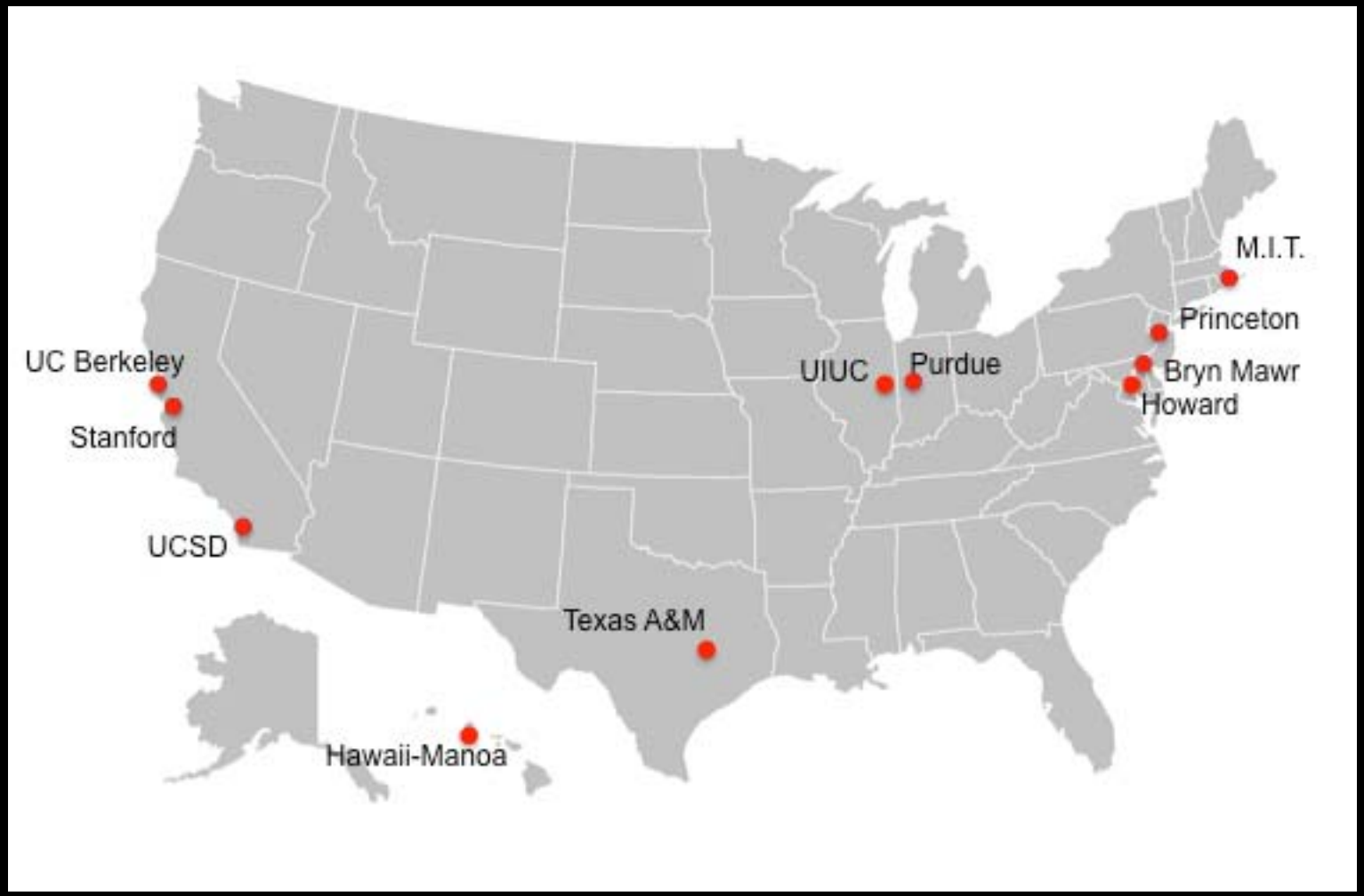
Case Study of Interdisciplinary Student Research Teams: Factors, Outcomes, And Lessons Learned

Brent T. Ladd, Director of Education
Center for Science of Information, Purdue University

National Science Foundation
Science and Technology Center



Center for Science of Information

Create a Community of Practice of young scholars around the emerging field of Science of Information

Pathways for Student Collaboration

Informal

- CSoI Member in our Center Network
- Conferences
- Summer Schools
- Annual Center Meetings

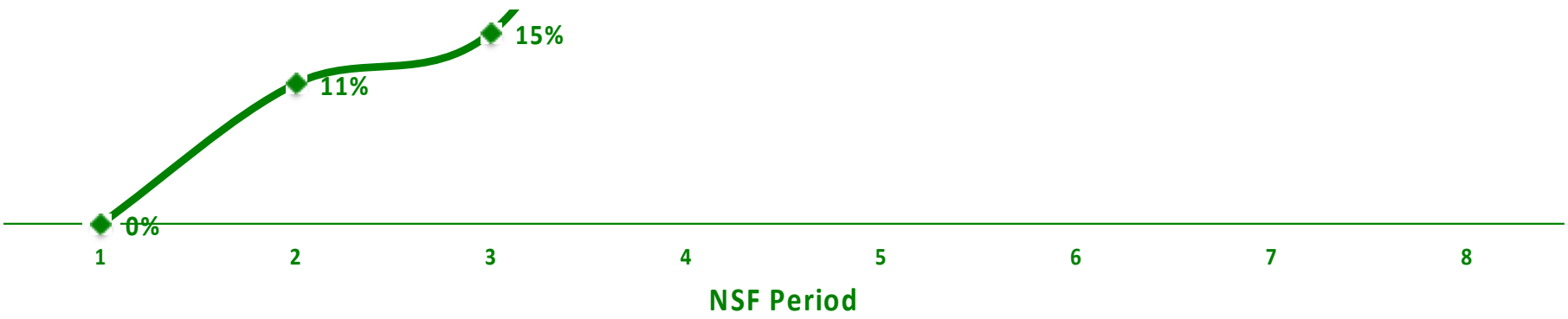
Formal

- Faculty “seed” Projects
- Co-Advisors
- Student Training Workshops
- Student-led Research Project Teams

Guiding Questions

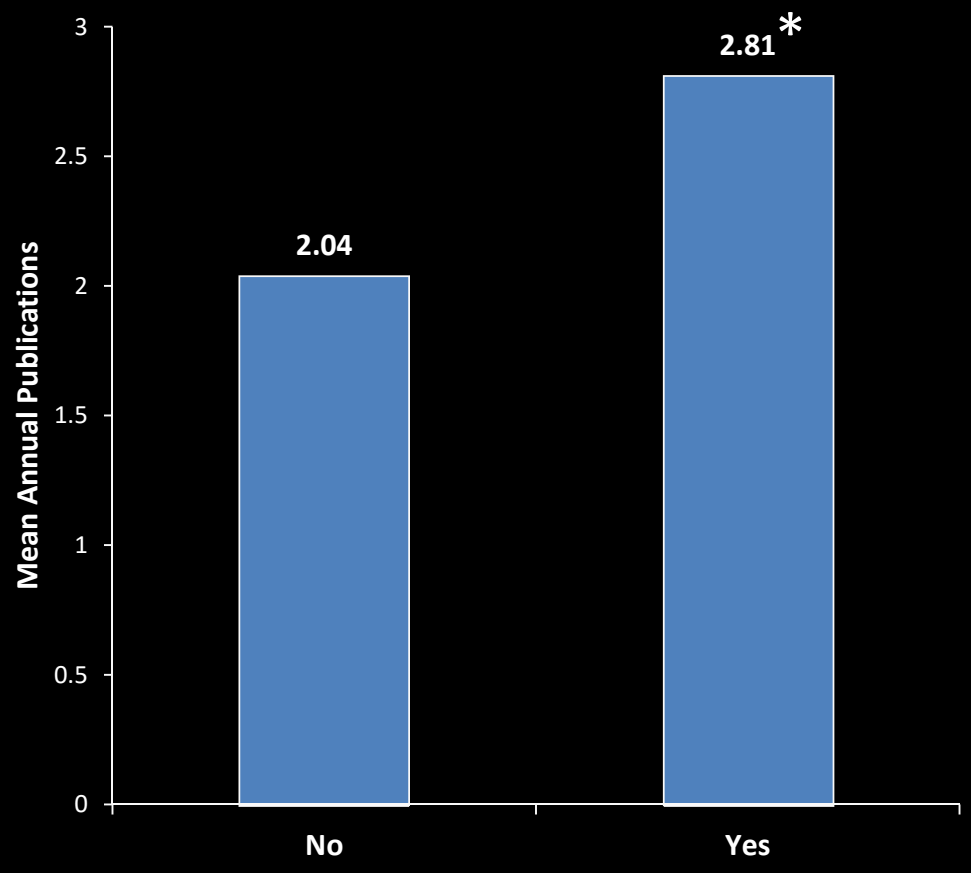
- Is there a significant relationship between community-based collaboration and publication productivity?
- Do factors of funding, gender, length of Center membership, participation level, or university influence publication productivity or rate of collaboration?
- What can be learned from student research team interactions, and their ability to address interdisciplinary questions?
- To what extent can a community of young scholars with large geographic distribution productively collaborate?

Percentage of Graduate Student Members Engaged in Community-based Collaboration



Factors included in the model for influence on publication productivity

- Community-based collaboration
- Gender
- Funding source
- Participation level
- Time spent as a Center member
- University



Community-Based Collaboration on Research Publications?

*F (1, 176) =11.89, $p < 0.001$

GLMM: BIC=473.87, random intercepts, Co-var. structure AR(1) Rho p=0.516, REML, Sequential Bonferroni

Pathways for Student Collaboration

Informal

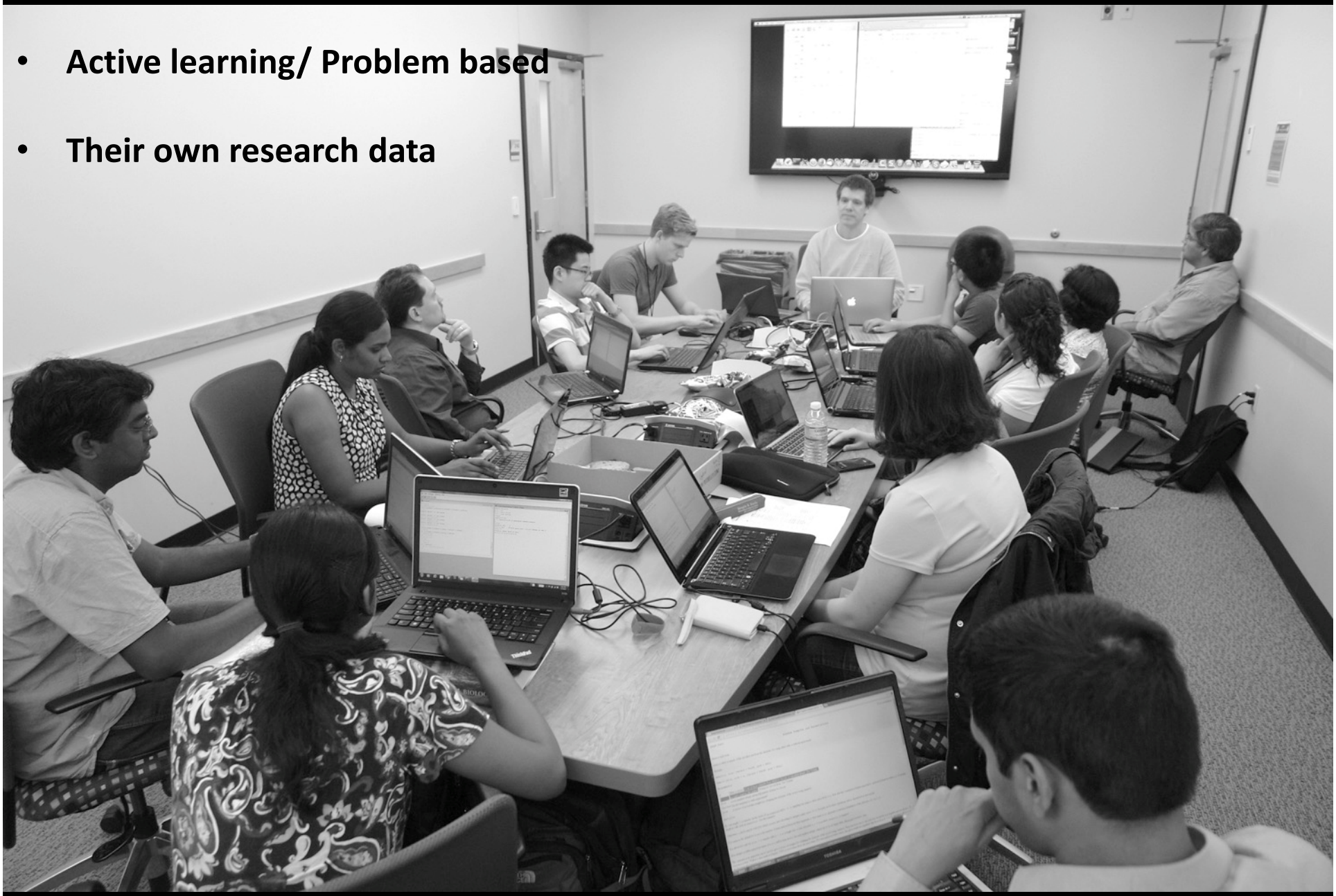
- CSoI Member in our Center Network
- Conferences (poster sessions)
- Summer Schools
- Annual Center Meetings

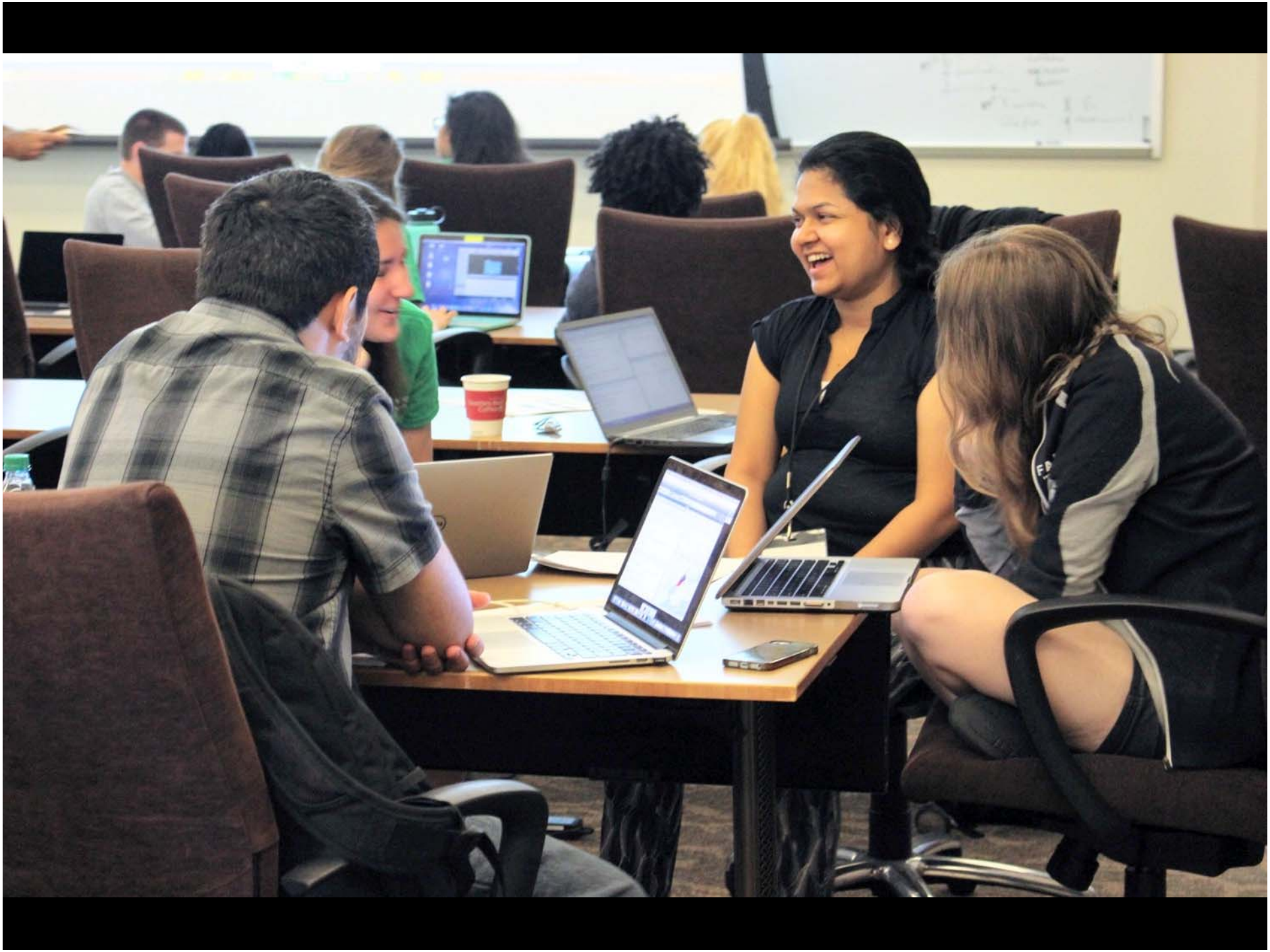
Formal

- Faculty Projects
- Co-Advisors
- Student Training Workshops
- Student-led Research Project Teams

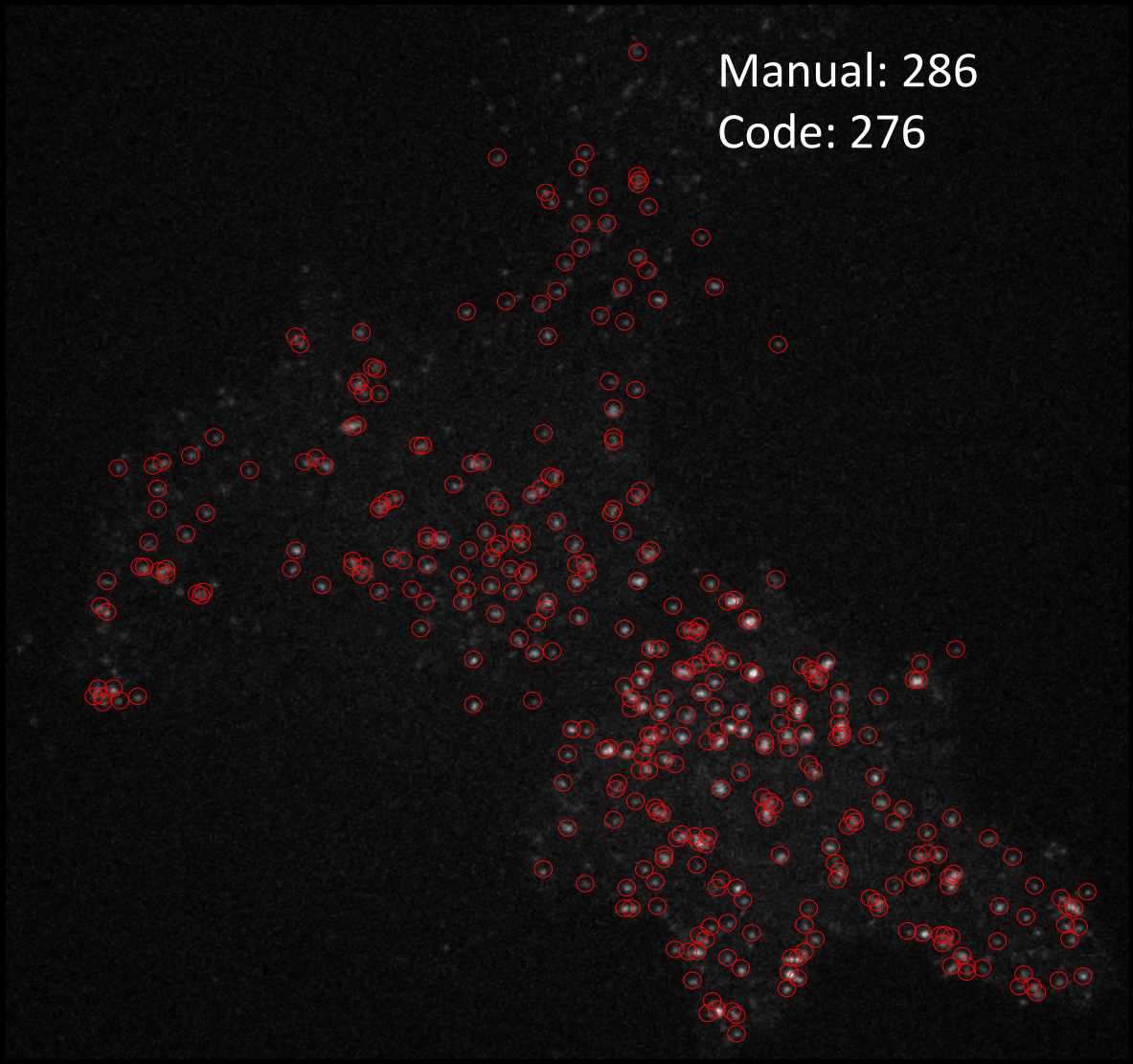

Training Workshop: Data Science &
Interdisciplinary Research Teams
Example

- **Active learning/ Problem based**
- **Their own research data**






GREY
BLACK
Internet background
1. R to G.
2. Fractions.
odwalla


Manual: 286
Code: 276

Professional Development Support

- NSF style proposal
- Roles of each participant
- How does it synergize with their graduate thesis goals
- Blessing and guidance of their respective major professors
- 6K per team for travel expenses per year

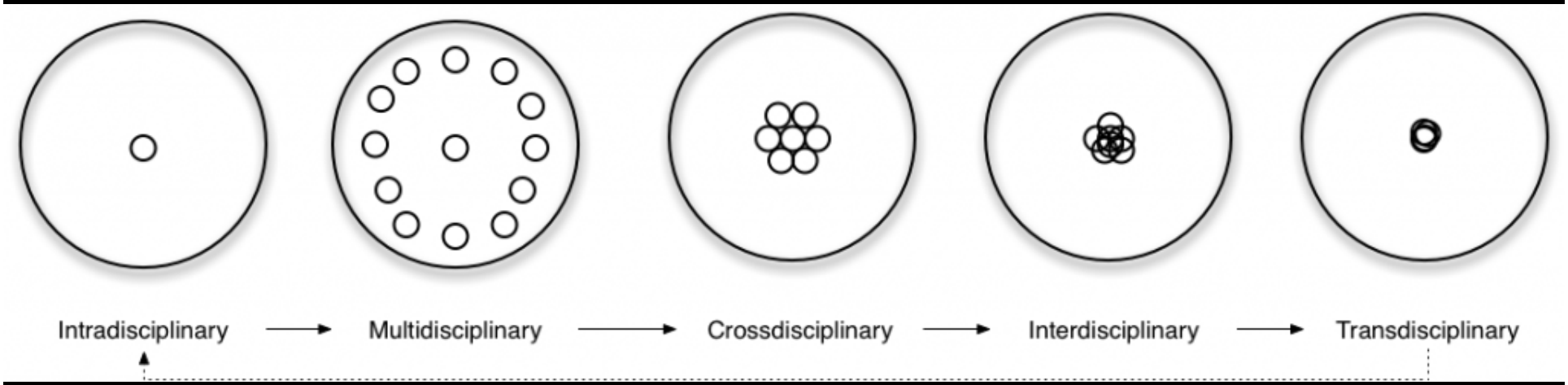
Year-Long Student-Led Teams

- 7 annual workshops
- 14 Student Research Teams
- 21 Universities
- 22 Departments
- 50/50 F/M ratio

Agronomy, Anthroplogy, Behavior and Brain Science, BioEngineering, Biology, Chemical Engineering, Civil Engineering, Computational Biology, Computer Engineering, Computer Science, Ecological Science and Engineering, Electrical and Computer Engineering, Electrical Engineering, Environmental Engineering, Forestry and Natural Resources, Geology, Languages, Mathematics, Medical, Physics, Sociology, Statistics

Student-Led Research Team Productivity

- 14 Student-Led Research Project Teams
- 44 Co-authored Conference Posters & Presentations
- 15 Co-authored Journal Papers



after Zeigler (1990), Jensenius (2012)

“Having access to the Center faculty and students, and participating in the workshops and schools benefited me greatly. It helped me to ask new and better questions about my own thesis topic, and gave me the chance to learn and work with other disciplines.”

Lessons Learned

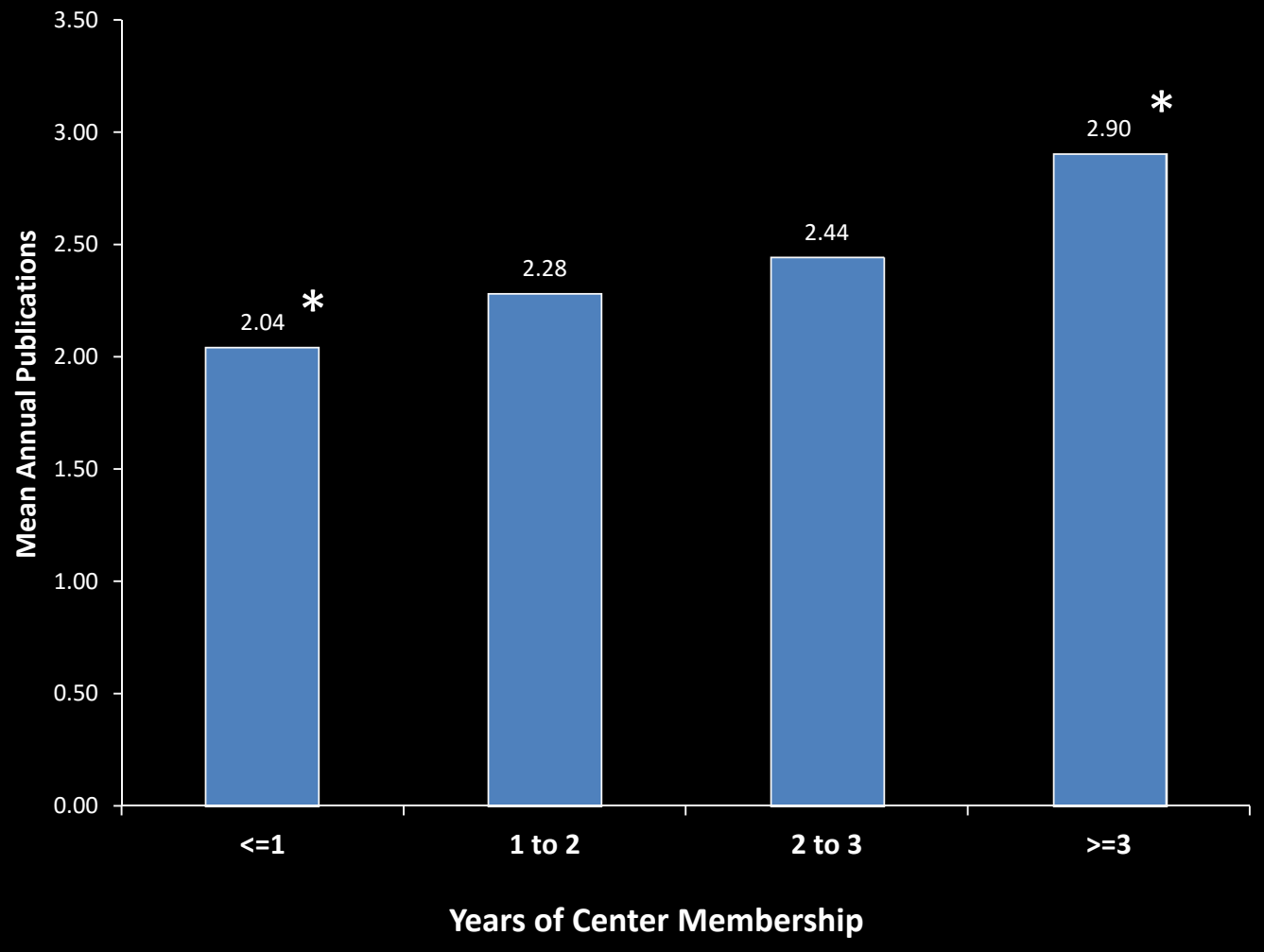
- Significant positive relationship exists between Center community-based collaborations by graduate students and their scholarly publication productivity
- Given a range of informal and formal pathways that encourage collaboration – graduate students in our community demonstrate capacity to successfully engage in interdisciplinary research
- Given even small amounts of travel and professional development support, our graduate students have successfully formed a community of practice (despite being geographically dispersed)
- This collaborative approach developed during graduate study appears to continue as they matriculate to post-doctoral and faculty positions

Questions?

Brent Thomas Ladd, Dir. of Education
Center for Science of Information
Computer Science Department
Purdue University
West Lafayette, Indiana, U.S.A.
laddb@purdue.edu
https://soihub.org

Made possible by grant NSF CCF-0939370

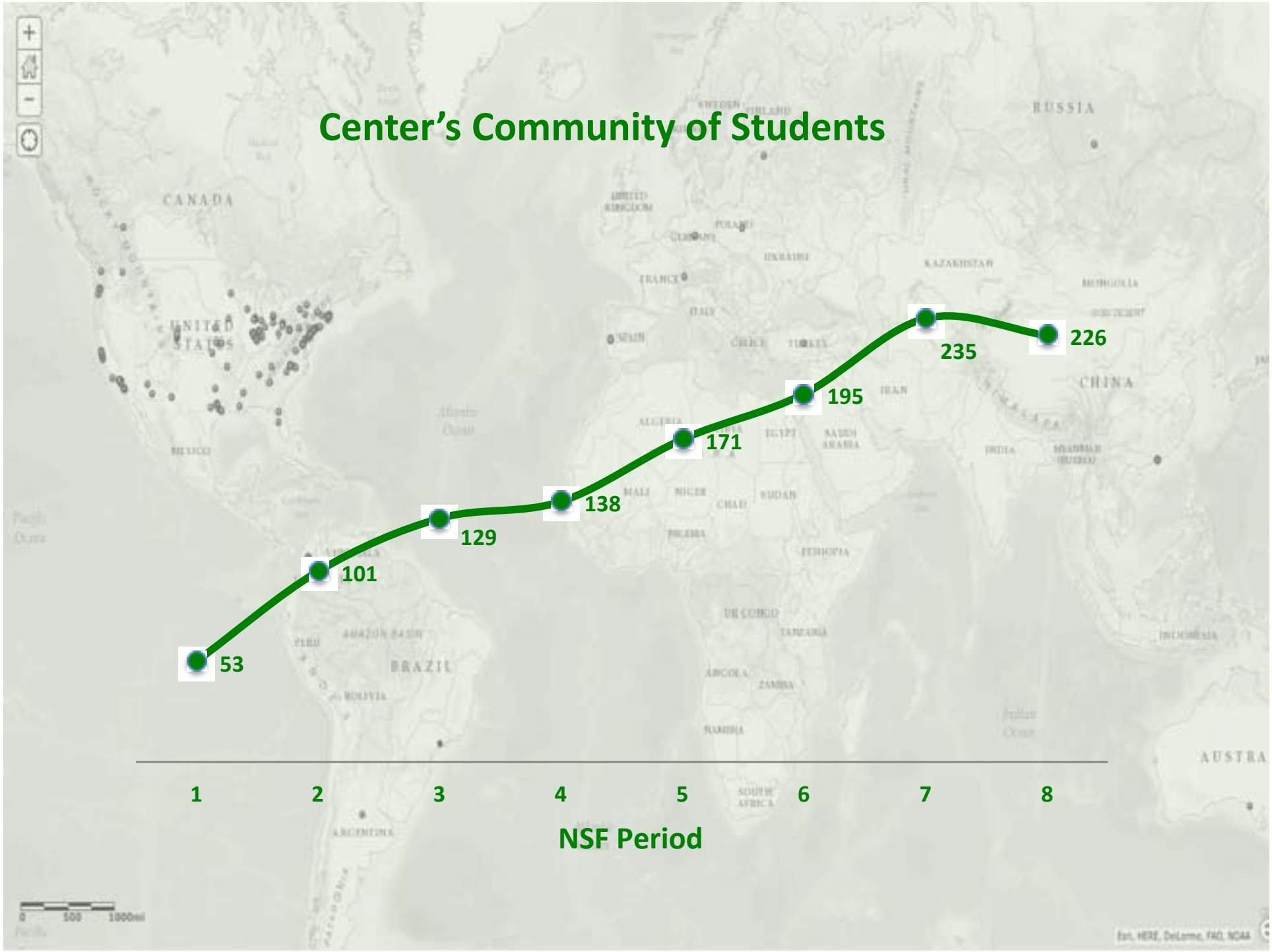
Additional Slides

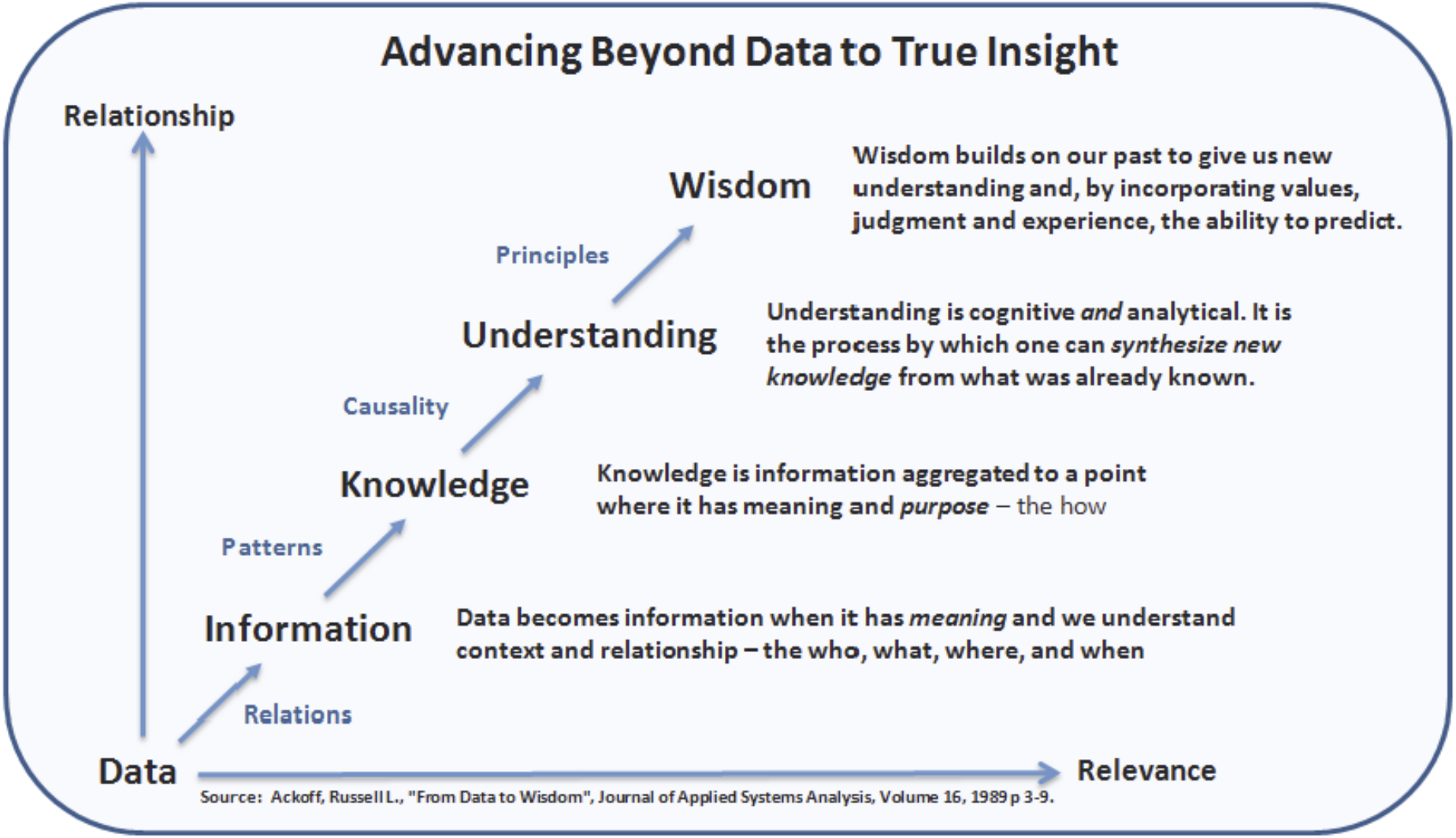


*n=256, F=3.25, p=0.023

Collaboration as Dependent Factor

- None of the independent factors measured revealed any significance in the model

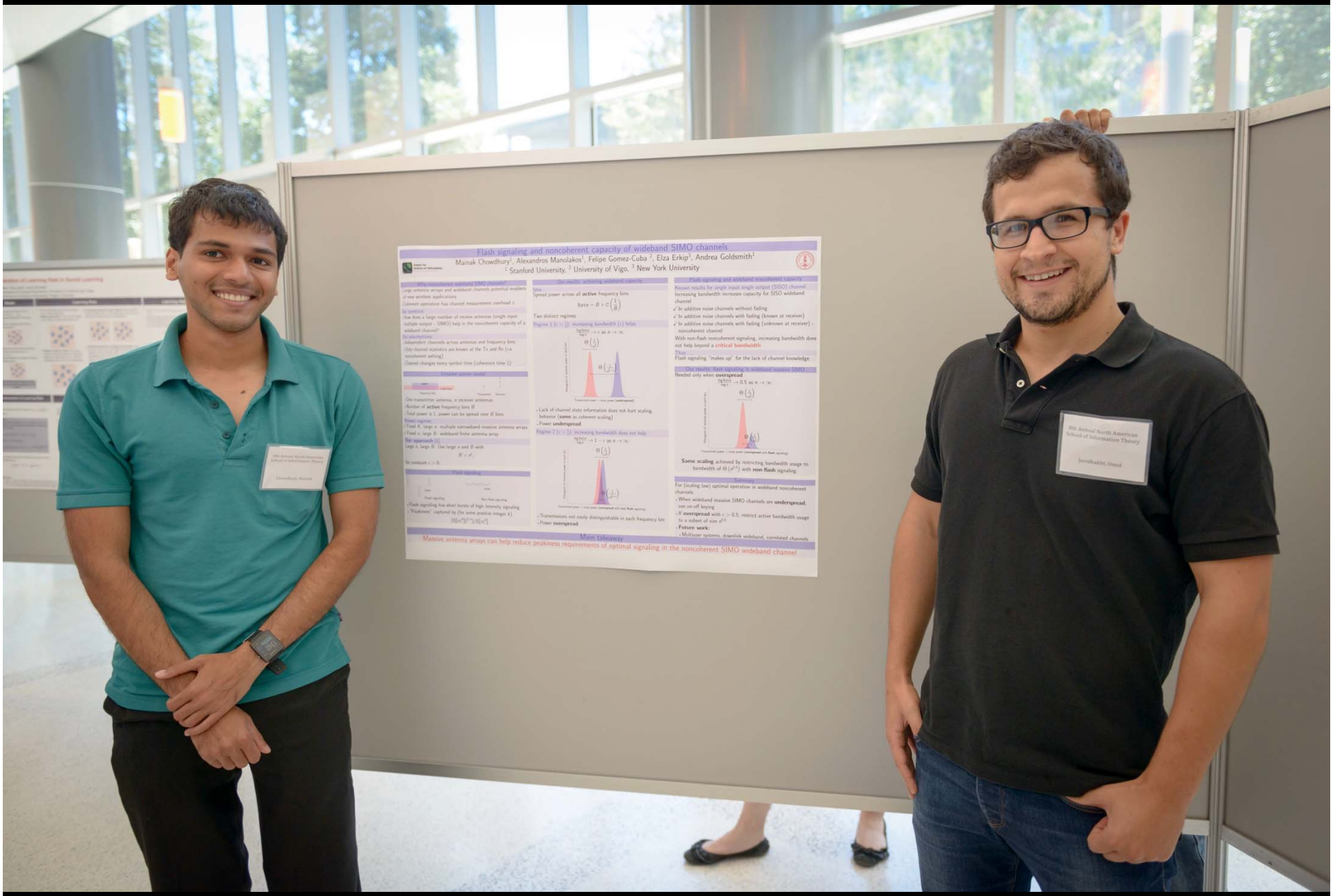

Center's Community of Students
53
101
129
138
171
195
235
226
1
2
3
4
5
6
7
8
NSF Period
CANADA
UNITED STATES
MEXICO
BRAZIL
RUSSIA
CHINA
KAZAKHSTAN
MONGOLIA
IRAN
INDIA
ETHIOPIA
SUDAN
NIGERIA
ALGERIA
EGYPT
SAUDI ARABIA
UKRAINE
FRANCE
ITALY
SPAIN
GREECE
TURKEY
ARGENTINA
BOLIVIA
PERU
ANGOLA
ZAMBIA
NAMIBIA
TANZANIA
DR CONGO
INDONESIA
AUSTRA


Advancing Beyond Data to True Insight
Relationship
Wisdom
Wisdom builds on our past to give us new understanding and, by incorporating values, judgment and experience, the ability to predict.
Principles
Understanding
Understanding is cognitive *and* analytical. It is the process by which one can *synthesize new knowledge* from what was already known.
Causality
Knowledge
Knowledge is information aggregated to a point where it has meaning and *purpose* – the how
Patterns
Information
Data becomes information when it has *meaning* and we understand context and relationship – the who, what, where, and when
Relations
Data
Relevance
Source: Ackoff, Russell L., "From Data to Wisdom", Journal of Applied Systems Analysis, Volume 16, 1989 p 3-9.


data
-
information
-
insights
www.soihub.org


Sheila
Rosenberg
Frank
DeVilbiss


DASANI


Flash signaling and noncoherent capacity of wideband SIMO channels
Mainak Chowdhury[1], Alexandros Manolakos[1], Felipe Gomez-Cuba [2], Elza Erkip[1], Andrea Goldsmith[1]
[1] Stanford University, [2] University of Vigo, [3] New York University
Why noncoherent wideband SIMO channels?
Large antenna arrays and wideband channels potential enablers of new wireless applications
Coherent operation has channel measurement overhead ☹
So questions:
How does a large number of receive antennas (single input multiple output - SIMO) help in the noncoherent capacity of a wideband channel?
So assumptions:
Independent channels across antennas and frequency bins
Only channel statistics are known at the Tx and Rx (i.e. noncoherent setting)
Channel changes every symbol time (coherence time 1)
Detailed system model
One transmitter antenna, n receiver antennas
Number of active frequency bins B
Total power is 1, power can be spread over B bins
Power regimes:
Fixed K, large n: multiple narrowband massive antenna arrays
Fixed n, large B: wideband finite antenna array
Our approach [1]:
Large n, large B: Use large n and B with
$B = n^{\epsilon}$,
for constant $\epsilon > 0$.
Flash signaling
Flash signaling
Non-flash signaling
Flash signaling has short bursts of high intensity signaling
"Flashness" captured by (for some positive integer k)
$\{\mathbb{E}[x^2]\}^{k/2}/\mathbb{E}[x^k]$.
Our results: achieving wideband capacity
Idea:
Spread power across all active frequency bins
$\text{Rate} = B \times C\left(\frac{1}{B}\right)$
Two distinct regimes
Regime 1 ($\epsilon < 1$): increasing bandwidth (ϵ) helps
$\frac{\log \text{Rate}}{\log n} \to \epsilon$ as $n \to \infty$
$\Theta\left(\frac{1}{n}\right)$
$\Theta\left(\frac{1}{n^{\epsilon}}\right)$
Transmitted power → noise power (underspread)
Lack of channel state information does not hurt scaling behavior (same as coherent scaling)
Power underspread
Regime 2 ($\epsilon > 1$): increasing bandwidth does not help
$\frac{\log \text{Rate}}{\log n} \to 1 - \epsilon$ as $n \to \infty$
$\Theta\left(\frac{1}{n}\right)$
$\Theta\left(\frac{1}{n^{\epsilon}}\right)$
Transmitted power → noise power (overspread with non flash signaling)
Transmissions not easily distinguishable in each frequency bin
Power overspread
Flash signaling and wideband noncoherent capacity
Known results for single input single output (SISO) channel
Increasing bandwidth increases capacity for SISO wideband channel
✓ In additive noise channels without fading
✓ In additive noise channels with fading (known at receiver)
✓ In additive noise channels with fading (unknown at receiver) - noncoherent channel
With non-flash noncoherent signaling, increasing bandwidth does not help beyond a critical bandwidth
Thus:
Flash signaling "makes up" for the lack of channel knowledge
Our results: flash signaling in wideband massive SIMO
Needed only when overspread
$\frac{\log \text{Rate}}{\log n} \to 0.5$ as $n \to \infty$
$\Theta\left(\frac{1}{n}\right)$
$\Theta\left(\frac{1}{n^{0.5}}\right)$
Transmitted power → noise power (underspread with flash signaling)
Same scaling achieved by restricting bandwidth usage to bandwidth of $\Theta(n^{0.5})$ with non-flash signaling
Summary
For (scaling law) optimal operation in wideband noncoherent channels
When wideband massive SIMO channels are underspread, use on-off keying
If overspread with $\epsilon > 0.5$, restrict active bandwidth usage to a subset of size $n^{0.5}$
Future work:
Multiuser systems, downlink wideband, correlated channels
Main takeaway
Massive antenna arrays can help reduce peakiness requirements of optimal signaling in the noncoherent SIMO wideband channel
Chowdhury, Mainak
9th Annual North American School of Information Theory
Javadtalebi, Omid