

Die Thesen sind aus dem Magazin „Zukunftsthemen: E-Commerce als Treiber der Digitalisierung“ zusammengestellt. Die aktuelle Ausgabe kann kostenfrei unter https://www.netz98.de/zukunftsthemen/?utm_source=pr&utm_medium=liste&utm_campaign=zukunftsthemen-2018 heruntergeladen werden.



Tim Hahn

1.4 RECHT

Ist die DSGVO das Ende der Data Science?

Der 25. Mai 2018 gilt als besonderer Tag für Unternehmen und ihre Datenanalysen, denn an diesem Tag wurden die Compliance-Regeln beim Datenschutz verschärft. Das Europäische Parlament hat dazu die Datenschutzgrundverordnung (DSGVO) erlassen, die seit jenem Tag verbindlich in allen EU-Mitgliedsländern gilt. Sie sieht unter anderem eine Erhöhung der Bußgelder auf bis zu 20 Millionen Euro bzw. bis zu vier Prozent des weltweiten Konzernvorjahresumsatzes vor. Um auf die neuen Vorgaben zu reagieren und daraus resultierende Risiken zu minimieren, müssen sich Unternehmen auf wirksame Systeme verlassen können. Doch die Menge an Daten, deren teilweise unklare Herkunft sowie die Komplexität der Datenstruktur führen zu einem Spannungsfeld zwischen datengetriebenen Anwendungen und der Umsetzung der DSGVO-Vorschriften.

Ein Großteil der Publikationen zur DSGVO konzentriert sich darauf, wie sich die neuen Regeln auf die Erfassung und Verwaltung personenbezogener Daten in Unternehmen auswirken. Gleichwohl ändert die DSGVO auch die Art und Weise, wie Unternehmen mit ihren Data Science Teams aus Daten ihre Erkenntnisse gewinnen werden. In diesem Kontext stellt dieser Artikel zwei Verfahren vor - Datenlöschungen und Analyse von personenbezogenen Daten – die zeigen, wie die aufstrebende Disziplin der Data Science sich im Spannungsfeld zwischen werthaltigen Analysen und Datenschutz nachhaltig aufstellen kann.

Das „Recht auf Vergessen“ in unstrukturierten Daten

Eine besondere Herausforderung für viele Unternehmen stellt das „Recht auf Vergessen“ gemäß Artikel 17 der DSGVO dar. Demnach hat die betroffene Person das Recht, dass ihre personenbezogenen Daten nach hinreichender Begründung auf Anforderung gelöscht werden.

In einer idealen Welt sind die Datensätze, welche mit einer Person verbunden sind, in einem Datenschema klar strukturiert nachvollziehbar. In diesem Fall kann man folglich die Daten relativ leicht auswählen. Auswirkungen auf die interne oder externe Berichterstattung durch Inkonsistenzen oder Unvollständigkeiten müssen auch bei der Löschung von Datenattributen vermieden werden. In diesem Kontext stellt sich daher die Frage, wie man Daten korrekterweise entfernen sollte. Um die referenzielle Integrität nicht zu verletzen, wird in der Praxis nur der Personenbezug entfernt und durch eine entsprechende Dummy-Variable ersetzt.

Anspruchsvoller hingegen gestaltet sich die Suche in unstrukturierten Daten, insbesondere wenn diese über verschiedene Datenquellen und Organisationseinheiten hinweg verteilt sind. Beispielsweise werden Kundendaten im Vertrieb erhoben, dann in der Auftragsfertigung und letztlich im Service verarbeitet oder neu erzeugt. Hier liegen dann womöglich Verträge, CRM-Daten, E-Mails, Briefe, etc. in unterschiedlichen Systemen.

Eine Zuordnung der Datensätze zu einem bestimmten Kunden ist teilweise nicht über eine exakte Suche möglich, da eine eindeutige Referenz nicht immer vorliegt – beispielsweise, wenn zu einem Kunden mehrere Emailadressen bzw. nicht alle (historischen) Adressen eindeutig im CRM-System hinterlegt sind. Der Komplexitätsgrad erhöht sich zusätzlich um ein Vielfaches, wenn noch verschiedene Schreibweisen, eine unterschiedliche Anzahl an Vornamen, der Wechsel der Nachnamen über die Lebenszeit oder – bei Firmenkunden - Umfirmierungen auftreten.

Jede „Suche“ nach personenbezogenen Daten birgt daher das Risiko, dass die Treffermenge nicht vollständig ist und somit nicht alle Daten vollständig identifiziert wurden. Idealerweise integriert man daher die Zuordnung von Inhalten zu Personen über einen eindeutigen Identifier in bestehende Prozesse und Systeme bzw. führt bestenfalls eine Anreicherung der bestehenden Daten durch. Es gibt hier unterschiedliche Lösungsstrategien, beispielsweise die Verknüpfung von personenbezogenen und personenbeziehbaren Daten, Triangulierung von Daten aus verschiedenen Quellen bis hin zur Auflösung von Handschriften oder Unterschriften zur exakten Zuordnung zu Personen.

Insbesondere bei unscharfen Suchen gilt grundsätzlich hier eine Balance zwischen Genauigkeit und Wirtschaftlichkeit zu bewahren. Moderne Data Science-Methoden können vor allem in der Aufarbeitung der Daten einen hohen Mehrwert liefern. Zwar entsteht hier ein einmaliges Investment, welches sich aber später durch



Michael Niederée,
Partner, Lighthouse
Germany, Center of
Excellence for Data
& Analytics,
KPMG

effizientere Prozesse und bessere Qualität der Prozesse letztlich doch auszahlt.

Durch ausgefeilte Anonymisierungsverfahren stichhaltige Analysen beibehalten

Darüber hinaus beeinflusst die DSGVO die Anwendung von Data Science dahingehend, als dass sie für die Erstellung von Kundenprofilen Grenzen setzt. Gleichzeitig haben Unternehmen auf der einen Seite hohe Anforderungen, ihre Kunden besser verstehen zu wollen und immer individueller zugeschnittene Produkte und Dienstleistungen anbieten zu können. Ein Beispiel wäre ein Onlineshop, der Produkte vorschlägt oder bei der Produktsuche besonders passende Produkte – etwa weil sie zum Geschmack oder Budget des Kunden passen – mit einem höheren Ranking darstellt.

Auf der anderen Seite sind personenbezogene Daten grundsätzlich zu pseudonymisieren bzw. zu anonymisieren. Einzelne Angaben dürfen nicht mehr oder nur mit „mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbarer natürlichen Person zugeordnet werden können“. Da der Schutzbedarf der Daten von vielen Faktoren abhängt – etwa von der Sensibilität der Daten und der Menge der unterschiedlichen Betroffenen – muss man die Anonymisierungsverfahren mit sehr viel Sorgfalt auswählen.

Wichtig ist, dass man die Anonymisierung nicht erst in der Data Science Workbench umsetzt, sondern eine klare Funktionstrennung etabliert, so dass die Datenlieferanten bzw. Eigentümer im Unternehmen nur Daten ohne erkennbaren Personenbezug zu Analyse Zwecken herausgeben. Hierzu kann man beispielsweise Datenverdichtungen anwenden, so dass mindestens immer zwei Datensätze in eine Gruppe verdichtet werden, um keinen direkten Personenbezug mehr herzustellen.

Alternativ lassen sich Daten synthetisieren, d.h. man vertauscht oder verändert die individuellen Merkmale derart, dass dies nur geringen Einfluss auf die Qualität der Modelle hat. Grundsätzlich lässt sich also feststellen, dass die Verarbeitung von personenbezogenen Daten mit der richtigen „Data Governance“ ein lösbares Problem ist. Es gilt nur, die richtigen Schutzmechanismen zu verwenden und sich von dem Gedanken zu lösen, Kundendaten zu Data Mining-Zwecken zu verwenden. Es geht letztlich immer um „Repräsentanten“ oder Mikrosegmente, die als Basis verwendet werden. In Folge wird ein besseres Kundenverständnis immer mit einer höheren Kundenloyalität belohnt und zahlt sich daher aus.

Fazit: Die Data Science im Spannungsfeld zwischen werthaltigen Erkenntnissen und Datenschutz als Business-Partner

Die DSGVO bringt einerseits eine drastische Verschärfung der Bußgelder mit sich. Sanktioniert wird dabei nicht nur der Datenschutzverstoß als solcher, sondern insbesondere das Fehlen DSGVO-konformer Prozesse und Maßnahmen. Gleichzeitig stehen Unternehmen vor der Herausforderung, durch die konsequente Auswertung ihrer produzierten und gesammelten Daten einen kundenindividuellen Service anzubieten und Trends rechtzeitig erkennen zu können.

Es gilt ein hoher Governance-Maßstab, der im Sinne eines Management-Systems zu konzipieren und zu implementieren ist. Die vorgestellten Methoden eröffnen den Weg für die Gestaltung eines Data Science-Ansatzes, der gleichzeitig den Anforderungen des modernen Datenschutzmanagements erfüllt. Unabhängig von der Branche gilt für alle Unternehmen: Sie müssen ihre bisherigen Zuständigkeiten, Prozesse und Maßnahmen neu bewerten und die Elemente der DSGVO als Leitplanke für die Umsetzung von Data Science Anwendungen integrieren. Da die Data Science selbst bei der Modellierung und Umsetzung der Verfahren durch ihr technisches Know-how zum Einsatz kommen wird, sollte die DSGVO aus Data Science-Sicht weniger als Beschränkung und mehr als Akzelerator betrachtet werden. Schließlich kann sich dadurch der Stellenwert der Disziplin zusätzlich steigern und ein größeres Vertrauen für die Analyse von Daten schaffen lassen.

Michael Niederée

Wem gehören die Daten im Connected Car?

Moderne Fahrzeuge erheben und übermitteln eine wahre Flut an Daten über Auto und Fahrer – Daten, die häufig einen Personenbezug aufweisen und für die sich die Hersteller, aber auch Versicherungen, Werbetreibende und Regierungseinrichtungen interessieren. Schon jetzt sind laut einer Statista-Umfrage rund zehn Millionen Connected Cars auf deutschen Straßen unterwegs, Tendenz steigend.

Wohin das exzessive Data-Mining führt, ist indes unklar: Welche Daten genau erhoben und an wen sie weitergeleitet werden, darüber schweigen sich die meisten Fahrzeughersteller aus. Die europäischen Automobilclubs fordern daher im Rahmen der Kampagne MyCarMyData, dass die Hersteller transparent machen, welche Daten sie erheben und weitergeben. Mehr noch: „Die Fahrer sollten das Eigentum an den Daten behalten, die ihr Auto produziert, und kontrollieren, wie sie verwendet werden, solange sie das

Fahrzeug besitzen“, so ein Sprecher des ADAC im Tagesspiegel.

... und was sagt die DSGVO dazu?

Auch wenn der Gesetzgeber streng genommen gar kein Eigentumsrecht an personenbezogenen Daten kennt, verweist der Bundesverband Digitale Wirtschaft (BVDW) im Zuge der Debatte auf die EU-Datenschutz-Grundverordnung: In einer Ende März 2018 veröffentlichten Stellungnahme der BVDW-Fokusgruppe Connected Mobility heißt es, die ab Mai 2018 anwendbare DSGVO und der weitere bestehende Rechtsrahmen böten bereits einen ausreichenden Spielraum für digitale Geschäftsmodelle und würden gleichzeitig die Datenhoheit und die informationelle Selbstbestimmung der Betroffenen gewährleisten. Der Datenaustausch, so der Bundesverband, müsse dabei allerdings „sicher, diskriminierungsfrei und zuverlässig“ erfolgen.

Wie geht sicherer Datenverkehr?

Für einen sicheren und diskriminierungsfreien Datenaustausch, wie der BVDW ihn fordert, ist zunächst eine sichere Infrastruktur nötig, die zudem den Anforderungen der DSGVO entspricht. Herkömmliche Server-Infrastrukturen kommen demnach nicht in Frage: Liegen die Daten verschlüsselt im Rechenzentrum vor, können sie nicht verarbeitet und ausgewertet werden; liegen sie hingegen unverschlüsselt vor, besteht die Möglichkeit eines unbefugten Zugriffs, beispielsweise durch den Betreiber des Rechenzentrums oder sein Personal. Oft ist dieser Personenkreis lediglich durch organisatorische Maßnahmen vom Zugriff auf die Daten ausgeschlossen – ein nicht zu vernachlässigendes Risiko bleibt dabei also bestehen.

Gerade bei personenbezogenen Daten, wie moderne Kraftfahrzeuge sie erheben, braucht es klare Regeln, um Missbrauch und Manipulationen auszuschließen. Und es braucht eine technische Lösung, die zuverlässig sicherstellt, dass die gesetzlichen Anforderungen eingehalten und die erhobenen Daten rechtskonform gespeichert, übertragen und verarbeitet werden.

Anonyme Datenverarbeitung mit CAR-BITS.de

In Zusammenarbeit mit dem Bundesministerium für Wirtschaft und Energie (BMWi), dem Fraunhofer Institut für Angewandte und Integrierte Sicherheit (AISEC), der Hochschule Bonn-Rhein-Sieg (H-RBS) und Continental Automotive GmbH arbeitet Unicon derzeit an einer solchen Lösung: Das Datenschutz-Projekt CAR-BITS.de hat die Aufgabe, Daten, die über das Auto ermittelt werden, rechtskonform auszuwerten. Ziel des Pilotvorhabens ist die Verwirklichung einer Diens-

te-Plattform, die verschiedenen Anwendungsbereichen anonymisierte Daten zur Verfügung stellt, welche aus Systemen von Autos stammen.

CAR-BITS.de kann aus den gesammelten Daten wertvolle Informationen gewinnen und verfügbar machen. Allerdings weisen diese keinen Personenbezug mehr auf. Da so auch bei der Verarbeitung der Daten keine Identifizierung einer Person möglich ist, wird auf diese Weise Big Data im Straßenverkehr rechtskonform nutzbar.

CAR-BITS.de basiert auf Unicons patentierter Sealed Cloud-Technologie. Die Sealed Cloud sorgt durch einen Satz rein technischer Maßnahmen dafür, dass Daten und Metadaten nicht nur bei der Übertragung zuverlässig geschützt sind, sondern auch bei der Verarbeitung im Rechenzentrum. So ermöglicht die Sealed Cloud nicht nur die datenschutzkonforme Übermittlung und Auswertung von Autodaten, sondern fungiert auch als Enabler für weitere kritische IoT-, SaaS- und Industrie-4.0-Anwendungen.

Dr. Hubert Jäger



Dr. Hubert Jäger,
Mitgründer und
Geschäftsführer,
Unicon GmbH

Gezielter Einsatz eines Patentschutzes für Software als Wettbewerbsinstrument

1. Optionen eines Patentschutzes, insbesondere im Bereich Open Innovation und Crowdsourcing

Die Digitalisierungsbranche kennzeichnet sich durch neue Entwicklungsmöglichkeiten, wie die verteilte Entwicklung, der Einsatz von Cloud-Diensten und agilen Methoden, um nur einige zu nennen. Software eignet sich zur Kollaboration.

In der Wissenschaft sind dafür interessanten Ansätze, wie Open Innovation und Crowdsourcing entwickelt worden, die von einigen Firmen inzwischen auch erfolgreich implementiert werden konnten. Dabei spielen IP-Schutzrechte eine zentrale Rolle.

Ein Patent ist ein wichtiger Vertreter dieser IP-Schutzrechte. Das Patent ist grundsätzlich ein sehr gutes Instrument zum Schutz des geistigen Eigentums (Intellectual Property – IP), zum Schutz von Erfindungen. Es ist DAS Instrument schlechthin zur nachhaltigen Sicherung von Marktexklusivität und zur Durchsetzung eines Premiumpreises auch für Erfindungen im Bereich Digitalisierung.

Für die Bestimmung einer geeigneten IP-Strategie ist es entscheidend, sich darüber Gedanken zu machen, für welche Innovation ein Patentschutz sinnvoll ist und welche Innovation z.B. unter bestimmten marktspezifischen Bedingungen für Wettbewerber frei verfügbar gemacht werden sollten. Dafür ist es wichtig, zu wissen, was die Erfordernisse für einen Patentschutz von Software sind.

2. Wann ist ein Patentschutz für Software sinnvoll und wichtig: Gewußt wie!

Aus einer historischen Perspektive betrachtet, wird deutlich, dass das Patentrecht für gegenständliche Erfindungen, wie gegen Ende des 19. Jahrhunderts typisch vorwiegend für Erfindungen in den Ingenieurwissenschaften, im Maschinenbau und in der Chemie sozusagen maßgeschneidert für diese traditionellen Technikgebiete geschaffen worden ist. Dementsprechend sind die Regelungen des Patentrechts auch auf diese traditionellen Gegenstände zugeschnitten.

Die Informatik ist eine relativ junge und sehr abstrakte Disziplin, die sich nicht in diese traditionellen Wissenschaftsbereiche einordnen lässt. Erfindungen auf dem Gebiet der Informatik passen somit nur eingeschränkt in das traditionelle Technik-Konzept, auf dem das Patentrecht aufbaut. Denn die Informatik unterscheidet sich in einigen für das Patentrecht relevanten Aspekten von den traditionellen Ingenieurwissenschaften.

Folgende Aspekte fördern eine spezifische Behandlung von Software-Erfindungen.

Im Unterschied zu anderen Erfindungen hat eine computerimplementierte Erfindung meist einen Doppelcharakter: Sie kann grundsätzlich sowohl in Software als auch in Hardware implementiert werden. Entsprechend gibt es unterschiedliche Verletzungsszenarien, die beim Drafting der Claims Beachtung finden müssen.

Software ist einfach, schnell und verlustfrei kopierbar. Gerade dies erfordert eine spezifische Verwendung der verfügbaren rechtlichen Schutzzinstrumentarien.

Software ist ubiquitär und diesbezügliche Erfindungen liegen häufig auf diversen Anwendungsgebieten und werden für diverse Alltagsgegenstände eingesetzt (z.B. Mechatronik, Software-Steuerungen im Anlagenbau oder der Medizintechnik etc.). Dennoch gibt es keine „eigene“ IPC-Klassifikation (etwa für „Softwarepatente“), was die Recherche und Prüfung mitunter erschwert und neue Konfliktszenarien erzeugt (Stichwort: bei Freedom-to-Operate Gutachten sieht sich ein deutscher Mittelständler im Anlagenbau plötzlich mit Googles Patentschriften konfrontiert) und Wissen notwendig ist, wie damit umzugehen ist

Ein Patentrecht sollte einen möglichst umfassenden Schutzzumfang für die Erfindung bieten. Deshalb werden im Patentwesen generell beim Drafting der Claims abstrakte, allgemeine Begrifflichkeiten verwendet. Abstraktion also, um einen möglichst breiten Schutzzumfang zu erzielen. Die Informatik ist jedoch ohnehin bereits eine sehr abstrakte Wissenschaft. Algorithmen werden häufig auf einer höheren Abstraktionsstufe be-

schrieben. Abstraktion, um die Funktionalität des Algorithmus unabhängig von der konkreten Implementierung in einer Programmiersprache darzustellen (Stichwort: Schichtenmodell von Softwarearchitekturen). Soll nun für algorithmische Lösungen ein Patentschutz nachgesucht werden, so führt dies in Summe zu einer quasi doppelten Abstraktion. Dies wiederum hat konkret zur Folge, dass mit weiteren Einwänden im Prüfungsverfahren zu rechnen ist (z.B. in Hinblick auf das Erfordernis der Klarheit). Die Frage, an welchen Stellen, wieviel abstrahiert werden sollte, spielt somit in der Praxis eine enorme Rolle und kann letztendlich entscheidend für die Erlangung und Rechtsbeständigkeit des Schutzrechts sein. Diese Frage kann nur unter ständiger Beobachtung der Rechtsprechung beantwortet werden.

3. Ständige Analyse der Rechtsprechung ist unabdingbar

Ein Blick in das Patentgesetz (und zwar sowohl in das deutsche PatG als auch in das Europäische EPÜ) gibt kaum eine Leitlinie, bei der Beurteilung, unter welchen Voraussetzungen Erfindungen mit Software-Bezug nun patentfähig sind und wann nicht. Im Gegenteil – findet sich im Gesetz die Formulierung, dass „Programme für Datenverarbeitungsanlagen als solches“ nicht als Erfindungen zu werten sind. Nutzt man aber die Recherchertools der Ämter, so findet man eine Vielzahl von erteilten Software-Patenten. Insbesondere die aktuellen Zahlen der Ämter signalisieren einen vermehrten Trend der Firmen, ihr geistiges Eigentum im Software Bereich zu schützen. Es ist somit patentrechtliches Know-How notwendig, um erteilungsfähige Patentanmeldungen zu formulieren, was eine intensive Analyse der sich dynamisch entwickelnden Rechtsprechung erfordert.

Fakt ist, dass sowohl in Deutschland als auch in Europa bestimmte Kriterien einzuhalten sind, damit eine Patentanmeldung – egal auf welchem Technikgebiet und somit egal, ob mit oder ohne Software – zum Patent erteilt werden kann. Diese sind vor allem, dass die Erfindung neu und erfinderisch sein muss. Zudem muss sie das sogenannte Technizitätskriterium erfüllen. Nach einer gefestigten Rechtsprechungslinie des EPA werden bei der Beurteilung der erfinderischen Tätigkeit nur die technischen Merkmale eines Patentanspruchs berücksichtigt (sog. COMVIK approach nach der Entscheidung T641/00). Damit wird es wichtig, wie jedes einzelne Anspruchsmerkmal in Hinblick auf seinen technischen Charakter beurteilt wird. Letzteres ist in der Praxis häufig das Zünglein an der Waage. Das Technizitätskriterium wird somit zur Messlatte und zur Grenzlinie zwischen Patentfähigem und Nicht-Patentfähigem. Legt man auf



Dr. Claudia Schwarz,
Gründungspartnerin,
Schwarz + Baldus
Patentanwältin

einen hohen Qualitätsmaßstab seiner Patente bzw. seines Patent-Portefolios Wert, so ist es wichtig, bereits beim Drafting einer Patentanmeldung immer die aktuelle, sich dynamisch entwickelnde Rechtsprechung im Blick zu haben, um das bereitgestellte Budget für IP auch sinnvoll zu nutzen und rechtsbeständige Patente zu erlangen. Eine Patentanmeldung schreiben können Viele, aber dies so zu tun, dass sie sich auch als rechtsbeständig erweist, steht auf einem anderen Blatt.

Inzwischen sind von den Spruchkörpern eine Vielzahl von Entscheidungen zu diesem Thema erlassen worden. Eine überblicksartige Zusammenfassung der Entscheidungen nach den unterschiedlichen Gebieten der Technik gegliedert, findet sich in unserem Fachbuch „Computerimplementierte Erfindungen – Patentschutz von Software?“, von Dr. Claudia Schwarz / Sabine Kruspig, 2. Auflage 2018, Carl Heymanns Verlag.

Patentanmeldung für Software Gegenstände können leicht an mehrere Hürden scheitern – Hürden, die in dieser Form nur in diesem Gebiet Anwendung finden. Auf dem Gebiet der Informatik ist es zur Erlangung eines hoch-qualitativen Schutzrechts somit nicht ausreichend, die Erfindung aus einer technischen Perspektive zu verstehen und zu beschreiben, sondern dies kann nur dann erfolgreich sein, wenn es unter ständiger Beobachtung der sich dynamisch entwickelnden Rechtsprechung erfolgt.

4. Internationaler Vergleich

Die Informatik fordert das Patentrecht heraus. Interessant ist, dass alle verschiedenen nationalen Patentrechtssysteme an dieser Stelle kämpfen. Es werden unterschiedliche Ansätze entwickelt, um patentfähige Software von nicht-patentfähiger zu unterscheiden. In Deutschland und Europa ist das Kernkriterium die Prüfung, ob eine technische Aufgabe mit technischen Mitteln gelöst wird (die Technizität), während z.B. in den USA geprüft wird, ob die Erfindung sich lediglich in einer abstrakten Methode erschöpft. Für das Erstellen einer geeigneten Patentstrategie für das jeweilige Unternehmen mit seinen Wettbewerbern im internationalen Markt ist es von erheblicher Bedeutung, auch diese nationalen Voraussetzungen mit zu bedenken.

Für eine genauere, internationale Analyse der unterschiedlichen gesetzlichen Anforderungen und den Ansätzen der jeweiligen nationalen einschlägigen Rechtsprechung sei auf unser Buch „Legal Protection for Computer-Implemented Inventions- A Practical Guide to Software-Related Patents“, Kruspig/Schwarz verwiesen.

Dr. Claudia Schwarz

Data Stewardship – die operative Komponente für ein effizientes Data Governance

Die Nachfrage nach Datenanalytik in Unternehmen steigt. Allein die EU-Datenschutz-Grundverordnung (DSGVO), die ab 28. Mai 2018 in Kraft tritt und stufenweise neue Datenschutzverordnungen für personenbezogene Daten einführt, zwingt Unternehmen EU-weit zu mehr Data Governance. Damit einher gehen neue Aufgabenbereiche beim Datenmanagement.

Notgedrungen wurde mit der Gesetzesankündigung eine Entwicklung in Gang gesetzt, die einerseits die Schwierigkeiten offenlegt, mit denen Unternehmen bei der Beherrschung ihrer Daten tatsächlich zu kämpfen haben. Gleichzeitig setzt durch die verstärkte Auseinandersetzung mit der Thematik ein Umdenken ein. Mittlerweile zeichnet sich in den Organisationen deutlich ab, welche konkreten Maßnahmen notwendig sind, damit ein vertrauenswürdige Datenmanagement etabliert werden kann.

Wer Daten besitzt, trägt Verantwortung

Doch nicht allein neue, steuerliche Rahmenbedingungen bringen Schwung in die Entwicklung. Unternehmen haben erkannt, dass sie die enormen Datenmengen, die täglich größer werden und verstreut und unstrukturiert in verschiedenen Datensilos gespeichert sind, nur dann effizient und gewinnbringend verwerten können, wenn feststeht, dass es sich um korrekte und vertrauenswürdige Daten handelt und diese - sinnvoll zueinander in Beziehung gesetzt - neuen Mehrwert generieren können. Kein verantwortungsvolles Unternehmen, das sich über die Qualität der eigenen Daten unsicher ist, wird diese weitergeben oder verwerten. Nicht allein deshalb diskutiert die IT-Welt neuerdings verstärkt über die Identifizierung, Kategorisierung, Verwaltung und Speicherung von Daten sowie über eine 360-Grad-Sicht auf Daten.

Um den Schutz und die Verwaltung unternehmenseigener, korrekter und vertrauenswürdiger Daten zu steuern und diese in der Folge auch tatsächlich als Produktionsfaktor effizient nutzen zu können, führen immer mehr Unternehmen Data Governance ein.

Data Stewardship als Wegbereiter der Analytik

Damit Data Governance als wertschöpfende Maßnahme gelingt, müssen im Unternehmen neue Rollen wie Data Owner und Data Stewardship klar definiert und eingeführt werden. Laut einer aktuellen DACH-weiten Studie[1] von BARC gaben 78% der befragten Unternehmen an, ihr wichtigstes Ziel bei der Etablierung von Data-Governance Rollen sei der Aufbau einer verlässlichen und

konsistenten Datenbasis für die Analytik. Weitere 55 % erwarten sich in der Folge bessere Analyseergebnisse durch mehr Datenqualität.

Die Rollen von Data Owner und Data Stewardship, die eine Schlüsselposition bei der operativen Umsetzung von Data Governance einnehmen, sind jedoch bislang zu wenig etabliert. Zwar haben laut Studie bislang bereits 65% der befragten Unternehmen die Rolle eines Business Analyst und immerhin 40 % die Rolle des Data Owner bzw. des Data Architects eingeführt. Das sind Aufgabenbereiche, die sich mehr der strategischen Entwicklung von Data Governance widmen.

Allerdings verfügen erst 13% der Unternehmen über Data Stewardship. Hier besteht also erheblicher Nachholbedarf, damit Unternehmen auch das letzte Stück des Weges schaffen. Schließlich sehen Spezialisten gerade in dieser Rolle den Wegbereiter für Analytik: Data Stewardship setzt die strategischen Vorgaben für mehr Datenqualität fach- und abteilungsübergreifend um. Zu den Hauptaufgaben eines Data Steward gehören:

- Identifizierung und Bereitstellung neuer Datenquellen.
- Erstellung und Pflege von Referenzdaten und konsistente Stammdatendefinitionen
- Veröffentlichung von relevanten Unternehmensdaten an entsprechende Nutzer im Unternehmen, Überwachung der veröffentlichten Datenquellen

- Erstellen und Verwalten eines Metadaten-Managements für veröffentlichte Datenquellen, um sicherzustellen, dass sie leicht interpretierbar und wieder erkennbar sind
- Beheben von Datenintegritätsproblemen für Anfragen der Business-Anwender
- Analyse der Daten für Qualität und Vereinbarkeit von Datenproblemen.

Datenmanagement- und Data Governance-Software sowie Werkzeuge für die Datenintegration wie z.B. entsprechende Datenbanklösungen, unterstützen diese Rolle.

Organisationen sollten Data Governance als große Chance betrachten, die richtigen Maßnahmen ergreifen und in zukunftsfähige Technologien investieren. Damit sichern sich Unternehmen einen starken Wettbewerbsvorteil: Sie realisieren ihre Projekte schneller, mit geringerem Risiko und weniger laufenden Kosten. Doch auch das zeigt die Studie: Die größten Hürden bei der Umsetzung von Data Stewardship ist die mangelnde Unterstützung auf Managementebene. 58% der Befragten sehen darin, sowie in der aktuellen Kultur beim Umgang mit Daten, die größten Hindernisse.

Dr. Stefan Grotehans

Referenzen: [1] <https://barc.de/Artikel/data-stewardship-wegbereiter-fur-analytik>



Dr. Stefan Grotehans,
Director Sales
Engineering DACH,
MarkLogic

UNSERE AKTUELLEN BLOG-RUBRIKEN:

- ✓ Machine Learning
- ✓ Quantum Computing
- ✓ Internet of Things
- ✓ Blockchain
- ✓ Cyber Security

INTERESSE GEWECKT?

Dann melden Sie sich bei der **DIGITALE WELT**-Redaktion per
E-Mail: blog@digitaleweltmagazin.de oder telefonisch unter der +49 89 2180 9171.



Werden Sie Teil unserer hochkarätigen Autorenschaft und platzieren Sie Ihre Digitalthemen von morgen auf der Plattform von heute mit bislang **317.000*** Klicks.

*Unsere Beiträge wurden online unter www.digitaleweltmagazin.de/blog veröffentlicht und erzielten dabei die oben genannte Klickanzahl im Zeitraum 1. August 2017 – 5. November 2018.