

PPA Soundex

Getting Started

System Requirements
Installing Your Software
About this Guide

User Guide

Chapter 1 – Using PPA Soundex 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

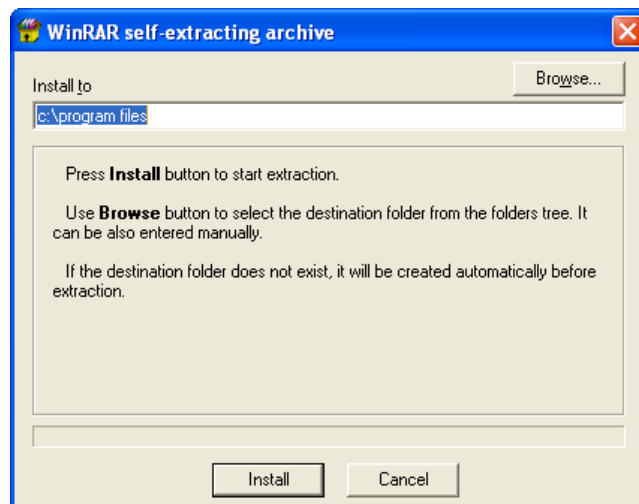
Getting Started

System Requirements

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

Installing Your Software

Setup PPA Soundex in your computer.



About this Guide

The PPA Soundex 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 Soundex Server Edition.

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

User's Guide

Chapter 1 – Using PPA Soundex Server Edition with the APIs

PPA Soundex 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 Soundex 2.xx\interface\Soundex.exe".

b. Processing step

- i. Open socket to post 44444 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://127.0.0.1:44444?text=สวัสดิ์คิรับทดสอบระบบตัดคำภาษาไทย”

c. Finale step

Exit PPA Soundex Server Edition by running "c:\program files\PPA Soundex 2.xx\interface\Soundex_exit.exe".

2. Executable File API

a) Initial step

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

Example:

"c:\program files\PPA Soundex 2.xx \interface\Soundex_process.exe
input.txt output.txt".

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

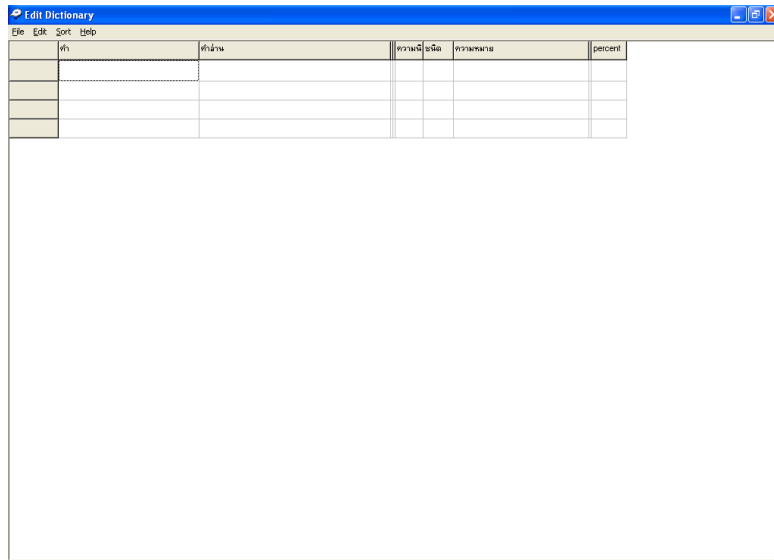
c) Finale step

Exit PPA Soundex Server Edition by running "c:\program files\PPA Soundex 2.xx\interface\Soundex_exit.exe".

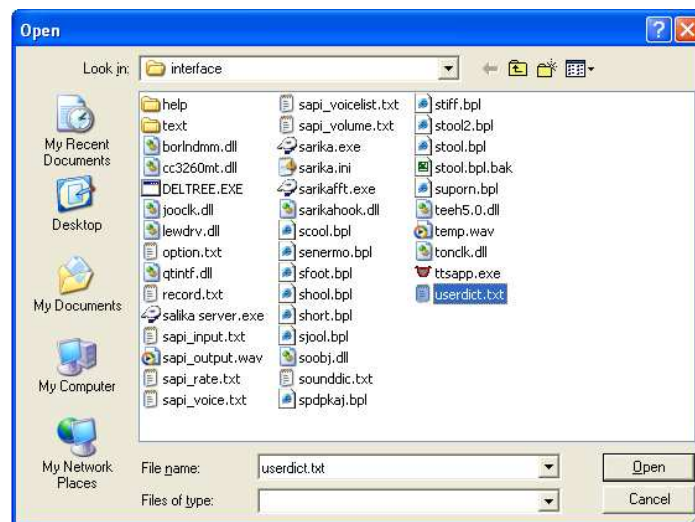
Chapter 2 - Edit Dictionary

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.



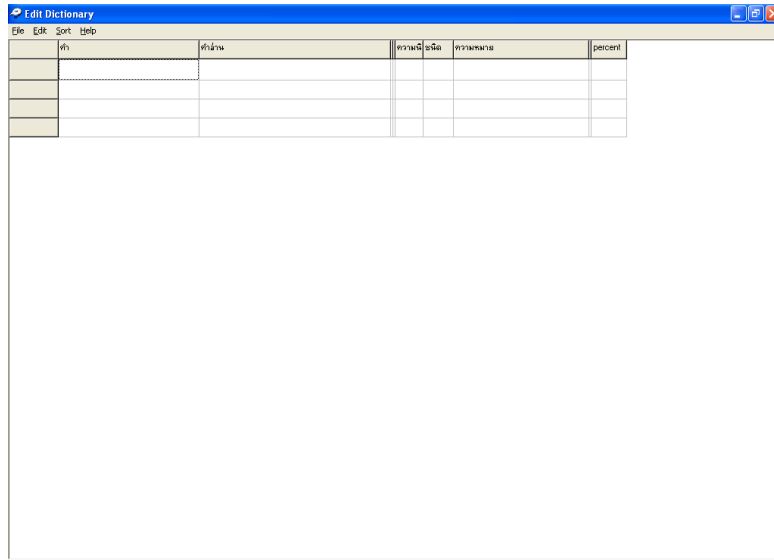
- Choose userdict.txt.
- Click at **“Open”** button, you will see the Check Dictionary window.

	คำ	คำอ่าน	ความถี่	ความหมาย	percent
1	กม.	กฺม. เกษ มกษ มช(กษ) ราช มช(กษ) ทร	1,000		
2					

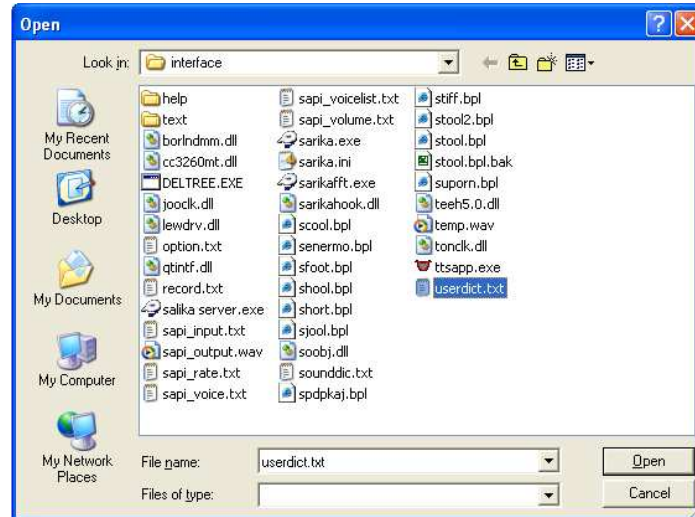
- Type your word in column “คำ”
- Type the reading symbol of the word in column “คำอ่าน”
- More information on how to type the reading symbol can be found in “**Help/เทคนิคการพิมพ์คำอ่าน**”
- After you add the new word, go to “**File/Save As**” you will see the “**Save As**” Dialog.
- Choose userdict.txt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



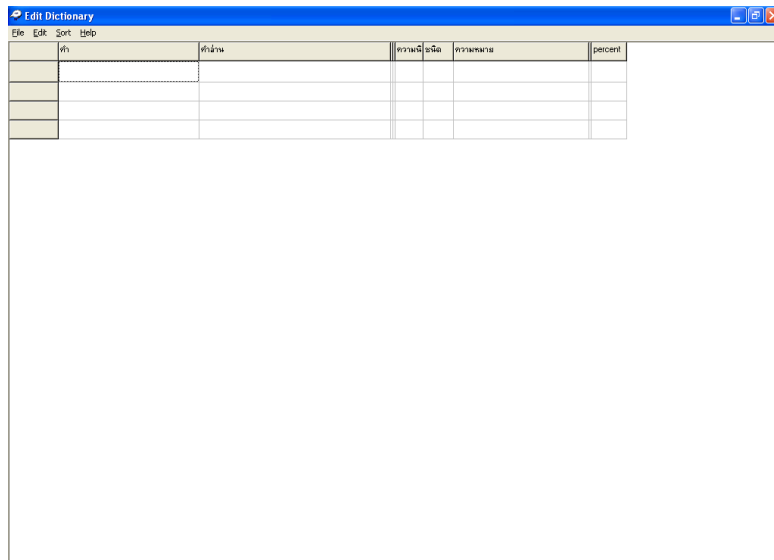
- Go to “**File/Open**”, it will show open dialog.



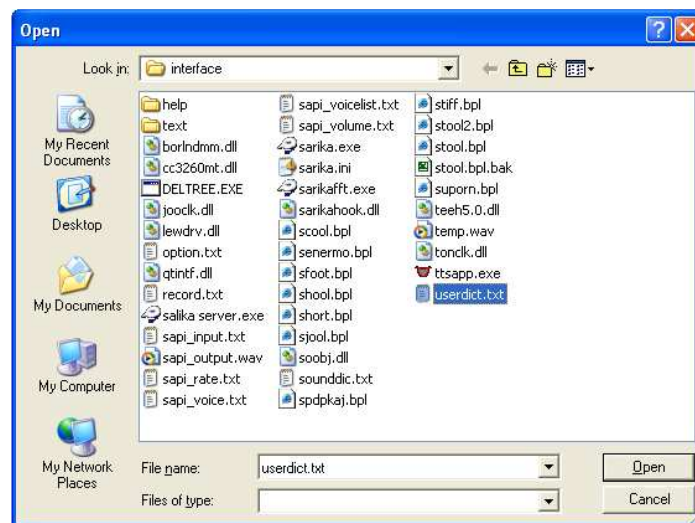
- Choose userdict.txt.
- Click at “**Open**” button, you will see the Check Dictionary window.

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.

id	word	definition	frequency	percent	percent
1	אמת.	אמת (hall) פנה אמת אמת	1,000		
2					

- Delete unwanted words.
- After you deleted unwanted words, 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.

Chapter 3 – Technical Detail

1. Performance

The rate of the converting time is about 28,000 characters/sec. or 6,600 words/sec. on Pentium IV 3.5GHz.

2. Specification

1. Language Supported
 - Thai
 - English (both Thai and English accent up to users)
 - Thai mixes English
 - Thai Slang, for examples, “มาเรี้ยวนี้ๆ” and “siam4you”
2. Accuracy
 - Thai known words: 99.387%
 - Thai unknown words: 96.881%
3. Known Dictionary Words
 - More than 60,000 words
4. Unknown, Specific, or New Words
 - Automatic Thai-Word Spelling System
 - Automatic English-Word Spelling System
5. Abbreviations and Acronyms
 - We use Thai abbreviation-and-acronym dictionary.
6. Numbers
 - Normal number, for example, "10000000000000000000000000000000"
 - Currency number, for example, "1,000,000,000 บาท"
 - Identification number, for example, "876238498239"
 - Real number, for example, "1000.123"
 - Negative number, for example, "-2349.324"
 - Number with commas, for example, "1,000,000"
 - Number with hyphens such as phone numbers, for example, "661-234-7777".
7. User Dictionary
 - Yes, we have a user dictionary for you. In order to add a new word, it's very simple. You don't need to use phonetic alphabet at all. Just use another word that pronounces like the word you want, for example, "พุดมพันธ์ พุด-ที-พันธ์"

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 Soundex on the same box as our solution?
- Yes.