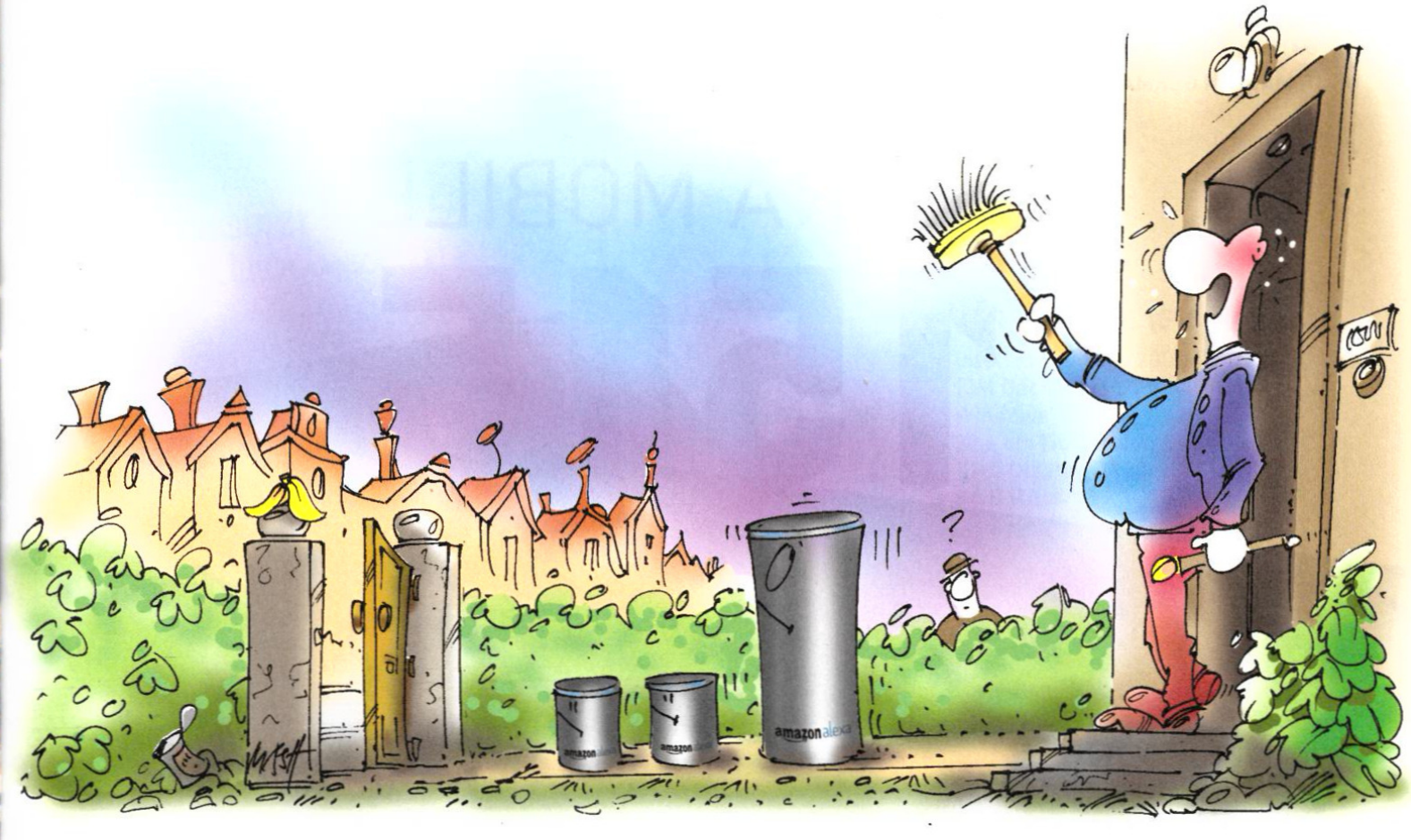


Die Wohnung entwanzen

"Ernsthaft? Ihr als Datenschützer nutzt Alexa? Wie kann man sich denn freiwillig so eine Abhörvorrichtung ins Haus stellen?" Ja, als Jurist bei Heise muss man sich von den Redakteuren einiges anhören. Und zugegeben, ständig von einer kleinen Dose belauscht zu werden, die Sprachdaten in US-Clouds funkt und dort bis zum St. Nimmerleinstag speichert, macht ein mulmiges Gefühl.

Doch das ignoriert der innere Zwölfjährige, der jeden Abend wieder auf und ab hüpft, wenn man die magischen Wort "Alexa, schalte das Licht im Wohnzimmer an!" sagen darf. Dazu noch das Lächeln der weniger technikaffinen Familienmitglieder, die nun nicht mehr mit vier Fernbedienungen jonglieren müssen, wenn sie Musik hören wollen, - oder die zum x-ten Mal den Pups-Generator aktivieren.

Vertrauliche Daten aus dem privatesten Umfeld einem Unternehmen anzuvertrauen funktioniert aber nur, wenn man darauf vertrauen kann, dass sie dort sicher verwahrt werden. Dieses Vertrauen hat Amazon durch eine prekäre Datenpanne dauerhaft zerstört, die Kollegen von c't aufgedeckt haben (siehe Seite 16): Tausende Aufnahmen eines Alexa-Nutzers sind in fremde Hände geraten.

Wer derart sensible Daten speichert, muss für ein entsprechend hohes Schutzniveau sorgen. Egal, was Amazon technisch und organisatorisch dazu unternahm: Es hat ganz offensichtlich versagt. Und warum muss Amazon die Sprachaufzeichnungen nach dem Anlernen von Alexa überhaupt noch speichern? Damit US-Ermittler sie anfordern können, wie jüngst geschehen? Solche Sammlungen bergen doch immer das Risiko einer Datenpanne. Die Amazon auch nicht zu überraschen scheint: "Ups, schade, ... Einzelfall?"

Den letzten Rest Vertrauen, auch bei den inneren Zwölfjährigen, hat Amazon mit seiner sehr schmallippigen Reaktion gegenüber c't und den Betroffenen der Datenpanne verspielt. Unsere persönliche Konsequenz lautet: "Alexa, pack deine Sachen und geh! Und lass bloß unsere Daten hier!"

Joerg Heidrich

Joerg Heidrich

Nicolas Maekeler

Nicolas Maekeler

Alexa, wer hat meine Daten?

Amazon gibt intime Sprachdateien preis

Irgendwann fallen alle Bedenken und Alexa steht auch im Schlafzimmer oder im Bad. Was soll schlimm daran sein, mit Amazons Sprachassistent auch im intimsten Bereich zu plaudern? Schlimm wird es dann, wenn all die Tondateien durch Schluderei von Amazon an fremde Personen geschickt werden. Nun ist Amazon.de genau das passiert, was Kritiker immer befürchteten.

Von Holger Bleich

Der Absatz von Sprachassistenten für die eigenen vier Wände boomt. Obwohl Amazon keine Absatzzahlen herausgibt, gilt als sicher: Echo-Lautsprecher des Konzerns haben sich in zig Millionen Haushalten eingenistet. Dieses Weihnachten werden wieder Hundertausende dazukommen.

Kritik von Datenschützern wischt Amazon mit den immer gleichen Argumenten beiseite. Der Nutzer wisse stets, was sein Echo wann aufzeichnet. Nur Amazon und er können laut Datenschutzerklärung die Aufzeichnungen abhören. Dieses Versprechen ist es, auf dem der große Vertrauensvorschuss der Alexa-Nutzer gegenüber dem Konzern fußt.

Und Amazon wäre von allen guten Geistern verlassen, wenn man dem Schutz dieser teils sehr intimen Sprachdaten aus den Wohn- und Schlafzimmern der Republik nicht größte Priorität beimessen würde. Doch ein Fall, den c't aufgearbeitet hat, legt nahe, dass es fatale Lücken in den Verarbeitungsprozessen der Alexa-Daten bei Amazon.de gibt. Es ist genau der Präzedenzfall eingetreten, vor dem Daten- und Verbraucherschützer so eindringlich warnen.

Vergiftete Selbstauskunft

Alles begann mit dem Recht auf Selbstauskunft aus der neuen EU-Datenschutz-Grundverordnung (DSGVO). Die Verordnung gewährt „betroffenen Personen" einen sehr weitgehenden Auskunftsanspruch über die gespeicherten Daten. Ende August dieses Jahres hat Amazon.de-Kunde Martin Schneider (Name von der Redaktion geändert) auf dieser Basis die Münchener Niederlassung des Online-Riesen um Auskunft gebeten. Rund zwei Monate später erhielt er einen Download-Link, unter dem Amazon die ihm zugeordneten Daten in einer rund 100 MByte großen ZIP-Datei zur Verfügung stellte.

Rund 50 der Dateien enthielten auf seine Person bezogene Daten, beispielsweise getrackte Suchverläufe auf der Amazon.de-Website. Allerdings fand Schneider auch cirka 1700 WAV-Dateien sowie eine PDF-Datei vor, die offensichtlich chronologisch unsortierte Transskripte darüber enthielt, was Amazons Sprachassistent Alexa aus Spracheingaben verstanden hat. „Darüber war ich sehr verwundert, da ich Amazon Alexa nicht verwende, geschweige denn ein Alexa-fähiges Gerät besitze. Deshalb hörte ich mir stichprobenartig einige dieser Audiodateien an und konnte die Stimmen keiner mir bekannten Person zuordnen", erzählte er uns später.

Am 8. November wandte sich Schneider an die Mailadresse des Amazon.de-Kundenservice. Er teilte mit, dass die ZIP-Datei offensichtlich Alexa-Sprachaufzeichnungen eines ihm fremden Amazon-Kontos enthielt und bat um Auskunft dazu. Eine Antwort auf diese Mail erhielt er nie, allerdings führte kurze Zeit später der Download-Link ins Leere. Die Datei ließ Amazon.de folglich verschwinden, doch Schneider hatte sie vorher gespeichert. Mitte November wandte er sich an die c't-Redaktion, weil ihn dieser Vorfall sehr besorgte und Amazon.de auf stumm geschaltet hatte: „Meiner Meinung nach sollten Kunden darüber informiert werden, dass möglicherweise fremde Personen Audiodateien von ihnen erhalten."

Verletzte Intimsphäre

Wir erbaten vertraulichen Zugriff auf die Tondateien, um den Fall einzuschätzen. Sie ergeben ein recht detailliertes Bild des Fremden und seines persönlichen Umfelds. Aus den Aufzeichnungen geht hervor, dass der unbekannte Amazon-Kunde

Alexa an mehreren Orten nutzt: Mindestens ein Echo steht in den Wohnräumen, am Fernseher hängt eine Fire-TV-Box, die via Amazon-Sprachfernbedienung gesteuert wird. Oft erhält Alexa Kommandos von einer Frauenstimme, es befindet sich also auch eine weibliche Person zumindest zeitweise im Haushalt.

Selbst unter der Dusche war Alexa hörbar dabei. An ein Thermostat gerichtete Kommandos legen nahe, dass Alexa auch mit Smart-Home-Gerätschaft verbunden ist. Überdies nutzt der Kunde Alexa auch auf dem Smartphone, zuhause und unterwegs. Die Aufzeichnungen decken wahrscheinlich den kompletten Monat Mai ab.

Das sehr unangenehme Gefühl, etwas Voyeuristisches, Unmoralisches zu tun, ließ uns beim Abhören die Nackenhaare sträuben: Unvermittelt fanden wir uns in der intimen Sphäre uns fremder Personen ohne deren Wissen wieder. Und die an Alexa gestellten Fragen verrieten viel über die Personen: Wecker-Kommandos oder Fahrplan-Abfragen etwa ließen konkret auf Job-Situation und Lebensgewohnheiten schließen, Lautstärke-, Radio- und Spotify-Steuerung auf die Musikgeschmäcker.

Es war nicht schwer, aus den Aufzeichnungen den Alexa-Nutzer zu indentifizieren. Wetter-Abfragen zu konkreten Orten, die Nennung von Vornamen und einmal sogar auch eines Nachnamens engte den Personenkreis schnell ein. Es dauerte nicht lange, bis wir uns anhand der Audiodateien und Transskripte nahezu sicher waren, den fremden Amazon-Kunden sowie die weibliche Person identifiziert zu haben. Informationen aus von uns durch weitere Recherche zugeordneten Facebook- und Twitter-Profilen passten und ergänzten das Bild.

Wir beschlossen, dem mutmaßlich Betroffenen via Twitter-Direktnachricht eine Bitte um Anruf zukommen zu lassen – eine Telefonnummer haben wir nicht herausgefunden. Er rief sofort an und wir erzählten, wie wir auf ihn gestoßen sind. Wir waren an der richtigen Adresse. Nils Schmidt (Name von der Redaktion geändert) schluckte hörbar, als wir ihm berichteten, welche seiner höchst privaten Daten Amazon.de an einen Fremden weitergereicht hatte. In seinem Kopf begann es zu rattern: Was haben wir Alexa alles gefragt? Welche Geheimnisse über mich sind geleakt? Er bestätigte überdies, dass wir auch bei der Person hinter der weiblichen Stimme korrekt recherchiert hatten.

Aus dem Transskript lassen sich jede Menge privater Informationen zum Betroffenen herauslesen.

```
2018-05-05T14:07:44.980Z,alexa alexa wie steht es gerade fÃ¼r den h. s. v.,b18e27146ad2374
2018-05-05T11:50:58.911Z,alexa stelle die schlafzimmer temperatur auf ein und zwanzig grad,
2018-05-05T19:31:26.171Z,alexa lass uns ein spiel spielen,1f710fb1ce3ccba6962f68791e2849c
2018-05-05T12:04:43.804Z,alexa starte starte shot oder spott,4fe7842b542614d8a4c2b78c95
2018-05-05T19:39:13.406Z,dinosaurier,6c4155896ee3e0a8b82f970c42122d070d2950b6.wav
2018-05-17T13:13:00.209Z,ja,af397b2f502d36daa909d0c8eb8725c18aa4dde7.wav
2018-05-15T13:18:14.800Z,alexa stell den timer auf fÃ¼nf und dreiÃŸig minuten,d73063b82ac
2018-05-05T11:19:32.700Z,nein,f132b16ef523ad933a2f938018b8ce32126d9f94.wav
```

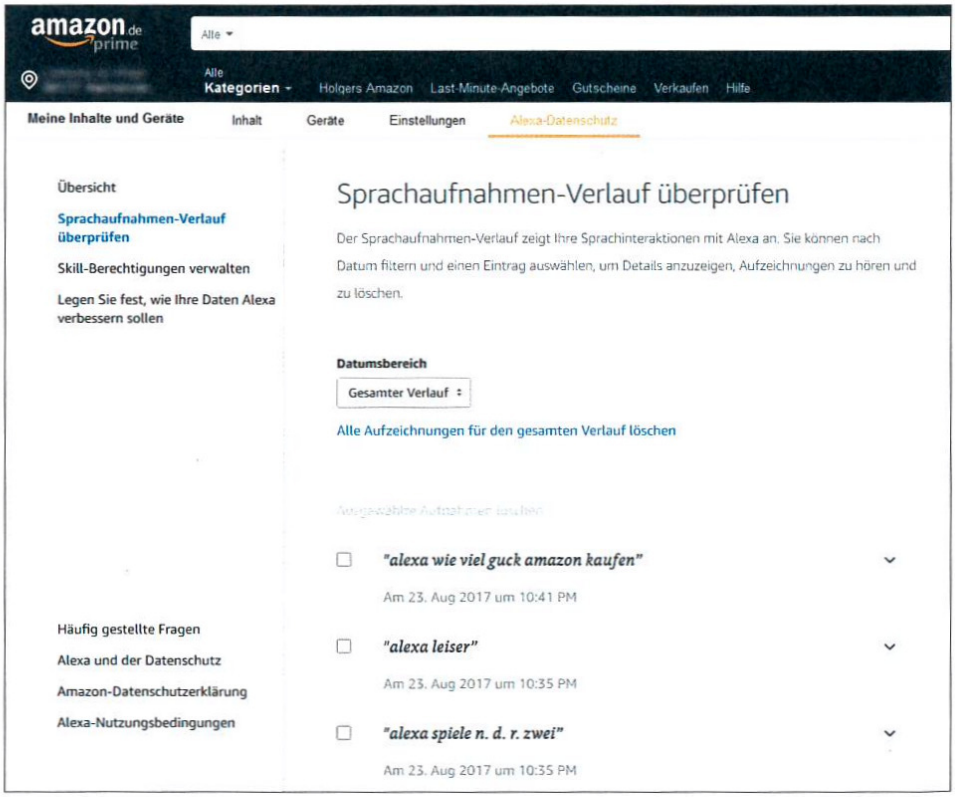
„Unglücklicher Fall"

Dass Amazon.de Kundendaten falsch zugeordnet hat und dies nicht aufgefallen ist, deutet auf mangelhafte Prüfmechanismen im Auskunftsverfahren hin. Eine Vier-Augen-Kontrolle hat da wohl kaum stattgefunden. Hinzu kommt, dass der Empfänger dieser Daten sogar auf den Fehler aufmerksam machte, darauf aber nicht einmal eine Antwort erhielt. Und auch Nils Schmidt, das Opfer des Leaks, gibt an, dass ihn Amazon.de vor unserem Anruf nicht informiert habe.

Amazon.de hat sich offensichtlich gleich auf mehreren Ebenen in einem hochsensiblen Bereich fehlerhaft verhalten. Wir schilderten dem Unternehmen den Vorfall, freilich ohne zu erwähnen, dass wir den Betroffenen zuvor bereits identifiziert und kontaktiert hatten. Unter anderem wollten wir wissen, ob Amazon die Panne innerhalb von 72 Stunden an die zuständige Datenschutzbehörde gemeldet hat – wie es in solchen Fällen Pflicht ist (siehe Kasten).

Antworten auf unsere konkreten Fragen blieb Amazon schuldig. Nach vier Werktagen folgte lediglich ein kurzes Statement, hier im kompletten Wortlaut: „Dieser unglückliche Fall war die Folge eines menschlichen Fehlers. Wir haben das Problem mit den beiden beteiligten Kunden geklärt und Maßnahmen zur weiteren Verbesserung unserer Prozesse ergriffen."

Wie diese „Klärung" konkret ablief, schilderten uns die Kunden: Martin Schneider wurde am 8. Dezember – also drei Tage nach unserer Anfrage – von einem Amazon-Mitarbeiter angerufen, der erklärte, dass da „einem Mitarbeiter ein einmaliger Fehler unterlaufen" sei. Nils Schmidt wurde am selben Tag telefonisch darüber informiert, dass Alexa-Sprachdateien von ihm in fremde Hände



Amazon speichert die Alexa-Sprachaufzeichnungen unbefristet. Sie lassen sich jedoch manuell im Kundenportal löschen.

Hohes Risiko

Rechtsanwalt Dr. Carlo Piltz findet, dass Amazon den Betroffenen über die irrtümliche Weitergabe seiner Alexa-Daten hätte informieren müssen. Piltz berät mittlere und große Unternehmen aus Industrie und Handel im nationalen und internationalen Datenschutzrecht. Außerdem betreibt er das Datenschutz-Blog delegedata.de.

Dr. Carlo Piltz: Meldung nach 72 Stunden ist Pflicht.

c't: Hätte Amazon den Betroffenen sofort informieren müssen, als das Unternehmen auf die Panne hingewiesen wurde?

Carlo Piltz: Sowohl gegenüber der Aufsichtsbehörde als auch gegenüber der betroffenen Person besteht grundsätzlich eine Melde- beziehungsweise Benachrichtigungspflicht, wenn es zu einer Verletzung des Schutzes von personenbezogenen Daten gekommen ist und entweder ein normales oder hohes Risiko besteht. Die DSGVO unterscheidet allerdings bei den Anforderungen den jeweiligen Adressaten der Unterrichtung.

Der Verantwortliche – in diesem Fall Amazon – muss nach der Feststellung des Risikos binnen 72 Stunden die Verletzung bei der Aufsichtsbehörde melden. Im Unterschied dazu muss es zu einer Unterrichtung der betroffenen Person nur unverzüglich – also ohne schuldhaftes Zögern – kommen, wenn voraussichtlich ein hohes Risiko besteht. Die Meldung an den Betroffenen hängt also nicht an einer fixen zeitlichen Frist. Was in der Praxis „unverzüglich" bedeutet, ist im Rahmen der DSGVO derzeit auf jeden Fall noch Auslegungssache und sicherlich eine Frage des Einzelfalls. Als deutsche Juristen würden wir mit Blick auf das deutsche Zivilrecht davon ausgehen, dass „unverzüglich" ungefähr 14 Tage bedeutet. Als europäisches Gesetz darf man die DSGVO aber natürlich nicht nur aus der deutschen Brille betrachten.

c't: Halten Sie die Datenpanne von Amazon.de nach DSGVO für sanktionswürdig?

Piltz: Die Weitergabe von Tondateien an einen unberechtigten Dritten kann mehrere DSGVO-Vorgaben verletzen, zum Beispiel die Vorschriften zur Sicherstellung einer angemessenen Datensicherheit oder auch die Pflicht, Daten nur zu übermitteln, wenn dafür eine Erlaubnis vorliegt. Die Aufsichtsbehörde muss sich also fragen, ob und wenn ja, in welcher Höhe ein Bußgeld zu verhängen ist und den Verhältnismäßigkeitsgrundsatz in ihren Maßnahmen berücksichtigen.

Die DSGVO listet dafür Bewertungsmaßstäbe auf, die in jedem Einzelfall zu beachten sind. Kam es vorsätzlich oder fahrlässig zu dem Verstoß? Wie umfangreich war die Zusammenarbeit mit der Aufsichtsbehörde, um dem Verstoß abzuhelfen und seine möglichen nachteiligen Auswirkungen zu mindern? Da spielen also sehr viele Faktoren eine Rolle, die wir derzeit nicht alle kennen. Es wäre nun Aufgabe der zuständigen Aufsichtsbehörde, den Vorfall genauer zu untersuchen. Erst in diesem Prozess ließe sich die Sanktionswürdigkeit final beurteilen und die Notwendigkeit eines Bußgelds klären.

geraten seien. Schmidt habe ungefähr zur selben Zeit eine DSGVO-Selbstauskunft beantragt, dann sei es zu einer teilweise falschen Zuordnung der Daten durch einen Mitarbeiter gekommen, teilte der Kundensupport mit. Dies habe man selbst festgestellt.

Speicherhunger

Zu diesem datenschutzrechtlichen GAU hätte es nie kommen können, wenn Amazon die Sprachaufzeichnungen zeitnah löschen würde, anstatt sie unbefristet in der Cloud zu speichern. Amazon aber will diese Daten, um „unsere Systeme zur Spracherkennung und zum Verstehen natürlicher Sprachen zu trainieren", wie in der Datenschutz-FAQ erklärt wird. Der Konzern insistiert, dass der Kunde ja alle „verbundene Sprachaufzeichnungen überprüfen und die Sprachaufzeichnungen einzeln oder alle auf einmal löschen" könne (im Amazon-Account unter amazon.de/alexaprivacy). Doch wie viele Kunden wissen von der Option und machen davon Gebrauch?

Datenschutzrechtlich ist dieses Gebaren zumindest problematisch. Der von uns in den Fall eingeweihte DSGVO-Experte Dr. Carlo Piltz erklärt: „Das Erfordernis von datenschutzfreundlichen Voreinstellungen ist fest in der DSGVO verankert." Dies bedeute aber nicht, per se null Daten verarbeiten zu dürfen, sondern es orientiere sich immer am erforderlichen Zweck: „Wenn hier der Zweck der Datenverarbeitung zur Anlernung der KI tatsächlich erforderlich ist, um auch den Dienst mit Spracherkennung so anbieten zu können, würde der Grundsatz der Datensparsamkeit diese Verarbeitung also nicht unbedingt verbieten. Es dürften nur nicht mehr Daten verarbeitet und länger gespeichert werden, als für diesen Zweck erforderlich."

Ob es sich bei der Datenpanne um einen Einzelfall handelt, bleibt unklar. Amazon.de wollte uns auch auf weitere Nachfrage nicht mitteilen, ob sie wenigstens pflichtgemäß an die zuständige Datenschutzaufsicht gemeldet wurde. Nils Schmidt hilft das ohnehin nicht mehr weiter. Immerhin hatte er das Glück, dass seine Sprachdaten bei uns statt in sozialen Medien landeten. Amazon.de zeigte sich ihm gegenüber äußerst erkenntlich: Als Entschädigung erhielt Schmidt nach eigenem Bekunden eine kostenlose Prime-Mitgliedschaft sowie einen Echo Dot und Spot. *(hob@ct.de)* c't