UA-TRIPODS

Helen Zhang
Statistics

Stephen Kobourov
CS

Joe Watkins
Math/Stat

David Glickenstein
Mathematics

What is TRIPODS?

Transdisciplinary **R**esearch **i**n **P**rinciples **o**f **D**ata **S**ciences

- **TRI**: **Computer Science, Mathematics, Statistics**
- Focus on theoretical foundations for data science

NSF TRIPODS-related Programs

- TRIPODS Phase I (2017-2020)
- TRIPODS+X (2018-2021)
- HDR TRIPODS Phase I (2019-2022)
- TRIPODS Phase II (2020-2025)

THE UNIVERSITY OF ARIZONA

What is Foundations of Data Science?

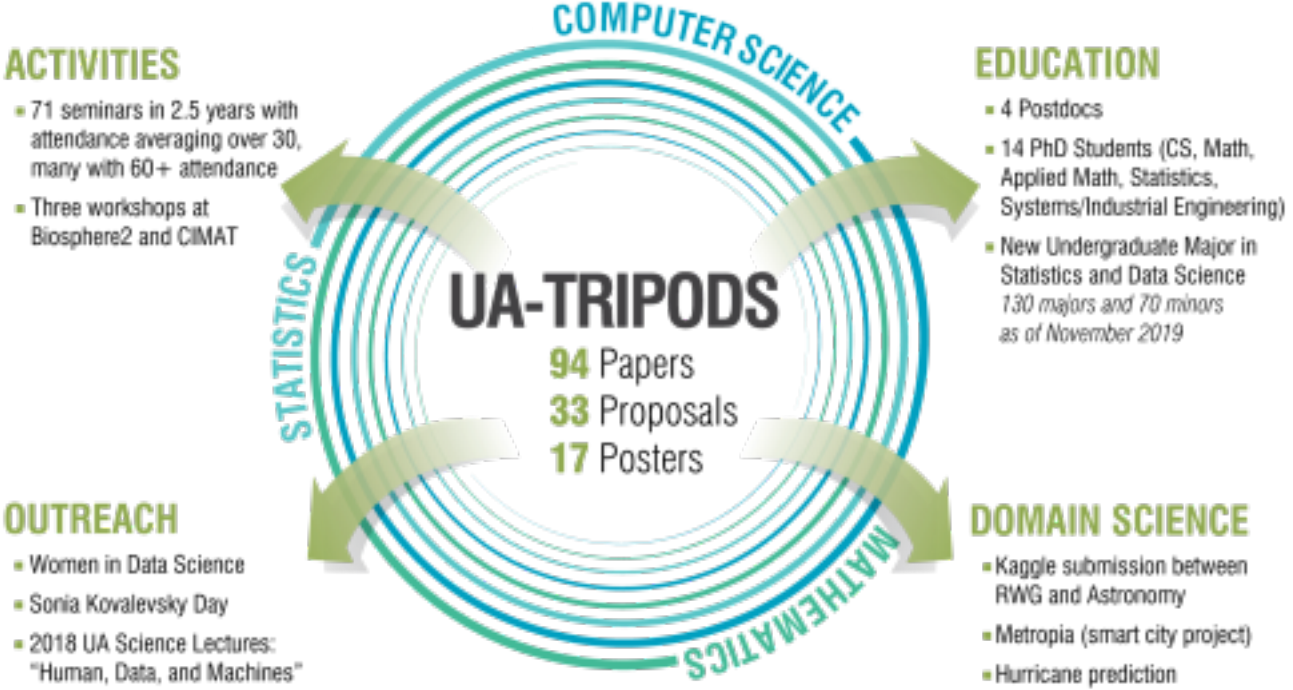
Our Perspective: Use principles of mathematics, statistics, and theoretical computer science to develop theory and methods to understand, analyze, and interpret data, focusing on

- How does it work? We want a rigorous theoretical framework.
- Why does it work? We want to be able to generalize.
- When will it work? We want to know the limitations.

Can the foundational disciplines work together?

Before 2017, few collaborations at UA among CS/Math/Stats.

Today:



THE UNIVERSITY OF ARIZONA

Research Infrastructure

8 Research Working Groups (RWGs)

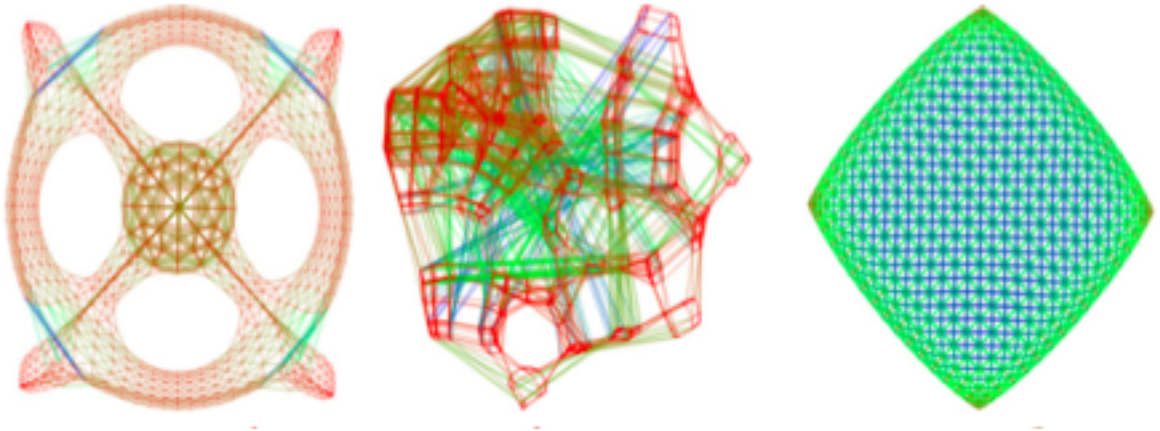
- Teams of CS/Math/Stats PIs and SPs
- Weekly meetings
- Clear research problems
- Postdocs, PhD students in each group

Joint Seminars and Conferences

- Weekly joint TRIPODS seminars (30 per year, 40-60 participants)
- Annual TRIPODS conferences (50-70 participants)

Research Working Groups

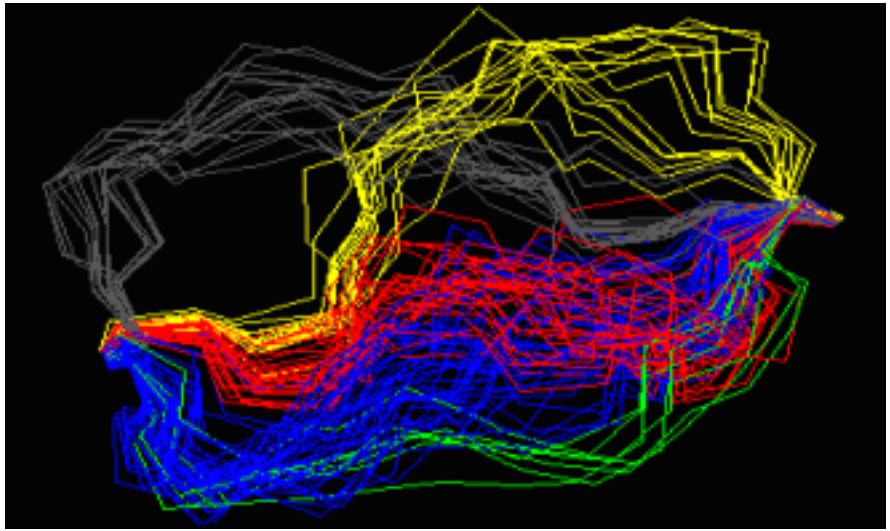
- **RWG 1: Theory of Large Network and Graphs**
- **RWG 2: Theory of Data Visualization/Interpretation**
- **RWG 3: Theoretical Research of Imaging Sciences**
- **RWG 4: Statistical Modeling for NLP**



THE UNIVERSITY OF ARIZONA

Research Working Groups

- **RWG 5: Theoretical Analysis for Optimization**
- **RWG 6: Analyzing Large-scale Point-set Data**
- **RWG 7: Bayesian Methods for Big Data**
- **RWG 8: Theory for Time-dependent Data Analysis**



THE UNIVERSITY OF ARIZONA

RWG1: Theory of Large Graphs & Networks

Key Participants:

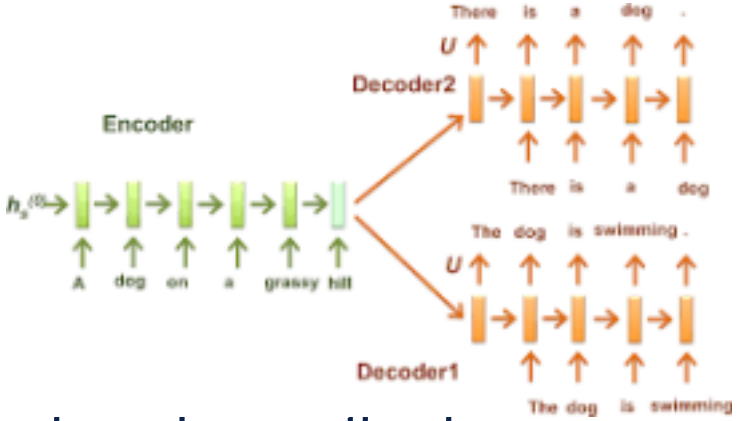
- Efrat (CS)
- Glickenstein (Math)
- Joly (CIMAT)
- Levine (CS)
- Lega (Math)
- Kobourov (CS)
- Sethuraman (Math)
- Tang (Stats)
- Watkins (Stats)

THE UNIVERSITY OF ARIZONA

RWG4: Natural Language Processing

Research Directions:

1. Trustworthy and explainable AI
2. Machines that reason over text



Research problems:

1. Human-in-the-loop ML via encoder-decoder methods
2. Information theory (informativeness) for inference
3. Debiasing methods to generalize beyond given data

THE UNIVERSITY OF ARIZONA

Research Highlights

Ahmed, Sahneh, Efrat, Glickenstein, Gronemann, Heinsohn, Kobourov, Spence, Watkins (2018). “Multi-Level Steiner Trees.” 17th Symposium on Experimental Algorithms (SEA).

First CS-Math-Stats Publication in UA History

THE UNIVERSITY OF ARIZONA

Additional Funding

Related Projects

- NSF TRIPODS+X Innovation Lab: *“from lemons to lemonade”*
- NSF TRIPODS+X Research
- NSF RTG
- NSF REU

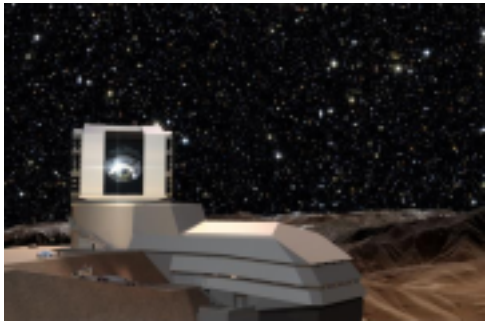
Highlights

Calendar

Conferences

Public Lecture Series

Penguin AI Club



LSST

Women in DS

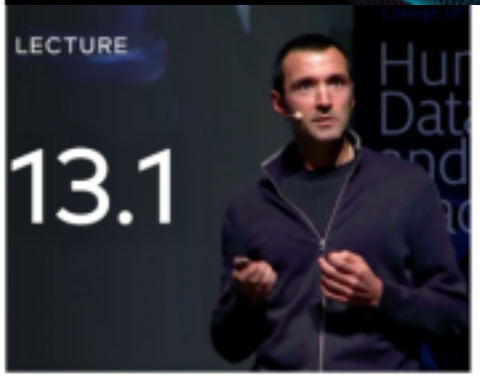
THE UNIVERSITY OF ARIZONA

Unique Public Lecture Series: 2018

Humans, Data and Machines

JANUARY 22, 2018 ——— FEBRUARY 26, 2018

In our automated lives, we generate and interact with unprecedented amounts of data. This sea of information is constantly searched, catalogued, analyzed and referenced by machines with the ability to uncover patterns unseen by their human creators. These new insights have far reaching implications for our society. From our everyday presence online, to scientists



LECTURE JANUARY 22, 2018

STEPHEN KOBOUROV

Problem Solving With Algorithms

LECTURE JANUARY 29, 2018

MIHAI SURDEANU

The Minds of Machines

LECTURE FEBRUARY 5, 2018

NIRAV MERCHANT

Working Alongside Thinking Machines

THE UNIVERSITY OF ARIZONA

TRIPODS Postdocs

Keaton Hamm

- Involved in three RWGs (RWG1: Large Scale Networks, RWG2: Visualization, and RWG3: Imaging)
- Applications to imaging and medicine
- Next year: Faculty at UT Arlington and TRIPODS-ADE partner for research and HSI DSC

Raymundo Navarrete

- Involved in two RWGs (RWG1: Large Scale Networks and RWG6: Large Scale Data Analysis)
- Applications to traffic and astronomy

metropia
Driving a better city

LSST
Large Synoptic Survey Telescope

The University of Arizona
Center for Gamma-Ray Imaging

The University of Arizona

Workforce Development: PhD Students

THE UNIVERSITY OF ARIZONA

Women in Data Science (WiDS)

Support women at all levels

- 33% of TRIPODS graduates were female students
- First WiDS-Tucson in 2019, and 2020
- High-school girls outreach

Women Leadership Role

- Caucus for Women in Statistics (CWS)
- Association for Women in Mathematics (AWM)
- Cox Scholarship Committee (for female graduates)

THE UNIVERSITY OF ARIZONA

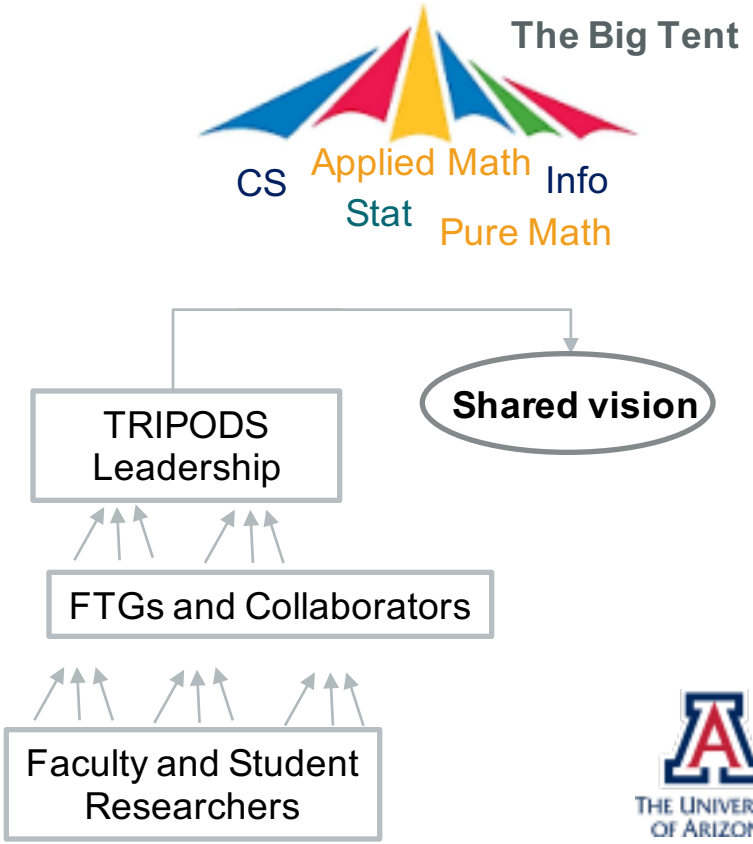
Challenges & Thoughts

- How to break culture barriers
- Where to publish the research results
- How to gain institutional support
- How to build coherent leadership
- How to make this sustainable and successful

THE UNIVERSITY OF ARIZONA

TRIPODS Leadership and Structure

- Bottom up, democratic design
- Teams come together as teams of teams
- Leadership organizes dissemination through talks, special issues, special sessions, white papers, and emissaries



THE UNIVERSITY OF ARIZONA