

Introducing Pinyinbase

<https://github.com/pfffy/pinyinbase>

The Pfffy Authors



Chinese-English Dictionary Problems

- **Community-maintained dictionary is not family-safe.**
- Most dictionaries go for quantity over quality.
- Too many entries for the average user.
- File size is too large for mobile devices.
- Under-managed dictionaries raise your risk profile.

Pinyinbase Provides Solutions

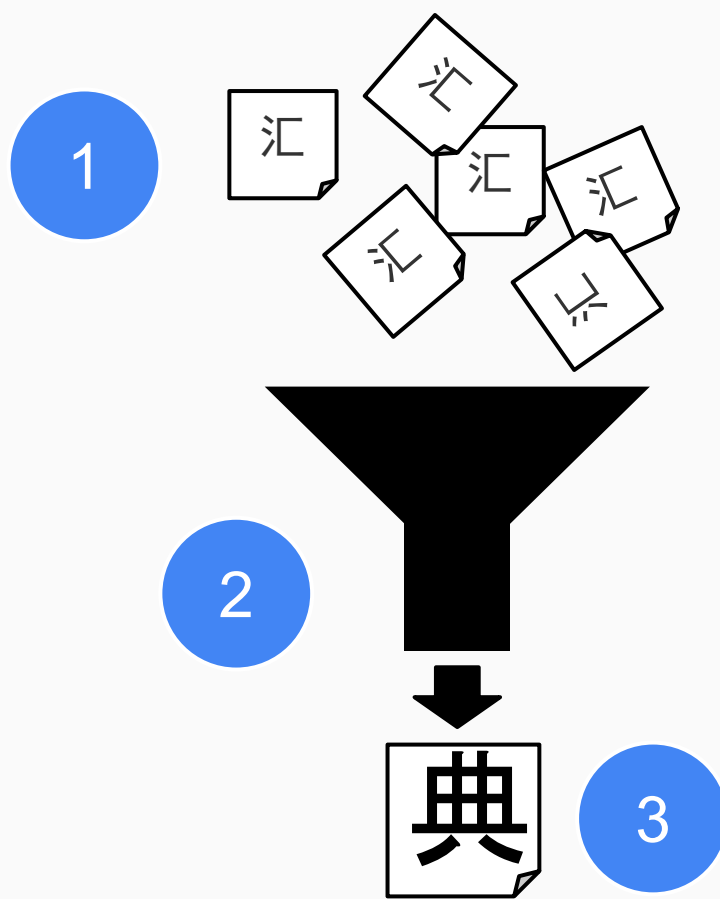
Pinyinbase is designed to:

- **Provide complete control of dictionary content.**
- Empower parents, teachers or clergy to filter content.
- Help developers comply with local laws.
- Facilitate rapid fixes with build-from-source tools.
- Reduce risk to your organization or brand.

Building a Pinyinbase File

Three steps to a Pinyinbase dictionary:

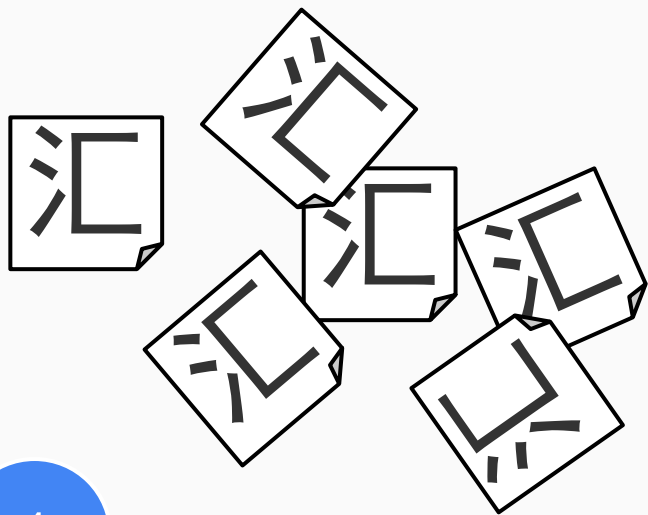
- Edit source files.
- Gather source files.
- Compile source files.



Building Your Pinyinbase File: 1. Edit; 2. Gather; 3. Compile.

Pinyinbase Source Files

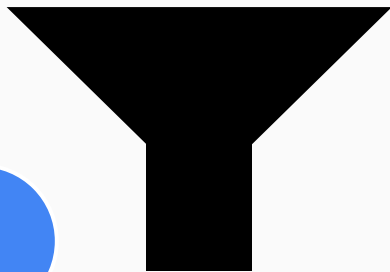
- Vocabulary lists grouped by category
- CEDICT-compatible dictionary format for legacy apps
- Modular design make adding, editing or omitting easy



1

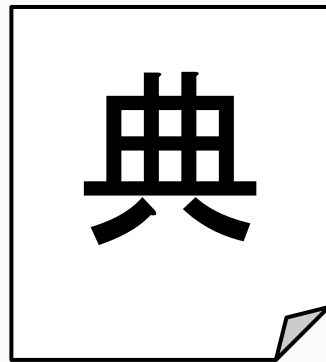
Pinyinpod Build Tools

- Gather Pinyinbase source files for compilation into a single JSON file
- Add additional hashing constructs to each entry for improved document search and retrieval



Pinyinbase Dictionary File

- Compiled JSON file contains:
 - basic t-s-p-d info
 - additional pinyin hashing constructs: **pmash**, **pbash**, **psmash**, **phash**
 - other indexing constructs: **dmash**, **src (source-line-key)**



Case Studies: App Developers

Dictionary app developers must understand:

- Google AdSense may suspend your account on grounds of prohibited content in your dictionary.
- Your app may be rejected or banned in local markets due to local laws on objectionable content.

Reduce your revenue at risk.

Thank you.

Learn more about Pinyinbase:

<https://github.com/pfffy/pinyinbase>

The Pfffy Authors