Data *de-identification* is the process of removing or masking personally identifiable information (PII) from data to prevent someone's identity from being revealed. This is important to reduce the risk of connecting subjects' identities with data and is a critical component of ethical human subject’s research. For example, data produced during research might be de-identified to preserve the privacy of participants