UA-TRIPODS

Helen Zhang
Statistics

Stephen Kobourov
CS

Joe Watkins
Math/Stat

David Glickenstein
Mathematics
What is TRIPODS?

Transdisciplinary Research in Principles of Data Sciences

- **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)
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:

**Activities**
- 71 seminars in 2.5 years with attendance averaging over 30, many with 60+ attendance
- Three workshops at Biosphere2 and C3MAT

**UA-TRIPODS**
- 94 Papers
- 33 Proposals
- 17 Posters

**Education**
- 4 Postdocs
- 14 PhD Students (CS, Math, Applied Math, Statistics, Systems/Industrial Engineering)
- New Undergraduate Major in Statistics and Data Science
  - 130 majors and 70 minors as of November 2019

**Domain Science**
- Kaggle submission between RWG and Astronomy
- Metropia (smart city project)
- Hurricane prediction

**Outreach**
- Women in Data Science
- Sonia Kovalevsky Day
- 2018 UA Science Lectures: “Human, Data, and Machines”
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
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
Key Participants:

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

First CS-Math-Stats Publication in UA History
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
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...
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
Workforce Development: PhD Students
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)
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
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