

# DEFINING DATA

## *in Human Subjects Research*



### ■ **OVERVIEW:**

In Human Subjects Research, the protection of private, personal data is both an ethical obligation and a regulatory requirement.

It is essential for researchers to understand the distinctions between the different types of data to ensure compliance with IRB regulations, data protection laws, and funding agencies.

## ■ The 4 Types of Data ■

Data gathered or used in human subjects research studies can be regarded as:

**Confidential**

**Anonymous**

**Coded**

**De-identified**

The following guidance document defines each of these 4 terms as they relate to an individual's participation in HSR, and the way that their data are collected and maintained for analysis.



















*See pages 4-6 for information about how to code and de-identify HSR data.*

# 1. CONFIDENTIAL DATA



Confidential data is Personally Identifiable Information that is provided by research participants to investigators under an expectation of privacy. Confidential data can be used to identify an individual—either directly or indirectly.

**Personally Identifiable Information (PII) includes any and all of the 18 identifiers listed below:**

 Names	 Social Security Numbers	 Vehicle Identifiers
 Geographic Elements	 Medical Record Numbers	 Website URLs
 Dates	 Health Plan Beneficiary Numbers	 IP Addresses
 Phone Numbers	 Account Numbers	 Biometric Elements
 Fax Numbers	 Certificate/License Numbers	 Facial images
 Email Addresses	 Device Identifiers & Serial Numbers	 Any other unique identifiers

*SDSU investigators must inform the IRB if they will access and/or record any of the 18 PII identifiers by listing them on the IRB Initial Application in the “Identifiers” table.*

**7.0 Identifiers**

7.1 Select any **PII** or **PHI** identifiers you will be **accessing** (i.e. using) or **recording** (i.e. collecting) among the following:

PII or PHI Identifier	Accessing & Recording
<input type="checkbox"/> Patient/Participant Names	<input checked="" type="checkbox"/> Accessing <input type="checkbox"/> Recording

*“PHI” is Personal/Protected Health Information, and relates to the health of an individual. PHI is a specific type of PII.*

## 2. ANONYMOUS DATA

Data is anonymous when it has not at ANY point been connected to an individual through direct identifiers (i.e., name, email address), OR indirect identifiers (age, race) that can be triangulated to ascertain the identity of an individual.

*If you directly access or record any identifiers through your study methods, your study is **NOT** anonymous.*

### EXAMPLE:

Peter designed a survey for college seniors on Qualtrics. The survey is open to individuals who identify as 18 years of age or older, are college seniors, and live in the United States. Participants' Qualtrics accounts will not be connected to their survey responses, and Peter has enabled "Anonymize Responses", so IP Addresses and location data will not be collected.



■ His survey is multiple choice, and includes Likert-style scales like this:

A.

When I am with other people, I feel included:

- 1 = Disagree
- 2 = Disagree a little
- 3 = Neither agree or disagree
- 4 = Agree a little
- 5 = Agree

...and demographic questions like this:

B.

What is your age?

- 18 - 25 years
- 26 - 35 years
- 36 - 45 years
- 46 - 55 years
- 56 - 65 years
- 65 + years

Question A does not collect any information that could directly identify an individual.

Question B is a basic demographic question, and even if Peter asks about age, gender, or other demographics, the information combined would not allow for an individual to be identified, because he cannot determine where the responses came from, or who they are associated with, as the participants did not sign up for the survey using their email address, phone number, name, or any other identifiable information.