**Research Proposal**

***Caifan Du*** | *PA 397C Data Management and Research Life Cycle*

It is widely believed that making research work open access (OA) can lead to broader impact. Hence, many higher education institutions develop their open access mandates and policies to encourage or require their researchers to make work open access. Some of them also host open access repositories to support open access deposits of scholarly work. For my final project, I propose a study to examine the effect of these institutional efforts promoting open access. My research question, research design, data sources and validation methods will be introduced in the following part.

**Research Question**: Do institutional efforts in promoting open access have an effect on institutional research impact?

The institutional efforts are operationalized as the age of institutional-hosted OA repositories and the number of years since the time when institutional open access policies became effective to date. These are defined as independent variables of the study. The dependent variables are the institutional research impact, operationalized as the citation impact of institutions.

**Research Design**: The study is designed in a cross-sectional manner. The effect of institutional hosted OA repositories and institutional OA policies will be examined and compared. The overall effect of the two types of OA promotion efforts on research impact will be explored as well.

**Data Sources**: Four data sources are used for answering my research question. The first two data sources are the Directory of Open Access Repositories (DOAR) and the Registry of Open Access Mandates and Policies (ROARMAP). These two sources are online databases that solicit self-reported data from organizations that host OA repositories or hold OA policies. The year that the OA repository was created will be retrieved from the first source DOAR for computing the age of OA repositories to date. The date when the OA policy became effective is acquired from the second source ROARMAP for calculating the number of years to date since the effective date of institutional OA policies.

The data for measuring research impact is gained from the citation indicator from Ranking Web of Universities, a publicly available world university ranking developed by a Webometrics research lab supported by Spanish National Research Council. For validating this data source, I also use the citation indicator from the university ranking developed by the Center for World University Rankings (CWUR).

**Validation Methods**: I seek to validate my study using similar measures developed by others to check the quality of the data being used. In other words, other measures of citation impact from another data source about university ranking will be used to run the same data analysis. So that I will have some clues about the quality of the datasets being used. However, this validation is limited since it only attends to one data source. For other data sources, I have not be able to find complementary data to validate their quality.