Enterprise Data Management Solutions
September 2008

IBM Information Management software

Best Practices für optimalen Datenschutz: Zeit zum Handeln!

Inhalt

Kurzübersicht

In der technologisierten Welt von heute sind Datenschutzverstöße nicht nur an der Tagesordnung, sondern können auch schnell hohe Kosten verursachen. Statistiken zu Datenschutzverstößen zeigen, dass die Anzahl solcher Ereignisse und die damit verbundenen Kosten kontinuierlich steigen. Dies ist ein Beleg dafür, dass Unternehmen in allen Branchen einen deutlich pragmatischeren Weg beim Schutz von Informationen einschlagen müssen. Insbesondere gilt dies für sehr anfällige Nichtproduktionsumgebungen (Entwicklung, Test und Schulung). Daten in solchen Umgebungen sind wesentlich empfänglicher für Datenschutzverstöße, wenn sie für Entwicklungs- und Testaktivitäten herangezogen werden, wenn mobile Mitarbeiter auf sie zugreifen oder wenn sie im Rahmen eines Outsourcingprojekts bereitgestellt werden.

Bei unvorhergesehenen Sicherheitsverletzungen ist eine umfassende Datenschutz- und Sicherheitsstrategie das beste Mittel, um den Schutz vertraulicher Informationen zu gewährleisten. Haben die Unternehmen einmal erkannt, dass Datenschutz mittlerweile keine Option, sondern ein Muss darstellt, häufen sich die Fragen: „Wo fangen wir an?“, „Was sind die Voraussetzungen?“ und „Welche Schritte müssen wir einleiten, um eine unternehmensweite Datenschutz- und Sicherheitsstrategie zu implementieren?“

Dieses White Paper erläutert die einzelnen Schritte, die bei der Entwicklung einer individuellen Datenschutzstrategie und der Implementierung Ihres ersten Datenschutzprojekts zu beachten sind. Mithilfe bewährter Datenmaskierungstechniken, wie sie beispielsweise in der IBM Optim Data Privacy Solution zu finden sind, kann Ihr Unternehmen Best Practices für den Datenschutz implementieren und Datenschutzprojekte vom Anfang bis zum Ende erfolgreich gestalten. Zum Schluss erfahren Sie dann noch, wie ein großes Einzelhandelsunternehmen mit der Optim-Lösung eine Best Practices-Strategie entwickelte und sehr erfolgreich ein Datenschutzprojekt durchführte. So konnten viele der Herausforderungen im Zusammenhang mit der unternehmensweiten Implementierung eines solchen Projekts vermieden werden.

Warum ist Datenschutz so wichtig? Die Fakten!

Durch immer mehr Hacker und Identitätsdiebstähle wirken sich Datenschutzverstöße in einer bisher unbekannten Dimension auf unser Privat- und unser Geschäftsleben aus. Zahlen belegen dies in überzeugender Weise. Laut Privacy Rights Clearing House erhöhte sich seit Januar 2005 alleine in den USA die Anzahl der Datensätze mit vertraulichen persönlichen Informationen, bei denen es zu Sicherheitsverstößen kam, auf 230.441.730.[1] Und diese Zahl steigt immer weiter.

Im heutigen Technologiezeitalter befinden sich viele der gefährdeten vertraulichen Informationen in den Geschäftsanwendungen und Computersystemen, die für unternehmensweite Geschäftsinitiativen wichtig sind. Ohne geeignete Maßnahmen für mehr Datenschutz und zur Vermeidung von Sicherheitsverstößen kann schon Ihr Unternehmen das nächste sein, das mit diesen Problemen zu kämpfen hat.

Datenschutz beginnt bereits mit dem Schutz der unterschiedlichen Arten vertraulicher Anwendungsdaten, und zwar unabhängig von deren Standort im Unternehmen. Dies gilt sowohl für Produktions- als auch andere Umgebungen (Entwicklungs-, Test- und Schulungsumgebungen). Vermehrt stellen Unternehmen jedoch fest, dass die Verfahren für den Datenschutz in Produktionsumgebungen in Nichtproduktionsumgebungen nicht umsetzbar sind.

Welche Bereiche sind am anfälligsten? Die Risiken!

Die Verfahren für den Datenschutz in Produktionsumgebungen und in Nichtproduktionsumgebungen sollten sich voneinander unterscheiden. In den meisten Produktionsumgebungen beispielsweise gibt es Sicherheits- und Zugriffsbeschränkungen, die helfen sollen, Datenschutzverstöße zu vermeiden. Standardsicherheitsmaßnahmen können auf Netzwerk-, Anwendungs- und Datenbankebene angewendet werden. Physische Eingangszugriffskontrollen können erweitert werden, indem Schemata für Mehrfachauthentifizierungen wie Schlüsseltokens oder Biometrie implementiert werden. Diese Schutzmaßnahmen können jedoch nicht einfach so in jeder anderen Umgebung repliziert werden. Wahrscheinlich erfüllen die Datenschutzverfahren in Produktionsumgebungen nicht die speziellen Datenschutzanforderungen für Nichtproduktionsumgebungen, in denen Entwickler, Tester und Schulungsleiter mehr – und nicht weniger – Zugriffsmöglichkeiten auf realistische Daten benötigen.

Einer Studie des Ponemon Institute und von Compuware aus dem Jahr 2007 zufolge verwendet die überwiegende Zahl der befragten Unternehmen aktuelle Daten für ihre Test- und Entwicklungszwecke (69 Prozent der Befragten nutzen aktuelle Daten für das Testen von Anwendungen und 62 Prozent für die Softwareentwicklung).[2] Unternehmen, die aktuelle Daten wie Kunden-, Mitarbeiter- und Lieferantendatensätze, Verbraucherlisten oder Kreditkarten-, Geschäftspartner- und andere Arten vertraulicher Informationen für Entwicklungs- und Testzwecke nutzen, erhöhten damit unweigerlich das Risiko von Sicherheitslücken.[3] Viele der untersuchten Unternehmen gaben an, dass die für die Softwareentwicklung eingesetzten aktuellen Daten nicht geschützt seien.

Die Studie legte zudem offen, dass ungefähr die Hälfte der Unternehmen ihre Anwendungstests auslagern und somit aktuelle Daten gemeinsam mit anderen Firmen nutzen. Die meiste Zeit wissen diese Unternehmen nicht, ob diese in ausgelagerten Testumgebungen verwendeten aktuellen Daten in irgendeiner Form gefährdet sind. Die einzige praktikable Lösung ist das Maskieren der Daten durch De-Identifizierung.

Beim De-Identifizieren von Daten in Nichtproduktionsumgebungen werden Datenelemente, die für die Identifizierung einer Person herangezogen werden könnten, systematisch entfernt, maskiert oder konvertiert. Durch die De-Identifizierung von Daten können Entwickler, Tester und Schulungsleiter realistische Daten verwenden, verwertbare Ergebnisse erzielen und dabei alle geltenden Datenschutzrichtlinien einhalten. De-identifizierte Daten können in der Regel problemlos in Nichtproduktionsumgebungen verwendet werden. So ist sichergestellt, dass selbst bei Diebstahl oder Verlust diese Daten für andere keinen Nutzen haben.

Planung eines Datenschutzprojekts

Durch branchenspezifische Datenschutzrichtlinien und -gesetze ist der Schutz von Daten längst keine Option mehr. Wie können Unternehmen also sicherstellen, dass die Informationen, die sie nach Übersee versenden, auf ihren Laptops speichern oder für interne Entwicklungs- und Testaktivitäten verwenden, ausreichend geschützt sind? Um solche sensiblen Informationen zu schützen, müssen Unternehmen bei ihrer Datenschutz- und Sicherheitsstrategie in erster Linie Verfahren zur De-Identifizierung von Daten in Betracht ziehen.

Darüber hinaus müssen sich die Unternehmen für eine bestimmte Entwicklungsmethodik entscheiden, die auf das Datenschutzprojekt abgestimmt ist. Im Verlauf des Projekts bilden sich in der Regel grundlegende Projektanforderungen und -voraussetzungen heraus. Aufgrund der Größe und Komplexität eines Datenschutzprojekts sowie der Anforderungen zur Einhaltung von Vorschriften ist es für Unternehmen ganz besonders wichtig, eine Methodik zu erarbeiten, die im gesamten Projektverlauf maximale Flexibilität bietet. Welche Kriterien sind also bei einem solchen Projekt zu berücksichtigen? Tabelle 1 zeigt sechs verschiedene Best Practices für die Implementierung eines erfolgreichen Datenschutzprojekts.

Tabelle 1. Management eines erfolgreichen Datenschutzprojekts

Schritt	Beschreibung
Organisieren	Zusammenstellung eines funktionsübergreifenden Datenschutzteams.
Anforderungen definieren	Definition der Anforderungen Ihres Datenschutzprojekts und Ermittlung der zu schützenden Anwendungen, Hardware und Daten.
Dateninventur durchführen	Analyse und Katalogisierung Ihrer Datenspeicher, -flüsse, -prozesse, -abhängigkeiten und Geschäftsregeln, um den Umfang Ihres Datenschutzprojekts zu reduzieren.
Lösung auswählen	Auswahl und Implementierung einer Datenschutzlösung, die die benötigten Verfahren aufweist, um den Datenschutz in allen Umgebungen zu gewährleisten.
Testen, testen, testen	Entwicklung eines Prototyps und einer Methodik für Ihr Projekt mit anschließender Gültigkeitsprüfung des Prototyps.
Geltungsbereich erweitern	Ausdehnung Ihres Datenschutzprojekts auf andere Anwendungen in Ihrem Unternehmen.

Diese Schritte bieten eine allgemeine Übersicht über das Management Ihres Datenschutzprojekts. Wir wollen nun die einzelnen Schritte näher betrachten.

Schritt 1 – Organisieren. Der erste Schritt beim Erstellen, Planen und Verwalten eines Datenschutzprojekts ist die richtige Organisation. Die Einrichtung verschiedener grundlegender Datenschutzdirektiven und -richtlinien hilft, den Umfang eines Projekts zu verdeutlichen und immer auf dem Laufenden zu bleiben. Für die effiziente Verwaltung eines Projekts und die Einhaltung solcher Direktiven ist die Bestimmung eines Projektleiters oder eines Projektteams erforderlich. Das Team muss aus folgenden Mitgliedern bestehen: Anwendungseigner und geschäftliche Nutzer, die direkt mit den Anwendungen arbeiten; Compliance-Manager, die sicherstellen, dass Ihr Unternehmen alle geltenden Datenschutzbestimmungen einhält; IT-Manager, deren Teams die Technologie für Ihre Datenschutzinitiativen implementieren; Organisationsleiter und Qualitätssicherungsmanager, die bei der Änderung automatisierter Prozesse und Testfälle hinzugezogen werden; andere Personen, die direkt mit dem Datenschutzprojekt zu tun haben. Ein solches funktionsübergreifendes Team setzt die abteilungsübergreifende Zusammenarbeit voraus, sodass alle involvierten Bereiche vertreten sind.

Während der gesamten Projektdauer arbeitet das Datenschutzteam bei projektspezifischen Entscheidungen eng zusammen, um Antworten auf alle aufkommenden Fragen zu finden, die das Projektziel gefährden könnten. Die Zusammenstellung eines solchen Teams bringt auch Vorteile bei anderen Projekten, da sich dadurch Datenschutzinitiativen auf das gesamte Unternehmen erstrecken.

Schritt 2 – Anforderungen definieren. Nachdem Sie Ihre grundlegenden Datenschutzdirektiven und -richtlinien definiert und ein Datenschutzteam für das Projekt zusammengestellt haben, müssen Sie als nächsten Schritt die Anforderungen hinsichtlich Datenschutz und De-Identifizierung der Daten im Unternehmen erkennen und definieren. Definieren Sie hierfür zunächst Ihre Zielsetzungen hinsichtlich der Compliance und berücksichtigen Sie dabei alle nationalen oder branchenspezifischen Datenschutzbestimmungen wie Health Insurance Portability and Accountability Act (HIPAA), Payment Card Industry Data Security Standard (PCI DSS) und andere. Jede Bestimmung weist ganz spezielle Anforderungen auf, die bei Ihrem Datenschutzprojekt zu beachten sind.

Als Nächstes erstellen Sie eine Liste der Anwendungen mit Angaben wie physische Position, Aufgabengebiet, unterstützende Datenbanken und Hardwareplattformen. Anwendungen, die für das Management und die Speicherung vertraulicher Kunden-, Mitarbeiter- und Geschäftsdaten verwendet werden, muss dabei hohe Priorität zugewiesen werden. Wenn Sie den Zweck jeder Anwendung verstanden haben, können Sie wesentlich einfacher die unterschiedlichen Datentypen, die in Programmbereichen maskiert werden müssen, und die Gründe dafür erkennen. Dann können Sie die Datentypen ermitteln, die de-identifiziert werden müssen, und abschätzen, welche dieser Datentypen maskiert und in Anwendungen, Datenbanken und Betriebsumgebungen verwendet werden müssen.

Schließlich müssen Sie beim Maskieren der Daten für Nichtproduktionsumgebungen wissen, wie präzise die Teilmengen der maskierten Ersatzdaten die Anwendungslogik in den originalen aktuellen Daten reflektieren müssen. Jeder Datentyp kann eine andere Maskierungsanforderung aufweisen. Wenn Ihre Produktionsdaten beispielsweise Sozialversicherungsnummern enthalten, stellt sich die Frage, ob die maskierten Ersatzdaten dem Format dieser Sozialversicherungsnummer entsprechen müssen.

Letztendlich ist festzuhalten, dass die De-Identifizierung von Daten das effektivste Mittel darstellt, um den Datenschutz und die Einhaltung von Vorschriften zu gewährleisten. Mit entsprechenden Funktionen für die De-Identifizierung vertraulicher Daten können Sie diesen Datenschutz gewährleisten und gleichzeitig die erforderlichen realistischen Daten für Entwicklungs-, Test-, Schulungs- und andere geschäftliche Zwecke bereitstellen.

Schritt 3 – Dateninventur durchführen. Dieser Schritt beinhaltet eine gewisse „Datenforensik“, da Sie die Metadaten analysieren (die Informationen zu Ihren Daten, die dafür sorgen, dass Ihre Daten leicht zu verstehen, zu verwenden und gemeinsam zu nutzen sind). Das Durchsuchen Ihrer Anwendungen, um den Inhalt zu ermitteln und zu bewerten, bietet Ihnen eine zuverlässigere Beschreibung der in Ihren Anwendungen verwalteten Daten. Dies gilt insbesondere für Bestandsdatenspeicher oder sequenzielle Dateien, bei denen persönliche Daten „verdeckt“ sein können.

Die Analyse des Datenflusses innerhalb Ihres Anwendungsportfolios kann ebenfalls bei der Klassifizierung der Daten in Gruppen und Hierarchien hilfreich sein. Alle Änderungen, die in den oberen Schichten der Hierarchie vorgenommen werden, werden auch in den unteren Schichten nachgezogen, sodass Sie Zeitaufwand, Umfang und Komplexität Ihres Projekts entsprechend minimieren können. Die Prüfung der Datenschutzanforderungen auf einer bestimmten Ebene kann Ihr Projekt ziemlich vereinfachen und beim Aufbau einer realistischeren Projektroadmap auf dem Weg zu Ihren Zielsetzungen helfen.

Schritt 4 – Lösung auswählen. Bei der Prüfung der Datenschutztechnologie sollten Sie nach einer Lösung suchen, die Ihren Anforderungen, wie in den Schritten 2 und 3 definiert, gerecht wird. Darüber hinaus sollten Sie entsprechende Änderungsumfänge einplanen. Da es immer wieder neue Datenschutzbestimmungen gibt und bestehende Verordnungen verstärkt werden, müssen Sie über Möglichkeiten verfügen, um entsprechende Änderungen vornehmen zu können. Gleichermaßen wird es bei jedem Upgrade oder jeder Erweiterung zu Änderungen an Ihren Geschäftsanwendungen kommen. Um mit diesen Änderungen Schritt halten zu können, muss Ihre Datenschutzlösung skalierbar sein und flexibel einsetzbare Funktionen aufweisen, über die Sie Datenmaskierungs- und Datenschutzroutinen nach Bedarf ändern können.

Die IBM Optim Data Privacy Solution beispielsweise bietet ein umfassendes Funktionsspektrum für die De-Identifizierung von Anwendungsdaten, das in Nichtproduktionsumgebungen effizient genutzt werden kann. Neben hoher Skalierbarkeit und Flexibilität zeichnen sich die Datenmaskierungstechniken von Optim durch Konsistenz und Reproduzierbarkeit aus, wodurch sie allen aktuellen und kommenden Anforderungen von Anwendungen, Datenbanken, Betriebssystemen und Hardwareplattformen gerecht werden.

Die Optim-Lösung weist die Bandbreite und Tiefe auf, die eine unternehmensweite Datenmaskierung ermöglicht. Sie ist keine Einzellösung, mit der nur punktuell einzelne Datenschutzprobleme behoben werden können. Als Unternehmenslösung lässt sich Optim in Ihre bestehenden laufenden Geschäftsprozesse einbinden.

Damit Unternehmen auch die komplexesten Datenschutzanforderungen erfüllen können, stellt die Optim-Lösung die folgenden grundlegenden Komponenten für eine effektive Datenmaskierung bereit:

- ***Beibehaltung der Anwendungslogik.*** *Die anwendungsorientierten Datenmaskierungsfunktionen von Optim erkennen, erfassen und verarbeiten Datenelemente mit hoher Präzision, sodass maskierte Daten die Anwendungslogik nicht beschädigen. Nachnamen werde beispielsweise durch wahlfreie, gültige Nachnamen und nicht durch bedeutungslose Zeichenfolgen ersetzt. Numerische Felder behalten ihre Struktur und ihre Muster. Bei vierstelligen Diagnosecodes, die im Bereich zwischen 0001 und 1000 liegen, wäre ein maskierter Wert von 2000 im Kontext von Anwendungstests ungültig. Prüfsummen behalten ihre Gültigkeit, sodass Funktionstests Gültigkeitsprüfungen für Anwendungen erfolgreich bestehen. Optim repliziert zudem alle maskierten Datenelemente konsistent in der gesamten Testdatenbank und in allen zugehörigen Anwendungen und Datenbanken.*

Das Unternehmen Direct Response Marketing Company, Inc. z. B. testet sein Auftragsabwicklungssystem und muss die Kundennamen de-identifizieren, um ein sicheres Testen zu gewährleisten. Mit der Optim-Funktion Random Lookup kann das Unternehmen ganz willkürlich Vor- und Nachnamen aus einer vordefinierten Kundeninformationstabelle generieren. Aus „Lucille Ball“ wird bei jedem Auftreten dieses Namens „Elena Wu“ usw. Diese Maskierung ist reproduzierbar und vorhersagbar, sodass dieselbe Änderung immer konsistent vorkommt und somit Ihren Anforderungen für Nichtproduktionsumgebungen gerecht wird.

- ***Schlüsseldatenelemente maskieren.*** *Die kontextorientierten, vordefinierten Datenmaskierungsroutinen von Optim de-identifizieren Schlüsseldatenelemente und stellen eine Vielzahl von bewährten Datenmaskierungstechniken bereit, die für die De-Identifizierung der unterschiedlichsten Arten von sensiblen Informationen herangezogen werden können. So können beispielsweise Geburtsdaten maskiert werden, um das richtige Alter einer Person präzise wiederzugeben. Gleichermaßen lassen sich Nummern von Bankkonten, nationale Kennungen (wie die Sozialversicherungsnummer in Deutschland oder die Codice Fiscale in Italien), Informationen zu Unterstützungsleistungen usw. maskieren.*

Vordefinierte Transformation Library-Routinen ermöglichen die präzise Maskierung komplexer Datenelemente wie Sozialversicherungsnummern, Kreditkartennummern und E-Mail-Adressen. Integrierte Suchtabellen unterstützen die Maskierung von Namen und Adressen. Sie können aber auch standortspezifische Datenkonvertierungsroutinen einbinden, mit deren Hilfe die Verarbeitungslogik aus verschiedenen zusammengehörigen Anwendungen und Datenbanken integriert und komplexe Datenmaskierungsanforderungen flexibler und kreativer gehandhabt werden können.

Die Kontonummern der Green Bill Bank beispielsweise sind im Format „999-9999“ formatiert, wobei die ersten drei Ziffern für einen Kontotyp (Bankkonto, Sparkonto oder Geldmarktkonto) stehen und die letzten vier Ziffern die Identifikationsnummer des Kunden angeben. Zu Testzwecken müssen diese Kontonummern maskiert werden. Die Optim-Lösung kann nun die ersten drei Ziffern der Kontonummer übernehmen und eine sequenzielle vierstellige Zahl generieren, die die letzten vier Ziffern der eigentlichen Kontonummer ersetzt. So wird z. B. aus der Kontonummer „001-4570“ die maskierte Nummer „001-1000“. Das Ergebnis ist somit eine erfundene Kontonummer, die weiterhin das Kontonummernformat der Bank aufweist.

- ***Maskierte Datenelemente präzise verbreiten.*** *Die persistenten Maskierungsfunktionen von Optim generieren für Quellenspalten konvertierte Ersatzwerte und replizieren diese konsistent und präzise für alle Anwendungen, Datenbanken, Betriebssysteme und Hardwareplattformen. Solche Maskierungsfunktionen stellen eine hohe Skalierbarkeit beim Datenschutz in mehreren Nichtproduktionsumgebungen sicher. Die Replizierung maskierter Primärschlüsselwerte in allen zusammengehörigen Tabellen ist erforderlich, um die referenzielle Integrität der Daten auch nach deren Maskierung zu erhalten, sodass auch vollständige Teilmengen an zusammengehörigen Daten intakt bleiben (siehe Abbildung 1).*

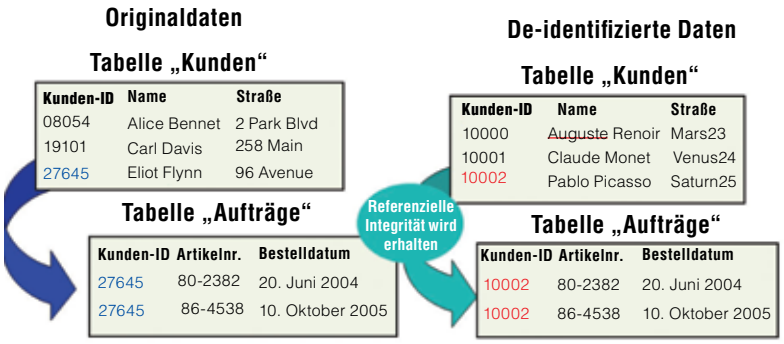

Abbildung 1. Die Schlüsselreplizierungsfunktionen von Optim erhalten die referenzielle Integrität der Daten.

Abbildung 1 zeigt ein einfaches Beispiel mit zwei zusammengehörigen Tabellen, bei dem die Tabelle „Kunden“ der Tabelle „Aufträge“ übergeordnet ist. Die Primärschlüsselspalte „Kunden-ID“ ist ein fünfstelliger numerischer Wert. Die Spalten „Kunden-ID“, „Name“ und „Straße“ sind maskiert. Der Name „Elliot Flynn“ wurde als „Pablo Picasso“ maskiert. Das Verfahren der sequenziellen Maskierung wurde verwendet, um die ursprüngliche Kunden-ID für Elliot Flynn von 27645 in 10002 zu konvertieren. Der maskierte Kunden-ID-Wert wird aus der Tabelle „Kunden“ in allen zusammengehörigen Tabellen repliziert. Die Schlüsselbeziehung zwischen den Tabellen „Kunden“ und „Aufträge“ in der Testdatenbank bleibt intakt.

Nach dem Entwurf Ihrer Datenschutzstrategie und der Entscheidung für eine Lösung erfolgt nun die Implementierung. Die anwendungs- und kontextorientierten Verfahren der Optim-Lösung sowie deren Replizierungsfunktionalität sind leistungsfähige Datenmaskierungsfunktionen, die Sie für eine erfolgreiche Implementierung brauchen.

Schritt 5 – Testen, testen, testen. Nach der Ausarbeitung Ihrer Datenschutzstrategie müssen Sie nun prüfen, ob diese auch funktioniert. Erstellen Sie hierzu einen Prototyp für ein Maskierungsszenario für eine bestimmte Anwendung oder Anwendungsgruppe, um sicherzustellen, dass Ihr Maskierungsverfahren und Ihre Testaktivitäten wie erwartet funktionieren. Dieses Szenario muss mit Ihren Anforderungen an die Anwendungstests übereinstimmen. Stellen Sie deshalb sicher, dass Ihre Datenschutzstrategie möglichst effizient ist, damit sie die definierten Anforderungen erfüllt. Vergleichen Sie beispielsweise Ihre Originaldaten mit den de-identifizierten Daten, um zu gewährleisten, dass Ihre Änderungen übernommen wurden. Zum Schluss überprüfen Sie dann noch, ob Ihre Datenschutzstrategie für Ihre Anwendungen und Ihre Nichtproduktionsumgebungen geeignet ist.

Jetzt wird der Prototyp getestet, um festzustellen, ob die jeweiligen Maskierungsverfahren angewendet wurden. Haben Sie dabei sichergestellt, dass alle Bereiche mit sensiblen Daten maskiert sind? Werden alle sensiblen Daten auch anwendungsübergreifend und in allen Testumgebungen maskiert? Entspricht Ihre Datenschutzstrategie den branchenspezifischen Regelungen und gesetzlichen Datenschutzbestimmungen?

Schritt 6 – Geltungsbereich erweitern. Nach der Implementierung können Sie den Geltungsbereich Ihres Datenschutzprojekts erweitern, damit auch andere Anwendungsgruppen im Unternehmen berücksichtigt werden. Ein proaktiver Ansatz beim Datenschutz im Unternehmen ist der Schlüssel zum Erfolg. Implementieren Sie Datenschutzstrategien auch in anderen Bereichen in Ihrem Unternehmen, in denen schutzwürdige Informationen gespeichert oder verwaltet werden. Kann es in diesen Bereichen eventuell zu Datenschutzverstößen kommen?

In der Regel ist Datenschutz nicht auf einen bestimmten Bereich im Unternehmen beschränkt. Durch die Ausdehnung Ihres Datenschutzprojekts auch auf andere Bereiche mit sensiblen Informationen können Sie Ihre Informationsressourcen noch besser schützen.

Ein übergreifender Datenschutz ist der beste und sinnvollste Weg, um sensible Daten effizient zu verwalten. Die De-Identifizierung von Daten ist ein sicheres Mittel für das Testen geschäftskritischer Anwendungen, mit dem Unternehmen zum einen hohen Datenschutz und zum anderen eine enge Kundenbindung sicherstellen. Optim verfügt über ein flexibel einsetzbares und skalierbares Funktionsspektrum, das alle aktuellen und auch zukünftigen Datenschutzanforderungen abdeckt. So wird die De-Identifizierung von Daten zum wesentlichen Bestandteil Ihrer gesamten Datenschutzstrategie. Mit Optim ist ein Unternehmen deutlich besser positioniert, wenn es um die Einhaltung der Datenschutzbestimmungen auf kommunaler, Landes-, nationaler und internationaler Ebene sowie von gesetzlichen und branchenspezifischen Vorgaben geht.

Weitere wichtige Informationen
Datenschutzprojekte – einschließlich Datenschutzteams, Prozesse, Umgebungen und Formeln – hängen immer vom jeweiligen Standort ab und können durchaus voneinander abweichen. Effizienz beim Datenschutz und bei der Datensicherheit kann von vielen Faktoren abhängen. Nachfolgend finden Sie einige weitere Punkte, die Sie bei Ihrem Datenschutzprojekt berücksichtigen sollten:

- *Physische Trennung des Datenschutzprozesses vom Testprozess. Die Implementierung des Datenschutzprozesses in einer produktionsähnlichen Testumgebung bietet ein Höchstmaß an Sicherheit für Ihr Unternehmen.*

- *Strikte „Rollentrennung" in Ihrem Projektteam. Wenn der Zugriff auf de-identifizierte Daten nur bedarfsorientiert erfolgt, ist das Risiko, dass der Maskierungsprozess ins Gegenteil verkehrt wird, geringer.*

- *Maßnahmen zum Schutz der internen Sicherheit durch leichte Veränderung Ihres „Datenschutzrezepts". Leichte Veränderungen an den Daten, bevor diese in die Produktionsumgebung eingebunden werden, können eine weitere Ebene des Zugriffsschutzes darstellen.*

- *Prüfung durch Dritte, um Ihre Datenschutzprozesse zertifizieren zu lassen und dadurch zu belegen, dass sie manipulationssicher sind, Sie die Rollentrennung wie erforderlich vornehmen, Ihre Roadmaps stimmen und Ihr Zeitfenster ausreichend Puffer zur Schließung von Lücken im Prozess beinhaltet.*

Optim-Implementierung aus Benutzersicht
Als Discounter mit hohen Wachstumsraten betreibt die Marzan Corporation eine Reihe von Einzelhandelsgeschäften in den USA und Großbritannien. Seit der Eröffnung der ersten Filiale in New Jersey im Jahr 1979 hat sich Marzan im Lauf der Jahre einen Namen als Anbieter von Waren für den täglichen Bedarf unter dem Motto „Gut und Günstig" gemacht. Vom Haushaltsartikel über Bekleidung und Elektronikartikel bis hin zum Kfz-Zubehör bietet Marzan dem Kunden ein bequemes, modernes, unterhaltsames Einkaufserlebnis zu günstigen Preisen.

Vielfältiges Anwendungsportfolio. Zur Unterstützung der laufenden Geschäftsinitiativen und der Geschäftsbeziehungen zu den Lieferanten setzt Marzan eine SCM-Anwendung von PeopleSoft Enterprise ein, die intern unter dem Namen IZZI läuft und die Lieferantenbestellungen und die Lagerbewegungen verwaltet. IZZI enthält sensible Informationen wie Lieferantennamen und IDs und läuft auf einer Oracle-Datenbank.

Marzan nutzt darüber hinaus eine CRM-Anwendung von Siebel, die den internen Namen Jasper trägt und über die Kundenbestellungen bearbeitet werden. Jasper läuft in einer DB2-Open-Systems-Umgebung und enthält vertrauliche Kundenbestelldaten wie Kundennamen, Adressen, Telefonnummern und Kreditkartennummern. Einige dieser Informationen nutzt Marzan für verkaufsfördernde Maßnahmen, um seinen Kundenstamm zu erweitern.

Als Drittes setzt Marzan eine eigenentwickelte Finanzanwendung namens Centz ein, die auf einem IBM System z Server läuft. Diese Rechnungsstellungsanwendung erfasst und speichert die Kreditkartenanwendungsdaten bei Marzan sowie die Rechnungsstellungsdaten für Kunden, die auch Angaben wie Sozialversicherungsnummern, Namen und Adressen enthalten können.

Herausforderungen beim Datenschutz. Für Marzan standen Kriterien wie Qualitätsprodukte und Kundenservice schon immer im Vordergrund. Um einen besseren Kundenservice anbieten zu können, wollte Marzan seinen Kunden bessere Funktionen für den Onlinezugriff auf Kontendaten und Onlinebestellungen bereitstellen. Für das Erreichen dieses Ziels war es erforderlich, dass die bestehenden Anwendungen Jasper und Centz optimiert wurden. Jede Anwendung enthält sensible Daten aus Onlineeingaben wie Kreditkartennummern, Sozialversicherungsnummern, Namen, Adressen und Telefonnummern. Entsprechende Verbesserungsmaßnahmen würden zukunftsweisende Onlinefunktionalität mit mehr Kundenorientierung in Bezug auf Kontoinformationen, Zugriffsmöglichkeiten und Sicherheit mit sich bringen.

Da in der IZZI-Anwendung auch Bestelldaten von Lieferanten mit Angaben wie Lieferantennamen und IDs abgelegt sind, muss Marzan dafür sorgen, dass diese Informationen maskiert werden, damit diese auch bei Test- und Entwicklungsaktivitäten sicher sind. Da die Tests auch auf Datenbank- und Anwendungsebene ausgeführt werden mussten, wollte Marzan zudem sicherstellen, dass die referenzielle Integrität der Daten intakt blieb.

Als große Einzelhandelskette musste Marzan letztendlich auch darauf achten, dass die Bestimmungen des PCI DSS (Payment Card Industry Data Security Standard) eingehalten wurden. Gemäß PCI DSS müssen große Einzelhandelsunternehmen und Großunternehmen, die Kreditkartendaten verarbeiten, persönliche Kundeninformationen, die in Anwendungstestumgebungen verwendet werden, entsprechend maskieren. Marzan musste also gewährleisten, dass alle Daten aus den Anwendungen Centz und Jasper, die für Test- und Entwicklungszwecke verwendet wurden, maskiert wurden und somit in Nichtproduktionsumgebungen sicher genutzt werden konnten.

Die Suche nach der geeigneten Lösung. Marzan benötigte eine Lösung, die alle Datenschutzanforderungen des Unternehmens erfüllte. Um die Datenschutzinitiativen des Unternehmens ausreichend zu unterstützen, entschied sich Marzan für die Optim Data Privacy Solution von IBM. Die Entscheidung für eine unternehmensweite Lösung für die Datenmaskierung wurde von einem kleinen Datenschutzteam angestoßen, das aus Anwendungsentwicklern und Testern, IZZI-, Jasper- und Centz-Benutzern und Datenschutzspezialisten bestand.

Marzan baute daraufhin ein Datenschutzteam auf, das dieses vielschichtige Projekt federführend betreuen sollte. Zielsetzung des Teams war, mit einem proaktiven Ansatz die Sicherheit der Kundeninformationen anzugehen. Optim bot hierfür die Funktionalität, um schutzwürdige Informationen im gesamten Unternehmen zu schützen. Die einzigartigen Maskierungs- und Konvertierungsfunktionen von Optim für Mainframe- und Open Systems-Anwendungsdaten bieten dabei Datenschutz über alle Anwendungen, Datenbanken, Betriebssysteme und Hardwareplattformen hinweg.

Wettbewerbsvorteile durch erfolgreiche Implementierung. Nach der erfolgreichen Implementierung der Optim-Lösung implementierte das für die De-Identifizierung der Daten zuständige Projektteam bei Marzan die aufwendig erarbeitete Methodik des Datenschutzprojekts. Beginnend mit IZZI leitete Marzan die Schritte für die De-Identifizierung der sensiblen Daten in den Nichtproduktionsumgebungen im Unternehmen ein.

Für die De-Identifizierung der Lieferantennamen in IZZI nutzte Marzan die Suchfunktionen von Optim, mit denen die Daten mithilfe von Substitutionswerten konvertiert wurden. Das Team konnte dabei den Wert in einer Quellenspalte maskieren, sodass ein entsprechender maskierter Wert in einer Zielspalte ausgegeben wurde. Die echten Namen der Lieferanten wurden also für die Test- und Entwicklungszwecke durch erfundene Namen ersetzt werden. Durch die Suchtabellen in Optim wurde damit aus „Dave Acme“ von der Acme Pencil Company in der Nichtproduktionsumgebung „Michael Craft“.

Die Lieferanten-IDs in IZZI bestehen aus sechs Zeichen, wobei das erste Zeichen der erste Buchstabe des Lieferantennamens und das zweite Zeichen ein Buchstabe aus der Stadt des Lieferanten ist. Die letzten vier Stellen sind numerische Werte und liegen im Bereich zwischen 1000 und 6999. Die Acme Pencil Company in Tucson, Arizona, hätte in diesem Fall also die Lieferanten-ID AT1453. Die anwendungsorientierten Funktionen in Optim stellen dabei sicher, dass eine maskierte Lieferanten-ID die Anwendungslogik beibehalten und eine maskierte Lieferanten-ID wie CD2047 – und nicht CD8945 – zurückgegeben wird.

Als Nächstes wandte Marzan das De-Identifizierungsverfahren auf Jasper, die CRM-Anwendung von Siebel, an. Die kontextorientierten Datenmaskierungsroutinen de-identifizierten in den Anwendungstest- und Entwicklungsumgebungen die Schlüsseldatenelemente wie Sozialversicherungsnummern, Kreditkartennummern und Geburtsdaten. Um die Entwicklungs- und Testaktivitäten korrekt ausführen zu können, mussten die Marzan-Mitarbeiter in der Entwicklung mit echten sechzehnstelligen Kreditkartennummern arbeiten. Die intelligenten Maskierungsfunktionen von Optim „erfanden“ für diese Zwecke kontextspezifisch korrekte, aber maskierte Kreditkartennummern. Bei einem Datenschutzverstoß sind die maskierten Kreditkartenummern für Diebe nutzlos, für den effizienten Einsatz in Nichtproduktionsumgebungen hingegen nicht (siehe Abbildung 2).

Abbildung 2. Optim generiert gültige und eindeutige, aber de-identifizierte Kreditkartennummern gemäß den Formatanforderungen des Ausstellers.

Vergleichbare kontextorientierte Maskierungsverfahren wurden auch für die Finanzanwendung Centz herangezogen. Diese Anwendung erfasst und speichert kundenspezifische Rechnungsstellungsinformationen. Marzan nutzte auch hier die Suchtabellen von Optim, um die in Centz gespeicherten Namen und Adressen zu maskieren. So wird beispielsweise aus „Beth K. Smith" nach der De-Identifizierung durchgängig „Claire P. Hamill". Diese Änderungen wurden von Optim so repliziert, dass die referenzielle Integrität beibehalten und der Datenschutz in der Nichtproduktionsumgebung von Marzan gewährleistet wurde.

Zum Schluss sollten die Onlinekunden von Marzan durch verschiedene Anwendungsverbessungen noch besser auf ihre Konten zugreifen können. Hierfür musste man verteilte Test- und Entwicklungsumgebungen aufbauen. Optim bot für diese Problematik Funktionen für den verteilten Zugriff, mit denen Entwickler Testdaten aus verschiedenen Datenquellen in einem Prozessschritt extrahieren und maskieren konnten. Die Subsetting-Funktionalität der Optim-Lösung stellte die Automatisierungs- und Reproduzierbarkeitsfunktionen für die Verarbeitung verteilter Extrakte aus den Anwendungen IZZI und Jasper zur Verfügung.

Mithilfe der Optim-Lösung erstellte Marzan realistische Teilmengen der Anwendungsdaten für seine Test- und Entwicklungsumgebungen. Sensible Kundeninformationen wie Kundennamen, Adressen, Telefonnummern, Kreditkartennummern und Zahlungsprotokolle wurden durch Optim maskiert. Die dadurch de-identifizierten Daten konnten somit problemlos in diesen Umgebungen eingesetzt werden. Anschließend wurden die Daten in allen Nichtproduktionsumgebungen repliziert, wobei die referenzielle Integrität der Daten beibehalten und somit ein zuverlässiges Testen gewährleistet werden konnte.

Die Optim-Funktionen zum Schutz von Kundeninformationen in den Test- und Entwicklungsumgebungen stellten zudem sicher, dass die PCI DSS-Anforderungen eingehalten wurden. Da sich in den SCM-Anwendungen von Marzan persönliche Kunden- und Kreditkarteninformationen befanden, stellte Optim auch hierfür Funktionen bereit, um solche vertraulichen Daten zu maskieren. Somit war Marzan in der Lage, Risiken von rechtlicher Seite und mögliche Geldstrafen zu vermeiden und die Bindung der Kunden an das Unternehmen zu stärken.

Mithilfe von Optionen wie „Substrings“, Ersetzung wahlfreier oder sequenzieller Zahlen, arithmetische Ausdrücke, „Date Aging“ und anderen Verfahren ersetzte Marzan aktuelle Kundendaten durch kontextspezifisch präzise, aber erfundene Daten, um genaue Testergebnisse zu erzielen. Diese Daten konnten problemlos und sicher in Nichtproduktionsumgebungen verwendet werden, waren jedoch für Diebe und Hacker völlig nutzlos.

Letztendlich half die Optim-Lösung dem Unternehmen dabei, Kosten und Risiken im Zusammenhang mit potenziellen Datenschutzverstößen zu vermeiden und den hervorragenden Ruf des Unternehmens zu untermauern. Für die Kunden und geschäftlichen Nutzer von Marzan bringen die zuverlässigeren und mit vielen Funktionen ausgestatteten Anwendungen zusätzliche Vorteile. Zudem muss noch erwähnt werden, dass Marzan mit dieser Lösung seinen hohen Standard beim Kundenservice aufrechterhalten konnte und mittlerweile die Früchte aus seinen umsatzfördernden Maßnahmen ernten kann.

Informationen zu IBM Optim

Im Mittelpunkt der IBM Optim Enterprise Data Management Solutions stehen kritische Geschäftsprobleme im Unternehmen wie Management des Datenwachstums, Einhaltung von Datenschutzrichtlinien, Testdatenmanagement, E-Discovery, Anwendungsupgrades, Migrationen und Außerbetriebnahmen. Bei Optim ist das Anwendungsdatenmanagement an den Geschäftszielen des Unternehmens ausgerichtet, um optimale Leistung zu erzielen, Risiken zu begrenzen und Kosten zu kontrollieren. Gleichzeitig stellt Optim eine Funktionalität bereit, die sich perfekt an die Anwendungen, Datenbanken und Plattformen im Unternehmen anpassen lässt. Mit Optim erzielen Unternehmen auf der ganzen Welt und aus allen Branchen einen hohen geschäftlichen Nutzen aus ihren Unternehmensanwendungen und -datenbanken, indem deren Datenbestände in allen Lebenszyklusphasen optimal verwaltet werden.

Weitere Informationen

Wenn Sie mehr über IBM Optim Enterprise Data Management-Lösungen erfahren möchten, wenden Sie sich an Ihren IBM Vertriebsbeauftragten oder besuchen Sie uns unter:

www.optimsolution.com

IBM Deutschland GmbH
Pascalstrasse 100
70569 Stuttgart
ibm.com/de

IBM Österreich
Obere Donaustrasse 95
1020 Wien
ibm.com/at

IBM Schweiz
Vulkanstrasse 106
8010 Zürich
ibm.com/ch

Die IBM Homepage finden Sie unter:
ibm.com

[1] *Privacy Rights Clearing House, www.privacyrights.org/ar/ChronDataBreaches.htm#2008*

[2] *Compuware and The Ponemon Institute LLC. „The Insecurity of Test Data: The Unseen Crisis.“ Compuware.com. Dezember 2007: Seite 3*

[3] *Ibd. Seite 4*

TAKE BACK CONTROL WITH **Information Management**

IMW14072-DEDE-00