Phil Castillo

LIS 545 B

Term Project - Repository Profile

10 February 2023

“Collaborative Psychiatric Epidemiology Surveys 2001-2003”

The “Collaborative Psychiatric Epidemiology Surveys 2001-2003” (CPES) repository permits access to information surrounding mental disorders and their relevant treatments. It is centered primarily on adult populations in the United States and intended to be utilized for further research relating to identity-based influences on mental health. This is similar to the dataset “Suicidality on Reddit” which was also created in order to collect data surrounding harmful implications and experiences in the deterioration of mental health with the intention to help create preventative measures. I determined this repository to be a good fit for this dataset not only because of their similarity in focus but also because it was broad enough to house a dataset that did not collect data of a specific group. “Suicidality on Reddit” did not contain data of groups of any particular race, gender, age, etc. Since this repository also did not specify any focus groups with mental disorders but rather centralized mental disorder-related data as a whole, I found that this was synonymous to the dataset.

The organization, Inter-university Consortium for Political and Social Research (ICPSR), to which the repository belongs, permits and encourages deposits of digital data submissions. Members and non-members of institutions are permitted use of ICPSR repositories and anyone is able to contribute data. ICPSR emphasizes that its repository services are open to be utilized by researchers, staff, students, citizen scientists, and the like.

ICPSR appears to encourage any sort of data to be submitted to the repository but requires depositors to include all necessary data and documentation to independently read and interpret the collection, including a study-level description. The only “restrictions'' are in regard to how certain data should be submitted. For example, for quantitative data, ICPSR advises those depositing data files to do so as SAS, SPSS, or Stata files with the expectation that “each variable in the data collection should have a set of exhaustive, mutually-exclusive codes” and that variable and value labels “clearly describe the information or question recorded in that variable” (About ICPSR, 2020). It is noted that datasets in other formats are accepted as well but they do not specify what these other formats are other than ASCII files if the data is accompanied with setup files. Additionally, depositors have the option to deposit data for public-use or have restricted-access.

ICPSR provides the deposit process which can be accessed through a video on the site as well as an informational PDF file describing deposit services and processes in full detail. Additionally, it is noted that depositing data submissions are free of charge with promised benefits of ensuring data meet FAIR principles, help with managing confidential human subject data, assistance with obtaining DOIs and citations, and reviews and improvement of data. In addition to the video and the informational PDF file, ICPSR provides user support through an email if depositors require any further assistance or consultation. This appears to provide opportunity for more specific inquiries regarding particular kinds of data that may not be addressed from the aforementioned resources due to its specificity.

The repository details indicate that the metadata standards included are Data Documentation Initiative (DDI) and Dublin Core with a DCC scheme. However, nowhere in the description of the deposition process does it indicate a requirement for data submissions to be deposited with these metadata standards. The protocol is unclear with the possibility of depositors submitting data not in this structure or according to DDI or Dublin Core specifically.

It is noted that the repository’s data access is available for public downloading through the ICPSR website, but only if the data is specified for public-use. Data in sponsored archives are free and accessible to everyone, but data in the membership archive are available for free for members and $550 per study for others. Membership archive data is able to be freely accessible if the curation’s cost is paid by the depositor or another funding entity. As mentioned, depositors have the option to have their data be restricted-access. If this is the case, then the data can only be accessed by “those who have executed Restricted Data Use Agreement with ICPSR and meet all the criteria by ICPSR” (About ICPSR, 2020).

The repository is accessible using a direct file download. There is an option to download only the documentation but it can also be downloaded as a SAS, SPSS, Stata, ASCII, or delimited file. It is noted specifically for this repository that its public-use data files are accessible to the general public and without a required affiliation with ICPSR. However, it also mentions that the repository contains one or more files that are specially restricted and therefore cannot be directly downloaded.

The repository’s DIP includes a general description of the project with its relevant citations, subject terms, and geographic coverage. Also included are the scope of the project which includes some data collection notes as well as the time period of the data collection (2001-2003) and its methodology.

References

Alegria, Margarita, et al. (2020) Collaborative Psychiatric Epidemiology Surveys 2001 - 2003. *re3data*, https://www.re3data.org/repository/r3d100010825.

Alegria, Margarita, et al. (2020) Collaborative Psychiatric Epidemiology Surveys (CPES), 2001-2003, *Inter-University Consortium for Political and Social Research*, https://www.icpsr.umich.edu/web/ICPSR/studies/20240.

About ICPSR. (2020) *ICPSR*, https://www.icpsr.umich.edu/web/pages/about/.