

## Re: examples

**Source:** <http://sci.tech-archive.net/Archive/sci.lang.japan/2004-10/1795.html>

---

*jim\_breen\_at\_hotmail.com*

**Date:** 10/27/04

Date: 27 Oct 2004 22:03:41 GMT

Max <max77@despammed.org> dixit:

>I would like to be able to use the edict dictionary looking for examples  
>sentences, just like in the online version but... I'd like to be able o do  
>it offline.

Yes, that would be great. Quite a few people have been asking for that.

>What do I need to install exactly? (if possible)

Software that no-one seems to have written yet.

>Thanks

Don't mention it.

I must admit to some mild surprise on this. It is nearly two years since I made the Tanaka corpus available for download, and apart from WWWJDIC and one other solitary WWW-based system no-one has touched it for a publicly-available tool (I know quite a few people have used it for other purposes.)

There were mutterings that it was too big, so I wrote an extractor that pulled out a subset about 35% the size of the full file, and which kept the same coverage of indexed words. This subset is updated daily and also freely available. As far as I know, no-one has used it.

If I could do more to encourage someone to add the examples to an "offline" dictionary program, I would. Perhaps people could send Alex Schonfeld (the JQuickTrans guy) pleading emails.

[I don't develop "offline" dictionary software any more, as I prefer to concentrate on the WWW server and the files. I am not going to make WWWJDIC "available" for local installation, as it's quite complex and locked into a \*nix/Apache environment, and I have no interest in doing installation documentation, etc. much less in providing support.]

--

Jim Breen <http://www.csse.monash.edu.au/~jwb/>  
Computer Science & Software Engineering,

sci.lang.japan: Re: examples

Monash University, VIC 3800, Australia  
\$B%8%`!&%V%j!<%s (B@ \$B%b%J%7%eBg3X (B