Exporting Data from KE Texpress

Description:

The KE Texpress *texforms* module supports a general purpose data exporting facility. From Display mode, two commands are available on the Report pull-down menu. These commands, *Export this match* or *Export all matches*, apply to the current record or all matching records, respectively.

The user is asked to select one of a list of different export formats. This list is extensible, with system and database administrators able to add new export formats to this list. The export formats are discussed below.

Having selected a format, the user is presented with a list of all columns (items) of the table, from which individual columns can be selected for exporting. Individual columns can be selected by positioning the cursor on the column's name or prompt and typing **F8**. The order in which columns are selected determines the order of the exporting of the column data. All columns can be selected for exporting by typing **F10**.

After some or all of the columns have been selected, the user simply types **<Enter>** and is then presented with another list of the column names from which sorting criteria can be defined. A column can be used to sort the exported records by positioning the cursor on the column's name or prompt and typing either **F8**, for ascending sort order, or **F9**, for descending sort order. Note that columns not specified for exporting can still be used to sort the records. No sort order need be defined.

After the sorting criteria have been determined, the user is then requested to select an existing file, enter a new file name or select a printer to which the exported data is sent.

The pre-defined export formats are as follows:

Comma delimited

> The selected columns are exported separated by the comma character. Text and string fields are surrounded by the double quote character. Any double quote characters within a text or string field are preceded by a double quote character.

Fixed width

> The selected columns are exported in fixed width fields, with the field width being that defined in the Insertion form. Data is padded with space characters.

Ids

> The selected columns are exported in the format:
>
> > `colid:fieldno=data`
>
> where *colid* is the column's Id, *fieldno* is the number of the field starting from **1** and *data* is the data in the field.
>
> Each field is output on a separate line. Each record is preceded by a line of the form:
>
> > `rownum=nnn`
>
> where *nnn* is the number of the record within the set of records being exported. Each record is then terminated by a trailing line of the form:
>
> > `###`

Ids (brief)

> The *Ids (brief)* export format is identical to the *Ids* format, except that where a field is empty, the entire line is omitted from the output.

Texpress data

> This export format produces KE Texpress internal data file format, also known as *dave* file format. The order of the columns in the data file produced is the order of the columns as selected by the user. This may be different from the internal data file format for the database and so care should be taken when using this option as the resulting data file may appear inconsistent.

Adding New Export Formats

The export formats are defined locally for a single database in the **export** file in the database directory and globally for all databases in the system in the file **~texpress/etc/export**. The five formats described above appear in the latter file.

Each export file consists of lines of the format:

> `label:command`

where *label* is the string displayed in the initial selection box of supported formats and *command* is the Unix command to which the data is passed for exporting. In this file, lines beginning with # are comment lines and are ignored, as are blank lines.

The command can contain various operators which alter the exporting sequence. These are evaluated from left to right and are as follows:

+c   This results in the user being asked to select one or more columns from the database, which are the columns to be exported. The *+c* sequence is replaced by a list of the form:

        colid,colid,…

where *colid* corresponds to the selected column name.

If no *+c* operator appears in the command, then the user is <u>not</u> requested to select columns to be exported.

+s   This results in the user being asked to select zero or more columns from the database, which are to be used to sort the records before they are exported. The *+s* sequence is replaced by a list of the form:

        colid,colid,…

where *colid* corresponds to the selected column name.

If no *+s* operator appears in the command, then the user is <u>not</u> requested to select columns to be used for sorting.

+o   This sequence is replaced by the name of a temporary file into which *texforms* places internal pointers to the selected set of records. Currently, this option is essential for all export formats as it is the only way to pass information about the records selected for export.

+d   This sequence is replaced by the name of the current database. This is useful when export formats are configured for the entire KE Texpress installation or when export formats are defined for a database which may be copied or renamed.

+f   This results in the user being requested to select a file or printer into which the output of the export is placed. If this option does not appear in the command then this selection box is <u>not</u> displayed.

If this option is included in the command, then the export of the data occurs in background.

+i   This option signifies that the export of data may be interactive and should be run in foreground.

All default export formats and, most likely, all user defined export formats, make use of a KE Texpress program called **texexport**. This program is described below.

Program Description: texexport

Usage:

texexport [-c*collist*] [-f*style*] [-s*sortlist*] [-m*modifiers*] [-nr] -ofile dbname

Arguments:

Optional arguments include:

-c*collist*	is the list of columns to export from the table. This is specified as a comma separated list of column Ids optionally followed by a colon and field number, e.g. (colid[:fldno],colid[:fldno]...)
-f*style*	is the style or format of the output (see below)
-s*sortlist*	is a list of the columns for sorting (in the same format as the *collist*). Columns used for sorting do not have to be included in the exported columns.
-m*modifiers*	defines modifier options for the selected style.
-n	means notify upon completion.
-r	means remove file at completion.

You must specify:

-ofile	file of offsets of rows in dbname table. This is generated by *texforms* when either the *Export this match* or *Export all matches* options is selected.

Formats include:

ids	lines of the format *colid:fieldno=data*
idsbrief	as for ids but with empty fields omitted

   optional modifiers
-ml	print Library items as a single value
-mw*n*	print column name in exactly n characters, blank padded on the right

texpressdata	KE Texpress internal data file format (dave format)
texql	KE Texql insert statements
delimited	character-separated columns

   optional modifiers
-md*ch*	use *ch* as text delimiter (dflt ")
-ms*ch*	use *ch* as column separator (dflt tab)
-mc	include column names on first line
-ml	print Library items as single value

fixedwidth	fixed width (blank-padded) output

format	column output is defined by a format string
formatbrief	as for format but with empty lines omitted

    optional modifiers
      -mf*str*   use *str* as the format string
      -ml   print Library items as single value

Examples:

The following examples are taken from the file, **~texpress/etc/export**.

```
Comma delimited:texexport -fdelimited -c+c -ms',' -s+s +f -r -o+o -n +d
```

> This selects the *delimited* option for *texexport*. The command sequence causes *texforms* to request the columns to be exported, which are substituted in place of +*c*, and the columns to be used for sorting the records (+*s* option). The sequence, +*f*, causes *texforms* to request a destination (file or printer) from the user, while the sequence, +*o*, is replaced by the temporary file containing the record pointers. Note that the -*r* option causes this file to be removed upon completion. The database name is substituted for the +*d* sequence.

> This command includes the modifier option, -*ms','*, which causes the field separator to be changed to a comma from the default tab character.

```
Fixed width:texexport -ffixedwidth -c+c -s+s +f -r -o+o -n +d
```

> This selects the *fixedwidth* option for *texexport*. The command sequence causes *texforms* to request the columns to be exported, the columns to be sorted and the output file or printer, as above.

```
Ids:texexport -fids -c+c -s+s +f -r -o+o -n +d
```

> This selects the *ids* option for *texexport*. The command sequence causes *texforms* to request the columns to be exported, the columns to be sorted and the output file or printer, as above.

```
Ids (brief):texexport -fidsbrief -c+c -s+s +f -r -o+o -n +d
```

> This option is identical to the *ids* option above, except that lines corresponding to empty fields are omitted from the output.

```
Texpress data:texexport -ftexpressdata -c+c -s+s +f -r -o+o -n +d
```

> This selects the *texpressdata* option for *texexport*. The command sequence causes *texforms* to request the columns to be exported, the columns to be sorted and the output file or printer, as above.

Other examples could include:

```
Mail Merge Format:texexport -fdelimited -mc \
               -cfirstnam,surname,addr:1,addr:2,suburb,pc \
               -ssurname +i -r -o+o people > /tmp/mailmerge.txt
```

This selects the *delimited* option for *texexport* with the optional modifier of *-mc* to produce a first line containing the column names. The columns are then explicitly stated as part of the command and so the user is not asked to select columns (i.e. no *+c* operator). Also there is no request of the user for sorting criteria. Instead the records are automatically sorted by surname. The output is directed to the Unix file */tmp/mailmerge.txt*.

This export option requires no further input from the user.