

Python

プログラミング言語 Python はバージョン 2.4 から JIS X 0213 を標準でサポートしています。codec で、EUC-JIS-2004, ISO-2022-JP-2004, Shift JIS-2004 が使用できます。

以前から、Python で日中韓の文字コードを扱う CJKCodecs というモジュールが配布されていましたが、それが Python 本体に取り込まれたということのようです。

コード例

例えば、下記のようなコードで、「EUC-JIS-2004 で符号化されたファイルを 1 行ずつ読み込み、各行の先頭にピュレット (箇条書きの頭に打つ、中ぐらいの大きさの点。面区点番号 1-03-32) を付加し、Shift JIS-2004 に変換して出力する」といった処理が可能です。

```
#!/usr/local/bin/python
import sys, codecs

dec = codecs.lookup('euc-jis-2004')[1]

f1 = open(sys.argv[1], 'r')
f2 = open(sys.argv[2], 'w')
while 1:
    s = f1.readline()
    if not s:
        break
    unicode_obj = dec(s)[0]
    unicode_obj = u'¥u2022' + unicode_obj
    f2.write(unicode_obj.encode('shift_jis-2004'))
f1.close
f2.close
```

参考

上のコードは以下のページを参考にしました。

- ・ [ファイル入出力](#)
- ・ [XML.com: Unicode Secrets](#)

関連項目

- ・ [PHP](#) - Web 開発に用いられる軽量言語
- ・ [Ruby](#) - Web 開発に用いられる軽量言語