The European
DAS DEBATTEN-MAGAZIN

30.06.2016

Neuer Trend aus dem Valley: KEINE Daten sammeln

Datenschutz made in Germany – das war mal. Jetzt kommt Privacy aus dem Silicon Valley. Mit neuen Methoden und Technologien versuchen die Techunternehmen, mehr für die Privatheit ihrer Kunden zu tun. Der aktuelle Privacy Champion, Apple, experimentiert dabei mit neuartigen Forschungsergebnissen.

Fast täglich erreichen uns Nachrichten, die die Hypothese zu bestätigen scheinen, dass man sich im Internet nicht anonym bewegen kann. Weder in den sozialen Netzwerken, die konsequenterweise die Klarnamenpflicht durchsetzen möchten, noch bei der Nutzung der Smartphones, über die systematisch und ununterbrochen Daten über die jeweiligen Nutzer gesammelt werden. So auch die Auswertung öffentlich zugänglicher, aber anonymisierter Daten der New Yorker Taxiunternehmen durch Anthony Tockar von Neustar Research. Er hat diese mit dem – ebenfalls öffentlichen – Wählerverzeichnis verglichen und konnte so prominente Besucher des Hustler Club identifizieren. Samt ihren Adressen.

Ein ähnliches Kunststück ist einer Gruppe von MIT-Wissenschaftlern rund um Yves-Alexandre de Montjoye gelungen, die anonymisierte Kreditkartenabrechnungen einer Queranalyse unterzogen haben. Ausgestattet mit lediglich einer Reihe von Zahlenreihen aus drei Monatsabrechnungen, ohne Namen der Halter, Kreditkartennummern, jegliche Informationen, die man als personenbezogene Daten bezeichnen würde, konnten die Wissenschaftler durch geschicktes Matching der Transaktionsdaten mit öffentlich verfügbaren Informationen die Namen und Adressen der Kunden identifizieren – und ihre gesamte Käufergeschichte rekonstruieren. Nur vier Abrechnungspositionen haben gereicht, um etwa 90 Prozent der Kartenbesitzer eindeutig zu identifizieren. Wie? Oft reichten schon Informationen, wie ein Foto vom gemeinsamen Kaffeetrinken mit Freunden auf Instagram, ein Tweet über ein gerade gekauftes Smartphone in Verbindung mit einer einzigen Zahlung.

Anonymität im Internet unmöglich?

Wie Napoleon schon sagte: das Wort „unmöglich“ gibt es nur im Wörterbuch

von Narren. Dieser Satz könnte auch von Craig Federighi, dem Vice President of Software Engineering bei Apple, stammen, der in seiner Keynote auf der Apple's Worldwide Developers Conference in San Francisco den Einsatz von differential privacy bei Apple ankündigte.

Sehr konkret wurde Federighi allerdings nicht, so berichtet das Wired Magazin von der Konferenz. Er verwies vielmehr nur auf Anfrage auf das Perview Guide für iOS 10. Doch klar ist, dass Apple mit neuartigen Methoden und Technologien seine Position als „Privacy Champion" im Silicon Valley verteidigen und weiter ausbauen möchte. Zuerst mit der End-to-end-Verschlüsselung für iMessage und Facetime – und nun mit differential privacy. Zusätzlich zu dem bereits von Tim Cook angekündigten Bestreben des Unternehmens, so viele Nutzerdaten wie möglich auf den Smartphones zu belassen und so wenige wie möglich zentral auf den Apple-Servern zu speichern.

Mission: ANON

Hinter der Ankündigung der differential privacy steckt allerdings auch der Hinweis, dass Apple die Nutzerdaten dennoch speichert. Denn die Methode sollte eine Auswertung von großen Datenmengen ermöglichen, ohne dass dabei Rückschlüsse auf die einzelnen Individuen gezogen werden können. Bei dieser Methode – wobei sie aufgrund komplizierter mathematischer Grundlagen noch nicht für den breiten Einsatz operationalisierbar ist – wird den Datensätzen eine Art Rauschen hinzugefügt, das diese verfremdet, aber nicht das Ergebnis der statistischen Auswertung beeinflusst. In seiner einfachsten Form kann differential privacy mithilfe von Hashing (Quersummen) umgesetzt werden. Die Pionierin der Methode ist Cynthia Dwork, die für Microsoft Research tätig ist.

Der britische The Economist wertete im August 2015 im Artikel Data Privacy. We'll see you, anon aktuelle – technische, theoretische und formelle – Methoden und Ideen aus, die die Anonymität im Internet wiederherstellen helfen könnten. Eine dieser Methoden ist „homomorphic encryption", also homomorphe Verschlüsselung, eine Art Heiliger Graal der Kryptografie, bei der die Datenbankabfragen verschlüsselt und der Analyst bzw. Auswertungsalgorithmus nie die Originaldaten zu sehen bekommt. Die Idee der anderen Methode, der „secure mulitparty computation", erinnert an das System der Bibelübersetzung, bei welchem der Datensatz zerstückelt und an verschiedene Stellen (bspw. Datenbanken) verteilt wird. Niemand hat daher Zugang zu der gesamten Datenbasis bzw. zum vollständigen Datensatz.

Aus Altem mach Neues

Neben Standards und neuen mathematisch-informatischen Methoden gibt es auch juristische (Aus-)Wege. Eine formelle Lösung könnte laut The Economist wie folgt aussehen: „Data might come with what have been called ‚downstream contractual obligations', outlining what can be done with a given data set and holding any onward recipients to the same standards. One perhaps draconian idea, suggested by Daniel Barth-Jones, an epidemiologist at Columbia University, in New York, is to make it illegal even to attempt re-identification."

Der erste Teil des Vorschlags dürfte einem hierzulande bekannt vorkommen. So steht es sinngemäß (noch) im Bundesdatenschutzgesetz, und zwar in Artikel 31 und Artikel 39. Es handelt sich um den Grundsatz der Zweckbindung bei der Verarbeitung personenbezogener Daten. Neu an dem Vorschlag ist lediglich die individualrechtliche, einzelvertragsbezogene Umsetzungsform.

Wie man Daten nicht sammelt

Während die Unternehmen in Deutschland und Europa immer noch auf das große Geschäft mit Big Data hoffen, das die Ausweichung der EU-Datenschutzverordnung und folglich des Bundesdatenschutzgesetzes erst möglich machen sollte, setzen die Unternehmen aus dem Silicon Valley bereits auf neue Geschäftsmodelle. Neu und ganz hip ist jetzt, wenn die Kundendaten NICHT gesammelt werden.

Mithalten kann dennoch die deutsche Forschung und Wissenschaft. Von der Gesellschaft für Informatik (GI) beispielsweise werden jährlich herausragende Doktorarbeiten ausgezeichnet. Der Träger des GI-Dissertationspreises, der auf der Informatik 2015 verliehen wurde, befasste sich gerade mit den Informatikaspekten von Privacy und Datenschutz. Der Träger des Preises, Dr. Dominik Herrmann, nahm das Domain Name System (DNS) unter die Lupe, das im Internet die Translation von Domainnamen in IP-Adressen übernimmt. Er konnte im DNS Beobachtungsmöglichkeiten nachweisen, die die Privatsphäre von Internetnutzern gefährden. Untersuchungen mit mehr als 12.000 Nutzern belegten, dass die Aktivitäten vieler Internetnutzer ohne deren Wissen über längere Zeit nachvollzogen werden können, anders als bisher angenommen. Herrmann schlug Techniken zum Selbstdatenschutz vor, mit denen sich Nutzer gegen unerwünschte Beobachtung verteidigen können. Seine Erkenntnisse seien nicht nur im DNS-Kontext von Bedeutung, schreibt Herrmann, sie stellten vielmehr die bisherigen Annahmen über die im Internet grundsätzlich erreichbare Privatheit infrage.

Sicher, weil privat

Eine sehr klare praktische Zielsetzung von mehr Sicherheit in der Cloud durch mehr und besseren Datenschutz verfolgt auch ein anderes Projekt: VeriMetrix. Das steht für „Definition und Verifikation von Kennzahlen für den Datenschutz in Cloud-Anwendungen“, erklärt Professor Rainer Böhme von Security and Privacy Lab der Universität Innsbruck. Warum ist das wichtig? „Einerseits lockt das Cloud Computing Unternehmen und Behörden an, nicht zuletzt mit dem Versprechen, die eigenen IT-Kosten zu reduzieren. Andererseits haben viele von ihnen Bedenken, ob in der Cloud ihre Daten geschützt und sicher sind“, sagt Rainer Böhme, „nicht zuletzt müssen Datenschutz und Datensicherheit auch gemäß Regulierungsvorgaben und Gesetzen gewährleistet werden.“ Es bestünde ein großer gesellschaftlicher Bedarf nach einer Lösung, die es den Cloud-Anwendern ermöglicht, eine solche hinsichtlich Sicherheit und Schutz für Daten – a priori und a posteriori – bewerten und beurteilen zu können.

Ein Unternehmen, das seine Daten im Inland lagern möchte und eine entsprechende Cloud-Lösung dafür einkauft, muss beispielsweise die

Möglichkeit haben, zu überprüfen, ob seine Daten sich tatsächlich in einem bestimmten Land befinden. Sprich: „Er muss prüfen können, ob Datenschutzvorgaben während der Nutzung der Cloud eingehalten werden“, bestätigt Rainer Böhme, „das gilt übrigens für den gesamten Lebenszyklus von Cloud-Diensten: von der Erstellung der Datenschutz-Anforderungen bis hin zur überprüfbaren Löschung der Daten nach Nutzungsende der Cloud-Anwendung.“ Das Konsortium um VeriMetrix wird vom Bundesministerium für Bildung und Forschung gefördert.

Konsequent nur, dass GI Themen wie Systemdatenschutz und Selbstdatenschutz in einer Fachgruppe Privacy-Enabling Technologies, kurz PET, aufgreifen möchte. Es gibt dafür sogar einen deutschen Namen: datenschutzfördernde Technik. Ob man ihn wirklich braucht, sei dahingestellt. Die Abnehmer und Interessenten für deutsche Erfindungen und Forschung werden sich vermutlich jenseits der deutschen Grenzen finden. Und das Silicon Valley holt in Sachen Datenschutz – in gewohnter Manier – extrem schnell auf. Datenschutz als Unique Selling Point für Produkte und Services? Das gab es schon mal, auch in Deutschland.

Aleksandra Sowa

Aleksandra Sowa leitete zusammen mit dem deutschen Kryptologen Hans Dobbertin das Horst-Görtz-Institut für Sicherheit in der Informationstechnik. Sie ist Autorin diverser Bücher und Fachpublikationen, Mitglied des legendären Virtuellen Ortsvereins (VOV) der SPD und aktuell für einen Telekommunikationskonzern tätig. In dieser Kolumne äußert sie ihre private Meinung.