

NAME

gensprep - compile StringPrep data from files filtered by filterRFC3454.pl

SYNOPSIS

gensprep [**-h, -?, --help**] [**-v, --verbose**] [**-c, --copyright**] [**-s, --sourcedir** *source*] [**-d, --destdir** *destination*]

DESCRIPTION

gensprep reads filtered RFC 3454 files and compiles their information into a binary form. The resulting file, **<name>.icu**, can then be read directly by ICU, or used by **pkgdata(8)** for incorporation into a larger archive or library.

The files read by **gensprep** are described in the **FILES** section.

OPTIONS

-h, -?, --help

Print help about usage and exit.

-v, --verbose

Display extra informative messages during execution.

-c, --copyright

Include a copyright notice into the binary data.

-s, --sourcedir *source*

Set the source directory to *source*. The default source directory is specified by the environment variable **ICU_DATA**.

-d, --destdir *destination*

Set the destination directory to *destination*. The default destination directory is specified by the environment variable **ICU_DATA**.

ENVIRONMENT

ICU_DATA

Specifies the directory containing ICU data. Defaults to **\${prefix}/share/icu/73.2/**. Some tools in ICU depend on the presence of the trailing slash. It is thus important to make sure that it is present if **ICU_DATA** is set.

FILES

The following files are read by **gensprep** and are looked for in the *source* /misc for rfc3454_*.txt files

and in *source* /unidata for NormalizationCorrections.txt.

rfc3453_A_1.txt Contains the list of unassigned codepoints in Unicode version 3.2.0....

rfc3454_B_1.txt Contains the list of code points that are commonly mapped to nothing....

rfc3454_B_2.txt Contains the list of mappings for casefolding of code points when
Normalization form NFKC is specified....

rfc3454_C_X.txt Contains the list of code points that are prohibited for IDNA.

NormalizationCorrections.txt

Contains the list of code points whose normalization has changed since
Unicode Version 3.2.0.

VERSION

73.2

COPYRIGHT

Copyright (C) 2000-2002 IBM, Inc. and others.

SEE ALSO

pkgdata(8)