

# C.A.R.E (Cohort Assessment & Retrieval Environment)

Project Team: Kyle Ellis, Payal Shah, Jordan Tang

Computer Science

Project Number CS 312

Faculty Advisor: Bridget McInnes PhD

Sponsor: VCU Computer Science

Mentor: Bridget McInnes PhD, Antonio Abbate MD, PhD

**Abstract:** The purpose of clinical trials is to explore whether a medical treatment is safe and effective for humans or to enhance preexisting methods. The identification of patients who satisfy a set of predefined criteria for the trial is instrumental. However, the process of distinguishing these patients on the basis of their clinical records is a challenging task since it can have structured (ex: precise measurements) and unstructured data (ex: physician notes). One difficulty with this is data normalization; there are many ways to describe a single concept. For example, “heart attack” and “myocardial infarction” both refer to the death of the heart muscle. The goal of this project is to develop a system that will process clinical records for the purpose of cohort discovery and make a visual framework to allow researchers to view and explore the associations between biomedical terms and their characteristics.

**Steps:** 1) The user inputs criteria on what they want/do not want in their patients’ medical records. 2) The criteria and patient records run through a system using MetaMap and MetaMap- DataStructures that measures the association between biomedical terms and links other terms to that word or phrase. 3) The patient records are ranked based on the user input criteria. Therefore, the records that have more prevalent information regarding the input criteria receive a higher score. 4) The user is able to see the records in the order they were ranked. This makes the process of finding patients for a clinical trial more manageable.

**Keywords:** Normalization, MetaMap, MetaMap-DataStructures, Concept-Unique-Identifier (CUI)

