International Journal of Population Data Science

Journal Website: www.ijpds.org

Proof of Concept for a Privacy Preserving National Mortality Register

Schnell, R[1] and Borgs, C[1]

[1]University of Duisburg-Essen

Introduction

National mortality registers are essential for medical research. Therefore, most nations operate such registers. Due to the administrative structure and data protection legislation, there is no such registry in Germany. We demonstrate that a national mortality registry is technically feasible under the given constraints with privacy preserving record linkage (PPRL).

Objectives and Approach

Getting the legal permission to operate a national mortality registry for research will be easier if the linkage can be done without revealing personal identifiers by using PPRL. To estimate precision and recall of different encodings, we used two settings: (1) matching a local mortality registry (n = 14,003) with mortality data of a university hospital (n = 2,466); (2) matching 1 million simulated records from a national database of names with a corrupted subset. This corresponds to a match of all deceased persons with the deceased persons in the largest federal state (n = 205,000).

Results

Linkage results for clear-text identifiers show very high recall and precision. Bloom-Filter based encryptions yield comparable results. Neither precision nor recall declines more than 2%. Phonetic codes yield high precision but low recall. Some variants of Bloom Filter-based encodings yield better results than probabilistic linkage on clear-text identifiers. This is mainly due to the rarely mentioned detail of using different passwords for different identifiers in the same Bloom Filter. Therefore, implementation details of Bloom Filters are more important than commonly thought. Overall, we recommend the use of salted Bloom Filter-based methods with different passwords for different identifiers to increase security and to prevent all known attacks on identifier encryptions.

Conclusion/Implications

Although most PPRL techniques would yield acceptable results in the given setting of a national register, salted Bloom filter encodings are more secure against attacks while still showing high precision and recall. Therefore, we consider a national mortality register using only encrypted identifiers of deceased persons as feasible.

http://dx.doi.org/10.23889/ijpds.v3i4.732