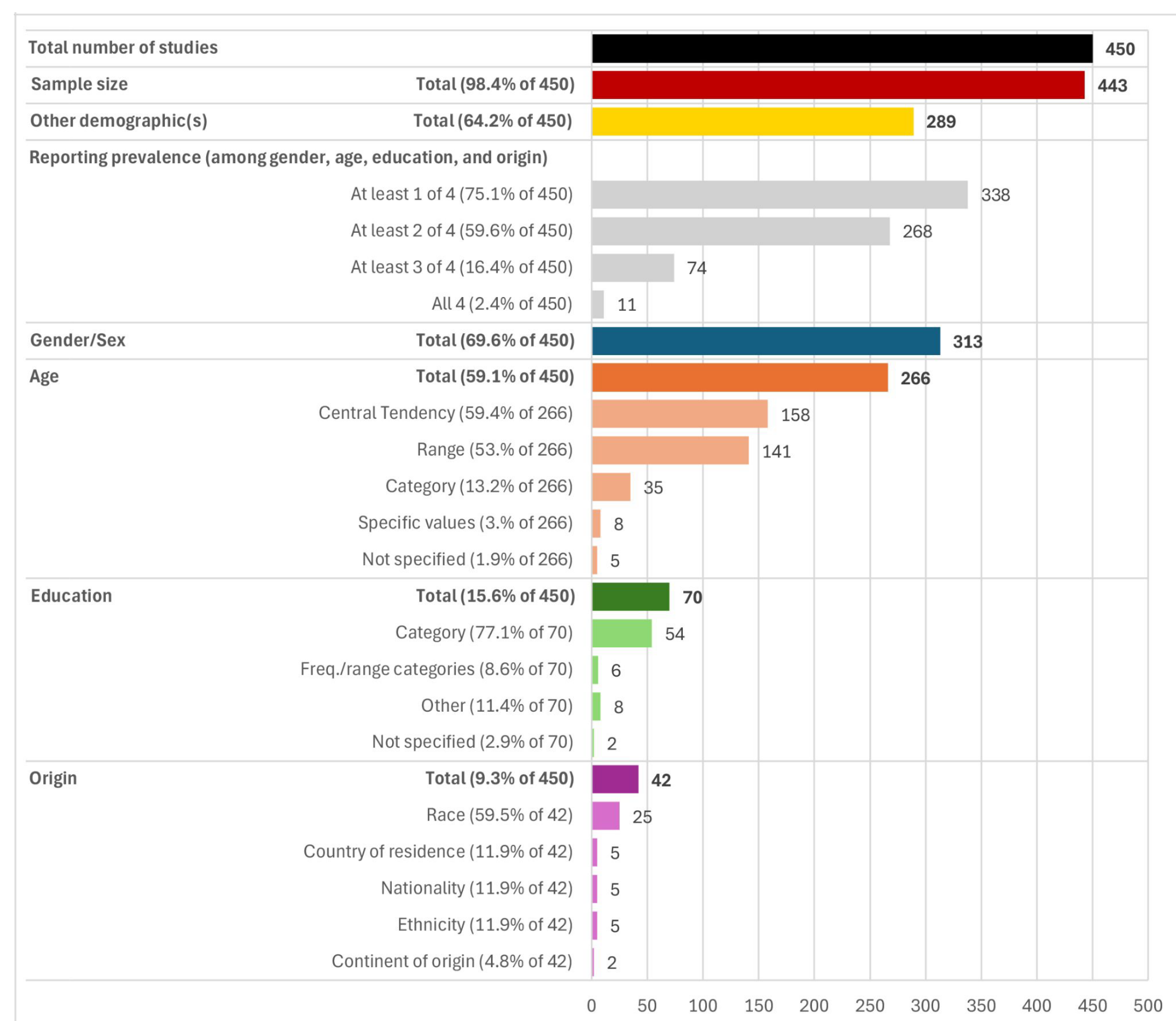


# Maximizing Data Utility and Participant Privacy through Usable, Secure Data Workflows for Human-Centered AI Research

PIs: Kelly Caine, Bart Knijnenburg, Susan McGregor

With help from: Emily Sidnam-Mauch (Research Scientist), Hansen Lee and Laila Shafiee (PhD Students), and Sarah Mendoza and Khadija Wane (REUs)

## Our project provides support for:



Demographic data collection

1. Which method best describes your overall study?

Survey  
Experiment  
Interview *Not yet available*

2. Which statistical test do you need a sample size for?

T-test  
ANOVA *Not yet available*  
Other *Not yet available*

2. A. What type of t-test will you use?

Independent samples T-test  
Dependent samples T-test

2. B. Is the size of the participant groups you are comparing equal?

Yes, each group will have the same number of participants [Recommended]  
No, there will be a different number of participants in each group

2. C. Will you use a one-tailed or a two tailed test?

Two-tailed test  
One-tailed test

## Get a Sample Size Recommendation

Built by researchers specializing in study design, privacy, and usability, this tool will help you decide what sample size is right for your study.

Get started

Method: survey -> Test type: T-test -> Balance: balance=yes -> Independence: independent -> Tails: two-tail

Alpha: 0.05  
Effect size: 0.5  
Power: 0.8

Participant count: 128

Calculate  
Export Log

Effect size: 0.5 Alpha: 0.05 Power: 0.8 Number of participants: 128 Timestamp: Thu, 29 Jan 2026 02:33:02 GMT

Sample size calculation

**Our Challenge:** Designing human subjects studies that preserve research participants' privacy and security while still generating robust results is tricky. AI researchers need support to do this effectively and efficiently.

**Approach:** Apply usable security techniques and iteratively test the prototype with human-centered AI researchers.

**Solution:** Develop a prototype to aid privacy- and security-preserving research design. The system helps AI researchers:

- Generate privacy-maximizing **demographic** questions
- Determine a study's optimal **sample size** and composition
- Follow recommended guidelines for **ethical data-sharing**
- **Justify privacy implications** of the study with targeted explanations

**Our Scientific Impact:** help AI researchers secure data collected about people while at the same time improving their ability to make new scientific discoveries. This will **enhance data integrity** and sample **representativeness**.

**Our Broader Impact:** our system improves the quality, security, and efficiency of human-centered AI research by reducing the amount of data collected about people and helping to create AI systems that are usable, fair, accurate, and trustworthy.

## Upcoming publications/presentations:

**Who are the Humans in HCAI?** Systematic Review of Demographic Reporting Practices in AI Research with Recommendations for Transparency and Accountability  
*ACM Conference on Fairness, Accountability, and Transparency (FAccT '26), June 25-28, Montreal.*

**SampleSavvy:** A Web Tool for Guiding Research Sample Size Decisions  
*HFES International Annual Meeting (ASPIRE), October 19-23, Reno, NV.*



PI Kelly Caine - caine@clemson.edu  
Co-PI Bart Knijnenburg - bartk@clemson.edu

