Privacy-Enhancing Computation

Weniger Risiken, mehr Chancen durch Datenschutz

Wie Unternehmen sensible Daten besser schützen können – und davon wirtschaftlich profitieren.

→ VON THOMAS HAFEN

DER AUTOR

Thomas Hafen
ist seit mehr als 15 Jahren als Journalist tätig. Er schreibt häufig über Cloud-Computing, Digitalisierung, Künstliche Intelligenz, Big Data Analytics und Virtual Reality.

Es ist 1995. Die Europäische Union schafft mit der Richtlinie 95/46/EG eine gesetzliche Grundlage, die EU-weit den Schutz der Privatsphäre sicherstellen und die Verarbeitung personenbezogener Daten regeln soll. Im selben Jahr veröffentlichen die Datenschutzexperten Ronald Hes und John Borking ein Werk, das die technischen Voraussetzungen dafür beschreibt: „Privacy-enhancing technologies – The path to anonymity". Der Zweck von Privacy-Enhancing-Technologien (PET) ist es, Daten von sensiblen Informationen zu reinigen oder sie zu verschlüsseln, sodass keine Rückschlüsse auf eine natürliche Person möglich sind.

Bis vor Kurzem interessierte sich allerdings kaum jemand für deren Einsatz. „Privacy-Enhancing-Technologien sind technisch ausgereift, werden aber nur sehr vereinzelt angewendet", berichtet David Harborth, Lehrstuhl Mobile Business & Multilateral Security an der Goethe-Universität Frankfurt. Doch das ändert sich nun drastisch, wenn das Analystenhaus Gartner recht hat. Seit zwei Jahren listet es in seinen jährlichen Top-Trends unter dem Begriff „Privacy-Enhancing Computation" (PEC) Techniken und Technologien, die den Datenschutz und die Privatsphäre fördern. Bis 2025, so die Gartner-Prognose, werden 60 Prozent aller großen Unternehmen eine oder mehrere PEC-Techniken für Analytik, Business Intelligence oder Cloud-Computing einsetzen, 40 Prozent der Befragten, die am „Gartner Global Security and Risk Management Survey 2021" teilnahmen, zählten PEC zu den Top-Drei-Investitionen für die nächsten zwölf Monate.

Diesen Trend spiegeln auch die Investitionen wider. Laut der Datenbank Crunchbase stiegen die Investitionen in Privacy-Tech-Start-ups in den vergangenen zwei Jahren fast exponentiell auf 4,6 Milliarden Dollar. „Datensparsame und datenschutzfreundliche Lösungen werden nicht nur technologisch, sondern auch gesellschaftlich in den kommenden Jahren eine große Rolle spielen", sagt Marian Gläser, CEO und Co-Founder von Brighter AI und Sprecher der Arbeitsgruppe Datenschutz im KI Bundesverband.

Bild: Shutterstock / PLMT

WAS PEC SO WICHTIG MACHT

Die Renaissance datenschutzfreundlicher Technologien und Konzepte hat auch mit der DSGVO zu tun. Bei allen Unzulänglichkeiten setzte das Regelwerk, das 2016 in Kraft trat und 2018 wirksam wurde, einen Standard für den Datenschutz. Zu den wichtigsten Neuerungen gehört das Marktortprinzip. Es bedeutet, dass alle Unternehmen, die im EU-Markt Geschäfte machen, an die Regelungen gebunden sind, egal ob sich ihr Firmensitz innerhalb oder außerhalb der EU befindet. Bei Verstößen gegen die DSGVO drohen zudem hohe Bußgelder, die bis zu 4 Prozent eines Jahresumsatzes betragen können. Auch das stellt einen wichtigen Unterschied zu den bislang sehr moderaten Sanktionsmöglichkeiten im Datenschutzrecht dar.

Viele Staaten haben Regelwerke nach dem Vorbild der DSGVO erlassen oder ihre Gesetze angepasst, darunter Brasilien, Japan, die Schweiz und der US-Bundesstaat Kalifornien. „Die DSGVO hat eine regulatorische Trendwelle ausgelöst“, konstatiert Marian Gläser von Brighter AI.

Die hohen Bußgeldandrohungen bei Verstößen und die universale Gültigkeit der Datenschutzgesetze in immer mehr Märkten macht Investitionen in datenschutzfreundliche Technologien attraktiv. Gleichzeitig steigen die Gefahren für die Privatsphäre durch KI, Big Data und die →

Bild: Brighter.ai

„Die DSGVO hat eine regulatorische Trendwelle ausgelöst.“

Marian Gläser
CEO und Co-Founder von Brighter AI und Sprecher der Arbeitsgruppe Datenschutz im KI Bundesverband

Allgegenwart von Sensoren im öffentlichen und privaten Raum. „Durch die Digitalisierung entstehen immer mehr Daten", erklärt Gläser, „die vor allem für das Training tiefer neuronaler Netze gebraucht und auch genutzt werden."

Die Industrie sucht daher nach Wegen, wenigstens ein Mindestmaß an Privatsphäre und Datenschutz sicherzustellen, um so noch schärfere Regulierungen zu verhindern, die neue, vielversprechende Geschäftsmodelle gefährden könnten. Datenschutzfreundliche Techniken und Technologien bieten darüber hinaus die Möglichkeit, sensible Daten in unsicheren Umgebungen wie der Public Cloud bearbeiten zu können. Dieses Confidential Computing genannte Konzept eröffnet neue Wege nicht nur für die Analyse von personenbezogenen Daten, sondern auch von Patenten, Finanzzahlen und anderen Geschäftsgeheimnissen.

Wenig Druck kommt dagegen von den Nutzern. Der Wissenschaftler David Harborth forschte im Projekt AN.ON-Next mit Partnern über datenschutzfreundlichere Möglichkei-

„Die Nachfrage nach datenschutzkonformen Lösungen wird massiv steigen"

Research VP Bart B. Willemsen erklärt, warum Gartner Privacy-Enhancing Computation ein enormes Wachstum vorhersagt – und wieso man mit Speichern in Blockchains sehr vorsichtig sein sollte.

com! professional: Herr Willemsen, Gartner verwendet in seinen Dokumenten die Formulierung „Techniken für Privacy-Enhancing Computation (PEC)" statt den gebräuchlicheren Begriff „Privacy-Enhancing Technologies" (PET). Was unterscheidet PEC von PET?
Bart B. Willemsen: Ich bevorzuge den Begriff Privacy-Enhancing Computation, weil er mehr als nur spezifische Technologien umfasst. Das mag zwar wie semantische Haarspalterei klingen, ist aber meiner Meinung nach ein wichtiger Unterschied. So ist beispielsweise das Trusted-Third-Party-Konzept (TTP) keine Technologie, sondern ein architektonischer Ansatz. Diese Entwicklung hin zu neuen Architekturen, Techniken und Methoden hat sich in den vergangenen zwei Jahren deutlich beschleunigt.

com! professional: Was sind die wichtigsten dieser neuen Ansätze?
Willemsen: Neben TTP gehören dazu Protokolle für verteiltes Rechnen (Secure Multi-Party Computation), Null-Wissen-Beweise (Zero-Knowledge Proof), homomorphe Verschlüsselung und Confidential Computing. Auch einige der altbekannten PET wie synthetische Daten oder Differential Privacy haben wieder an Bedeutung gewonnen.

com! professional: Laut Ihrer Prognose werden 60 Prozent der größeren Unternehmen bis 2025 mindestens eine Form von Privacy-Enhancing Computation implementiert haben. Was sind die Hauptursachen für dieses Wachstum?
Willemsen: Einer der Hauptgründe ist die zunehmende Verbreitung von Datenschutzgesetzen. Vor zwei Jahren waren vielleicht 20 Prozent der Weltbevölkerung davon betroffen, im kommenden Jahr werden es rund 75 Prozent sein. Allein das Personal Information Protection Law (PIPL) Chinas stellt die Daten von 1,4 Milliarden Menschen unter Schutz, mit dem Inkrafttreten der Privacy Data Protection Bill (PDPB) in Indien kommen noch einmal 1,4 Milliarden hinzu. Damit fallen riesige Datenmengen unter gesetzliche Vorgaben, die Nachfrage nach datenschutzkonformen Lösungen wird massiv steigen.

com! professional: Wo liegen aktuell die wichtigsten Einsatzszenarien für PEC?
Willemsen: Analytik und Business Intelligence sind immer noch ganz klar die Nummer eins. Dabei muss man interne und externe Szenarien unterscheiden. Innerhalb einer Organisation hat es wenig Sinn, Daten für interne Analysen zu pseudonymisieren oder personenbezogene Daten zu verschlüsseln. Ein Mitarbeiter braucht nur mit den Metadaten über den Flur zu gehen und einen Kollegen im Vertrieb oder Kundensupport zu bitten, die Person dahinter zu identifizieren. Hier sind Methoden wie Differential Privacy oder synthetische Daten zielführender.

com! professional: Und bei externen Analysen über Firmengrenzen hinweg?
Willemsen: Da gibt es verschiedene Ansätze, etwa Daten vor der Weitergabe zu bereinigen, homomorph zu verschlüsseln oder TTP-Umgebungen zu benutzen. Auch Zero-Knowledge Proofs (ZKP) können eine wichtige Rolle spielen. Sie ermöglichen es, das Vorhandensein und die Korrektheit einer bestimmten Information zu verifizieren, ohne diese

> „Je mehr Daten eine Organisation sammelt und an einem Ort zusammenführt, desto größer ist das Risiko."

übertragen zu müssen. Die Anwendungsmöglichkeiten sind vielfältig und reichen von der Altersüberprüfung und Nutzeridentifikation bis hin zur Bekämpfung von Betrug und Geldwäsche im Bankenumfeld oder dem Zugriff auf Gesundheitsdaten.

com! professional: Welche weiteren Use-Cases sind besonders bemerkenswert?
Willemsen: Der zweite wichtige Anwendungsbereich ist das Training von KI-Modellen. Je wichtiger Machine Learning und anderen Methoden der Künstlichen Intelligenz werden, desto mehr steigt die Nachfrage nach Trainingsdaten. Oft sind die für das Lernen notwendigen Informationen im eigenen Unternehmen nicht vorhanden oder unvollständig. Daher gewinnen Methoden wie föderales Lernen und synthetische Daten an Bedeutung. Der dritte Treiber ist die zunehmende Ein-

ten der Internetnutzung wie VPN oder Onion-Router wie TOR. Seine Erfahrung: „Nutzer wollen für ihre Privatsphäre kein Geld ausgeben, selbst Cent-Beträge sind schon zu viel."

DIE VIER PRINZIPIEN VON PEC

Bei Privacy-Enhancing Computation gibt es vier Möglichkeiten, sensible Daten zu schützen: verschlüsseln, verschleiern (Obfuscation), entfernen oder erst gar nicht erheben. Bei der homomorphen Verschlüsselung (Homomorphic Encryption, HE) etwa bleiben die Eigenschaften der Basisdaten trotz Verschlüsselung erhalten. Laut Gartner wird es aber noch fünf bis zehn Jahre dauern, bis es zu einem großflächigen produktiven Einsatz von HE kommt. Einige Provider bieten jedoch bereits Services an (siehe Tabelle auf Seite 32). Berichten zufolge experimentiert auch Facebook/Meta mit HE. Ziel ist es, individuelle Werbung in Whatsapp einblenden zu können, ohne die verschlüsselte Kommunikation aufbrechen zu müssen. ➔

Bart B. Willemsen
Research Vice President bei Gartner

Bild: Gartner

schränkung des grenzüberschreitenden Datentransfers. Das betrifft nicht nur die Übermittlung zwischen der Europäischen Union und den USA, sondern zunehmend auch andere Regionen. So beschränken beispielsweise neue Gesetze in China wie das Cybersicherheitsgesetz (Cyber Security Law, CSL), das Datensicherheitsgesetz (Data Security Law, DSL) oder das bereits erwähnte Datenschutzgesetz PIPL den Datentransfer. In Indien, Russland und dem Mittleren Osten gibt es ähnliche Bestrebungen.

com! professional: Was können Unternehmen tun, um trotz dieser Einschränkungen transnationale Bereitstellungsmodelle wie Public Cloud weiter nutzen zu können?

Willemsen: Die Schlüsseltechnologie, die sich vor allem für Infrastructure as a Service (IaaS) eignet, nennt sich Confidential Computing. Sie ermöglicht es, Daten sicher verschlüsselt zu speichern und zu übertragen und Anwendungen in einem geschützten Bereich des Rechenkerns, der sogenannten Enklave, auszuführen.

com! professional: Welche Rolle spielen PEC-Techniken für die Digitalisierung in der öffentlichen Verwaltung, etwa bei der elektronischen Patientenakte?

Willemsen: Je mehr Daten eine Organisation sammelt und an einem Ort zusammenführt, desto größer ist das Risiko. Handelt es sich dabei um so sensible Informationen wie Arztberichte, Diagnosen und andere Patientendaten, potenziert sich die Gefahr. Wenn diese sensiblen Daten von einer KI ausgewertet werden, um auf dieser Basis individuelle Entscheidungen zu treffen, dann ist das wahrscheinlich das größte Risiko, das es im Bezug auf Datenschutz und den Schutz der Privatsphäre gibt. Der Einsatz von PEC ist in einem solchen Fall absolut anzuraten.

com! professional: Welche Bedeutung haben Blockchain-Technologien für die Entwicklung und den Einsatz datenschutzfreundlicher Lösungen?

Willemsen: Es stecken sicher Chancen in diesem Ansatz. So arbeiten Anbieter bereits daran, Zero-Knowledge Proofs in ihre Plattformen zu integrieren. Das Prinzip der Unveränderbarkeit, das einer Blockchain ja zugrunde liegt, stellt jedoch ein großes Problem für den Datenschutz dar. Informationen, die auf ihr gespeichert sind, bleiben dort für immer. Personenbezogene Daten müssen aber gelöscht werden können, wenn bestimmte Fristen ablaufen oder der Betroffene sein Recht auf Löschung wahrnimmt. Werden solche Daten auf einer Blockchain gespeichert, ist es daher praktisch nicht mehr möglich, moderne Datenschutzgesetze einzuhalten. Das Problem verschlimmert sich noch dadurch, dass sich Personen nicht nur über Name, Adresse und Geburtsdatum, sondern auch über Metadaten identifizieren lassen. Man sollte also sehr, sehr vorsichtig sein, was man in einer Blockchain speichert und was nicht.

com! professional: Würde es nicht reichen, personenbezogene Daten zu verschlüsseln?

Willemsen: Das löst das Problem nicht, dass es auf der Blockchain kein Verfallsdatum gibt. Heutige Verschlüsselungstechniken sind spätestens dann leicht zu knacken, wenn Quantencomputer verfügbar werden. Ein solcher Schutz hält also höchstens ein paar Jahre.

com! professional: Welche Strategie empfehlen Sie Unternehmen, die datenschutzfreundliche Technologien in bestehende Strukturen integrieren wollen?

Willemsen: Als Erstes würde ich die Fälle identifizieren, in denen das Risiko für Datenschutzverstöße besonders hoch ist. Hier ist der Nutzen von PEC offensichtlich. Dann würde ich nach Fragestellungen suchen, die sich aktuell nicht beantworten lassen, ohne gegen bestehende Datenschutz- oder Vertraulichkeitsregeln zu verstoßen. Vor allem aber rate ich Organisationen, langfristig zu planen und Datenschutz von Anfang an in ihre Datenstrategie zu integrieren, statt im „Patch-Modus" kurzfristig Sicherheitslöcher zu stopfen.

com! professional: Wann lohnen sich die Investitionen in PEC?

Willemsen: Mit PEC können Sie aus sensiblen Daten neue Erkenntnisse gewinnen, ohne rechtliche Risiken einzugehen oder unethisch zu handeln. Sie können also etwas tun, was vorher undenkbar war. Dafür lässt sich kein ROI berechnen.

Ein Beispiel für Verschleierung ist Data Poisoning. Dieses Verfahren kommt in erster Linie bei der Gesichtserkennung zum Einsatz. Dabei werden in Bildern nur wenige Pixel verändert. Dies genügt jedoch, um Gesichtserkennungssysteme wie Microsoft Azure Face API oder Amazon Rekognition auszutricksen. Das SAND Lab (Security, Algorithms, Networks and Data) der Universität Chicago hat mit Fawkes1 eine Software entwickelt, mit der Anwender ihre öffentlich zugänglichen Bilder verschleiern können.

Bei der Anonymisierung werden sensible Informationen vor der Weitergabe oder Verarbeitung von Datensät-

Lösungen und Services für Privacy-Enhancing Computation (Auswahl)

Anbieter	Produkt	Bereich	Beschreibung
Apheris AI www.apheris.com/platform	Apheris	Secure Multi-Party Computation, Differential Privacy, Homomorphic Encyption	Platform für abteilungs- oder unternehmensübergreifende Data-Science-Workflows
Brighter AI www.brighter.ai/product	Brighter Redact	Deep Natural Anonymization	KI-basierte Anonymisierung von Gesichtern oder Nummernschildern in Fotos und Videos. Lässt sich per REST-API in andere Lösungen integrieren
Duality www.dualitytech.com	Duality Secure Plus	Homomorphic Encryption, Secure Multi-Party Computation	Duality Secure Plus ermöglicht es mehreren Parteien, sensible Daten gemeinsam zu verarbeiten, ohne Datenschutz oder Vertraulichkeit zu gefährden
Evernym www.evernym.com/verity	Verity	Zero Knowledge Proof	Plattform zur Erstellung und zum Austausch von Identitätsnachweisen
Google www.github.com/Google/private-join-and-compute	Private Join and Compute	Homomorphic Encryption, Secure Multi-Party Computation	Open-Source-Projekt, das Secure Multi-Party Computation ermöglicht; zwei Parteien können ihre Daten auf Gemeinsamkeiten untersuchen, ohne dass sie sensible Informationen über die Identität der verwendeten Datensätze austauschen müssen
IBM www.ibm.com/de-de/security/services/homomorphic-encryption	Homomorphe Verschlüsselungsservices	Homomorphic Encryption	Ermöglicht es dank vollhomomorpher Verschlüsselung, sensible Daten in Cloud-Umgebungen zu analysieren, ohne dass der Provider Einblick in die Informationen erhält
Intel www.intel.de/content/www/de/de/architecture-and-technology/software-guard-extensions.html	Software Guard Extensions (SGX)	Confidential Computing	SGX bildet die technische Grundlage für Confidential Computing, indem es die Ausführung von Applikationen in einem geschützten Bereich des Prozessors (Enklave) ermöglicht
Microsoft https://azure.microsoft.com/de-de/solutions/confidential-compute	Azure Confidential Computing	Confidential Computing	Public-Cloud-Umgebung für die Verarbeitung vertraulicher und personenbezogener Daten
Microsoft www.microsoft.com/en-us/research/project/microsoft-seal	Microsoft SEAL (Simple Encrypted Arithmetic Library)	Homomorphic Encryption	Open-Source-Bibliothek, mit der Entwickler homomorphe Verschlüsselung in ihre Software integrieren und so durchgängig verschlüsselte Services entwickeln können
Mostly AI www.mostly.ai	MOSTLY AI	Synthetic Data	Plattform, mit der sich vollautomatisch synthetische Daten für unterschiedliche Anwendungsfelder erstellen lassen
Secustack www.secustack.com/index.de.html	SecuStack	Confidential Computing	Cloud-Betriebssystem, das eine sichere und geschützte Rechenumgebung für Confidential Computing auf Public-Cloud-Infrastrukturdiensten (Infrastructure as a Service) ermöglicht
Statice www.statice.ai/product/synthetic-data-software	Synthetic Data Software	Synthetic Data	Tool zur Generierung synthetischer Daten; kann als webbasierte Plattform eingesetzt oder per SDK als Library genutzt werden
Xayn AG www.xaynet.dev	XayNet	Federated Learning	Open-Source-Framework für föderales Lernen

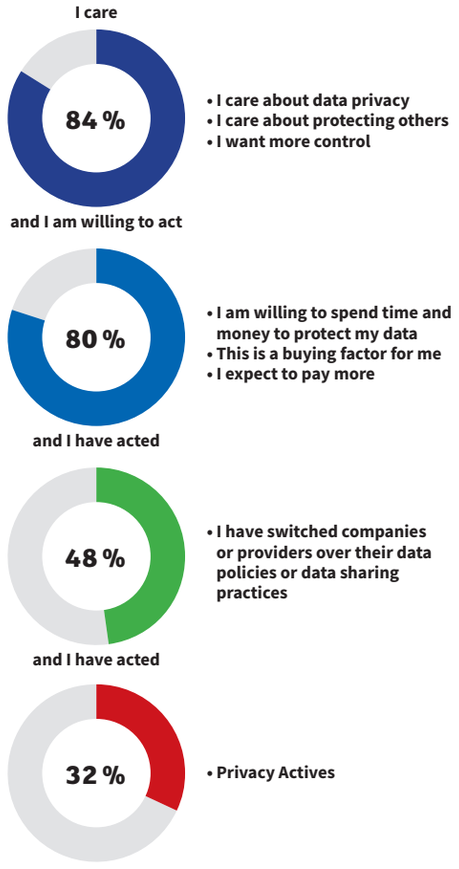
com! professional 4/2022

zen gelöscht. Die Systeme erkennen die zu entfernenden Daten entweder anhand vorgegebener Templates oder lernen selbstständig (Deep Natural Anonymization). So bietet der deutsche KI-Spezialist Brighter AI mit Brighter Redact ein System, das eine KI-basierte Anonymisierung von Bildern und Videos ermöglicht. Die Lösung kommt unter anderem in einem Pilotprojekt der Deutschen Bahn zum Einsatz, wo sie für die Auslastungsmessung in Zügen und S-Bahnen verwendet wird.

Erst gar nicht übertragen werden Informationen in Konzepten wie Federated Learning. Dabei lädt sich ein Teilnehmer beispielsweise einen Algorithmus auf sein Gerät, der dort mit den persönlichen Daten trainiert wird. Der Anwender spielt nur den trainierten Algorithmus zurück, personenbezogene Daten werden nicht übertragen. Google nutzt den Ansatz bereits, um die Suchvorschläge auf Android-Geräten zu optimieren.

Privacy-Paradox

In Umfragen schreiben Nutzer der Privatsphäre einen hohen Stellenwert zu. Tatsächlich wollen jedoch nur die wenigsten dafür bezahlen.



com! professional 4/2022
Quelle: „Cisco Consumer Privacy Study 2019", n = 2601

Bild: Goethe-Universität Frankfurt

„Wir müssen Menschen ermächtigen, sich gegen den Missbrauch ihrer Daten zu schützen."

Dr. David Harborth
Lehrstuhl Mobile Business & Multilateral Security an der Goethe-Universität Frankfurt

PEC IN GESUNDHEIT UND VERKEHR

Große Bedeutung haben datenschutzfreundliche Technologien besonders im Gesundheitswesen und im Verkehrssektor. Es braucht große Mengen Patientendaten, um neuronale Netze für die Krebsdiagnostik oder die Entwicklung von Medikamenten zu trainieren. Die Risiken, dass sie in falsche Hände gelangen, sind hoch. Das Start-up Curai, ein Anbieter KI-basierter Medizin-Apps, nutzt deshalb synthetische Daten, um seine Systeme ohne sensible Patienteninfos trainieren zu können. Dafür wurden 400.000 Krankenakten simuliert und mit einem gefalteten neuronalen Netzwerk (Convolutional Neural Network, CNN) verarbeitet. Das so trainierte System soll zur Diagnostik eingesetzt werden und Patienten remote beraten können.

Das deutsche Start-up Ebenbuild setzt in einem Forschungsprojekt, das mit dem Münchner Klinikum Rechts der Isar durchgeführt wird, auf Confidential Computing. Ziel ist es, einen digitalen Zwilling der menschlichen Lunge zu entwickeln, um Beatmungstherapien individueller gestalten und die mit der Beatmung verbundenen Schäden minimieren zu können. Die Verarbeitung erfolgt in der Public Cloud, die Informationen sind aber durchgehend verschlüsselt. Berechnungen werden in einem besonders geschützten Bereich des Prozessors, der Enklave, ausgeführt.

Im Straßenverkehr ist es vor allem der Trend zum autonomen Fahren, der die Analyse großer Datenmengen erfordert. Schon heute sind moderne Fahrzeuge mit einer Vielzahl von Sensoren und Kameras ausgestattet, die große Mengen an sensiblen Daten erfassen. Wegen dieser Sammelwut erhielt der E-Mobilhersteller Tesla 2020 den Big Brother Award, einen Preis, der jährlich für eklatante Datenschutzverstöße vergeben wird. Besonders sauer stieß den Juroren auf, dass Tesla nicht nur permanent Daten im Fahrzeug, sondern auch in dessen Umgebung erfasst. „Wenn Menschen gefilmt und aufgezeichnet werden, die nur an einem Auto vorbeigehen, ohne dass sie sich konkret verdächtig machen, ist dies klassische illegale Vorratsdatenspeicherung", sagte Datenschutzexperte Thilo Weichert anlässlich der Verleihung.

Eine Bereinigung oder Verschleierung der von einem Fahrzeug erfassten Daten könnte das Problem lösen. →

So nutzt der französische Autozulieferer Valeo Deep Natural Anonymization, um personenbezogene Informationen aus Bildern zu entfernen, die mit Fisheye-Fahrzeugkameras aufgenommen wurden. Marian Gläser sieht in diesem Weg eine Chance für die europäische Autoindustrie: „Wenn es gelingt, sensible Daten in Fahrzeugen zu verarbeiten, ohne dabei gegen Datenschutzvorschriften zu verstoßen, dann ist das ein echter Wettbewerbsvorteil."

FAZIT & AUSBLICK

Nach gut 20 Jahren Dornröschenschlaf erleben datenschutzfreundliche Technologien einen regelrechten Boom. Aus guten Gründen. Zum einen stellen personenbezogene und andere sensible Daten ein hohes Risiko für Organisationen dar. Je mehr davon gesammelt und gespeichert werden, desto größer ist die Gefahr, Opfer eines teuren, reputationsschädigenden Datendiebstahls zu werden. Zum anderen lassen sich vielversprechende Analyseansätze und Geschäftsmodelle nicht oder zumindest nicht rechtskonform umsetzen, wenn die verwendeten Daten personenbezogene Informationen enthalten. Architekturen und Technologien wie föderales Lernen, sichere Mehrparteiensysteme oder Confidential Computing erschließen hier Möglichkeiten, die ohne PEC nicht möglich wären.

Ob datenschutzfreundliche Technologien allerdings auch zu mehr Privatsphäre für private Anwender führen, ist fraglich. „Selbst wenn wir vieles in die Infrastruktur verlagern können, werden wir der technischen Entwicklung noch jahrelang hinterherrennen", fürchtet David Harborth von der Uni Frankfurt. Er sieht den Hebel daher eher auf der Nutzerseite: „Wir müssen Menschen ermächtigen, sich gegen den Missbrauch ihrer Daten zu schützen." •

Die wichtigsten Lösungen für Privacy-Enhancing Computation

Diese Verfahren sollten Unternehmen kennen, die auf Datenschutz Wert legen und zugleich Daten gewinnbringend verwerten wollen:

- **Confidential Computing:** Confidential Computing ermöglicht es, sichere, datenschutzkonforme Berechnungen in öffentlichen Ressourcen wie der Public Cloud durchzuführen. Code und Daten werden dazu in einen speziellen Bereich des Prozessors geladen. Dieses Trusted Execution Environment (TEE) oder Enklave genannte Segment ist von der Umgebung abgeschottet und kann nur von autorisierten Anwendungen genutzt werden. Das stellt nicht nur den Schutz der Daten sicher, sondern auch die Integrität des für die Bearbeitung genutzten Codes.
- **Data Perturbation/Poisoning:** Dabei werden Bilder auf Pixelebene so subtil verfälscht, dass die Veränderungen einem menschlichen Betrachter nicht auffallen, KI-basierte Systeme aber verwirrt werden und etwa Personen nicht mehr identifizieren können.
- **Data Sanitization:** Verfahren, die Informationen vollständig und unwiederbringlich von Datenträgern löschen oder aus Systemen entfernen.
- **Deep Natural Anonymization:** Erkennt personenbezogene Daten wie Gesichter oder Nummernschilder in Fotos und Bewegtbildern und ersetzt sie irreversibel durch synthetische Daten, die keine Rückschlüsse mehr auf die Person zulassen.
- **Differential Privacy:** Dieses Verfahren erlaubt möglichst genaue Berechnungen auf einem Datenbestand mit personenbezogenen Informationen, ohne dass einzelne Personen identifizierbar sind. Dazu wird den Daten kontrolliert Rauschen hinzugefügt, das Rückschlüsse auf Einzelne erschwert, die relativen Verhältnisse der Informationen aber nicht nennenswert beeinträchtigt.
- **Dynamic Data Masking (DDM):** Bei DDM werden sensible Informationen in Echtzeit aus Datenabfragen entfernt, bleiben aber in den gespeicherten Daten erhalten.
- **Federated Machine Learning:** Beim föderalen Lernen erfolgt das Training von Algorithmen verteilt auf lokalen Knoten. Das können zum Beispiel Smartphones, autonome Fahrzeuge oder IoT-Geräte sein. Die Daten bleiben auf den jeweiligen Geräten, was den Datenschutz erhöht.
- **Homomorphic Encryption (HE):** Die homomorphe Verschlüsselung bewahrt die Eigenschaften von Daten in einer Weise, dass darauf weiterhin Berechnungen durchgeführt werden können. Bei einer vollständigen Homomorphic Encryption gibt es keine Einschränkungen für die Operationen, partielle HE erlaubt dagegen nur bestimmte Berechnungen, beispielsweise Multiplikation oder Addition.
- **Privacy by Design:** Dieser Ansatz hat das Ziel, Datenschutz von Anfang an in die Entwicklung von Applikationen, Prozessen oder Plattformen einzubeziehen. Die Erfassung und Weiterverbreitung personenbezogener Daten soll auf das absolut notwendige Minimum begrenzt werden.
- **Secure Multi-Party Computation:** Kryptografische Methoden, mit denen sich Daten von mehreren Parteien gemeinsam bearbeiten und analysieren lassen. Sensible Informationen werden dabei so verschlüsselt, dass sie von den jeweils anderen Beteiligten nicht eingesehen werden können.
- **Synthetic Data:** Synthetische Daten beruhen nicht auf Messungen oder Datenerhebungen, sondern werden künstlich erzeugt. Dabei kommen meist statistische Modelle, regelbasierte Ansätze oder Deep-Learning-Algorithmen zum Einsatz.
- **Zero-Knowledge Proofs (ZKP):** Mit sogenannten Null-Wissen-Beweisen können Unternehmen oder Einzelpersonen nachweisen, dass bestimmte Informationen korrekt sind, ohne diese direkt preisgeben zu müssen. ZKP können beispielsweise für die Altersverifikation oder zum Nachweis einer Fahrerlaubnis eingesetzt werden.

Quellen: Gartner „Hype Cycle for Privacy", 2021; KI Bundesverband „State-of-the-Art Report Privacy Tech", 2021 (ergänzt, verändert)