PSEUDONYMISIERUNG PERSONENBEZOGENER DATEN

Faktenblatt

Adam Knoop

VIACRYP B.V. Danzigerkade 19, NL-1013 AP Amsterdam

Inhaltsverzeichnis

1 Dokumentinformationen 2
1.1 Dokumenthistorie 2
2 Einleitung 3
2.1 Hintergrund 3
2.1.1 Viacryp B.V. 3
2.1.2 Rechtslage bezüglich personenbezogener Daten 3
2.1.3 Pseudonymisierung 4
2.2 Ziel dieses Dokuments 4
3 Die Pseudonymisierungsstraße 5
3.1 Einleitung 5
3.2 Spaltung der Daten 5
3.3 Architekturschema 6

1 Dokumentinformationen

1.1 Dokumenthistorie

Version	Verfasser	Bezeichnung	Datum
0.9	Arjo Hooimeijer	Konzeptfassung	03.03.2014
1.0	Arjo Hooimeijer	Endfassung	05.03.2014
1.1	Arjo Hooimeijer	Erläuterung zur Anonymität und Pseudonymisierung überarbeitet	16.02.2015
1.2	Adam Knoop	Übersetzung auf Deutsch	01.06.2017

2 Einleitung

2.1 Hintergrund

2.1.1 Viacryp B.V.

Viacryp B.V. ist eine Gesellschaft niederländischen Rechts und sie hat sich auf die Unterstützung von Betrieben spezialisiert, die mit dem Datenschutzgesetz zu tun haben. Marktteilnehmer, die für ihre Kommunikation mit der Zielgruppe und für ihre innerbetrieblichen Prozesse, personenbezogene Daten zur Erreichung ihrer Zielsetzungen brauchen, unterliegen strengen Richtlinien in Bezug auf den Datenschutz. Die Lösung von Viacryp B.V. dient dazu, diese personenbezogenen Daten nutzen zu können und dabei das Datenschutzrecht der Betroffenen zu wahren.

Viacryp B.V. ist ein unabhängiges Unternehmen, das sich seit dem 1. Juli 2013 als *Trusted Third Party* auf dem Gebiet der Pseudonymisierung personenbezogener Daten engagiert.

2.1.2 Rechtslage bezüglich personenbezogener Daten

Das Datenschutzgesetz stellt strenge Anforderungen an die Verarbeitung von Daten, die sich auf jegliche Weise auf natürliche Personen beziehen. Dabei gilt, dass diese Herstellung des Personenbezugs im weitesten Sinne des Wortes aufgefasst werden muss. Das beinhaltet, dass dabei sowohl die direkt bestimmbaren Daten, wie die Steuernummer, Name, Anschrift oder IP-Adresse wie auch indirekt bestimmbare Daten, wie das Geburtsdatum oder eine vollständig ausgefüllte Postleitzahl berücksichtigt werden müssen.

Für die Verarbeitung personenbezogener Daten gelten bestimmte Grundlagen, die im Datenschutzgesetz aufgeführt werden. Das Datenschutzgesetz schreibt ferner eine Reihe allgemeiner Ausgangspunkte im Hinblick auf das Speichern und Verarbeiten personenbezogener Daten vor:

- Datenminierung (nicht mehr speichern als erforderlich)
- Nicht länger aufbewahren als erforderlich
- Angemessene Sicherheitsmaßnahmen, um die unnötige Erhebung und weitere Verarbeitung personenbezogener Daten zu verhindern[1]

Wenn keine zulässige Rechtsgrundlage für die Verarbeitung vorliegt, dürfen Daten ausschließlich anonym verarbeitet werden.

Die zuständige niederländische Aufsichtsbehörde AP erklärt, dass bei der Anwendung von Pseudonymisierung die folgenden Voraussetzungen erfüllt werden sollen[2]:

I. Die Pseudonymisierung wird auf kompetente Weise durchgeführt. Dabei findet die erste der beiden durchgeführten Verschlüsselungen beim Anbieter der Daten statt.
II. Es wurden technische und organisatorische Maßnahmen zur optimalen Rücknahmefestigkeit des Verschlüsselungsverfahrens getroffen.
III. Die verarbeiteten Daten sind nicht indirekt identifizierend.
IV. Diese drei Voraussetzungen unterliegen regelmäßig abzuhaltender Audits.

Außerdem ist die Pseudonymisierungslösung auf klare und vollständige Weise in einem aktiv veröffentlichten Dokument darzustellen, damit jeder Betroffene in Erfahrung bringen kann, welche Garantien die gewählte Lösung bietet.

[1] CBP Richtlinie zum Schutz personenbezogener Daten (Februar 2013)
[2] Siehe http://cbpweb.nl/sites/default/files/downloads/uit/z2006-1382.pdf

2.1.3 Pseudonymisierung

Viacryp B.V. erbringt verschiedene Dienstleistungen, die zum Schutz personenbezogener Daten beitragen, indem die Menge der lesbar gespeicherten und verarbeiteten personenbezogenen Daten auf ein Mindestmaß reduziert und pseudonymisiert wird.

Unsere Prozesse und Verfahren entsprechen den in den Richtlinien der niederländischen Aufsichtsbehörde CBP genannten Kriterien, wenn anonymer Output das Ziel der Verarbeitung ist.

2.2 Ziel dieses Dokuments

Dieses Dokument ist zu veröffentlichen, um der Anforderung V., dass die Pseudonymisierungslösung auf klare und vollständige Weise in einem aktiv veröffentlichten Dokument dargestellt wird, gerecht zu werden.

Mit diesem Ziel wird in Kapitel 3 die Pseudonymisierungslösung in Form der **Pseudonymisierungsstraße** ausführlich dargestellt.

3 Die Pseudonymisierungsstraße

3.1 Einleitung

Eine **Pseudonymisierungsstraße** besteht aus einer oder mehreren **Quellen**, die dem **Pseudonymizer** über eine **Supply**-Plattform Daten liefern. Wenn die Daten dieser Quellen pseudonymisiert sind, werden sie über eine **Delivery**-Plattform einem **Abnehmer** geliefert, der auf Grundlage von **Pseudonymen** Daten aus verschiedenen **Quellen** kombinieren und Analysen hinsichtlich des Verhaltens durchführen kann, ohne über die entsprechenden personenbezogenen Daten zu verfügen.

Bei Bedarf werden verschiedene **Quellen** in bestimmten Konfigurationen kombiniert, bevor sie dem **Abnehmer** bereitgestellt werden. In einem solchen Fall werden dem **Abnehmer** keine **Pseudonyme** geliefert und die Daten werden zu Analysezwecken vorbereitet, um eine indirekte Wiederherstellungsmöglichkeit des Personenbezugs dieser Daten zu verhindern.

3.2 Spaltung der Daten

Die Straße, mit der Viacryp B.V. als Trusted Third Party arbeitet, führt sowohl zu einer „Spaltung der Daten“, dabei werden personenbezogene Daten (Wer) und das zu analysierende Verhalten (Was) in einem frühen Prozessstadium voneinander gespalten wie auch zu einer „Trennung der Daten“. Dabei werden die gespaltenen **Wer** und **Was-Daten** des Hashings und der Verschlüsselung unterzogen. Damit wird erreicht, dass an keiner Stelle im Prozess (mit Ausnahme bei der Quelle der ursprünglichen Daten) sowohl das **Wer** wie auch das **Was** in lesbarer Form herangezogen werden können.

Das kann anhand der folgenden Tabelle erläutert werden:

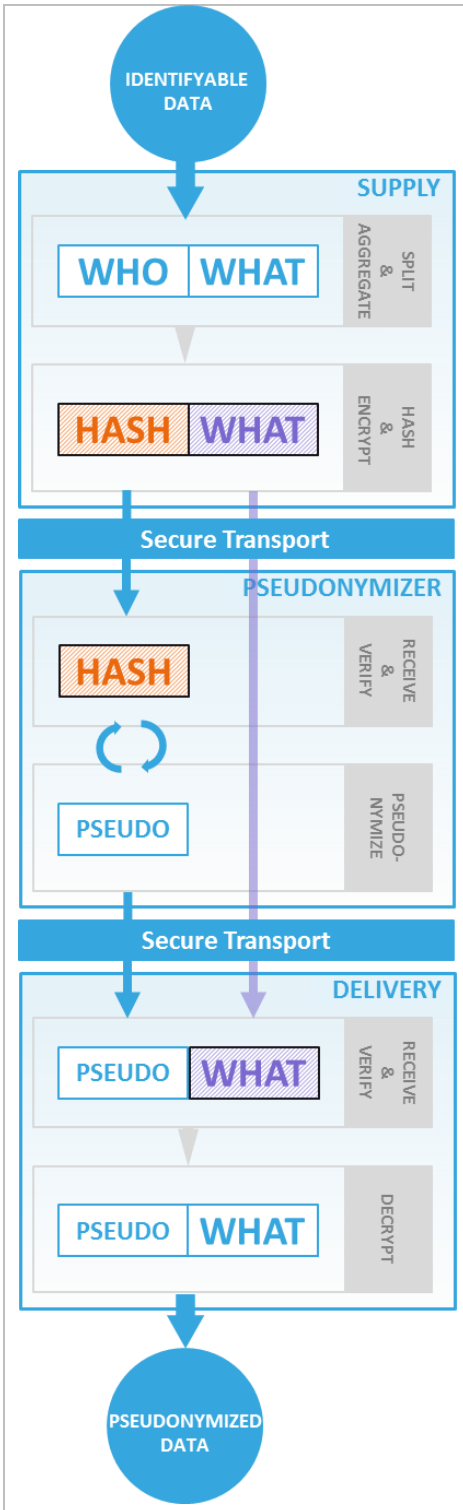
	Quellen	Supply	Pseudonymizer	Delivery	Abnehmer
Wer	Original	Hashed	Hashed	Pseudonym*)	Pseudonym*)
Was	Original	Encrypted	Encrypted	Encrypted	Original

*) Optional

- Nur die **Quelle** verfügt über das originale **Wer** und **Was**
- Auf der **Supply**-Plattform werden die **Wer-Daten** gehasht und verschlüsselt und die **Was-Daten** verschlüsselt, bevor die Daten die **Quelle** verlassen.
- **Pseudonymizer** verfügt ausschließlich über die gehashten **Wer-Daten**, um daraus **Pseudonyme** erstellen zu können, die entweder (zusammen mit den gehashten **Was-Daten**) dazu eingesetzt werden können, Daten zu Analysezwecken vorzubereiten oder mit den verschlüsselten **Was-Daten** kombiniert werden können**.** Anschließend werden diese Daten der **Delivery**-Plattform beim Abnehmer zugeführt.
- Auf der **Delivery**-Plattform werden die eingegangenen Daten entschlüsselt und dem Abnehmer zur Verfügung gestellt, der daraufhin Analysen durchführen kann, ohne über personenbezogene Daten zu verfügen.

3.3 Architekturschema

Die gesamte Straße, die jede Datei einer **Quelle** durchläuft, wird anhand des folgenden Architekturschemas dargestellt:



Quelle

Führt Daten in einem vereinbarten Format einem **Supply**-Modul zu, das sich im Bereich der Quelle befindet.

Split & Aggregate

Quelldaten werden in bestimmbare Daten (**Wer**) und Verhalten (**Was**) getrennt. **Was** wird auf einem derartigen Niveau aggregiert, dass der (indirekte) Personenbezug nicht wieder hergestellt werden kann.

Hash & Encrypt

Wer wird mithilfe einer Einwegverschlüsselung (**Hash**) unleserlich gemacht. **Was** wird derart verschlüsselt, dass es ausschließlich vom **Abnehmer** entschlüsselt werden kann.

Receive & Verify

Diese Daten werden über eine gesicherte Verbindung an die **TTP** geschickt, wo das **Pseudonymizer**-Modul die Herkunft der Daten validiert.

Pseudonimyze

Im **Pseudonimyze**-Schritt werden **Hashes** in **Pseudonyme** umgesetzt.

Receive & Verify

Pseudonyme und (immer noch) verschlüsselte **Was-Daten** werden über eine gesicherte Verbindung dem **Delivery**-Modul zugeführt, das beim **Abnehmer** der pseudonymisierten Daten steht. Dort wird wiederum die Herkunft der Daten validiert.

Decrypt

Im letzten Schritt wird **Was** vom **Abnehmer** entschlüsselt.

Abnehmer

Verfügt jetzt über pseudonymisierte Daten und kann die entsprechenden Analysen durchführen, ohne über personenbezogene Daten zu verfügen.