

基于 **Excel** 的英语词汇辅助学习系统
V1.0

使用说明书

目录

1	引言	1
1.1	编写目的	1
1.2	软件概述	1
2	运行环境与安装	1
2.1	硬件要求	1
2.2	软件环境	2
2.3	安装与目录结构	2
3	软件配置指南	2
3.1	Excel 列映射配置	2
3.2	网络请求配置	3
4	操作流程详解	3
4.1	第一步：准备数据源	3
4.2	第二步：启动程序	3
4.3	第三步：监控运行进度	4
4.4	第四步：自动备份机制	4
4.5	第五步：查看最终结果	5
5	特色功能演示	5
5.1	无损格式化技术	5
5.2	失败重试与日志记录	5
6	常见问题与故障排除	5

1 引言

1.1 编写目的

本手册旨在为“基于 Excel 的英语词汇辅助学习系统 V1.0”的用户提供详细的安装、配置及操作指南。本系统专为英语学习者、教育工作者及语言研究人员设计，旨在解决传统单词学习中人工查询释义繁琐、数据整理效率低下、复习资料格式混乱等痛点。通过阅读本手册，用户可以掌握软件的全流程操作，实现个性化词汇库的自动化构建。

1.2 软件概述

本系统是一款集成了数据采集、自然语言处理与文档自动化管理的桌面端软件。系统采用 Python 语言开发，深度整合了 Excel 文档处理技术与网络爬虫技术。其核心功能是读取用户指定的 Excel 单词表，自动连接在线词典接口（如 Youdao API），精准抓取单词的音标（Phonetic）、释义（Definition）及权威例句（Example），并将数据智能回填至原表格中。

此外，系统还具备以下企业级特性：

- 数据安全机制：**在每次写入操作前自动生成带时间戳的备份文件，彻底杜绝数据丢失风险。
- 断点续传能力：**自动识别已处理状态，网络中断或程序崩溃后重启即可接着上次进度继续运行。
- 无损格式化：**在美化排版的同时，智能保留用户原有的单元格颜色标记，尊重用户的学习笔记习惯。

2 运行环境与安装

2.1 硬件要求

本软件对硬件配置要求较低，普通办公电脑即可流畅运行：

- 处理器：**Intel Core i3 或同级别 AMD 处理器及以上。
- 内存：**建议 4GB RAM 及以上。
- 硬盘空间：**软件主体占用约 50MB，建议预留 500MB 以上空间用于存储备份文件。
- 网络连接：**必须连接互联网（推荐宽带或稳定的 Wi-Fi），用于实时获取词典数据。

2.2 软件环境

- 操作系统: Windows 10 / Windows 11 (64 位)。
- 运行依赖: Python 3.8 及以上版本。
- 必要运行库: OpenPyXL, Requests, BeautifulSoup4, PyYAML。
- 办公软件: Microsoft Excel 2016+ 或 WPS Office (用于查看结果)。

2.3 安装与目录结构

本软件为绿色免安装版, 用户只需将软件包解压至任意无中文路径的文件夹即可。标准的目录结构如下所示:

如图 ?? 所示:

- **config/**: 存放系统配置文件 `config.yaml`。
- **xlsx/**: 存放用户的数据源文件 `words.xlsx` 及自动生成的备份目录。
- **src/**: 存放软件的核心源代码及日志文件。

3 软件配置指南

在使用软件前, 用户可根据自身需求调整 `config/config.yaml` 文件。该设计使得软件具有极高的灵活性, 能够适应不同格式的 Excel 表格。

3.1 Excel 列映射配置

软件允许用户自定义单词、音标、释义等信息在 Excel 中的列位置。默认配置如下:

- **file_path**: 指定待处理的 Excel 文件相对路径。
- **columns**: 定义各数据项对应的列号 (A 列对应 1, B 列对应 2, 以此类推)。
 - **word**: 1 (A 列存放英文单词)
 - **phonetic**: 2 (B 列用于回填音标)
 - **definition**: 3 (C 列用于回填释义)
 - **example**: 4 (D 列用于回填例句)
 - **status**: 6 (F 列用于记录处理状态, 避免重复抓取)

3.2 网络请求配置

为了应对网络波动及防止被目标服务器封禁，软件内置了智能请求控制策略：

- **delay**: 请求间隔时间（默认 0.3 秒），适当的延时可以模拟人类操作。
- **retries**: 失败重试次数（默认 3 次），当网络超时，系统会自动重试。
- **user_agent**: 浏览器伪装标识，确保请求的合法性。

4 操作流程详解

4.1 第一步：准备数据源

打开 `xlsx` 目录下的 `words.xlsx` 文件。在第一列（或配置文件中指定的 Word 列）纵向输入需要学习的英文单词。其他列留空即可，软件会自动处理。

注意：数据准备完成后，请务必关闭 Excel 文件，否则软件无法获得写入权限。

4.2 第二步：启动程序

双击运行 `run.bat` 或在命令行中执行 `python src/main.py`。程序启动后，系统将按照下图所示的逻辑进行处理：

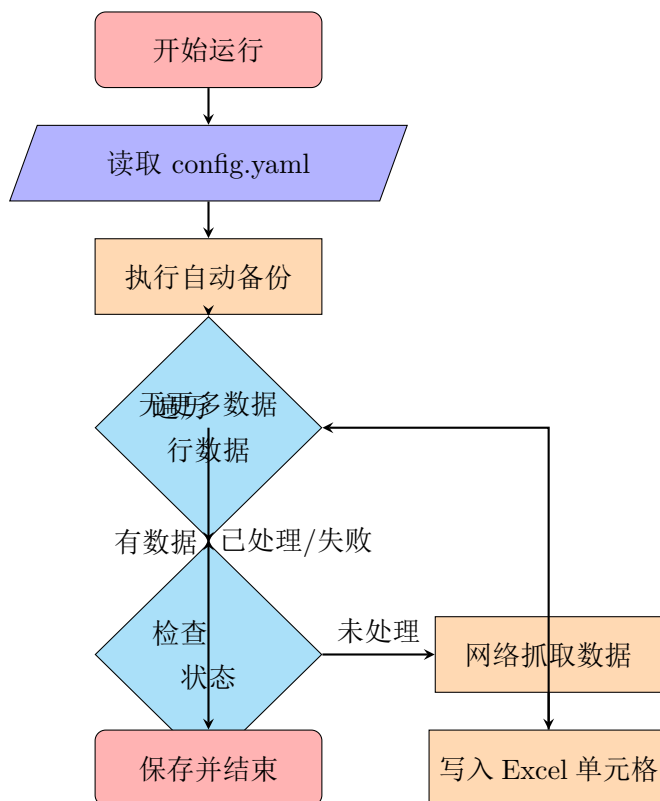


图 1: 软件核心处理逻辑流程图

4.3 第三步：监控运行进度

程序运行过程中，控制台会实时输出当前的进度日志。

- **INFO:** 已处理 [单词]: 表示该单词获取成功。
- **WARNING:** 部分成功: 表示可能缺失例句或音标，但核心释义已获取。
- **ERROR:** 处理失败: 表示单词拼写错误或词典中未收录。

4.4 第四步：自动备份机制

为了防止意外情况导致用户辛辛苦苦收集的单词丢失，软件在每次运行的瞬间，都会在 `xlsx/backup` 目录下生成一个以当前时间命名的副本文件。

4.5 第五步：查看最终结果

待控制台显示“处理完成”后，打开 `words.xlsx`。此时，表格中的音标、释义、例句列均已自动填充完毕，且所有单元格均已应用了统一的字体和自动换行格式。

5 特色功能演示

5.1 无损格式化技术

本系统的一大特色是“无损格式化”。在调整字体（如统一为 Arial 12 号）和对齐方式时，系统会通过深拷贝技术，完整保留用户手动设置的单元格背景色（如标记为红色的生词）和字体颜色。

5.2 失败重试与日志记录

对于因拼写错误或网络原因导致获取失败的单词，系统会自动将其记录在 `failed_words.txt` 文件中，并在 Excel 的“错误历史”列进行标记，方便用户后续进行人工核对。

6 常见问题与故障排除

表 1: 常见问题解决方案表

问题现象	可能原因	解决方案
Permission denied 报错	Excel 文件正在被其他程序占用	请关闭所有打开的 Excel 窗口后重试。
Connection Error	网络连接中断或不稳定	检查网络设置，增加 <code>config.yaml</code> 中的 <code>retries</code> 次数。
所有单词均失败	IP 地址可能被目标服务器暂时限制	增加 <code>config.yaml</code> 中的 <code>delay</code> 值（如设为 1.0），或稍后重试。
中文乱码	配置文件编码格式错误	确保 <code>config.yaml</code> 保存为 UTF-8 编码格式。