

6.6 Big Data und das Anonymisierungsproblem

Der Begriff «Big Data» steht für grosse Datenmengen aus vielfältigen Quellen, die mit hoher Verarbeitungsgeschwindigkeit erfasst, gespeichert und für unbestimmte Zwecke auf unbestimmte Zeit für Auswertungen und Analysen verfügbar gemacht werden. Der Umgang mit personenbezogenen Daten ist in der Datenschutzgesetzgebung von Bund und Kantonen geregelt. Big Data stellt bisherige Datenschutzsysteme, die auf dem Ansatz der individuellen Kontrolle basieren, sowie die Datensouveränität, also die Kontrolle jeder einzelnen Person darüber, welche Daten sie herausgeben möchte, allerdings vor erhebliche Herausforderungen. Dies insbesondere, weil Datensätze verschiedener Herkunft miteinander verknüpft und durch automatisierte, lernfähige Softwaresysteme in kürzester Zeit verarbeitet werden können (*machine learning*, *deep learning*, künstliche Intelligenz).

Die Verwendung anonymisierter Daten galt bislang – gerade im Gesundheitsbereich – als Lösung zur Vermeidung datenschutzrechtlicher Konflikte, da anonymisierte Daten nicht unter das Datenschutzrecht fallen. Die Anonymisierung selbst ist jedoch eine Datenbearbeitung im Sinne des Datenschutzgesetzes. Anonymisierung eines Datensatzes bedeutet, dass die Zuordnung einer Einzelangabe zu einer konkreten Person nicht mehr oder nur noch mit einem unverhältnismässigen Aufwand möglich ist. Dazu müssen Daten verändert werden, z.B. durch die komplette Entfernung von identifizierenden Merkmalen aus einem Datensatz oder im Rahmen einer Generalisierung, z.B. durch die Angabe von Wertebereichen. Bei der Zuordnung von Daten zu einer Person kommt es insbesondere darauf an, wie charakteristisch die Daten sind. Dies allein bestimmt, welcher Aufwand betrieben werden muss, um eine De-Anonymisierung (oder Re-Identifikation) zu ermöglichen. Gerade bei genetischen Daten, die hochindividuell sind, ist eine Anonymisierung nur in seltenen Fällen möglich, nämlich dann, wenn keine personenbezogenen Daten vorliegen oder greifbar sind.

Generell gilt: Je mehr Quellen zur Verfügung stehen, aus denen Hintergrundinformationen über Personen bezogen werden und die für einen Abgleich verwendet werden können, desto grösser ist die Wahrscheinlichkeit einer Re-Identifizierung resp. De-Anonymisierung. Zudem darf das Risiko der De-Anonymisierung auch mit Blick auf die sich immer weiter entwickelnden Analysetechniken und die wachsenden Möglichkeiten des Austauschs von Daten nicht unterschätzt werden. Im Zusammenhang mit Big Data Analytics, insbesondere der Weiterentwicklung selbstlernender Algorithmen, besteht eine grosse und weiter zunehmende Wahrscheinlichkeit, dass eine irreversible Anonymisierung von Daten grundsätzlich und auf Dauer technisch nicht mehr möglich ist. Entsprechend weisen Expertengruppen und Fachleute ausdrücklich darauf hin, dass heute aus technischer Sicht keine absolut sichere Anonymisierung möglich sei.

Letzte Aktualisierung am 12.06.2026