DSGVO – 11/2020 – Daten löschen in komplexen Systemlandschaften

DSGVO – 11/2020 – Daten löschen in komplexen Systemlandschaften

[expand title="mehr lesen…"]

Daten löschen in komplexen Systemlandschaften

Ballast abwerfen

Frank Heisel

Viele Unternehmen setzen das in der DSGVO geforderte Löschen obsoleter Daten entweder gar nicht oder nur unzureichend um. Da sich die Datenschutzaufsichtsbehörden in jüngster Zeit verstärkt für das Thema interessieren, müssen Firmen umdenken.

-tract

- Das datenschutzkonforme Löschen personenbezogener Informationen in verteilten Systemen ist eine anspruchsvolle Aufgabe, denn es gibt meist zahllose Abhängigkeiten zwischen den Daten.
- Dabei müssen die Verantwortlichen nicht nur rechtliche,

sondern auch viele technische Aspekte berücksichtigen.
- Bislang haben viele Unternehmen in dieser Hinsicht wenig getan. DSGVO, bissige Datenschutzbehörden und hohe Bußgelder zwingen sie nun zum Umdenken.

Personenbezogene Daten lassen sich in einem einzelnen System noch recht einfach löschen. In komplexen, heterogenen IT-Welten sieht die Sache anders aus. Hier gilt es, die Integrität, Verfügbarkeit und Konsistenz der Informationen nicht durch übereiltes Hantieren in den Datenbeständen zu gefährden.

Art. 17 Abs. 1 DSGVO räumt einem Betroffenen beim Vorliegen bestimmter Gründe das Recht ein, etwa die Betreiber eines Onlineshops dazu aufzufordern, seine persönlichen Daten unverzüglich aus den Systemen zu entfernen. Das anlassbezogene Löschen muss beispielsweise bei Widerruf einer erteilten Einwilligung erfolgen [Art. 17 Abs. 1 lit. b) DSGVO] oder bei - einem erfolgreichen Widerspruch gegen die Verarbeitung [Art. 17 Abs. 1 lit. c) DSGVO]. Der Artikel 17 regelt das sogenannte Recht auf Vergessenwerden.

In diesem Fall ergreift der Betroffene die Initiative. Um seinen Wunsch zu erfüllen, muss das Unternehmen einen Prozess implementiert haben, der das sofortige Prüfen der Anfrage einleitet. Dieser Prozess erfasst alle zugehörigen Datensätze, die über mehrere Systeme verteilt sein können. Nach der Überprüfung, ob und welche Daten gelöscht werden können, startet die zuständige Fachabteilung den Prozess manuell. Falls die Entwickler eine Löschfunktion eingerichtet haben, sollte man sie natürlich verwenden. Eine solche einfache Löschung funktioniert allerdings nur bei Systemen, die keine oder nur wenige Schnittstellen zu vor- und nachgelagerten Rechnern haben.

Das anlasslose Löschen verlangt das Entfernen von Daten, ohne dass jemand eine explizite Anfrage stellt. Diese Pflicht ergibt sich aus Art. 17 Abs. 1 lit. a) DSGVO. Danach müssen

Unternehmen personenbezogene Daten löschen, wenn der Zweck der Verarbeitung entfällt und keine Gründe gemäß Art. 17 Abs. 3 DSGVO dagegensprechen.

Die Aufgabe besteht nun darin, Löschfristen für einzelne Datensätze und -felder festzulegen, Löschregeln zu definieren und sie regelmäßig anzuwenden. Die Bereinigungsfrequenz dürfte sich in der Regel an den Aufbewahrungsfristen orientieren – je kürzer sie sind, desto höher die Rate der Wiederholungen. Bei wenigen und einfach strukturierten Daten lässt sich dieser Vorgang vergleichsweise einfach umsetzen, indem man einzelne Skripte zeitgesteuert ausführt oder manuell eingreift. Im Verbund mit anderen Systemen und verteilter Ablage der Daten kann das Löschen dagegen nur als Gesamtkonzept funktionieren, da ein ungeregelter Datenfluss zwischen den Systemen gelöschte Daten mit der nächsten Übertragung wieder einspielen kann.

Daten liegen überall verstreut

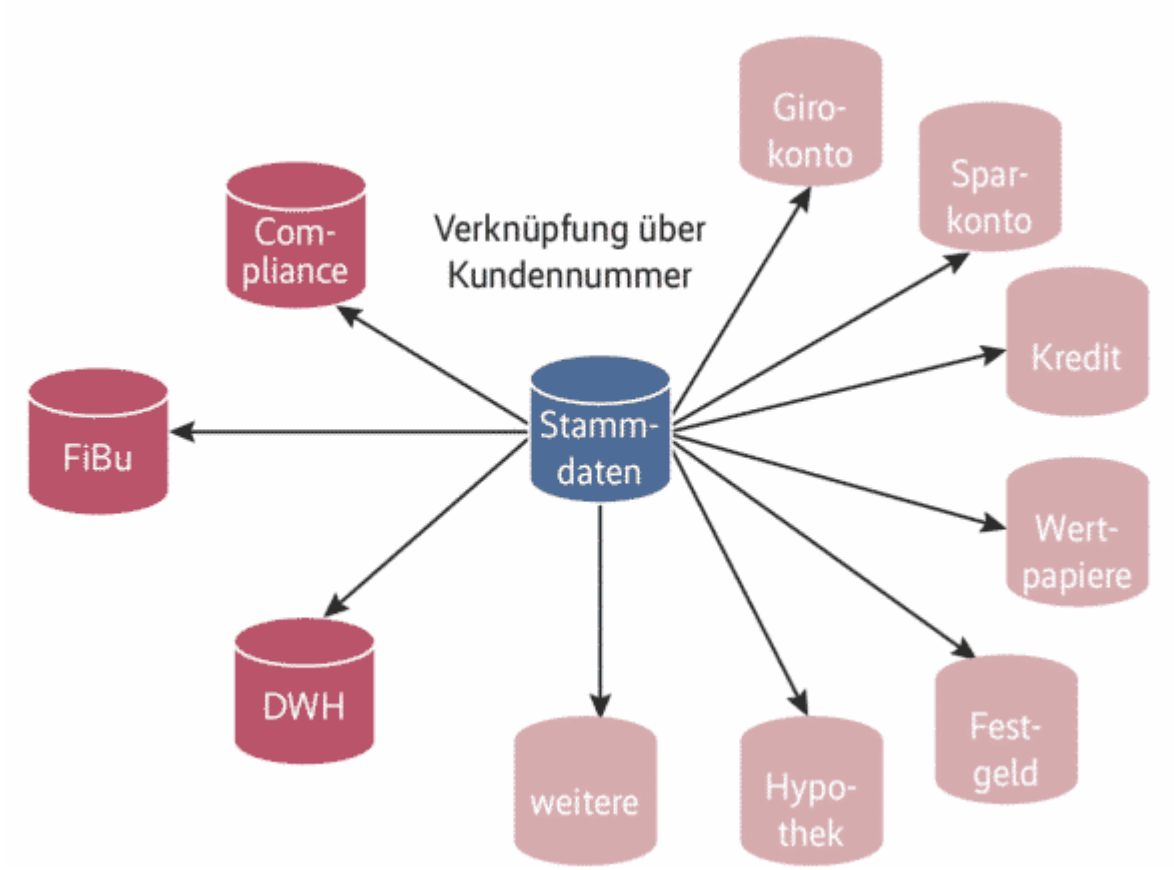
In größeren Organisationen besteht die IT-Landschaft gewöhnlich aus zahlreichen heterogen organisierten Anwendungen. Personenbezogene Daten werden zu verschiedenen Zwecken in vielen Datenbanken mehrfach oder verteilt gespeichert. Und nicht nur in den Produktivsystemen: Auch Entwicklungs- und Testumgebungen arbeiten oft damit.

In einem Webshop beispielsweise liefert der Kunde mindestens Name und Anschrift ab. Die Rechnungen werden als PDF gespeichert, per E-Mail versendet und an nachgelagerte Applikationen wie die Finanzbuchhaltung weitergereicht. Der Webshop will meist mehr wissen. Rechnungen müssen jedoch nur die in § 14 Abs. 4 UStG genannten Informationen enthalten. Alle erhobenen Daten landen normalerweise in einem CRM-System oder in einem Data Warehouse. Das Löschen einzelner Einträge muss daher in verschiedenen Systemen zu unterschiedlichen Zeitpunkten erfolgen.

Der Umfang der gespeicherten Daten verdient ebenfalls

Beachtung. Shop- oder CRM-Anwendungen erfassen außer dem Namen weitere Daten wie Kundennummer, Lieferanschrift, Rechnungsanschrift, Telefonnummer und E-Mail. Diese Daten sind notwendig, um die Bestellung zu bearbeiten, etwa um die Lieferung durch eine Spedition zu veranlassen und den Versand per E-Mail anzukündigen. Nach erfolgreicher Zustellung braucht man einiges davon zur Erfüllung des Vertrags jedoch nicht mehr. So sind für das Erstellen der Rechnung neben den Informationen zu den gelieferten Waren oder Leistungen lediglich der Name und die Anschrift des Empfängers, nicht jedoch die E-Mail-Adresse oder Telefonnummer erforderlich. Dafür entfällt dann der Verarbeitungszweck.

Komplexe Systeme verbinden die personenbezogenen Daten beispielsweise über ein Merkmal wie die Kundennummer (Abbildung 1). Die eigentlichen Daten liegen in der Stammdatenbank, alle anderen Systeme greifen lediglich über die Verknüpfung darauf zu. Zudem existieren viele Schnittstellen für den Datenaustausch. Bei einer Analyse des Datenflusses gehören alle Schnittstellen, über die personenbezogene Informationen laufen, in die Betrachtung. Wesentlichen Einfluss auf das Einsetzen einer datenschutzkonformen Löschung haben Faktoren wie der Umfang der personenbezogenen Daten, die Häufigkeit der Übertragung und die Art der Schnittstelle.



So etwa sieht die Systemlandschaft in einer Bank aus. Jedes Datenhaltungssystem speichert nur Teile eines kompletten Datensatzes. Löschen ist hier kompliziert (Abb. 1).

Akribische Untersuchungen sind notwendig

Die Fachabteilungen müssen für jede Schnittstelle zunächst erheben und dokumentieren, welche Daten sie wohin übermitteln. Entscheidend ist dabei, ob sie ausschließlich über ein systemweit eindeutiges Merkmal wie die Kundennummer verknüpft werden oder ob weitere Parameter im Spiel sind. Hier sollte man im Sinne einer guten „Privacy by Design“ darauf achten, dass nur notwendige Daten fließen. Bei Ausgangsrechnungen wäre es ausreichend, lediglich den Namen und die Anschrift des Rechnungsempfängers zu übertragen, nicht jedoch die weiteren Kontaktdaten.

Ein weiterer Aspekt ist die Häufigkeit der Übermittlung. Die meisten Schnittstellen dürften regelmäßig Daten erhalten und weitergeben. Falls das empfangende System eigene autonome Löschroutinen anbietet, müssen Übermittlungs- und

Löschfrequenz aufeinander abgestimmt sein, da es sonst vorkommen kann, dass Daten erneut übertragen werden und jemand sie im Zielsystem wieder löschen muss. Bei der Implementierung von Löschprozessen spielt daher die Art der Schnittstelle eine entscheidende Rolle.

Werden die Daten in Dateien verschickt, muss sichergestellt sein, dass auf dem abgebenden System oder einem eventuell zwischengeschalteten Fileserver keine Kopien verbleiben, da der Verarbeitungszweck nun erfüllt ist. Gegebenenfalls benötigt der Fileserver ein eigenes Löschkonzept, und er darf nicht als Backup oder Archiv missbraucht werden.

Eine Möglichkeit, Abhängigkeiten zu berücksichtigen, bieten zentrale Löschsysteme, die ihre Arbeit über alle angeschlossenen Systeme koordinieren und dafür sorgen, dass keine Daten verschwinden, die anderweitig noch benötigt werden. Man spricht in solchen Fällen von einer Orchestrierung der Löschung. Es gibt einige kommerzielle Produkte, etwa das Modul SAP ILM für die SAP-Welt, eine Alternative dazu namens - Cronos von der gleichnamigen Unternehmensberatung aus Münster und das universal einsetzbare Customer Data Deletion Management (CDMS) von Impetus.

Die zentrale Löschinstanz sammelt von allen Systemen die personenbezogenen Daten ein und hält sie in einer eigenen Datenbank. Die Löschregeln informieren bei Bedarf die Systemverantwortlichen über zu löschende Datensätze. Administratoren können dann das Löschen manuell oder automatisiert anstoßen und einen Nachweis an das zentrale Löschsystem schicken, das diesen archiviert. Hier angesiedelte Regeln sollten das unkontrollierte Löschen in den betroffenen Ablageorten verhindern. So lässt sich beispielsweise sicherstellen, dass der Stammdatensatz eines Kunden, der noch laufende Verträge hat, erst dann freigegeben wird, wenn alle Verträge beendet und die jeweiligen Aufbewahrungsfristen ausgelaufen sind.

Weiterer Vorteil einer zentralen Instanz: Unternehmen können sich bei einer Anfrage schnell einen Überblick darüber verschaffen, ob und welche Informationen zu dem Betroffenen vorliegen. Das Löschen lässt sich bei gesetzlichen Aufbewahrungsfristen mit einer entsprechenden Regel unterdrücken.

Manchmal müssen Daten verweilen

In bestimmten Situationen muss das Löschen grundsätzlich verhindert werden. Als Beispiel seien hier eine laufende Betriebsprüfung oder ein andauernder Rechtsstreit genannt. Bei einer Betriebsprüfung müssen die Daten bis zum Abschluss vorliegen. Hier bietet es sich an, die Frist zu verlängern. Relevante Daten zur Rechnungslegung liegen oft in Jahresarchiven, deren Löschung sich aussetzen lässt. Sie muss nach Abschluss der Prüfung manuell angestoßen werden. In jedem Fall sollte der verantwortliche Fachbereich vor dem Auslösen der regulären Löschung über den Vorgang mitbestimmen. Um nicht zu viele Daten auszunehmen, ist eine genaue Prüfung der Systeme und Daten notwendig.

Während eines Rechtsstreits müssen die benötigten Daten ebenfalls zur Verfügung stehen. In diesem Fall sollte die Rechtsabteilung die Fälle, deren Daten nicht gelöscht werden sollen, an den für das Löschen verantwortlichen Fachbereich übermitteln. Der nimmt die betroffenen Daten von der Löschung aus und entfernt sie nach Abschluss des Verfahrens einzeln. Alternativ bietet es sich an, alle benötigten Daten in einem separaten Archiv zu sichern, das sich in der Verantwortung der Rechtsabteilung befindet. Ist der Rechtsstreit beendet, stößt sie die datenschutzkonforme Löschung an.

Um die Rechenschaftspflicht gemäß Art. 5 Abs. 2 DSGVO zu erfüllen, ist das Unternehmen verpflichtet, die erfolgreiche Löschung nachweisen zu können. Dabei muss es darauf achten, keinen neuen Fundus personenbezogener Daten zu erzeugen. Das Ausführen eines SQL-Löschbefehls und das Speichern der

Ergebnismenge wäre beispielsweise ein neuer Datenbestand, dem allerdings die Rechtsgrundlage fehlt.

Bei einer anlassbezogenen Löschung sind die Verantwortlichen verpflichtet, den gesamten Vorgang der Anfrage zu dokumentieren. Hier kann auch das erfolgreiche Löschen von Daten einfließen. Mit Screenshots, den SQL-Befehlen oder dem negativen Ergebnis einer Suchabfrage ließe sich der technische Vorgang nachweisen.

Schwieriger ist dagegen das Dokumentieren des anlasslosen Löschens nach Zeitablauf. Das erledigen in der Regel automatisierte Skripte, die in den Datenbanken parametrisierte Routinen ausführen. Als Nachweise können hier die Systemeinstellungen gelten, die zeigen, dass die Löschroutinen regelmäßig arbeiten, sowie die Skripte selbst. Zusätzlich müssen die Zuständigen die Parametrisierung der einzelnen Aufrufe und das Ergebnis belegen. Dabei dürfen nur die Informationen im System verbleiben, die die erfolgreiche Ausführung dokumentieren, nicht jedoch die gelöschten Daten selbst.

Backup und Restore sind außen vor

Anders als in Archiven darf in Backups niemand etwas ändern, da deren Zweck in der Wiederherstellung von Daten liegt. Das Herausnehmen einzelner Daten aus einem Backup kann die Konsistenz der Daten verletzen und das Wiederherstellen verhindern. Bei einem Restore geraten jedoch auch die gewollt gelöschten Daten zurück ins System. Um diesen Vorgang datenschutzkonform abzuwickeln, müssen sie im zurückgespielten Datenbestand nochmals gelöscht werden. Das lässt sich beispielsweise anhand der Löschnachweise erledigen. Das Vorgehen beim Wiederherstellen sollte im entsprechenden Verfahren dokumentiert und getestet sein.

Genormtes Vorgehen

Eine gute Grundlage für Löschkonzepte findet sich in der Norm DIN 66398. Sie zeigt eine Vorgehensweise zum Erstellen solcher Konzepte und bietet Orientierungshilfe für Unternehmen aller Größen. Projekte, die das erfolgreich umsetzen wollen, kommen ohne die betroffenen Fachbereiche, die IT-Abteilung und den Datenschutzbeauftragten nicht aus.

Ein solches Projekt könnte nach folgenden Schritten vorgehen (Abbildung 2):

Schritt für Schritt: Vorgehensmodell zum regelkonformen Löschen personenbezogener Daten (Abb. 2)

System-Scoping: Zunächst müssen die Zuständigen alle Systeme ermitteln, die personenbezogene Daten verarbeiten. Hierzu sollte das Unternehmen in einem Regelwerk Datenschutzklassen vorgeben. Häufig erfolgt das Klassifizieren im Rahmen der Schutzbedarfsfeststellung, die neben den Kriterien Vertraulichkeit, Integrität und Verfügbarkeit auch erfasst, ob personenbezogene Daten verarbeitet werden.

Untersuchung der Daten: Weiterhin ist für alle erkannten Systeme detailliert aufzunehmen, um welche Daten es geht. Sie lassen sich dann in vorgegebene Kategorien einsortieren.

Fachliche Löschregeln: Aus den vorliegenden Informationen können die Administratoren nun fachliche Löschregeln ableiten. Dabei ermitteln sie für jede Datenkategorie, ob Aufbewahrungsfristen existieren und welche Löschfristen sich daraus ergeben.

Analyse der Schnittstellen und Abhängigkeiten: Zu den einzelnen Systemen müssen die Fachabteilungen weiterhin die Schnittstellen zu vor- und nachgelagerten Systemen finden und analysieren. Dabei erfassen sie die Quell- und Zielsysteme, die Datenkategorien, die Häufigkeit der Übertragung und die Art der Schnittstelle. Mit dieser Dokumentation können sie Abhängigkeiten untersuchen.

Fachliches Löschkonzept: Aus den gesammelten Informationen wird ein Löschkonzept erstellt.

Entscheidung für eine Löschlösung: Abhängig von der Analyse muss das Unternehmen eine Löschstrategie formulieren – etwa, ob es eine zentrale, eine dezentrale oder eine Mischlösung einführen will.

Technische Löschkonzepte: Sie beschreiben die konkrete Durchführung des Löschens. Auch Produktion und Ablage der Löschnachweise werden hier dokumentiert.

Test: Vor der Einführung der Löschroutinen müssen diese ausführlich geprüft werden. Dabei sind sowohl die vollständige Löschung innerhalb des Systems als auch die Abhängigkeiten zu anderen Systemen einzubeziehen.

Umsetzung: Nach erfolgreichem Test können die Zuständigen die Löschroutinen auf die einzelnen Systeme loslassen. In der Regel kann die Einführung in Stand-alone-Systemen, die keinerlei Schnittstellen nach außen haben, unabhängig vom Rest der IT-Welt erfolgen. Die meisten Systeme sind jedoch mit anderen verbunden und müssen gleichzeitig in Produktion genommen werden.

Ablage der Nachweise: Darüber muss ebenfalls eine Entscheidung fallen. Eine zentrale Löschlösung bietet oftmals eine Ablagemöglichkeit an. Wird dezentral gelöscht, müssen die Zuständigen über eine gemeinsame Ablagemöglichkeit und einen Prozess nachdenken, der die Nachweisführung sicherstellt.

Fazit

Das Ganze ist genauso kompliziert, wie es sich anhört. Die Vielzahl der Systeme, die Abhängigkeiten, der Umfang der Daten und ihre Redundanz erschweren die Umsetzung. Hinzu kommt, dass mit dem Löschen von Daten ein Kulturwandel einhergeht, da der jederzeitige Zugriff auf Daten in Vor-DSGVO-Zeiten wichtiger schien als das datenschutzkonforme Löschen.

Die mit der Einführung der Datenschutz-Grundverordnung drastisch gestiegenen und von den Aufsichtsbehörden inzwischen auch durchgesetzten Bußgelder sollten reichen, den Unternehmen das Umdenken zu erleichtern. Bei entsprechenden Projekten muss man sich auf möglicherweise lange Laufzeiten von mehreren Monaten bis zu einigen Jahren einstellen. (jd@ix.de) Frank Heisel

war bei einer Big-Four-Wirtschaftsprüfungsgesellschaft als verantwortlicher Prüfungsleiter für System-, Prozess- und IKS-Prüfungen tätig. Er ist als externer Datenschutzbeauftragter für mehrere Unternehmen benannt und berät Unternehmen zum Thema Datenschutz und internes Kontrollsystem.

[/expand]