

# Data Re-Identification

**Data re-identification or de-anonymization** is the practice of **matching anonymous data with publicly available information, or auxiliary data, in order to discover the individual to which the data belong to.** (Wikipedia)

“It was found that **87%** of the population in the US had reported characteristics that likely made them **unique** based only on {**ZIP-code, gender, date of birth**}. About **50%** of the U.S. population are likely to be **uniquely identified** by only {**place, gender, date of birth**}, where place is basically the city, town, or municipality in which the person resides” (Sweeney, 2000).

“...**four** spatio-temporal points [e.g., **GPS locations**] are enough to **uniquely identify 95%** of the individuals” (de Montjoye et al., 2013)

- Sweeney, L. (2000). Simple demographics often identify people uniquely. *Health (San Francisco)*, 671(2000), 1-34
- De Montjoye, Y. A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*, 3(1), 1-5.