Uživatelská příručka

Aplikace GEM Winch

Příručka pro verzi aplikace 3.1

Září 2018

Obsah

Historie dokumentu

Verze	Autor	Datum	Popis
1.3	Martin Kreidl	6.5.2016	Popis pro verzi aplikace 1.3
1.4	Martin Kreidl	13.5.2016	Aktualizace pro verzi aplikace 1.4
1.4	Martin Kreidl	26.5.2016	Aktualizace pro verzi aplikace 1.4 – upraven popis parametrů a tříd / funkcí
3.0	Jiří Mlejnek	6. 9. 2017	Aktualizaci pro verzi aplikace 3.0
3.1	Jiří Mlejnek	16.1.2018	Aktualizace pro verzi 3.1
3.1.1	V. Davidík	20.4.2018	Revize a doplnění návodu pro spuštění
3.1.2	Jiří Mlejnek	13.6.2018	
3.1.3	V. Davidík	10.8.2018	Doplnění návodu pro PostgreSQL prostředí.

1 Úvod

Tato dokumentace popisuje postup při implementaci a použití nástroje GEM Winch pro anonymizaci a řezy dat. Anonymizace je proces, který nenávratným způsobem odstraňuje z evidovaných dat osobní údaje, jehož cílem je zajistit řešení těchto požadavků:

- Anonymizace dat v produkčních systémech, která již nejsou aktivní (pominul účel jejich zpracování), ale jejich odstranění by vedlo k nežádoucím dopadům na statistiky, konzistence dat, apod.
- Vytváření balíčků anonymizovaných dat pro vývojové a testovací prostředí obchodních systémů (aplikací) s možností předání interním i externím vývojářům a testerům.

Pro splnění uvedených požadavků je navržen koncept řešení procesu anonymizace dat s využitím nástroje GEM Winch. Implementace je v kontextu požadavků rozdělena na následující částí:

- Implementace procesů datových integrací pro přenos dat pro anonymizace a jejich následnou propagaci do produkčního prostředí včetně kontrolních reportů.
- Dodávka nástroje GEM Winch a konfigurace nástroje pro anonymizaci dat včetně návrhu způsobu anonymizace pro jednotlivé datové položky dle datového modelu ve variantách pro produkci a test.

2 Anonymizace a řezy dat

2.1 Důvody a požadavky pro anonymizaci a řezy dat

Základními důvodem pro řešení anonymizace dat je splnění legislativních a regulačních opatření na ochranu dat v souladu se Zákonem 101/2000 Sb. o ochraně osobních údajů.

Tento zákon vymezuje klíčové pojmy z hlediska účely tohoto zákona následovně:

a) osobním údajem je jakákoliv informace týkající se určeného nebo určitelného subjektu údajů. Subjekt údajů se považuje za určený nebo určitelný, jestliže lze subjekt údajů přímo či nepřímo identifikovat zejména na základě čísla, kódu nebo jednoho či více prvků, specifických pro jeho fyzickou, fyziologickou, psychickou, ekonomickou, kulturní nebo sociální identitu,
b) citlivým údajem je osobní údaj vypovídající o národnostním, rasovém nebo etnickém původu, politických postojích, členství v odborových organizacích, náboženství a filozofickém přesvědčení, odsouzení za trestný čin, zdravotním stavu a sexuálním životě subjektu údajů a genetický údaj subjektu údajů; citlivým údajem je také biometrický údaj, který umožňuje přímou identifikaci nebo autentizaci subjektu údajů,
c) anonymním údajem je takový údaj, který buď v původním tvaru, nebo po provedeném zpracování, nelze vztáhnout k určenému nebo určitelnému subjektu údajů,
d) subjektem údajů je fyzická osoba, k níž se osobní údaje vztahují,

Dalšími důvody pro anonymizaci jsou:

- Možnost poskytnutí konzistentních dat a přitom neobsahujících osobní údaje pro vývojová a testovací prostředí uvnitř i vně společnosti
- Úspora času/financí a zvýšení kvality testování použitím reálných množin dat oproti generování „umělých“ testovacích dat

- Ochrana před zneužitím dat společnosti v neprodukčních prostředí, ve kterých nemusí být zajištěn přístup k datům se stejnými restrikcemi, jako na produkčních systémech.

Hlavním účelem datových řezů je minimalizace dat. Smyslem použití této techniky je snížení požadavků na hardware v neprodukčních prostředích. Datové řezy jako takové nejsou přímo určeny k zabezpečení dat, současně však omezením objemu dat v testovacím prostředí částečně chráníme data před jejich zneužitím.

2.2 Anonymizační metody

Anonymizace je trochu jiný přístup k ochraně dat, než je například šifrování (pseudoanonymizace). Šifrováním data pouze převedeme do nečitelné podoby, ale stále existuje klíč, kterým lze získat data původní. Kdežto anonymizovaná data stále vypadají jako reálná a co je hlavní, neobsahují žádné osobní/citlivé informace. Správně provedená anonymizace neumožňuje návrat k původním datům. Obecné techniky, používané pro anonymizaci jsou uvedeny v příloze č. 1.

Proces anonymizace a řezu dat by měl z obecného pohledu splňovat následující podmínky:

- Anonymizační proces musí anonymizovat všechny datové položky, podle kterých by mohlo dojít k identifikaci fyzické případně i právnické osoby
- Pokud datové položky splňují nějaká matematická nebo formální kritéria, musí stejná pravidla platit i na anonymizovaných datech.
- Stejné datové položky v různých datových objektech mají být anonymizovány stejným způsobem.
- Řez dat musí být primárně proveditelný přes jednotlivé aplikace.

Výběr anonymizační metody závisí nejvíce na typu datové položky (osobního údaje). V kontextu informačních systémů jsou nejčastějšími osobními údaji pro anonymizaci:

- Rodné číslo
- Datum narození

- Jméno a příjmení
- IČ
- DIČ

a další údaje, které ve spojení s výše uvedenými mohou představovat osobní případně i citlivé údaje, například:

- Bankovní spojení
- Telefonní spojení
- Mailová adresa
- Adresa trvalého bydliště
- Datová schránka
- …

3 GEM Winch – nástroj pro anonymizaci a řez dat

3.1 Popis nástroje GEM Winch

GEM Winch je modulární systém pro anonymizaci a řez dat. Je tvořen dvěma komponentami:

- Winch Add-In pro Enterprise Architect, který slouží k načtení struktury datových objektů pro anonymizaci a řez dat, uživatelskému nastavení konfigurace / parametrů. Dále umožňuje spouštět jednotlivé kroky vlastního procesu anonymizace, jakými jsou: vygenerování příslušných skriptů, nasazení skriptů do db a jejich spuštění. Vlastní provedení těchto kroků však provádí GEM Winch Actor.
- GEM Winch Actor: Program spustitelný z příkazové řádky, který na základě konfigurace v EA modelu provádí příslušné kroky.

Anonymizace a řez dat probíhají v následujících krocích:

- Instalace nástroje GEM Winch Add-In do Enterprise Architect
- Instalace nástroje GEM Winch Actor do jednotlivých databází (provádí se zvlášť pro každý databázový systém: MSSQL, Oracle, PostgreSQL, DB2, atd)
- Import existujících datových modelů do Enterprise Architect s využitím ODBC připojení – standardní funkce nástroje Enterprise Architect
- Konfigurace – nastavení předpisu pro anonymizaci a řez dat na základě připravených anonymizačních tříd a funkcí
- Vytvoření příslušných skriptů pro provedení anonymizace
- Nasazení a spuštění skriptů pro provedení anonymizace

3.2 Podmínky a postup pro užití nástroje Winch

Základní podmínkou pro užití nástroje GEM Winch pro anonymizaci a řez dat je, jak vyplývá z výše uvedeného popisu, instalace a licence SW Enteprise Architect. Dodání

SW Enterprise Architect a licence k jeho užití není předmětem dodávky GEM System. Dalšími předpoklady jsou:

- Instalace a konfigurace 32 bitové verze ODBC driveru pro všechny požadované databáze.

3.3 Práce s DB a DB schématy v nástroji GEM Winch při anonymizaci

Pokud chceme provádět anonymizaci a řez dat prostřednictvím nástroje GEM Winch, je třeba zvážit koncept rozvržení databází a schémat.

Do procesu anonymizace a řezu dat vstupují potenciálně až tři oddělená databázová schémata:

- DB a schéma pro vlastní GEM Winch Actor: obsahuje slovníkové tabulky, anonymizační funkce a dále se do něho nasazují vlastní anonymizační skripty v podobě PLSQL/TSQL procedur.
- DB a schéma se zdrojem dat pro anonymizaci
- DB a schéma pro uložení anonymizovaných dat

Ke zdrojové databázi je nutné zajistit read-only přístup, současně zde musí být připravena schémata objektů pro anonymizaci. Pro vlastní anonymizaci by mělo platit, že schéma nástroje GEM Winch a schéma pro uložení výsledků by nikdy neměla být totožná s produkční databází.

Z hlediska obsazení prostor v databázi je třeba počítat s tím, že v cílovém DB/schématu bude uložena kopie anonymizovaných dat.

4 Předpoklady užití nástroje GEM Winch pro anonymizaci dat

4.1 Analýza prostředí

Aplikace informačního systému pracují s datovými položkami, obsahujícími osobní údaje v různých datových podobách.

Na základě analýzy datového modelu aplikací, určených pro anonymizaci musí být identifikovány všechny typy položek, obsahujících osobní údaje, které je nutné anonymizovat:

- Jméno fyzické osoby, které může nabývat podob:
    - Jméno
    - Příjmení
    - Příjmení a jméno, titul
    - Příjmení a jméno, titul, další identifikátor
- Telefon/Fax
- Bankovní spojení, které může nabývat podob:
    - Předčíslí účtu
    - Číslo účtu
    - Předčíslí + číslo účtu
    - Směrový kód banky
    - IBAN
- Adresa, která může nabývat podob:
    - Ulice, číslo
    - Obec, číslo
    - PSČ Obec
    - PSČ
- Email
- Datum narození, které může nabývat podob:
    - YYYYMMDD (řetězec)
    - YYYY-MM-DD (datum)
    - CC + YY + MM + DD (řetězce)
- Rodné číslo, které může nabývat podob:

- YYMMDD/NNNK resp. YYYMMDD/NNN
- YYMMDDNNNK resp. YYMMDDNNN
- NNNNNNNN (jiné ID)

Pro každou uvedenou položku pak je navržena anonymizační třída, která s užitím anonymizační funkce a jejích parametrů anonymizuje data podle definovaných pravidel.

5 Postup práce s nástrojem GEM Winch Add-In při anonymizaci

5.1 Instalace a odinstalace nástroje GEM Winch jako Add-Ins do EA

5.1.1 Instalace nástroje GEM Winch Add-in

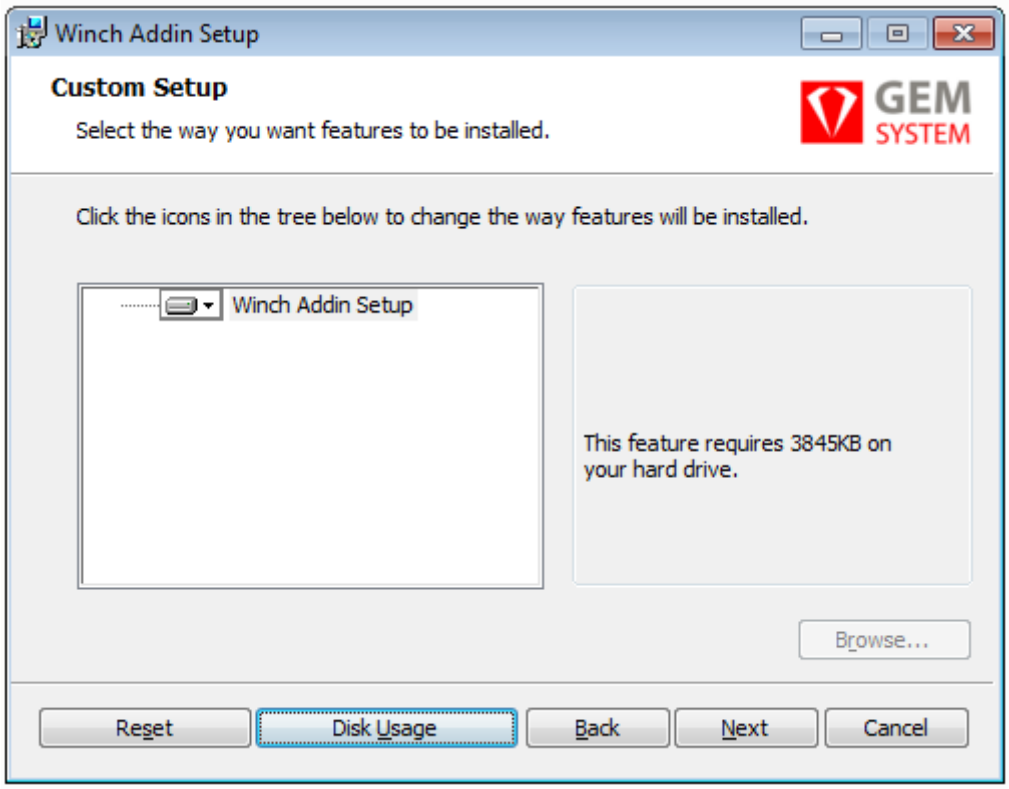
Pro instalaci modulu GEM Winch Add-In do SW Enterprise Architect je určen balíček WinchSetup-version.msi. K provedení instalace musíte mít práva administrátora.

Postup instalace:

- Před instalací zavřete aplikaci Enterprise Architect.
- Z instalačního média se spustí soubor WinchSetup-version.msi.

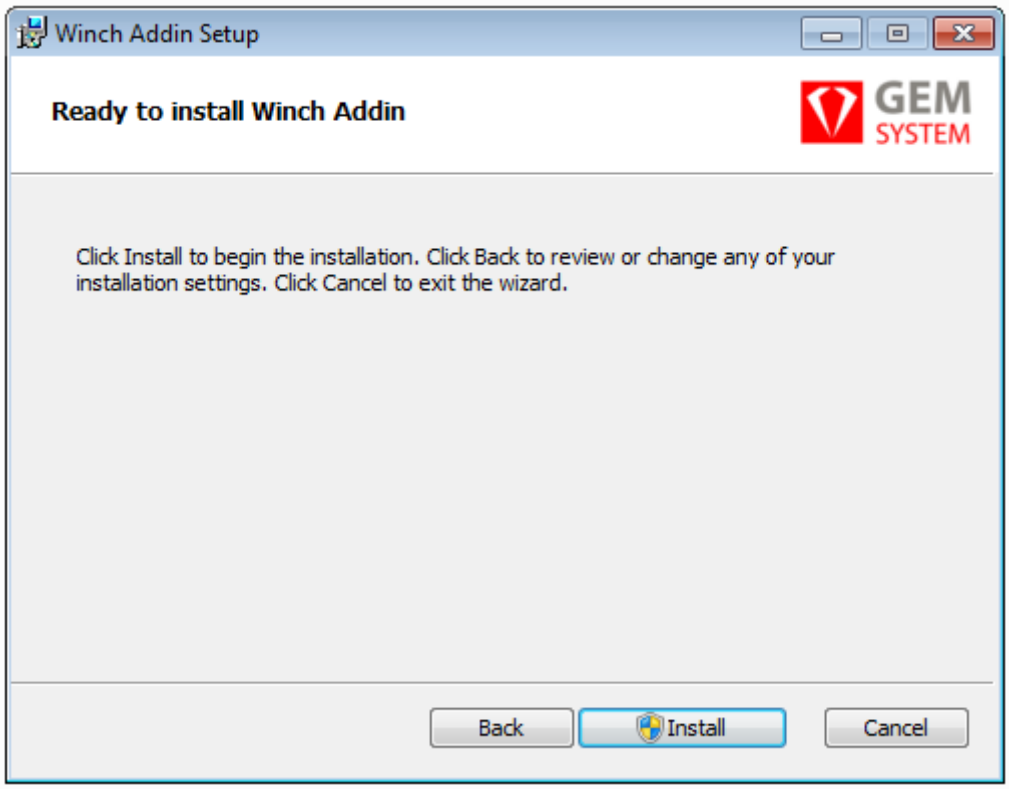
Obr. 1:Spuštění instalace Winch Add-in.

- Zobrazí se průvodce instalací.

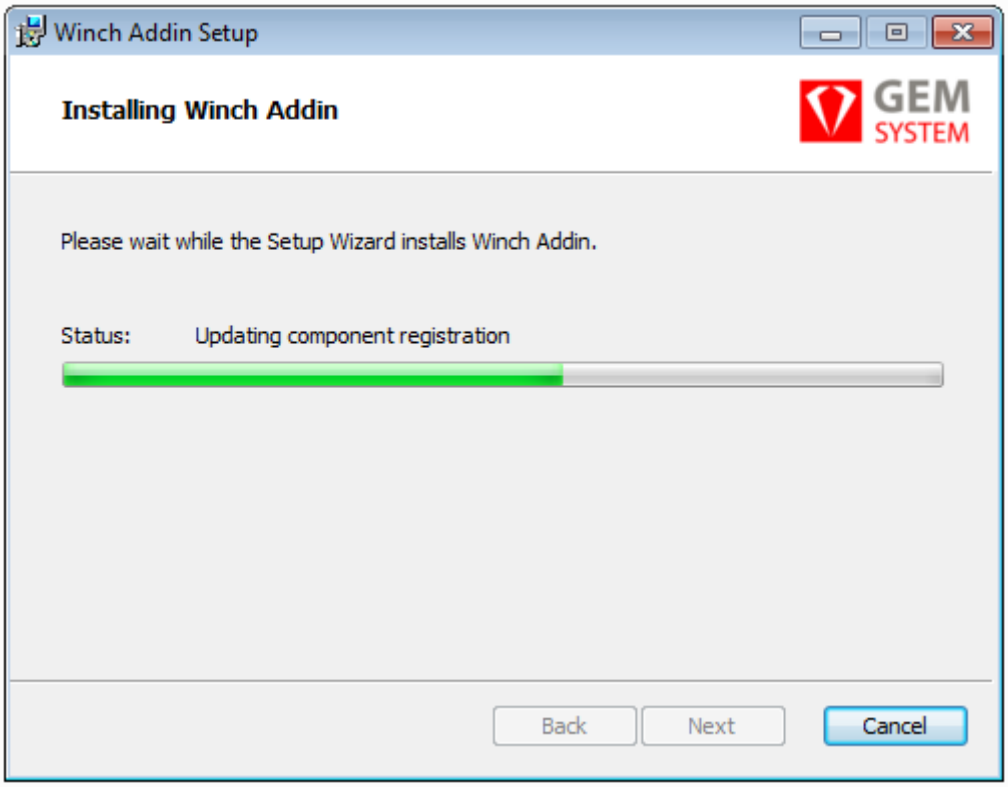


Obr. 2:Nastavení parametrů instalace Winch Add-in.

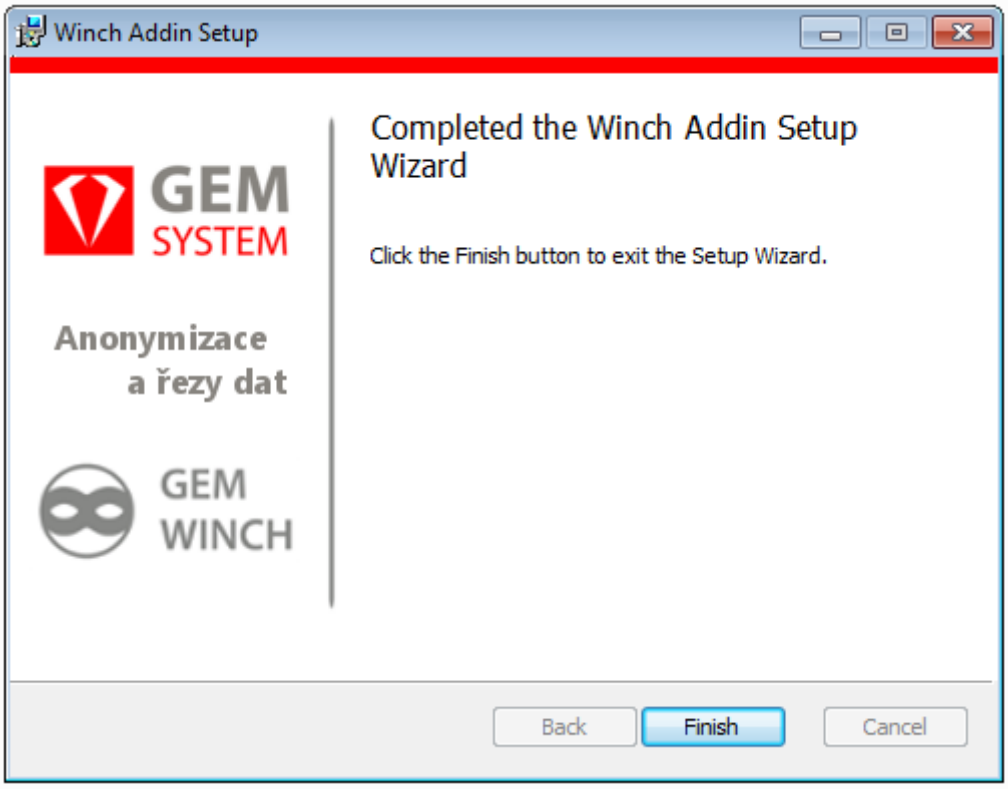
- Po kliknutí na tlačítko ***Next*** (Další) se objeví obrazovka se stručnou charakterizací instalovaného addinu a informací o průběhu instalace. Potvrďte tlačítkem ***Next*** (Další). V posledním okně stiskněte tlačítko ***Install*** a instalátor provede instalaci.



Obr. 3:Potvrzení instalace Winch Add-in.



Obr. 4:Zobrazení průběhu instalace Winch Add-in.



Obr. 5:Ukončení instalace Winch Add-in..

- V posledním kroku instalátor zobrazí zprávu o úspěšném provedení instalace. Stiskněte tlačítko ***Finish***

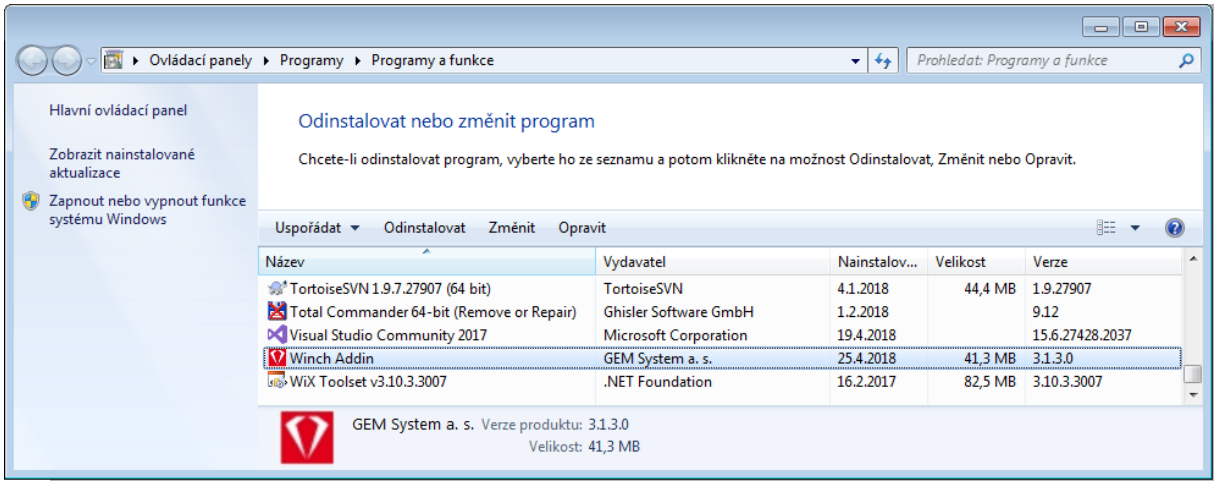
V rámci instalace instalátor nahraje knihovnu (WinchAddin.dll) s Add_in pro Enterprise Architekta do „C:\Program Files (x86)\GEM System a. s\Winch Addin“, tuto knihovnu zaregistruje do registrů Windows a doplní údaje potřebné ke spolupráci s Enterprise Architektem.

Po provedení instalace je nutné ještě aktivovat zakoupenou licenci a tím zpřístupnit jednotlivé funkčnosti nástroje. Licence je distribuována v souboru „license.txt“ a je nutné ji nahrát do složky, kam byl nástroj nainstalován (C:\Program Files (x86)\GEM System a. s\Winch Addin\license). Tento licenční soubor obsahuje informace o majiteli licence a zakoupených komponentách. Licenční soubor byste měli obdržet společně s instalačním souborem nástroje GEM Winch.

5.1.2 **Odinstalace Winch Add-in**

Odinstalování modulu Winch Add-in je možné provést standardním způsobem prostřednictvím nástroje Ovládací panely / Programy / Odinstalovat program nebo instalátorem aktuální instalované verze WinchSetup-version.msi a zvolením Remove. K odinstalování aplikace musíte mít práva administrátora.

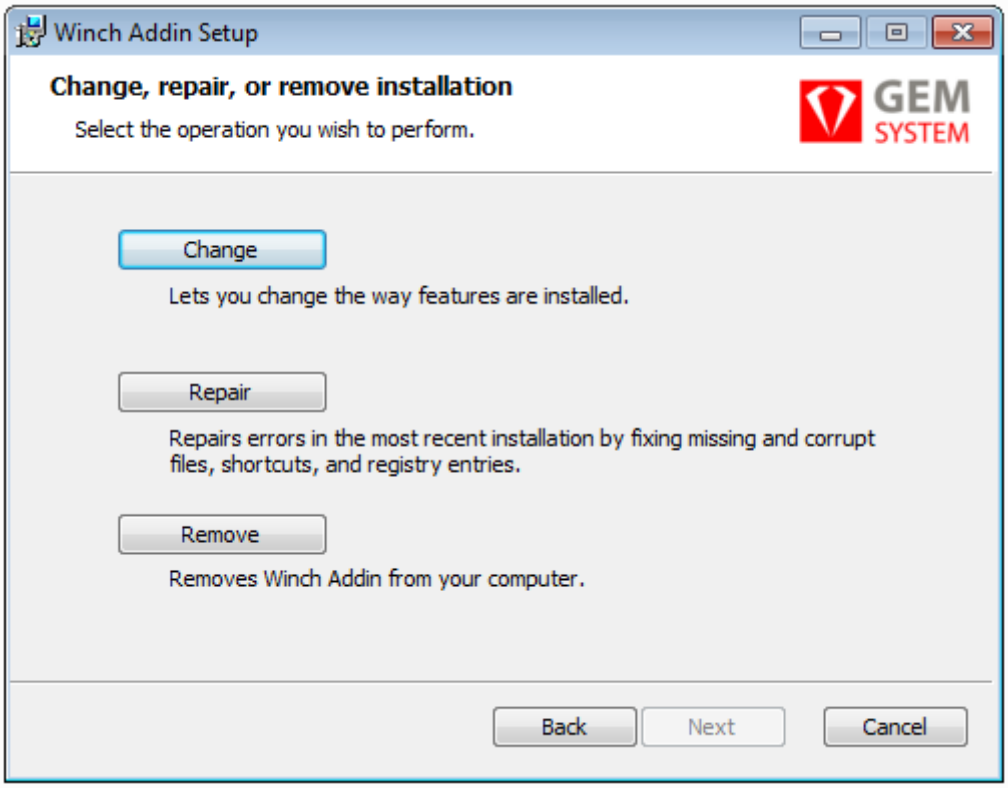
Při použití standardní odinstalace zvolíte WinchAddin:



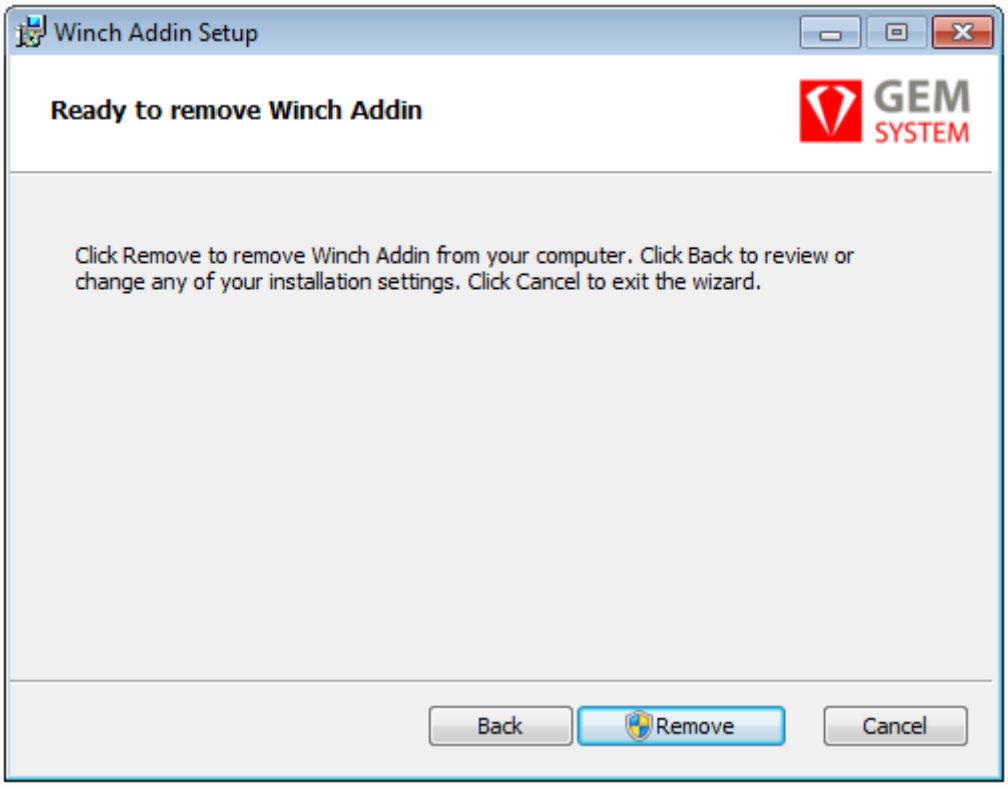
Obr. 6:Odinstalování Winch Add-in.

Použití WinchSetup-version.msi

Spusťte soubor WinchSetup.msi (instalovaná verze). Zobrazí se průvodce instalací. Po kliknutí na tlačítko ***Next*** se objeví okno s možnostmi ***Repair*** nebo ***Remove.*** Kliknutím na tlačítko ***Remove*** a následným potvrzením tlačítkem ***Remove*** bude aplikace odstraněna.



Obr. 7:Odinstalování pokračování.



Obr. 8:Odinstalování pokračování 2.

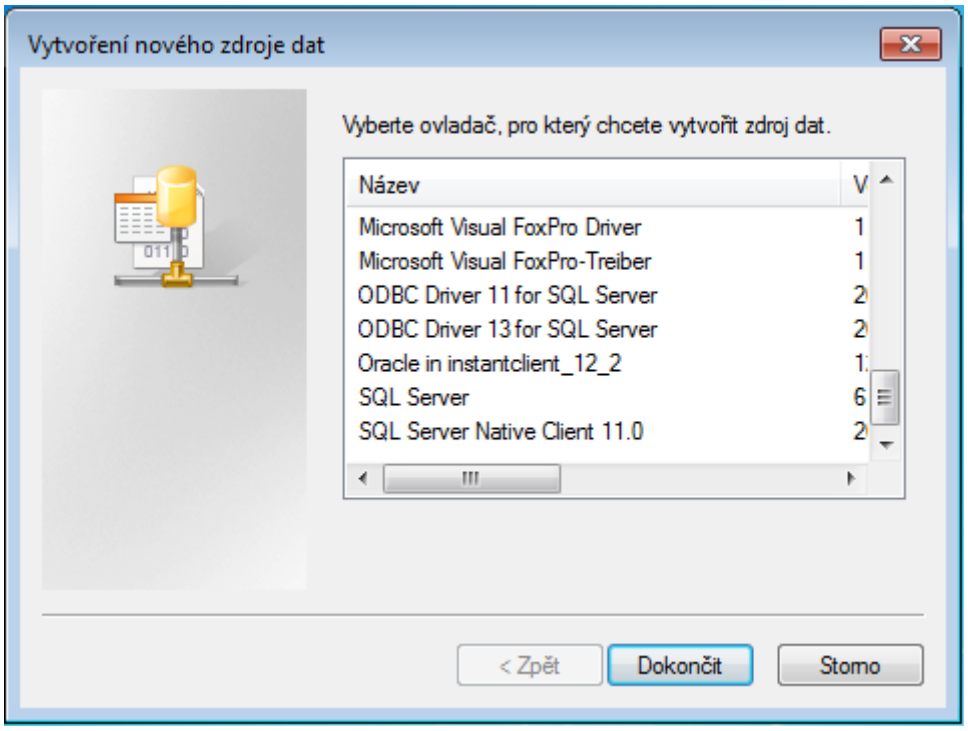
5.2 Načtení schémat pro anonymizaci prostřednictvím ODBC

Postup pro vytvoření ODBC připojení je odlišný, dle typu databáze. Následující kapitoly obsahují specifické popisy dle typu databáze.

5.2.1 Vytvoření ODBC připojení pro MSSQL

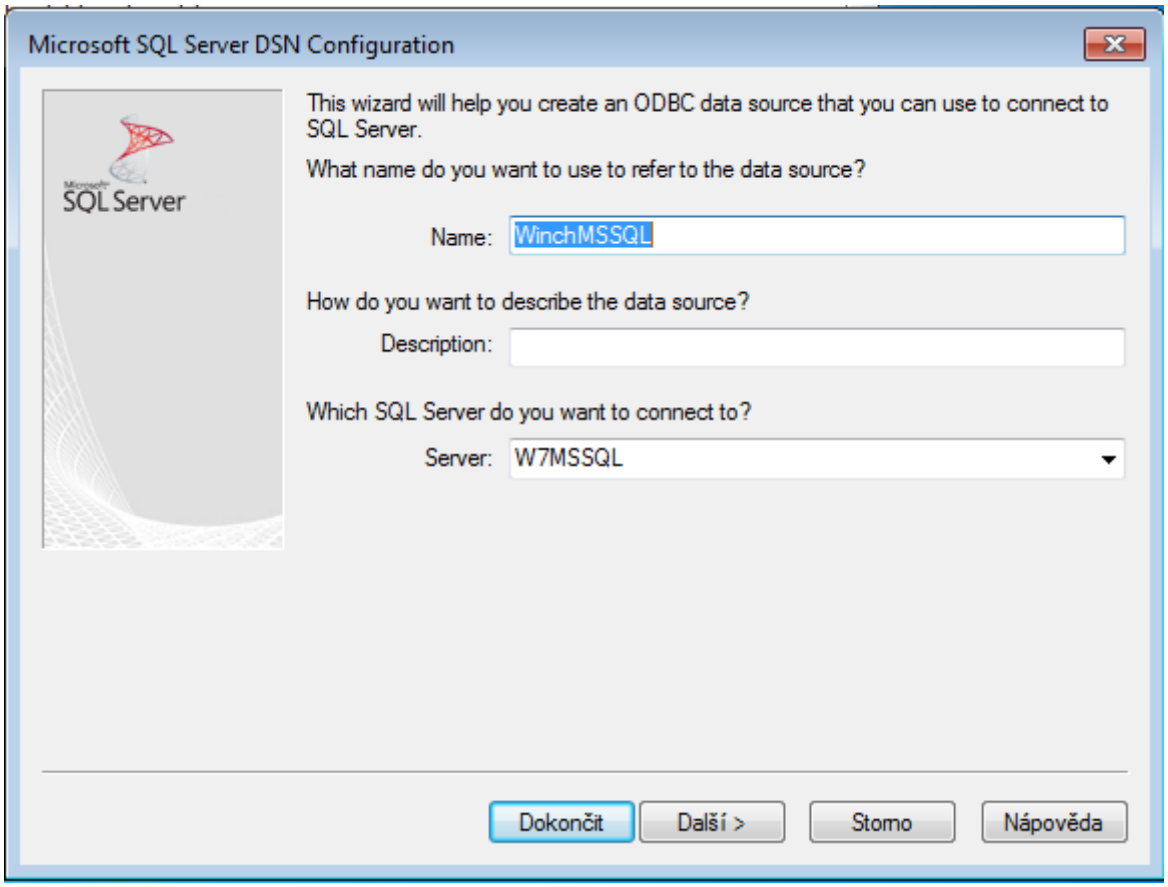
V následujícím odstavci si ukážeme postup pro načtení struktury anonymizovaných objektů z databáze.

Pokud chceme prostřednictvím ODBC načíst do Enterprise Architekta strukturu zdrojových tabulek, musíme založit nové ODBC spojení pro SQL Server. Spustíme správce zdrojů dat pro ODBC (je nutné využít 32 bitovou verzi). Po stisku tlačítka ***Přidat*** se objeví následující okno:



Obr. 9: Výběr druhu ODBC připojení.

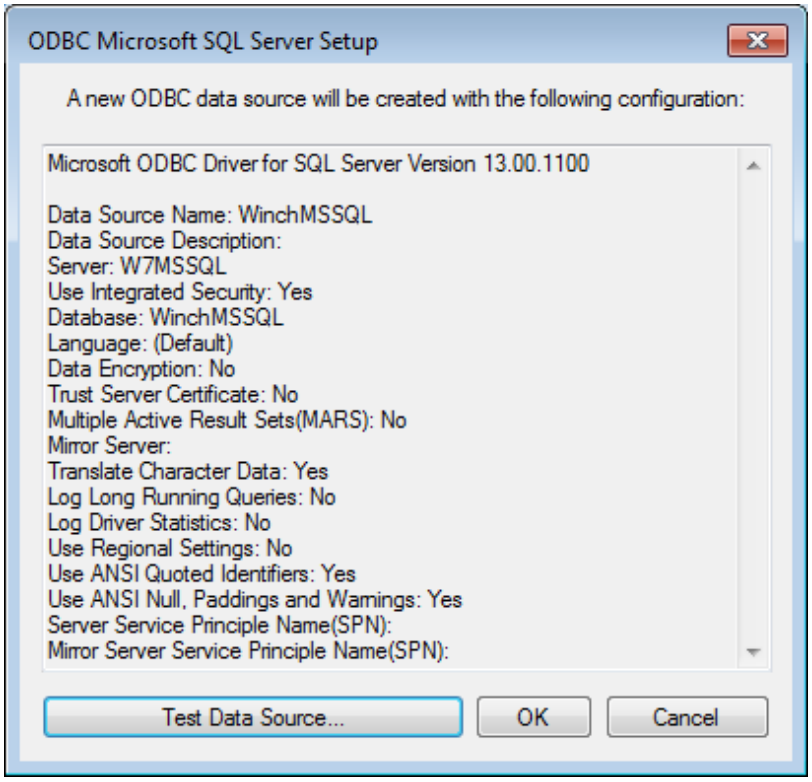
Vybereme ovladač, který chceme použít např. ODBC Driver13 for SQL Server nebo Oracle in instantclient_12_2, případně jiný podle verze db a instalovaných ovladačů. Následně se zobrazí obrazovka pro zadání parametrů ODBC připojení.



Obr. 10:Nastavení parametrů ODBC připojení.

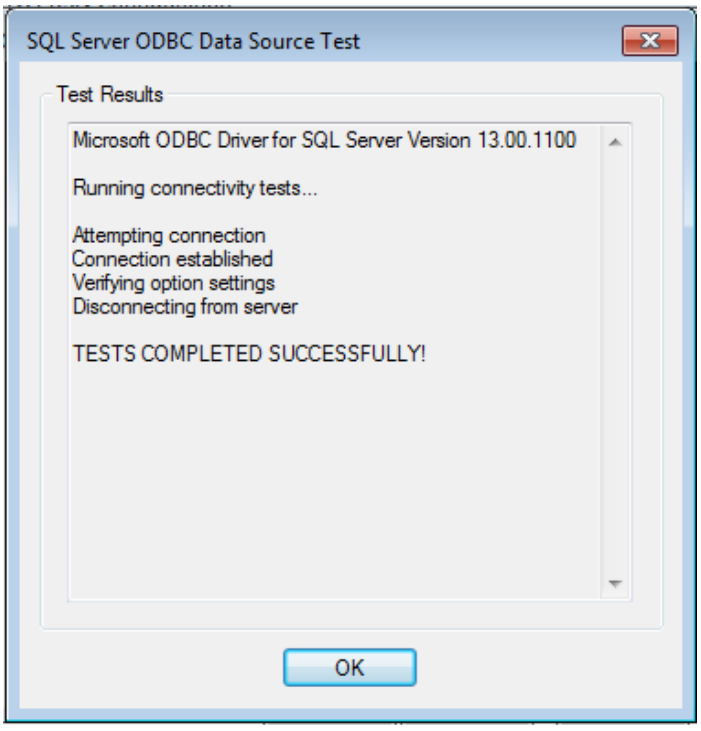
Zde vyplníme název nebo adresu serveru, na kterém je dostupná databáze resp. struktura objektů pro načtení do EA. Vyplňované údaje se mohou lišit dle typu databáze a verze ODBC ovladače.

Po stisku tlačítka ***Dokončit*** se objeví okno s rekapitulací parametrů připojení.



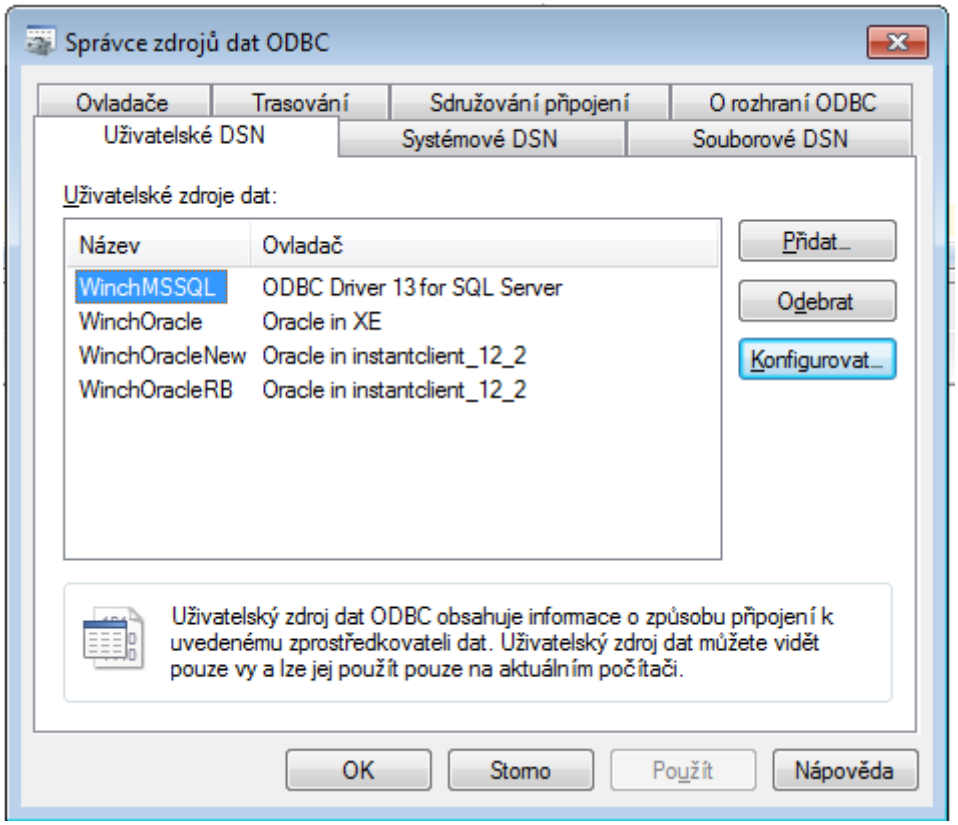
Obr. 11:Rekapitulace parametrů ODBC připojení.

Pomocí tlačítka můžeme provést ***Test zdroje dat*** a ověřit si tak, že připojení je nastaveno správně.



Obr. 12:Výsledek testu ODBC připojení.

V okně správce zdrojů pro ODBC pak přibude námi definované připojení.



Obr. 13:Zobrazení nového ODBC připojení.

5.2.2 **Vytvoření ODBC připojení pro ORACLE**

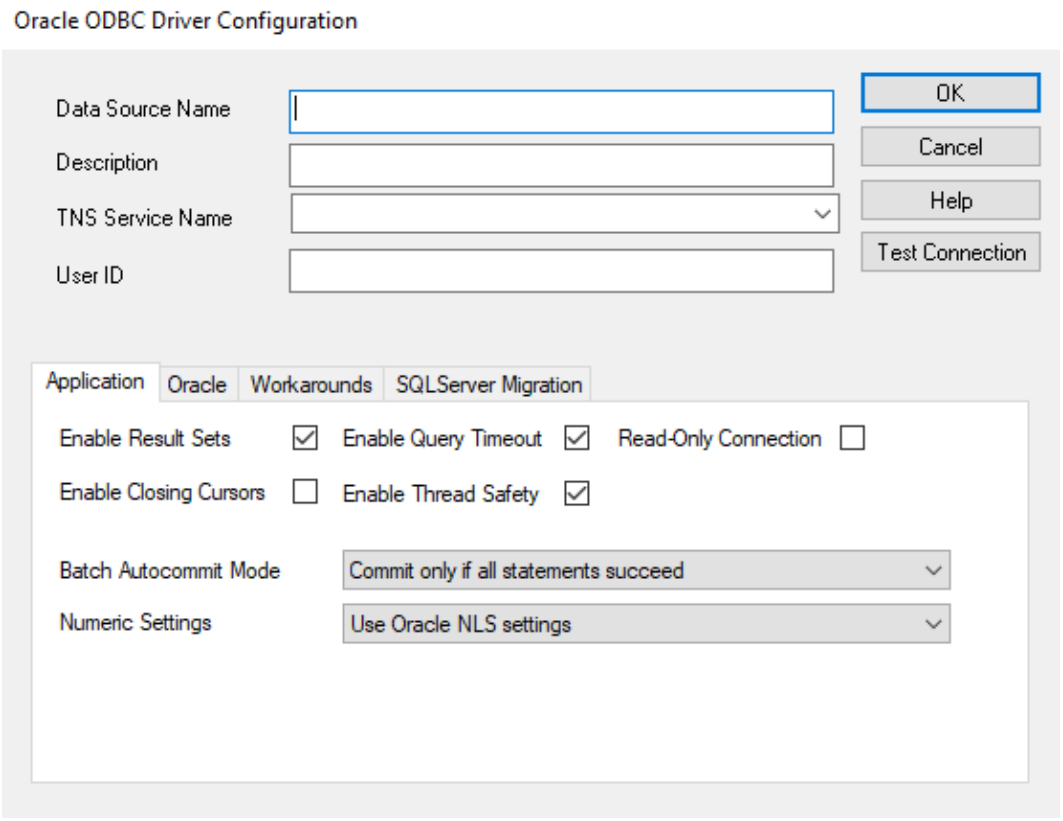
Vytvoření ODBC připojení pro ORACLE je podobné. Spustíme správce zdrojů dat pro ODBC (je nutné využít 32 bitovou verzi)

Jestliže se pracuje s WIN 7 64 bit systémem, je nutné pro správné fungování Enterprise Architectu mít nastaveno 32 bit ODBC připojení do databáze. Ve výchozím nastavení se pouští 64 bitová verze pro nastavování připojení (ODBC), aby se pustila 32 bit verze, je potřeba jej pustit z této cesty: C:\Windows\SysWOW64/odbcad32.exe a zde poté nastavit potřebné připojení

Pro ODBC připojení je potřeba mít instant klienta ve 32 bit verzi. Toho lze stáhnout z oficiálních stránek Oracle: http://www.oracle.com/technetwork/topics/winsoft-085727.html . Zde je poté potřeba stáhnout Instant Client Package - Basic a Instant Client Package - ODBC. Tyto 2 ZIP archivy poté rozbalit do stejné složky a spustit odbc_install.exe. Pro stahování z Oracle stránek je nutné mít vytvořenou registraci a souhlasit s licenčními podmínkami.

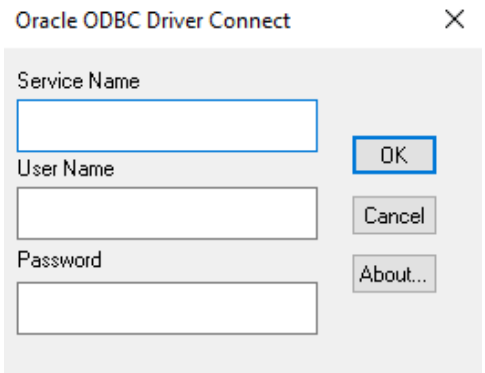
Dále je potřeba nastavit cestu systémové proměnné ORACLE_HOME. Výchozí nastavení: "%instantclient-odbc-nt-XXX\instantclient_XXX%".

Při vytváření ODBC spojení vyplníme název spojení, TSN Service Name v následujícím řetězci:"název serveru":"port"/"název db". A jméno uživatele, pod kterým budeme navazovat spojení.



Obr. 14:ORACLE ODBC konfigurace.

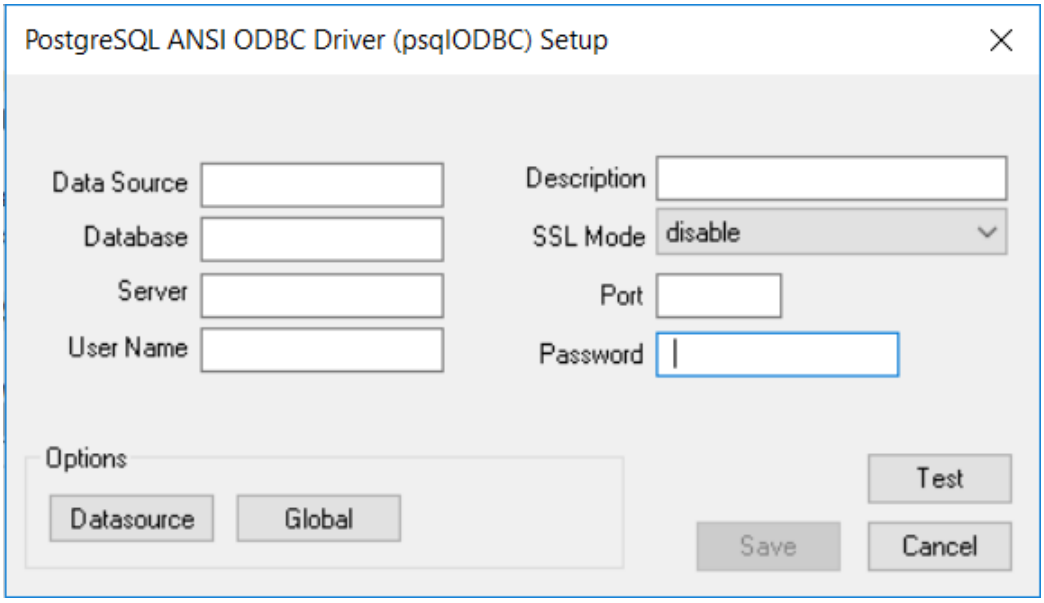
Po vyplnění všech potřebných parametrů a kliknutí na Test Connection, je možné spojení otestovat vložením hesla.



Obr. 15:Otestování ODBC ORACLE spojení.

5.2.3 **Vytvoření ODBC připojení pro PostgreSQL**

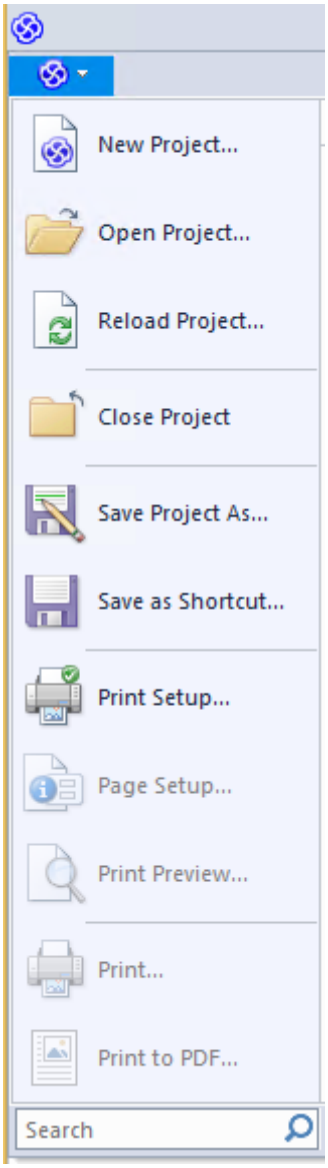
Spustíme správce zdrojů dat pro ODBC (je nutné využít 32 bitovou verzi) Na oficiálních stránkách postgreSQL: https://www.postgresql.org/ftp/odbc/versions/msi/ si stáhneme ODBC driver. Po instalaci by měl být driver dostupný v komponentech ODBC. Po spuštění se zobrazí okno, ve kterým se nadefinuji parametry pro propojení do DB.



Obr. ODBC postgreSQL konfigurace

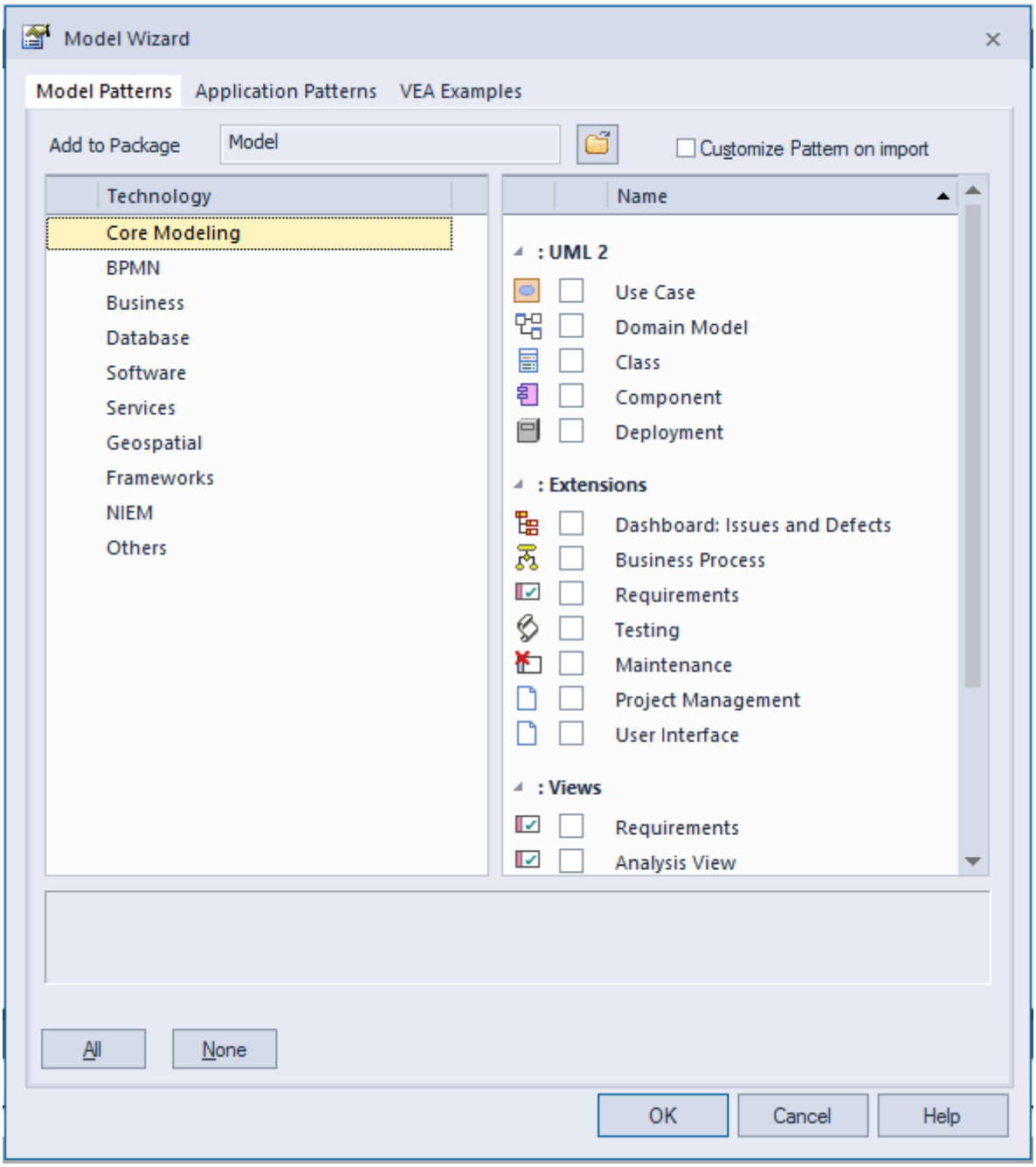
5.2.4 **Načtení struktury datových objektů pro anonymizaci do EA.**

Pro anonymizaci je v EA vhodné standardním způsobem založit nový projekt.



Obr. 16:Menu pro založení nového projektu v EA.

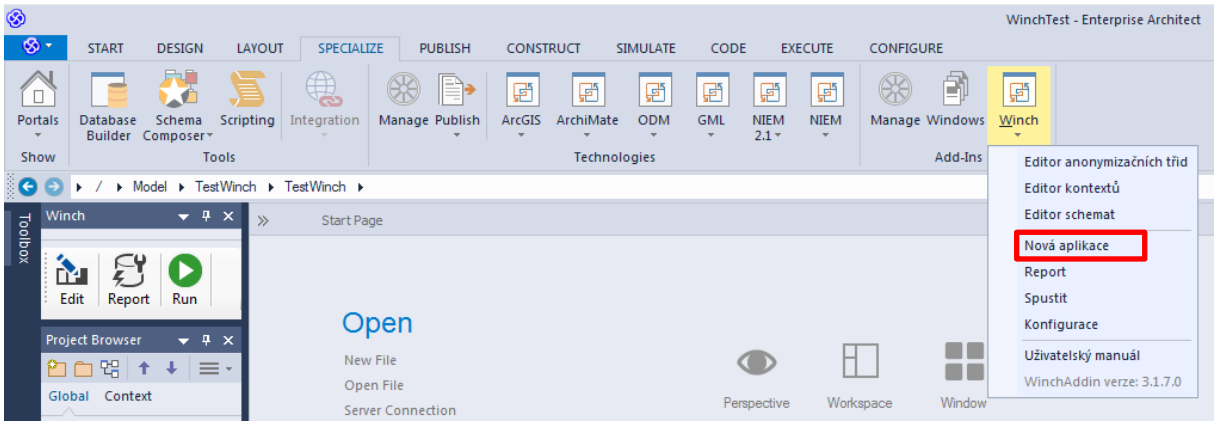
Nový model ponecháme prázdný. Na následující obrazovce stiskneme ***Cancel.***



Obr. 17:Zadání Data Modelu.

5.3 Vytvoření entity (aplikace) pro anonymizaci

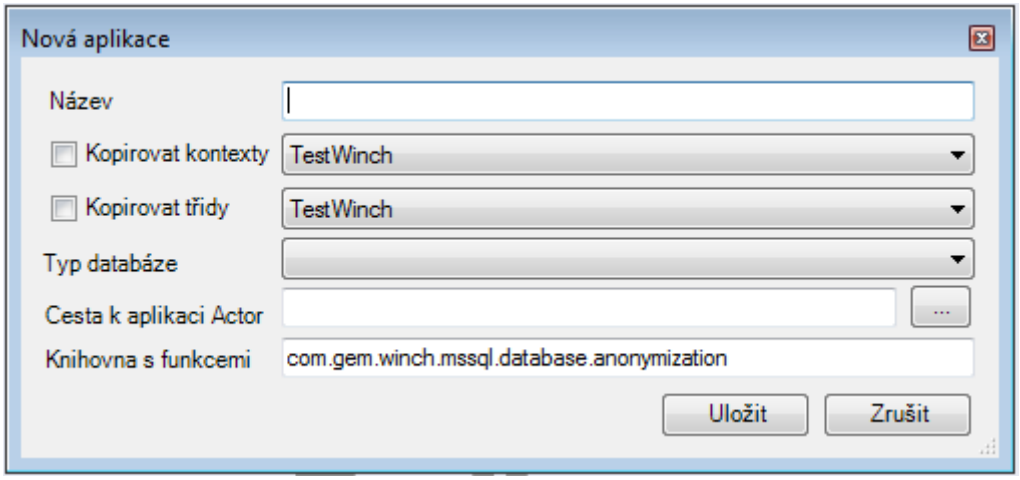
V nově vytvořeném modelu nejprve musíme připravit požadovanou strukturu pro nastavení anonymizačních tříd. Vybereme záložku ***SPECIALIZE*** a tlačítko ***Winch***. Ze zobrazeného menu vybereme položku ***Nová aplikace***



Obr. 18:Umístění ovládání Winch addinu.

Aplikace se založí pod package, na kterém stojí kurzor v Project Browseru.

Objeví se okénko pro zadání názvu nové aplikace, kterou chceme anonymizovat.



Obr- 19:Založení nové aplikace.

Vyplníme název aplikace a vybereme typ databáze, kterou chceme anonymizovat. Dále je nutné vyplnit cestu ke komponentě GEM Winch Actor pro danou databázi (Mssql/Oracle...). Jedná se o cestu k souboru disl-winch-mssql.bat nebo disl-winch-

oracle.bat, popřípadě jiné (dostupné defaultně: c:\Program Files (x86)\GEM System a. s\Winch Addin\disl-winch-mssql\bin\disl-winch-mssql.bat nebo c:\Program Files (x86)\GEM System a. s\Winch Addin\disl-winch-oracle\bin\disl-winch-oracle.bat.).

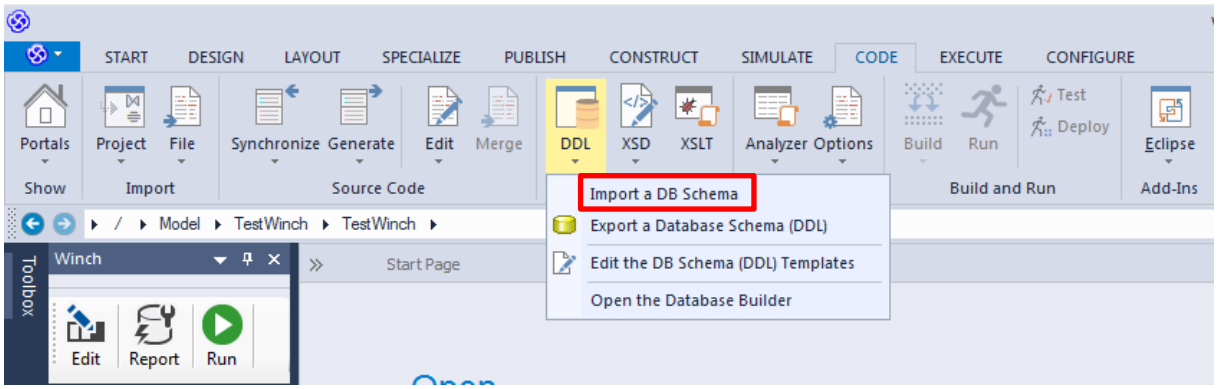
Po kliknutí na OK se nová aplikace zobrazí jako package v Project Browseru.

Pokud již v modelu existují jiné aplikace, máme možnost si do nové aplikace překopírovat existující anonymizační třídy a kontexty z této aplikace.

Následujícím krokem je import struktury databáze do modelu v EA. Tento krok je popsán v následující kapitole.

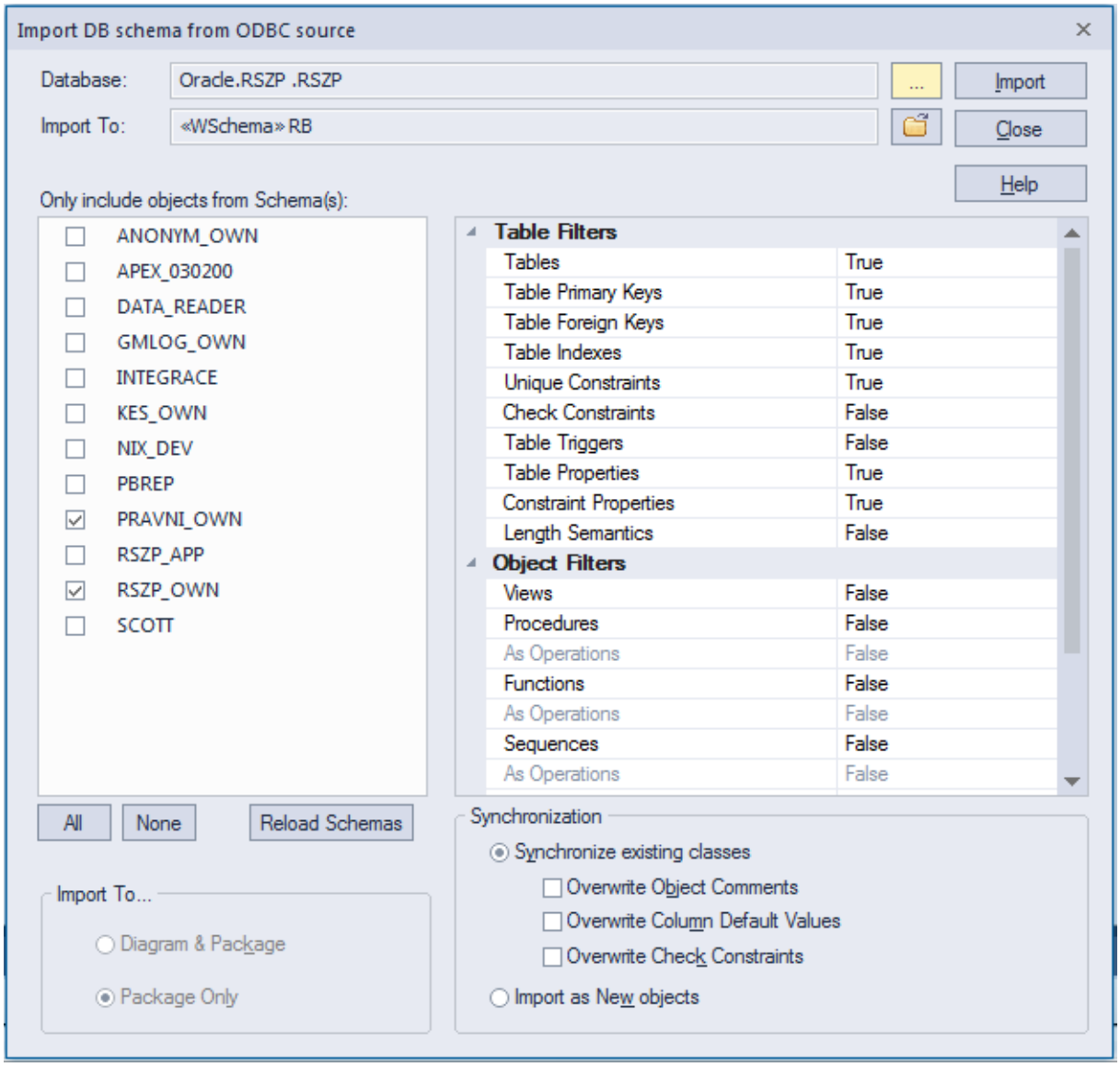
5.4 Import struktury db do ea

Pro import struktury databáze do modelu v EA lze využít funkci standardně dostupnou v nástroji Enterprise Architect. Nejprve označíme ve vytvořené aplikaci schéma/balíček do kterého chceme import provést a následně spustíme proces importu. To zle provést přes menu CODE-DDL-Import a DB Schema Import.



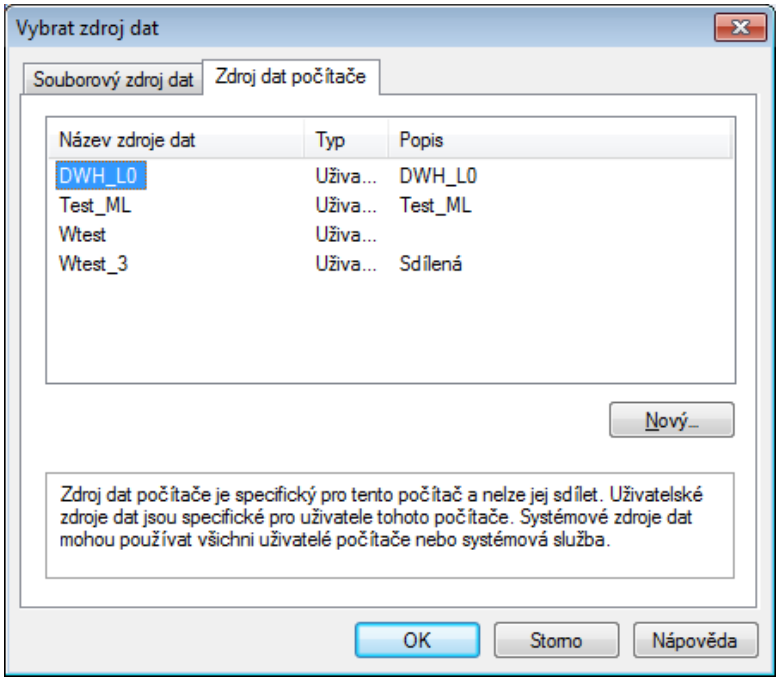
Obr. 20:Import schéma do DB.

Objeví se okno pro výběr ODBC připojení a schémat, které chceme do EA nahrát.



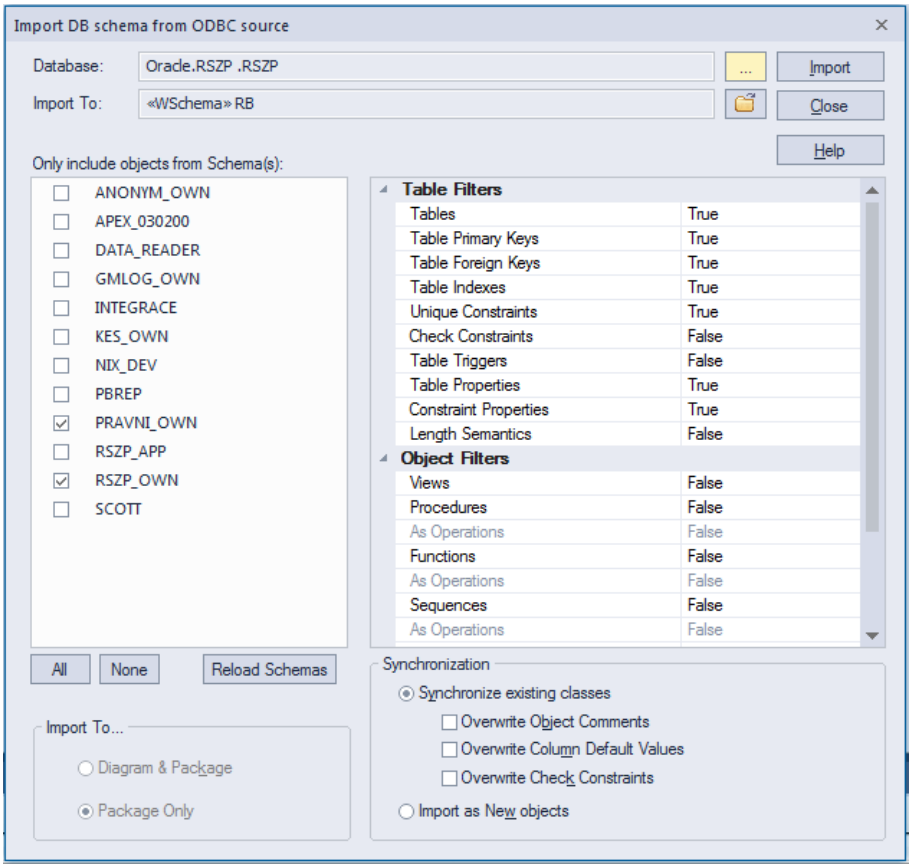
Obr.21:Nastavení ODBC připojení pro import do EA.

V případě, že není nastaven správný zdroj dat, klikneme na tlačítko [**...**] na řádku ***Database*** a vybereme požadované připojení na kartě ***Zdroj dat počítače***.



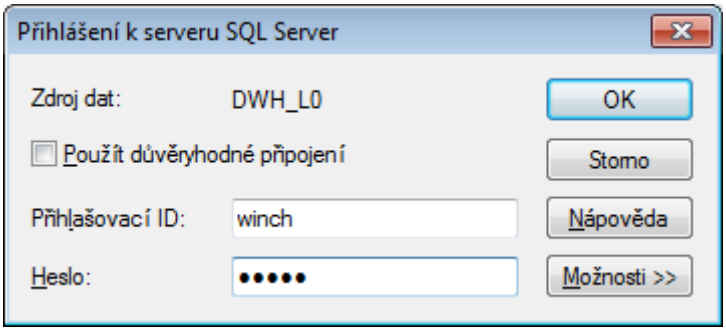
Obr.22:Výběr ODBC připojení pro import do EA

Po výběru se vrátíme do výchozí obrazovky



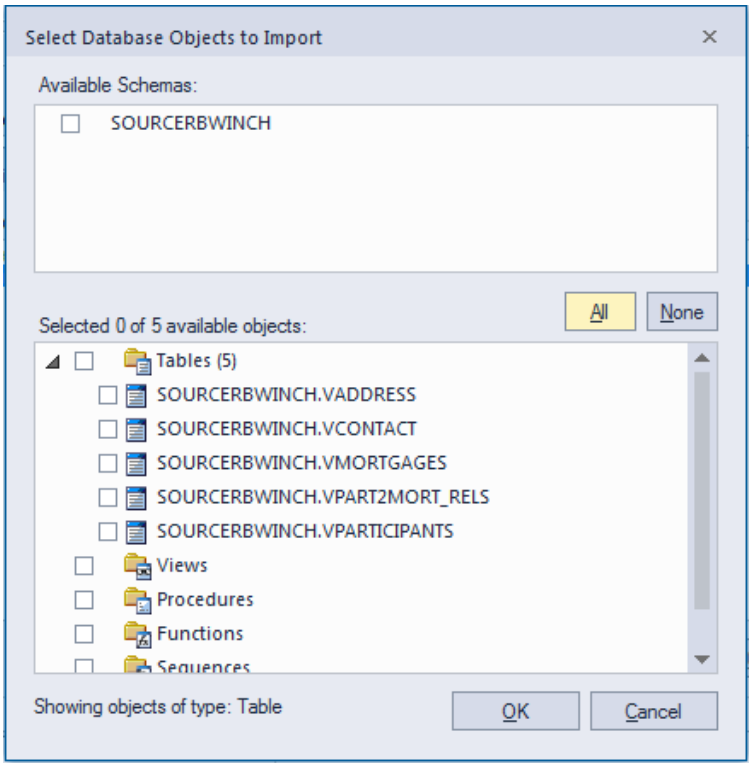
Obr.23:Výběr DB a objektů pro import.

Po nastavení zdroje klikneme na tlačítko ***Import*** a objeví se okno pro zadání přihlašovacích údajů.



Obr. 24:Zadání přihlašovacích údajů

Objeví se okno pro výběr konkrétních schémat nebo jednotlivých datových objektů. Pokud zatrhneme položky v podokně ***Available Schemas***, vyberou se všechny tabulky příslušné k těmto schématům. Pokud nezatrhneme, máme možnost vybrat jednotlivé tabulky.



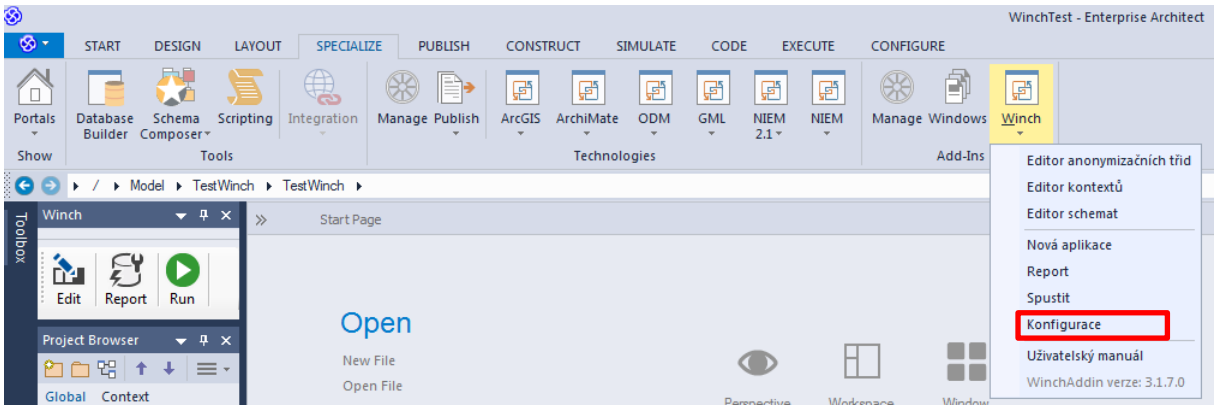
Obr. 25:Výběr objektů pro import do EA.

V následujícím okně ponecháme nastavení beze změny a stiskneme ***Import***.

Tímto jsme úspěšně završili přípravu části EA a můžeme přejít k vlastnímu nastavení parametrů anonymizace.

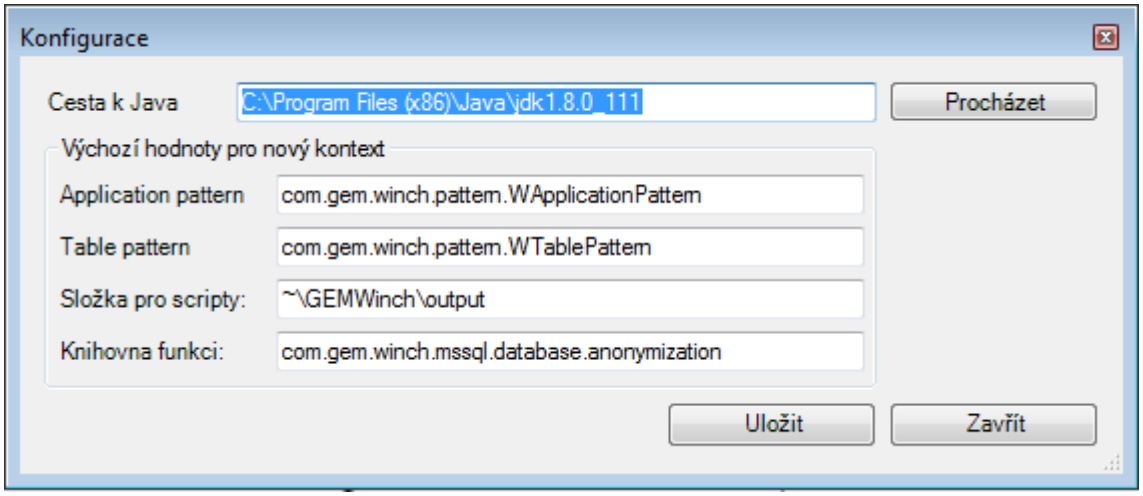
5.5 Konfigurace nástroje GEM Winch Add-in.

Předtím, než začneme se samotným nastavením anonymizačních tříd, nástroje GEM Winch Add-in, je nutné zkontrolovat konfiguraci. Dialog pro konfiguraci je dostupný v menu: SPECIALIZE-Winch-Konfigurace.



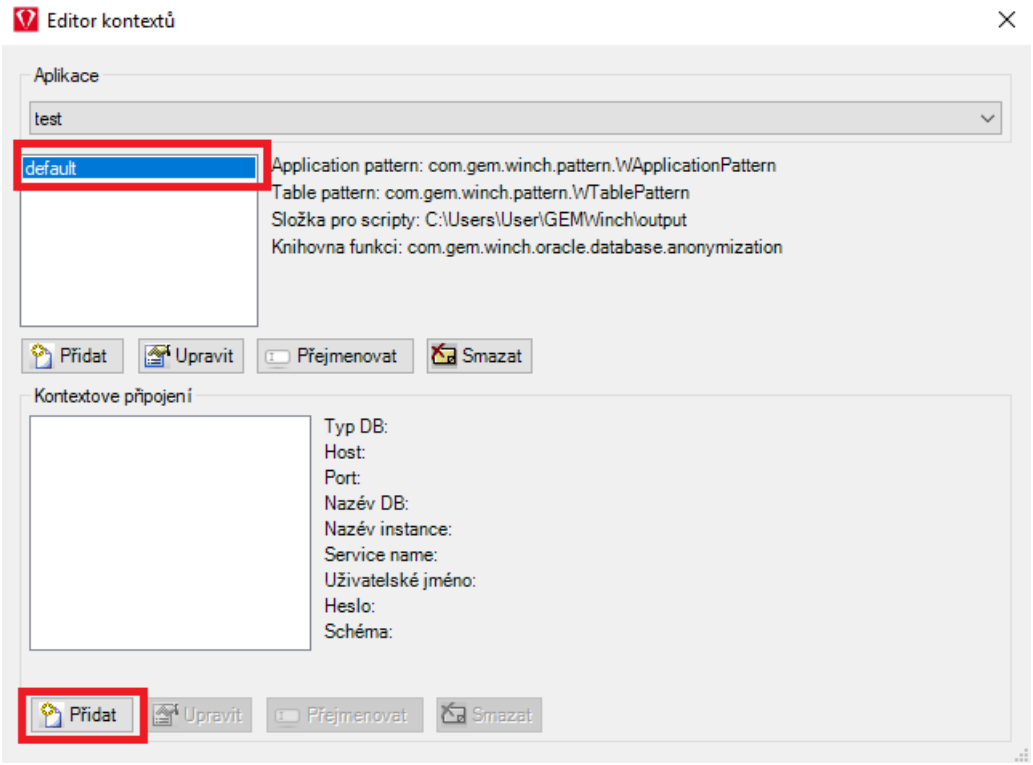
Obrázek 1Konfigurace nástroje

Zde je nutné nastavit cestu k Javě (vyžadována je 32 bitová verze) a zkontrolovat, zda se používá správná knihovna funkcí dle typu databáze, ve které pracujeme. Složka pro scripty obsahuje cestu, kam budou ukládané výstupní generované skripty pro anonymizaci. Výchozí nastavení předpokládá ukládání skriptů do domovského adresáře přihlášeného uživatele do složky GEMWinch/output. Tlačítkem Uložit vytvoříme novou aplikaci.



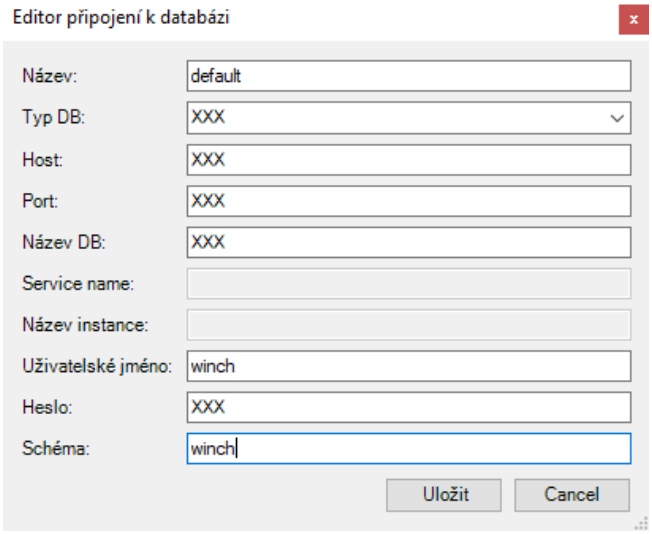
Obr. 26:Konfigurace winch add-in.

Dále je potřeba nastavit kontextové připojení aplikace. Pokud není k dispozici defaultní nastavení, zkontrolujeme přístup aplikace k JAVE.



Obr. 27:Editor kontextu.

Následně vytvoříme dvě aktivní kontextová připojení přes tlačítko přidat. Jedno s názvem **default**, kam se nasadí všechny potřebné struktury včetně zvolených anonymizovaných tabulek.

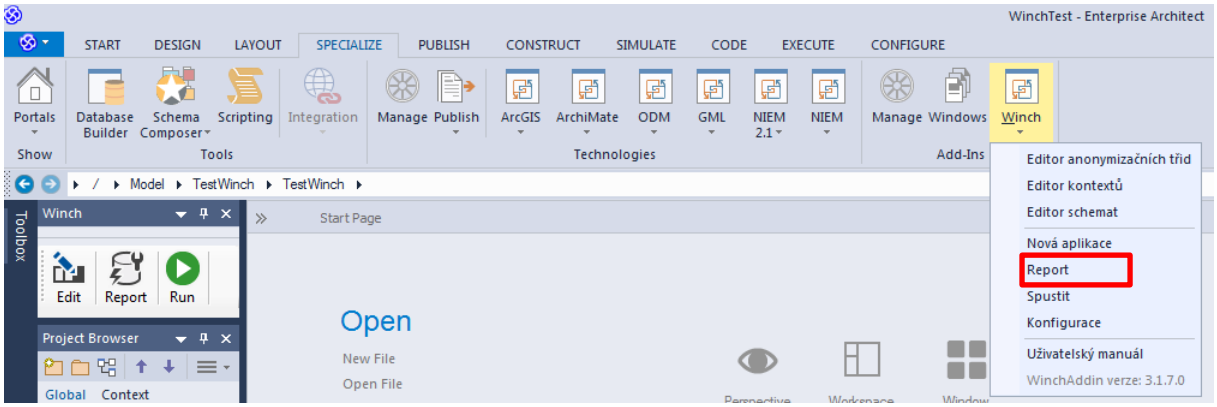


Obr. 28:Editor připojení k databázi.

A druhé pro konkrétní databázové schéma, ve kterém se nachází schéma jako zdroj tabulek pro anonymizaci.

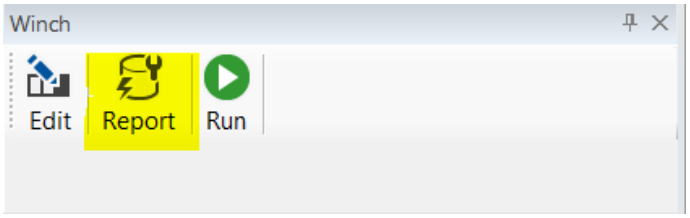
5.6 Nastavení konfigurace (parametrů) pro anonymizaci tabulek v rámci aplikace.

V dalším kroku přejdeme na obrazovku s přehledem dostupných tabulek pomocí menu SPECIALIZE-Winch-Report



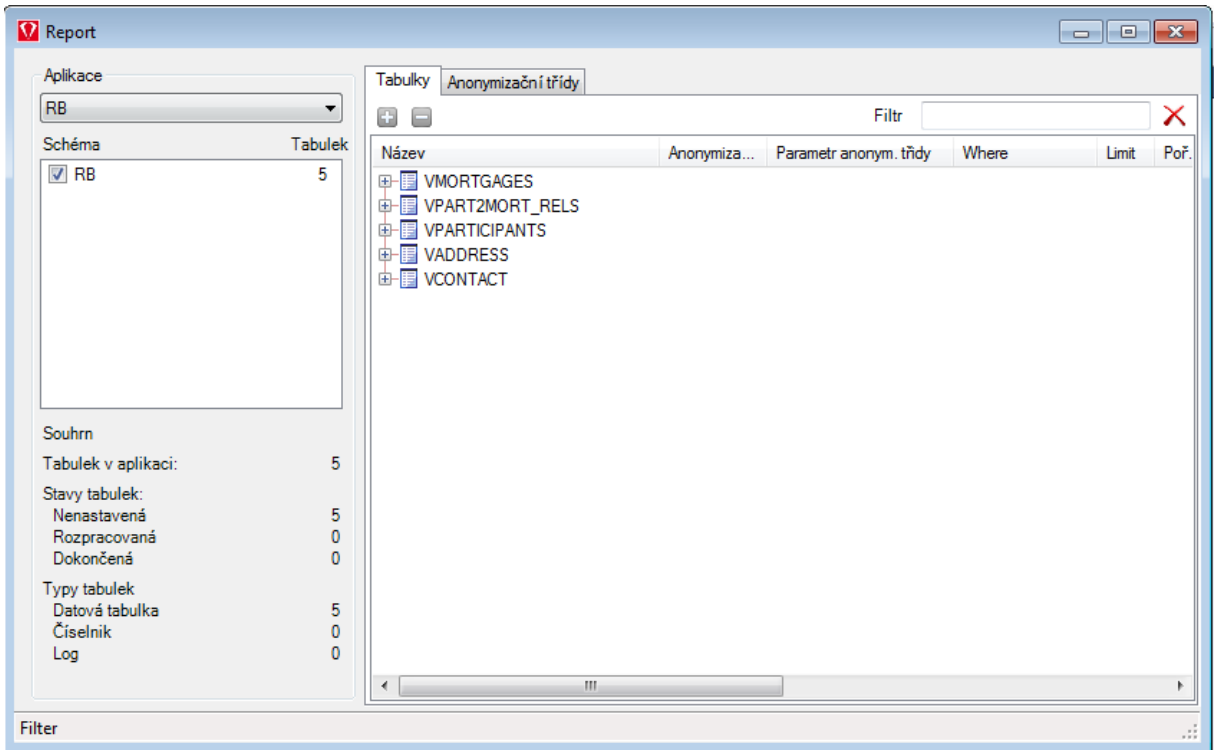
Obr. 29: Nastavení konfigurace.

Případně pomocí ikonky Report z panelu Addinu.



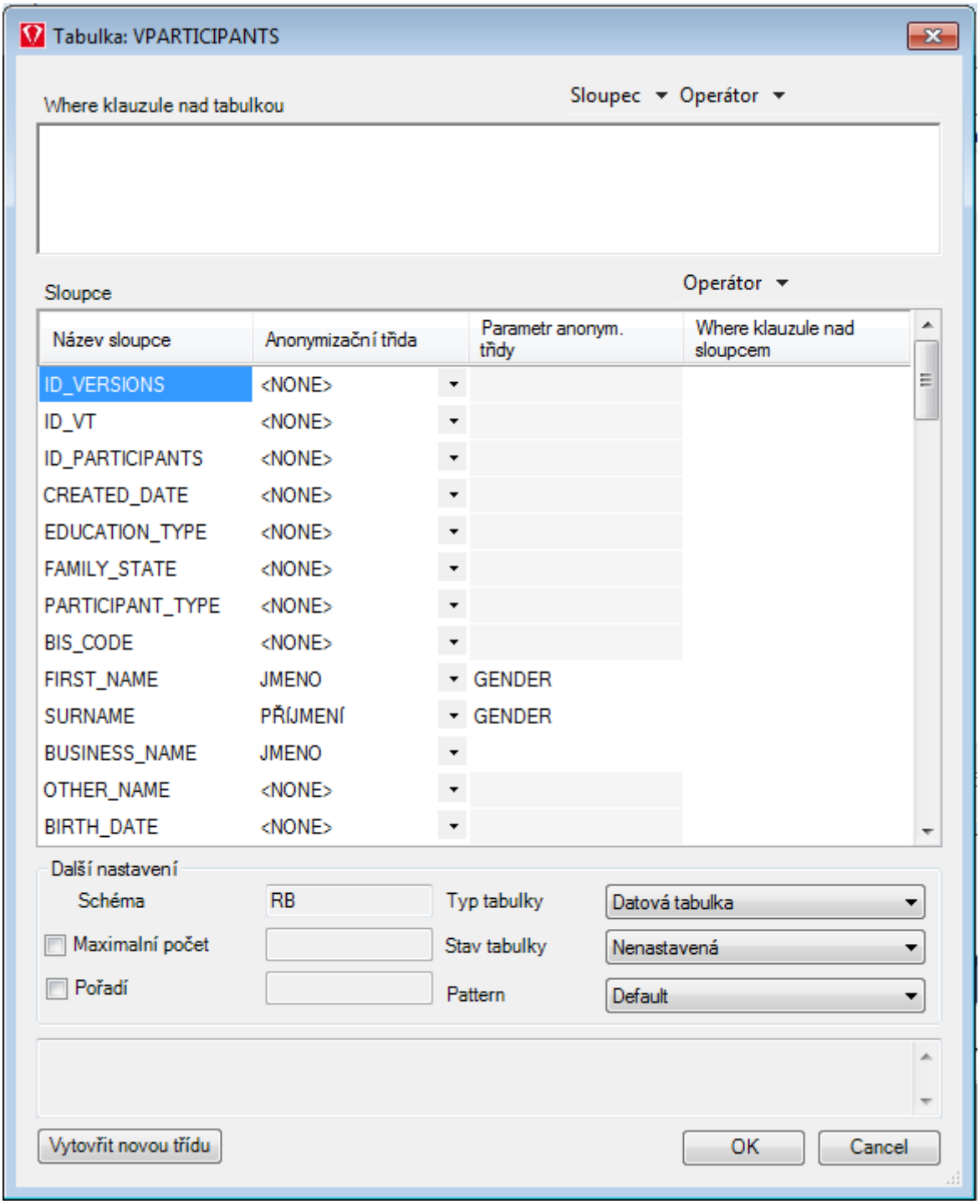
Obr 30:Winch add-in report.

Zobrazí se přehled tabulek a jejich nastavení.



Obr. 31:Okno s přehledem tabulek a jejich nastavení.

Dvojklikem na název tabulky se objeví okno pro zadání a editaci parametrů anonymizace (a řezu) pro zvolenou tabulku.



Obr. 32: Okno Add-In pro nastavení parametrů anonymizace pro zvolenou tabulku.

Editační okno se skládá ze tří částí:

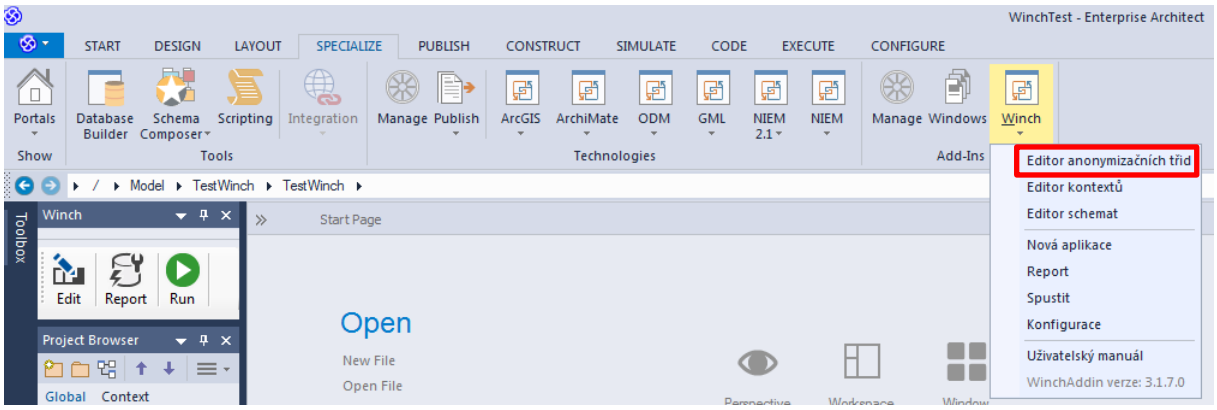
1. V první části okna (***Where klauzule nad tabulkou***) můžete přímo nastavit SQL klauzuli WHERE do textového pole, která slouží k řezům (výběru) dat nad zvolenou tabulkou. Klauzule může být i složitějšího charakteru. V tomto textovém poli je k dispozici podpora intelli-sense, která se zobrazí buď po zadání celého názvu tabulky a znaku tečky. Toto okno tedy slouží k omezení výběru záznamů, které budou výstupem provedené anonymizace.
2. V druhé části okna (***Sloupce***) můžete nastavit pro každý sloupec anonymizační třídu ze seznamu anonymizačních tříd
   - Sloupec „***Název sloupce***" je převzat z načtené struktury dané tabulky.
   - Sloupec „***Anonymizační třída***" slouží k nastavení (výběru) anonymizační třídy a parametru pro daný sloupec, pokud je pro přiřazenou funkci podporován.
   - Sloupec „***Parametr anonymizační třídy***" slouží jako doplňující parametr vybrané anonymizační třídy resp. funkce, pokud funkce pracuje s parametrem. Funkce, které tento parametr používají, většinou očekávají název sloupce, ve kterém je rozlišení na základě kterého se rozlišuje způsob anonymizace, nebo hodnotu, podle které se anonymizační funkce řídí. Například pro anonymizaci jména může vstoupit jako druhý parametr název sloupečku s pohlavím, který určí, zda bude vytvořeno ženské nebo mužské jméno.
   - Do sloupce „***Where klauzule nad sloupcem***" můžete vložit klauzuli WHERE, která slouží jako doplnění resp. alternativa k „***Where klauzule nad tabulkou***". Umožňuje snadněji zadat řez podle hodnoty v nějakém sloupci. Mezi jednotlivými podmínkami ve „***Where klauzule nad sloupcem***" a podmínkou zadanou v „***Where klauzule nad tabulkou***" se používá logický operátor AND.
3. V třetí části (***Další nastavení***) můžete nastavit
   - ***Maximální počet*** přenášených ***řádků***. (Hodnota od nuly výše nebo proměnná.)
   - ***Pořadí*** zpracovávání tabulek v databázi, pokud aplikace obsahuje více tabulek a anonymizace má proběhnout ve striktně zvoleném pořadí.
   - ***Typ tabulky*** - slouží k informačnímu rozlišení různých typů tabulek.
   - ***Stav tabulky*** – slouží k informačnímu rozlišení stavu zpracování parametrů pro danou tabulku. Stav tabulky si uživatel nastavuje sám pro vlastní

orientaci (stav může být Nenastavená, Rozpracována, nebo Dokončena). Výchozí hodnota je nastavena na Nenastavená

- ***Pattern*** – umožňuje ovlivnit, vlastní proces anonymizace
    - *Zda vytvořit novou tabulku pro anonymizovaná data (Create as new)*
    - *Změnit zdrojová data (Update Source Data)*
    - *Smazat zdrojová data (Delete SourceData)*

5.7 Využití hodnoty ve sloupci pro parametrizaci funkce.

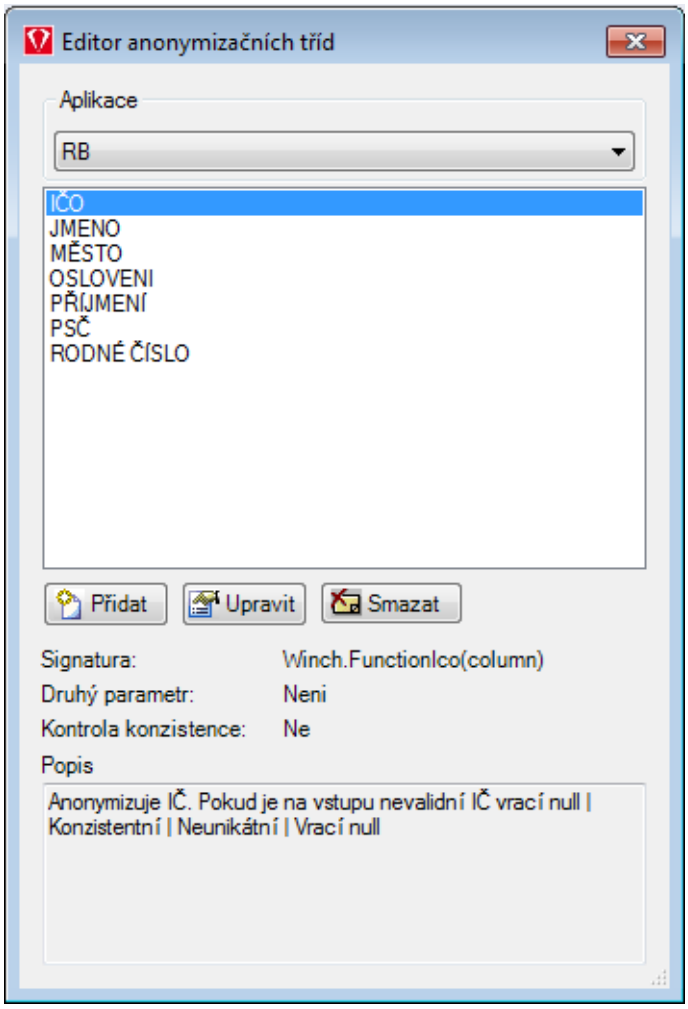
Pokud některý ze sloupců obsahuje údaje ve formátu, který může být použit jako parametr pro některou z anonymizačních funkcí, lze do parametru funkce zadat jméno tohoto sloupce. Předpokladem je, že tato vazba je nastavena prostřednictvím editoru anonymizačních tříd, který je dostupný v menu: SPECIALIZE-Winch-Editor anonymizačních tříd.



Obr. 33: Editor anonymizačních tříd.

Zobrazí se nové okno. V něm nalezneme všechny existující anonymizační třídy.
V případě, že nám žádná z uvedených tříd nevyhovuje, můžeme si nastavit vlastní třídu pomocí tlačítka ***Přidat***.

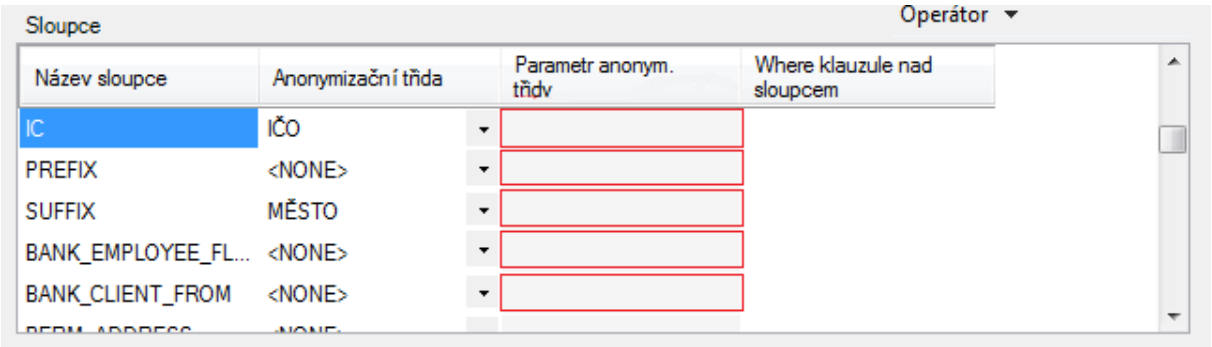
Tlačítkem ***Odstranit*** můžeme smazat již vytvořenou označenou anonymizační třídu.



Obr. 34: Editor anonymizačních tříd.

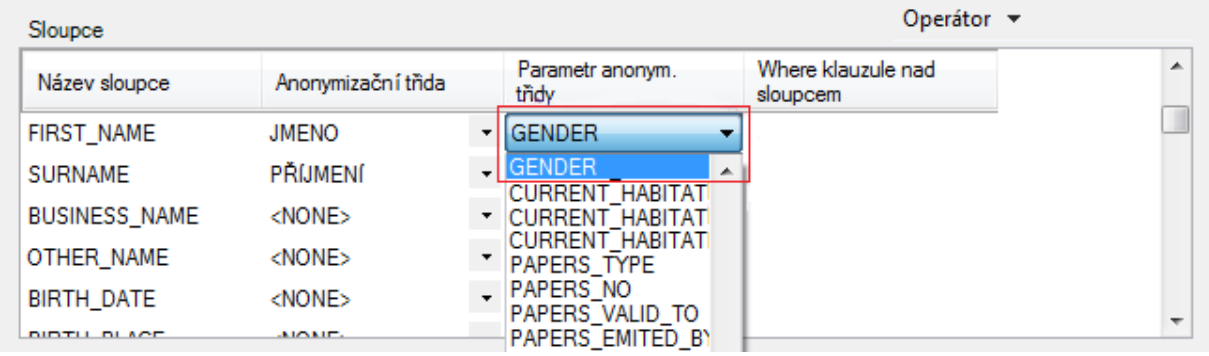
V případě přidávání nové třídy ji nejprve pojmenujeme, určíme anonymizační funkci, seznam anonymizačních funkcí zobrazíme tlačítkem ***Vybrat funkci ze seznamu***. Typ druhého parametru je možné vybrat z těchto možností:

- ***Žádný*** – při přiřazení anonymizační třídy ke sloupečku tabulky nebude možné zadat žádnou hodnotu druhého parametru



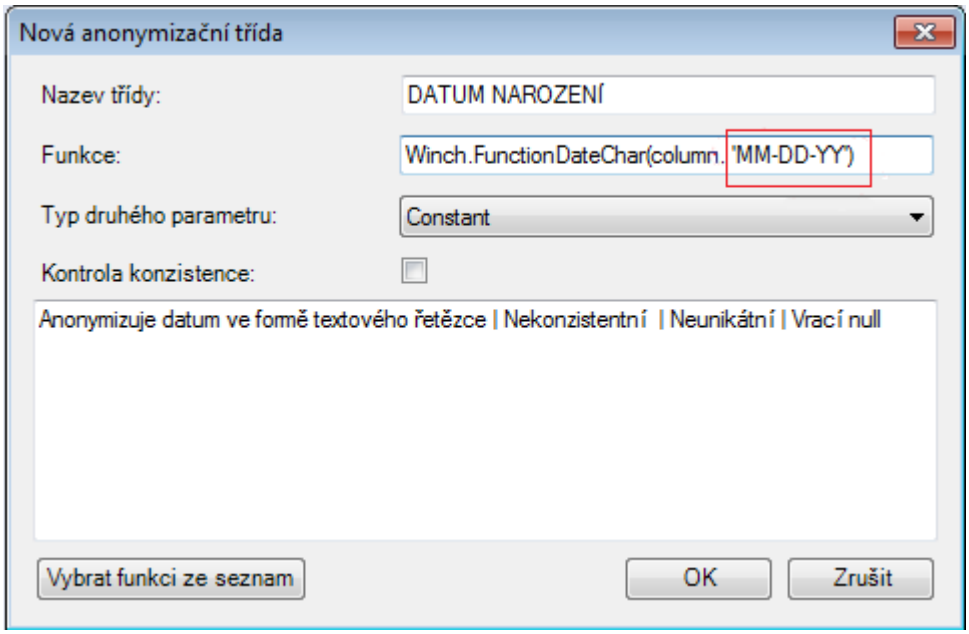
Obr 35:Anonymizační třída sloupce.

- ***Sloupeček*** - při přiřazení anonymizační třídy ke sloupečku tabulky bude možné jako druhý parametr vybrat ze seznamu jeden ze sloupečků této tabulky. Například výběr sloupečku obsahující pohlaví, dle kterého se má generovat jméno odpovídající danému pohlaví.



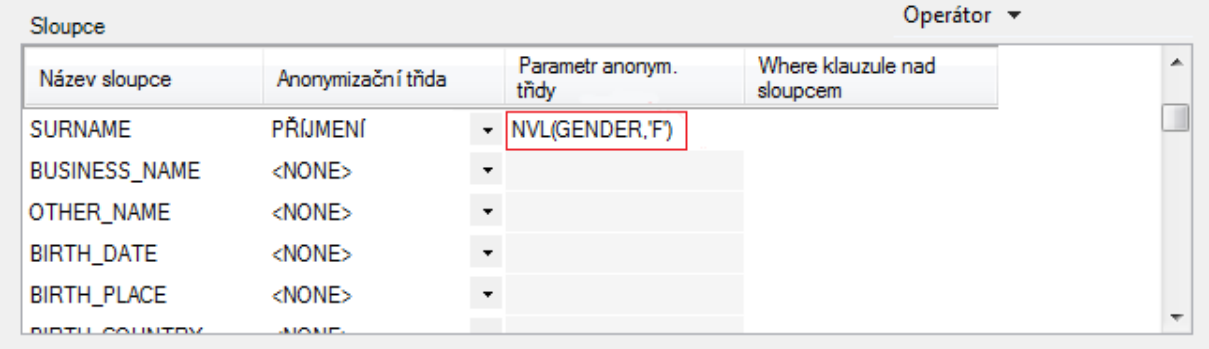
Obr. 36: Anonymizační třída slupec s parametrem.

- ***Konstanta*** – při přiřazení anonymizační třídy ke sloupečku tabulky nebude možné zadat žádnou hodnotu druhého parametru. Hodnota druhého parametru může být zadána jako konstanta přímo v definici anonymizační funkce. Například požadovaný formát data pro datumovou funkci. V tomto případě je konstanta shodná pro použití anonymizační třídy ve všech tabulkách.



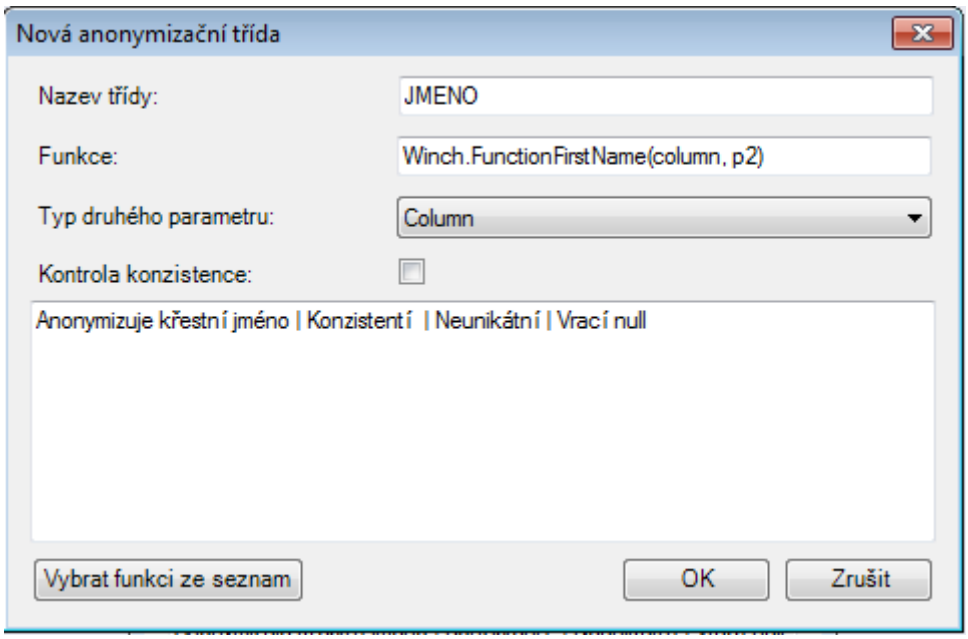
Obr. 37: Nová anonymizační třída.

- ***Proměnná*** - při přiřazení anonymizační třídy ke sloupečku tabulky bude možné zadat jako druhý parametr celý sql výraz. Například v případě, že není pohlaví uvedeno, použij ženské jméno.



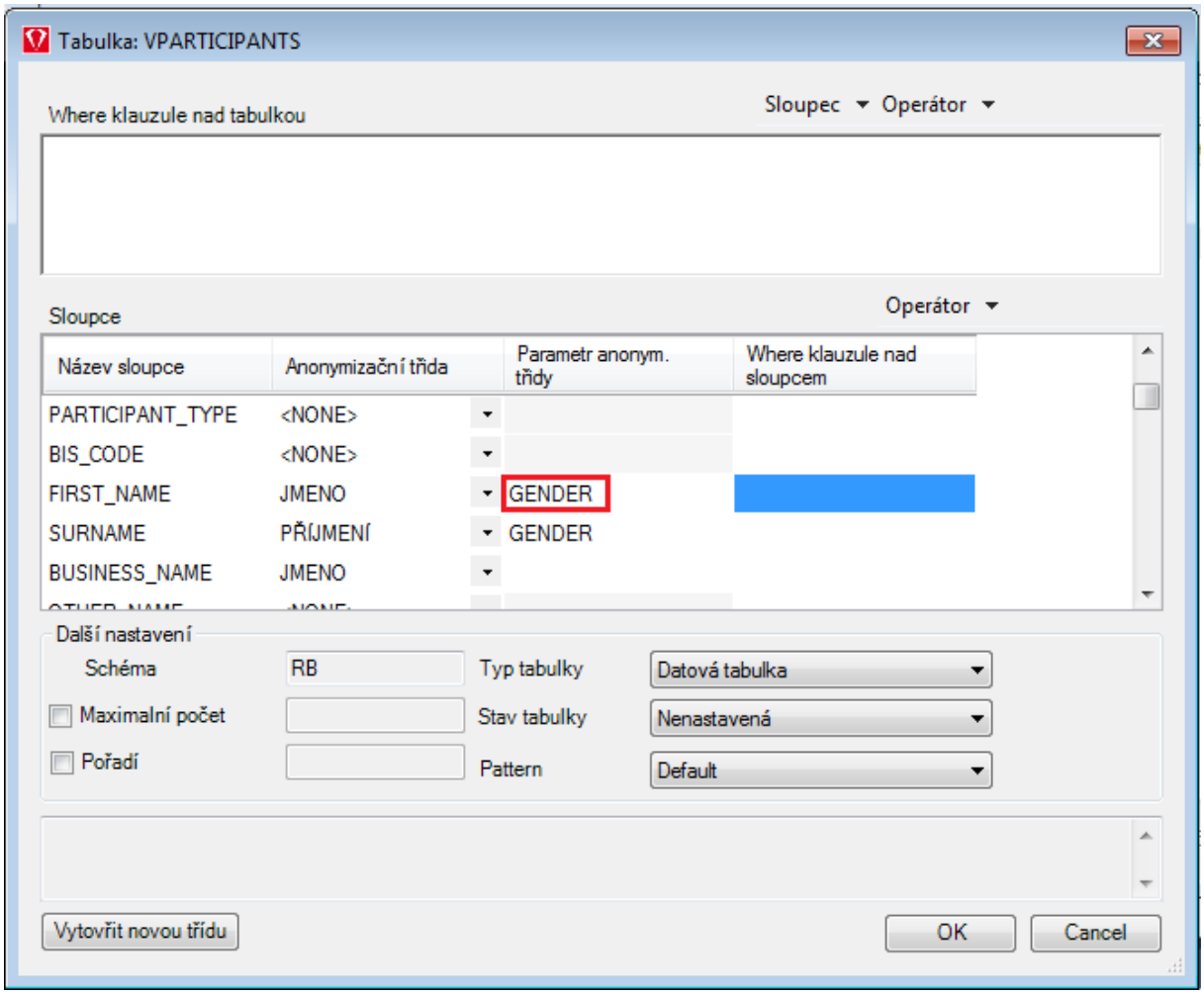
Obr. 38: Nastavení parametru.

V případě nastavení typ druhého parametru jako ***Column.***



Obr. 39:Nastavení parametru jméno.

V nastavení parametrů anonymizace tabulky potom můžeme pro třídu jméno vybrat v dalším sloupci jiný sloupec obsahující pohlaví.

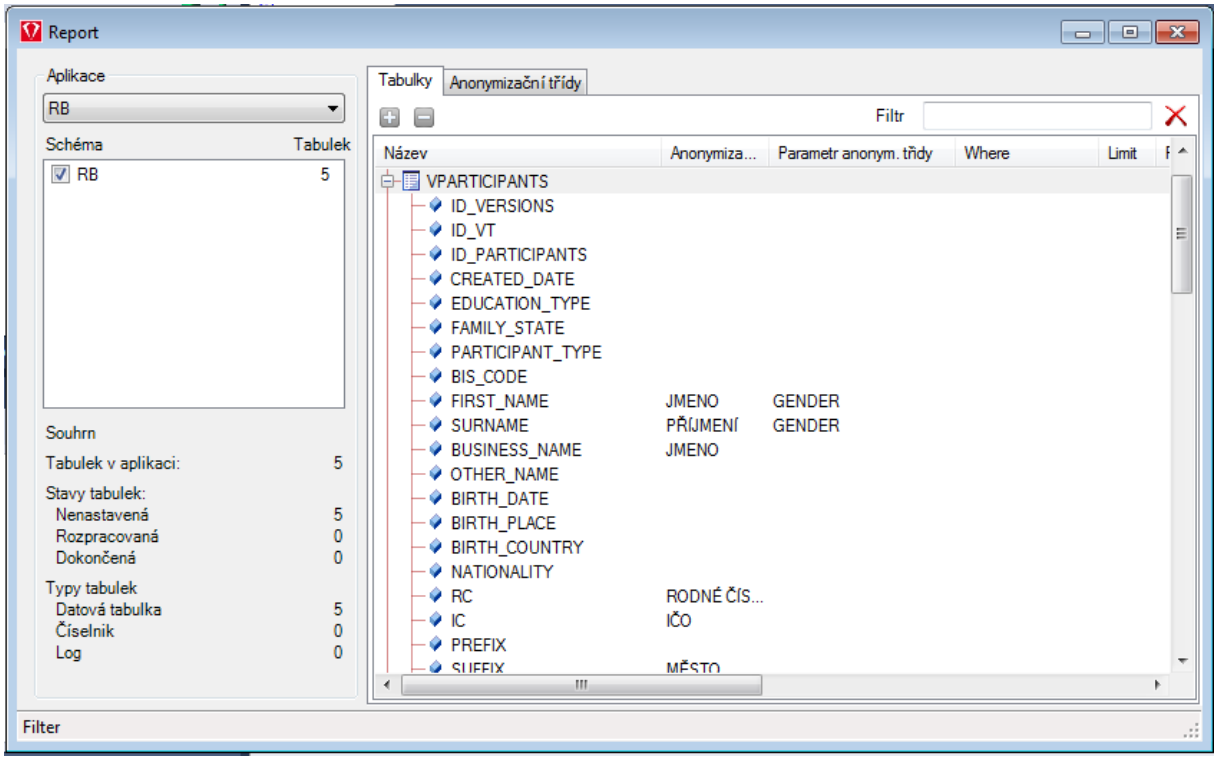


Obr. 40: Vparticipants.

Při takto provedené konfiguraci pak jsou jména anonymizována pomocí slovníku, odpovídajícímu danému pohlaví klienta, uloženému ve zvoleném sloupci.

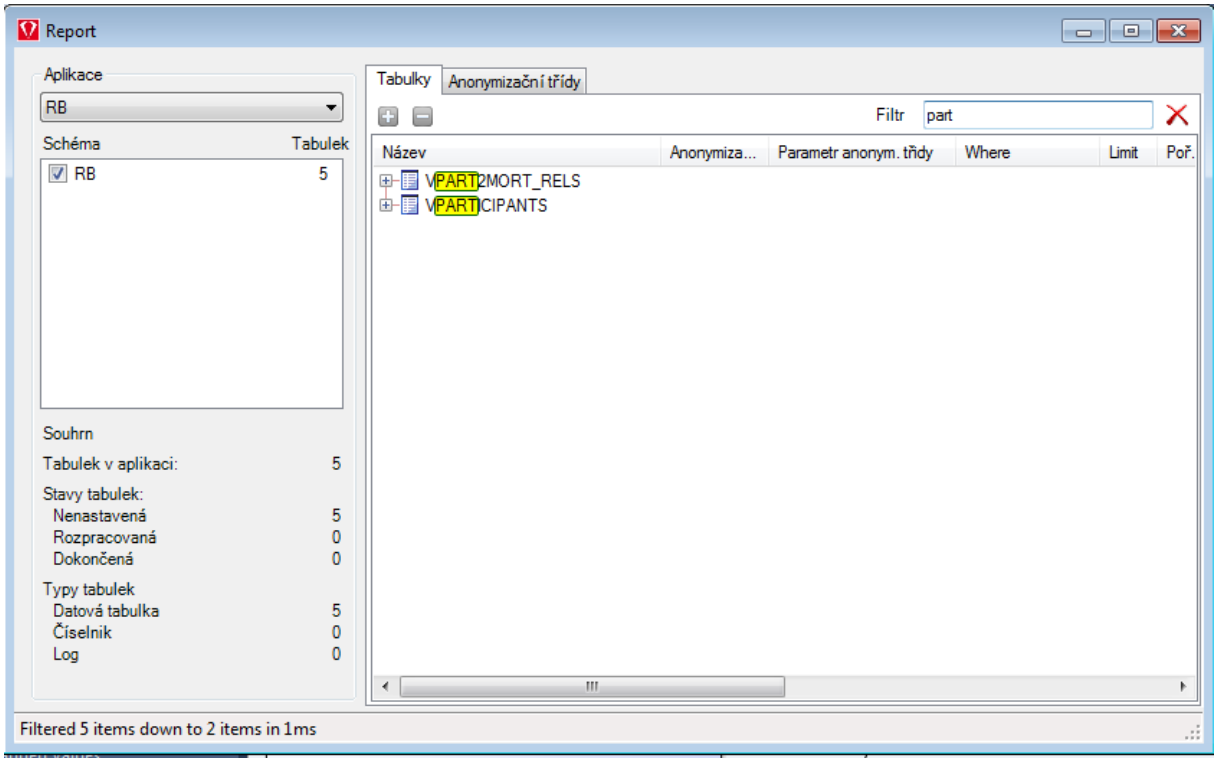
5.8 Zobrazení obrazovky pro přehled stavu a nastavení anonymizačních tříd a řezů

Okno report zobrazuje aktuální přehled nastavených anonymizačních tříd. Levá část okna obsahuje prvky určené k výběru aplikace a jednoduchému výběru zobrazených schémat společně se souhrnem vybraných položek. Pravá část okna obsahuje dvě záložky. První záložka zobrazuje všechny tabulky a jejich sloupce korespondující s vybranými schématy aplikace. U každé tabulky a sloupce jsou zobrazeny jejich nastavené hodnoty.



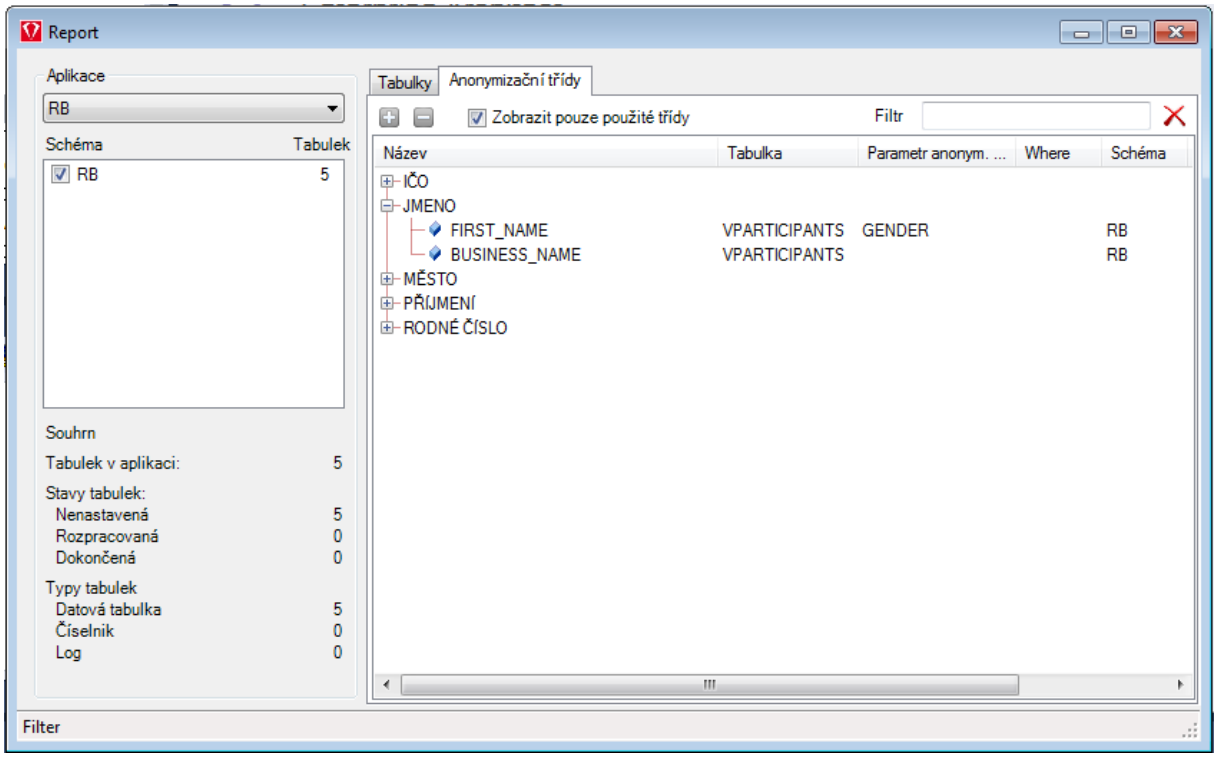
Obr. 41:Zobrazení parametrů anonymizace.

Pomocí filtru v pravém horním rohu je možné velmi snadno vyhledat požadovanou tabulku. Zadáním části názvu se v okně zobrazí pouze tabulky, které splňují zadané vyhledávací kritérium.



Obr. 42:Zobrazení parametrů anonymizace.

Můžeme také přejít na záložku ***Anonymizační třídy***. Tato druhá záložka obsahuje seznam dostupných anonymizačních tříd a k nim přirazené sloupce a tabulky. Slouží k rychlému dohledání např. kde všude je v databázi dostupné rodné číslo, apod.

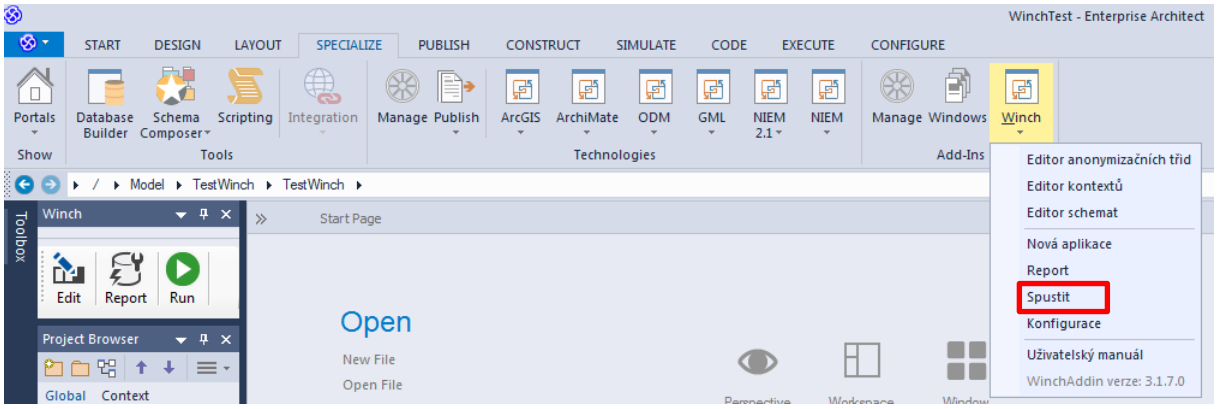


Obr. 43: Zobrazení přehledu anonymizace formou reportu

Zaškrtávací políčko „***Zobrazit pouze použité třídy***" umožňuje zobrazit pouze anonymizační třídy, které jsou nastaveny u některého ze sloupců. Dvojklikem nad vybranou položkou lze otevřít editační okno vybrané tabulky či sloupce. Obě záložky lze jednoduše filtrovat zadáním výrazu do textového pole v pravém horním rohu okna. Ve stavovém řádku okna se zobrazuje počet vyfiltrovaných položek. Zobrazené údaje lze libovolně řadit kliknutím na hlavičku sloupce. Po kliknutí pravým tlačítkem na hlavičce sloupce se zobrazí pokročilejší nastavení zobrazení. Lze využít pokročilého filtru a vybrat sloupce zobrazované v tabulce. Tlačítka „**+**" a „**–** " v horní části okna slouží k rozbalení a sbalení všech položek v tabulce.

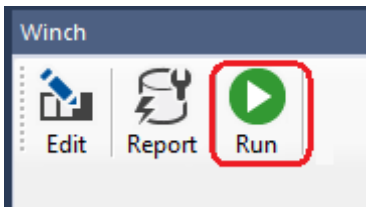
5.9 Export řídícího souboru pro DB server

Po dokončení nastavení anonymizace lze přejít k dalšímu kroku, kterým je provedení vlastní anonymizace. Lze ho spustit z menu SPECIALIZE-Winch-Spustit



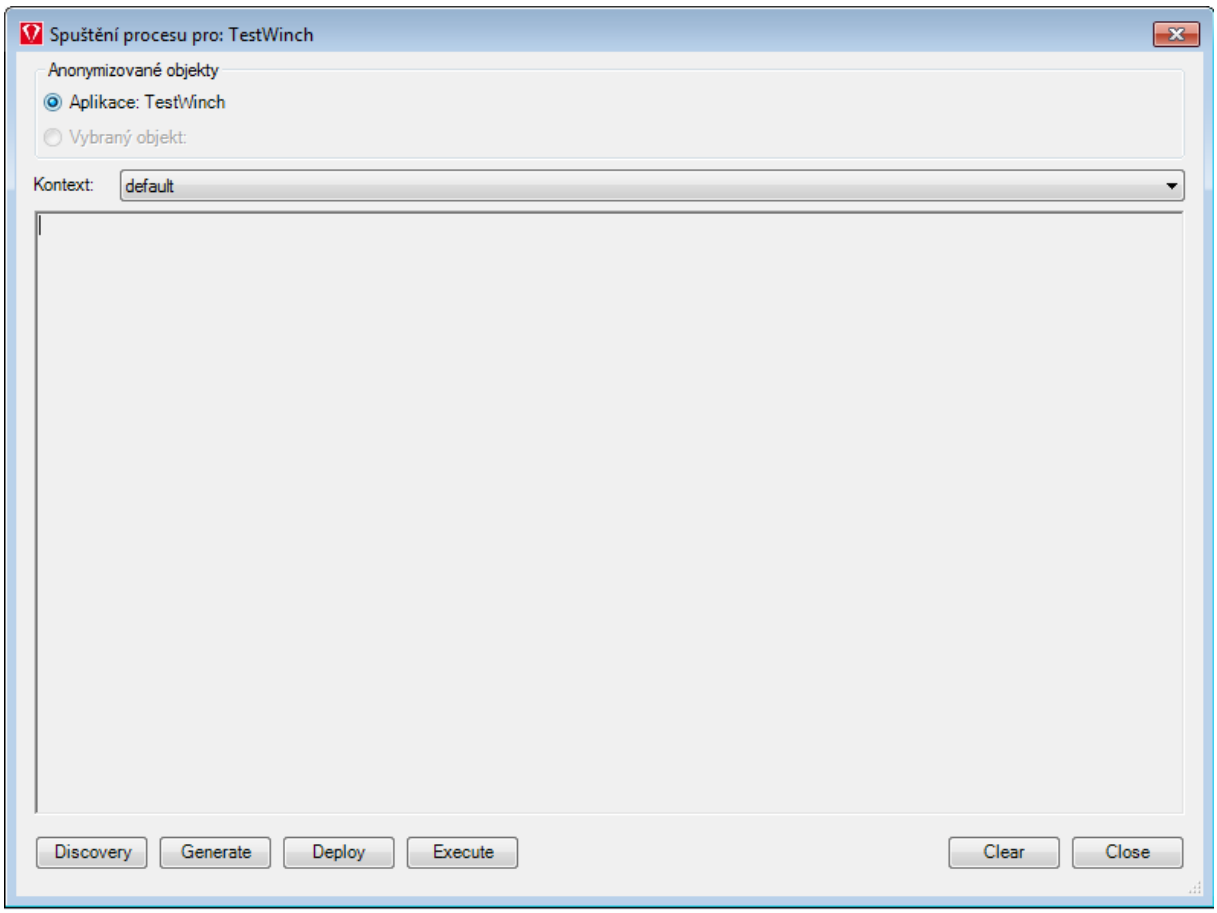
Obr. 44: Spouštění.

Nebo pomocí ikonky Run:



Obr. 45 Spouštění widget.

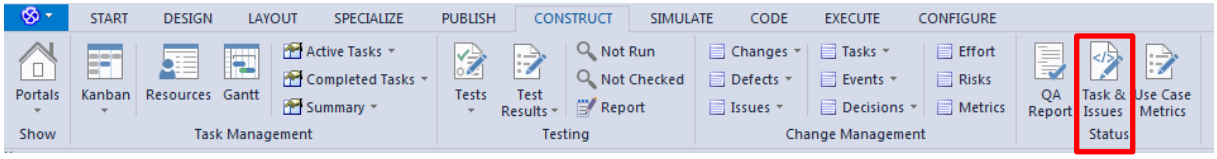
Zobrazí se okno umožňující provést různé operace



Discovery – projde strukturu zdrojové databáze a její data a pokusí se v nich vyhledat osobní/citlivé údaje. Tomuto kroku je věnována následující kapitola.
Generate – vygeneruje anonymizační skripty a uloží je na disk
Deploy – nasadí anonymizační skripty do databáze
Execute – spustí vlastní proces anonymizace – tomuto kroku musí předcházet Deploy, který připraví prázdné struktury pro uložení anonymizovaných dat

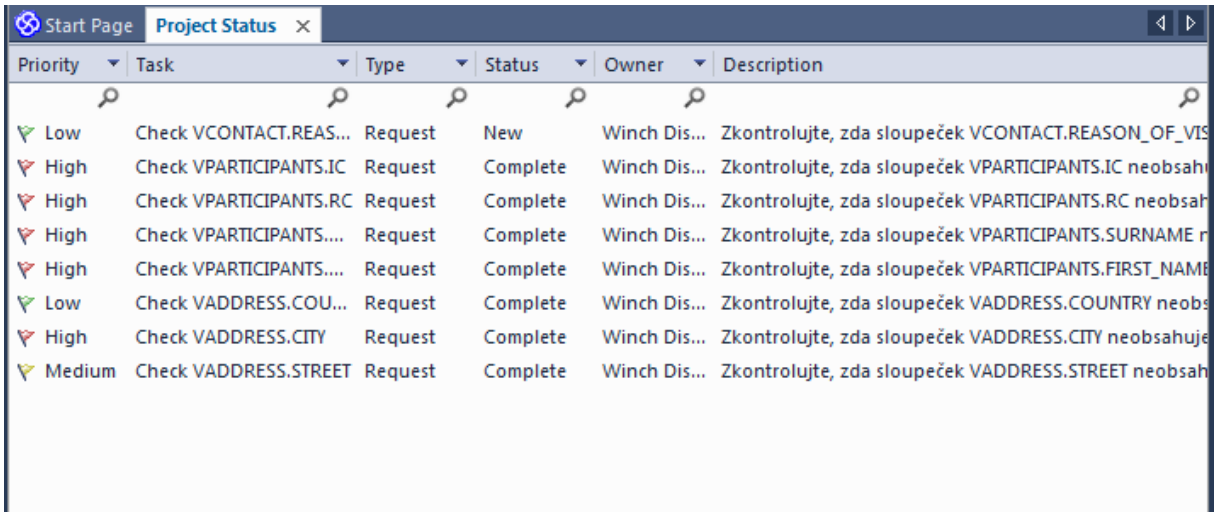
5.10 GEM Winch Discovery – vyhledání osobních/citlivých údajů

Funkce umožňuje uživateli pomoci s vyhledáváním osobních/citlivých dat v databázi. Pokud u některého sloupečku dospěje k názoru, že by mohl osobní/citlivé údaje obsahovat, pak doporučí uživateli jeho kontrolu. Výsledek prohledávání zadává jako jednotlivé úkoly (tasky) přímo do nástroje Enterprise architect. Tento seznam úkolů je možné zobrazit pomocí menu: CONSTRUCT-Task&Issues



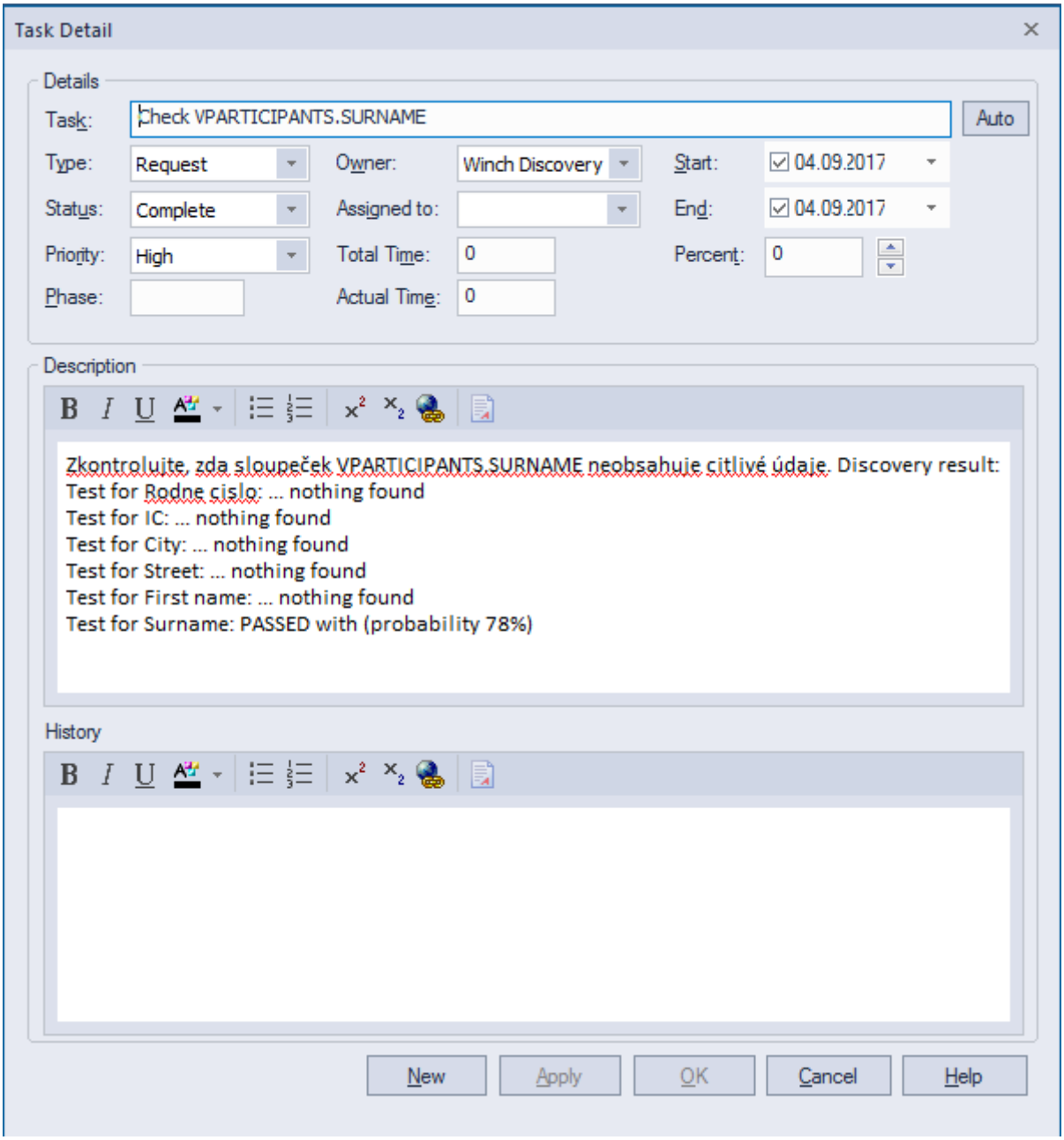
Obr. 46: Výsledek discovery procesu.

Zobrazí se přehled vygenerovaných úkolů



Obr. 47:Výsledek discovery fcí.

V detailu každého úkolu jsou uvedeny detailní výsledky testu



Obr. 48: Task detail.

6 Provedení anonymizace

Proces provedení anonymizace lze rozdělit do čtyř částí:

1. Instalace objektů pro Winch DB Actor (pouze jednou)
2. Nasazení vygenerovaných skriptů do databáze
3. Spuštění SQL skriptů

Opakované kroky 2 a 3 lze pohodlně spouštět přímo z grafického rozhraní Winch-Addin nebo z příkazové řádky.

6.1 Instalace objektů pro Winch DB Actor

Prvním krokem a předpokladem pro provedení anonymizace je instalace objektů potřebných pro provádění anonymizace. K této instalaci je dodáván sql skript specifický pro konkrétní databázi:

- oracleDbInitScript.sql pro Oracle databáze
- mssqlDbInitScript.sql pro Microsoft SQL Server
- postgresqlDbInitScript.sql pro PostgreSQL Server
- db2DbInitScript.sql pro DB2 databázi

V případě, že jsou implementovány specifické slovníky pro konkrétní zemi, je možné využít instalační skript, který tyto slovníky obsahuje. Skript používá jako prefix dvoupísmenný kód dané země např. sk_oracleDbInitScript.sql pro použití v rámci Slovenské republiky.

6.1.1 Oracle databáze

Nutná nastavení pro schéma winch. Uživatel/schéma, pod kterým se budou nasazovat pomocné db objekty pro nástroj GEM Winch:

create user winch identified by “winch“;

grant connect to “winch“;

GRANT UNLIMITED TABLESPACE TO winch;

grant create table to winch;

grant create view to winch;

grant create procedure to winch;

grant create sequence to winch;

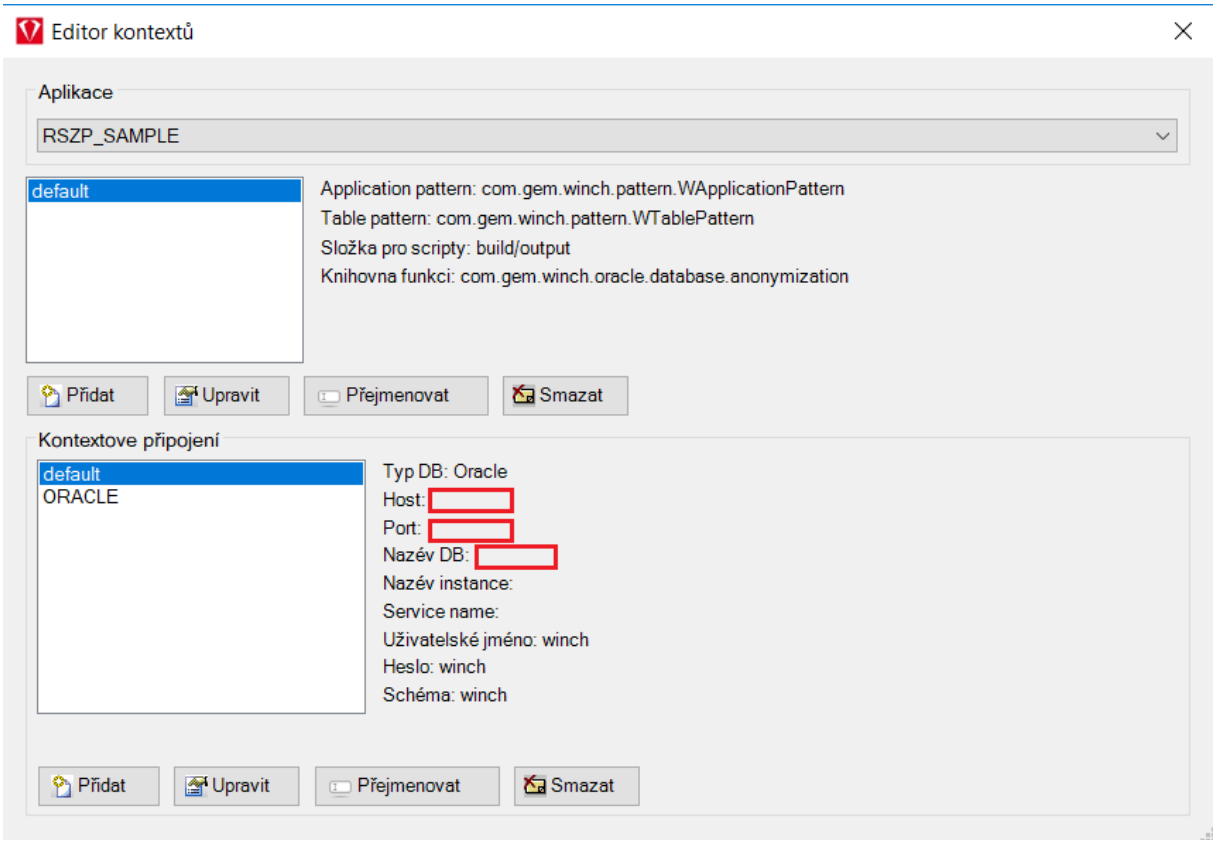
grant create trigger to winch;

Pro případné ladění je vhodné mít i práva:
GRANT DEBUG CONNECT SESSION TO WINCH ;
GRANT DEBUG ANY PROCEDURE TO WINCH ;

Pro zpřístupnění tabulek, které mají být anonymizovány lze spustit níže uvedený skript pod uživatelem, který vlastní tabulky určené k anonymizaci:

```
BEGIN
FOR x IN (SELECT * FROM user_tables )
LOOP
EXECUTE IMMEDIATE 'GRANT SELECT ON ' || x.table_name || ' TO WINCH';
END LOOP;
END;
```

Default kontextové připojení se nastaví na výše vytvořeného uživatele s příslušnými granty. Nastavíme host, port a název DB, ke které přistupujeme. Service name volíme stejné jako pro název DB.

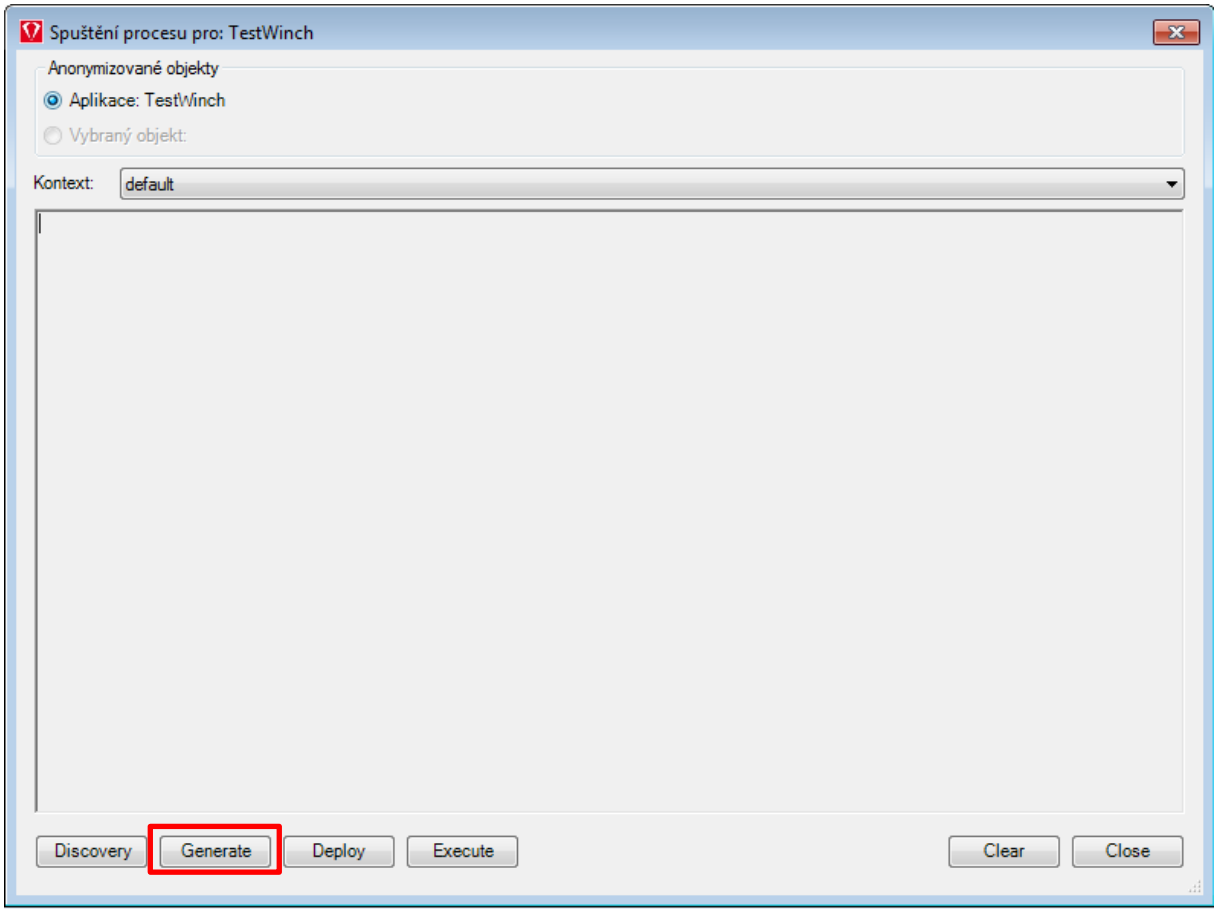


|Obr. 49:Add-in Editor kontextů

ORACLE kontextové připojení obsahuje nastavení na zdrojová data, nad kterými budeme pouštět proces anonymizace.

6.2 Nasazení vygenerovaných skriptů do databáze

Nasazení vygenerovaných skriptů do databáze se provádí stiskem tlačítka Deploy ve WinchAddinu.



Obr. 50:Deploy

Případně je možné provést nasazení z příkazové řádky pomocí příkazu:

`disl-winch-[mssql/oracle/postgresql/db2].bat -d`

6.3 Spuštění SQL skriptů

Spuštění SQL skriptů v databázi se provádí pomocí tlačítka Execute.

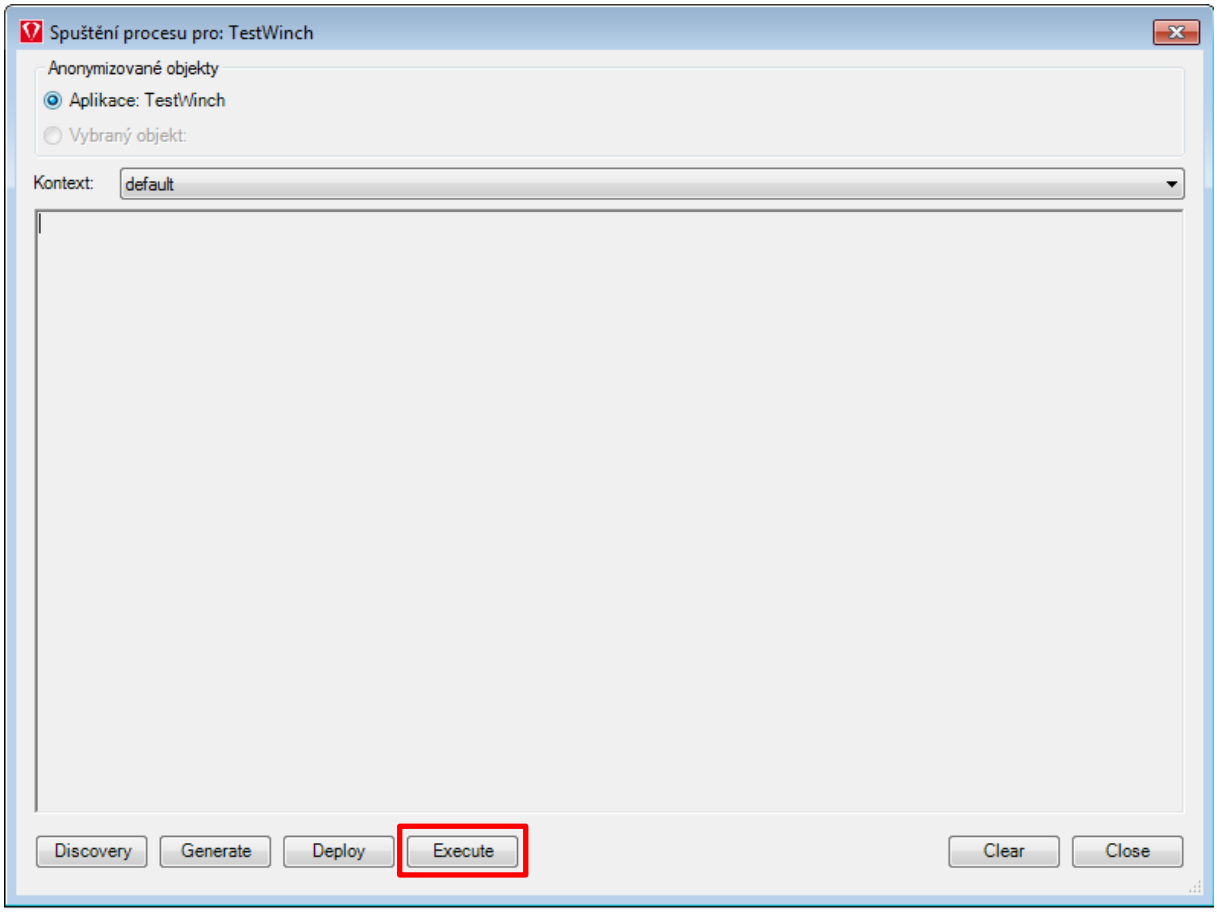


Table 51:Execute

Případně je možné provést spuštění z příkazové řádky pomocí příkazu:

`disl-winch-[mssql/oracle/postgresql/db2].bat -e`

7 Řezy dat

Možnost provádět řezy dat je realizována prostřednictvím „where“ klauzule, kterou je možno zadat pro celou tabulku nebo pro sloupec v okně pro editaci konfigurace.

7.1 Konfigurace řezů

Na úrovni tabulky je možné definovat následující kritéria pro uplatnění řezů:

- table.where – slouží pro omezení počtu řádku v tabulce. Muže obsahovat SQL příkaz v podobě where podmínky (např. datod>(sysdate-120))
- table.limit – slouží pro absolutní omezení počtu řádků v tabulce. Hodnota udává maximální počet řádků v tabulce. (rownum<anonym.limit).

  Pozn. Limit je určen bez ohledu na řazení dat.

- table.order – slouží k definici, v jakém pořadí mají být tabulky zpracovávány - číselná hodnota mezi 1 až 99999

Na sloupci mohou být definovány tyto kritéria pro uplatnění řezů:

- column.where – slouží pro omezení počtu řádku v tabulce. Muže obsahovat SQL příkaz v podobě where podmínky (např. datod>(sysdate-120))

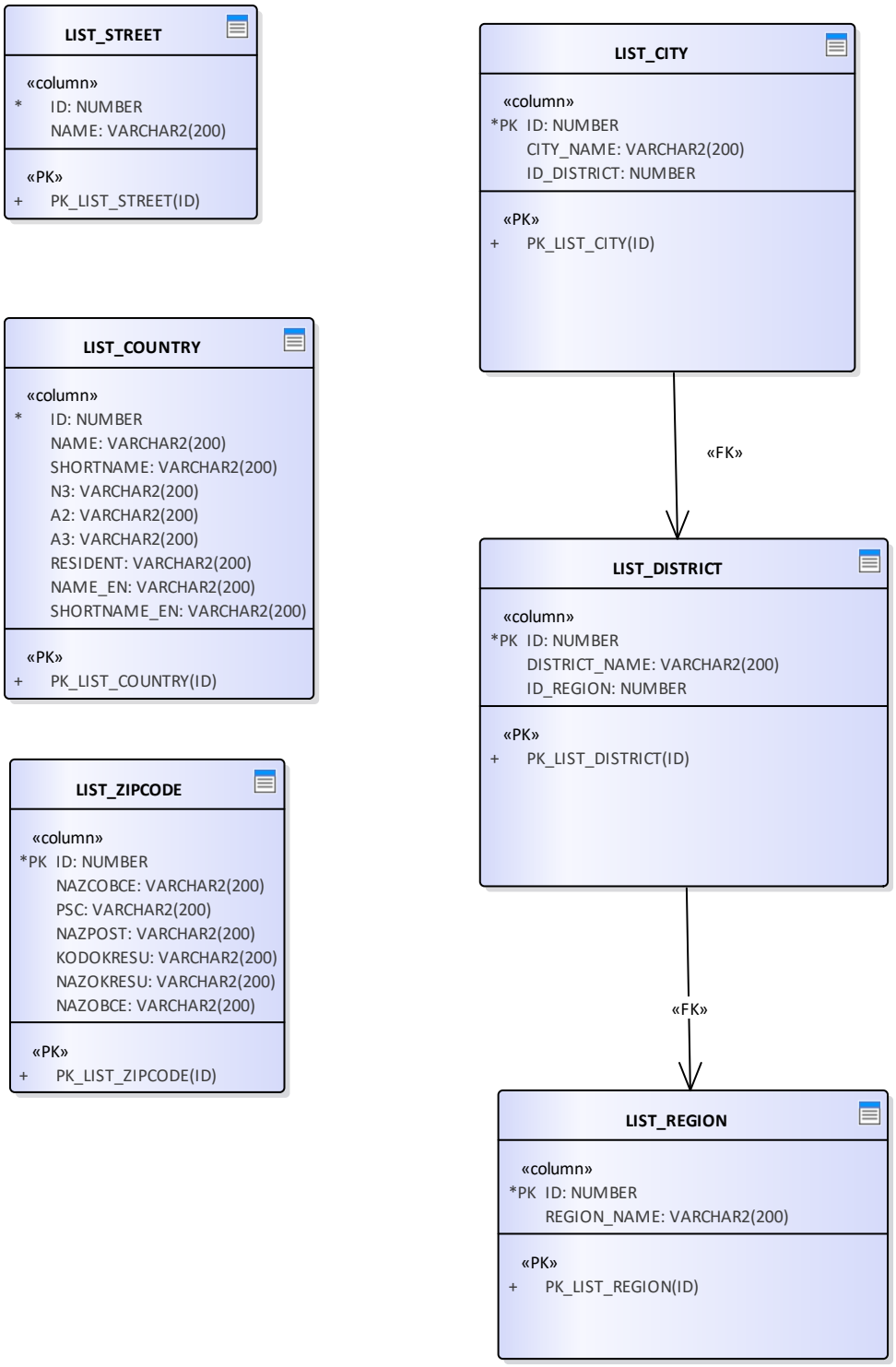
8 Popis klíčových datových objektů (tabulek) vytvořených Winch DB Actor

8.1 Slovníky

Obsahuje slovníkové tabulky pro jednotlivé anonymizační funkce. Slovníky obsahují kompletní výčet možných hodnot pro danou anonymizační třídu, případně výběr nejčastěji se vyskytujících hodnot. Používají se jak pro anonymizaci údajů, tak pro discovery funkce.

8.1.1 **Adresy**

Databázové tabulky související s adresami.



Obrázek 1: Adresy

8.1.1.1 LIST_CITY

Seznam českých obcí - výběr.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
CITY_NAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
ID_DISTRICT	**NUMBER** Length: Precision: 0 Scale: 0	PK: False FK: False NOT NULL: False UNIQUE: False	

Reference

Název	Cílová tabulka	Podmínka spojení
	LIST_DISTRICT	

8.1.1.2 LIST_COUNTRY

Seznam zemí

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: False FK: False NOT NULL: True UNIQUE: False	

Název sloupce	Datový typ	Vlastnosti	Popis
NAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
SHORTNAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
N3	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
A2	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
A3	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
RESIDENT	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

Název sloupce	Datový typ	Vlastnosti	Popis
NAME_EN	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
SHORTNAME_EN	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.1.1.3 LIST_DISTRICT

Seznam všech okresů v ČR

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
DISTRICT_NAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
ID_REGION	**NUMBER** Length: Precision: 0 Scale: 0	PK: False FK: False NOT NULL: False UNIQUE: False	

Reference

Název	Cílová tabulka	Podmínka spojení
	LIST_REGION	

8.1.1.4 LIST_REGION

Seznam všech krajů v ČR.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
REGION_NAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.1.1.5 LIST_STREET

Seznam názvů ulic

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: False FK: False NOT NULL: True UNIQUE: False	

Název sloupce	Datový typ	Vlastnosti	Popis
NAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

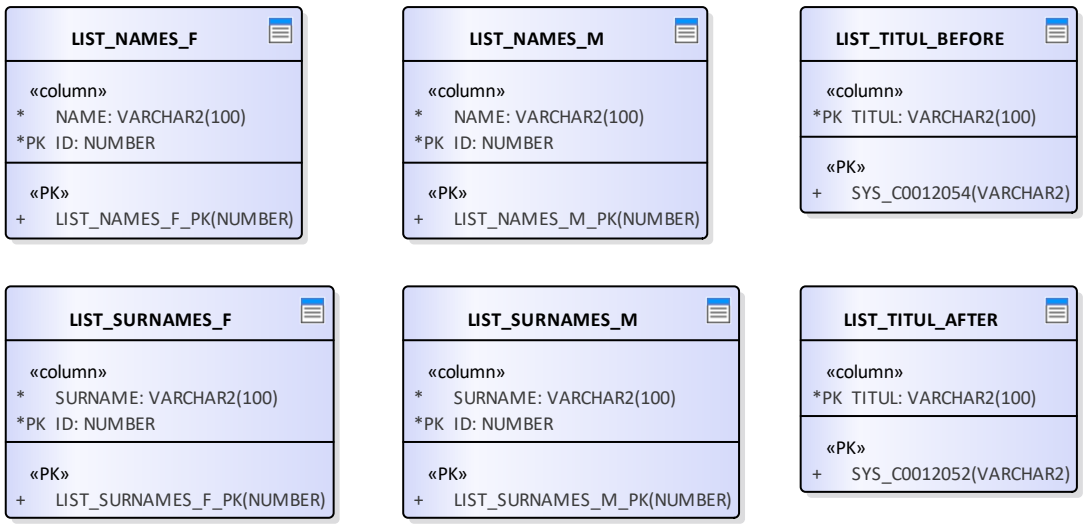
8.1.1.6 LIST_ZIPCODE

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
NAZCOBCE	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
PSC	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
NAZPOST	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

Název sloupce	Datový typ	Vlastnosti	Popis
KODOKRESU	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
NAZOKRESU	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
NAZOBCE	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.1.2 **Jména**



Obrázek 2: Jména

8.1.2.1 LIST_NAMES_F

Seznam nejčastějších ženských křestních jmen

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
NAME	**VARCHAR2** Length: 100 Precision: Scale:	PK: False FK: False NOT NULL: True UNIQUE: False	
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	

8.1.2.2 LIST_NAMES_M

Seznam nejčastějších mužských křestních jmen

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
NAME	**VARCHAR2** Length: 100 Precision: Scale:	PK: False FK: False NOT NULL: True UNIQUE: False	
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	

8.1.2.3 LIST_SURNAMES_F

Seznam nejčastějších ženských příjmení

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
SURNAME	**VARCHAR2** Length: 100 Precision: Scale:	PK: False FK: False NOT NULL: True UNIQUE: False	
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	

8.1.2.4 LIST_SURNAMES_M

Seznam nejčastějších mužských příjmení.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
SURNAME	**VARCHAR2** Length: 100 Precision: Scale:	PK: False FK: False NOT NULL: True UNIQUE: False	
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	

8.1.2.5 LIST_TITUL_AFTER

Seznam titulů používaných před jménem.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
TITUL	**VARCHAR2** Length: 100 Precision: Scale:	PK: True FK: False NOT NULL: True UNIQUE: False	

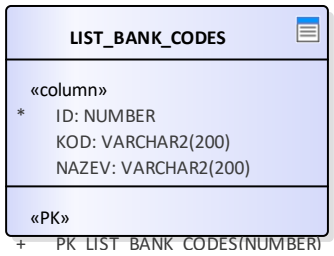
8.1.2.6 LIST_TITUL_BEFORE

Seznam titulů používaných za jménem.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
TITUL	**VARCHAR2** Length: 100 Precision: Scale:	PK: True FK: False NOT NULL: True UNIQUE: False	

8.1.3 Bankovní spojení



Obrázek 3: Bankovní spojení

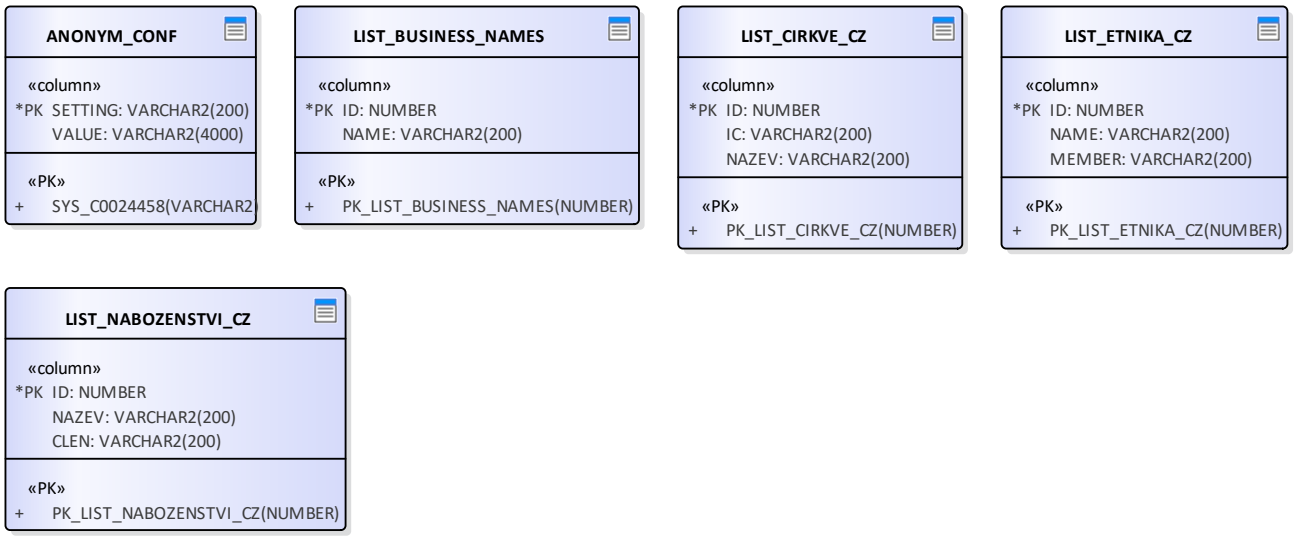
8.1.3.1 LIST_BANK_CODES

Seeznam kódů bank.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: False FK: False NOT NULL: True UNIQUE: False	
KOD	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
NAZEV	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.1.4 **Ostatní**



Obrázek 4: Ostatní

8.1.4.1 ANONYM_CONF

Tabulka obsahujici konfiguracni udaje pro anonymizacni funkce.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
SETTING	**VARCHAR2** Length: 200 Precision: Scale:	PK: True FK: False NOT NULL: True UNIQUE: False	
VALUE	**VARCHAR2** Length: 4000 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.1.4.2 LIST_BUSINESS_NAMES

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
NAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.1.4.3 LIST_CIRKVE_CZ

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
IC	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
NAZEV	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.1.4.4 LIST_ETNIKA_CZ

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
NAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

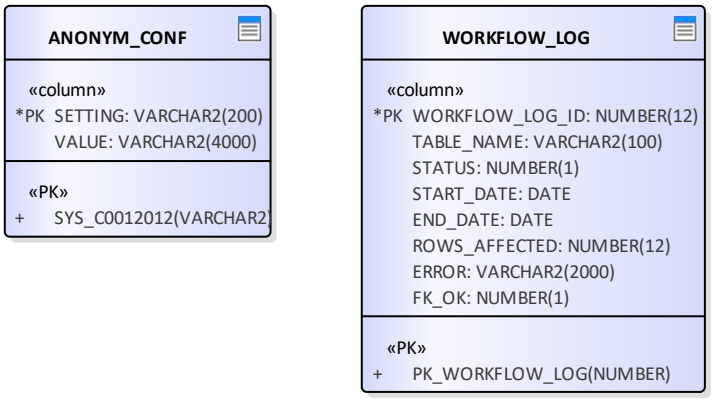
Název sloupce	Datový typ	Vlastnosti	Popis
MEMBER	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.1.4.5 LIST_NABOZENSTVI_CZ

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
ID	**NUMBER** Length: Precision: 0 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
NAZEV	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
CLEN	**VARCHAR2** Length: 200 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.2 Společné



Obrázek 5: Společné

8.2.1 **ANONYM_CONF**

Tabulka obsahující konfigurační údaje pro anonymizační funkce.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
SETTING	**VARCHAR2** Length: 200 Precision: Scale:	PK: True FK: False NOT NULL: True UNIQUE: False	
VALUE	**VARCHAR2** Length: 4000 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	

8.2.2 **DICTIONARY_SIZES**

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
TABLE_NAME	**VARCHAR2** Length: 200 Precision: Scale:	PK: True FK: False NOT NULL: True UNIQUE: False	
TABLE_SIZE	**NUMBER** Length: Precision: 10 Scale: 0	PK: False FK: False NOT NULL: True UNIQUE: False	

8.2.3 WORKFLOW_LOG

Tabulka obsahuje zaznamy o prubehu spustene anonymizace.

Sloupce

Název sloupce	Datový typ	Vlastnosti	Popis
WORKFLOW_LOG_ID	**NUMBER** Length: Precision: 12 Scale: 0	PK: True FK: False NOT NULL: True UNIQUE: False	
TABLE_NAME	**VARCHAR2** Length: 100 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
STATUS	**NUMBER** Length: Precision: 1 Scale: 0	PK: False FK: False NOT NULL: False UNIQUE: False	

Název sloupce	Datový typ	Vlastnosti	Popis
START_DATE	**DATE** Length: Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
END_DATE	**DATE** Length: Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
ROWS_AFFECTED	**NUMBER** Length: Precision: 12 Scale: 0	PK: False FK: False NOT NULL: False UNIQUE: False	
ERROR	**VARCHAR2** Length: 2000 Precision: Scale:	PK: False FK: False NOT NULL: False UNIQUE: False	
FK_OK	**NUMBER** Length: Precision: 1 Scale: 0	PK: False FK: False NOT NULL: False UNIQUE: False	

9 Slovník zkratek a pojmů

Zkratka	Význam
DB	Databáze
ODBC	Open Database Connectivity je standardizované softwarové rozhraní (API) pro přístup k databázovým systémům (DBMS).
WINCH	Nástroj anonymizace
WINCH ADD-IN	Součást nástroje pro anonymizaci, doplněk (Add-In) pro Enterprise Architect. Slouží pro nastavení parametrů (konfiguraci) anonymizace případně řezu dat na základě datového modelu.
WINCH DB ACTOR	Součást nástroje pro anonymizaci, databázová část. Slouží pro vlastní vykonání anonymizace případně řezu dat.
EA	Enterprise Architect

Příloha č. 1) Obecné anonymizační metody

Níže jsou uvedeny techniky a principy anonymizace dat použité v anonymizačních funkcích:

a) Náhodné generování znaků: nejjednodušší způsob anonymizace je vygenerování sekvence náhodných znaků, kterými je následně nahrazena původní hodnota. Výsledkem funkcí tohoto typu jsou řetězce, které na první pohled nedávají žádný smysl. Splňují sice účel dokonalé anonymizace, ale ztrácí základní vypovídající hodnotu reálných dat.
b) Použití předpřipravených slovníků: tento typ funkcí ke své činnosti potřebuje předem připravený slovník řetězců, ze kterých náhodně vybírá výslednou hodnotu. Funkce kontroluje, zda se výsledná hodnota odlišuje od hodnoty původní, popřípadě musí provést nový náhodný výběr. Tyto funkce využíváme například při anonymizaci jmen, názvu obcí, ulic a dalších typů dat, ke kterým existuje známý seznam hodnot.
c) Zamíchání dat (Shuffling): tato metoda je zajímavá tím, že ke své práci používá původní data. Hlavní princip spočívá v „promíchání" dat v jednom sloupci. Výsledek je pak takový, že žádný záznam nezůstane na svém místě. Vzhledem k tomu, že se jako nahrazované hodnoty používají původní data, pouze výsledná hodnota je vybrána z jiného náhodně vybraného záznamu, anonymizovaná data vypadají velice věrohodně.
d) Přidávání šumu: patří mezi méně bezpečné metody. Principem je přidávání náhodných hodnot k existujícím hodnotám, např. přičítání náhodné hodnoty ke stavu účtu klienta.
e) Aplikace pravidel pro dodržení formátu dat: funkce z této kategorie použijeme v případě, kdy máme pevně stanovené požadavky na anonymizované hodnoty dat. Tyto funkce jsou většinou jednoúčelové a nelze je použít na jiný typ dat, než pro který jsou určeny. Jako příklad můžeme uvést anonymizaci rodného čísla, IČ, DIČ, IBAN, kde je nutné dodržet formát dat.
f) Anonymizace binárního obsahu: v mnoha případech se v datech vyskytují celé dokumenty. Může jít například o vystavené faktury, fotografie osob apod. Tyto dokumenty mohou v některých případech obsahovat údaje ekvivalentní s údaji v databázi. Pokud bychom tedy anonymizovali všechna data kromě binárního

obsahu, může se stát, že velká část údajů zůstane zpětně dohledatelná na základě těchto dokumentů. Nejjednodušší řešení je nahrazení všech záznamů sloupce hodnotou NULL, případně generovat validní dokumenty odpovídajícího typu.

g) Mikro agregace: jedná se o funkci použitelnou na číselné údaje. Data určená k anonymizaci se seřadí a rozdělí na několik skupin. V rámci každé skupiny je použita agregační funkce (typicky aritmetický průměr) pro výpočet finální hodnoty a tato hodnota je použita k anonymizaci hodnot dané skupiny.
h) Maskování: jedná se o jednoduchý způsob nahrazení obsahu položky uživatelsky zadanou konstantou, pro každý záznam stejnou. Tento způsob lze použít například pro skrytí neaktuálních dat charakteru osobních údajů, při kterém nezáleží na jinakosti a je vhodné vidět, že se jedná o tento druh dat.