

# Rosette Unicode 対応ライブラリ

## 文字コードが異なるデータを Unicode に変換して処理したい、 外字人名漢字もきちんと表示 したい方などにお勧めのツール

ソフトウェアやインターネットアプリケーションを国際市場に対応させるには、言語にかかわる事項、特に文字コードが大きな課題となります。先進的な Rosette® Unicode 対応ライブラリ (RCLU: Rosette Core Library for Unicode) は多言語対応の際に直面する文字コード処理の課題を解決します。

各国固有の文字コードを世界共通の Unicode に変換することにより、どのような文字コードのテキストデータも処理できるようになります。またデータ出力などで Unicode 以外の文字コードのサポートが必要な場合には、その文字コードへ変換して処理できます。

### 旧来の文字コードへの対応

旧来の文字コードから Unicode へ、またその逆の変換を、確実かつ効率的におこないます。

### マルチプラットフォーム対応

Microsoft Windows、Sun Solaris、Linux 等、現行の主要プラットフォームすべてに対応します。

### 簡単なテキスト操作

データのソート、正規化、文字列の一致等の機能を容易に実現するため、さまざまな文字変換及び文字属性分類の機能を備えています。

### 日本語対応

ISO-2022-JP (JIS)、Shift-JIS、EUC-JP などといった日本固有の文字コードでのデータ処理が必要となる場合、これら日本語固有の文字コードから Unicode への変換、あるいはその逆をおこなうためのツールとしてもご利用いただけます。JIS X0213 (JIS 第1～第4水準)にも対応しています。

### 文字コード変換

Unicode 5.0 と160種類以上の旧来の文字コードとの間の変換をサポートしています。

### UNICODE 文字の分類と変換

176種類の文字属性および文字種によって Unicode 文字の分類と置き換えを行います。

### 高い実行性能

高速かつ効率的で、スレッド・セーフ。

### 文字コードテーブルのダイナミック・ロード

メモリー必要量を最小限に抑えます。

### 仕様

- 160種類以上の旧来の文字コードに対応
- Unicode 対応エンコーディング  
Unicode標準 1.2、2.X、3.X、4.X、5.X  
UCS-2 (big-endian または little-endian)  
UTF-7、UTF-8 フォーマット  
バイトオーダーマーク (BOM) 有/無  
Java 及び SGML の文字参照
- 30種類の文字変換
- コンパクトなライブラリサイズ - 別に500KB - 10MB の変換テーブル

### お問合せ

さらに詳しい製品情報ならびに評価版のご利用をご希望の方は下記へご連絡ください。

[info2011@basistech.jp](mailto:info2011@basistech.jp)

[www.basistech.jp](http://www.basistech.jp)

電話 03-3511-2947



Basis Technology は Unicode Consortium のメンバー企業として、Unicode 標準の策定や普及等にも貢献を続けています。

## 対応文字コード

アラビア語 — CP10004	ヘブライ語 — CP10005	ラテン文字 — ISO 8859-15
アラビア語 — CP1256	ヘブライ語 — CP1255	ラテン文字 — ISO 8859-2
アラビア語 — CP20420	ヘブライ語 — CP28598	ラテン文字 — MacRoman
アラビア語 — CP28596	ヘブライ語 — CP38598	ラテン文字 — NextStep
アラビア語 — CP720	ヘブライ語 — CP862	ラテン文字 — Adobe-Standard
アラビア語 — CP864	ヘブライ語 — ISO 8859-8	ラテン文字, カナダフランス語 — CP863
アラビア語 — ISO 8859-6	アイスランド語 — CP10079	ラテン文字, 中央ヨーロッパ — CP28592
アラビア語 — CP708	アイスランド語 — CP861	ラテン文字, 中央ヨーロッパ — MacCentralEuropean
バルト文字 — CP1257	アイスランド語 — MacIcelandic	ラテン文字, 東ヨーロッパ — CP1250
バルト文字 — CP28594	日本語 — CCSID 1027	ラテン文字, エスペラント — CP20905
バルト文字 — CP775	日本語 — CCSID 290	ラテン文字, ポルトガル語 — CP860
バルト文字 — ISO 8859-4	日本語 — CCSID 930	ラテン文字, 南東ヨーロッパ — ISO 8859-3
バルト文字 — ISO 8859-13	日本語 — CCSID 939	ラテン文字, 英語 — ASCII
ケルト文字 — ISO 8859-14	日本語 — CCSID 942	ラテン文字, 英語 — CP037
中国語(簡体字) — CCSID 935	日本語 — CP10001	ラテン文字, 英語 — CP1026
中国語(簡体字) — EUC-CN	日本語 — CP20290	ラテン文字, 英語 — CP1252
中国語(簡体字) — GB2312	日本語 — CP21027	ラテン文字, 英語 — CP20105
中国語(簡体字) — HZ-GB-2312	日本語 — EUC-JP	ラテン文字, 英語 — CP437
中国語(簡体字) — CP936	日本語 — ISO-2022-JP	ラテン文字, 英語 — CP500
中国語(簡体字) — MacChineseSimplified	日本語 — JIS_X_0201	ラテン文字, 英語 — CP875
中国語(繁体字) — CCSID 937	日本語 — JIS_X_0208	マラーヤラム語 — CP10017
中国語(繁体字) — CNS-11643-1986	日本語 — MacJapanese	Nordic — CP865
中国語(繁体字) — CNS-11643-1992	日本語 — Shift-JISMS	Nordic — ISO 8859-10
中国語(繁体字) — EUC-TW	日本語 — Shift-JIS78	ルーマニア語 — MacRomanian
中国語(繁体字) — GB12345	朝鮮語 — CP10003	スラブ文字 — CP852
中国語(繁体字) — Big5	朝鮮語 — CP1361	シンボル — Adobe-Symbol
中国語(繁体字) — CP10002	朝鮮語 — CP949	シンボル — Adobe-Zapf-Dingbats
中国語(繁体字) — CP950	朝鮮語 — EUC-KR	シンボル — CP10008
中国語(繁体字) — MacChineseTraditional	朝鮮語 — ISO-2022-KR	シンボル — MacDingbats
クロアチア語 — MacCroatian	朝鮮語 — Johab	シンボル — MacSymbol
キリル文字 — CP10007	朝鮮語 — KS_C_5601-1987	タイ語 — CP20838
キリル文字 — CP1251	朝鮮語 — KS_C_5861-1992	タイ語 — CP874
キリル文字 — CP20866	朝鮮語 — MacKorean	タイ語 — ISO 8859-11
キリル文字 — CP20880	ラテン文字 — CP10000	タイ語 — MacThai
キリル文字 — CP21025	ラテン文字 — CP10029	トルコ語 — CP10081
キリル文字 — CP21866	ラテン文字 — CP10082	トルコ語 — CP1254
キリル文字 — CP28595	ラテン文字 — CCSID 1047	トルコ語 — CP28599
キリル文字 — CP855	ラテン文字 — CP20261	トルコ語 — CP857
キリル文字 — CP866	ラテン文字 — CP20269	トルコ語 — ISO 8859-9
キリル文字 — ISO 8859-5	ラテン文字 — CP20273	トルコ語 — MacTurkish
キリル文字 — MacCyrillic	ラテン文字 — CP20277	Ukrainian — MacUkrainian
デーヴァナーガリー文字 — MacDevanagari	ラテン文字 — CP20278	Unicode — BMP
ギリシャ語 — CP10006	ラテン文字 — CP20280	Unicode — Java
ギリシャ語 — CP1253	ラテン文字 — CP20284	Unicode — UCS2
ギリシャ語 — CP20423	ラテン文字 — CP20285	Unicode — Unicode Big-endian
ギリシャ語 — CP28597	ラテン文字 — CP20297	Unicode — Unicode Little-endian
ギリシャ語 — CP737	ラテン文字 — CP20833	Unicode — Unicode11-UCS2
ギリシャ語 — CP869	ラテン文字 — CP20871	Unicode — Unicode11-UTF7
ギリシャ語 — ISO 8859-7	ラテン文字 — CP28591	Unicode — Unicode11-UTF8
ギリシャ語 — MacGreek	ラテン文字 — CP28593	Unicode — UTF7
グジャラート語 — MacGujarati	ラテン文字 — CP850	Unicode — UTF8
グルムキー文字 — CP10010	ラテン文字 — CP870	Unicode — UTF-EBCDIC
グルムキー文字 — MacGurmukhi	ラテン文字 — ISO 8859-1	ベトナム語 — CP1258

## 対応プラットフォーム

以下のプラットフォーム対応のSDKを提供します。その他のプラットフォームのサポートも、ご要望に応じ対応します。

AIX 6.1, PPC  
HP-UX 11i, IA64  
Linux CentOS 4.x/5.x, IA32/AMD64  
Linux Debian 5.x, IA32/AMD64  
Linux Red Hat 4.x/5.x, IA32/AMD64

Linux Ubuntu 10.x/11.x, IA32/AMD64  
MacOS  
Solaris 10, SPARC32/64, IA32/AMD64  
Windows XP/Vista/7, IA32/AMD64  
Windows Server 2003, 2008

詳細 [www.basistech.jp](http://www.basistech.jp) お問い合わせ [info2011@basistech.jp](mailto:info2011@basistech.jp) 電話 03-3511-2947

〒102-0084  
東京都千代田区二番町9-6

One Alewife Center  
Cambridge, MA 02140

2553 Dulles View Drive  
Herndon, VA 20171

171 Second Street  
San Francisco, CA 94105

