

Gregory Serapio-García

**Graduation Year:** Junior

**College:** Arts & Letters

**Major(s):** Psychology

**Minors(s):** Computing and Digital Technologies, Glynn Family Honors Program

**Scholar Group Membership:** Glynn

**Did you received other funding for this project?:** Glynn

**Could you have completed this project without CUSE funding? No**

**More details on CUSE funding assistance?**

**Project Title:** Big Data & the Big Five: Computationally Predicting Personality and Psychopathology from Social Media

**Project Location:** Center for the Advanced Measurement of Personality and Psychopathology; Center for Research Computing; University of Notre Dame

**ND Faculty Mentor:** David Watson

**Project Type:** Research

**Why did you undertake this project/experience?** Deepen your knowledge of a topic or issue, Research/experience necessary for senior thesis or capstone project, Prepare for graduate school (MA or PhD), Prepare for national fellowships, Career discernment and/or preparation

**Did your funded experience help you:**

**[Deepen your understanding of your coursework or field of study]:** Very Much

**[Discern your interests and post-bac goals]:** Very Much

**[Become confident in your ability to set and achieve your goals]:** Yes

**[Gain a more nuanced view of local, national, or global communities]:** Yes

**[Improve your written and verbal communications skills]:** Very Much

**Tell us about your experience.**

This summer, I completed a fellowship with the National Science Foundation and Notre Dame's Center for Research Computing (CRC) which aimed to apply high performance computing and data analytics techniques to the social sciences. My project, in particular, aimed to predict scores for the 'Big Five' traits of personality from unstructured social media data (e.g., textual data). I actively developed a website backend with Python and Flask, designed relational database schema for storing my data, and conducted item concordance analyses between existing Facebook personality data and offline psychopathology data. Future work and analyses are to continue into the Fall of 2017.

**Describe the impact this project had, both on you as a student-scholar and on the people you worked with.**

My project was an academic formation experience, especially with regards to conducting research that stretches across several disparate disciplines. Psychology is often criticized for

being a 'soft science' based off of empirically unsound principles. Not only was my research steeped in psychological theory (supported by decades of data!), but it also incorporated many 'hard science' principles that are used in computer science and high performance computing. Thus, I ended up working with computer engineers, clinicians, sociologists, and data scientists alike to make this project come together. In all, it was empowering to bring a spectrum of disciplines together.

**Describe how this experience is connected to your plans as a student or future professional.**

This experience helped me envision my future life in graduate school. I wasn't given a set schedule or concrete goals by my mentors and advisers; all of my work on a day-to-day basis was self-directed.

Practically, my project provides a sound basis for a senior honors thesis in psychology and the Glynn Family Honors Program. At the Undergraduate Student Summer Research Symposium, I was able to present a poster and practice my conference presentation skills.

**What advice would you give other students who are planning to pursue similar projects?**

Meticulous planning is key, but unexpected deviations are bound to happen in any long-term project. Use these deviations to your advantage! For me, hiccups in my data storage and analysis plans led me to new methods that will greatly improve the outcomes of my project.

I acknowledge that this form has been filled out truthfully and to the best of my ability. I understand that this information will be shared with many different CUSE constituencies. As such, I have provided as much useful information as I was able. I understand that CUSE will not complete my award disbursement until this form is successfully completed. If I have any questions or concerns, I will contact CUSE before submitting this form. To illustrate that you understand all of these points, please enter your Notre Dame email in the box below.  
gserapio@nd.edu

