

Re: kanjdic parser in Perl?

Re: kanjdic parser in Perl?

Source: <http://sci.tech-archive.net/Archive/sci.lang.japan/2005-08/msg00805.html>

- *From:* Ben Bullock <usenet@xxxxxxxxxxx>
 - *Date:* Tue, 16 Aug 2005 09:46:45 +0900
-

Gabor Farkas wrote:

Ben Bullock wrote:

"Gabor Farkas" <gabor@xxxxxxxxxxxxxxxx> wrote in message
[news:1f516\\$42ffa6b0\\$55d8898b\\$6438@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx](news:1f516$42ffa6b0$55d8898b$6438@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx)

Ben Bullock wrote:

David Alexander Ranvig wrote:

Ben Bullock <usenet@xxxxxxxxxxx> writes:

| Does anyone know of a parser for Jim Breen's kanjdic written
| Perl?

<URL: <http://search.cpan.org>> is a nice tool for finding all the
perl. Maybe you can use the module `Lingua::JP::Kanjidic` by Simon
Cozens?

Thanks for the tip. I had a

Re: kanjdic parser in Perl?

look at it, and it seems to need some work. Doesn't parse all the fields in the dictionary yet, unfortunately. I'll try editing it up a bit.

sorry, but what are you trying to achieve?
i mean, what advanced features you need from a kanjdic parser?

for me it seems that 2-3 lines of perl (mostly splits) would completely parse the kanjdic database for you...

Thanks for your input.

very funny ;))

the problem is that i can read perl relatively well, but i'm not good enough to be able to write it (i use python to solve text processing problems)

but i know that perl can split a string into by the whitespace, and that should be enough for your needs (if you know perl).