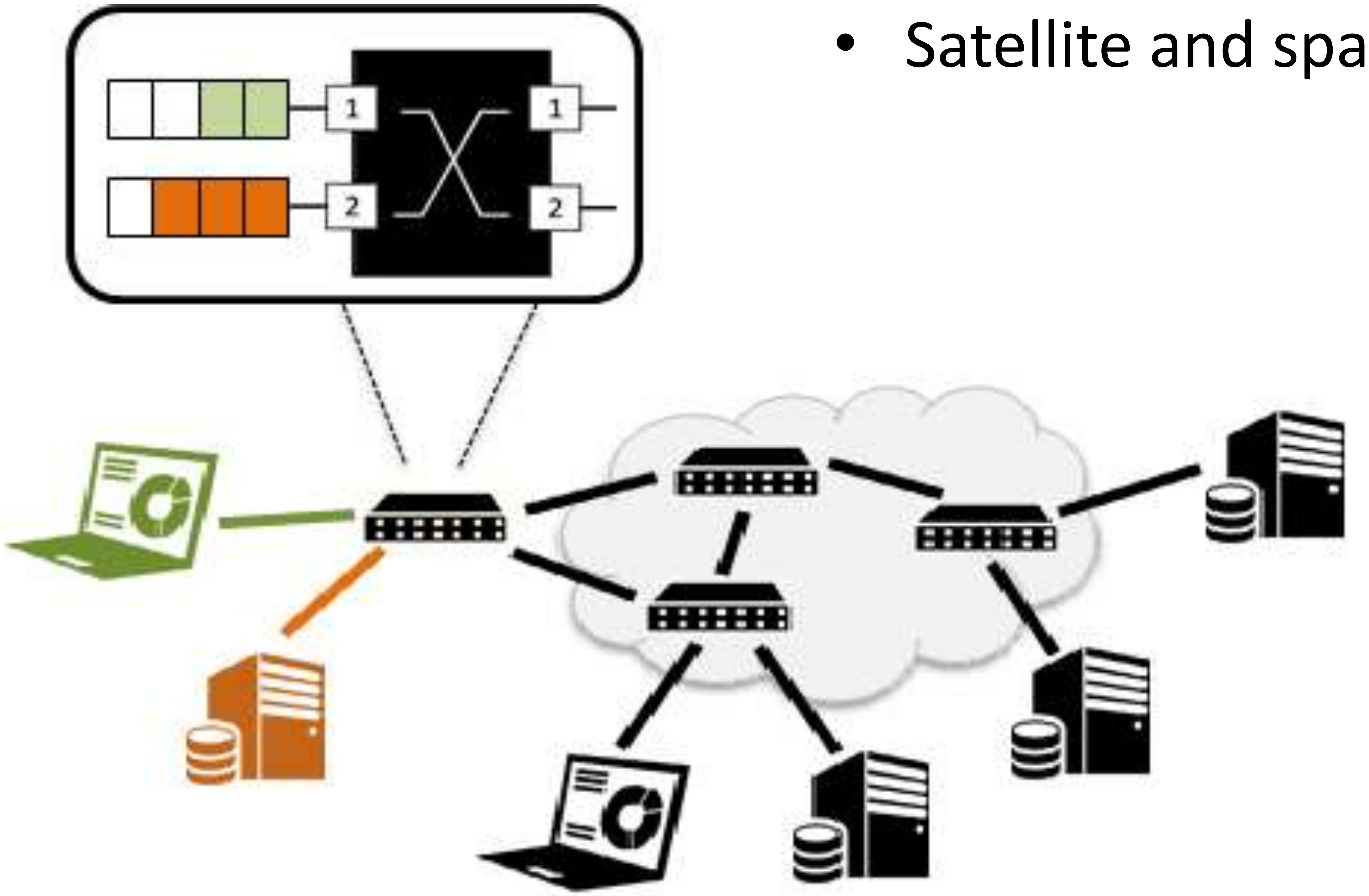
A new approach to teaching feedback control systems allows the students to instantly test their control design on a palm-size drone in the comfort of their room. Each student enrolled in 16.30 will get a Parrot mini drone.

LIDS
RESEARCH GROUPS
2024-2025

# Communications and Networking Research
# Eytan Modiano

We develop architectures and algorithms for communication networks, including:

- Machine learning in networks
- Autonomous network control
- Communication for UAVs
- Wireless networks
- Satellite and space networks



**LIDS**
**RESEARCH GROUPS**
**2024-2025**

# Asu Ozdaglar

**Research Focus**: Developing new models, mathematical tools, and algorithms for the analysis and optimization of large-scale data-driven systems and for machine learning.

## Areas:

- Optimization
- Machine Learning
- Economics
- Networks

## Current Projects:

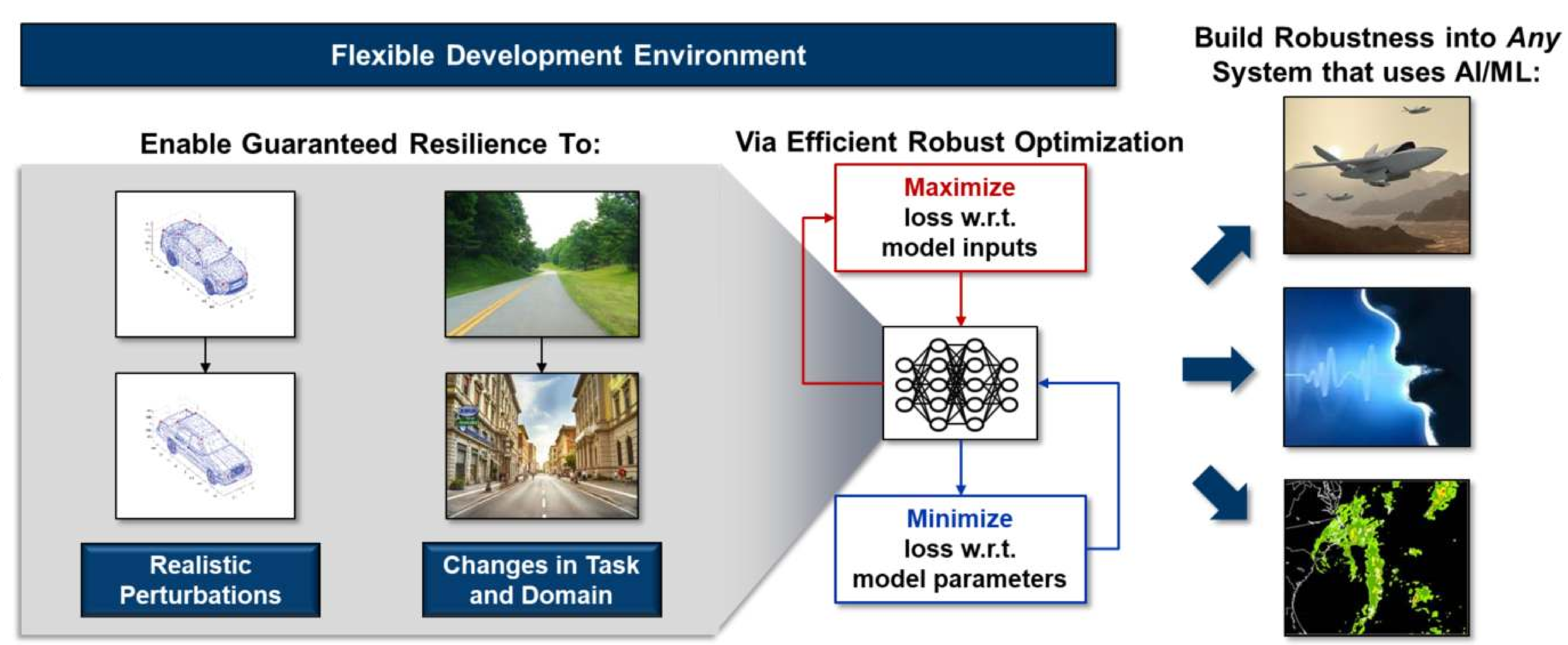
Privacy, Data Ownership and Markets

AI-Driven Social Media: Algorithms and Regulations

Robust and Decentralized Machine Learning

Information Design and Learning in Networks

…



# Sasha Rakhlin

## Machine Learning, Statistics, and Optimization

**Statistical Learning:** We study the problem of building a good predictor based on an i.i.d. sample. While much is understood in this classical setting, our current focus is large overparametrized models, such as those employed in deep learning. In particular, we study various measures of complexity of neural networks that govern their out-of-sample performance. Our recent focus is on statistical and computational aspects of interpolation methods, as well as understanding the phenomenon of benign overfitting in overparametrized models.

**Contextual Bandits and Reinforcement Learning:** In these problems, data are collected in an active manner and feedback is limited. Our work focuses on understanding the sample complexity and on developing computationally efficient methods. Among the highlights is a recent reduction from these decision-making problems to Supervised Learning.

**Online Learning:** We aim to develop robust prediction methods that do not rely on the i.i.d. or stationary nature of data. In contrast to the well-studied setting of Statistical Learning, methods that predict in an online fashion are arguably more complex and nontrivial. This field has some beautiful connections to Statistical Learning and the theory of empirical processes.

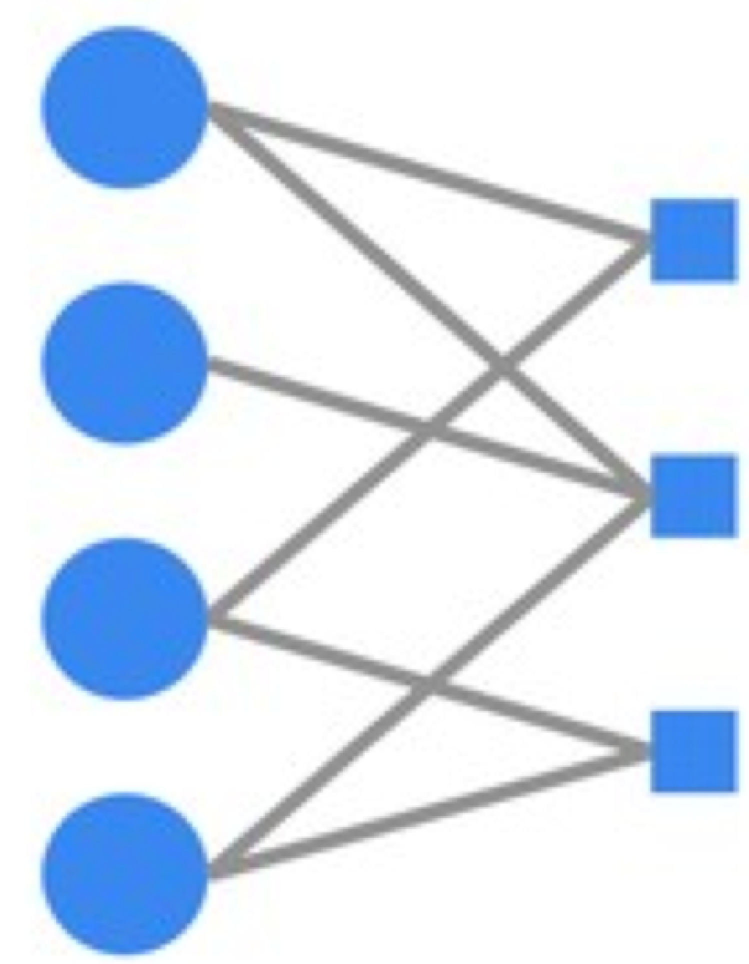
LIDS
RESEARCH GROUPS
2024-2025

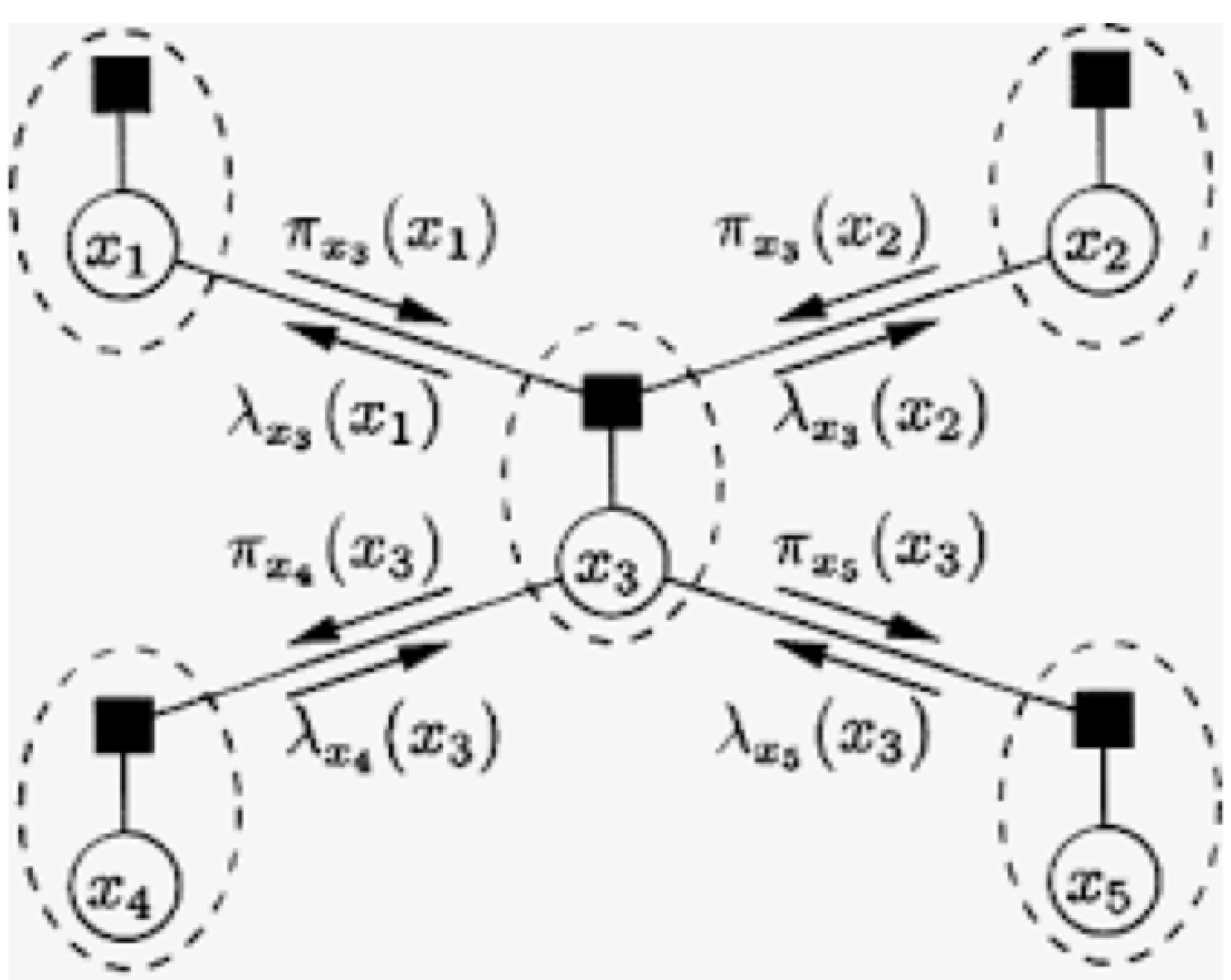
# Devavrat Shah

## Social Data Processing

Networks

Probabilistic Graphical Models







## Causal Inference

# David Simchi-Levi

## MIT Data Science Lab

MIT DSL DATA SCIENCE LAB

## MIT Data Science Lab: Executive Summary

### Theoretically Elegant & Practically Relevant Research

| Supply Chain Resiliency | Price Optimization | Personalized Offering | Inventory, Transportation & Procurement | Online Resources Allocation | Supply Chain Digitization |
|---|---|---|---|---|---|
| Ford | Rue La La | RYANAIR | BlueYonder | IBM | accenture |
| accenture | GROUPON | accenture | THE HOME DEPOT | Alibaba.com | ABInBev |
| Sigma alimentos | ORACLE | XL Group Insurance Reinsurance | accenture | | zalandolounge |
| CAMPOFRIO FOOD GROUP | Coppel | ERICSSON | MANGO | | |
| DENSO | zalando | | | | |
| Schneider Electric | starwood Hotels and Resorts | | | | |
| | B.TECH | | | | |

**Strategic intent:** Develop solutions to leading edge problems for lab partners through research that brings together data, modeling, and analysis to improve business performance

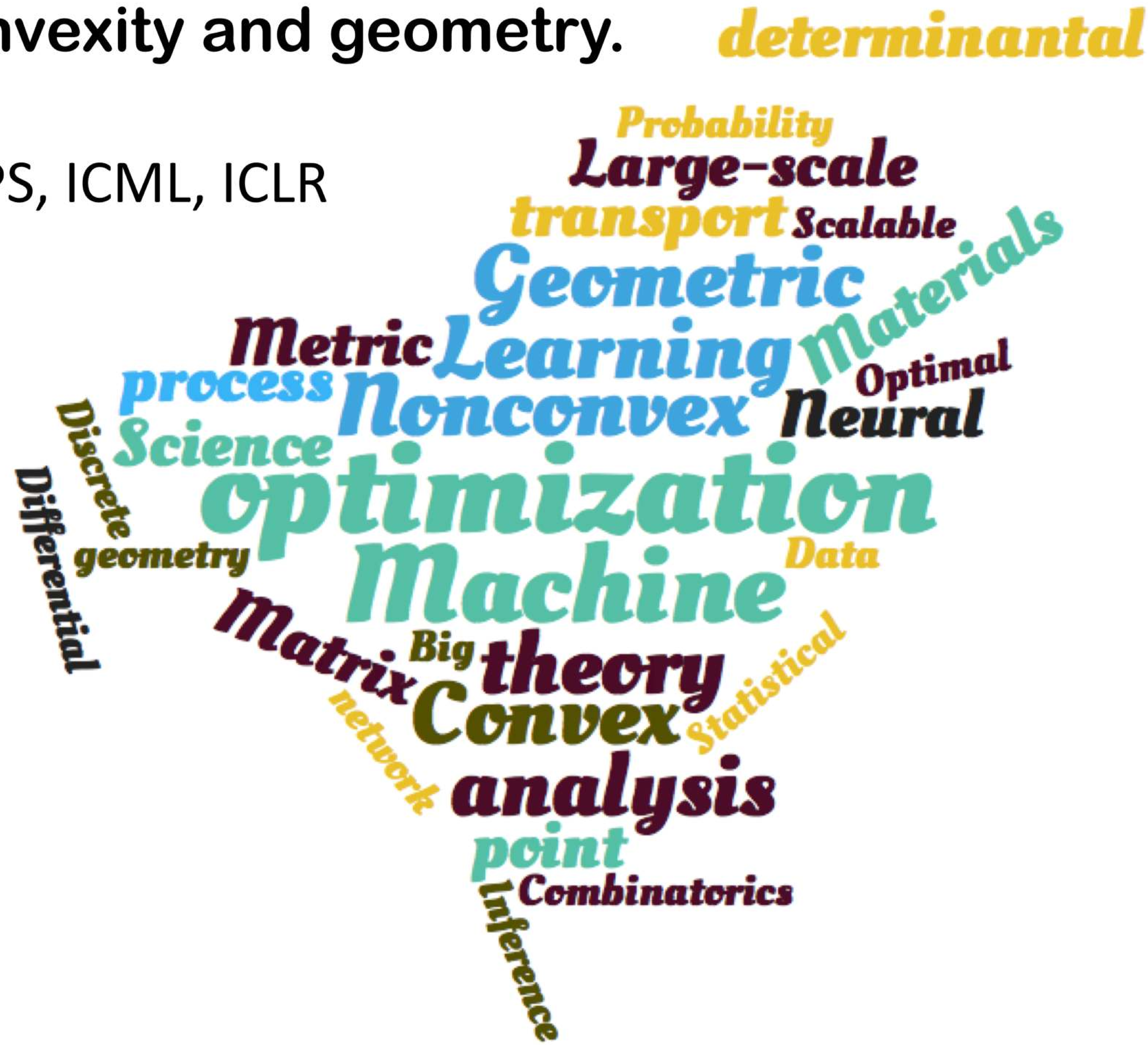
**Cross-industry:** Oil / Gas, Retail, Financial Services, Government, Insurance, Airlines, Industrial Equipment, Software

**Global footprint:** NA, EU, Asia, LA

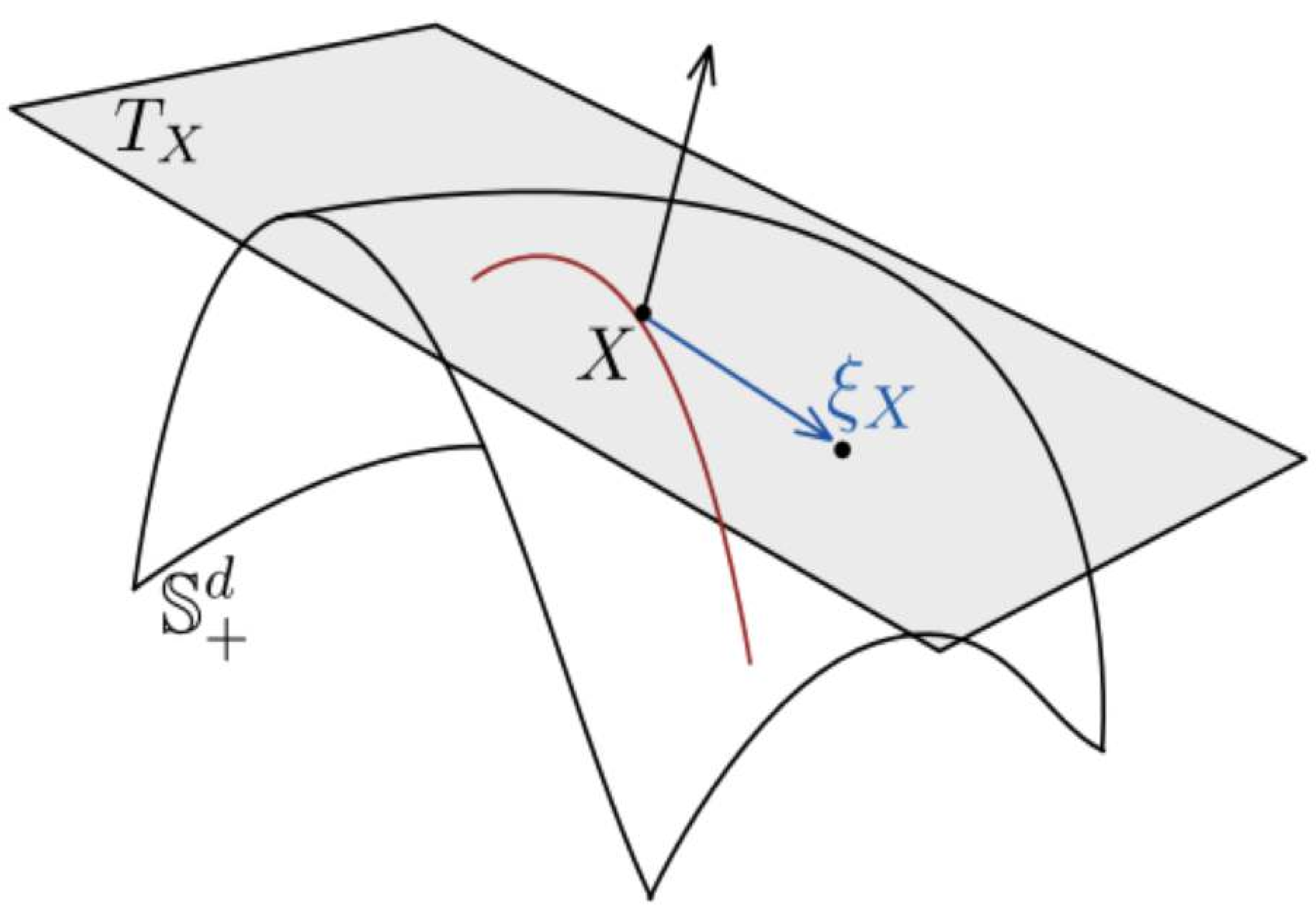
MIT DSL DATA SCIENCE LAB

# OPTML: Optimization for Machine Learning
## Suvrit Sra

**We work on theory, analysis, and development of mathematical models for optimization, sampling, and machine learning with a particular focus on non-convexity and geometry.**




**Main conferences:** COLT, NeuRIPS, ICML, ICLR




http://optml.mit.edu

http://ml.mit.edu

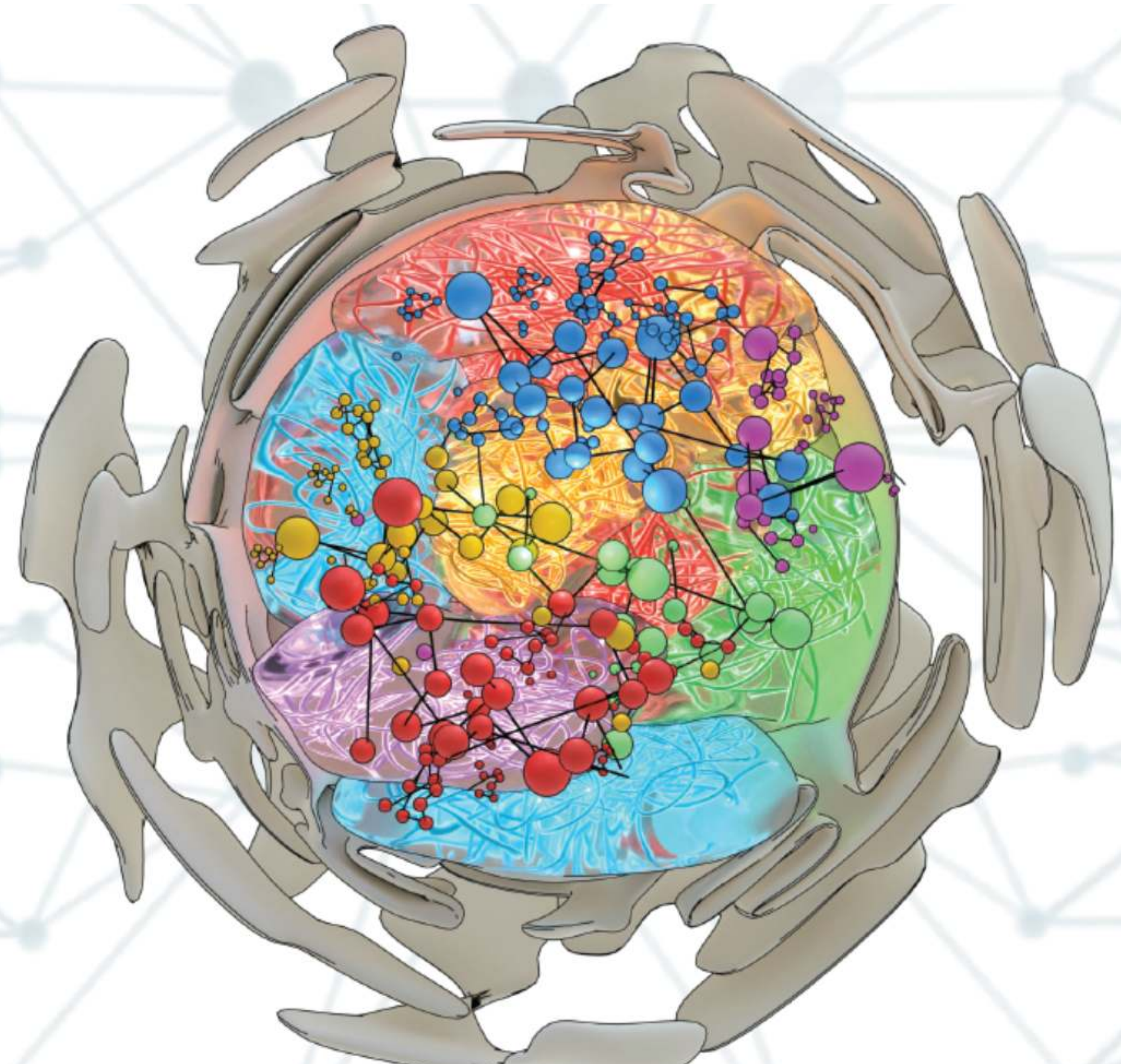
# Caroline Uhler

*«The Uhler Lab develops machine learning foundations and methods for integrating different data modalities and inferring causal relationships from such data. The developed methods and algorithms are applied to discover the regulatory circuits underlying the programs of cells and tissues in health and disease.»*

**CAUSAL INFERENCE**

**REPRESENTATION LEARNING**

**ACTIVE LEARNING**



**GENOME ORGANIZATION**

**CELL STATE TRANSITIONS**

**DIAGNOSTICS & THERAPEUTICS**

LIDS
RESEARCH GROUPS
2024-2025

# Data to AI Group
# Kalyan Veeramachaneni

## Systems for Machine Learning

Automatic - feature engineering, machine learning task generation, modeling and creating interactive developers tools.

## AI for cyber security

Spanning the gamut of malware detections exfiltration, explainable and adverserial AI.

## AI for software engineering

How can we transform software engineering using machine learning?

## Applications

Ranging from monitoring health of satellites, water pipes to healthcare and education.

## Our open source tools

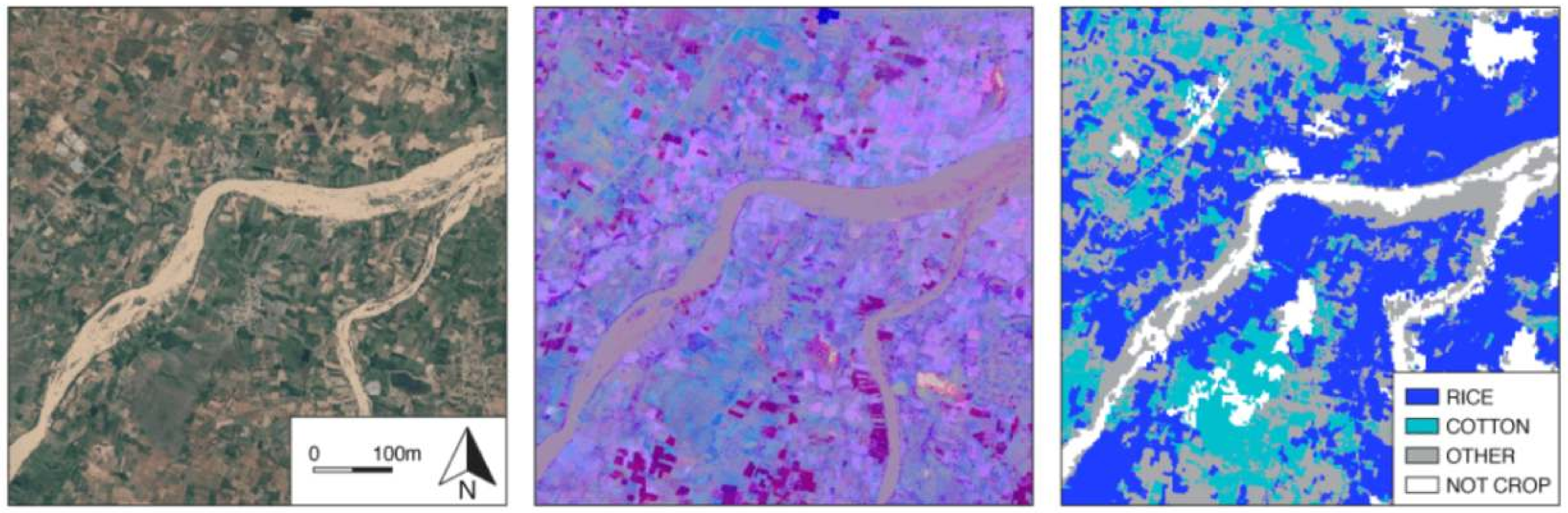
Use our software to build your own AI applications.

LIDS
RESEARCH GROUPS
2024-2025

# earth intelligence lab

# Sherrie Wang

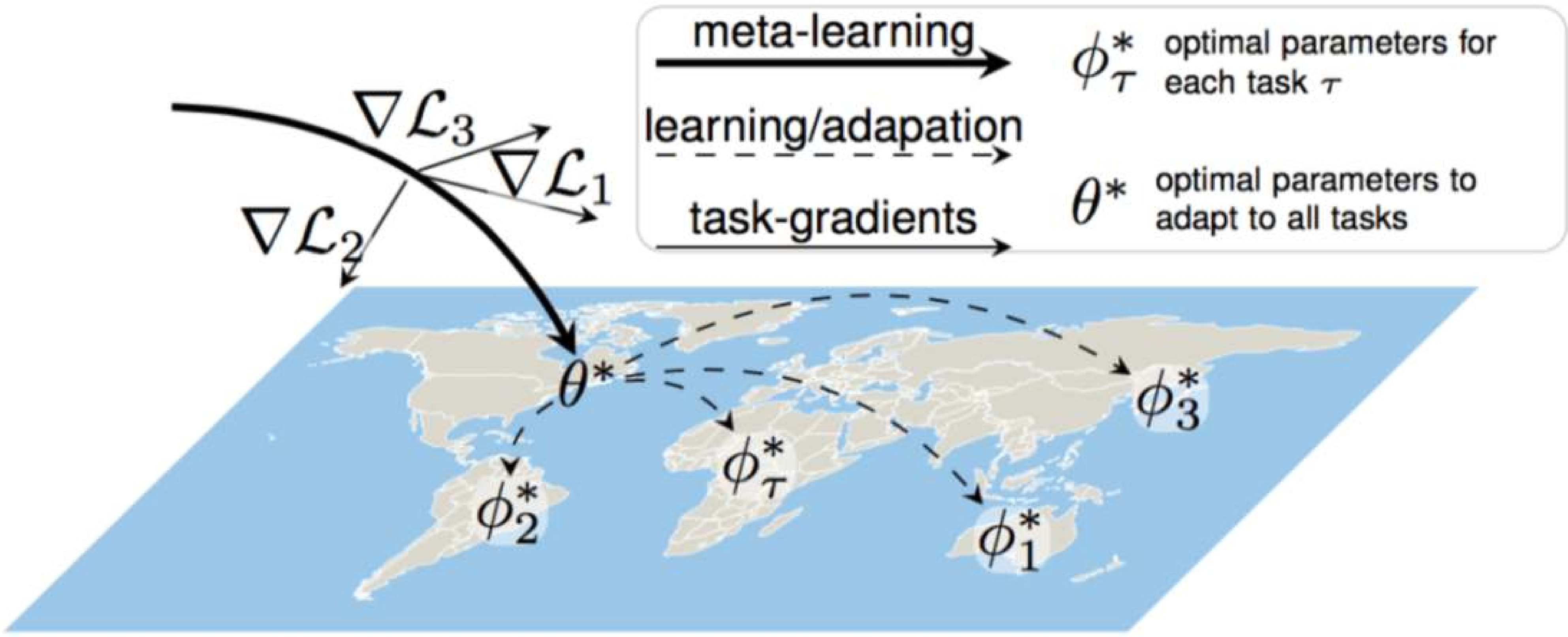
## Mapping agriculture



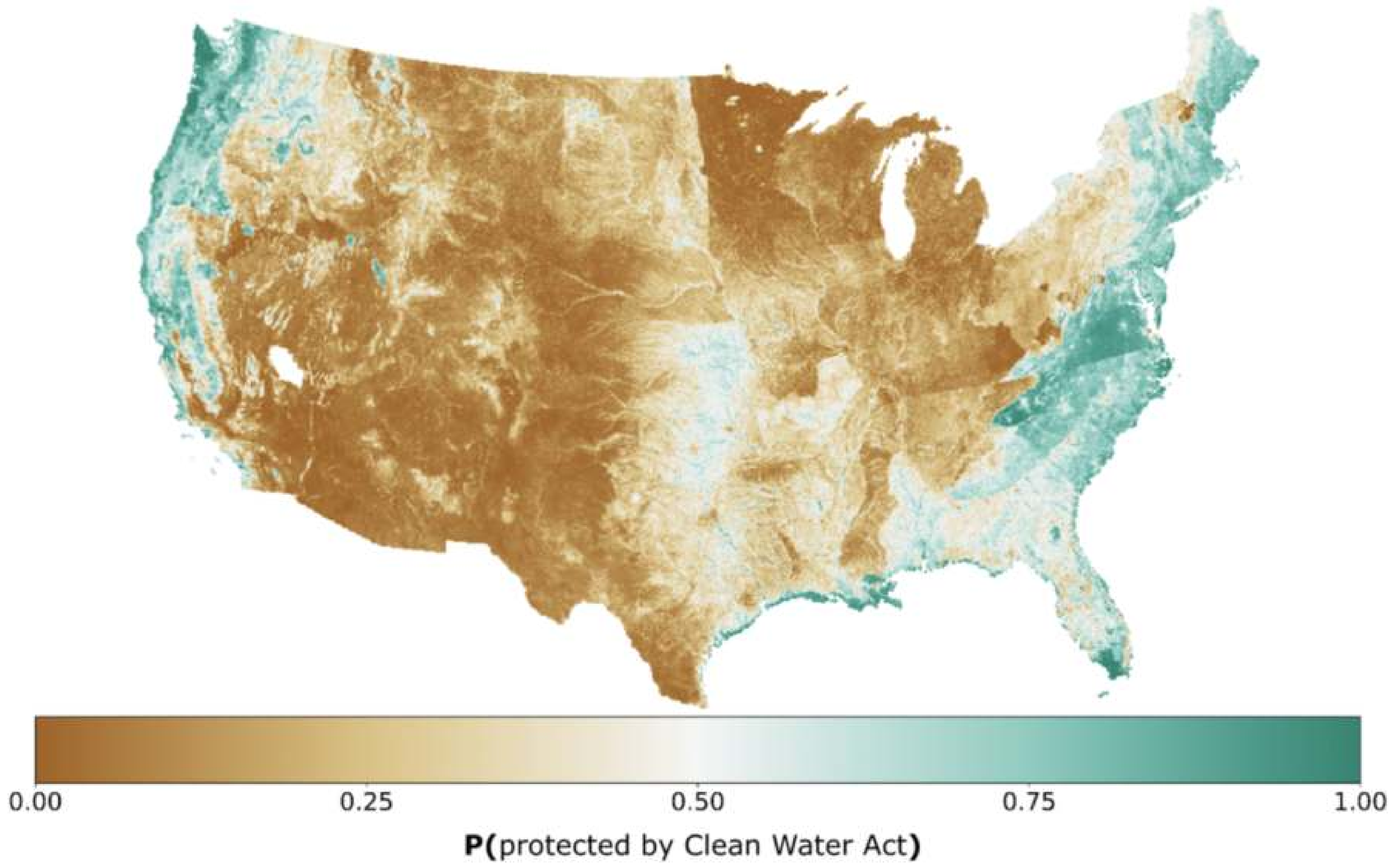

What crops are farmers growing?
How is crop choice changing over time?
Which agricultural practices, technologies, and policies lead to better outcomes for farmers and for the environment?

## Geospatial machine learning

Geospatial data is a new frontier for machine learning. Unlike natural images, satellite imagery is multi-spectral, multi-temporal, and multi-modal. We develop machine learning motivated by these characteristics.




## Remote sensing for causal inference

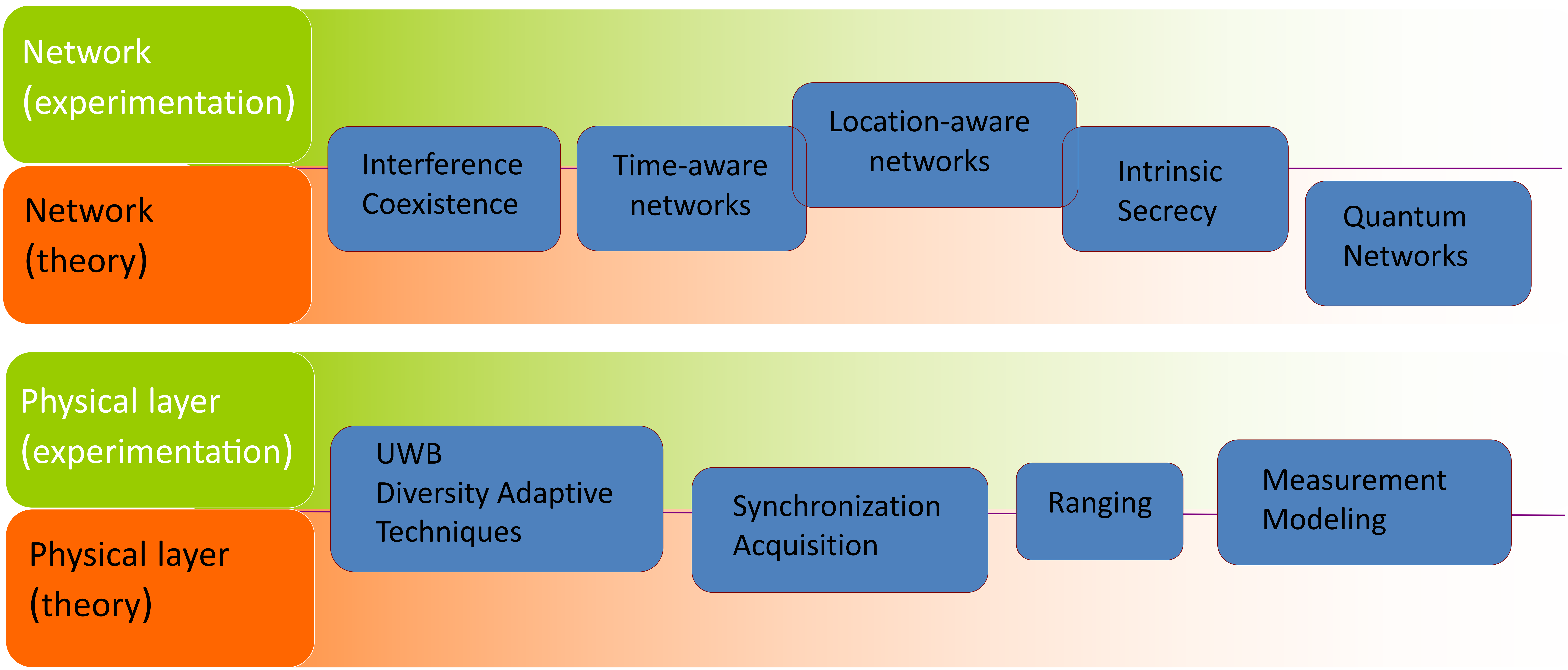



Ultimately, we want geospatial data to help us make better choices. We answer questions like: Do protected areas reduce deforestation? Which crop rotations improve soil health? How many stream miles and wetland acres are regulated by the Clean Water Act?

# Wireless Information & Network Sciences
## Moe Win

Research Vision: Our research combines

- ✓ theoretical analysis for determination of fundamental performance limits;
- ✓ the design of practical algorithms that approach such ultimate limits; and
- ✓ experimentation, both for validation and for developing realistic models



LIDS
RESEARCH GROUPS
2024-2025

# Cathy Wu

**We study AI & ML to solve hard optimization problems for next-generation mobility systems.**

Our vision is mobility systems that just work. Safe, efficient, sustainable, equitable.

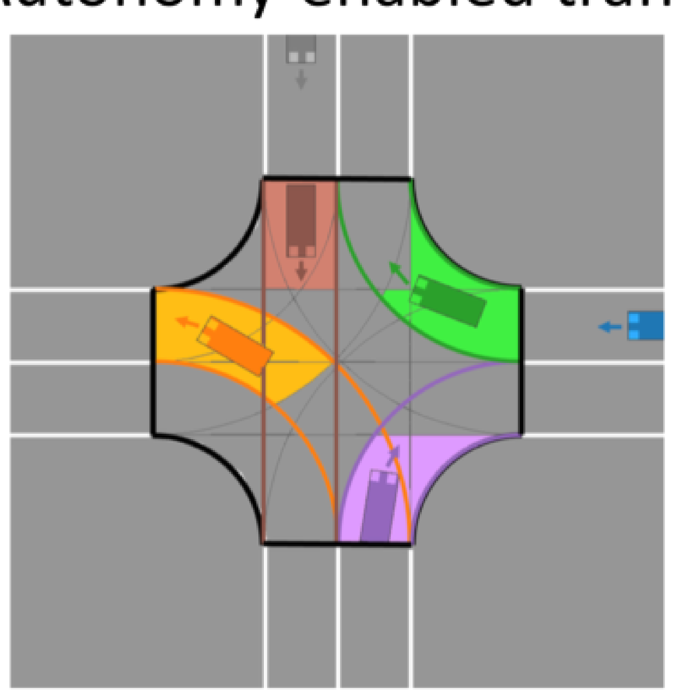
Designing and operating these systems requires solving many hard optimization problems.
Too many problems to derive algorithms by hand. **Can we automatically derive algorithms?**

## Research Topics

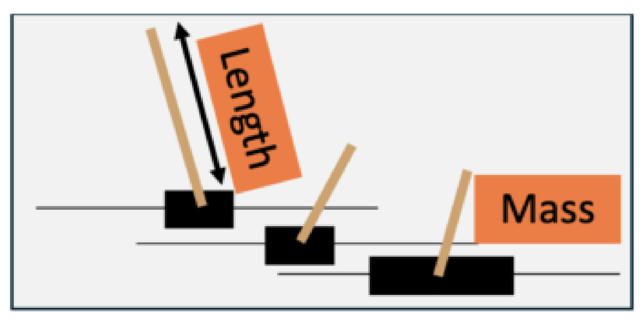
- Neural combinatorial optimization
  - Learning-guided optimization [1-2]
  - Foundation models
- Learning for generalization [3-4]
  - Contextual reinforcement learning
  - Long-term traffic simulation
- Multi-agent coordination [5-9]
  - Controllable congestion at scale
  - Multi-agent path finding & motion planning [10]
  - Offline reinforcement learning for traffic
- Reproducible research in engineering [11]

## Applications & Testbeds

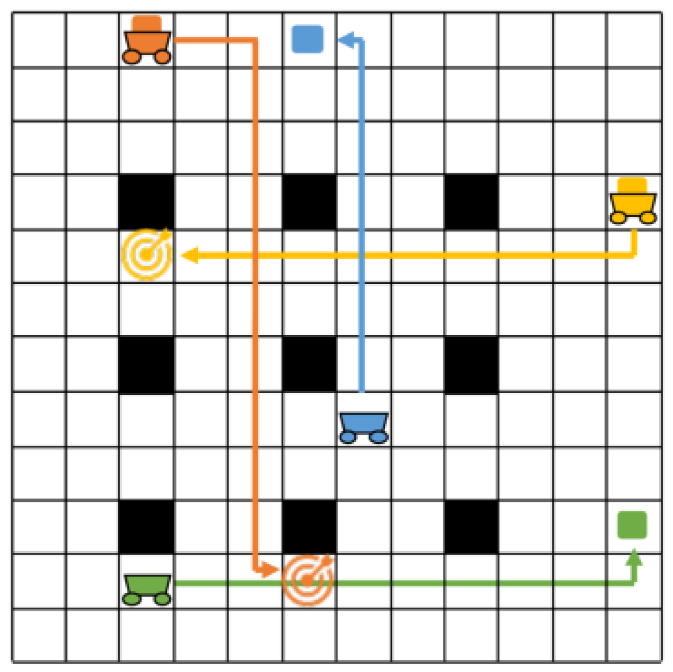
Autonomy-enabled traffic



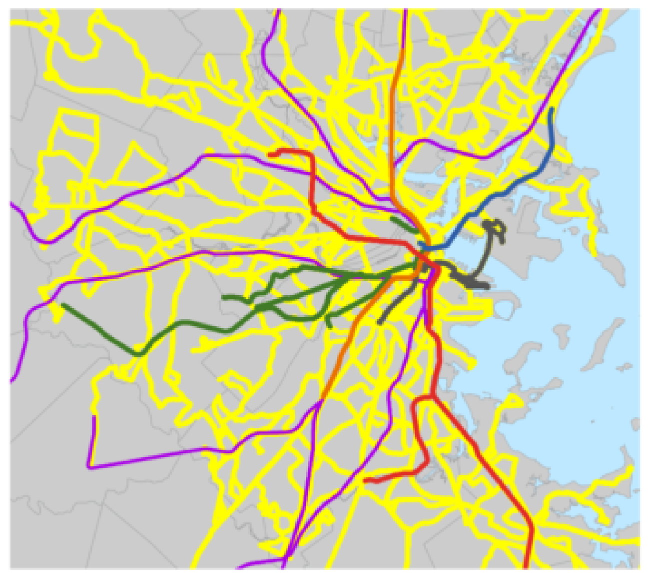
Common control systems



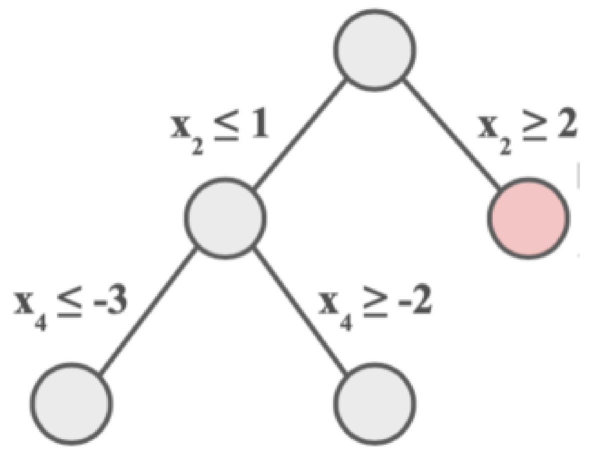

Multi-robot warehousing



Vehicle routing



Transit network planning




Mixed integer linear programming

[1] *NeurIPS*, 2021. Spotlight (<3%). [2] *NeurIPS*, 2023. [3] NSF CAREER Award, 2023. [4] *NeurIPS*, 2022. [5] *T-RO*, 2021. [6] *T-ASE*, 2022. [7] *T-RO*, 2023. [8] *T-CNS*, 2023. [9] *T-RO*, 2024. [10] ICLR, 2024. [11] Tutorial at ITSC 2024.

LIDS
RESEARCH GROUPS
2024-2025

# Zardini Lab

Driven by **societal challenges**, we develop **efficient computational tools**, and **algorithms** to formulate and solve complex, **compositional system design** and **autonomous decision-making problems**

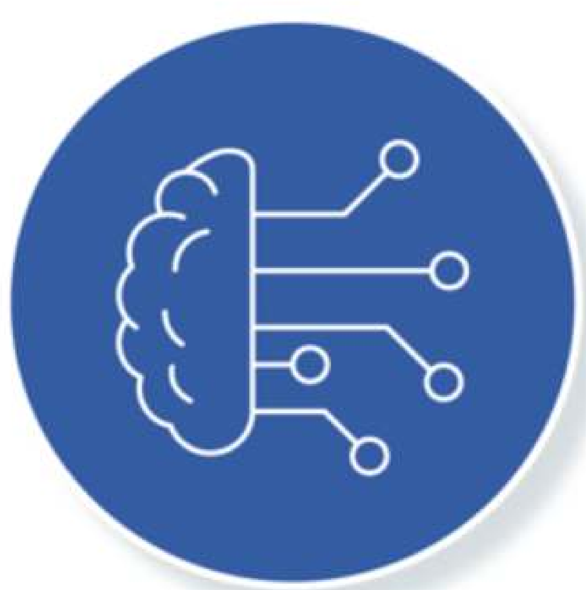
## Complexity when designing sociotechnical systems

### Large systems

- Many components
- Heterogeneous natures
- Multiple objectives

### Strategic interactions

- Many agents
- Heterogeneous interactions
- Conflicts/collaborations



*Modeling and Algorithmic Foundations*

- Compositional design optimization via domain theory and applied category theory
- Strategic interactions via game theory
- Planning and decision making processes under uncertainty

*Societal Applications*

- Networks, Logistics & Society-critical Infrastructure (e.g., urban transportation and maritime shipping)
- Embodied Intelligence, Robotics
- Aerospace, Automotive

*User-friendly Tools*

- Engaging with public authorities and industry partners
- Developing tools accessible to various stakeholders

# Other LIDS Faculty/PIs

**Stephen Bates** — *Statistical inference with AI systems; Data impacted by strategic behavior and information asymmetry; Shifting distributions and feedback loops*

**Priya Donti** — *Machine learning for forecasting, optimization, and control in high-renewables power grids*

**Marzyeh Ghassemi** — *The "Healthy ML" group at MIT focuses on creating and applying machine learning to understand and improve health in ways that are robust, private and fair. Health is important, and improvements in health improve lives*

**Kuikui Liu** — *High-dimensional geometry and analysis of Markov chains*

***Youssef Marzouk*** — *New methodologies for uncertainty quantification, Bayesian modeling and computation, data assimilation, experimental design, and machine learning in complex physical systems*

**Alexandre Megretski** — *Nonlinear system identification and model reduction*; *Nonlinear dynamical system analysis*; *Design and validation of hybrid control algorithms*; *various topics in Optimization*

**Sendhil Mullainathan** — *Uses machine learning to understand complex problems in human behavior, social policy, and medicine*

**Pablo Parrilo** — *Mathematical optimization, Machine learning, Control and identification, robustness analysis and synthesis, and the development and application of computational tools based on convex optimization and algorithmic algebra to practically relevant engineering problems*

**Manish Raghavan** — *Application of computational techniques to domains of social concern, including online platforms, algorithmic fairness/discrimination, and behavioral economics*

**Ashesh Rambachan** — *The intersection of econometrics and machine learning, with focuses on improving decision-making in high-stakes settings and improving our understanding of behavior*

**Philippe Rigollet** — *The intersection of Statistics, Machine learning, and Optimization, focusing primarily on the design and analysis of statistical methods for high-dimensional problems*

**Martin Wainwright** — *High-dimensional statistics, Information theory and statistics, and Statistical machine learning*

**Ashia Wilson** — *Development of statistical tools for aligning AI with social goals, methodological foundations of optimization and related topics*

For more details on these and other research being done at LIDS, please refer to the LIDS homepage.

**lids.mit.edu**