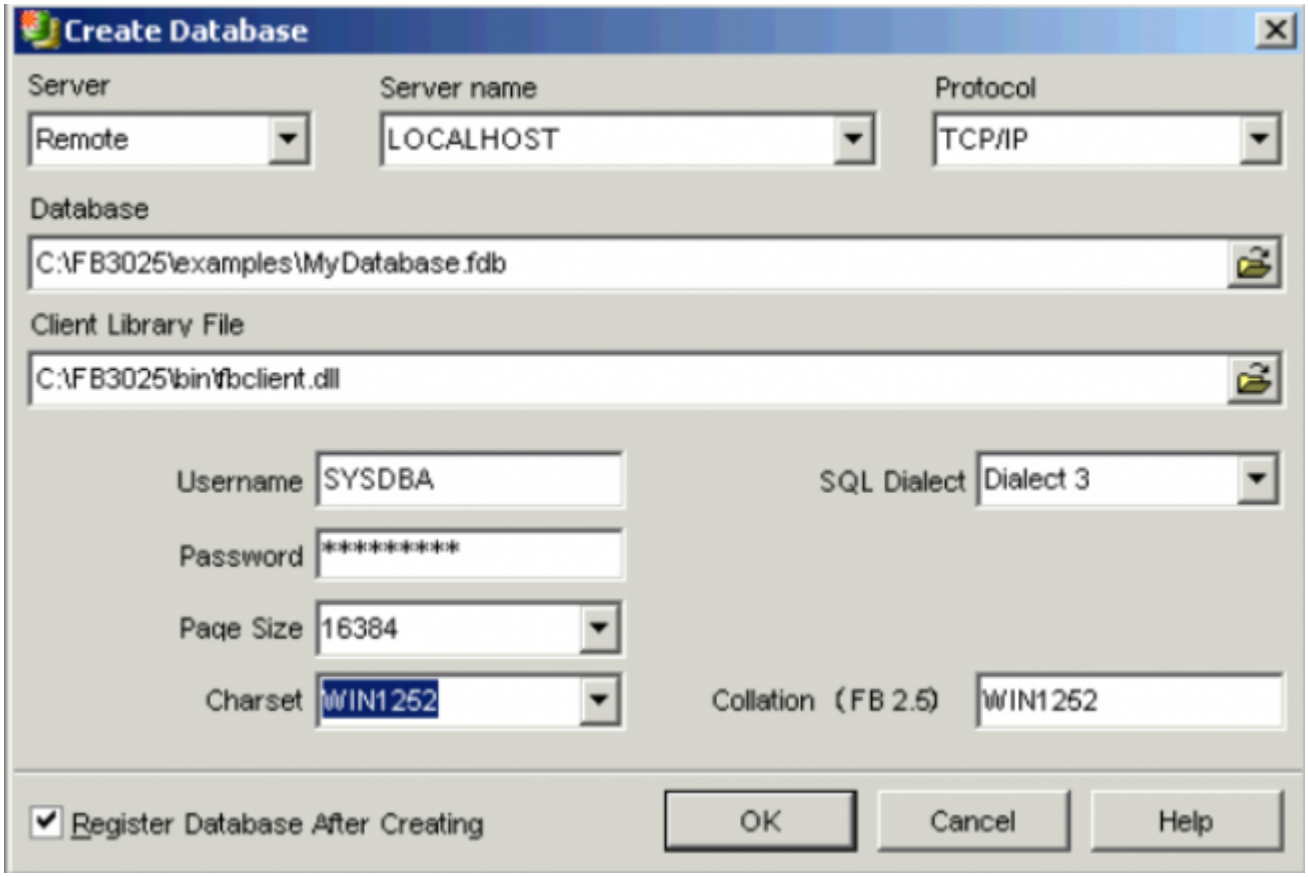
Default character set

The default character set is the character set defined when creating the database, and applicable for all areas of the database unless overridden by the domain or field definition. It controls not only the available characters that can be stored and displayed, but also the collation order, and influences casing (UPPER(), LOWER()). Given a specific character set, a specific collation order can be specified when data is selected, inserted, or updated in a column.

Firebird/InterBase® supports multiple character sets for use around the world. If no special character set is specified for individual columns, the database default character set is assumed. The default character set is defined in IBExpert in the Create Database dialog:



If a character set is defined as the default character set when creating the database (and since Firebird 2.5 also the default collation), it is not necessary to define this again for individual columns.

Firebird/InterBase® supports more than 20 different character sets directly. The chosen character set is also of importance when importing and exporting data with different character sets. This needs to be taken into consideration when applications are developed with multiple language versions.

The ASCII character set is not synonymous with a non-defined character set. If no character set is defined, Firebird/InterBase® chooses the character set NONE, i.e. values are stored exactly as typed. You can load any character set into a column defined with NONE, but you cannot load that same data into another column that has been defined with a different character set. In that case, no transliteration is performed between the source and destination character sets, and transliteration errors may occur during assignment. When the ASCII character set is specified, all characters are translated into the ASCII equivalents from the character set under which they were input.

Character sets may have a number of collations (for example, ISO8859_1 offers English, German, French, Spanish, Italian, and so on, and so on). From Firebird 2.5 upwards it is also possible to specify a default collation which can be altered as needed. For those working with older Firebird versions, it is necessary to specify the collation at field level or in your queries when a collation sorting order is required.

Generally this default character set cannot be altered at a later date (only using the command line tool IBEScript). Alternate character sets can however be defined for individual domains and tables, which override the default character set.

IBExpert offers full Unicode support. The internal representation of all texts in the code editors is Windows Unicode (UTF-16LE, two bytes per character). This allows you to use multilingual characters in procedures, queries, database object descriptions etc., if you use the UTF8 character set when connecting to your database.

When you're working with a database using the UTF8 character set IBExpert performs automatic conversion from UTF8 to Windows Unicode (for example, when you open a stored procedure for editing) and backwards (when you compile a stored procedure). This applies to Firebird 2.1 and 2.5 databases. For other databases you will need to enable this behavior manually (if you really need this!) by flagging the *Do NOT perform conversion from/to UTF8* checkbox in the Database Registration Info.

If you use ANSI character sets to work with your databases there are no changes to previous versions of IBExpert, except it is now possible to enter characters that are not presented in your default system locale. Such characters will be replaced by converting from Unicode to ANSI representation.

As a rule IBExpert knows when it must convert strings from windows unicode to UTF8 but sometimes it is necessary to specify the conversion type manually (please refer to the IBExpert Tools menu item, Script Executive for further information.

Specifying the right default character set for your application and needs is vital. We therefore strongly recommend first reading the following articles before deciding which character set to define as the default when creating a database:

- Character Set
- Character sets and Unicode in Firebird
- Convert your Firebird applications to Unicode
- Firebird 2.1.3 Release Notes: International character sets

and, depending on your database version:

- Overview of the main character sets in Firebird
- InterBase® character sets and collation orders
- New character sets in Firebird 2.5
- New character sets in Firebird 2.1
- New character sets in Firebird 2.0
- Firebird 2.0 Language Reference Update: Character set NONE
- Firebird 2.1 Release Notes: International Language Support (INTL)
- Firebird 2.1 Release Notes: Appendix B: International character sets

From:
https://www.ibexpert.com/docu/ - **IBExpert**

Permanent link:
https://www.ibexpert.com/docu/doku.php?id=01-documentation:01-13-miscellaneous:glossary:default-character-set

Last update: **2023/08/14 16:23**