






Faculty Member Contact Information

Name	Dr. John Matta
Contact Info	
SIUE Email	jmatta@siue.edu
Campus Box	1656
Department	Computer Science

1 Funded, Unfunded URCA Assistant

X	This position is ONLY open to students who have declared a major in this discipline.	M
X	This project deals with social justice issues.	
	This project deals with sustainability (green) issues.	
X	This project deals with human health and wellness issues.	
	This project deals with community outreach.	
	This mentor's project is interdisciplinary in nature.	

Are you willing to work with students from outside of your discipline? If yes, which other disciplines?

- No

How many hours per week will your student(s) be required to work in this position?

(Minimum is 6 hours per week; typical is 9)

- 9 hours

Will it be possible for your student(s) to earn course credit?

- Yes—CS 495 (3 credit hours)

Location of research/creative activities:

- SIUE Edwardsville Campus

Brief description of the nature of the research/creative activity?

The research activity involves using the tools of computing to do genetic analysis. NIH has opened a new dataset called "All of Us" with over 100,000 DNA samples and many hundreds of associated phenotypes. This data will revolutionize genetic research. Students will apply computational methods to this data to develop new methods for doing a Phenome-Wide Association Study (PheWAS).

Brief description of student responsibilities?

Students will be responsible for cleaning and assembling the "All of Us" data, and for using tools provided with the dataset to perform PheWAS studies. We will also apply algorithms that worked successfully in a previous paper (written by the URCA's of a few years ago!) that attempted to find genes implicated in Autism Spectrum Disorder (ASD). The previous successful project was to find links between genes involved in ASD and Cleft Palate birth issues and Gender Dysphoria. We will attempt to use the same methods to find links between ASD and several other conditions. These activities will involve skills such as Python Programming, R programming, database creation and querying. We will also be writing the results into papers, so that the students will have experience doing technical writing, researching related work, etc.

URCA Assistant positions are designed to provide students with *research or creative activities* experience. As such, there should be measurable, appropriate outcome goals. What exactly should your student(s) have learned by the end of this experience?

Students will learn rules and procedures for doing research involving human medical data.

Students will learn and use Python and R libraries, and will write programs using those languages.

Students will learn about PheWAS studies and computational genetics work.

Students will learn technical research and writing.

Requirements of Students

If the position(s) require students to be available at certain times each week (as opposed to them being able to set their own hours) please indicate all required days and times:

- Students will be expected to attend one research meeting per week. A time will be chosen that accommodates everyone's schedules. Generally students may set their own hours.

If the location of the research/creative activities involves off campus work, must students provide their own transportation?

- No transportation requirements.

Must students have taken any prerequisite classes? Please list classes and preferred grades:

- I generally prefer students who have taken CS325 (Software Engineering) and CS340 (Algorithms and Data Structures) and done well in both.

Other requirements or notes to applicants:

- None