

Anonymous, Confidential, or De-identified?

The IRB often finds that the terms *anonymous*, *confidential*, and *de-identified* are not used correctly. These terms are described below as they relate to an individual's participation in the research and the way that their data are collected and maintained for analysis.

(a) **Anonymity addresses both subject participation (knowing any individual was involved in the research project) and data (whether data gathered through the project can be connected in any way to an individual participant).** **Anonymous** indicates it is impossible to determine whether any individual participated in the research project. Data are *anonymous* if no one, **not even the researcher**, can connect the data to the individual who provided it. No identifying information is collected from the individual, either as part of the study methodology or documentation such as informed consent forms or other releases. A study should not collect identifying information of research participants unless it is essential to the study protocol. Anonymity cannot be guaranteed if any personally identifiable (PII) information will be collected. Researchers should be aware however, that collection of information regarding other unique individual characteristics (indirect identifiers) could make it possible to identify an individual from a pool of participants.

(b) **Confidential** indicates that the research team knows that any particular individual has participated in the research but is obligated not to disclose that information to others outside the team. When data are *confidential*, there continues to be a link between the data and the individual who provided it. The research team is obligated to protect the data from disclosure outside the research according to the terms of the research protocol and the informed consent document. Note that coded data are not *anonymous*.

(c) **De-identified** data are those data that have NO direct identifiers or codes linking it to an individual subject. For data to be deemed '*de-identified*', all direct or indirect identifiers or codes linking the data to the individual subject's identity are destroyed.