

Anonymisierung von Datenbanken

Zu statistischen Zwecken werden regelmäßig anonymisierte Datenbanken veröffentlicht, z.B. anonymisierte Patientendatenbanken von Krankenhäusern. Forscher können diese dann auf statistische Korrelationen hin untersuchen, die z.B. Hinweise auf Nebenwirkungen von Medikamenten geben können. Aufgrund der sehr sensiblen personenbezogenen Daten in solchen Datenbanken ist eine gute Anonymisierung unbedingt notwendig.

Es gibt jedoch zahlreiche Fälle von nicht hinreichend anonymisierten Datenbanken. Mit verschiedenen Methoden konnten einzelne Personen oder ganze Personengruppen in den anonymisierten Datenbanken re-identifiziert („deanonymisiert“) werden. In der Folge entstanden wissenschaftliche Arbeiten, die Anonymisierungsverfahren untersuchten. Inzwischen gibt es eine Vielzahl von Anonymitätsbegriffen und Anonymisierungsverfahren, die gewisse Garantien bieten.

Im Rahmen der Seminararbeit zum Stand der Forschung sollen ein Vergleich der Anonymisierungsleistung von verschiedenen Verfahren durchgeführt werden. Hierbei sollen sowohl Verfahren aus der Forschung als auch real eingesetzte Verfahren berücksichtigt werden.

Im weiteren Verlauf von „Praxis der Forschung“ können dann z.B. gängige Verfahren genauer untersucht werden oder bestehende Ansätze weiterentwickelt werden.

Ansprechpartner: Gunnar Hartung (ITI Crypto) gunnar.hartung@kit.edu