ASU IRB Guidance for Human Subjects Research Involving Secondary Data

What is secondary research?

- Secondary research refers to research using archival data or data that has already been collected by someone else. This type of research could also involve records, information or biospecimens that were collected for non-research purposes, such as materials that are left over from routine clinical diagnosis or treatments. [Refer to 45 CFR 46.116(d) of the revised Common Rule.]

When is secondary research considered human subjects research?

- The answer to this depends on the types of data that are being requested, how those data were obtained, how it is being accessed, and how it is being used. Secondary data must already be in existence at the time of submission. **To protect researchers, the ASU IRB reviews all protocols involving the use of secondary data at individual level that include the intent to generalize knowledge from living individuals.**
    - Individual vs. aggregate (group) level data:
        - If your analysis involves **individual student educational records or individual patient records**, then you are conducting secondary data analysis of records on an individual level. IRB review is required.
        - If your analysis involves data only in **aggregate form**, such as unemployment rates or infant mortality rates, then IRB review is not required. This means not at the individual level.

Types of data	Considerations
Publicly available data: Information that can be freely used, reused and redistributed by anyone with no restrictions on access or usage.	Research involving secondary human subjects data that are publicly available are eligible for the IRB exempt wizard. This type of data is generally available on a website with no restriction on who can access the data. (No data provider application process with restrictions on data use.)
De-identified data: All identifiable information has been removed prior to ASU receiving it. Data is de-identified when any direct or indirect identifiers linking the data to the subject's identity have been destroyed, or if received from another researcher, the link to re-identify is not disclosed to the receiving researcher.	How were the data originally obtained? Who owns the data and permission(s) needed to access de-identified data? When data is de-identified but not publicly available, the IRB application should explain the data provider's requirements to access the data.
Anonymous data: Anonymous data is data where at the time of the data collection neither the researcher nor a third party can connect the data to the research participant.	Is the data actually anonymous? Are there any direct identifiers such as IP address or a worker ID #? In online studies, often a third party such as Prolific or Qualtrics will have assigned a participant a worker ID # which is linked to the participant's identity. The participant's identity is not known by the researcher however it is known to the third party and should not be described as anonymous.

Restricted use data: Contains sensitive information or information that enables the potential identification of respondents through inference. Data may also be restricted due to confidentially promises or for proprietary reasons.	A Data Use Agreement (DUA) is generally required whenever human subject data is being sent or received between ASU and an Outside Party and there is a restriction on the data recipient. See Agreements with ASU for information about Contracting Services. See Nondisclosure agreements, data use agreements and material transfer agreements for explanation on when DUAs are needed. Please note these are reviewed by industryagreements@asu.edu.
Limited Data Set: This is a dataset where direct patient identifiers are removed. As defined by the Health Insurance Portability and Accountability Act, better known as "HIPAA," a data set of patient information may disclosed without patient's authorization if certain conditions are met. Limited data sets may include only the following indirect identifiers: Dates (such as admission, discharge, service, and date of birth (DOB)), geographical information (city, state, and zip code), age, and any unique code or identifier that is not listed as a direct identifier.	Similar to a Restricted Use Data Set, a Limited Data Set requires a Data Use Agreement which places restrictions on the recipient of the data. See HIPAA Research Guidance Flowchart
Coded private data: Existing human subjects data previously collected for a separate purpose whereby an external partner maintains a key linking identifying information to the private information or specimens, but the current investigator(s) cannot readily ascertain the identity of the individual(s) to whom the coded private information or specimens pertain.	Coded private data is not considered human subject research. See Coded Private Information or Biospecimens Used in Research, Guidance (2018) from the U.S. Department of Health and Human Services (HHS). We recommend reaching out to ASU IRB to receive confirmation that no IRB review is required.

Researchers who are unsure whether their project requires IRB review should contact ORIA at asuirb@asu.edu for consultation.