GLB RODO Tool

Dokumentacja administratora

Autor: Globema

Data: kwiecień-maj 2018

www.globema.pl
EXPERIENCE. INNIVATIONS. SMART SOLUTIONS.

Globema Sp. z o.o.
ul. Wita Stwosza 22
02-661 Warszawa

tel.: 22 848 73 13
fax: 22 848 68 83
kontakt@globema.pl

Spis treści

1 Przeznaczenie modułu

Moduł **GLB RODO Tool** jest przeznaczony dla administratorów systemu Smallworld i służy do anonimizacji treści zawartych w bazie danych Smallworld.

Typowym użyciem tego narzędzia jest anonimizacja danych osobowych w ramach przygotowania kopii bazy produkcyjnej do przekazania podmiotowi świadczącemu usługę serwisową, testową lub rozwojową.

Anonimizacja jest procesem nieodwracalnym, dlatego zawsze powinna być wykonywana na kopii bazy.

Moduł jest dostarczany jako moduł Magikowy o nazwie `glb_rodo` i nie wymaga do działania innych modułów. Został przetestowany na platformie Smallworld w wersji 4.2 i 4.3.

2 Funkcjonalność modułu

2.1 Anonimizacja

Podstawową funkcją modułu jest anonimizacja wskazanego pola w bazie danych, poprzez podanie nazwy banku danych, nazwy kolekcji (tabeli), nazwy pola i sposobu anonimizacji. Anonimizacji podlega wskazane pole we wszystkich rekordach wskazanej tabeli.

Dostępne są następujące sposoby anonimizacji:

1. Wyczyszczenie zawartości pola – pole dostaje wartość `_unset`.
2. Wpisanie w pole losowej wartości z podanej listy.
3. Wygenerowanie wartości z generatora liczb naturalnych, począwszy od podanej liczby, z krokiem 1, z opcjonalnym przedrostkiem, np. „user5”, „user6”, „user7”, ...

Anonimizacji mogą podlegać zarówno pola numeryczne jak i alfanumeryczne, należy jednak zadbać o zgodność sposobu anonimizacji i typu pola. Na przykład, próba wpisania wartości z listy tekstowej do pola numerycznego zakończy się błędem.

Pola, które są puste nie podlegają przetworzeniu – pozostają puste bez względu na wybrany sposób anonimizacji.

Anonimizacji nie można zastosować do pól kluczowych.

Należy pamiętać, że anonimizacja działa w danej alternatywie w bieżącym punkcie kontrolnym. Dane pozostaną bez zmian w pozostałych alternatywach.

2.2 Inne narzędzia

Dodatkowo są dostępne inne narzędzia wspomagające:

1. Anonimizacja użytkowników zdefiniowanych w bazie autoryzacyjnej – można:
    a. usunąć wszystkich użytkowników oprócz root
    b. usunąć wszystkich użytkowników oprócz wskazanych
    c. wyczyścić pole *Pełna nazwa*, nie usuwając rekordów użytkowników
2. Usuwanie alternatyw – można usunąć:
    a. wszystkie alternatywy (oprócz szczytowej)
    b. wskazane alternatywy
    c. wszystkie alternatywy oprócz wskazanych
    d. wszystkie alternatywy pod wskazaną
3. Usuwanie wszystkich punktów kontrolnych dla wskazanych alternatyw

3 Korzystanie z narzędzia

3.1 Typowa procedura użycia

Typowa procedura przygotowania bazy pod kątem pozbawienia jej danych osobowych wygląda następująco:

1. Wykonać kopię produkcyjnej bazy danych. Zaleca się użycie narzędzi ds_transfer w celu kompresji tej bazy z opcją przeniesienia tylko jednej alternatywy.
2. Uruchomić obraz otwarty aplikacji.
3. Załadować moduł: `sw_module_manager.load_module(:glb_rodo)`
    a. W przypadku, gdy moduł nie jest umieszczony w strukturze modułów lecz został dostarczony osobno, można go załadować za pomocą `sw_module_dialog.open()` jako *standalone module*.
4. Wyczyścić (zminimalizować) drzewo alternatyw w każdym banku danych. Jeśli nie zrobiono tego w kroku 1, można użyć narzędzia modułu `glb_rodo`. W tym celu należy:
    a. Przygotować plik konfiguracyjny o nazwie `alternatives_configuration.xml` (patrz rozdz. 3.2). W ramach konfiguracji określa się, które alternatywy maja być usunięte i czy punkty kontrolne w pozostałych/wskazanych alternatywach mają być usunięte.
    b. Użyć narzędzia `glb_rodo_tool.remove_alternatives()` podając jako argument katalog, w którym znajduje się plik konfiguracyjny.
5. Zanonimizować wskazane pola w bazie danych. W tym celu należy:
    a. Przygotować plik konfiguracyjny o nazwie `engines_configuration.xml` (patrz rozdz. 3.2), w którym definiujemy sposoby anonimizacji..
    b. Przygotować plik konfiguracyjny o nazwie `fields_configuration.xml` (patrz rozdz. 3.2). W ramach konfiguracji określa się, które pola moją podlegać anonimizacji i wskazuje sposób anonimizacji osobno dla każdego pola.
    c. Użyć narzędzia `glb_rodo_tool.anonymize_fields()` podając jako argument katalog, w którym znajduje się plik konfiguracyjny.
6. Zanonimizować użytkowników w bazie autoryzacyjnej. W tym celu należy:
    a. Przygotować plik konfiguracyjny o nazwie `users_configuration.xml` (patrz rozdz. 3.2). W ramach konfiguracji określa się, czy mają być usunięci użytkownicy, czy tylko zanonimizowane ich pełne (opisowe) nazwy zawierające zwykle imię i nazwisko przy użyciu wskazanego silnika z pliku `engines_configuration.xml`.
    b. Użyć narzędzia `glb_rodo_tool.anonymize_users()` podając jako argument katalog, w którym znajduje się plik konfiguracyjny.
    c. Uwaga: po anonimizacji użytkowników należy powtórnie uruchomić aplikację Administratora, ponieważ okno z użytkownikami nie odświeży się.

3.2 Konfiguracja

Narzędzie wymaga do działania zdefiniowania plików konfiguracyjnych o następujących nazwach:

`engines_configuration.xml`	Definicja możliwych silników (sposobów) anonimizacji pól, np. czyszczenie pola, przypadkowa lista, generator sekwencyjny.
`fields_configuration.xml`	Określenie pól w poszczególnych bankach danych i tabelach, które mają podlegać anonimizacji oraz wskazanie silnika anonimizacji dla każdego z nich.
`users_configuration.xml`	Określenie sposobu anonimizacji użytkowników bazy danych Smallworld.
`alternatives_configuration.xml`	Określenie sposobu usuwania alternatyw i punktów kontrolnych.

Pliki te mogą być umieszczone w dowolnym katalogu, który należy potem wskazać uruchamiając metody anonimizacji.

Przykładowe pliki są zamieszone wewnątrz modułu `glb_rodo`. Są one opatrzone komentarzami objaśniającymi działanie poszczególnych parametrów.

3.3 Przykład użycia

Przykładowe wywołania realizujące wyczyszczenie alternatyw i anonimizację bazy:

```
glb_rodo_tool.remove_alternatives("C:\Moja_konfiguracja_RODO")

glb_rodo_tool.anonymize_fields("C:\Moja_konfiguracja_RODO")

glb_rodo_tool.anonymize_users("C:\Moja_konfiguracja_RODO")
```

4 Środki ostrożności

Działanie narzędzi i modułu wykonuje masowe operacje na bazie danych prowadzące do nieodwracalnego usunięcia wielkiej ilości danych lub struktur. Dlatego należy bezwzględnie uruchamiać je na kopii bazy.

W celu upewnienia się, że jest się we właściwej bazie danych (na właściwej konsoli), zaleca się wywołanie następującej metody przed każdym uruchomieniem operacji wymienionych w rozdz. 3.3:

```
glb_rodo_tool.check()
```

która wypisze na konsoli ścieżkę do produktu i poszczególnych baz danych.

5 Wskazówki

1. Jeśli nazwy użytkowników (loginy) mają postać czytelnych danych osobowych, np. jkowalski, i chcemy usunąć tych użytkowników z bazy autoryzacyjnej, należy pamiętać także o usunięciu tych nazw z innych tabel, które mogą je przechowywać. Dotyczy to w szczególności rejestru zmian (glb_object_history). Stosowny przykład jest podany w przykładowym pliku konfiguracyjnym.
2. Moduł jest w stanie wychwycić i obsłużyć niektóre błędy popełnione w pliku konfiguracyjnym – błąd zostanie wypisany na konsoli i moduł przejdzie do wykonywania następnej czynności (np. anonimizacji kolejnego pola). W takim przypadku można poprawić błąd, zakomentować w konfiguracji poprawnie wykonane czynności i ponownie uruchomić anonimizację. W przypadku innego rodzaju nieprzewidzianych błędów nastąpi przerwanie działania modułu i dalsze czynności nie zostaną wykonane. Trzeba więc śledzić postęp operacji wypisywany na konsolę w celu stwierdzenia, które operacje powiodły się lub nie.
3. Zaleca się konsultować sposób przygotowania bazy (konfigurację narzędzia) z odbiorcą tej bazy, aby uniknąć sytuacji, że nadmierne usunięcie alternatyw lub danych utrudni lub uniemożliwi świadczenie usługi przez tego odbiorcę.
4. Osoba korzystająca z narzędzia powinna mieć odpowiednie szerokie uprawnienia do zmian w bazie danych, aby nie natknąć się na komunikat o braku dostępu. Typowo powinien być to użytkownik *root*.
5. W przypadkach wykraczających poza funkcjonalność modułu, należy zwrócić się do dostawcy aplikacji w celu stworzenie specyficznego algorytmu anonimizującego.