

# **PPA WordSegment**

## **Getting Started**

System Requirements  
Installing Your Software  
About this Guide

## **User Guide**

### **Chapter 1 – Using PPA Wordsegment Server Edition with the APIs**

1. HTTP API
2. Executable File API

### **Chapter 2 – Edit Dictionary**

1. Add Word
2. Edit Word
3. Delete Word

### **Chapter 3 – Technical Detail**

1. Performance
2. Specification
3. Frequency Asked Question

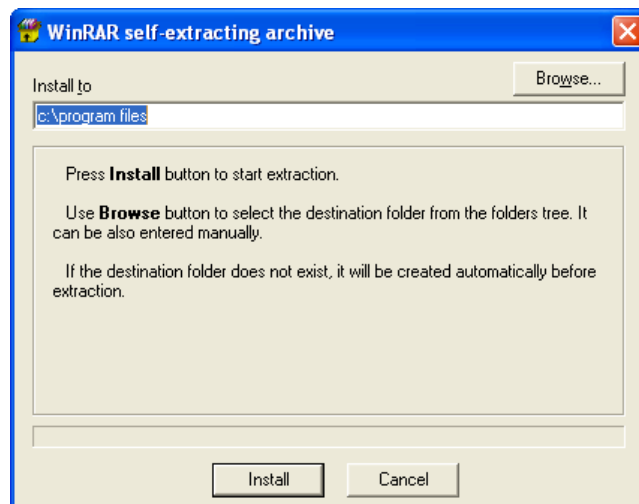
## Getting Started

### *System Requirements*

- 1.0GHz CPU or higher
- 100MB RAM or higher
- 100 MB Harddisk
- Windows 2003 Server / 2000 / XP

### *Installing Your Software*

Setup PPA Wordsegment in your computer.



### **About this Guide**

The PPA Wordsegment Server Edition Guide provides important information to help you start your work quickly with 2 main sections.

**Getting Started** – Provides basic information on the system requirements and how to install PPA Wordsegment Server Edition.

**User's Guide** – Provides step-by-step instructions about using PPA Wordsegment Server Edition.

## User's Guide

### Chapter 1 – Using PPA Wordsegment Server Edition with the APIs

PPA Wordsegment Server Edition has 2 APIs. You may choose only 1 API which you prefer.

#### 1. HTTP API

##### a. Initial step

run "c:\program files\PPA Wordsegment 2.xx\interface\wordsegment.exe".

##### b. Processing step

- i. Open socket to post 55555 with http protocol.
- ii. Send "GET" operation with one argument, "text", and the input text value to the socket.
- iii. Receive the output text from the socket.

Note: You can manually test this API by using your browser, for example, type the follow text in the address box.

"http://127.0.0.1:55555?text=สวัสดีครับทดสอบระบบตัดคำภาษาไทย"

##### c. Finale step

Exit PPA Wordsegment Server Edition by running "c:\program files\PPA Wordsegment 2.xx\interface\wordsegment\_exit.exe".

## 2. Executable File API

### a) Initial step

run "c:\program files\PPA Wordsegment 2.xx\interface\Wordsegment.exe".

### b) Processing step

- i. Save your input text in a text file.
- ii. Run the follow command with 2 parameters: the input file name and the output file name.

"c:\program files\PPA Wordsegment 2.xx\interface\wordsegment\_process.exe".

Example:

```
"c:\program files\PPA Wordsegment 2.xx  
\interface\wordsegment_process.exe input.txt output.txt".
```

Note: You can manually test this API by using the “Command Prompt”

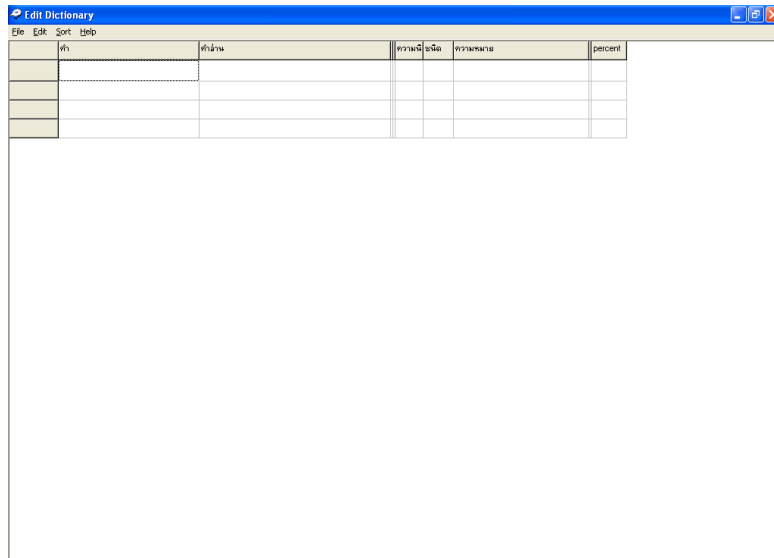
### c) Finale step

Exit PPA Wordsegment Server Edition by running "c:\program files\PPA Wordsegment 2.xx\interface\wordsegment\_exit.exe".

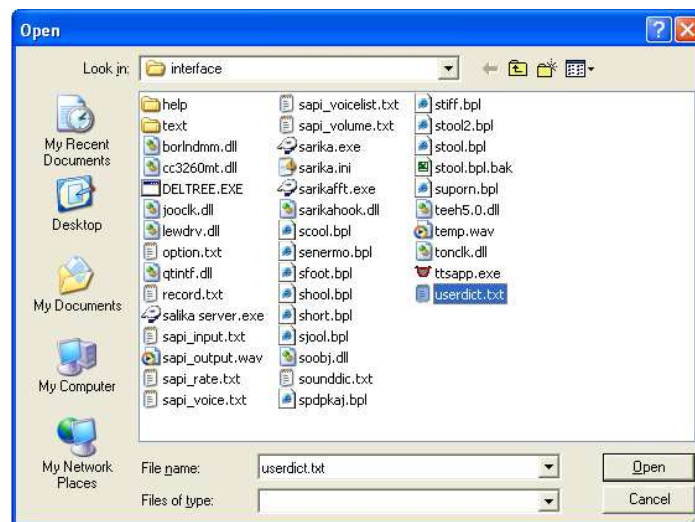
## Chapter 2 - Edit Dictionary

### 1. Add Word

- Click **“Menu”** button then go to **“File/Edit Dictionary”**. The Edit Dictionary window will appear on your screen.



- Go to **“File/Open”**, it will show open dialog.



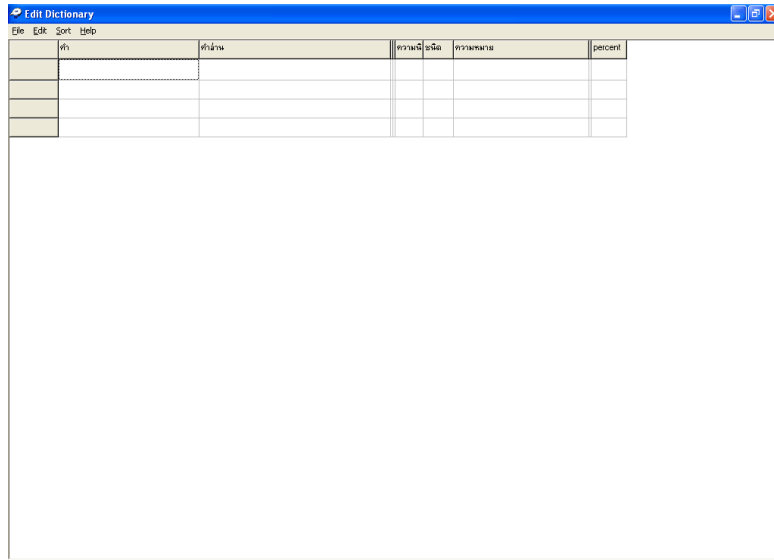
- Choose userdict.txt.
- Click at **“Open”** button, you will see the Check Dictionary window.

	คำ	คำอ่าน	ความถี่	ความหมาย	percent
1	กบฏ.	กบฏ เกษ มอ (bhal) กว มอ (bhal) กอ	1,000		
2					

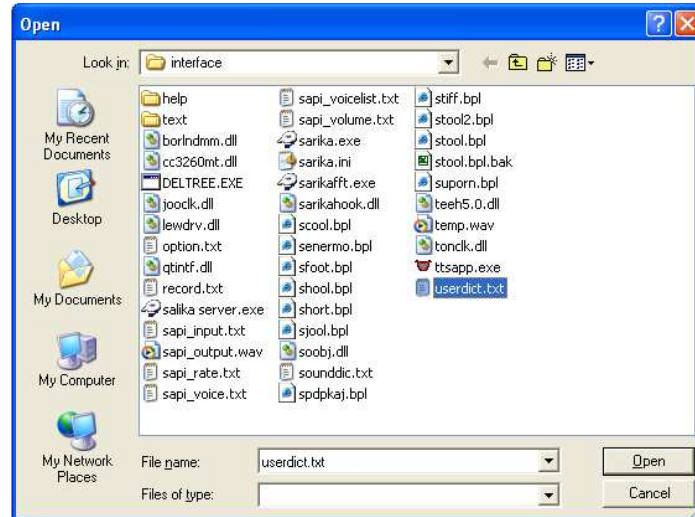
- Type your word in column “คำ”
- Type the reading symbol of the word in column “คำอ่าน”
- More information on how to type the reading symbol can be found in “**Help/เทคนิคการพิมพ์คำอ่าน**”
- After you add the new word, go to “**File/Save As**” you will see the “**Save As**” Dialog.
- Choose userdict.txt then click **Save**.
- You must restart your program after you made change to dictionary.

## 2. Edit Word

- Click “**Menu**” button then go to “**File/Edit Dictionary**”. The Edit Dictionary window will appear on your screen.



- Go to “**File/Open**”, it will show open dialog.

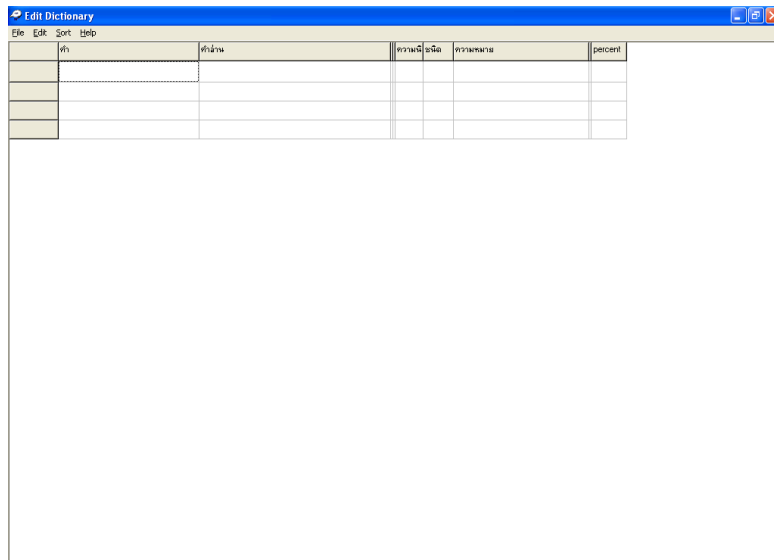


- Choose userdict.txt.
- Click at “**Open**” button, you will see the Check Dictionary window.

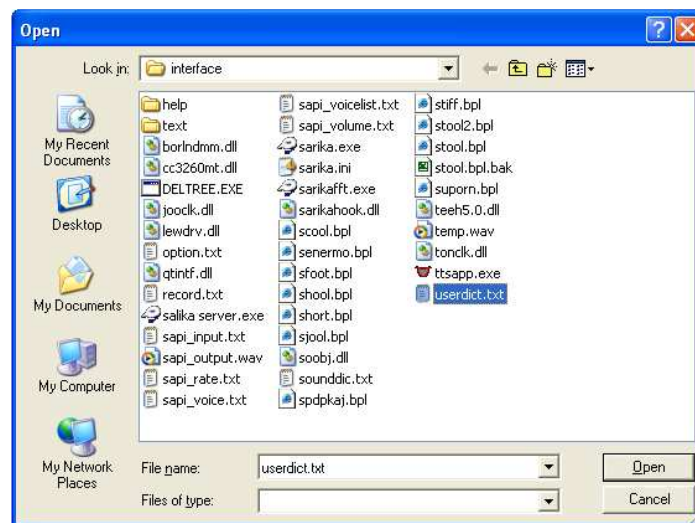


### 3. Delete Word

- Click “**Menu**” button then go to “**File/Edit Dictionary**”. The Edit Dictionary window will appear on your screen.



- Go to “**File/Open**”, it will show open dialog.



- Choose userdict.txt.
- Click at “**Open**” button, you will see the Check Dictionary window.





8. Input Format

- ANSI (default) / UTF-8 (default) / Unicode / Unicode big enddian
- Unlimited Input Text Length

9. Output Format

- A text file containing a series of words
- Customizable

**3. Frequency Asked Question**

- Can we host PPA Wordsegment on the same box as our solution?  
- Yes.