

# Institute for Data, Econometrics, Algorithms, and Learning (IDEAL)



## Participating Universities and Phase I Institutes

### IDEAL (Phase I)

Northwestern University (NU)

Toyota Technological Institute Chicago (TTIC)

University of Chicago (UC)

### Foundations of DS Institute (Phase I)

University of Illinois at Chicago (UIC)

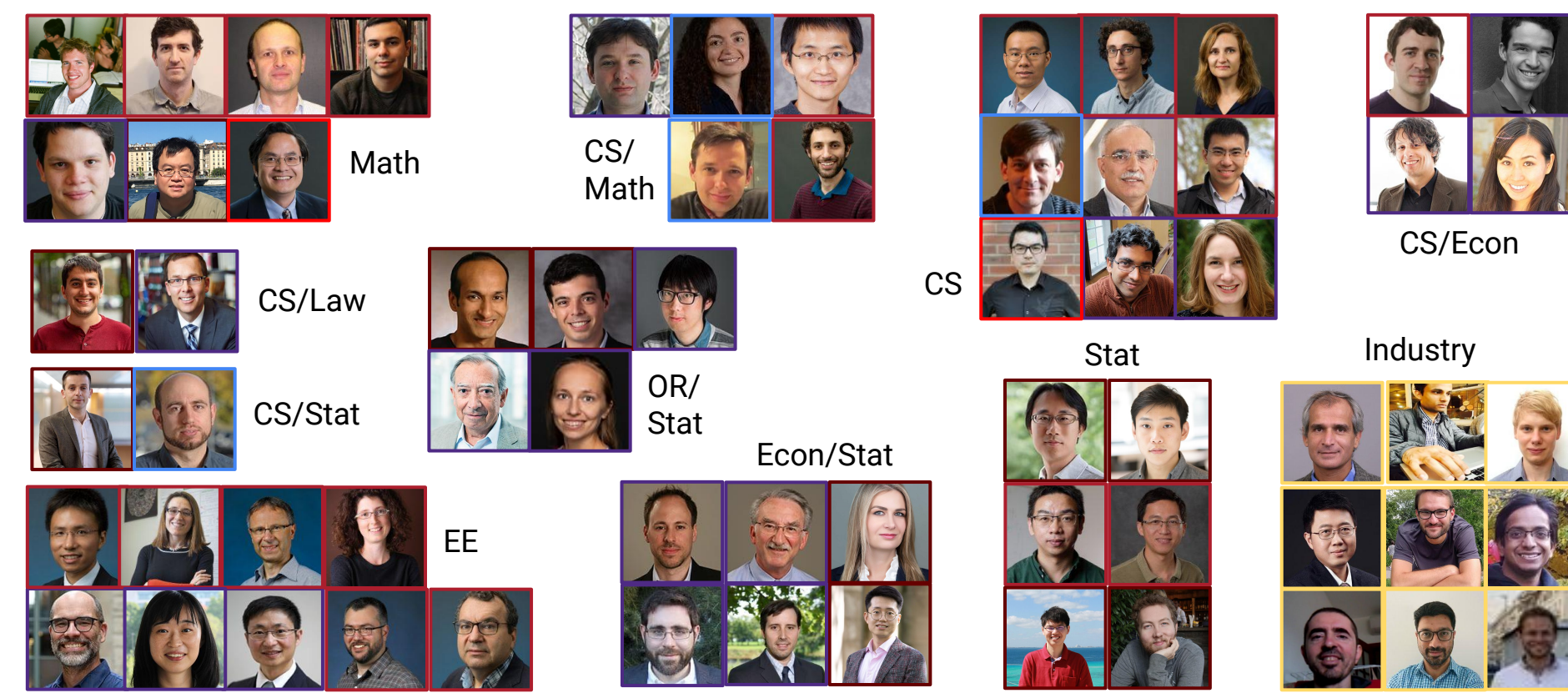
### External

Illinois Institute of Technology (IIT)

### Industrial Partner

Google Research

## Team Composition, by Field



## Goals of the Phase II Institute

The institute's research agenda focuses on solving key foundational problems in data science, ranging from the **core foundations** of data science to its **interfaces with other disciplines**.

Through its activities the institute will **broaden participation** in data science locally and nationally, build a **lasting research and educational infrastructure**, and **foster strong connections** within the Chicago community.

## A Comprehensive View of Data Science

We leverage the strong ties we've built among world-class research groups in **core-areas** of data science (CS, EE, probability, statistics) and exceptional researchers **outside the traditional center** of data science (econ, law, logic, operations research).

Additionally, the involvement of **Google Research** adds to our technical strengths and will allow us to have more **real-world impact** with a **direct connection to industry**.



## Combining and Expanding on our Strengths

Our institute will fully combine and expand on the strengths of both Phase I institutes to advance data science **locally** and **nationally**.

- **UIC Phase I** found novel ways to promote data science and transdisciplinary collaboration **locally**, within one university.
- **IDEAL Phase I** brought nearby universities together and fostered interactions among them and other institutions **across the country**.
- The addition of **IIT** will create a cohesive **community** in Chicago that will serve as a national resource for data science.

## Some Phase I Highlights

**UIC Phase I** found novel ways to promote data science and transdisciplinary collaboration **locally**, within one university.

- creation of data science degree
- new cross-departmental seminars
- 3 long-term visitors
- workshops for high schoolers
- strong growth in data science
- >20 funded students + 1 postdoc
- multiple research breakthroughs

**IDEAL Phase I** brought nearby universities together and fostered interactions among them and other institutions **across the country**.

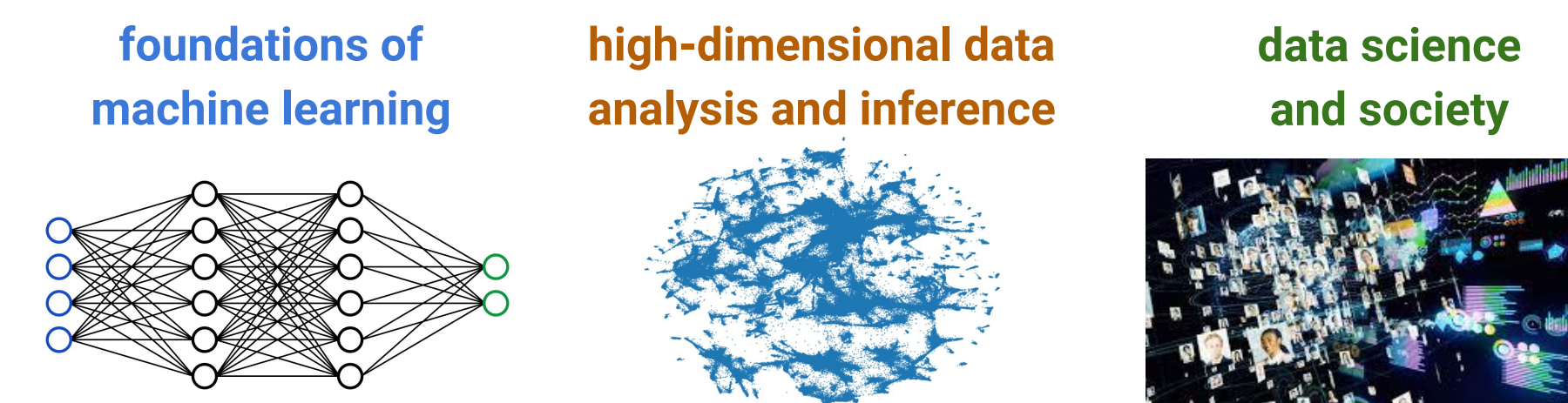
- 5 special quarters
- many short-term visitors
- 9 virtual workshops
- dozens of online seminars
- over 50 public videos of events
- many funded students + 3 postdocs
- multiple research breakthroughs

## Key Initiatives

Research Programs	Educational Programs
special programs	pre-REU workshops
summer workshops	teacher workshops
problem sessions	public lectures
cross-institutional seminars	exhibits at the MSI
Recurring Events	Personnel
annual meeting	undergraduate supervision
industry affiliates day	graduate fellows
weekly team meetings	postdoctoral program
cross-institution courses	visiting fellows

## Research Thrusts

The institute's research agenda focuses on solving key foundational data science problems in three main thrusts:



## Research Activities

Our research activities build on earlier successes and include:

- Activities structured around **topical special programs**. Two programs each year, each focused on one of the **8 research topics** and the final one on **future plans**.
- Each year, we will have **summer workshops**, each of 2-3 days duration. Some organized within IDEAL, some by external researchers who can nominate topics.
- **Problem brainstorming sessions** in the summer where students and team members split up into groups based on topic. Collaborations will continue after.
- Hybrid cross-department and cross-institutional **seminars** with online participation throughout the country.

## Research Thrusts and Topics

**Research goal:** Data science foundations with perspectives from related disciplines.

**Research thrusts** focus on most pressing problems in data science

**Research topics** overlap with research interests of team members

- weekly meetings from Sept 2021 of team members to learn research interests

Leveraged specific **strengths** of our universities

- CS+econ at NU, logic at UIC, high-dim stats at UC, ML theory at TTIC, scientific comp. at IIT

Encourage **collaboration** across multiple TRIPODS disciplines

## Research Topics and Topical Special Programs

One topical special program (~ 10 weeks) for each topic

### (T1) Foundations of Machine Learning

1. Deep Learning and Optimization
2. Reinforcement Learning and Control
3. Machine Learning and Logic

### (T2) High-dimensional Data Analysis and Inference

4. Networks and Statistical Inference
5. High-dimensional and Complex Data Analysis

### (T3) Data Science and Society

6. Trustworthy and Reliable Data Science
7. Interpretability, Privacy and Fairness
8. Data Science with Strategic Agents

## Broadening Participation

- Leveraging the lead institution **UIC's position as a Minority Serving Institution (MSI)**, a **Hispanic Serving Institution (HSI)**, and an **Asian American and Native American Pacific Islander-Serving Institution (AANAPISI)** through every aspect of our proposed activities, including community engagement, evaluation, student **recruitment**.
- Partnering with existing programs for undergraduates, including UIC LASURI, NU Women in OR/MS (WORMS), NU Society for Women Engineers.
- Reaching high school students through the Young Scholars Program, Math Circles of Chicago, Women in Engineering Summer Program.
  - 80 high school students participated last summer in the YSP run by Frietag and Perkins (UIC)

## Impact on Local Community

Museum of Science and Industry

- hands-on museum exhibits on how algorithms can derive useful insights from data, career-day events

Public lectures

- each research workshop will feature a public outreach lecture

Teacher workshops

- organized with Math Circles of Chicago for high school math and science teachers

High school lectures

- Institute faculty and postdocs will deliver lectures at Chicago high schools



## Applications in Science and Industry

- Google is an integral part of the center
- Industry affiliates day and industry collaborations
- Coordination with Argonne Labs, DPI, CQuB
- Applications of research, e.g.
  - deep-learning to error correcting codes
  - network analysis to diffusion of new agricultural technologies in developing countries
  - adversarially robust machine learning techniques in practical systems for computer vision
- Application to other disciplines (law, social sciences)



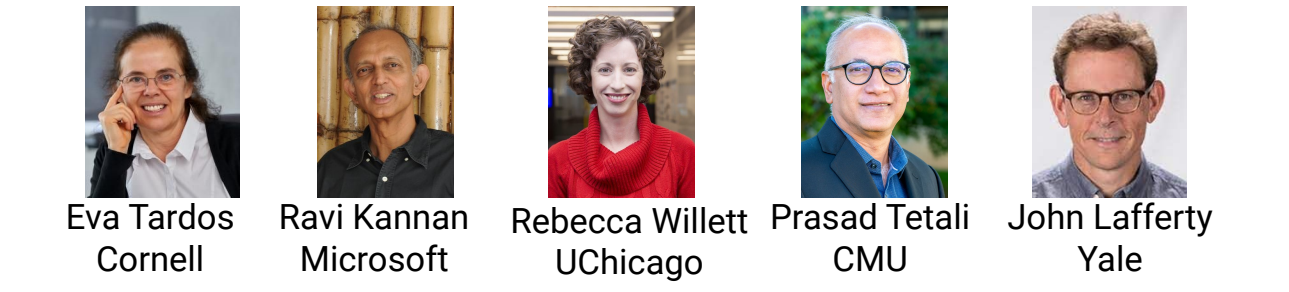
## Organization of the institute

Executive Committee consisting of 20 team members (including Site Directors)

Spread across institutions and all the disciplines

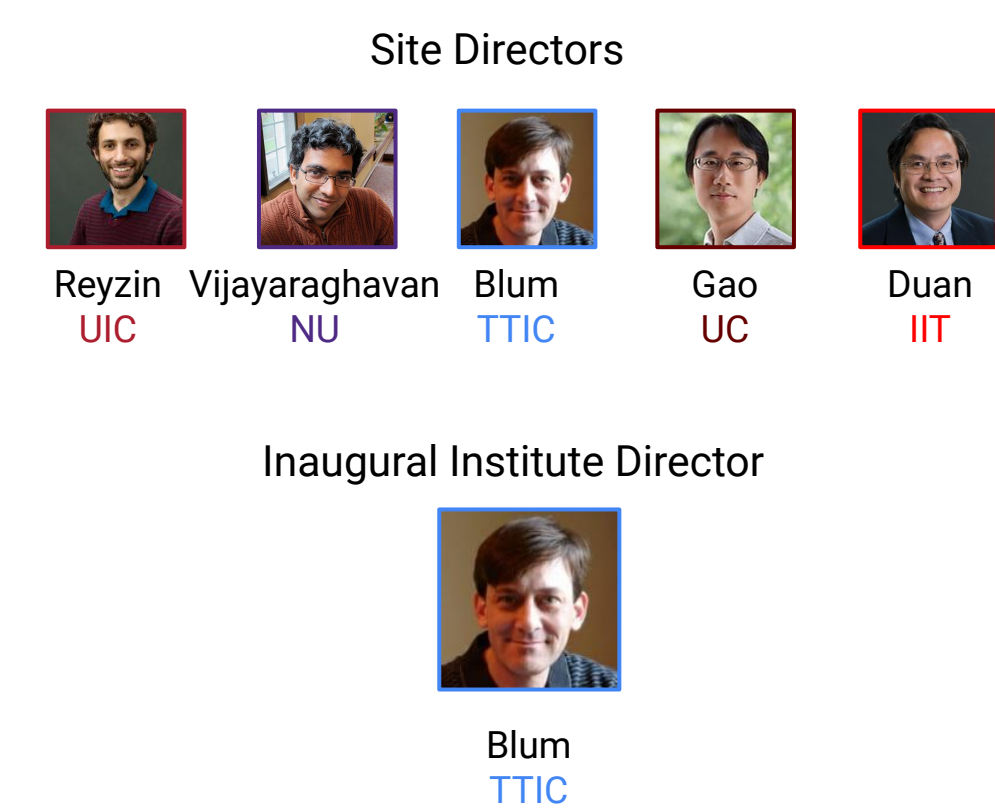
	UIC	NU	TTIC	UC	IIT
Site Director	Reyzin	Vijayaraghavan	Blum	Gao	Duan
Directors of Participation & Outreach	Perkins	Khalil			
Research Committee	Devoze, Y. Wu	Auffinger, Canay, Hart-Im, Novicki	Sebeo	Lim	
Education Committee	Kash	Obhucosian	K. Makarychev		
Workshops & Special Programs Director		Berry			
Executive Director			Y. Makarychev		
Director of Foundational Program	Tanwani	Khalil (alternate)			
Outreachperson					
Industry Coordinator	Zheleva (UIC), Awasthi (Google)				
Graphic Team Head					

External Advisory Board



## Site Directors and Institute Director

- one Site Director for each of the 5 participating universities
- same as the PIs of the grant
- in charge of overall administration and functioning of institute
- Institute Director
  - Rotating position with 1-yr term (TTIC, NU, and UIC site directors for first 3 years)
  - Rotation gives broad set of perspectives
  - Stability from site directors



## Evaluation of Outcomes

Involves both internal and external evaluation. Two components:

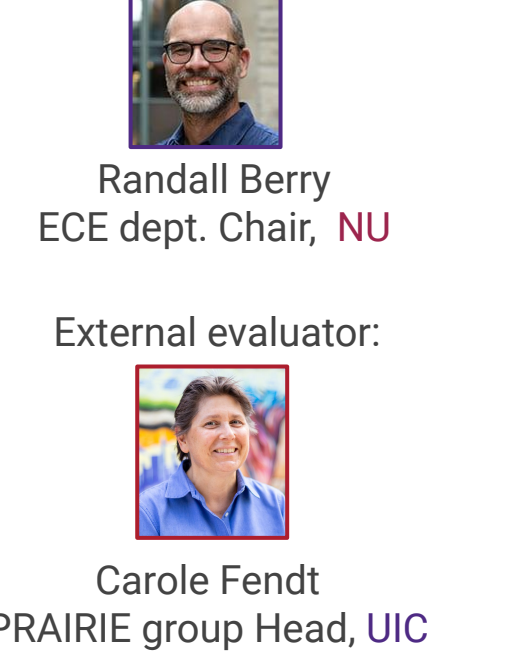
Research and interdisciplinary collaborations

- Research committee and site directors compile, evaluate surveys and quantifiable metrics like participation numbers, video views, publications, number of interdisciplinary and inter-university collaborations.
- External review of annual report by External Advisory Council

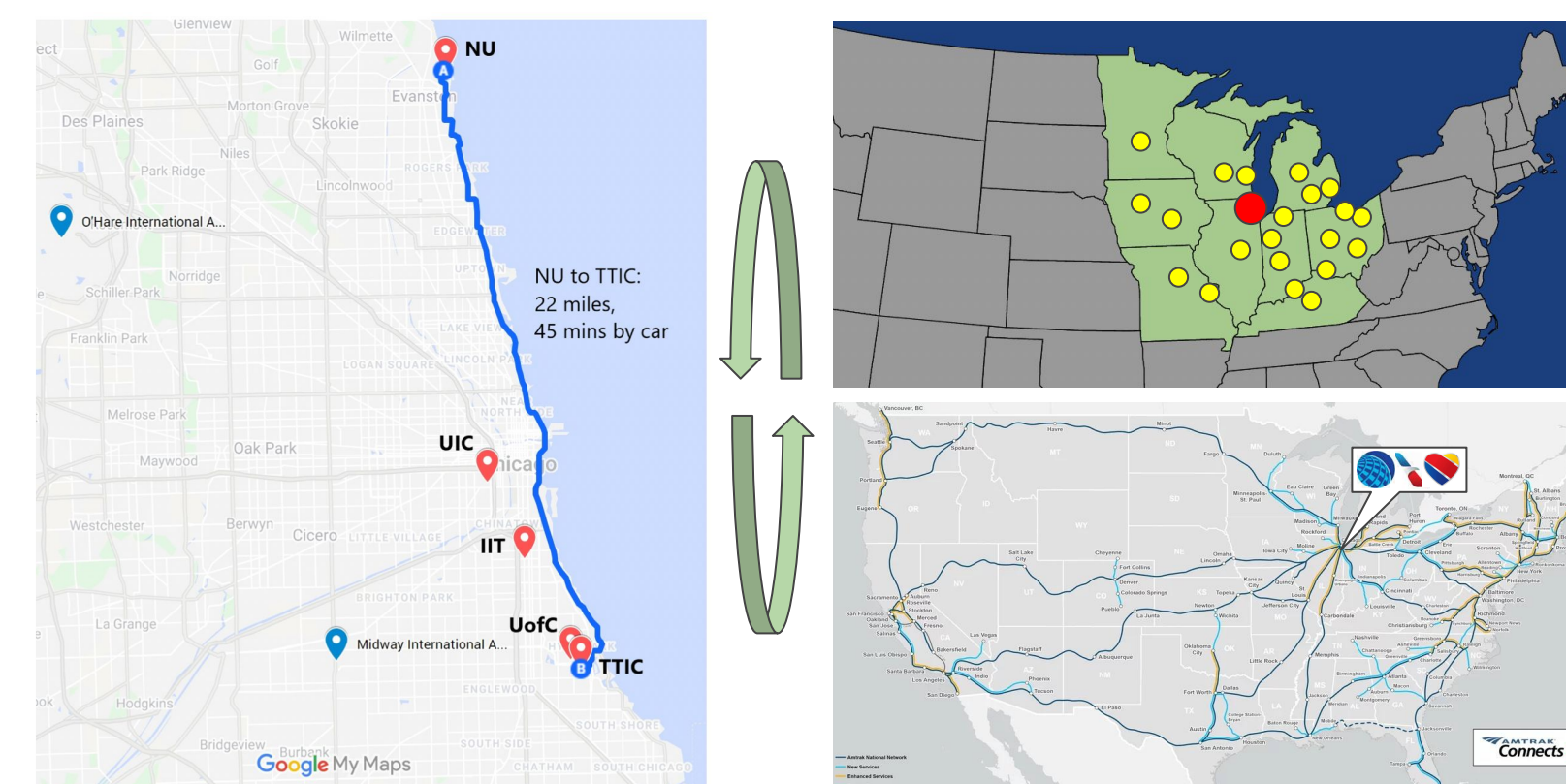
Education, training and outreach

- Education committee, directors of participation & outreach, site directors and postdoc program direction compile data
- External evaluation by experts: PRARIE group (UIC)

Evaluation director oversees the evaluation efforts:



## Local Community, Regional Connections, National Impact



## Broader Impact in the Chicago Area

- Connections to Chicago-area industry
  - Industry Affiliates day
  - Data science startups by Discovery Partners Institute, P33 Chicago
- Connections to applied labs and centers
  - Argonne, CQuB.
- Impact through outreach efforts
  - Museum of Science and Industry, Young Scholars Program, Math Circles of Chicago etc.



## Vision for the Institute

- New data science foundations bringing in perspectives from wide-range of disciplines.
- Creating new multidisciplinary research areas.
- Changing lives through Pre-REU workshops and community outreach activities.
- Creating a diverse workforce with multidisciplinary expertise contributing to advancements in data science nationwide
- Establishing a sustainable infrastructure that outlasts the funding.