DATENSCHUTZ UND BIG TECH : "Ihre Macht hat mich überrascht"

Das Konzept "Datenschutz durch Technik" ist jahrzehntealt. Nach den Snowden-Enthüllungen hob es ab, nun fürchten Kritiker eine feindliche Übernahme durch große Techkonzerne.

15. Februar 2024, 12:02 Uhr
Viele Unternehmen nehmen die Datenfährte der Internetnutzer auf.
Überall im Netz hinterlassen Nutzer eine Datenspur, können leicht verfolgt und ausgespäht werden. Um sie besser zu schützen, gibt es seit langem das Konzept der Privacy by Design, bei der Datenschutz von vornherein in die Technik integriert wird. Umgesetzt wird es durch Privacy Enhancing Technologies (PETs), die einigen als Wundermittel gelten. Daten sammeln und teilen und gleichzeitig deren Vertraulichkeit schützen – das soll damit möglich sein.

Nur: PETs zu entwickeln, ist nicht einfach. Wer aber die Ressourcen dafür hat, sind Big-Tech-Firmen wie Google, Apple und Amazon. Sie haben PETs bereits als Geschäftsmodell entdeckt. Für den Datenschutz heißt das nichts Gutes, sagen Kritiker.

Der Mensch als Summe der von ihm hinterlassenen Datenspuren war anfangs nur in Hackerfilmen wie Das Netz mit Sandra Bullock Realität. Mittlerweile gehört diese einstige Filmfiktion zum Alltag. Ob beim Einkaufen mit Karte, beim Telefonieren mit dem Handy oder beim Surfen im Internet – der User hinterlässt überall eine Datenfährte. "Die Angriffe auf die Privatsphäre kommen von allen Seiten", wusste Helmut Bäumler, der von 1992 bis 2004 Datenschutzbeauftragter von Schleswig-Holstein war, schon vor Jahrzehnten.

Dem Big-Brother-Staat sind viele kleine Geschwister in der Privatwirtschaft nachgewachsen. Daten werden vor allem von Unternehmen gesammelt, die auf "One-to-One-Marketing" mit gezielter Werbung schwören. Ihre Hauptressource ist eine mithilfe von Data Mining durchforstbare Datenbank, die jeden Kontakt mit dem Kunden festhält.

Als Eldorado der Datenschürfer hat sich das Internet erwiesen. Mit jedem Mausklick verraten die Nutzer automatisch, aus welcher Ecke des Cyberspace sie kommen, mit welchem Gerät sie online gehen oder welches "Fortbewegungsmittel" sie verwenden.

Zusätzlich setzen viele Webseitenbetreiber auf Cookies, unscheinbare Browserdateien, die sie den Besuchern auf die Festplatte krümeln. Sie erlauben es Anbietern, alle vom Nutzer auf der Site – und meist darüber hinaus – ausgeführten Operationen festzuhalten und bei Bekanntheit des Kunden – die beim Onlineshopping spätestens bei Angabe einer Lieferadresse gegeben ist – auch persönliche Profile aufzubauen.

Patternz: Profile von über fünf Milliarden Nutzer weltweit

Das System hat sich weiter professionalisiert: Kleine werbefinanzierte Apps wie 9gag, Kik und eine Reihe von Programmen, die eine angezeigte Nummer mit dem potenziellen Namen des Anrufenden verknüpfen (Caller-ID), sollen Teil eines globalen, staatlich-industriellen Überwachungskomplexes sein. Diesem kamen der Wiener Internetforscher Wolfie Christl von Cracked Labs und das Magazin 404 Media auf die Spur.

Die mobile Massenüberwachung startet demnach mit gezielten Anzeigen in Apps, die per Real Time Bidding (RTB) verkauft werden. Dieses Verfahren dient dazu, quasi in Echtzeit etwa über Auktionen personalisierte Banner zu vertreiben. Die zunächst für den Kommerz zu Profilen verdichteten Daten landen dann aber etwa auch in den Händen von Strafverfolgern und Geheimdiensten.

Christl hat zusammen mit Johnny Ryan vom Irish Council for Civil Liberties in einer Studie zu Europas verborgener Sicherheitskrise den Fall Patternz enthüllt. Dabei handelt es sich um ein Werkzeug der Israelischen Sicherheitsakademie (ISA Security), die nach eigenen Angaben als Verein firmiert, nach außen aber wie ein Überwachungsunternehmen auftritt.

Patternz hat die Organisation bis vor kurzem als Lösung beworben, die massive RTB-Daten von Anbietern wie Google und X analysiert und dabei Profile von fünf Milliarden Geräten und ihrer Nutzer erstellt.

Das sind nur wenige Beispiele, die zeigen, wie leicht Handy- und Internetnutzer heimlich verfolgt und ausgespäht werden können. Befürworter des Rechts auf informationelle Selbstbestimmung setzen daher seit mehr als 20 Jahren auf das Gegenkonzept Privacy by Design, wonach der Datenschutz von vornherein direkt in die Technik integriert werden soll.

Diesen Ansatz prägte unter anderem der niederländische Datenschutzbeauftragte John Borking in den 1990ern mit. Bäumler und das von ihm ins Leben gerufene Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein (ULD) veröffentlichten dazu 1998 das Buch Der neue Datenschutz.

Ein wegweisendes Gutachten zu PETs

Die Idee brachten der frühere Berliner Datenschutzbeauftragte Hansjürgen Garstka, der Kasseler Rechtsprofessor und aktuelle hessische Datenschutzbeauftragte Alexander Rossnagel sowie sein Dresdener Informatikkollege Andreas Pfitzmann 2001 in ein Gutachten für das Bundesinnenministerium ein.

Es verschwand zwar rasch in den Schubladen, da es nach den Anschlägen vom 11. September des gleichen Jahres nicht in die Zeit der Anti-Terror-Kataloge des damaligen Ressortchefs Otto Schily (SPD) zu passen schien. Trotzdem erwiesen sich viele Punkte daraus letztlich als prägend für die Debatte rund um den Persönlichkeitsschutz.

Um drei Stichworte kreist das Empfehlungspaket der Wissenschaftler: Systemdatenschutz, Transparenz sowie Datenschutz durch Technik. Den Herausforderungen durch dynamische Technikentwicklung, für den Einzelnen unübersichtliche Strukturen, unbemerkte Datenerhebungen und undurchschaubare Verarbeitungsformen kann den Experten zufolge vor allem durch Systemdatenschutz begegnet werden.

Hinter dem Begriff stecken grundsätzliche Organisationsverfahren. Sie sollen sicherstellen, dass "das technisch-organisatorische System" nur zu der Datenverarbeitung in der Lage ist, zu der es "rechtlich auch ermächtigt" ist.

Dienen sollen einschlägige Verfahren dem Zweck, den Personenbezug der Daten von Anfang an zu vermieden oder auf das absolut notwendige Maß zu begrenzen. "Dabei geht es nicht um

Sparsamkeit im Umgang mit Daten, denn Daten müssen in einer Informationsgesellschaft in breitem Umfang genutzt werden", betonten die Gutachter. Doch es sei oft gar nicht nötig, dass solche Messwerte "einen Personenbezug aufweisen".

In der Platform for Privacy Preferences (P3P), einem Standard des World Wide Web Consortium (W3C), sahen die Forscher eine erste technische Lösung. Voraussetzung für deren Anwendbarkeit sei, dass die datenverarbeitende Stelle ihre Spielregeln in einer allgemein zugänglichen "Privacy Policy" als prinzipiell auch maschinenlesbare Erklärung veröffentliche.

Aus der Forschung in die DSGVO

Der Staat müsse den einzelnen Bürgern und Unternehmen "durch technische Hilfsmittel und durch Infrastrukturleistungen in die Lage versetzen, sich selbst zu schützen", verknüpften die Gutachter diesen Gedanken mit dem Prinzip von Privacy Enhancing Technologies (PETs). Als solche Instrumente kommen ihnen zufolge Mittel zum Inhaltsschutz wie Kryptografie und Steganografie genauso infrage wie die technische geschaffene Anonymität, Pseudonymität und das Identitätsmanagement.

Programme, die Schlüssel, Identitäten und Pseudonyme verwalten und den Nutzer bei der Verwendung von Selbstschutztechniken unterstützen, müssten gefördert und die potenziellen Anwender im Rahmen einer "Bildungsoffensive" auf die Möglichkeiten vorbereitet werden. Für jeden Zweck sollte ein anderes Pseudonym verwendet werden.

Wem solche Grundsätze bekannt vorkommen: Das EU-Parlament griff sie 2007 in einer Entschließung auf, die Kommission brachte sie 2012 in ihren Entwurf für die Datenschutz-Grundverordnung (DSGVO) ein. Über diese sollten PETs & Co. eigentlich längst zum Alltag der Bürger gehören. Doch es gibt viele technische, rechtliche und organisatorische Hürden.

Hinzu kommt, dass Big-Tech-Konzerne wie Apple und Google, gegen deren Datensammelei solche Technologien eigentlich eine gewisse Brandmauer errichten sollten, PETs spätestens seit den Snowden-Enthüllungen 2013 selbst als Geschäftsmodell entdeckt und zu ihren Gunsten umgewandelt haben.

Auch auf Ebene der Industriestaaten ist das vermeintliche Wundermittel angekommen. Die OECD definiert PETs in einem Papier zur digitalen Ökonomie 2023 als "Sammlung digitaler Technologien und Ansätze, die das Sammeln, Verarbeiten, Analysieren und Teilen von Informationen ermöglichen und gleichzeitig die Vertraulichkeit personenbezogener Daten schützen". Insbesondere ermöglichten solche Anwendungen "einen relativ hohen Nutzen von Daten und minimieren gleichzeitig den Bedarf an Datenerfassung und -verarbeitung".

Datenverschleierung durch "Rauschen"

Erst mit den jüngsten Fortschritten in der Konnektivität und Rechenkapazität hätten PETs "zu einer grundlegenden Veränderung in der Art und Weise geführt, wie Daten verarbeitet und geteilt werden können", heißt es bei der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung. Auch wenn diese Entwicklungen noch in den Kinderschuhen steckten, "bergen sie ein enormes Potenzial, die Gesellschaft näher an den fortlaufenden Prozess und die Praxis des 'Privacy by Design' heranzuführen und dadurch das Vertrauen in den Datenaustausch" zu stärken.

PETs werden in vier Kategorien unterteilt. Zu den Verfahren zur Datenverschleierung gehören Zero-Knowledge-Proofs (ZKP), Differential Privacy, synthetische Daten sowie Werkzeuge zur Anonymisierung und Pseudonymisierung. Sie sollen den Schutz der Privatsphäre erhöhen, indem sie die Ausgangsdaten verändern, "Rauschen" ("Noise") beziehungsweise Unschärfe hinzufügen oder identifizierende Details entfernen.

Aggregierte Messwerte lassen sich etwa dank Hintergrundrauschen dann so veröffentlichen, dass der einzelne Nutzer nicht mehr einfach identifizierbar ist. Konzerne wie Apple und Google experimentieren mit solchen Methoden.

Eine entsprechende Verschleierung ermöglicht datenschutzfreundlicheres maschinelles Lernen und die Überprüfung von Informationen etwa im Rahmen einer Altersverifikation, ohne dass eine Offenlegung sensibler Daten erforderlich ist. Wenn die Verfahren nicht sorgfältig implementiert werden, ist es mit dem Schutz aber nicht weit her. Mithilfe von Big-Data-Analysen und ergänzenden Datensätzen können etwa "anonymisierte" Informationen vergleichsweise einfach wieder personalisiert werden.

Berechnungen auf verschlüsselten Daten

Im Bereich Verschlüsselung gelten vor allem homomorphe Verfahren, Mehrparteienberechnung (Multi-Party Computation) einschließlich privater Mengenschnitte sowie vertrauenswürdige Ausführungsumgebungen als PETs.

Ziel ist es, dass Daten während der Verarbeitung verschlüsselt bleiben ("In-Use-Encryption"). In großem Stil kamen solche Ansätze etwa in Covid-Tracing-Apps zur Anwendung. Auch diese Tools haben jedoch Einschränkungen: Sie verursachen etwa hohe Rechenkosten und benötigen leistungsfähige Infrastrukturen.

Föderierte und verteilte Analysen wiederum sollen das Bewältigen von Aufgaben auf Daten ermöglichen, die für die Ausführenden nicht sichtbar oder zugänglich sind. Beim föderierten Lernen etwa werden Informationen an der Datenquelle "vorverarbeitet".

Dabei werden nur zusammenfassende Statistiken und Ergebnisse an diejenigen übertragen, die die Aufgaben ausführen. Föderierte Lernmodelle werden etwa in prädiktiven Textanwendungen auf mobilen Betriebssystemen in großem Maßstab eingesetzt, um zu vermeiden, dass vertrauliche Tastenanschlagsdaten an den Datenkontrolleur zurückgesendet werden. Für den Betrieb von föderierten und verteilten Analysen ist eine ständige, zuverlässige und breitbandige Onlineverbindung erforderlich.

Zu den Instrumenten zur Datenverantwortung gehören rechenschaftspflichtige Systeme, die gemeinsame Nutzung von Geheimnissen unter Schwellenwerten und persönliche Datenspeicher. Sie zielen nicht in erster Linie auf den Schutz der Vertraulichkeit personenbezogener Daten auf technischer Ebene ab und gelten daher häufig nicht als PETs im engeren Sinne.

Auch solche Tools zielen aber darauf ab, die Privatsphäre zu verbessern, indem sie betroffenen Personen die Kontrolle über ihre eigenen Daten ermöglichen und Regeln für den Zugriff auf diese festlegen und durchsetzen. Die meisten einschlägigen Werkzeuge befinden

sich in einem frühen Entwicklungsstadium und verfügen bislang nur über begrenzte Anwendungsfälle.

PIMS als PETs: Die Politik tut sich schwer

Hierzulande startete das Bundesministerium für Digitales im Sommer 2022 mit dem Entwurf für eine Einwilligungsverwaltungs-Verordnung einen Testballon in diese Richtung. Durch die Einbindung anerkannter Dienste soll damit eine "anwenderfreundliche Alternative" zur Verfügung stehen, die Verbraucher "von vielen Einzelentscheidungen entlastet".

Festlegen will das Ressort von Volker Wissing (FDP) etwa die rechtlich-organisatorischen Anforderungen an Personal Information Management Systems (PIMS). Dem Plan zufolge, der bislang aber nicht weit gekommen ist, sollen Anwender "generelle Einwilligungen geordnet nach Kategorien für bestimmte Zugriffe auf Endeinrichtungen und Gruppen von Telemedienanbietern erteilen" können.

PETs werde generell ein großes Potenzial zugeschrieben, schreibt die OECD. Abgesehen von einer "noch begrenzten Anzahl solider und überzeugender" Anwendungsfälle bestehe aber auch Einigkeit darüber, dass der Reifegrad solcher Instrumente immer noch sehr unterschiedlich sei. Multi-Party Computation und Differential Privacy seien mittlerweile zwar weitgehend praxisreif, erklärte Simone Fischer-Hübner, Professorin an der Fakultät für Informatik der Uni Karlstad in Schweden Ende Januar auf einer Konferenz zum Europäischen Datenschutztag in Berlin. Es sei aber immer noch eine große Herausforderung, PETs zu designen, die einfach zu nutzen und zu konfigurieren sind.

Ernüchterung angesichts einer "feindlichen Übernahme" Privatsphäre-schonender Technologien durch Big Tech herrschte Ende Januar auf einer Diskussionsrunde auf dem Privacy Camp der Dachorganisation European Digital Rights (EDRi) in Brüssel. Denn vor allem IT-Giganten hätten die für PETs benötigte infrastrukturelle Leistungsfähigkeit.

Die Rede war von einem "sozialen Dilemma": Private und öffentliche Institutionen müssten sich verstärkt an eine Handvoll mächtiger Technologiekonzerne wenden, um die Datenschutzrechte der Endnutzer zu wahren. Die Anbieter wiederum könnten mit PETs werben, um noch mehr Kunden anzulocken und der mittelständischen Konkurrenz das Leben schwer zu machen. Sie würden so zu "Schiedsrichtern über die Privatsphäre".

Negativbeispiele: Google Privacy Sandbox und Amazon Sidewalk

Als Beispiel nannte Kris Shrishak von der Bürgerrechtsorganisation Irish Council for Civil Liberties (ICCL) die von Google als Ersatz von Cookies von Drittanbietern angepriesene Technik Privacy Sandbox. Auch wenn Browser selbst damit nicht mehr persönliche Daten für Auktionen Plätze für gezielte Werbung in Echtzeit aussendeten, bedeute das ihm zufolge "nicht das Aus für Targeting".

Google habe viel Geld in Versteigerungssysteme, Ad-Server, die Cloud und mobile Betriebssysteme investiert und wolle daher das Geschäftsmodell der "spionierenden Werbung" beibehalten. Für die bei Privacy Sandbox eingesetzten Themen (Topics) spiele der Browser-Verlauf mit angesteuerten Webseiten nach wie vor eine wichtige Rolle.

Dies könne durchaus gefährlich sein, erläutert Shrishak. Wenn das Thema etwa die sexuelle Orientierung von Nutzern sei, könnten Homosexuelle etwa in Saudi-Arabien rasch auffliegen. Individuen könnten zudem nach wie vor identifiziert werden über zusätzliche "Fingerprinting-Methoden". Insgesamt sei die Werbetechnik so "nicht viel besser als die heutigen Cookies in Chrome". Es bleibe dabei, dass Google die Inhalte mit "massiver Rechenpower" serviere und dafür alle gewünschten Informationen über die Nutzer bekomme.

Prinzipiell ähnlich sehe es bei Amazon Sidewalk aus, berichtete Thijmen van Gend, Masterstudent an der TU Delft. Das von dem Online-Handelsriesen entwickelte drahtlose Kommunikationsprotokoll für das Internet der Dinge agiere als Gateway zwischen Geräten wie den Echo-Lautsprechern, Videoklingeln der Amazon-Tochter Ring oder ähnlichen digitalen Helfern auch von Drittanbietern.

Der US-Konzern werde damit zu einer Art Netzwerk-Provider und könne besser in Erfahrung bringen, wie die angeschlossenen Apparate arbeiten und Funktionen anderer Firmen klonen. Über die hauseigene Cloud AWS würden zudem Nutzerdaten ausgetauscht. Dabei setze Amazon zwar auf PETs. Es bleibe aber fraglich, ob der Schutz wirklich funktioniere. Nutzer würden zudem verstärkt abhängig vom gesamten Amazon-Imperium.

Corona-Tracing-Apps als Muster-PETs

Datenschutztechnisch gut funktioniert hätten dagegen in Europa entwickelte Apps zur Corona-Kontaktnachverfolgung, erklärte Carmela Troncoso, leitende Forscherin an der École Polytechnique Fédérale de Lausanne (EPFL): "Die Zweckbindung stand im Mittelpunkt des Designs." Deswegen sei es auch gut, dass die Anwendungen inzwischen "wieder von den Telefonen verschwunden sind im Gegensatz zu den Big-Tech-Apps", sagte Troncoso.

Erstaunt habe sie aber, dass Google und Apple partout die Protokolle direkt in ihre eigenen Lösungen und Betriebssysteme integriert wissen wollten. Zudem hätten sie ein Konzept durchgedrückt, demzufolge keine Daten auf einem Server gespeichert werden durften, sondern nur auf den individuellen Endgeräten.

Die Bundesregierung liebäugelte anfangs mit dem App-Rahmenwerk PEPP-PT (Pan-European Privacy-Preserving Proximity Tracing). Eine Tracing-Anwendung mit einer zentralen Serverlösung, wie sie Entwickler damals verfolgten, hätte mit den Vorgaben von Apple und Google aber nicht kontinuierlich im Hintergrund auf iPhones und Android-Geräten laufen können und wäre folglich kaum praxistauglich gewesen. Die hiesige Exekutive schwenkte daher auf den Kurs der beiden Tech-Giganten um.

Dabei hätte ein zentraler Ansatz Forschern und Gesundheitsämtern mehr Einblicke in das Pandemiegeschehen verschafft, gibt Troncoso zu bedenken. Die beiden Torwächter hätten genau die Definition von Privatsphäre festgelegt und diese auch von einem Tag auf den anderen ändern können.

"Ihre Macht hat mich überrascht", sagt die Schweizerin. App-Entwickler hätten ihr aber erläutert, dass sie sich immer dem Willen von Apple und Google beugen müssten. Glücklicherweise sei das Design der letztlich entwickelten Covid-Apps prinzipiell sehr restriktiv gewesen, so dass es für beide Konzerne nur "limitierte Auswertungsmöglichkeiten" von damit generierten Daten gegeben habe.

Komplexe Technik, schwierige Umsetzung

Auch die schleswig-holsteinische Datenschutzbeauftragte Marit Hansen weiß von der Ambivalenz des Themas. "Gegenüber einem IT-System, in dem alles im Klartext verarbeitet wird, ist der Einsatz von kryptografischen PETs komplexer", erläutert sie Golem.de. "Das merkt man schon, wenn man sein Backup verschlüsselt hat und beim Wiedereinspielen nicht alles gleich reibungslos klappt. Auch für PETs, die mit Trennung von Domänen oder Verteilung von Geheimnissen arbeiteten, besteht eine erhöhte Komplexität, weil dann mehr Parteien an der Verarbeitung personenbezogener Daten beteiligt sind".

"Eine Festplattenverschlüsselung kann ein User selbst durchführen, wenn er über die Software verfügt", verdeutlicht die Datenschützerin. Asymmetrische Verschlüsselung für eine Kommunikation wie bei E-Mail über Programme wie PGP (Pretty Good Privacy) benötige schon zwei Parteien, die jeweils mitwirken. Für die Datenverschleierung über das Netzwerk TOR sei sogar eine ganze Crowd von Usern erforderlich, um den Schutz der Anonymisierung beim Senden von Nachrichten zu ermöglichen.

Die Verfügbarkeit einschlägiger Open-Source-Software wird Hansen zufolge prinzipiell zwar etwa über Code-Depots wie Github gewährleistet. Private Nutzer könnten ohne eigene Programmierumgebung und weitergehendes Know-how damit aber nicht viel anfangen. Infrastruktur-Anbieter machten es ihnen leichter, indem sie ganze "Erklärpakete" zusammen mit den Tools bereitstellten und eine optimale Verwendung über Schnittstellen zu ihren eigenen Systemen wie bei Cloudangeboten versprechen.

PIMS reloaded: Hoffnung auf KI-Assistenten

Das bedeutet der Kontrolleurin zufolge: "Wer sich sowieso schon für die Angebote der globalen Quasi-Monopolisten entschieden hat, bekommt dann auch PET-Module zur eigenen Verwendung oder kann die Funktionen gleich so nutzen, wie sie implementiert sind." Die Abhängigkeit von diesen Anbietern verfestige sich damit. Im Prinzip spreche zwar nichts dagegen, auch souveräne Lösungen in ähnlicher Form bereitzustellen. Aber die großen Plattformbetreiber hätten hier klar die Nase vorn.

Gefördert werden müssten Hansen zufolge daher generell Lösungen, in denen es leicht falle, die Sicherheitsanforderungen und Vorgaben aus der Datenschutz-Grundverordnung (DSGVO) etwa rund um Verantwortlichkeiten, Betroffenenrechte sowie Informations- und Nachweispflichten einzuhalten. Ein im Grundsatz guter Ansatz sei auch die künftige Pflicht zur Interoperabilität von Messengern wie Whatsapp, Signal oder Threema über den Digital Markets Act (DMA).

Aber auch hier stecke der Teufel im Detail: Die Vorgabe könne leicht dazu führen, dass damit Sollbruchstellen für die Schutzanforderungen wie ein Gateway zur Ent- und Wiederverschlüsselung der Daten eingeführt werden. Auch standardisierte maschinenlesbare Datenschutzerklärungen und Einwilligungen seien sinnvoll.

Die IT-Sicherheitsexpertin Fischer-Hübner sieht das ähnlich und will PETs ebenfalls nicht abschreiben. Künftig könnte künstliche Intelligenz (KI) verwendet werden, hofft sie, "um die Privatsphäre besser zu schützen". Ein einschlägiger Assistent wäre etwa imstande, durch die Datenschutzeinstellungen zu führen oder Informationen aus Datenschutzerklärungen auszulesen – angepasst an die individuellen Präferenzen des Nutzers. Eine solche Lösung

müsste aber selbst in einer datenschutzfreundlicher Umgebung implementiert werden. Genauso wichtig sei es, die Kontrolle des Nutzers über die Technik zu gewährleisten.