ARTIKEL

Datensicherheit

Lösung: **XDM**

Maskieren von Daten mit XDM

Das Masking-Tool von XDM wurde entwickelt, um die verschiedenen Herausforderungen zu meistern, die mit der Maskierung von Daten einhergehen. Es sind verschiedene Szenarien denkbar, und jedes hat einzigartige Anforderungen. Beispiele sind die Maskierung von personenbezogenen Daten (PII) oder sensiblen Gesundheitsdaten (PHI) zur Einhaltung von Vorschriften wie DSGVO/GDPR oder HIPAA. Unabhängig vom Szenario ist die Aufrechterhaltung der Datenvalidität entscheidend, um realistische Tests zu ermöglichen.

Dieser Gedanke gilt sowohl für bestehende Randfälle als auch für Datensätze, die im Laufe der Zeit gewachsen sind. Durch die Verwendung realer Fallstudien, die von (modifizierten) Produktionsdaten abgeleitet sind, können Fehler frühzeitig erkannt und gleichzeitig die Datenschutzbestimmungen eingehalten werden. Mit Tools wie XDM zur Datenmaskierung lassen sich bestimmte Daten effizient ändern, so dass die Einhaltung der Vorschriften ohne zusätzlichen Aufwand oder manuelle Anpassungen gewährleistet ist.

Ein weiterer entscheidender Punkt bei einem Datenmaskierungsprojekt in einem Unternehmen ist die bestehende Struktur. Ziel ist es, sich nahtlos in die bestehenden Strukturen zu integrieren, d.h. ein Maskierungsprojekt unabhängig von der verwendeten Datenbankinfrastruktur durchführen zu können. Ein Tool zur Datenmaskierung muss mit den vorhandenen Strukturen umgehen können, um den zusätzlichen manuellen Aufwand zu minimieren.

Das Masking-Tool von XDM kopiert und modifiziert Daten nahtlos über alle Datenbankgrenzen und Datenhaltungssysteme hinweg. So können beispielsweise Daten aus verteilten Datenbanksystemen zusammen mit Daten aus legacy VSAM-Systemen kopiert und maskiert werden. XDM kann auch in einer Cloud-basierten Umgebung eingesetzt werden und ermöglicht das Lesen und Schreiben von Daten in der Cloud.

XDM erleichtert hybride Lösungen, indem es verschiedene Arten von Datenbanksystemen unterstützt und konsistente Modifikations-Definitionen über verschiedene Plattformen hinweg ermöglicht. Dadurch entfällt die Notwendigkeit, für jeden Systemtyp separate Modifikationen zu erstellen. Ein mal definierte Modifikationen, die mit den Spezifikationen der Compliance-Abteilung übereinstimmen, können universell angewendet werden, was den manuellen Aufwand auf ein Minimum reduziert.

Umsetzung: Aufspüren relevanter Daten und Maskierung

A Ein Projekt zur Verfremdung von Daten beginnt häufig mit der Aufgabe, jene Daten zu finden, die überhaupt maskiert werden müssen. Dazu gehört die Bestimmung, welche

Datensicherheit

Lösung: **XDM**

UBS HAINER

Umsetzung (Fortsetzung)

Daten anonymisiert werden müssen, welche anonymisiert werden können und welche Daten für Test- oder Entwicklungszwecke benötigt werden.

XDM verfügt über ein Tool namens PII Finder, das dazu dient, personenbezogene Daten (PII, engl. für personal identifiable information) in einer Tabellenspalte zu identifizieren. Das Tool verwendet vorkonfigurierte Dictionaries, die Listen mit sensiblen Informationen wie Namen, Bankleitzahlen oder Straßennamen enthalten. Mit diesen Informationen wird geprüft, ob eine Spalte solche sensible Daten enthält.

Technisch gesehen verwendet XDM so genannte "Matcher", um die zu analysierenden Spalten und Daten zu definieren. Nachdem die Tabelle gescannt wurde, erstellt XDM einen Bericht, der die Wahrscheinlichkeit anzeigt, ob eine bestimmte Information in einer Spalte vorkommt. Diese Berichte bilden die Grundlage für die Identifizierung geeigneter Maskierungstechniken für die betreffenden Spalten.

Sobald die Ergebnisse der Analyse verfügbar sind, besteht der nächste Projektschritt darin, die zu maskierenden Daten und die zugehörigen Maskierungsmethoden zu definieren. In XDM unterscheidet man zwischen verschiedenen Komponenten, die für die Verfremdung von Daten notwendig sind.

Modifizierungsmethoden (engl. Modification Methods) werden verwendet, um die Art und Weise der Maskierung von Daten zu definieren. Eine einfache Methode kann zum Beispiel vorschreiben, dass der Ausgabewert unabhängig vom Eingabewert konstant ist. Alternativ können komplexere Methoden mit Programmiersprachen wie JavaScript oder Groovy entwickelt werden, um interne Werte entsprechend den spezifischen Anforderungen anzupassen.

Modifikationsregeln (engl. Modification Rules) legen fest, welche Modifikationsmethode für welche Felder zu verwenden ist. Dies geschieht durch Zuweisung der zuvor erstellten Methode zu einer Spalte innerhalb einer Tabelle.

Einzelne Modification Rules werden in einem Modification Set gruppiert. Ein Modification Set umfasst mindestens eine Modification Rule und wird im weiteren Verlauf des Projekts mit einer Datenkopie verknüpft.

Durch die Kapselung der Modification Methods können XDM-Benutzer diese wiederholt auf verschiedene Objekte anwenden, ohne jedes Mal dieselbe Maskierung neu definieren zu müssen. Die Maskierung muss nur einmal definiert werden und kann dann beliebig oft wiederverwendet werden.

Umsetzung (Fortsetzung)

Die Fähigkeit von XDM, Modification Methods zu parametrisieren, erhöht die Flexibilität. Benutzer können dieselbe Methode problemlos in verschiedenen Szenarien anwenden, indem sie einen oder mehrere Parameter für jede Anwendung anders einstellen.

Bei der Datenmaskierung sorgt das sog. Enforcement, das meist im Zusammenhang mit Berechtigungskonzepten verwendet wird, für eine konsistente Maskierung bestimmter Daten, sobald diese kopiert werden. In XDM wird die Modifikation dafür entsprechend auf einer höheren Ebene definiert (etwa direkt für die verwendete Datenbankverbindung). Das führt dazu, dass diese Modifikation immer angewendet wird, sobald die spezifizierten Daten Teil einer Datenkopie sind.

Anwendungsbereiche

XDM bietet verschiedene Möglichkeiten zum Kopieren und Verfremden von Daten:

Eine Möglichkeit besteht darin, Daten in-place zu kopieren und sie gleichzeitig zu ändern. Bei diesem Verfahren werden die Daten gelesen, geändert und dann zurückgeschrieben, ohne dass ein weiteres Datenspeichersystem erforderlich ist. Dies ist nützlich, wenn eine Datenkopie bereits vorhanden ist oder mit schnellen Klonmechanismen vorbereitet wurde und die Maskierung dann mit XDM durchgeführt werden soll.

Alternativ dazu kann XDM die Daten auch auf ein separates Zielsystem kopieren. In diesem Fall werden die Quelldaten gelesen, verfremdet und dann in ein anderes System kopiert. XDM kann auch automatisch fehlende Strukturen im Zielsystem anlegen.

Wenn Sie die benötigten Daten eingrenzen wollen, ohne ganze Tabellen oder Schemata zu kopieren und zu verändern, können Sie eine Teilmenge der Daten selektiv kopieren und maskieren. XDM kann dieses selektive Kopierverfahren nutzen, um einzelne Datensätze für bestimmte Testfälle zu kopieren.

Zusätzlich zu den Kopiermethoden schreiben die GDPR/DSGVO-Anforderungen auch die Möglichkeit vor, maskierte Daten zu verfolgen und bei Bedarf zu löschen. Ein Beispiel: Ein Kunde verlangt, dass Daten, die zu Test- oder Entwicklungszwecken bereitgestellt wurden, nach Abschluss der Arbeiten gelöscht werden. Um dem nachzukommen, muss es möglich sein, die Daten für eine vollständige Löschung abzurufen.

Anwendungsbereiche (Fortsetzung)

Eine weitere Anforderung besteht darin, bestimmte Daten auf eine schwarze Liste zu setzen, wodurch sie vom Maskierungsprozess ausgeschlossen werden – das sogenannte Black-listing von Daten. Dadurch wird sichergestellt, dass die Daten nicht in den maskierten Datensatz des Zielsystems aufgenommen werden.

Definieren und Anpassen von Modifikationsmethoden

XDM stützt sich auf o.g. Modifizierungsmethoden, um die Art und Weise der Datenmaskierung festzulegen. Standardmäßig enthält XDM eine Reihe von vordefinierten Methoden, die in jedem Maskierungsprojekt sofort verwendet werden können. Diese vorkonfigurierten Methoden eignen sich gut für gängige Szenarien wie die Maskierung von Namen, Adressen, Bankinformationen und Geolocation-Daten.

Darüber hinaus gibt es die Möglichkeit, benutzerdefinierte Modifcation Methods zu erstellen, die auf die Anforderungen der vorhandenen Datenstruktur zugeschnitten werden können. Diese Anpassung kann die Arbeit mit Hash-Algorithmen, die Integration von benutzer-definierten Bibliotheken (JavaScript, Groovy oder Java) oder die Erstellung von benutzerdefinierten Lookup-Tabellen umfassen. Sie können die Daten, die sich bereits in der Anwendungsdatenbank befinden, zur Erstellung von Lookup-Tabellenverwenden, um Sonderfälle in den Originaldaten zu erfassen.

Durch die Verwendung speziell berechneter Hash-Werte stellt XDM sicher, dass die Daten-maskierung bei Bedarf wiederholt werden kann und die Werte konsistent anonymisiert werden. Es ist auch möglich, Hash-Werte auf der Grundlage anderer Felder zu erzeugen, entweder in derselben oder in einer anderen Tabelle. Durch die zufällige Anordnung der Lookup-Tabellen können die Ergebnisse des Maskierungsverfahrens verändert werden. So ist es möglich, in einem bestimmten Zeitraum deterministische Maskierungsergebnisse zu erhalten, die jedoch zu einem anderen Zeitpunkt anders ausfallen.

XDM bietet eine Reihe von Reports, die die in den Datenkopien verwendeten Datenmaskierungsverfahren zusammenfassen und dokumentieren.

Maskieren persönlicher Informationen

Bei der Datenmaskierung werden vorhandene Informationen so verändert, dass es unmöglich ist, sinnvolle Rückschlüsse auf die ursprünglichen Daten zu ziehen. Für eine wirksame Datenmaskierung ist es entscheidend, Richtlinien wie die DSGVO/GDPR zu befolgen. Projekte, die sich auf Datenmaskierung konzentrieren, benötigen ein Tool, das in der Lage ist, vorhandene Daten gemäß diesen Richtlinien zu maskieren. Bei diesem Prozess müssen jedoch verschiedene Herausforderungen bewältigt werden. Nachfolgend werden diese Herausforderungen anhand von Beispielen untersucht und gezeigt, wie XDM Datenmaskierung effektiv handhaben kann.

Maskieren von Namen

Eines der häufigsten Beispiele für Datenmaskierung beginnt mit dem Namen einer Person, der in der Regel aus einem Vor- und einem Nachnamen besteht. Die Änderung eines Vornamens beispielsweise erfordert häufig, dass der geänderte Name weiterhin als Name einer Person identifizierbar ist. Dadurch wird sichergestellt, dass die Daten und die aus den geänderten Daten erstellten Reports für technische und explorative Tests verwendbar bleiben.

In XDM kann dies mit Hilfe von Lookup-Tabellen erreicht werden, die einen Satz von Vor- und Nachnamen enthalten. Anhand einer mathematischen Berechnung wird dann ermittelt, welcher Name einen anderen ersetzt. Ein identischer Name muss in der gesamten Datenbank einheitlich geändert werden, wenn er an mehreren Stellen erscheint. Dies kann entweder durch eine feste Zahl (z. B. eine Personenkennzahl) oder durch den eingegebenen Namen selbst erreicht werden. Die Verwendung des Eingabenamens macht eine zusätzliche Referenznummer überflüssig, kann aber zu völlig unterschiedlichen Namen führen, wenn die Schreibweise an verschiedenen Stellen der Datenbank leicht variiert.

Außerdem muss oft sichergestellt werden, dass das Geschlecht der Person unverändert bleibt. Wenn beispielsweise der ursprüngliche Name "John" lautet, der in der Regel männlich ist, sollte der maskierte Name ebenfalls männlich sein, z. B. "Mark". Dazu sind Informationen über das Geschlecht der Person erforderlich, die entweder direkt aus der Tabelle extrahiert oder über eine Lookup-Tabelle ermittelt werden können. Wenn beides nicht möglich ist, kann XDM einen zufälligen Namen als maskierten Wert generieren.

Maskieren von Namen (Fortsetzung)

Es gibt Fälle, in denen Nachnamen Doppelnamen oder Leerzeichen enthalten oder in denen Vor- und Nachname zusammen als ein einziger "Name" in einer Tabellenstruktur gespeichert sind. Es ist wichtig, diese Fälle genau zu erkennen und zu maskieren und dabei die vordefinierten Regeln für die Änderung zu befolgen.

Ein weiteres Beispiel für spezielle Maskierungsanforderungen sind Geburtsdaten. In einzelnen Anwendungsfällen muss eine Person einer bestimmten Altersgruppe angehören, auch wenn der Name maskiert wurde. Das folgende Beispiel stammt aus der Versicherungsbranche: John, der 18 Jahre alt ist, hat sein Auto versichert. Als Fahranfänger ist er in eine bestimmte Tarifgruppe für die Kfz-Versicherung eingestuft worden. Damit John in dieser Tarifgruppe bleibt, auch wenn seine realen Daten zu Testzwecken maskiert werden, kann sein Geburtsdatum nicht beliebig geändert werden, sondern muss nach bestimmten, von einem XDM-Benutzer festgelegten Kriterien maskiert werden. So dürfen beispielsweise nur der Tag und der Monat maskiert werden, während das Jahr unverändert bleibt.

Maskieren von Adressen

Ein weiteres Beispiel aus der Versicherungsbranche ist die Verfremdung von Adressen der Versicherungsnehmer. Bei der Verfremdung solcher Daten ist es wichtig, die Art der versicherten Immobilie zu berücksichtigen, um sicherzustellen, dass sie unverändert bleibt. So ist beispielsweise ein Einfamilienhaus anders versichert als ein Mehrfamilienhaus. Wenn Datensätze von Personen, die ein Einfamilienhaus versichert haben, für einen Testfall benötigt werden, müssen die maskierten Daten diese Informationen genau wiedergeben.

Dabei geht es um das spezifische Risiko von Überschwemmungen (oder anderen potenziellen Gefahren wie Lawinen) und die entsprechende Einstufung von Gebäuden in verschiedene Risikogruppen. Diese Faktoren werden beim Abschluss einer Versicherung berücksichtigt und müssen für Testszenarien entsprechend maskiert werden. Die Adresse kann also nicht willkürlich geändert werden. So sollte eine Person, die in Risikobereich A wohnt, auch nach einer Maskierung der ursprünglichen Daten noch als in Risikobereich A wohnend eingestuft werden.

XDM bietet verschiedene Möglichkeiten, Adressen zu maskieren. Sie können festlegen, dass die maskierte Adresse innerhalb einer bestimmten Entfernung oder im selben Stadt- oder Postleitzahlengebiet wie die ursprüngliche Adresse liegen muss.

Maskieren von Adressen (Fortsetzung)

Alle oben genannten Beispiele erfordern die Kenntnis der Daten der jeweiligen Anwendung und müssen entsprechend betrachtet werden. Wenn zum Beispiel eine Police für eine bestimmte Klassifizierung vollständig gebucht ist, kann die Berechnung der Gebühren nur auf den Informationen in der Police beruhen, auch wenn die Adressen nicht einem bestimmten Kriterium entsprechen, das während der Maskierung geändert wurde. Möglicherweise führt eine Erneuerung einer solchen Police dann zu anderen Ergebnissen, als sie in der Produktion erzielt würden. Diese Szenarien sollten mit dem Test- und dem Anwendungsteam besprochen werden, um die gewünschten Ergebnisse dieser Vorgänge bei der Arbeit mit den maskierten Daten zu klären.

Maskieren von Bankinformationen

Neben Namen und Adressen werden bei Maskierungsprojekten häufig auch Bankdaten wie IBANs, Swift-Codes oder Routing-Nummern unkenntlich gemacht. Bankleitzahlen können mit Hilfe von vordefinierten Algorithmen maskiert werden, die mit XDM ausgeliefert werden, während zur Maskierung von Banknamen Lookup-Tabellen verwendet werden.

Zusammenfassung

Die in diesem Artikel vorgestellten Funktionen verdeutlichen die zahlreichen Vorteile von XDM als Datenmaskierungswerkzeug in jedem Projekt. Mit seinen benutzerfreundlichen vordefinierten Maskierungsmethoden, der Möglichkeit, diese Methoden anzupassen, und der Fähigkeit, geeignete Daten im Voraus zu identifizieren, ist XDM das ideale Werkzeug für die effiziente Durchführung von Projekten.

Die benutzerfreundliche Oberfläche, der minimale Bedarf an manuellen Eingriffen und die nahtlose Integration in bestehende Umgebungen machen XDM zu einer überzeugenden Wahl für alle Datenmaskierungsanforderungen.

Datensicherheit

Lösung: **XDM**

Besuchen Sie uns unter https://ubs-hainer.com/de/product/xdm/masking/ und kontaktieren Sie uns noch heute, um Ihre erfolgreiche Datenmaskierung zu starten.

Für weitere Informationen kontaktieren Sie bitte:

info@ubs-hainer.com