

MGconvert.exe - Simple MS Windows CMD line tool
Converts a text file to an encoding or code page
(c) Markus Gnam 2019. Version 1.0.6.0 20191130

USAGE: MGconvert <input file> <from> <to> [out]
Output file name is the input file name + <to>
If an [out] name is given this is used instead.

Allowed values for <from> and <to>:

UTF-8 or UTF8
UTF-8BOM or UTF8BOM
UTF-16LE or UTF16LE or UNICODE
UTF-16BE or UTF16BE or BIGENDIANUNICODE
UTF-32LE or UTF32LE
UTF-32BE or UTF32BE
OEM or CP_OEMCP
ANSI or CP_ACP
LATIN-1 or ISO-8859-1 or 28591
CP437 or 437
CP850 or 850
CP1252 or 1252
... or any other valid code page ...

Additional notes:

Example:

```
mgconvert test.txt utf8 1252
```

⇒ File test.txt successfully converted from UTF-8 to CP1252 into file test.txt.1252

The same example with given [out] name "result.txt" as output file name:

```
mgconvert test.txt utf8 1252 result.txt
```

⇒ File test.txt successfully converted from UTF-8 to CP1252 into file result.txt

Hints:

- This program takes advantage of MS Windows Best-fit Mapping API functions. Example: mgconvert test.txt utf8 1252

with UTF-8 encoded content: Antonín Dvořák

⇒ Best fit conversion to ANSI code page 1252: Antonín Dvorák instead of Antonín Dvořák

- Note that any existing files will be overwritten, without warning.

- Be careful to choose the correct input encoding. If the input is wrong the output will be wrong too. Also, be aware that specific code pages that are not installed on a PC will most probably fail to work.

If a not valid code page is recognised, there is an error message, e.g.: [ERROR]: Code page 4232 not allowed

- Special cases: Some characters can't be converted (e.g. Chinese characters won't fit into code page 437). These characters are replaced by "?". If the transliteration is longer than the original character this character is replaced by "_". Example: "Þorvaldur" does not fit into CP 437. The transliteration "Thorvaldur" is too long so it gets "_orvaldur".

- If the <from> value is the same as the <to> value there is a warning:
[WARNING]: <from> value is the same as <to> value
However, this may be useful to check for the correctness of an encoding or code page: If the output file is identical with the input file, this means that the input data contains valid characters according the value.
- If the [out] value is the same as <from> or <to>, there is a message:
[ERROR]: [out] value is the same as <from> or <to>
This is done to avoid accidentally confusing one of these parameters.
- All Unicode encodings except UTF-8 (UTF-16LE, UTF-16BE, UTF-32LE, UTF-32BE) are created with BOM. The default for UTF-8 is to use no BOM. If you want to generate a file with BOM for UTF-8, use UTF-8BOM as <to> value.
- I would like to thank Ralf Junker from Delphi Inspiration (<https://www.yunqa.de>) for some valuable tipps about UTF-32.
- This program was written with Embarcadero Delphi XE7.
- If you want to contact the author please have a look at the file "history.txt" inside the file MGconvert.zip.
- This program is distributed as "freeware," copyright reserved by the author. There are no warranties of any kind, nor any liability by the author. Users accept full responsibility for the use they make of the software and for any damage caused thereby.

Version history:

File name: MGconvert.exe

SHA256: 3d3b0b761310e7eee15395ed4cc93303fec6106ac515482b2c80f617c0ebaaf0

- Converts a text file to an encoding or code page -
- (c) Markus Gnam 2019. Version history:

Version 1.0.6.0 20191130

- Encoding UTF-32LE and UTF-32BE added.

Version 1.0.5.0 20191109

- Support for encoding UTF-7 dropped.

Version 1.0.4.0 20191003

- Support for any valid code page added.

Version 1.0.3.1 20191002

- Optional [out] parameter added.

Version 1.0.2.0 20190928

- Encoding UTF-7 added.

Version 1.0.1.1 20190922

- Encoding UTF-16BE added.

Version 1.0.0.0 20190914

- Initial version.