

# Detecting and Addressing **BIAS** in Data, Humans, and Institutions

## About

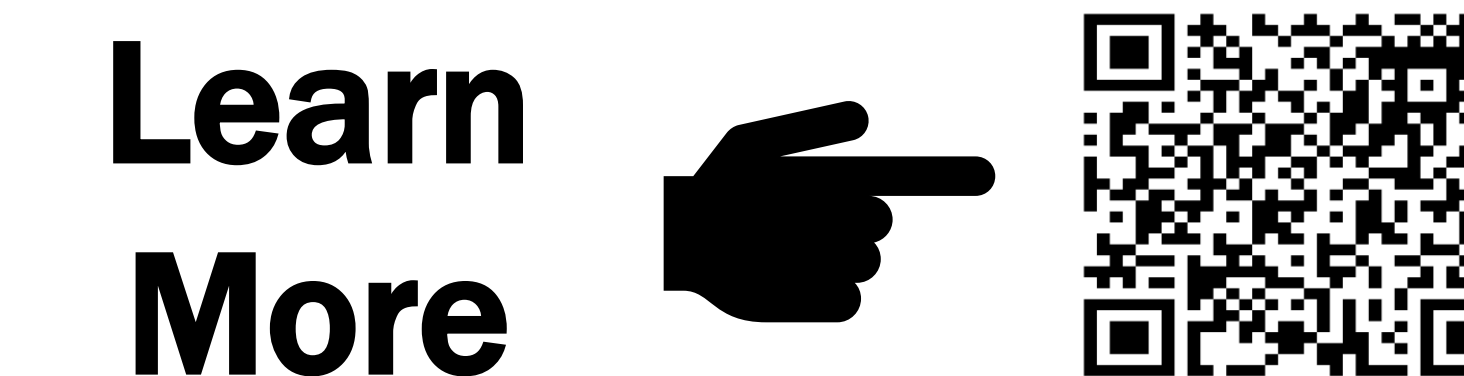
Data science and AI are powerful tools for generating new knowledge, fueling innovation, and dealing with society's most pressing problems. However, "big data" and machine learning tools can perpetuate biases that advantage some people, and disadvantage others. This training project (NSF 2125295) bridges perspectives from the human-centered sciences with those from the data sciences in support of convergent research projects.



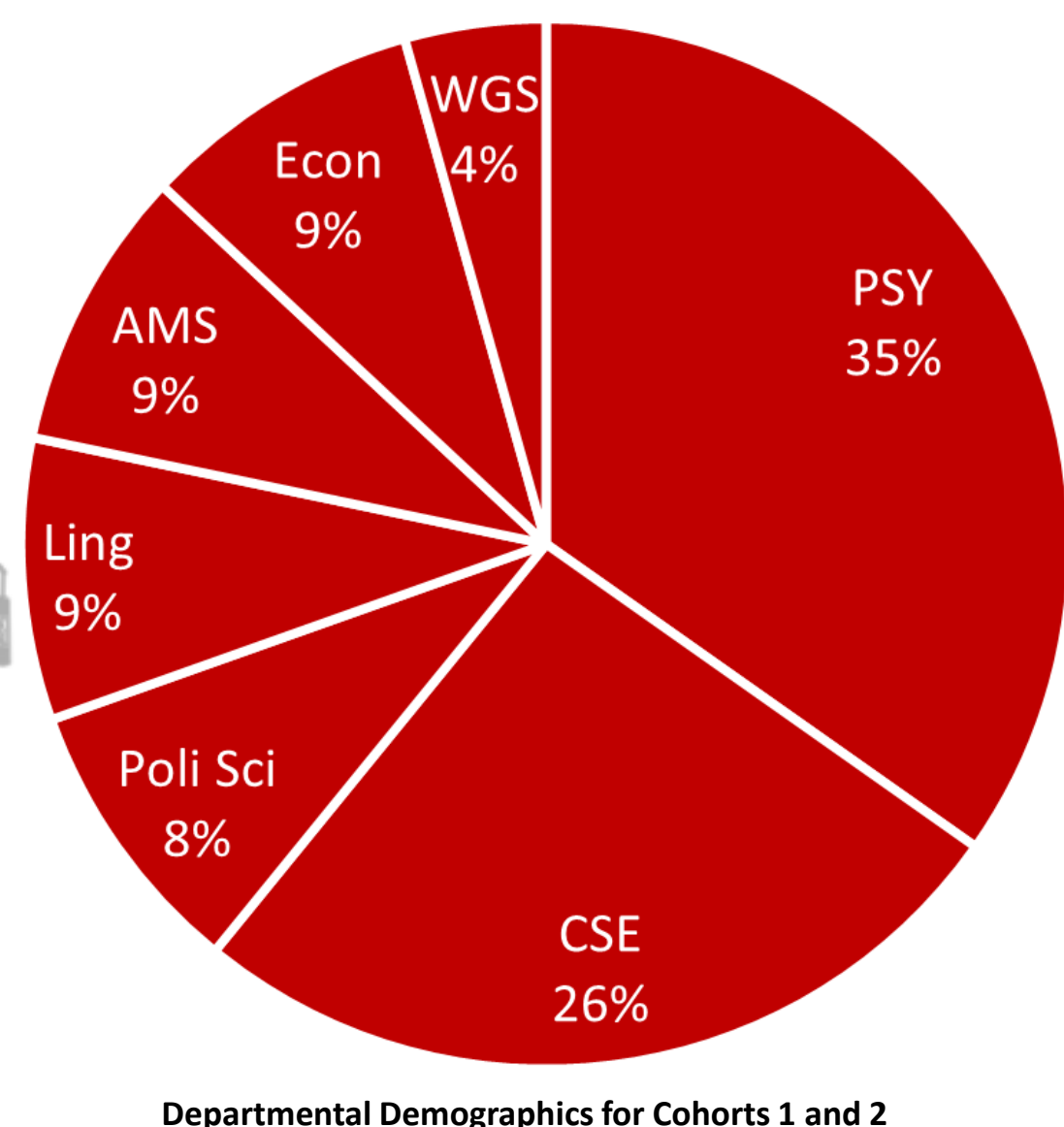
**Leadership Team:** Susan E. Brennan (PI), C.R. Ramakrishnan, Wei Zhu, Bonita London, Jeffrey Heinz  
**Project Coordinator:** Kristen Kalb-DellaRatta  
**Project Evaluation:** Catherine Good, Elevate Learning, LLC



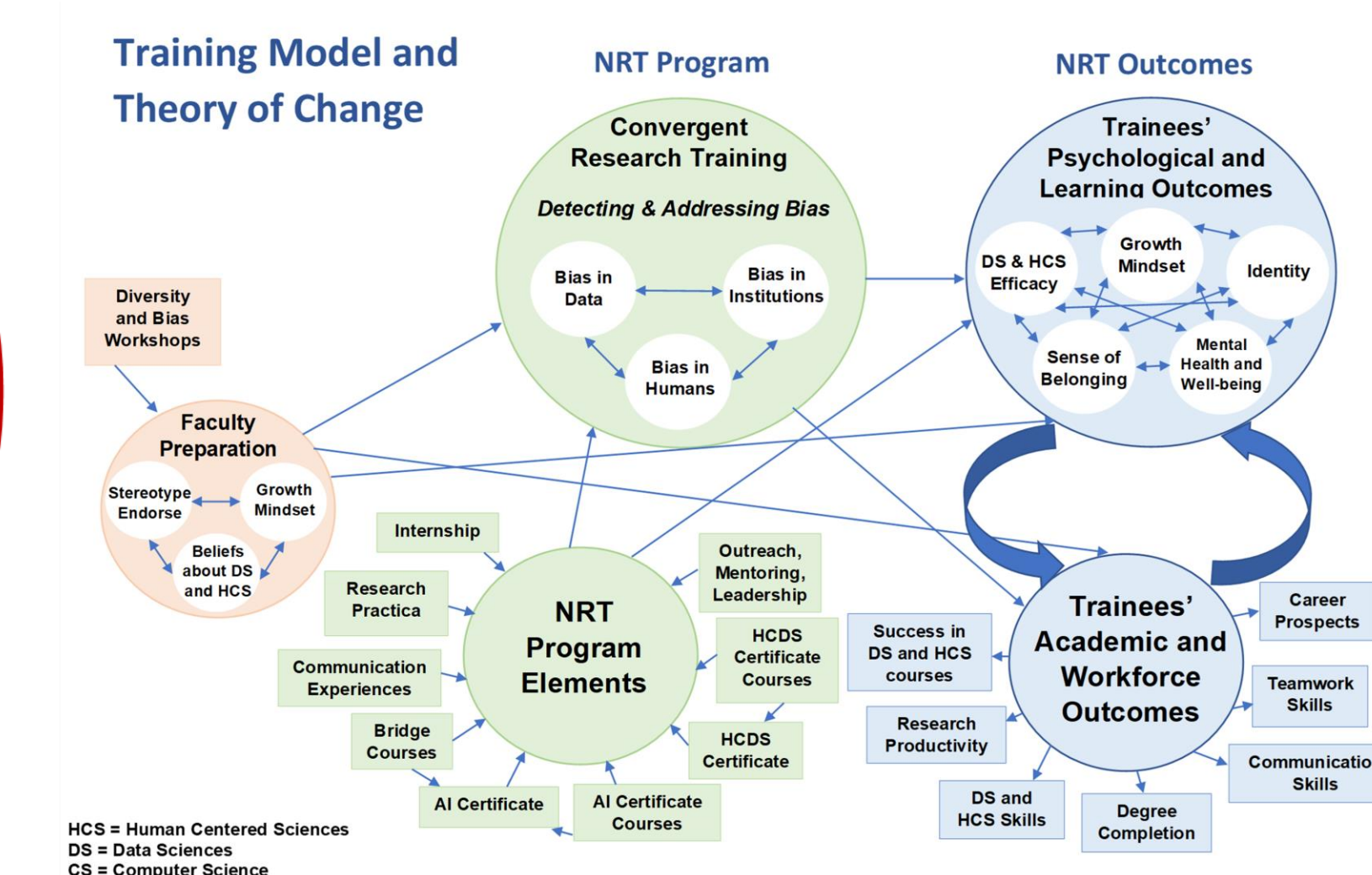
**Mission:** to seed a generation of researchers trained to identify and mitigate biases that arise when data-centric methods are applied to real-world problems



### Features of the Traineeship Model



### Y2 Evaluation



Bias NRT Program-Related Measures	Mean	SD	n
Support from Bias-NRT faculty	5.00	1.61	11
Support from Bias-NRT staff	5.09	1.58	11
Support from Bias-NRT peers	5.09	1.78	11
Support from Bias-NRT friends	4.91	1.58	11
Satisfaction: Mentorship	5.31	1.37	4
Satisfaction: Leadership skill development	4.80	.89	8
Satisfaction: Communication skill development	5.46	.69	9
Satisfaction: Research Practicum	5.46	.67	10
Satisfaction: Bridge Course	2.51	.80	3
Satisfaction: AI Certificate	4.41	1.17	7
Satisfaction: Human Centered and Data Science Certificate	5.60	.35	5

Science Environment Measures	Mean	SD	n
perceptions of other people's fixed mindset beliefs in data science	5.03	1.45	8
perceptions of other people's fixed mindset beliefs in human centered science	3.33	1.99	10
perceptions of other people's gender stereotype endorsement in data science	3.47	1.50	7
perceptions of other people's gender stereotype endorsement in human centered science	1.67	.55	9
perceptions of faculty growth mindset beliefs in data science	3.83	.92	8
perceptions of faculty growth mindset beliefs in human centered science	4.62	.93	9
meta-beliefs about peers' confidence in data science	4.28	1.38	11
meta-beliefs about peers' data science confidence in human centered science	5.06	.91	11
perceptions of support within the community (data science)	4.03	1.48	8
perceptions of support within the community (human-centered science)	5.48	.68	10

- Recommendations Moving Forward**
1. Infuse programmatic elements with content that research has shown to be effective interventions for constructs like growth mindset, identity, efficacy, and well-being.
  2. Provide on-going support for mentors.
  3. Work with the research practicum presenters to intentionally include topics that could help foster the development of healthy psycho-social and well-being outcomes.
  4. Implement additional program-related activities that could foster the development of healthy psycho-social and well-being outcomes.

### Y2 Achievements

- Fall 2022 Topic Based Research Practicum; 20 meetings held with over 18 topics discussed
- Internships secured for 3 trainees by way of faculty mentorship and affiliation with the NRT
- Travel opportunities provided for trainees, faculty, and staff, establishing a vast network of connections
- Applications for two Advanced Graduate Certificates, *Artificial Intelligence* and *Human-Centered Data Science*, submitted to SUNY and NYSED for approval
- 13 new trainees recruited as Cohort 2
- 23 total trainees going into Y3, 8 NRT Funded and 15 Non-NRT Funded

### Cohort 1 Research Projects

- Post-Conviction Project-with data from the National Registry of Exonerations and advisors from the Innocence Network.
- Bias in Large Language Models
- Bias in Facial Emotion Recognition (clinical applications)

### Post-Conviction Project

Kalina Kostyszyn, Carl J. Wiedemann, Rosa Bermejo, Amie Paige, Kristen W. Kalb-DellaRatta, Nancy Franklin, & Susan E. Brennan

Using Decision Trees to predict latent class membership of 3223 cases in the National Registry of Exonerations (NRE) database

Decomposing hidden patterns in the data to develop a **transparent** and **user-friendly** framework, supporting decisions by intake staff at Innocence organizations

Below, a Decision Tree trained on exoneration data to predict trends associated with latent class membership. This branch organizes cases marked as 'murders.' Depending on the features associated with each case, the case is labeled as one of the latent classes.

Using machine learning to support Innocence Organizations' intake decisions

The Post-Conviction Project will present their paper, *A Computational Decision-Tree Approach to Inform Post-Conviction Intake Decisions*, at the Innocence Project's 2023 Just Data: Advancing the Innocence Movement conference and will be subsequently published in *The Wrongful Conviction Law Review*

