

III. Wie funktionieren Hash-Werte?

Hash-Werte sind eine weitverbreitete Technik im Sicherheitsbereich. In der Informatik auch als Prüfsumme bekannt, stellen sie eine einfache Maßnahme zur Gewährleistung von Datenintegrität bei der Datenübermittlung oder -speicherung dar. Hash-Werte werden durch mathematische Formeln berechnet und haben die besondere Eigenschaft – wie Prüfsummen – dass man allein anhand des Hash-Wertes nicht mit praktikablen Aufwand auf den ursprünglichen Wert, von welchem der Hash-Wert abgeleitet ist, zurückschließen kann.

Ein Beispiel für einen Hash-Wert, hier über die Identität des Autors dieses Papiers: Der Wert
Matthias, Mehl dau

wird beispielsweise nach dem SHA256-Algorithmus zu folgendem Hash-Wert umgerechnet:

823b4194cd6cdb88e2298d54dbeb7ea54ade328a16d5500f1bdbcfd2ab62731b

Um von dem Hash-Wert wieder zurück auf den Namen schließen zu können, verbliebe als einzige Möglichkeit alle möglichen Kombinationen, in unserem Beispiel aus Vor- und Nachnamen, durchzuprobieren. Dieser Aufwand wäre technisch nur schwer zu realisieren.

Hash-Werte werden etwa für digitale Signaturen, die gemäß dem Deutschen Signaturgesetz bindend sein müssen, verwendet. Das Bundesamt für Sicherheit in der Informationstechnik (BSI) empfiehlt zu jeder Zeit, welche Algorithmen zur Berechnung von Hash-Werten als sicher gelten.

IV. Wie können Hash-Werte einer Schülerdatenbank dienlich sein?

Speichert man nun in der landesweiten Schülerdatenbank ausschließlich den Hash-Wert sowie die Schule, von welcher die Anmeldung stammt, genügt dies, um zu erkennen, dass ein Schüler an zwei Schulen angemeldet wurde. Die Landesstelle kann sich mit dem Hash-Wert an die betreffenden Schulen wenden, um die Identität des Schülers zu erfahren.

Wichtig für dieses Verfahren ist die einheitliche, eindeutige Erfassung der für die Erzeugung des Hash-Werts relevanten Daten. So ist die Zusammenstellung und Formatierung der verwendeten Felder – etwa die zu erfassenden zusätzlichen Vor- und Nachnamen, der Jahrgang oder die Namen der Eltern – genau zu bestimmen. Eine einheitliche Software an allen Schulen sowie ein klares Prozedere zur Datenerfassung ist ausreichend, um die Eindeutigkeit der Daten sicherzustellen. Hierfür ist es unabdingbar, dass man sich auf eine einheitliche Kodierung einigt, um etwa Sonderzeichen korrekt zu erfassen.

V. Fazit

Dieses Verfahren würde ferner die Eigenschaft mit sich ziehen, dass die Landesstelle – sofern die Daten nach Beginn des Schuljahres noch vorliegen – an Hand einer bekannten Identität eines Schülers feststellen kann, an welcher Schule er oder sie angemeldet ist.

Um den postulierten Zweck einer Schülerdatenbank sicherzustellen, ist das hier skizzierte Verfahren vollkommen ausreichend. Ob eine Datenschutz-unfreundlichere Umsetzungen einer verfassungsrechtlichen Prüfung standhalten würde, ist anzuzweifeln.

Berlin, den 12. Januar 2009