
UWM Institutional Review Board Guidance on De-identified Data

A de-identified data set would not contain any of the following:

- Names
- Street address, city, zip code
- Dates- birth or death dates, appointment, hospitalization, claims, etc...
- Telephone numbers
- Fax numbers
- Email addresses
- ID numbers such as Student ID number, Social Security number, Medical record number, Health plan numbers, Account numbers, Certificate/license number
- Vehicle identifiers including license plate numbers
- Device identifiers/serial numbers
- URL
- IP (Internet protocol) addresses
- Full face photographs and comparable images
- Biometric identifiers including voices and finger prints
- Genomic sequence data
- Any other unique identifying number, characteristic or code
- Website or social media user names or handles- twitter, facebook, Instagram ,etc...

Be mindful of indirect identifiers- these are not direct identifiers but depending on the size of your population may become identifiers either on their own or when combined with each other. The UWM IRB will take into consideration indirect identifiers and sample size when determining if a dataset is truly deidentified:

- Names or other identifiers of the individual's relatives, employers or household members Direct quotes taken from websites or social media sites as these can be searched and traced back to original setting
- Medical conditions (ex.the 40-year old male who had a brain tumor), hospitalizations, accidents
- Job titles, number of years with an employer, education, income

- Gender, race, ethnicity, age, marital status, household composition, number of children, place of birth, etc...
- Dates- marriage, divorce, graduation, arrest, crime, trial or conviction
- Non-randomly assigned ID numbers- assigning the first participant ID #1, making codes based on personal characteristics- birthdates, etc...