Prof. Dr. David Basin: Risikofolgenabschätzung zur Verwendung der AHV-Nummer als Personenidentifikator

5 Risikoanalyse

Im Folgenden werden zunächst die relevanten Kategorien motivierter und qualifizierter Angreifer aufgezeigt. In einem zweiten Schritt werden die Systemmerkmale, die für die hier betrachteten Szenarien massgebend sind, und deren Auswirkungen auf die Datensicherheits- und -schutzrisiken geprüft. Schliesslich werden die Datenschutzrisiken in Verbindung mit den verschiedenen Szenarien untersucht.

5.1 Wer könnte die Schweizer Register angreifen?

Wie unter Kapitel 4.1 erläutert, entstehen Datenschutzrisiken bei Attacken durch Angreifer. In diesem Kapitel wird dargelegt, dass es tatsächlich Angreifer gibt, die genügend motiviert und qualifiziert sind, um sich Zugriff auf die Daten in verschiedenen Schweizer Registern zu verschaffen.

Bei Datenregistern können *interne Benutzer* Urheber problematischer Datenoperationen sein. Im Allgemeinen sind Attacken und Vergehen interner Benutzer ein ernsthaftes und oft unterschätztes Problem [14]. Ein interner Benutzer kann seine Zugriffsberechtigung für unberechtigte Zwecke verwenden. Er kann zum Beispiel aus Neugier den Status einer bestimmten Person abfragen oder einem Kollegen in einer anderen Behörde helfen wollen. Verärgerte Arbeitnehmer wollen womöglich Schaden anrichten, z. B. indem sie aus Rache den Ruf ihres Arbeitgebers schädigen wollen. Wenn Personendaten für den Identitätsdiebstahl oder für gezielte Werbung verkauft werden können, könnten interne Benutzer mit kriminellen Absichten ihre Zugriffsberechtigung dafür missbrauchen, Daten zu beschaffen, um sie auf dem Schwarzmarkt zu verkaufen. Schliesslich können interne Benutzer auch nachlässig sein. Unterlassen es Benutzer, ihre Systeme angemessen zu sichern, die Speicherung besonders schützenswerter Daten auf ein Minimum zu beschränken, Daten zu löschen usw., stellt dies an sich noch keine problematische Datenoperation dar, es kann jedoch externen Angreifern den Zugriff auf die Daten erleichtern.

Ein System kann auch von *externen Angreifern* attackiert werden. Hacker sehen es als Herausforderung, in Systeme einzubrechen, und die Online-Veröffentlichung besonders schützenswerter Daten kann als Zeichen des Erfolgs angesehen werden. Politisch motivierte Hacktivisten können Attacken ausführen, um ihre Gegner zu demütigen. Kriminelle brechen in Systeme ein, um gestohlene Daten zu verkaufen. Nationalstaatliche Feinde können an besonders schützenswerten Daten einschliesslich Finanzdaten, Gesundheitsdaten oder Strafregisterdaten interessiert sein, um Informationen über Personen zu beschaffen, um z. B. Steuerflüchtlinge zu fassen oder Wahlen zu beeinflussen. Ein Beispiel für die Fähigkeiten und die Bestrebungen, mit denen nationalstaatliche

Angreifer erfolgreich in gehärtete, schwer gesicherte Systeme einbrechen, liefert der jüngste MELANI-Bericht zum Spionagefall bei der RUAG [8].

Anhand zweier konkreter Beispiele soll nun aufgezeigt werden, welche Anreize bestehen, Personendaten zu sammeln und zu aggregieren (erstes Beispiel), und welche Schäden bereits aus der Kompromittierung eines einzigen Systems resultieren können (zweites Beispiel).

Moneyhouse.ch Das erste Beispiel aus [5] zeigt, wie Personendaten in Geld umgemünzt werden können. Es illustriert auch, wie einzelne Datenquellen miteinander zu Persönlichkeitsprofilen verknüpft werden können, sodass die Daten zusammen einen wesentlichen Teil des Lebens einer Person offenbaren. Schliesslich geht aus dem Beispiel hervor, dass solche Risiken auch entstehen können, ohne dass aktiv in Systeme eingebrochen wird. Es genügt, öffentlich verfügbare nicht pseudonymisierte Daten zu verknüpfen.

Das Schweizer Unternehmen itonex AG betrieb die Plattform moneyhouse.ch, auf der Daten über Unternehmen und Privatpersonen gesammelt und verkauft wurden. Zum Zeitpunkt des Berichts [5] bot das Unternehmen eine Reihe kostenloser und kostenpflichtiger Informationsdienste, deren Wert darin bestand, verschiedene Datenquellen einschliesslich Daten aus Verwaltungsregistern wie dem Handelsregister miteinander zu verknüpfen. Durch die Verknüpfung der Daten erstellte itonex Persönlichkeitsprofile, die über ihre Plattform oder über eine Google-Suche zugänglich waren. Zahlende Kunden konnten auf die Daten über die Personen einschliesslich Vor- und Nachnamen, Wohnort, Postleitzahl, Alter, Geburtsdatum, Beruf, Haushaltsmitglieder (einschliesslich Kindern) und Nachbarn zugreifen. Die Daten zur Wohnsituation umfassten Luftaufnahmen und Google-StreetView-Bilder der Liegenschaft sowie Angaben zu Gebäudetyp, Anzahl Haushalte im Gebäude, Baukosten, Bauperiode und Anzahl Etagen. Alles in allem lieferten die verknüpften Daten ein umfassendes Bild der Lebenssituation identifizierbarer Personen und damit einen wesentlichen und sensiblen Aspekt ihres Persönlichkeitsprofils.

Equifax Das zweite Beispiel stammt aus den Vereinigten Staaten. Im September dieses Jahres berichtete Equifax, eine der grössten Wirtschaftsauskunfteien der USA, von einem Datenleck aufgrund einer Sicherheitslücke in einer Webapplikation [19, 10]. Hacker nutzten diese aus, um sich Zugriff auf die Namen, Adressen, Sozialversicherungs-, Kreditkarten- und Fahrausweisnummern sowie Dokumente über Streitigkeiten mit Personendaten von 143 Millionen Amerikanern – fast die Hälfte der US-Bevölkerung – zu verschaffen.

Die Folgen dieses Lecks wiegen schwer. Die Daten werden noch für Jahrzehnte auf dem Schwarzmarkt im Umlauf sein und verschiedene Formen des Identitätsdiebstahls ermöglichen. Wie in [19] erläutert, kombinieren Betrüger die gestohlenen Sozialversicherungsnummern und allenfalls weitere Informationen über die Inhaber typischerweise mit einer fremden E-Mail-Adresse und

beantragen neue Kreditkarten, die sie dann selbst verwenden können. Geduldigere Schwindler benutzen die neuen Identitäten sogar, um sich zusätzliche Kreditkarten oder Darlehen zu beschaffen. In der Folge belasten sie diese gleichzeitig bis ans Limit und lassen womöglich zehntausende Dollar mitgehen.

Dieses Beispiel zeigt ebenfalls die Risiken, welche die Aggregation von Personendaten birgt. Es gibt motivierte Angreifer, die einen starken Anreiz haben, die Daten zu stehlen und zu Geld zu machen. Ausserdem ist die Sicherheit der Systeme nicht perfekt und Zugangskontrollen und andere Mechanismen können geknackt werden. Für eine Kreditauskunftei mit Informationen über die Konsumenten wie Equifax ist die Vertrauenswürdigkeit zentral, und es kann davon ausgegangen werden, dass ein solches Unternehmen hohe Sicherheitsstandards einhält. Nicht zuletzt hat der Fall langfristige Folgen: Sozialversicherungsnummern (wie die AHVN13) werden fest zugewiesen und andere persönliche Merkmale lassen sich nur schwer ändern. Diese Datenkompromittierung wird den Betroffenen noch lange Probleme bereiten.

5.2 Systemmerkmale mit einem Einfluss auf die Datenschutzrisiken

In diesem Kapitel werden einige allgemeine Bemerkungen angebracht, die in der Folge auf den Vergleich der verschiedenen Szenarien angewendet werden.

5.2.1 Absolute Sicherheit ist eine Illusion

Computersysteme sind in der Praxis nie zu hundert Prozent sicher. Sicherheit ist eine Vorbedingung für den Datenschutz, denn der unerlaubte Zugriff auf Daten muss verhindert werden. Leider besteht immer die Gefahr, dass Attacken erfolgreich sind, wenn Schwachstellen in der Hardware, den Betriebssystemen, den Anwendungen, beim Personal, in den Prozessen usw. ausgenützt werden. Systeme können folglich kompromittiert werden und die Daten, die darin gespeichert sind, können von den Angreifern ausgelesen werden. Dies selbst dann, wenn sie durch Sicherheitsmassnahmen auf dem neusten Stand der Technik geschützt sind.

5.2.2 Nicht alle Systeme sind gleich sicher

Die Sicherheit eines Systems hängt davon ab, wie viel in die Sicherung investiert wird und wie wirksam die getroffenen technischen und organisatorischen Massnahmen sind. Zur Abklärung, ob die Massnahmen auch tatsächlich wie beabsichtigt funktionieren, sind auch Qualitätssicherungsmassnahmen wichtig. Die Systeme des Bundes (siehe die Besprechung unter Kap. 2.2.2) entsprechen aufgrund der hohen Standards und der Qualitätssicherungsprozesse der für sie zuständigen Organisationen wie dem ISB in der Regel relativ hohen Sicherheitsanforderungen.

Anders präsentiert sich die Ausgangslage bei Systemen ausserhalb des

Bundes, die von Organisationen verwaltet werden, die zur systematischen Verwendung der AHVN13 ermächtigt sind – sei dies auf Stufe Kanton und Gemeinde oder in nichtstaatlichen Organisationen wie Spitälern und Schulen. Der Bund stellt niedrigere Anforderungen an den Schutz dieser Systeme und die betreffenden Qualitätssicherungsprozesse [16] (obwohl die Anforderungen des Bundes zum Teil durch zusätzliche Anforderungen anderer Verwaltungen ergänzt werden). Sie sind im Allgemeinen wesentlich weniger sicher.

5.2.3 Umgang mit Identitätsinformationen

Wie unter Kapitel 3.1 erläutert gibt es verschiedene Möglichkeiten, Registerdaten in Informationssystemen zu organisieren. Generell birgt es Datenschutzrisiken, von Teilen eines UPI-Registers lokale Kopien zu erstellen. Wenn der Angreifer genügend (weniger gut geschützte) Systeme kompromittieren kann, so kann er grosse Teile der UPI-Datenbank rekonstruieren. Darüber hinaus ist es mit einem grossen Aufwand verbunden, die Konsistenz redundanter Kopien zu gewährleisten.

Ein weiteres Datenschutzrisiko stellt die Tatsache dar, dass UPI-Daten in Verwaltungsregistern allgemein zusammen mit Identitätsattributen gespeichert werden (siehe Kap. 2.2.3) und die Identitätsattribute an sich bereits genügen, um Daten und Verknüpfungstabellen sehr präzis zu reidentifizieren (siehe Kap. 2.2.3). Die Datenschutzrisiken werden somit nicht wesentlich vermindert, wenn lediglich die AHVN13-Attribute oder die Quasi-Identifikatoren aus den Verwaltungsregistern entfernt werden. **Ein erhöhter Datenschutz kann sichergestellt werden, indem administrative und organisatorische Datensätze von der AHVN13 *und* Quasi-Identifikatoren, anhand welcher grosse Teile der Bevölkerung reidentifiziert werden können, entkoppelt werden.** Dafür muss bei der Organisation der Daten allerdings ein anderer Ansatz verfolgt werden als jener, der den Organisationen, die die AHVN13 systematisch verwenden dürfen, gegenwärtig empfohlen wird (siehe Kap. 3.1.1).

5.2.4 Wo und wie Daten gespeichert werden

Dieser Punkt hängt eng mit dem vorherigen zusammen. Daten können nicht nur auf verschiedene Weisen in einem Informationssystem *organisiert* werden, die Datentabellen können auch über separate Datenbanken und Plattformen in einem verteilten Informationssystem *verteilt* werden. Wenn die Verteilung sauber aufgesetzt ist, kann sie dieselbe Funktionalität bieten wie eine zentrale Speicherung (die Datensätze können z. B. je nach Bedarf zu Geschäftszwecken verknüpft werden). Unter dem Aspekt der Sicherheit und des Datenschutzes weist eine Verteilung der Daten jedoch bessere Eigenschaften auf als eine zentrale Speicherung.

Im Folgenden ein Beispiel auf Grundlage von Abbildung 4 und anhand zweier Szenarien:

Szenario 1: In jedem Sektor wird das sektorielle Register in der *gleichen* Datenbank gespeichert wie die sektorielle Verknüpfungstabelle.

Szenario 2: In jedem Sektor werden die beiden Tabellen in verschiedenen Datenbanken auf verschiedenen Computern gespeichert. Zur Unterstützung der Geschäftsprozesse wird die verteilte Abfragebearbeitung unterstützt, z. B. kann ein sektorspezifischer Identifikator auf eine AHVN13 gemappt werden, indem Daten der beiden Datenbanken kombiniert werden, wenn dies für die Geschäftszwecke (absolut) notwendig ist. Für eine erhöhte Sicherheit erfolgt jegliche Kommunikation über eine angemessene Authentifizierung.

Im ersten Szenario könnte ein Angreifer eine Softwareschwachstelle in einer Datenbank mit einem sektoriellen Register ausnützen, um *beide* Datentabellen des Sektors zu kompromittieren. So könnte er den gesamten Inhalt auslesen, die beiden Tabellen verbinden und so das sektorielle Register mit AHVN13-Nummern verknüpfen. Im zweiten Szenario wäre es auf Grundlage der Kompromittierung des sektoriellen Registers nicht möglich, die Daten zu reidentifizieren oder mit anderen Daten ausserhalb des Sektors zu verknüpfen. Dafür müsste der Angreifer *zusätzlich* (i) die Datenbank mit der Verknüpfungstabelle oder (ii) den (eine Authentifizierung erfordernden) Kommunikationskanal kompromittieren, der für die verteilte Abfragebearbeitung verwendet wird, damit er im verteilten System gefälschte Abfragen vornehmen kann. Die Reidentifizierung oder Kombination von Daten verschiedener Register ist im zweiten Szenario folglich schwieriger und die Sicherheits- und Datenschutzrisiken sind entsprechend geringer.

Der Unterschied zwischen den beiden Szenarien kann noch deutlicher ausfallen, wenn im zweiten Szenario weitere Sicherheitsmassnahmen in Betracht bezogen werden, z. B. gehärtete Systeme, verschlüsselte Speicherung, Schlüsselverwaltung unter Verwendung von Hardware-Sicherheitsmodulen usw. In einem angemessen designten System können solche Massnahmen das Risiko für bestimmte Angriffe und damit die Datenschutzrisiken allgemein vermindern. Dem stehen eine höhere Komplexität des Systems und höhere Kosten gegenüber.

5.3 Szenarienspezifischer Vergleich

Im Folgenden werden die unter Kapitel 1.2 aufgeführten Szenarien wieder aufgegriffen und deren Risiken verglichen.

5.3.1 Breitere Verwendung der AHVN13

Zurzeit sind über 14 000 Organisationen berechtigt, die AHVN13 systematisch zu verwenden, viele davon mit relativ unsicheren Systemen. Auch wenn es beim Status quo bleibt, besteht somit eine erhebliche Gefahr, dass Angreifer Systeme kompromittieren und Datensätze mit AHVN13-Nummern und persönlichen

Attributen auslesen. Folglich kann selbst der Verlust der Daten eines einzigen Registers dazu führen, dass besonders schützenswerte Personendaten offengelegt werden. Werden mehrere Systeme kompromittiert, können deren Daten aufgrund der Verwendung der AHVN13 sogar noch einfacher kombiniert werden.

Je mehr Organisationen die AHVN13 in ihren Informationssystemen verwenden, desto mehr werden diese Risiken zunehmen. Das erhöhte Risiko ist auf zwei Faktoren zurückzuführen:

1. Je mehr Organisationen Personendaten in Verbindung mit der AHVN13 sammeln, speichern und bearbeiten, desto höher ist die Wahrscheinlichkeit, dass einige ihrer Systeme kompromittiert werden.

2. Je mehr Organisationen ausserhalb der Bundesverwaltung die AHVN13 systematisch verwenden dürfen, desto mehr Daten, die anhand der AHVN13 direkt und eindeutig mit Personen verknüpft werden können, werden in relativ unsicheren IT-Systemen ausserhalb des Bundes gesammelt, gespeichert und bearbeitet werden. Das erhöhte Risiko geht darauf zurück, dass bei diesen Systemen die Sicherheitsstandards tiefer und die Qualitätssicherungsprozesse schwächer sind und dass beim Design dieser Systeme typischerweise nicht die Sicherheit im Vordergrund stand (siehe Kap. 5.2.2).

In Bezug auf beide Faktoren ist zu sagen, dass hinsichtlich der Daten von Organisationen, die bereits aktuell Personendaten sammeln und bearbeiten und zusammen mit qualitativ hochstehenden Quasi-Identifikatoren (z. B. Vorname, Name, Geburtsdatum) speichern, schon jetzt das Risiko besteht, dass sie über die Quasi-Identifikatoren reidentifiziert und verknüpft werden. Eine *zusätzliche Speicherung* der AHVN13 birgt in diesem Fall ein kaum höheres Risiko (siehe Kap. 2.2.3 und 5.2.3), da die Reidentifizierung und Verknüpfung bereits möglich ist.

Gemäss diesen Überlegungen könnte auch argumentiert werden, dass es keinen Grund gibt, die AHVN13 *nicht* in Registern *zu speichern*, da dies ja minimale Auswirkungen auf die Datenschutzrisiken hat. Dieses Argument wurde denn auch vorgebracht, um die gegenwärtige Verwendungsweise und deren Ausweitung zu rechtfertigen. Dieser Schluss ist grundsätzlich korrekt, *aber nur in der aktuellen Ausgangslage,* in der die AHVN13-Nummern immer mit Identitätsattributen verbunden sind. Doch die Identitätsattribute müssen nicht zwingend im gleichen Register gespeichert werden. Die redundante Speicherung von Identitätsinformationen entspricht vielmehr weder einem guten Datenbankdesign noch einer guten Datenschutzpraxis.
Die wichtige Frage ist folglich, wie hoch die Datenschutzrisiken aufgrund der gegenwärtigen Verwendungsweise und deren Ausweitung verglichen mit anderen Szenarien sind, in denen die Risiken potenziell reduziert werden können, indem höhere Standards angewandt werden (Kap. 5.2.2), die Speicherung der

Identitätsattribute und der damit verbundenen Personendaten reorganisiert und vermindert wird (Kap. 5.2.3) und die Daten verteilt gespeichert werden (Kap. 5.2.4). Dieser Frage wird in den nächsten beiden Fällen nachgegangen.

5.3.2 Andere nichtsprechende, sektorspezifische Identifikatoren

Die AHVN13 ist ein nichtsprechender Identifikator, der sektorübergreifend verwendet wird. Sektorspezifische Identifikatoren sind ebenfalls nichtsprechend, deren Verwendung ist jedoch *lokal* auf einen bestimmten Sektor beschränkt. Wenn ein sektorspezifischer Identifikator verwendet wird, so können Register dieses Sektors nicht direkt mit Registern anderer Sektoren verknüpft werden. Durch eine geeignete Umsetzung dieser Massnahme kann folglich das Risiko vermindert werden, dass ein Angreifer grosse Mengen von Personendaten sektorübergreifend aggregieren kann, sodass sie einen bedeutenden Teil des Lebens der Betroffenen offenbaren.

Dabei ist zu betonen, dass die Datenschutzrisiken angesichts des *heutigen Ansatzes*, Registerdaten zusammen mit Identitätsattributen aufzubewahren, nicht wesentlich vermindert würden, wenn die AHVN13-Nummern einfach durch sektorspezifische Nummern ersetzt würden. Die Verbindung zwischen den Registerdaten und den Personen kann weiterhin hergestellt werden, indem die Identitätsattribute als Quasi-Identifikatoren verwendet werden. Anhand der Identitätsattribute können auch Datenbanktabellen sehr präzis miteinander verknüpft werden.

Es ist möglich und erstrebenswert, als Alternative zur AHVN13 sektorspezifische Identifikatoren einzuführen. Dies muss allerdings auf eine geeignete Weise geschehen. Es ist insbesondere erforderlich, die Speicherung von Identitätsattributen auf ein Minimum zu beschränken und sämtliche Verknüpfungstabellen angemessen zu schützen (siehe Kap. 5.2.2–5.2.4). **Durch die entsprechend gestaltete Einführung sektorspezifischer Identifikatoren könnten die Sicherheitsrisiken aufgrund der gegenwärtigen Verwendungsweise der AHVN13 und deren kontinuierlichen Ausweitung reduziert werden.**

Dafür müssen die aktuellen Softwarearchitekturen womöglich angepasst werden. Ein Beispiel dafür bietet das Szenario 2 unter Kap. 5.2.4, in dem vorgeschlagen wird, für manche Geschäftsfälle eine verteilte Abfragebearbeitung einzuführen. Als weiteres Beispiel kann das unter 3.2.3 besprochene Gesundheitswesen dienen. Die Identitätsattribute sollten in diesem Bereich nicht in Master Patient Indices der Gemeinschaften von Gesundheitsdiensten gespeichert werden, und die Speicherung der Identitätsattribute in Datenbanken der Dienstleister im Gesundheitswesen und in den Gesundheitsakten selbst sollte auf ein Minimum beschränkt werden. Stattdessen sollte hauptsächlich mit dem damit verbundenen Pseudonym gearbeitet werden. Die Patienten ihrerseits sollten sich mit ihrem Pseudonym identifizieren müssen (z. B. auf einer «e-health»-Karte); andere Identitätsattribute könnten ebenfalls bekannt gegeben werden, die zusätzlichen Attribute sollten jedoch nicht gespeichert werden.

Wenn sektorspezifische Identifikatoren für administrative Prozesse mit Identitätsinformationen verknüpft werden müssen, kann dies stets durch eine (verteilte) Abfragebearbeitung unter Verwendung der Verknüpfungstabellen erfolgen. Die Anforderungen an den Schutz dieser Tabellen sollten überdies sehr hoch sein und es sollten besondere Massnahmen getroffen werden, um sie zu sichern und um zu gewährleisten, dass sie ausschliesslich zu angemessenen Zwecken verwendet werden. Dafür bestehen zahlreiche Designmöglichkeiten. Die Verknüpfungstabellen können beispielsweise in der ZAS, in einer sektorspezifischen Verwaltung wie dem EHRA oder einer privaten Organisation wie einer Gemeinschaft von Gesundheitsdienstleistern gespeichert werden. Sofern die Verknüpfungstabellen angemessen geschützt sind, können diese Möglichkeiten mit deutlich weniger Datenschutzrisiken umgesetzt werden als wenn direkt die AHVN13 verwendet wird.

5.3.3 Interne Kombination der AHVN13 mit extern verwendeten Identifikatoren

Das Handelsregister ist ein Beispiel für eine derartige Kombination. Grundsätzlich ist der einzige Unterschied zwischen einem intern und einem extern verwendeten sektorspezifischen Identifikator, dass es für Angreifer ein Leichtes ist, extern verfügbare Daten zu beschaffen. Bei weniger schützenswerten Daten wie jenen im Handelsregister kann ein solches Design sinnvoll sein. Dies insbesondere, wenn die Daten öffentlich sein sollen. **In diesem Fall wird das Risiko, dass Angreifer die Daten aus Registern ausserhalb des Sektors aggregieren, durch die Verwendung externer sektorspezifischer Identifikatoren im Vergleich mit der unmittelbaren Verwendung der AHVN13 vermindert.** Das Ausmass der Risikominderung hängt jedoch davon ab, wie leicht und erfolgreich Personen anhand anderer ebenfalls veröffentlichter Identitätsattribute reidentifiziert werden können.

27. September 2017