Löschkonzepte für Datenbanksysteme – Alle Daten im Griff?

Autor: Dr. Axel-Michael Wagner

[März 2021]

ZUSAMMENFASSUNG

Das Thema „Datenmanagement und Datenschutz in Unternehmen“ lässt sich in zwei Teilbereiche aufteilen: Frontend und Backend. Im Frontend geht es um dasjenige, mit dem ein Unternehmen bzw. Verantwortlicher im datenschutzrechtlichen Sinne nach außen hin sichtbar wird. Da geht es um datenschutzrechtliche „Populärthemen“ an der Schnittstelle zwischen Verantwortlichem und Betroffenem wie etwa Pflichtinformationen gegenüber Betroffenen, Einwilligungserklärungen, Cookies, Auskunftserteilung, Facebook-Fanpages, Werbe-E-Mails und Ähnliches. Das Backend wird demgegenüber häufig stiefmütterlich behandelt, obwohl dort doch das „Öl des 21. Jahrhunderts“ in den großen „Tanks“ der Unternehmen lagert: die sagenumwobenen, riesigen Datenschätze.

Das unübersichtliche Backend als gigantische „Daten-Baustelle“

Da liegen sie nun, die Daten, meist personenbezogen, in Tera-, Peta- und Exa-Byte-Größe, in unübersichtlichen, historisch gewachsenen Datenbanken, entweder auf „remote“-Serverstrukturen in der Cloud oder „on premise“ beim Unternehmen selbst. In diese Datenbanken können Angestellte wie durch kleine Fenster von ihren Desktop-PCs hineinsehen. Unentwegt werden Daten an andere Systeme exportiert und Daten aus anderen Systemen importiert, teils über Unternehmensgrenzen hinweg, teils als „Konzerndaten“ ohne Zuordnung zu einzelnen Konzerngesellschaften. Nur die Datenbanksoftware und das grundsätzliche Abfrage- und Verarbeitungs-Framework stammen vom Softwarehersteller, während die Datenstrukturen und die tatsächlich einsetzbaren Funktionalitäten meist das Ergebnis eines aufwendigen, immer „under construction“ befindlichen Customizing-Prozesses des Unternehmens selbst und/oder seiner IT-Berater sind. Denn letztlich geht es darum, sowohl die Daten als auch die dahinterliegenden Datenstrukturen und die Abfrage- und Verarbeitungslogiken an die sich beständig ändernden Prozesse des Unternehmens anzupassen. Sämtliche größeren IT-Systeme in Unternehmen basieren auf solchen Datenbankstrukturen, seien es „Enterprise Resource Planning“-Systeme (ERP), seien es „Enterprise Content Management“-Systeme (ECM) wie etwa Dokumenten-Management-Systeme (DMS), seien es „Customer Relationship“-Systeme (CRM), „Human Capital Management“-Systeme (HCM) oder Management-Informations-Systeme (MIS). Und keines dieser Systeme ist statisch, sondern alle sind ständig in Bewegung, sowohl was den Inhalt, als auch was die Strukturen angeht.

Nicht nur „kundendatenintensive“ Branchen wie Social Media, Cloud-Dienste, die Versicherungs-, Werbe- und Unterhaltungsindustrie sowie Online-Angebote jeder Art verwenden derartige Systeme, sondern auch klassische Industrien wie die Finanzdienstleistungsindustrie, Automotive, Maschinenbau oder Logistik. Selbst Unternehmen, die hartnäckig behaupten, keine personenbezogenen Daten zu verarbeiten, müssen spätestens bei der Diskussion um die (umfangreichen) Personaldaten ihrer eigenen Arbeitnehmer einsehen, dass Handlungsbedarf besteht, sich mit dem Thema der gewachsenen „Altdatenbestände“ zu befassen.

Im Hinblick auf die schier unerschöpflichen Datenbestände der heterogenen IT-Landschaften gibt es mindestens zwei Ordnungsanforderungen: Datenmanagement und Datenschutz. Datenmanagement ist das, was das Unternehmen mit den Daten machen *will*; Datenschutz ist das, was das Unternehmen mit den Daten machen *darf*. Man kann sich leicht ausmalen, dass das nicht immer deckungsgleich ist. In der Praxis geht allerdings die bewusste Bildung des Willens des Unternehmens darüber, was man mit den Daten noch oder keinesfalls mehr anstellen will, oft unter den Anstrengungen verloren, dass die Sys-

teme überhaupt „laufen“. Sogar bei großen DAX-Konzernen gibt es nicht selten einen historisch gewachsenen Flickenteppich von Backend-Systemen, die ungerne angefasst werden und daher auch jede Menge Compliance-Risiken bergen, denn: *„never touch a running system“*.

Befragt man Fachabteilungen, mit welchen personenbezogenen Daten sie umgehen, so kommt es für die (Qualität der) Antwort erfahrungsgemäß wesentlich darauf an, ob die Prozesse der Fachabteilung schriftlich niedergelegt wurden. Wer seine Prozesse im Griff hat, hat auch meist seine (aktuellen) prozessbezogenen Daten im Griff. Eine konkrete Rechtspflicht zur Niederlegung einer „schriftlichen Ordnung“ und damit von entsprechenden Prozessbeschreibungen gibt es aber nur in regulierten Industrien wie im Finanzsektor. Die Konkretisierung allgemein gehaltener gesetzlicher Vorgaben durch Verwaltungsanweisungen wie den GoBD im steuerlichen Bereich, die dann Verfahrensdokumentation bzw. Prozessbeschreibungen einfordern, sind nicht auf personenbezogene Daten fokussiert. Die DSGVO selbst enthält ein paar „schwammige“ Formulierungen (wie Art. 5 Abs. 2 DSGVO), die eine Dokumentationspflicht interner Prozesse nicht ausdrücklich vorsehen. Das verwundert nicht, weil die Spannbreite der Unternehmen, die hier angesprochen werden, so groß ist, dass durch generelle Prozess-Dokumentationspflichten weithin „mit Kanonen auf Spatzen geschossen“ würde.

Soweit es in Unternehmen überhaupt Prozessbeschreibungen gibt, sind diese erfahrungsgemäß häufig veraltet, nur auf Teilprozesse bezogen und/oder nur für Eingeweihte lesbar (die das Unternehmen dann im Zweifel schon verlassen haben). Zumeist handelt es sich bei den aktuellen abteilungsinternen Abläufen und ebenso beim Informationsaustausch mit anderen Abteilungen, Systemen und externen Dritten um (mündlich) „überliefertes“ Wissen, selbst wenn das Unternehmen eigentlich der Meinung ist, schon viel verschriftlicht zu haben. Häufig lautet dann die Aussage, dass es dazu zwar ein Dokument gibt, das aber nicht das beschreibt, was in der Praxis tatsächlich getan wird. Ein „self-assessment“ durch die Fachabteilung, welche personenbezogenen Daten denn in den IT-Systemen für sie „gehostet“ werden, scheitert daher häufig daran, dass erst umfangreiches, nur in verschiedenen Köpfen vorhandenes Wissen zusammengezogen und verschriftlicht werden müsste. Dies ist ein Aufwand, der unter dem Tagesgeschäft begraben zu werden droht und daher ist das Ergebnis solcher Bemühungen oft „suboptimal“.

Spricht man hingegen mit den IT-Abteilungen, die für die Backend-Systeme verantwortlich sind, erhält man seitenlange Listen mit – mehr oder weniger kryptischen – Bezeichnungen von Feldern, Datensätzen und Ablauflogiken. Die dahinterliegende Software-Struktur generiert pausenlos aus unterschiedlichsten Daten noch mehr Daten und Reports, die dann wieder an unterschiedlichste weitere Orte verbracht werden. Die Fachabteilungen äußern

in diesem Kontext punktuell aufgrund tagesaktueller Bedürfnisse Wünsche zur Weiterentwicklung der Systeme, die dann auch umgesetzt, aber nicht systematisch dokumentiert werden, und deren Anlass, Zusammenhang und Zielsetzung später kaum jemand mehr rekonstruieren kann.

Überhaupt kann die IT-Abteilung häufig das genaue Funktionieren einzelner Algorithmen oder die genauen Verknüpfungen von Datenfeldern, also „was tatsächlich aktuell passiert", nur nach einer Analyse aktueller Programm- und Datenbankstände erklären, sprich nur auf konkrete Anforderung hin, nicht aber in Tiefe und Breite gleichermaßen. Das ist verständlich, weil eine vollständige, stets aktuelle (IT-) Dokumentation „über alles hinweg" mindestens so aufwendig zu pflegen wäre wie der zugrunde liegende Code bzw. der Datenbankaufbau selbst. Vor diesem Hintergrund ist es einfacher bzw. ressourcenschonender, bei konkretem Bedarf „einfach noch einmal reinzuschauen" als tausende von Seiten Dokumentation komplexer Systeme so zu pflegen, dass sie stets aktuell und auch für Dritte nachvollziehbar ist. Und selbst wenn es eine solche Dokumentation geben würde, so würde sie dennoch erst einmal nur IT-Zusammenhänge, nicht aber die dahinterstehende „Business-Logik", geschweige denn die Herleitung und Bewertung aus Compliance-Sicht erläutern. So dümpeln in den Systemen viele Daten umher, die weder die Fachabteilung (nach vielen Personalwechseln) noch die IT-Abteilung (nach ebenso vielen Personalwechseln) kennen und erst recht nicht erklären können, wofür diese (noch) wichtig sind.

Praxistipp: Die notwendigen Informationen für die Erstellung eines Löschkonzepts für ein komplexes IT-System müssen in einem geführten Dialog mit Fach- und IT-Abteilung generiert werden. Der Ansatz über die Fachabteilung ist üblicherweise „top down" (welche Daten „verwendet" die Fachabteilung wie), übersieht aber oft die Existenz weiterer (Einzel-) Daten bzw. Kopien/Speicherorte. Der Ansatz über die IT-Abteilung ist üblicherweise „bottom up" (welche Daten sind im System), kann aber häufig aufgrund eingeschränkter (Meta-) Datenqualität Daten nicht korrekt nach rechtlichen Kriterien (Betroffenengruppe, Zweckbestimmung, Kategorie des Einzeldatums) filtern bzw. selektieren. Es muss daher erwogen werden, beide Ansätze zu kombinieren, bzw. begründbar sein, weshalb davon abgesehen wurde. Im Idealfall entsteht im Rahmen eines derartigen Bestandsaufnahme-Projekts nicht nur eine „Datenlandkarte", sondern auch bereits eine Erkenntnis, welche Daten von welcher internen oder externen Funktion bzw. Stelle für welchen Zweck überhaupt noch benötigt werden.

Wo kein Kläger, da kein Richter?

Wer viele personenbezogene Daten anhäuft, sollte auch ihre Löschung im Griff haben. Wer das Löschen nicht im Griff hat, sollte lieber nicht zu viele Daten anhäufen. Aber das

ist nur die Theorie. Die Deutsche Wohnen musste hier eines der ersten größeren deutschen DSGVO-Bußgelder erleiden, weil sie viele personenbezogene Daten angehäuft, aber nicht gelöscht hatte. Angeblicher Grund: Die eingesetzte Software ermöglichte keine automatisierte Löschung. Der entsprechende Bußgeldbescheid wurde zunächst einmal von einem Gericht aufgehoben, aber nicht deshalb, weil das Löschgebot der DSGVO doch nicht so ernst zu nehmen sei, sondern weil die Datenschutzaufsichtsbehörde aus Sicht des Gerichts das falsche Ordnungswidrigkeitenverfahren angewandt hatte. Derartige Verfahrensfehler der chronisch unterbesetzten Datenschutzaufsichtsbehörden ebenso wie die üblichen „Anwaltsscharmützel" können jedoch nicht darüber hinwegtäuschen, dass in der Sache selbst die Einschläge näher kommen.

Vor diesem Hintergrund ist es kein Geheimnis, dass die mittelständische Praxis – und teilweise selbst Großunternehmen – immer wieder nach der Formel „Wo kein Kläger, da kein Richter" verfahren, also: Wo keine Aufsichtsbehörde kommt, da muss auch nichts gelöscht werden. Man kann diesen Ansatz implizit „leben" oder sich eine Feigenblattargumentation in die Schublade legen. Wenn man in der Sache weitermacht wie bisher, muss auch niemand der Geschäftsleitung erklären, weshalb „potentiell noch für das Unternehmen wichtige Daten" gelöscht werden müssen. Der klassische Konflikt zwischen einem konservativen Datenmanagement (alles aufheben, weil man nie wissen kann, wofür es noch gut sein kann) und dem datenschutzrechtlichen Gebot der Datenminimierung und der Speicher(zeit)begrenzung wird so – wenn auch nicht offen ausgesprochen – im Sinne des ersteren gelöst. Das ist angesichts des Vollzugsdefizits des Datenschutzrechts auch kein Wunder. Um diesem Vorgehen auf die Spur zu kommen, müssten „zigtausende" zusätzliche Mitarbeiter in den Datenschutzaufsichtsbehörden mit profunden Kenntnissen der komplexen Backend-Systeme der Unternehmen eingestellt werden.

Wer sich über die Reichweite datenschutzrechtlicher Pflichten keine vertieften Gedanken macht, wird in der Praxis desto erfolgreicher sein, je weniger die betroffenen Personen wissen können, dass und welche Daten von ihnen in derartigen Systemen schlummern. Im Fall der Deutsche Wohnen waren das historische Bewerbungsunterlagen abgelehnter Mieter, in einem spanischen Fall Kontaktdaten eines Bankkunden, der schon seit vielen Jahren kein Kunde der Bank mehr gewesen war. Nur dann, wenn die Existenz von Daten, die schon längst hätten gelöscht werden müssen, irgendwie für den Betroffenen „ans Licht kommt", sich also das Unternehmen „verrät", besteht eine erhöhte Gefahr, dass sich insbesondere Betroffene darüber bei der Datenschutzaufsichtsbehörde beschweren. Die stark eingeschränkte Kontrolldichte im Datenschutzrecht lässt die Backend-Systeme weiter „blühen".

Eine hier nicht weiter zu vertiefende Frage in diesem Zusammenhang ist, ob der Geschäftsleiter eines Unternehmens im Anwendungsbereich der Compliance-Grundsätze bzw. der

Organhaftung explizit oder implizit damit hantieren darf, dass das Entdeckungsrisiko unter bestimmten Umständen minimal ist. Man wird das in der Praxis natürlich immer so wenden, dass rechtlich nicht ganz klar war, was man nun wann hätte löschen müssen. Der Umgang mit einer unklaren Rechtslage ist auch nach Jahren der Entwicklung im Compliance-Bereich immer noch unklar. Grundsätzlich darf sich das Unternehmen bzw. die Geschäftsleitung bei unklarer Rechtslage und nach Ausschöpfung aller zumutbaren Erkenntnisquellen auf einen für das Unternehmen günstigen Rechtsstandpunkt stellen. Und (höchst-)richterliche Entscheidungen zum Thema Löschen, die bei unklarer Rechtslage verbindlich das Richtige festlegen würden, gibt es bislang kaum. Im Fall der Deutsche Wohnen war das Problem jedoch nicht, dass nicht jedes Einzeldatum im Einklang mit einer genau durchstudierten Rechtslage gesetzlicher Aufbewahrungspflichten gelöscht wurde, sondern, dass die Aussage des Unternehmens gegenüber den Datenschutzbehörden offensichtlich sinngemäß lautete: „Wir löschen gar nicht, weil wir nicht löschen können“. Das taugt natürlich als Feigenblatt-Argument wenig. Besser wäre es da schon gewesen, in die Unternehmenssatzung hineinzuschreiben, dass sämtliche Daten des Unternehmens 100 Jahre aufzubewahren sind, und sich vor diesem Hintergrund auf den – europarechtlich natürlich umstrittenen – § 35 Abs. 3 BDSG zu berufen, der „satzungsgemäßen Aufbewahrungsfristen“ einen (pauschalen) Vorrang vor den datenschutzrechtlichen Löschpflichten einräumt. Da könnte man wirklich mit einer unsicheren Rechtslage argumentieren, denn Urteile zu § 35 Abs. 3 BDSG gibt es bislang nicht.

Am Anfang jeder Datenschutz-Compliance im Unternehmen muss sich die Unternehmensführung also darüber Gedanken machen, ob und inwieweit sie die Einhaltung des Datenschutzrechts insgesamt mehr als Lippenbekenntnis oder mehr als Herzensangelegenheit leben will. Natürlich wird kein Geschäftsführer offen zugeben, dass er in Kauf nimmt, gegen geltendes Recht zu verstoßen, und deshalb den Handlungsbedarf nicht umfassend bestimmen lässt. Die Praxis zeigt dennoch: Ebenso wie man höchst ungerne Versicherungen über Risiken abschließt, von deren Nicht-Materialisierung man subjektiv überzeugt ist, gibt ein – zumal mittelständisches – Unternehmen ungerne Geld dafür aus, um zukünftige Ermittlungen wegen Rechtsverstößen zu vermeiden, an deren Einleitung man gar nicht glaubt. Gute Compliance-Arbeit kostet Geld, zumal wenn am Beginn erst einmal Altlasten identifiziert und aufgeräumt werden müssen. Vor dem Hintergrund sich potenzierender Compliance-Pflichten in vielen Bereichen auch außerhalb des Datenschutzes, werden den Unternehmen hier zunehmend Kosten aufgebürdet. Entscheidet die Geschäftsleitung, dass ein „restriktiver Ansatz“ gefahren bzw. „Datenschutz mit Augenmaß“ betrieben werden soll, so mag das zwar nicht unbedingt im Sinne des Gesetzgebers sein, kann aber verhindern, dass im Unternehmen unterschiedliche Ansätze (mit der Folge noch höherer Aufwände) kollidieren.

Praxistipp: Am Anfang ist weniger Datenschutz besser als gar kein Datenschutz. Die Praxis mittelständischer Unternehmen zeigt deutlich, dass der notwendige Umfang der von der DSGVO geforderten Anstrengungen zum Schutz personenbezogener Daten – und damit auch genaue Überlegungen zum Löschen nicht mehr erforderlicher Daten – nur Stück für Stück in das Bewusstsein der Verantwortlichen vordringt. Durch mehr mediale Aufmerksamkeit, durch „spektakuläre" Bußgeldfälle, durch die wachsende Sensibilität gegenüber Datenschutzrisiken im Bereich der privaten Daten, durch datenschutzrechtliche Compliance-Klauseln in Dienstleistungs- und Beschaffungsverträgen, durch das Verhalten anderer Unternehmen im Umfeld, durch neue Mitarbeiter, durch vermehrte Ansprache aus dem Kreis der Beschäftigten und durch viele weitere Einzelumstände wird sich zunehmend die Erkenntnis – auch im Management – durchsetzen, wie viel zum Schutz personenbezogener Daten im eigenen Haus getan werden muss. Man kann dies als „Kulturwandel zum Datenschutz" bezeichnen, was auch die Einsicht beinhaltet, dass Datenschutz Aufwand produziert, aber auch positive Außenwirkung. Hier gilt es, die Erkenntnis der wachsenden Bedeutung („An der Stelle müssen wir mehr tun als bisher") in die richtigen Bahnen zu lenken, d. h. die richtigen Schlüsse und Einzelmaßnahmen abzuleiten, auch und gerade bei der Bestandsaufnahme und dem Löschen nicht mehr erforderlicher Daten.

Volle Kontrolle über Daten ist gar nicht so einfach

Das Datenschutzrecht, genauer gesagt viele Datenschutz-Apologeten in Parlamenten, Verwaltungen und Gremien, fordert hundertprozentige Kontrolle des Verantwortlichen über alle von ihm verarbeiteten personenbezogenen Daten. Es fordert dies in Bezug auf IT-Verantwortliche, Programmierer, Fachabteilungen, Administratoren, kurz von Beschäftigten in Unternehmen, die auch viele andere Aufgaben haben – nämlich insbesondere den Betrieb am Laufen zu halten und möglichst hohen Umsatz mit zu generieren und/oder die Kosten niedrig zu halten. Aber halt: Sind nicht auch Datenschutzbeauftragte, Compliance-Organisationen und Rechtsabteilungen gefordert? Eigentlich schon, aber diese sind nicht die operativen Herren der Daten in den Unternehmen und verheddern sich leicht in gegenseitigen Abhängigkeiten von IT-Systemen, im (Un-)Verständnis technischer Notwendigkeiten und nicht zuletzt in Datenbank-Feldbezeichnungen (wenn sie sich überhaupt dorthin verirren). Kurz: Sie tun sich schwer damit, überhaupt herauszufinden, „was der Fall ist".

Damit kommen wir zu einem wichtigen Grund, warum es in der Praxis schwierig ist, das Backend datenschutzkonform zu gestalten. Wer einmal eine Compliance-Richtlinie eines großen Unternehmens gelesen hat, die (auch) auf das Thema Datenschutz im Backend anwendbar ist, bemerkt schnell, worum es hier in erster Linie geht: Technikferne Experten aus der Compliance-Abteilung verlagern die „Subsumtionslast" auf Fachabteilungen, die häufig nicht einmal wissen, was Subsumtion überhaupt bedeutet. Das kulminiert dann in

dem Satz einer Compliance-Richtlinie (nach vielen Definitionen und Gesetzeswiederholungen), dass die – chronisch überlasteten – technischen und organisatorischen Verantwortlichen für die einzelnen IT-Systeme das Datenschutzrecht in „ihrem" IT-System umzusetzen haben. Die „Guidance" der Datenschutzbehörden ist da wenig besser, weil sich natürlich keine Aufsichtsbehörde in diesem „großtechnischen Minenfeld" vorwerfen lassen will, sie hätte konkret dies oder das empfohlen, und dann führt das zu einer Praxis, die man so gar nicht wollte. Da bleibt man lieber im Ungefähren. Ein Paradebeispiel ist der Baustein 60 des Standard-Datenschutzmodells („Löschen und Vernichten"), der sagt, was ein Unternehmen tun soll, aber letztlich doch nicht sagt, was ein Unternehmen wirklich tun soll. Im Ergebnis wissen die, die tatsächlich die Systeme „zum richtigen Löschen befähigen" sollen, nicht, was von ihnen genau verlangt wird.

Selbst wenn alle mustergültig „an einem Strang" – und in dieselbe Richtung – ziehen, ist ein Löschkonzept zwangsläufig das Ergebnis vieler (hoffentlich bewusster) Weichenstellungen, die aufgrund von technischen Notwendigkeiten, Budgetgrenzen, organisatorischen Gegebenheiten, Entscheidungen über rechtliche Sichtweisen etc. getroffen wurden, und diese müssen transparent gemacht werden. Ein Löschkonzept, das aus Sicht des später Prüfenden (Auditors) falsche Weichenstellungen beinhaltet, ist besser als ein Löschkonzept, das diese Weichenstellungen nicht benennt oder nicht erklärt, warum die Weiche wie gestellt wurde (denn das lässt das Löschkonzept als etwas „Beliebiges" erscheinen).

Praxistipp: Jedes Projekt, mit dem ein Löschkonzept erstellt wird, sollte zum Ziel haben, für einen Dritten nachvollziehbar das Löschkonzept selbst zu beschreiben und – ebenso wichtig – warum es so ausgefallen ist, wie es ausgefallen ist. Dies zwingt einerseits die Verfasser, sich genügend Gedanken über Begründungen zu machen, und zeigt andererseits dem Leser, dass nicht „einfach so" ein Muster abgeschrieben oder Löschfristen „ausgewürfelt" wurden.

Was sind personenbezogene Daten?

Bisweilen führt die Erstellung eines Löschkonzepts dazu, dass juristische Problemfelder an ganz anderen Stellen der „Prozesskette" bzw. des „Life Cycles" von personenbezogenen Daten in den Blickpunkt geraten. So haben sich viele Unternehmen zu der Zeit, als sie Daten in IT-Systeme importiert haben, keine Gedanken darüber gemacht, welche davon eigentlich personenbezogen sind, und müssen dies anlässlich der Arbeit am Löschkonzept nachholen. Über diese Frage der Einordnung unter den Gesetzestext (Subsumtion) haben zwar schon viele kluge Juristen viele kluge Texte geschrieben auf großer Abstraktionshöhe, die entweder von anderen Juristen gelesen und verstanden – und oft auch widerlegt – werden oder von Nichtjuristen gelesen und nicht verstanden werden. Kurzer Exkurs dazu aus Wikipedia: *„Die logische Struktur der Subsumtion eines konkreten Falles unter*

die Begriffe einer Rechtsnorm ist nicht unproblematisch. Denn im streng logischen Sinn kann nur ein Begriff unter einen Begriff subsumiert werden. Nach Karl Engisch kann die Subsumtion eines konkreten Sachverhalts unter einen Begriff nur als Einordnung des Sachverhalts in die Klasse der durch den Rechtssatz bezeichneten Fälle gedeutet werden. Es gehe dabei um die Gleichsetzung des Falles mit denjenigen Fällen, deren Zugehörigkeit zu der Klasse bereits feststeht." Der kluge Jurist Karl Engisch wusste es: Dies ist der – sogar von Juristen oft vergessene – Grund dafür, dass man eine Rechtslage nicht „ausrechnen" kann. Es geht immer darum, ob sich Fälle inhaltlich ähneln, und Ähnlichkeit ist kein mathematischer Begriff, sondern etwas, das man „so oder so" sehen kann.

Man stelle sich also nun den Systemverantwortlichen für ein SAP-System einer größeren mittelständischen Unternehmensgruppe vor, die u. a. Finanzdienstleistungen erbringt und in diesem Kontext Konten für Kunden führt. Die Kunden sind fast immer – aber nicht immer – Gesellschaften. Sie verfügen also über dahinterstehende wirtschaftlich Berechtigte, die vom Unternehmen aus geldwäscherechtlichen Gründen zu erfassen sind. Daneben werden die Personen im System erfasst, die den Kunden gegenüber dem Unternehmen vertreten dürfen. Diese Personen wechseln aufgrund der üblichen Fluktuationen. Sämtliche Änderungen im IT-System werden aufgezeichnet mit dem ursprünglichen Datum, dem neuen Datum und der Kennung des Mitarbeiters, der die Änderung einpflegt.

Schon befindet man sich im datenschutzrechtlichen Dickicht, obwohl es hier nur darum gehen soll, ein Löschkonzept aufzustellen. Nur ein Detailthema zur Illustration: Die Daten über Kontoinhaber, die juristische Personen (z. B. GmbHs) sind, sind „eigentlich" keine personenbezogenen Daten. Was aber, wenn sich aus dem Kontostand eine wesentliche Information über den Mehrheits- oder Alleingesellschafter ablesen lässt? Liegt dann ein personenbezogenes Datum vor oder ist das davon abhängig, ob der Verantwortliche über die Zusatzinformation verfügt, dass es sich bei dem Kontostand der juristischen Person um eine ganz wesentliche Information im Gesamtvermögensstatus der natürlichen Person „hinter" der juristischen Person handelt und damit eine relevante Aussage über die wirtschaftlichen Verhältnisse dieser natürlichen Person getroffen werden kann? Und verfügt ein Finanzdienstleister aufgrund der geldwäscherechtlichen Identifizierungspflichten und aufgrund von Bonitäts-Selbstauskünften etc. nicht regelmäßig über diese Zusatzinformation?

Man sieht: Es ist nicht so einfach zu entscheiden, welche Daten hier nun genau Personenbezug aufweisen und wo die Grenzen liegen. Wenn ein Jurist nun sagt, dies sei eine Frage der „besonderen Umstände des Einzelfalles", dann hat er nicht verstanden, dass damit weder eine Fachabteilung noch eine IT-Abteilung, die auf der Suche nach einem Löschkonzept für personenbezogene Daten in großen IT-Systemen ist, etwas anfangen kann.

Dies alles ist Teil des ersten großen Puzzles, das gelöst werden muss, bevor man ein Löschkonzept erstellen kann: Welche personenbezogenen Daten liegen denn in der Datenbank? Es gibt dazu den schönen Begriff der Datenlandkarte, die eigentlich jeder, der auch nur Datenmanagement betreibt, vor sich liegen haben sollte. Dennoch haben die wenigsten Unternehmen eine einigermaßen detaillierte und aktuelle Datenlandkarte, sondern eher eine vage Vorstellung davon, was auf ihren Servern schlummert. Das mag auch daran liegen, dass sich die Inhalte der Datenbank sekündlich ändern und auch Änderungen der Struktur oft im Wochen- oder Monatsrhythmus angestoßen werden. Im obengenannten Beispielsfall würde eine Datenlandkarte auflisten, welche Daten zu den einzelnen Personengruppen vorliegen, also zu den Kunden, den wirtschaftlich Berechtigten, den Verfügungsberechtigten und vielleicht weiteren Personengruppen wie den Vermittlern des Geschäfts. In einer guten Datenlandkarte könnte man sehen, in welchen Datenbankbereichen diese Daten liegen, vielleicht sogar im Sinne einer Datenflussanalyse, aus welchen Systemen sie stammen und in welche anderen Systeme sie fließen. Wer „Herr der Daten" sein will, sollte also – wie ein König hinsichtlich seines Volkes – hin und wieder eine Volkszählung machen.

Praxistipp: Gerade in Systemen, in denen bislang nie etwas gelöscht wurde, kann es einfacher sein – wenngleich das auf den ersten Blick aufwendiger erscheint –, die noch relevanten Daten in eine neue Datenbankinstanz zu überführen, als uralte, gewachsene Datenbestände, die niemand im Unternehmen mehr kennt, mühselig zu analysieren und punktuell zu bereinigen.

Löschen oder Anonymisieren?

Auch wenn die Umsetzung in der Praxis nicht immer trivial ist, kann es eine Alternative zu einer Löschung von Daten sein, diese „nur" zu anonymisieren. Zwar geht die Mehrheit der Juristen (und auch die österreichische Datenschutzbehörde) im Grundsatz davon aus, dass ein Löschen auch durch Anonymisierung umgesetzt werden kann, aber im Detail ist das rechtlich womöglich nicht so einfach. Dabei liegt das Problem nicht darin, dass nach Ansicht der meisten Juristen das Anonymisieren – ebenso wie das Löschen – eine Verarbeitungstätigkeit ist, für die es eine datenschutzrechtliche Legitimationsgrundlage geben muss. Denn diese Legitimationsgrundlage ist beim „Löschen durch Anonymisieren" ein „Gesetzesbefehl" (Art. 6 Abs. 1 S. 1 lit. c) DSGVO) in Gestalt der Löschverpflichtung (Art. 17 DSGVO). Es mag sein, dass mit dieser Erkenntnis die rechtlichen Themen bereits abgehandelt sind, aber möglicherweise fangen sie hier auch gerade erst an. So genau weiß das im Moment niemand.

Nach einem Positionspapier des Bundesdatenschutzbeauftragten gibt es neben der Anonymisierung zur Löschung (auf der eben genannten Legitimationsbasis) auch eine Anonymisierung aus anderen Gründen. Letztere kann z. B. aufgrund eines „weitergedachten" ursprünglichen Erhebungszwecks gerechtfertigt sein (Zweckänderung), etwa im Falle statistischer (Big-Data-) Analysen. Diese Zweiteilung berücksichtigt allerdings nicht, dass in beiden Fällen sowohl der technische Vorgang als auch das Ergebnis dasselbe sind: Personenbezogene Daten werden ihres Personenbezuges entkleidet und fallen danach nicht mehr unter das Datenschutzrecht, sie sind datenschutzrechtlich „weg". Vor diesem Hintergrund bleibt fraglich, wie man diese beiden Anonymisierungsfälle (bzw. Anonymisierungszwecke) überhaupt sinnvoll voneinander abgrenzen kann. Der Bundesdatenschutzbeauftragte sieht den Unterschied darin, dass in einem Fall „der datenschutzrechtlich relevante Zweck der Anonymisierung nicht die Aufhebung des Personenbezugs ist, sondern das dahinter stehende tatsächliche Interesse des Verantwortlichen", also der Folgezweck. Aber auch ein Verantwortlicher, der Daten durch Anonymisierung löschen möchte, kann danach noch etwas „mit den Daten vorhaben", also einen Folgezweck beabsichtigen. Allenfalls könnte man faktisch unterscheiden zwischen dem Fall, in dem die ursprünglichen Daten anonymisiert und dadurch die Ausgangsdaten endgültig gelöscht werden und dem Fall, in dem die ursprünglichen Daten neben der „anonymisierten Fassung" weiter als personenbezogene Daten verarbeitet, also die anonymisierten Daten nur gleichsam „abgezweigt" werden.

Was diese Unterscheidung für einen Nutzen und was für Folgen hat, ist unklar. Der Bundesdatenschutzbeauftragte verlangt im Falle einer Anonymisierung in aller Regel eine aufwändige Datenschutz-Folgenabschätzung, in deren Rahmen auch die weitere Verarbeitung der anonymisierten Daten mit einbezogen werden muss. Es ist offen, ob damit auch die „Anonymisierung als Löschersatz" gemeint ist. Wenn man von der Notwendigkeit einer Datenschutz-Folgenabschätzung in jedem Fall der Anonymisierung ausgehen wollte – der Bundesdatenschutzbeauftragte geht davon aus, dass hier regelmäßig die Gründe „Verarbeitung in großem Umfang" und/oder „neue Technologien" gegeben sind –, dann wird die Vorbereitung solcher Anonymisierungsvorgänge eine aufwendige Angelegenheit. Auch ist offen, wie eine spätere Löschung durch Anonymisierung in den Pflichthinweisen, die dem Betroffenen bereits bei der Erhebung zur Verfügung gestellt werden müssen, zu reflektieren sind (bzw. ob eine Zweckänderungsnachricht an den Betroffenen gehen muss). Dieses Thema ist durchaus ernst zu nehmen, denn es gibt genügend Juristen, die davon ausgehen, dass personenbezogene Daten zu jeder Zeit rechtswidrig verarbeitet werden, wenn bereits die Pflichthinweise bei der Erhebung unrichtig bzw. unvollständig waren.

Im schlimmsten Fall kommt ein Gericht also dereinst zu dem Schluss, dass die anonymisierten und „eigentlich" gelöschten Daten vorab daraufhin geprüft (und vielleicht sogar

später daraufhin überwacht) werden müssen, was mit diesen Daten nach der Anonymisierung passiert, obwohl die „Welt der anonymen Daten“ vom Datenschutz eigentlich gar nicht mehr erfasst wird und unklar ist, auf welcher Grundlage eine spätere missbräuchliche Verwendung der anonymisierten Daten überhaupt geahndet werden könnte. Nach dem Grundsatz „Nachher ist man immer schlauer“ könnte ein Gericht die später hinzugetretenen Nutzungshandlungen an den anonymen Daten als von vornherein mit „berücksichtigungspflichtig“ sehen. Nimmt man diese Überlegungen ernst, so ist die Alternative Löschen oder Anonymisieren nicht ganz so alternativ, wie man dies gerne hätte. Auch dies zeigt, wie schwierig – um nicht zu sagen unhandlich – das Datenschutzrecht bei der Planung von Löschkonzepten ist.

Der Hauptschwerpunkt bei der Beantwortung der tatsächlichen bzw. Informatik-Fragestellung, ob Anonymisieren als „Löschersatz“ in Betracht gezogen werden kann, liegt aber beim Begriff der Anonymisierung selbst. Wer Namen und Geburtsdaten aus einem Dokument löscht, aber stehenlässt, dass der Betroffene „von 2013 bis 2018 Leiter der IT-Abteilung war“, der anonymisiert nicht. Sogar der Schreibstil eines Betroffenen kann hinreichend wahrscheinliche Rückschlüsse auf dessen Identität zulassen, sodass dann, wenn nur wenige Personen möglicherweise der Betroffene sind, die Wahl zwischen diesen nicht schwerfällt. Hierzu gibt es mittlerweile viel theoretische Literatur und keine klare Linie. In der Praxis kann man sich daher kaum je wirklich sicher sein, massenhafte Datenbankdaten tatsächlich anonymisiert zu haben, was die Anonymisierung als „Löschersatz“ nachhaltig in Frage stellt.

Praxistipp: Wer am Ende der Laufzeit von Daten (statt löschen) anonymisieren möchte, sollte sich überlegen, dies bereits von Anfang an im Rahmen der Erhebung transparent zu machen. Salopp gesagt kann das, was man mit Daten machen kann, dadurch (natürlich in bestimmten Grenzen) erweitert werden, indem man es dem Betroffenen von vornherein klar und deutlich kommuniziert.

Wann muss gelöscht werden?

Das Ziel eines Löschkonzepts ist es, einerseits das Datenschutzrecht durch rechtzeitiges – also nicht zu frühes und nicht zu spätes – Löschen einzuhalten, andererseits aber auch die Herleitung und Vorgehensweise dieses Löschkonzepts ausreichend, nachvollziehbar und auf aktuellem Stand zu dokumentieren. Was „rechtzeitig“ ist, ist vermeintlich leicht zu erklären: Sobald ein personenbezogenes Datum nicht mehr für den Erhebungszweck erforderlich ist, muss es vorbehaltlich gesetzlicher Aufbewahrungsfristen gelöscht werden. Eigentlich sollte das eine einfache und leicht anwendbare Regel sein. Aber wie immer steckt der Teufel im Detail. Es ist keine kleine Aufgabe, die genannte, sehr abstrakt formu-

lierte, auf Art. 17 DSGVO beruhende „Löschregel“ in ein konkretes, automatisiert ablaufendes (Datenbank-) „Programm“ aus vorrangig technischen, nachrangig organisatorischen Maßnahmen zu übersetzen.

Wer sich dem Thema Löschkonzept isoliert nähert bzw. sogar den Datenschutz vorrangig „vom Löschkonzept her denkt“, wird sich nicht unbedingt zuerst mit der Herkunft der Daten befassen, sondern diese sind „schon da“, warum auch immer. Aber das ist kein gelungener Startpunkt für ein Löschkonzept. Die datenschutzrechtlichen Vorfragen ergeben sich weit weg vom Löschzeitpunkt schon im Zusammenhang mit der Erhebung und der produktiven Verwendung von Daten. So ist der früheste Zeitpunkt des Löschens das Ende des ursprünglichen Erhebungszwecks, der bisweilen aus der Perspektive des Löschens anders verstanden werden kann als aus der isolierten Sicht der Erhebung selbst. Wer Daten erhebt, um einen Kaufvertrag mit dem Betroffenen abzuschließen und anschließend zu erfüllen, wird sich zurecht fragen, ob dieser ursprüngliche Erhebungszweck nicht auch nach der eigentlichen Erfüllung der Hauptpflichten noch die Gewährleistungsphase oder die Weitergabe der Kaufvertragsdaten an den Jahresabschlussprüfer des verantwortlichen Unternehmens mit abdeckt.

Das Ende des ursprünglichen Erhebungszwecks führt dazu, dass die Daten eine Zweckänderung in Richtung Aufbewahrung erfahren. Über die sich an die „Produktivphase“ anschließenden Aufbewahrungszwecke und deren Dauer ist der Betroffene im Grundsatz schon bei der Erhebung zu informieren, und zwar auch dann, wenn „nur“ gesetzliche Aufbewahrungsfristen bestehen. Deshalb muss ein Löschkonzept eigentlich schon stehen, bevor die ersten Daten erhoben werden, weil sonst die notwendigen Bestandteile für die Pflichtinformationen gegenüber dem Betroffenen nicht vorliegen.

Auch wenn ein Löschkonzept wesentlich mehr umfasst als nur eine Auflistung gesetzlicher Aufbewahrungspflichten, ist doch die genaue Analyse, welche Daten wie lange unter welche gesetzlichen Regelungen fallen, nicht trivial. Das gilt beispielsweise für die „Lieblings-Aufbewahrungsnorm“ der Unternehmen, § 147 AO, deren Anwendungsbereich gerade von Steuerabteilungen extrem weit veranschlagt wird, um einem Betriebsprüfer nicht vor den Kopf zu stoßen. Nicht nur im Bereich der E-Mail-Archivierung, um den es hier nicht geht, sondern auch und gerade in komplexen Datenbanksystemen sind natürlich jede Menge „steuerrelevante“ Daten, die unterschiedlichen Aufbewahrungsvorschriften (insbesondere aber dem genannten § 147 AO) unterliegen. Die genauen Grenzen zwischen steuerlich aufbewahrungspflichtigen und nicht aufbewahrungspflichtigen Daten sind bislang in den Randbereichen wenig erforscht. War eine vorsorglich umfängliche und lange Aufbewahrung vor Geltung der DSGVO geübte und unbeanstandete Praxis, sieht die DSGVO – jedenfalls in der Theorie – durchaus kritischer auf die genaue Abgrenzung zwi-

schen steuerlichen „Allaufbewahrungstendenzen“ und datenschutzrechtlicher „Speicherbegrenzung“. Hier liegt ein Minenfeld jedes Löschkonzepts mit vielen Weichenstellungen im Detail, die im Einzelfall zu entscheiden und zu dokumentieren sind.

Ob man die „Aufbewahrungsphase“ IT-technisch durch eine Verlagerung der Daten in ein Archivsystem umsetzt, ist datenschutzrechtlich nicht vorgegeben. Wichtig ist vielmehr, dass die veränderte – verengte – Zweckbindung umgesetzt wird. Werden etwa Daten nur noch aufgrund steuerrechtlicher Aufbewahrungspflichten weiter gespeichert, dürfen auch nur diejenigen im Unternehmen darauf zugreifen, die sich mit Themen wie Steuererklärung und Betriebsprüfung beschäftigen – und natürlich der Betriebsprüfer des Finanzamts. Man sieht, dass das Löschkonzept und das Berechtigungskonzept hier aufeinander abgestimmt sein müssen.

Praxistipp: Das Löschkonzept sollte nicht nur Aufbewahrungsfristen bzw. Löschzeitpunkte vorsehen, sondern auch den Zweck der Aufbewahrung im Blick haben. Am Anfang der Aufbewahrungsphase können Daten noch aus mehreren Gründen aufbewahrt werden, die dann zum Ende der Aufbewahrungsphase hin immer weiter wegfallen. Entsprechend dürfen am Anfang der Aufbewahrungsphase mehr Mitarbeiter auf die Daten zugreifen als zum Ende hin.

Aufbewahrungsinteressen als Aufbewahrungsgrund?

Ähnlich wie die Unterscheidung zwischen Datenmanagement und Datenschutz gibt es auch die Unterscheidung zwischen Aufbewahrungsinteressen und Aufbewahrungspflichten. Liegt ein Aufbewahrungsinteresse vor, so möchte der Verantwortliche nicht löschen, liegt eine Aufbewahrungspflicht vor, so darf der Verantwortliche nicht löschen. Ungünstig nur, dass die DSGVO im Grundsatz lediglich Aufbewahrungspflichten anerkennt und Aufbewahrungsinteressen datenschutzrechtlich nicht so einfach zu begründen sind.

Der einfachste Fall betrifft die Frage, ob die Daten nicht „wenigstens“ für die Zeit aufgehoben werden dürfen, in der der Verantwortliche möglicherweise noch Partei eines (Rechts-)Streits werden könnte. Denn ein Löschen von Beweismitteln vor dem Ende von Verjährungsfristen führt dazu, dass man als Kläger nicht mehr viel vortragen kann und als Beklagter nicht mehr viel Verteidigungsmittel in der Hand hat. Paradoxerweise gibt die DSGVO auf diese wichtige Frage keine klare Antwort. Die Antwort der Datenschutzbehörden lautet, dass „eigentlich“ ein Rechtsstreit konkret im Einzelfall drohen muss, damit die Daten, die in diesem (konkreten) Rechtsstreit eine Rolle spielen könnten (?), noch weiter aufbewahrt werden dürfen. Alleine das theoretische Risiko, sich mit jemandem noch einmal streiten zu müssen – also noch nicht abgelaufene Gewährleistungs- bzw. Verjährungsfristen –, rechtfertigt nicht die „Massenaufbewahrung“ aller potentiell einschlägigen

Daten. Spricht man dann hinter vorgehaltener Hand, dass man doch die Verantwortlichen nicht rechtlos (bzw. beweislos) stellen darf, weil man sie damit zur Selbstschädigung zwingt, und dass es doch „nur" um drei Jahre Regelverjährungsfrist nach dem Ende des maßgeblichen Geschäftsjahres geht, nicken die Mitarbeiter der Datenschutzbehörden meist mit dem Kopf und sagen, sie hätten jetzt eine zehn- oder gar dreißigjährige Aufbewahrung befürchtet. Das lässt aber Zweifel entstehen, ob das Gesetz das Maß aller Dinge ist oder irgendein „Augenmaß", weil „die DSGVO das ja nicht so strikt gemeint haben kann". Richtig ist: Die DSGVO hat überhaupt nichts „gemeint", denn niemand hat sich vor dem Erlass der DSGVO mit solchen Fragen konkret beschäftigt. Dennoch bleibt die Erkenntnis, dass im Kontext von Verjährungsfristen zumindest zur Vermeidung von Restrisiken jede Aufbewahrung außerhalb konkret bevorstehender Streitigkeiten (also über Art. 17 Abs. 3 lit. e) DSGVO hinaus) wohl der Legitimation durch ein – genauer zu begründendes – Aufbewahrungsinteresse bedarf.

Ein (anderes) Aufbewahrungsinteresse sind organisatorische bzw. betriebliche Belange des Verantwortlichen. Wer sein Unternehmen so strukturiert, dass er bestimmte Daten für bestimmte Zwecke benötigt, hat ein berechtigtes Interesse, die Daten solange aufzubewahren, wie es für diese Belange erforderlich ist. Doch die Leitlinie ist hier weniger die DSGVO, deren Formulierungen ja oft dort aufhören, wo es interessant wird, sondern eher der gesunde Menschenverstand oder das „Rechtsgefühl", also ein sehr unpräziser Maßstab. Damit lassen sich große Aufbewahrungszeiträume nicht rechtfertigen. Überhaupt ist die Frage, welche an den Primärzweck anschließenden Aufbewahrungsfristen noch durch ein betriebliches Interesse gerechtfertigt werden können, nur sehr unbefriedigend durchdrungen. Die notwendigen Maßstäbe für die Abwägung nach Art. 6 Abs. 1 S. 1 lit. f DSGVO fehlen völlig. Es wäre dann auch immer zu befürchten, dass der Betroffene widerspricht und die weitere Aufbewahrung im Einzelfall geprüft werden muss, was jedes Archivsystem durcheinanderbringt. An die Pflicht, das Aufbewahrungsinteresse schon bei der Erhebung zu benennen oder später eine Zweckänderungsnachricht an den Betroffenen abzusetzen, möchte man sich an dieser Stelle schon gar nicht mehr erinnern.

Praxistipp: Jedes Unternehmen sollte sich frühzeitig, also eigentlich vor der Erhebung von Daten, darüber Gedanken machen, welche Daten aus welchen (betrieblichen) Gründen außerhalb von gesetzlichen Aufbewahrungsfristen und dem Ende des Primärzwecks noch benötigt werden. Die datenschutzrechtliche Legitimation bzw. Interessenabwägung ist zu dokumentieren und die Rückkopplung im Hinblick auf die Pflichtinformationen für den Betroffenen ist zu beachten.

Praktische Umsetzungsprobleme in IT-Systemen

Das Oberlandesgericht Köln hat in einer Entscheidung in einem anderen datenschutzrechtlichen Kontext (nämlich zum Auskunftsrecht des Betroffenen) geurteilt, dass es – zumindest in Bezug auf die Strukturierung von IT-Systemen – die Sache des Verantwortlichen ist, seinen Betrieb so auszugestalten, dass die rechtlichen Vorgaben eingehalten werden. Das klingt zunächst so, als müsse man jegliche Ressourcen einsetzen, um die DSGVO umzusetzen, und das Argument wirtschaftlicher Unverhältnismäßigkeit bzw. Unmöglichkeit sei unzulässig. Ob auch andere Gerichte bei der Linie bleiben, wird sich zeigen. Im Grundsatz besteht aber schon aufgrund der „privacy by design"-Anforderungen der DSGVO Einvernehmen darüber, dass datenschutzrechtswidrige Nichtlöschungen kaum mit „unfähiger Technik" begründet werden können. Es gibt viele Fälle, in denen Software nach dem Datenschutzrecht gebotene Maßnahmen nicht oder nur sehr „holprig" umsetzen kann. Das mag an veraltetem Design oder an einer gewissen Renitenz des Herstellers liegen, der ja selbst insoweit nicht zu den Verpflichteten der DSGVO zählt. Meist ist die Anpassung von Altsoftware auch schwierig, wie jüngst ein IT-Spezialist auf golem.de erklärte: *„Wenn eine Software neue Anforderungen erfüllen soll, scheuen viele Firmen aus Angst vor Bugs tiefgreifende Code-Änderungen. Was dabei herauskommt ist aber schlimmer als Bugs: erstarrte, nutzlose Legacy-Systeme*". Man könnte an erstarrte und nutzlose noch „datenschutzwidrige" anfügen.

Kann eine Software weder verbessert (durch individuelle Änderungen seitens des Verantwortlichen oder durch Druck auf den Hersteller) noch ausgetauscht werden, müssen ersatzweise organisatorische Maßnahmen definiert werden. Man könnte zwar meinen, dass es datenschutzrechtlich irrelevant ist, ob ein Programm zeitgesteuert auf den „Löschen-Knopf" drückt oder ein Mensch, der nach einer entsprechenden Handlungsanweisung vorgeht. Aber die europäischen Datenschutzbehörden gehen (nachvollziehbarerweise) davon aus, dass eine technische, automatisierte Lösung einer Umsetzung durch Menschen vorzuziehen ist. Löschkonzepte sollten also nicht von vornherein nur auf Anweisungen an Mitarbeiter beruhen.

Komplexe IT-Landschaften bestehen aus verschiedenen datenhaltenden Systemen, die miteinander regelmäßig Daten austauschen. Geschieht dies über die Unternehmensgrenzen hinweg, kann dies datenschutzrechtlich eine Auftragsverarbeitung oder eine Übermittlung von Daten an einen gemeinsam Verantwortlichen oder an einen selbstständigen Verantwortlichen – beispielsweise die Steuerbehörden – darstellen. Geschieht dies unternehmensintern, so wird die Übermittlung an andere Systeme meist anderen Abteilungen Zugriffsmöglichkeiten eröffnen. Um die hieraus resultierenden „need to know"-Fragen geht es hier aber nicht, sondern darum, dass sich nach einer solchen „Distribution" von

Daten in den unterschiedlichen IT-Systemen viele Kopien derselben Daten befinden können. Dies wirft die Frage auf, ob Löschzeitpunkte „vorkommensspezifisch“ oder „inhaltsspezifisch“ zu bestimmen sind – löscht man das einzelne konkrete Datum in einem konkreten System oder das Datum insgesamt, gleich, wo (in welchem System) und wie oft es im Unternehmen vorkommt? Letzteres würde ja bedeuten, dass ein einheitliches Löschkonzept für sämtliche Systeme für dasselbe inhaltliche Datum zum selben Löschzeitpunkt gelangt oder sich die System untereinander bezüglich der Löschung synchronisieren. Solche Lösch-Synchronisationsprozesse zwischen verschiedenen IT-Systemen gehen meist mit erheblichem „Bastelaufwand“ einher, weil es hierzu keine genormten Schnittstellen bzw. Protokolle gibt.

Die Art und Weise der Speicherung bzw. Indexierung von Daten variiert von System zu System. In den meisten Fällen werden Daten in Datenbankstrukturen gespeichert, die aber unterschiedlichen Organisationskriterien unterliegen. Auch die technischen Löschmöglichkeiten variieren erheblich. So können bestimmte Datenfelder zu einem „Löschobjekt“ zusammengefasst sein. In einem System kann ein spezifisches Datum gelöscht werden, im anderen System führt die Löschung desselben (inhaltlichen) Datums zu system- bzw. datenmodellbedingten Inkonsistenzen. Auch wenn dies aus datenschutzrechtlicher Perspektive eigentlich unerheblich ist – die IT-Systeme haben sich dem Datenschutzrecht zu fügen, nicht das Datenschutzrecht den IT-Systemen –, sieht doch zumindest die Umsetzung in jedem System anders aus. Hinzu kommt, dass selbst Systeme, die eine sehr flexible Definition und Steuerung von Löschzeitpunkten erlauben, die Vielzahl der historisch gewachsenen gesetzlichen Aufbewahrungs-, Auskunfts- und Vorlagepflichten kaum abbilden können. Wenn etwa eine Aufbewahrungsfrist an den Zeitpunkt der letzten sozialversicherungsrechtlichen Prüfung anknüpft, wird dieses Datum nicht ohne Weiteres im System vorhanden bzw. im Rahmen der Fristdefinition referenzierbar sein. Noch schwieriger wird es, wenn das Ende eines Rechtsverhältnisses oder der letzte „Kontakt“ den Beginn der Frist markieren sollen. In der Praxis behilft man sich in diesen Fällen mangels Verfügbarkeit der entsprechenden „Metadaten“ zum Löschzeitpunkt mit Pauschalierungen. Die dabei entstehenden Fragestellungen werden als nächstes behandelt.

Praxistipp: Grundsätzlich sollten Leitlinien unternehmensweit und anhand der Unternehmenswirklichkeit vorgeben, innerhalb welcher Fristen welche Datenkategorien zu löschen sind, wer für die Umsetzung in welchem System verantwortlich ist und welche (Ermessens-) Spielräume im Einzelnen bestehen. Die Umsetzung im konkreten System ist aber individuell anzugehen und zu beschreiben. Das klingt allerdings einfacher, als es ist.

Risikoabhängigkeit des Löschkonzepts?

Das Kernproblem der DSGVO ist die Granularität. Wie granular „denkt" man die DSGVO? Wenn man Datenschutz auf „atomarer Ebene" angeht, wird man nie fertig. Beispiel Personalakte: Wer in der Personalakte „ein" einziges personenbezogenes Datum sieht und dafür „eine" Löschfrist vorsieht, ist schnell fertig. Wer sich hingegen darüber Gedanken macht, ob der Praxisstempel eines Arztes bei einer Arbeitsunfähigkeitsbescheinigung zu einem anderen Zeitpunkt (vorab) geschwärzt werden muss als der Rest der Arbeitsunfähigkeitsbescheinigung, wird für ein Löschkonzept sehr viel Zeit aufwenden müssen. Dasselbe Problem stellt sich bei Datenbanken, die in Tabellen, Infotypen, Einzelfelder, Subtypen etc. gegliedert sind. Auf welcher dieser Ebene man Löschkonzepte entwirft und umsetzt, ist natürlich auch eine Frage des Aufwands, der betrieben werden kann bzw. soll.

Die Gretchenfrage der Granularität gilt aber auch für den Löschzeitpunkt selbst. Die DIN 66398 zu Löschkonzepten sieht beispielsweise die Definition verschiedener Löschzeitpunkt und Löschfristen (zu „Löschklassen") vor und erlaubt in diesem Rahmen eine „Clusterung". Wie bei jeder Rasterung wird dies dazu führen, dass für manche Daten ein früherer oder späterer Löschzeitpunkt angenommen wird als derjenige, der sich bei einer genaueren Einzelfallbetrachtung ergeben mag. Das richtige Datum wird also immer mal wieder „durchs Raster fallen". Wenn hier verschiedene Empfehlungen für mittelgroße Unternehmen zwischen 4 und 7 unterschiedliche Löschklassen annehmen, mag das eine gute Pauschalisierung sein oder auch nicht; die wirkliche Vielfalt der unterschiedlichen gesetzlichen Fristen, etwa im Bereich der Beschäftigtendaten, dürfte das nicht abdecken.
Dies führt zur Grundsatzfrage jedes Löschkonzepts, ob bzw. inwieweit dieses risikoabhängig bestimmt werden darf bzw. muss. Bei „Hochrisikodaten" wäre der Löschzeitpunkt dann sehr genau zu bestimmen (bzw. die Notwendigkeit des Löschens müsste häufig geprüft werden) und die Daten wären sehr selektiv zu löschen. Bei „Normalrisikodaten" oder bei „Niedrigrisikodaten" dürften hingegen auch mal mehr Daten in einen Topf geworfen, d. h. derselben pauschalen Löschfrist unterworfen werden. Wenn man so vorgeht, würde am Anfang jeden Löschkonzepts eine Risikoeinschätzung über sämtliche personenbezogenen Daten, um die es geht, stehen. Diese müsste der Verantwortlich aber schon zu Beginn der Verarbeitung in der Tasche haben, denn eigentlich darf eine Verarbeitung personenbezogener Daten gar nicht erfolgen, bevor nicht das Risiko dieser Verarbeitung bewertet wurde: Sowohl die Einhaltung der DSGVO insgesamt (Art. 24 DSGVO) als auch spezifisch die Justierung der technisch-organisatorischen Maßnahmen zum Schutz der Daten (Art. 32 DSGVO) sind explizit risikoabhängig. Auch die Beantwortung der Frage, ob und für welche Verarbeitung eine Datenschutz-Folgenabschätzung notwendig ist, erfordert eine Risikoanalyse, die in diesem Kontext meist als „Schwellwert-Analyse" bezeichnet wird (Art. 35 DSGVO). Aber wie steht es mit dem Löschen?

Das Standard-Datenschutzmodell der Aufsichtsbehörden stellt Risiko und „Löschselektivität", nicht aber Risiko und Löschzeitpunkt in Beziehung zueinander: *„Die Granularität, in der Daten gespeichert werden und vor allem löschbar sind, hängt maßgeblich vom Risiko, das aus der Verarbeitung der Daten resultiert, vom Zweck der Erhebung und von der weiteren Verwendung der Daten ab. Je höher das Risiko für betroffene Personen ist, desto präziser müssen Daten löschbar sein. Das schließt jedoch nicht aus, dass etwa besonders hohes Risiko auch ein summarisches Löschen aller auf eine Person bezogener Daten erfordern kann, das ein feingranulares Differenzieren entbehrlich macht."* Letztlich umfasst diese Aussage zur Rasterung bzw. Kategorienbildung im Rahmen der „Löschselektivität" aber auch die zur Anwendung kommenden Fristen: Wer mehr Daten beim Löschen in einen Topf wirft, weil er nicht hochgranular löschen kann, löscht bei „hochgranularer" Betrachtung einzelne Daten zu spät (und damit nicht DSGVO-konform) oder zu früh (und damit nicht im Einklang mit Aufbewahrungspflichten). Man kann eine zu späte Löschung als „Sicherheitsaufschlag" deklarieren; ob das datenschutzkonform ist, ist eine andere Frage.

Eigentlich ist die Entscheidung, ob und wann zu löschen ist, aber gar keine Frage des Risikogehalts von Daten oder Verarbeitungen. Nach dem Wortlaut von Art. 17 DSGVO gibt es für jedes einzelne Datum – und damit bei hochgranularer Betrachtung – einen einzigen, richtigen Löschzeitpunkt, der allerdings nicht mit mathematischer Präzision, sondern allenfalls im Rahmen einer letztinstanzlichen Gerichtsentscheidung mit mehr oder weniger Überzeugungskraft festgelegt wird. In der Praxis ist die Ermittlung dieses richtigen Löschzeitpunkts im Vorhinein nicht mit letzter Sicherheit möglich, was aber solange kein praktisches Problem ist, solange der ernstgemeinte Versuch, es richtig zu machen, honoriert wird.

Praxistipp: Es empfiehlt sich, bei der Definition des Löschkonzepts auf eine bereits bestehende Risikoeinschätzung bezüglich der zu löschenden Daten aufzusetzen. Nach den Vorgaben der Datenschutzbehörden darf mit der Verarbeitung eigentlich gar nicht begonnen werden, ohne eine detaillierte Risikoanalyse durchgeführt zu haben. Soweit solche Risikoanalysen nicht vorliegen, ist vielleicht die Ausarbeitung eines Löschkonzepts eine gute Gelegenheit, dies nachzuholen.

Protokolldaten Beschäftigter

Datenbanksysteme erfassen meist neben den eigentlichen Inhalts- bzw. Nutzdaten auch die Zugriffe auf die Datenbank selbst in Form von Protokolldaten: *„geändert am [...] um [...] von [...]"*. Natürlich ist das ein personenbezogenes Datum des jeweiligen Beschäftigten. Mit vielen solcher Datensätze ließen sich gute Aktivitätsmuster generieren (Profiling). Wann hat wer wie viel gearbeitet?

Im Rahmen des Standard-Datenschutzmodells gibt es umfangreiche Vorgaben der Datenschutzbehörden zum Thema Protokollierung, die nicht immer ganz einleuchtend sind. Hier soll nicht auf die Frage eingegangen werden, in welchem Maße die Protokollierung – als Teil von technisch-organisatorischen Maßnahmen der Datensicherheit (Nachvollziehbarkeit von Zugriffen) – erforderlich ist, auch und gerade in Bezug auf Lesezugriffe. Wichtiger ist die Löschung, zu der es heißt: *„Für Protokolldatenbestände müssen Löschfristen festgelegt werden, wenn diese personenbezogene Daten enthalten. In der Regel sind zwei Löschfristen zueinander ins Verhältnis zu setzen: Zum einen die Löschfrist, die aus der Fachlichkeit abzuleiten ist, zum zweiten die Löschfrist, die aus funktionalen Gründen auf der jeweiligen Protokollierungsebene bestehen kann. Die fachlich begründete Löschfrist ist maßgebend.“* Das ist deswegen interessant, weil sich für den einzelnen Mitarbeiter, dessen Aktivität – normalerweise personenbezogen auf den Mitarbeiter – hier protokolliert wurde, die Frage nach der datenschutzrechtlichen Legitimation der Protokolldaten stellt. Normalerweise würde jeder davon ausgehen, dass die „Datenspuren“ der Mitarbeiter bei der Arbeit mit den IT-Systemen dem Arbeitgeber „gehören“ und im Extremfall „nie“ gelöscht werden müssen. Das sollte aber aufhorchen lassen. Alles, was nicht direkt mit § 26 BDSG, der Standard-Legitimationsgrundlage für die Verarbeitung von Beschäftigtendaten, gerechtfertigt werden kann, verdient eine genauere Betrachtung, auch hinsichtlich der Speicherdauer. Angenommen, der Mitarbeiter, dessen Aktivitäten protokolliert wurden, verlässt den Arbeitnehmer und verlangt Löschung sämtlicher Protokolldaten, weil diese nicht mehr für den Zweck des (beendeten) Beschäftigungsverhältnisses „erforderlich“ sind, was dann?

Die Vorgaben des Standard-Datenschutzmodells geben darauf keine Antwort, sondern erklären sybillinisch: *„Für die Protokollierung der Tätigkeiten von Beschäftigten, Administrationstätigkeiten sowie der Aktivitäten von IT-Systemen und Diensten sowie an Schnittstellen gelten ebenfalls die datenschutzrechtlichen Grundsätze der Zweckbindung und der Datenminimierung sowie die Regelungen des Beschäftigtendatenschutzes. In der Regel dürfen Protokolldaten nur zu den Zwecken, die Anlass für ihre Speicherung waren, ausgewertet werden.“*

Ein weiterer Aspekt der Protokollierung ist auch noch die Protokollierung, dass (andere) personenbezogene Daten gelöscht wurden. Hier sollte natürlich das ursprüngliche Datum nicht Teil des Protokolldatums sein, sonst wird der Löschzweck nicht erreicht. Man muss also eine „Krücke“ definieren, wie man den gelöschten Datensatz im Protokoll so individualisieren kann, dass er erkennbar ist, und so vage halten kann, dass nicht die ursprünglichen Inhaltsdaten erkennbar sind. Ein Protokolleintrag „37 Datensätze gelöscht“ hat wenig Aussagekraft, und ein Protokolleintrag (am Beispiel von Beschäftigtendaten) „Sämtliche 37 Einträge zur Behinderung von Herrn Florian Huber gelöscht“ hat zu viel Aussagekraft.

Praxistipp: Protokollierungen sollten nicht „einfach so“ von Technikern initiiert bzw. strukturiert werden. Zu jeder Verarbeitungsaktivität ist zu überlegen, ob und in welchem Umfang diese protokolliert werden sollte. Dabei ist schon beim Prozess- und Programmdesign an die Möglichkeit zu denken, die Identität des Beschäftigten, dessen Aktivitäten protokolliert werden, zu einem beliebigen Zeitpunkt isoliert löschen bzw. anonymisieren zu können. Daneben ist die turnusmäßige Löschung von Protokolldaten zu prüfen.

Zombies aus dem Back-up

Hat ein Unternehmen in Umsetzung eines Löschkonzepts Daten aus den Produktivsystemen gelöscht, stellt sich natürlich die Frage, wie die Löschung in Backup-Systemen gehandhabt wird. Das hochselektive Löschen von Datenbank-Einzeldaten ist auf Backup-Medien nicht so einfach möglich. Außerdem wird häufig der Verlust der Datenintegrität im Backup bei selektivem Löschen von Einzeldaten moniert (falsche Checksummen etc.). Backups sind also wichtig, um die Verfügbarkeit der Daten im Katastrophenfall sicherzustellen, stellen aber andererseits eine ständige, latente Gefahr der Verletzung von Löschpflichten dar, wenn dann der Katastrophenfall eintritt.

Meist wird akzeptiert, dass Backup-Medien Daten enthalten können, die im Produktivsystem bereits gelöscht wurden. Je schneller die Backup-Zyklen sind, desto weniger sammeln sich gelöschte Altdaten in den Backups an. Laut mündlicher Aussage eines Mitarbeiters der Datenschutzaufsicht sollte das Aufheben eines Backups für ein Jahr noch angemessen sein; wie eine solche Faustregel mit dem Gesetzeswortlaut in Einklang gebracht werden kann, ist eine andere Frage.

Auch hier sollten eigentlich Softwarehersteller kluge Lösungen bereitstellen, die beispielsweise Datensätze „hashen“ und beim Wiedereinlesen von Backup-Daten im „Disaster Recovery“-Fall prüfen, ob Datensätze mit einem bestimmten Hash-Fingerabdruck im Produktivsystem gelöscht wurden. Das setzt freilich voraus, dass die Datenbank nicht als Ganzes in das Produktivsystem reimportiert wird, sondern dass eine entsprechende Prüfung auf Datensatzebene stattfindet. Das wäre zwar datenschutzfreundlich und solche Produkte wären vom Verantwortlichen im Rahmen der Beschaffungsentscheidung vorzuziehen (Art. 25 Abs. 1 DSGVO), aber in der Praxis geben dann doch häufig andere Gesichtspunkte den Ausschlag.

Praxistipp: In einem vollständigen Löschkonzept sollte auch das Thema Backups hinreichend durchdacht sein. Immerhin enthalten Back-ups regelmäßig personenbezogene Daten, die der Verantwortliche eigentlich gar nicht mehr haben dürfte. Dabei geht es nicht-nur darum sicherzustellen, wie das „Wiederaufleben" eigentlich gelöschter Daten durch das Einspielen von Backups verhindert werden kann, sondern auch um das Berechtigungskonzept, wer eigentlich auf Backup-Daten zugreifen darf.

Was nun?

Das Vorstehende zeigt: Datenschutzrechtliche Löschkonzepte sind nicht einfach und sie sind mit vielen anderen Aspekten des Datenschutzes verknüpft. Das gilt auch und besonders für Datenbanksysteme, die personenbezogene Daten in derart großen Mengen und aus so vielen Jahren und in solchen Veränderungszyklen beinhalten, dass praktisch niemand mehr sämtliche Details kennen kann. Außerdem gibt es „das" Löschkonzept nicht, sondern die Art und Beschaffenheit der konkreten Daten, die Möglichkeiten der Systeme, die unternehmenseigenen Prozesse und vieles mehr sind miteinander zu verknüpfen, um zu einem maßgeschneiderten und nachvollziehbar hergeleiteten Löschkonzept zu gelangen.

Diese Erkenntnisse dürfen nicht zur Folge haben, dass der Verantwortliche den Kopf in den Sand steckt. Wichtig ist in jedem Fall, dass „zielgerichtetes Tätigwerden" entfaltet wird. Ein Löschkonzept kann auch in „Layern" entstehen, d. h. man setzt zunächst ein sehr grobes Löschkonzept auf und verfeinert dies immer weiter, oder man beginnt mit einem Löschkonzept für ein System und transferiert dies mit den gewonnenen Erkenntnissen und gemäß dessen Besonderheiten auf ein weiteres System.

Die praktische Erfahrung zeigt, dass der Schlüssel für ein erfolgreiches Vorgehen in der Bildung interdisziplinärer Teams aus Technikern, Fachabteilungen (die mit den Daten inhaltlich „hantieren") und Juristen bzw. Datenschützern liegt. Eine einzelne Disziplin kann die immensen Herausforderungen der Analyse, der Planung, der Dokumentation und der – auch technischen – Umsetzung nicht alleine stemmen. Wer hier seine datenschutzrechtlichen Hausaufgaben ordentlich gemacht hat, braucht sich nicht nur vor Bußgeldern und Haftungsansprüchen nicht zu fürchten, sondern sollte seine weiße Weste auch ordentlich gegenüber Mitarbeitern, Kunden, Lieferanten und Dritten zur Schau stellen gemäß dem Motto: Tue Gutes und sprich darüber. Denn datenschutzrechtlich „compliant" zu sein, ist ebenso wie Korruptionsbekämpfung, Mindestlohneinhaltung, Geldwäscheprävention und ähnlichen Themen ein wichtiger „soft factor", dessen (auch Geld-)Wert kaum unterschätzt werden kann.

Am Ende steht aber auch die Erkenntnis, dass, wer sein Unternehmen zunehmend digitalisiert (hat), auch die rechtlichen (Compliance-) Konsequenzen daraus ziehen muss. Man kann nicht die Vorteile der Digitalisierung für sich in Anspruch nehmen, dann aber die für die Digitalisierung geltenden rechtlichen Rahmenbedingungen ignorieren. Dass diese Rahmenbedingungen komplex, vielschichtig und mit hoher Rechtsunsicherheit behaftet sind und daher in der konkreten Anwendung anstrengend, aufwendig und mit unsicheren „Haftungsminimierungswirkungen" für die Verantwortlichen umgesetzt werden müssen, liegt möglicherweise an einer (entgegen den öffentlichen Beteuerungen) nur eingeschränkten Digitalkompetenz und praktischen Erfahrung der „Staatsgewalt" – also Legislative, Exekutive und Judikative in Deutschland und der EU – mit der Unternehmenspraxis. Die DSGVO ist nicht gerade vor dem Hintergrund einer profunden Kenntnis der Unternehmenspraxis entstanden – vermutlich wurden alle jetzt zutage tretenden Auslegungsprobleme nicht einmal im Ansatz gesehen –, aber das wird natürlich im Nachhinein trotzig im Sinne eines *„it's not a bug, it's a feature"* sowie mit dem Argument „so viel hat sich doch gar nicht verändert" gerechtfertigt. Dementsprechend wurde auch der bei den Unternehmen durch die Umsetzung entstehende Aufwand im Gesetzgebungsverfahren beinahe vollständig übersehen und massiv unterschätzt. Es nutzt aber nichts, sich über die Qualität von Gesetzen zu beschweren, die von demokratisch legitimierten Institutionen erlassen wurden, von dürftig ausgestatteten Behörden publikumswirksam umgesetzt werden sollen und dann von mit wenig Fachkenntnis versehenen „Allgemeingerichten" in punktuellen Einzelentscheidungen verbindlich ausgelegt und interpretiert werden. Die (nicht nur monetären) Folgekosten der Unternehmen, die langfristig entstehen werden, wenn die Anforderungen ignoriert werden, dürften – trotz manchem Unverständnis über die Regelungen – den Aufwand einer sorgsamen Planung und Implementierung weit übersteigen.