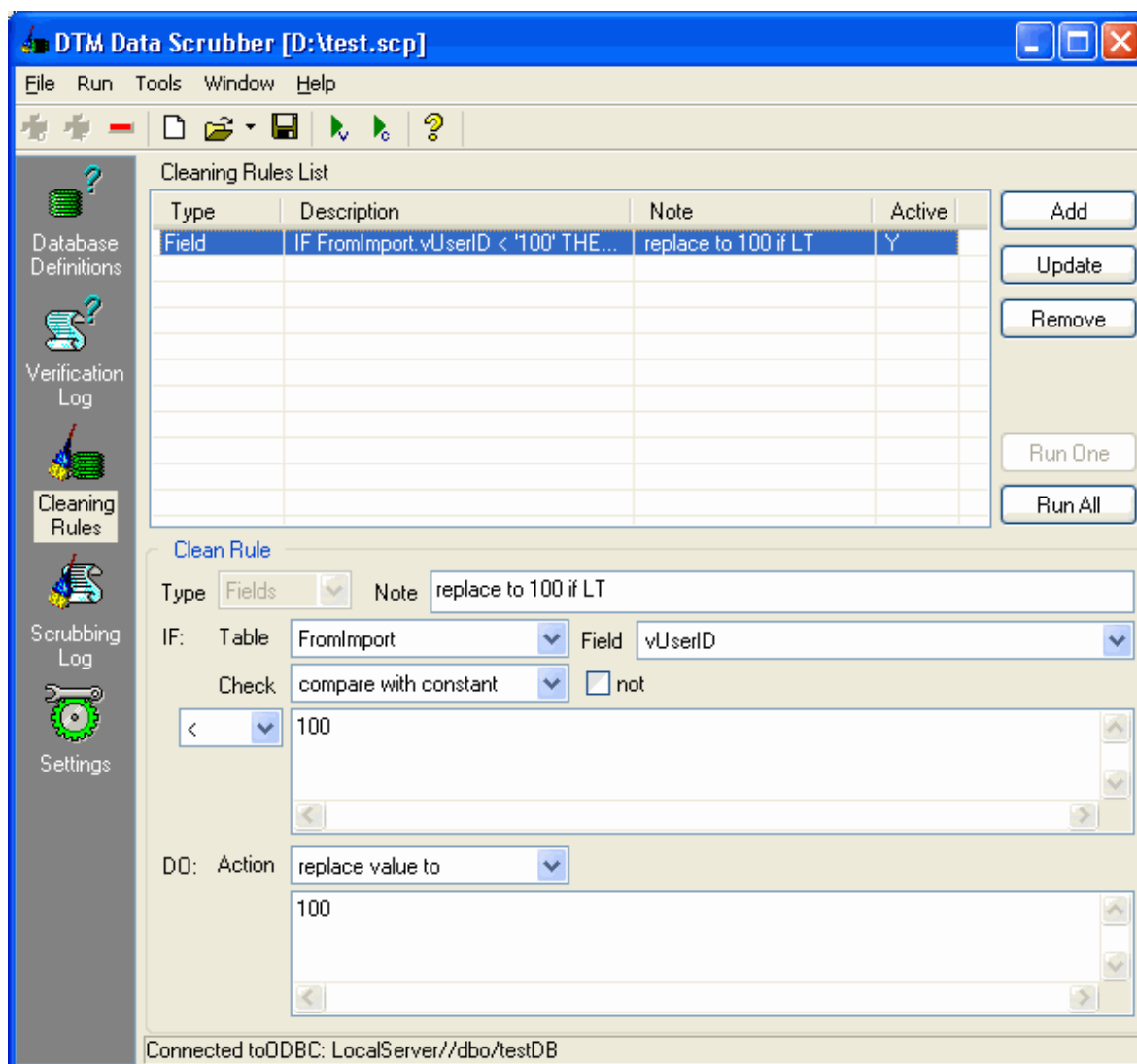


DTM Data Scrubber is a set of intelligent tools for data verification (audit) and scrubbing (cleaning). Depending on user-defined rules and data properties, the program either creates a report about the actual state of affairs or performs database data correction.

The program supports two sets of rules - for [verification](#) and for [cleaning](#). The rules of the first type describe what the correct data should look like so each table field can have only one rule of this type. The rules of the second type describe the types of problems that can happen to data values. Together with this type you can specify the action the program should perform for problem correction. There can be several rules of this type for each field and the program will perform the [action](#) corresponding to the detected conditions.

The program generates a report based on the results of verification or scrubbing (the current version offers the HTML format only).



DTM Data Scrubber is a true Win32 application, known to be compatible with the following operating systems: Windows 98 and ME, Windows 2000 family, Windows XP (Home and Professional), Windows 2003 Server family. To be able to work with Windows 95 and Windows NT 4.0, you will have to download an additional library set from our [site](#).

Verification Rules

Data verification rules describe what the correct data should look like. Each table field can have only one rule of this type. The program will add all field values of the selected table that don't match the criterion set in the rule, to the report.

Verification Rule

Type: Fields

Table Name: FromImport

Type of Check: compare with constant

100

There are five ways to connect to a database:

1. a direct connection
2. a connection to desktop files
3. a data source with ODBC, IDAPI or Oracle Call Interface (OCI)
4. a DSN file
5. an OLE DB connection

See also: [Troubleshooting guide](#).

Direct connection

The direct connection method allows you to connect to most popular databases ([MS SQL Server](#), [Oracle](#), [Interbase/Firebird](#), [MySQL](#), [PostgreSQL](#) and [DB2](#)).

Enter the server name and the database name, if required. The user name and password are optional. Their necessity depends on the settings of your database. The owner name (schema) is optional too. The list of visible database objects depends on the choice of the owner. If the owner is empty, you will access all objects.

If you do not find the required database type in the list or cannot connect directly, use a connection through the predefined data source. If DBMS is in the list, but unavailable, it means that either the required ODBC driver is not installed or it is not configured properly.

During its use, the program stores the entered values of server names, users and owners. You can select a value from the stored list using the corresponding combo box. For some DBMS types (MS SQL, for example), the program can fill the list of available databases. Use the button with two arrows for this purpose.

Direct Database Connections to most popular databases

Microsoft SQL Server
 Oracle database
 MySQL
 Interbase/Firebird
 DB2

Note: If you do not find the required database type in the list or cannot connect directly, use a connection through the predefined Data Source.

Server: DTM-ACER
User (login): sa
Password: *****
Owner/schema: dbo
Database: Northwind

Desktop files

The second way is designed for connecting to desktop data files. Select the required format and specify the file name or the directory where the data is located. Other parameters are optional.

Connections to Desktop data formats

Text file (*.txt, *.csv). Format: TabDelimited, ANSI
 Microsoft Access file (*.mdb)
 dBase, FoxBase or FoxPro file (*.dbf)
 Microsoft Excel file (*.xls)
 Paradox file (*.db)

Location: D:\SAP.mdb [Browse...]
Optional authentication information:
User (login): Admin
Password: *****
 Read Only mode

Preconfigured data sources

A connection with the use of a data source is the most universal one. You can select an interface (ODBC, IDAPI or OCI (if installed)) and the preconfigured data source name. Other options in this case are similar to those of a direct connection. The "Manage" button allows you to get access to the external configuration utility if it is available. When you want to access the tables belonging to single data scheme (or owner), you should fill the "owner" entry; otherwise all tables will be accessed.

Connections to existing and configured data sources

Interface: ODBC Manual commit

Data source: LocalServer

User (login): sa

Password: *****

Owner (optional): SCOTT

Database: Frameshop

DSN file

The fourth way is using a DSN file. In this case just select the file name with DSN definition.

Connections to existing and configured file DSN

File DSN name: d:\dsn\LocalDB2.dsn

OLE DB connection

Use 'Configure' button to specify connection information. Password and owner fields are optional.

General information

In all modes the "Test" and "Information" buttons, as well as tools for working with connection profiles are available. "Test" button allows you to check information you entered and/or data source (or alias) configuration.

Connection profile helps you to save information about your connection (interface, data source or alias name, user name (login), password and database name, etc) and get access it by the one click.

"Save" and "Load" buttons allow you to save or load profiles to external file.

The program provides detailed database, connection and driver information and properties. "Information" button allows you to view it.

Data source information

DBMS: Microsoft SQL Server [08.00.0534] Server name: TOSHIBA

Driver: SQLSRV32.DLL [03.81.9030] Manager: 03.52.9030.0000 / 03.52.0000

General information about the driver and data source associated with a connection:

Description	Vers	Value
SQL_KEYWORDS	2.0	BREAK,BROWSE,BULK,CHEC...
SQL_LIKE_ESCAPE_CLAUSE	2.0	Y
SQL_MAX_ASYNC_CONCURRENT_STATEME	3.0	1

SQL_MAX_BINARY_LITERAL_LEN	2.0	524288
SQL_MAX_CATALOG_NAME_LEN	1.0	128
SQL_MAX_CHAR_LITERAL_LEN	2.0	524288
SQL_MAX_COLUMN_NAME_LEN	1.0	128
SQL_MAX_COLUMNS_IN_GROUP_BY	2.0	0
SQL_MAX_COLUMNS_IN_INDEX	2.0	16
SQL_MAX_COLUMNS_IN_SELECT	2.0	4096
SQL_MAX_COLUMNS_IN_TABLE	2.0	1024
SQL_MAX_CONCURRENT_ACTIVITIES	1.0	1
SQL_MAX_CURSOR_NAME_LEN	1.0	128
SQL_MAX_DRIVER_CONNECTIONS	1.0	0

DBMS-specific connection options

Microsoft SQL Server

- Empty user name means use Trusted Connection
- "(local)", empty or "." server name means local server

Oracle

Use connect string for the Oracle Server that you want to access as a Server name.

Interbase and Firebird

Examples:

- Server: **localhost** and Database **c:\interbase\myDb.fdb** - connect to specified DB on local system.
- Server: **172.17.2.10/3051** and Database **/usr/local/db/myDb.fdb** - connect to specified server with alternate port 3051 on remote system 172.17.2.10

MySQL

- Use **localhost** for local MySQL
- example.com;port=3306 means MySQL at example.com on 3306 port

DB2

ServerName;port=5000;protocol=TCPIP as a server name means connect to ServerName, use 5000 port and TCP/IP protocol.

PostgreSQL

ServerName as a server name means connect to ServerName, use 5432 port and TCP/IP protocol. Database name is required. localhost as a server name is acceptable.

Troubleshooting guide

Problem description	Reasons	Solutions
Required database type not found in the list at Direct Connection and Desktop Connection pages		Switch to "data source" connection mode and to select data source from the list or to configure new using "Manage" button.
Required format is in the list, but unavailable (disabled).	ODBC driver for your database does not installed or not configured properly.	Install required driver. If it is already present in the system, try to contact our support staff .
Errors during direct connection.	Compatibility problems.	Try to create data source for your database connection.
Login error for correct user name and password.	Read-only desktop data file.	Try to change file mode to 'read and write'.

Scrubber Rules

Data Scrubbing rules describe the types of problems that can happen to data values. You can specify the [action](#) the program should perform for problem correction.

The program will perform the action for all records where field values satisfy the specified criterion. If the rule is marked as terminate and its conditions are fulfilled, this field of the current record is not checked anymore.

For example, you want to replace all empty field values with the figure 10. In this case you should select "Is this value empty" for checking, and specify replace as an action, with 10 as a parameter. For most popular DBMS you can use built-in functions as parameters and in expressions.

Clean Rule

Type: Fields Note: replace to 100 if LT

IF: Table: FromImport Field: vUserID

Check: compare with constant not

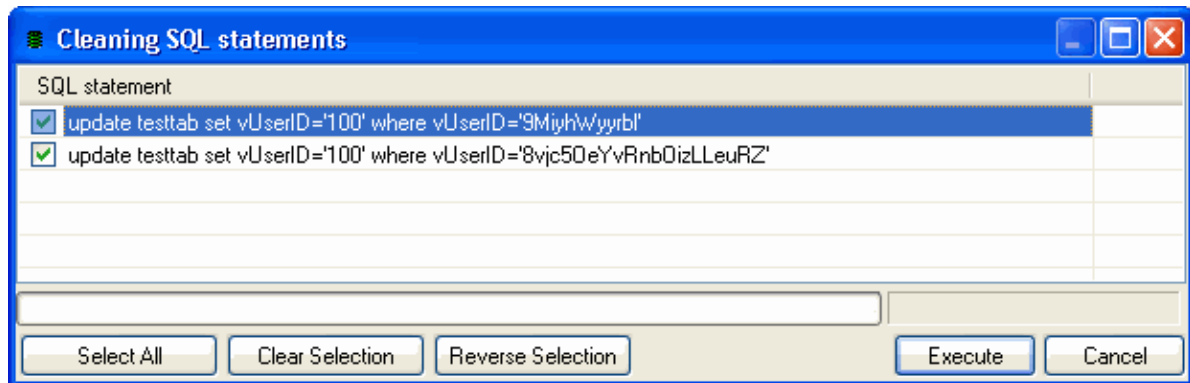
< 100

DO: Action: replace value to

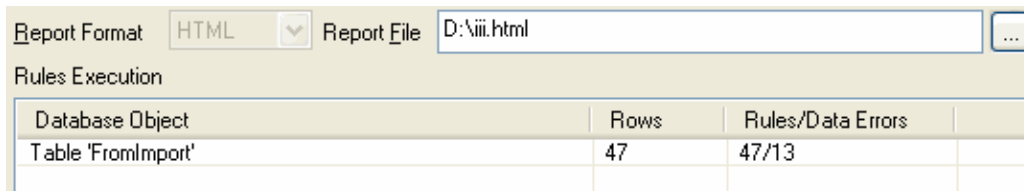
100

Execution Results

In the data scrubbing mode the program creates a set of update statements for database correction. The user can select the necessary statements or forbid the unnecessary ones in a special dialog box.



Execution log (console) is a list of the executed items of the project. It allows you to view execution process and results of the project execution. There are two logs present (for verification and for data scrubbing mode).



The screenshot shows a software interface with a 'Report Format' dropdown set to 'HTML' and a 'Report File' text box containing 'D:\iii.html'. Below this is a section titled 'Rules Execution' containing a table with the following data:

Database Object	Rows	Rules/Data Errors
Table 'FromImport'	47	47/13

The program supports the following types of data checks:

- Compare the field value with a constant
- Compare the field value with a SQL statement execution result
- Are there values matching a regular expression
- Is the value NULL
- Is the value empty
- Is the value within the range of values
- Search substring

A regular expression is a special template type allowing you to create effective [checks](#) and [substring changes](#) in data. An expression consists of usual characters and meta symbols.

Most popular meta symbols and language clauses are:

Item	Match with
^	the beginning of line
\$	the end of line
.	any character
[]	characters in set
[^]	characters not in set
	previous or next pattern
?	previous pattern 0 or 1 times
+	previous pattern 1 or more times
*	previous pattern 0 or more times
()	group pattern and remember to \n variable

If part of the expression is enclosed in brackets, after the matching is over, the substring value gets into a variable of the type \n, where n is the number of the expression from the beginning of the string. These variables can be used only in replace [actions](#).

See also: [How to...](#)

The program supports the following types of actions:

- Changing the value to a constant or an expression (only for match with a [regular expression](#))
- Setting a value for a NULL field
- Do nothing (usually uses for terminate rules)
- Replace substring (first match)
- Replace substring (all matches)

Q: How can I replace substring AAA in field value to BBB?

A: Create rule using "match regular expression" condition.

Use `(.*)AAA(.*)` as a condition value and `\1BBB\2` as a replace string.

A: Create rule using "find substring" condition.

Use AAA as a condition value and BBB as a replace string.

The program has some settings and options aided to make your work more comfortable.

General Settings

- Restore last connection at program startup
- Ask for connect if not present
- Load recent project during program startup
- Default interface mode is Verify

This License Agreement covers all existing versions of DTM Data Scrubber: demo and commercial. License Agreement is a legal agreement between the end-user and DTM soft. By installing or otherwise using DTM Data Scrubber, you agree to be bound by the terms of this Agreement. If you do not agree to the terms of License Agreement, please do not install or use DTM Data Scrubber.

General Information

1. DTM soft is exclusive owner of all DTM Data Scrubber copyrights. The product is licensed, not sold. DTM Data Scrubber is protected by copyright laws and international copyright treaties.
2. Demo version. Anyone may install and use demo version of DTM Data Scrubber for evaluation and testing purposes.
3. Commercial version. You may install and use one copy of DTM Data Scrubber on a single computer. The primary user of the computer on which DTM Data Scrubber is installed may make a second copy for his or her exclusive use on a portable computer.
4. You may not reverse engineer, modify, decompile, or disassemble DTM Data Scrubber. The Software Product is licensed as a single product. Its component parts may not be separated for use on more than one computer.
5. You may not rent, lease, or lend DTM Data Scrubber. Also, You may not resell, or otherwise transfer for value, DTM Data Scrubber.
6. Without prejudice to any other rights, DTM soft may terminate this License Agreement if you fail to comply with the terms and conditions of this Agreement. In such event, you must destroy all copies of DTM Data Scrubber and all of its component parts.
7. You may permanently transfer all of your rights under this license, provided you retain no copies, you transfer all of DTM Data Scrubber (including all component parts), and the recipient agrees to the terms of this license.
8. DTM Data Scrubber IS DISTRIBUTED "AS IS". NO WARRANTY OF ANY KIND IS EXPRESSED OR IMPLIED. YOU USE DTM Data Scrubber AT YOUR OWN RISK. DTM soft WILL NOT BE LIABLE FOR DATA LOSS, DAMAGES, LOSS OF PROFITS OR ANY OTHER KIND OF LOSS WHILE USING OR MISUSING THIS SOFTWARE.

Support and Upgrades

During one year after ordering, you are entitled to free technical services and support for DTM Data Scrubber which is provided to you by DTM soft. During this period, e-mail support is unlimited and includes technical and support questions. Also, during one year, you may access to free updates to DTM Data Scrubber when and as DTM soft publishes them on www.sqledit.com. After end of the described period you may continue to use the software product in accordance with the terms of this Agreement except free support and upgrades. After end of the free support and updates period (one year), you may purchase annual Upgrade and Support subscription. If you ordered a few DTM licenses, you will access to free upgrade and support period and will use subscriptions independently.

Trademarks information

DTM Data Scrubber is trademark of DTM soft.

If You have some question or unusual problem feel free to contact the DTM Data Scrubber technical support at support@sqledit.com

When you contact technical support, you should be prepared to provide the following information:

- DTM Data Scrubber version (you can find this information from About menu item of Help menu).
- Type and version of the ODBC or IDAPI driver or OLE DB provider.
- DBMS version and operating system version (including service pack version, if applicable).
- DTM Data Scrubber Log file.
- A description of what you do before the problem occurs.
- Error messages you see when the problem occurs.
- Your name, company name and how to contact you.

Log file

When running the program, you have to select one of menu items from "**Tools->Log file**" in order to view or truncate your Log file. Log file contains detailed description of any errors and other events that occurred while processing script.



Program Installation

- Click one the windows Start menu and select "Run" item
- Select or enter installation file name and path (scr.exe or scr_demo.exe) and click OK



Uninstall the Software

The Uninstall program removes all installed DTM Data Scrubber components and all records in the Windows registry made by the installation script. You can invoke this program by selecting the "**DTM Data Scrubber**" item in "**Add/Remove Programs Dialog**" in "**Control Panel**".

Important! Uninstall feature of the program does not remove files and objects created by users such as configuration files, registry records etc.



How to upgrade your copy of DTM Data Scrubber?

Please contact our support staff at support@sqledit.com. When you upgrade your copy of this software please send us the following information:

- You name, company name and how to contact you
- Payment information (at least "ORDER No" and "Date")

Demo version is available for [download](#) free of charge.

[DTM SQL Editor](#) is a set of powerful database management tools that allow you to achieve two goals - to have unified access to different types of databases and to have a set of solutions that makes processing your data easy. DTM SQL Editor gives database users, developers and administrators an ability to access different databases, whether desktop or client-server ones (provided you have ODBC driver installed). This is very convenient, since most organizations use several different types of databases installed and each stores data in different formats and with varying parameters. Having a program that can get data from various sources is often essential. Furthermore, in addition to letting you quickly switch between different data sources, DTM SQL Editor lets you see database schema and results of the query execution.

[DTM Migration Kit](#) is a powerful yet simple data migration tool that comes in handy if you run multiple databases. Use it to import, export or migrate data between different data sources (ODBC, IDAPI, or Oracle Call Interface supported). The program is fully automatic and supports all popular database formats. Simple visual interface lets you set own transformation and flow control rules to give you added flexibility.

[DTM Schema Reporter](#) is a reporting tool for database schema. The program creates reports in RTF, HTML, XML or plain text formats and supports all common database interfaces - ODBC, IDAPI, or even Oracle Call Interface. This utility helps technical writers and database administrators create a report of any complexity level within seconds. Also, you can alter table order in the report and manually add annotations to the individual tables.

[DTM Data Generator](#) is a simple, powerful and fully customizable utility that generates data for database testing purposes. Currently, database developers and administrators often have to spend hours of dull work to create test data sets before examining database performance. This tool makes all this unnecessary by automatically creating database objects AND sets of SQL statements, if necessary.

[DTM Data Editor](#) is a data viewer and editor for database professionals who are tired of wasting their time on mundane tasks. The program uses form-based interface and works with any ODBC data source. SQL statements are generated automatically and can be modified later. For data that has foreign key - primary key relation, there are options to enter values manually or select them from a list, which is much faster.

[DTM DB Stress](#) is a utility for stress testing the server parts of information systems and applications, as well as DBMSs and servers themselves. This tool allows you to create and configure a continuous set of requests to the server of the OLAP (query execution) and OLTP (adding, modifying and deleting data in the database) types. At the same time, the user can flexibly change both the number and the priority of this or that type of requests to a database or an application.

[DTM Data Modeler](#) is a CASE tool for database developers that supports both forward and reverse engineering. It is an easy-to-use tool allowing you to work both with logical and physical data models in the form of an entity-relationship diagram. The product is intended for database architects and developers and works with data sources via the ODBC interface, which means compatibility with all modern DBMS. Along with basic model properties (sets of entities and relationships between them), the program allows you to create indexes and triggers on the physical level corresponding to the tables of the database that is modeled.

[DTM Data Scrubber](#) is a set of intelligent tools for data verification (audit) and scrubbing (cleaning). Depending on user-defined rules and data properties, the program either creates a report about the actual state of affairs or performs database data correction.

[DTM Data Comparer](#) is a visual tool for data compare and synchronization. The program successively views the contents of both tables basing on the order of ascending of unique key values and shows differences or creates synchronization script.

[DTM Schema Comparer](#) is a tool for comparing database schemas. The comparison process supports tables, views, indexes, triggers and stored procedures. The visual representation of database schemas as a tree makes the comparison process more comfortable.

[DTM Query Repoter](#) is a reporting tool for database query. This utility helps technical writers, developers and database administrators create a report based on database query within seconds.

Database catalog - The collection of system tables, tables that store metadata about that specific database.

Database record - one row in a table (table can be a result of SQL-query).

Database schema - logically connected, usually owner-based, set of DBMS objects (tables, views, procedures etc).

DBMS - database management system.

DBMS connection - the fact that both client and DBMS server have signed a contract and ready to query and data communications.

Drag-n-drop - the file manipulation technique when mouse is used to move the file from the place of storage to the program, which performs processing.

SQL language - declarative language used to manipulate the data and its' structure in the modern DBMS and their client applications.

IDAPI - Integrated Database Application Program Interface, unified DBMS access interface.

OCI - Oracle Call Interface, access interface for Oracle Server.

ODBC - Open Database Connectivity, unified DBMS access interface.

Metadata - information about data. See also: database schema

SQL statement - single SQL operator having the complete role in a data manipulation script.

SQL-server - program or program complex, which is able to execute the SQL-queries.

