新版

雪人 CAT 操作手册



佛山市雪人计算机有限公司

网站：http://www.gcys.cn

邮箱：xueren@gcys.cn

电话：0757-83216691

13380230293

日期：2025 年 5 月 9 日

前言

佛山市雪人计算机有限公司是一家专注于软件开发及销售的高科技公司，公司成立于 2009 年，一直以开发和推广自主知识产权的软件为目标，致力于行业软件的开发。公司坚持走自主创新、稳步发展之路，以“做国内最好用的行业软件”为目标，以产品自身的品质、先进的技术和良好的服务取信于用户。公司的软件产品在实际应用中具有高效、稳定、易用、运行速度快等特点，深受用户好评。 公司不断追求卓越，坚信质量是企业发展的根本，不断地改进和完善产品，以满足用户的需求。同时积极引进先进的技术，不断创新，努力打造更具竞争力的产品。

雪人计算机辅助翻译软件(Snowman Computer Assisted Translation Software——缩写为：雪人 CAT)是一种充分利用计算机的超强计算能力、记忆能力、机器自学习能力和人的创造能力相结合的人机互动的辅助翻译软件，由译者把握翻译质量，计算机提供记忆翻译、机器自动翻译、AI 大语言模型翻译辅助，相同的句子永远不须要翻译第二遍，相似的句子参考着翻译，并可很好地保留原文排版风格。它能够辅助译者优质、高效、轻松地完成翻译工作，帮助企业及个人积累自己的翻译资产并充分利用已有的翻译资源、降低成本、提高工作效率。 雪人 CAT 自从 2009 年推出的第一个版本至今已经走过了 16 年头，雪人软件的稳定性、搜索结果的快速响应性能、操作的简单易用性、双语对齐建语料库的准确性、对 Windows 各个版本操作系统的适应性等都经得起广大用户的长期实践检验。 雪人 CAT 在 2023 年隆重推出新版本，这是全新研发的新版本软件，不是旧版本软件的升级改版。新版本在继承旧版本优点的基础，并在 CAT 技术中融入 AI 技术，使得 CAT 更加智能化、自动化和高效化。

“删繁就简，易学易用”是我们的软件设计理念，我们始终认为，一款好的软件应该具备简洁易用的特点。在软件设计和研发过程中，我们注重软件的易用性和易学性，在做到功能完备的前提下，尽可能地简化软件的界面和操作流程，使软件更加易于使用和上手。 “智享高效，引领未来”是我们对工作的态度和目标，我们相信，只有不断提高自己的技能和知识，不断探索新的技术和方法，才能在竞争激烈的软件行业中立于不败之地。 “勇于创新，追求卓越”是我们的核心价值观，我们会对计算机辅助翻译技术不断地改进和创新，特别是在 CAT 技术中融入 AI 技术，使得 CAT 技术更加智能化、自动化和高效化；在创新的过程中，我们注重细节和质量，努力把每一个细节都做到极致。 “恪守承诺，诚信为本”是我们的一直坚守的原则，我们坚信，只有恪守承诺、诚信为本，才能赢得客户的信任和信赖。我们相信，我们的产品能够为用户带来真正的价值！

为了让更多的人了解雪人的新版软件，目前开放免费申请使用，请登录雪人官网 http://www.gcys.cn 下载。

目录

封面 0
前言 1
目录 2
新版软件特点 6
软件持续更新记录 9
第一章 登录雪人 CAT 12
第一节 下载软件 12
第二节 登录软件 12
第三节 修改密码 13
第四节 创建桌面图标 13
第五节 关联项目文件 14
第二章 快速入门 15
第一节 新建翻译项目文件 15
第二节 外挂记忆库和词典 19
第三节 快速全文预翻译 24
第四节 导出译文及双语对照 26
第五节 常用热键 31
第六节 扩展支持更多的语言 32
第三章 在线机器翻译引擎 39
第一节 机器翻译引擎的设置 39
第二节 逐句机器翻译提示 42
第三节 同时显示多个机译引擎结果 43
第四节 机器翻译时进行术语替换 45
第五节 按鼠标划选的内容进行机译 49
第六节 申请机器翻译 API 接口 51
第四章 本地机器翻译引擎 62
第一节 支持本地机器翻译 62
第二节 本地化机器翻译模型定制与训练服务 67
第五章 预翻译 68
第一节 全文预翻译 68
第二节 参照翻译 73
第三节 记忆库匹配分析 77
第四节 文档相似度分析 79
第六章 翻译各种格式的文档 81

第一节 翻译 TXT 格式的文档 81
第二节 翻译 WORD 格式的文档 88
第三节 样式码 90
第四节 翻译 EXCEL 格式的文档 94
第五节 翻译 PPT 格式的文档 98
第六节 剪贴板翻译 100
第七章 支持复杂的项目文件结构 101
第一节 单层项目文件结构 101
第二节 多层项目文件结构 102
第三节 项目文件的整理 104
第八章 词典和术语 107
第一节 定义术语和定义新词 107
第二节 词典的导入导出 109
第三节 词典的外挂 113
第四节 整批设置术语/词语 114
第五节 在线词典和本地词典 115
第六节 导出项目所使用的术语 118
第九章 记忆库 120
第一节 记忆库的导入导出 120
第二节 记忆库的外挂 122
第三节 记忆库的上下文精确匹配 123
第四节 直接引用翻译的项目文件作记忆库 125
第五节 项目文件的记忆库导出 128
第六节 记忆库搜索 129
第七节 片段搜索 133
第八节 将片段搜索的结果排序 134
第十章 双语对齐建记忆库 135
第一节 创建多层结构的双语项目 135
第二节 各种文档格式的对齐 137
第三节 对齐后的编辑 143
第四节 双语对齐记忆库的导出 146
第五节 双语对齐二次拆分对齐 154
第十一章 查找和替换 159
第一节 词语的整批查找和替换 159
第二节 灵活准确的查找 172
第十二章 协同翻译 173
第一节 客户端如何连接服务器 174

第二节 组长给项目组添加词典及记忆库 180
第三节 译员连接项目组进行翻译 183
第四节 审校连接项目组进行审核 187
第五节 团队成员的权限 195
第六节 服务器项目组中数据的导出 197
第十三章 在雪人 CAT 中接入 AI 203
第一节 AI+MTPE 人机交互翻译模式 203
第二节 注册 AI 账号 204
第三节 面向翻译的提示语(Prompt)工程 210
第四节 AI 助手 212
第五节 AI 大模型的 API 调用 229
第六节 用 AI 大模型进行全文批量翻译 243
第七节 依据用户术语库，让 AI 译得更准确 249
第八节 AI 大模型 API 的申请 257
一、智谱 AI 大模型 API 申请 257
二、阿里通义千问 AI 大模型 API 申请 259
三、百度文心一言 AI 大模型 API 申请 263
四、Kimi AI 大模型 API 申请 267
五、ChatGPT AI 大模型 API 申请 270
六、接入腾讯混元 AI 模型的 API 272
七、DeepSeek 深度求索 AI 大模型 API 申请 274
第九节 在百度平台调用 DeepSeek 277
第十节 云端 AI 通用模型接口 282
第十一节 Ollama 接口 287
第十二节 本地 AI 模型的部署 288
第十三节 升级为 VIP 用户 291
第十五章 审校与 AI 批量审校 292
第一节 审校 292
第二节 AI 逐句审校 296
第三节 AI 批量审校 300
第四节 导出修订项目 302
第五节 外部审校 309
第十六章 软件设置与界面调整 314
第一节 调整界面字体和大小 314
第二节 自动存盘 317
第三节 精确匹配时自动填写译文 319
第四节 保存界面设置 320

第五节 窗口的放大、缩小 320
第十七章 其他 323
第一节 导入、导出双语文件 323
第二节 导出未翻译的句子 326
第三节 正则表达式词典 329
第四节 自定义断句规则 331
第五节 词频统计 334
第六节 质量检查 338
第七节 全角/半角转换 339
第八节 自动填充句尾空格 340
第九节 字数统计及显示翻译进度 341
第十节 常用符号工具条 342
第十一节 修改原文 343
第十二节 词语输入自动提示 345
第十三节 缩略语表 347
第十四节 雪人 CAT 账号级别 348

新版软件特点

雪人 CAT 自从 2009 年推出的第一个版本至今已经走过了 16 年头，软件的稳定性、搜索结果的快速响应性能、操作的简单易用性、双语对齐的准确性、对 Windows 各个版本操作系统的适应性等都经得起广大用户的长期实践检验。但是，由于雪人 CAT 旧版软件采用的是嵌入 OFFICE 的方式，OFFICE 的版本众多且更新频繁，造成了对有些用户的 OFFICE 的兼容性问题；由于对 OFFICE 的依赖和受限于 OFFICE 处理速度，导致雪人旧版本在原文导入、译文的导出速度方面表现不佳。旧版软件字体、界面在新的高清屏幕下也会显得不够清晰。

新版本主要改进如下：

1. 新版软件彻底解决了对 OFFICE 版本的依赖，用户机器上即使没有安装 OFFICE 版本，也不会影响雪人 CAT 的预览原文、译文，而且导入、导出速度变得飞快，支持批量导入（即：一次性导入多个需要翻译的文档）。
2. 对导入的文本格式的文件，也保留了原有的格式。
3. 新版软件解决了对高清屏幕的支持，界面更清晰。
4. 在左侧的项目文件中，支持分层管理的多层文件夹的项目目录结构，项目文件分多少层不作限制，每个文件夹下面存放多少个翻译文档不作限制，对导入的每一份文件大小亦不作限制，方便用户对大型翻译项目的分层次的项目管理。
5. 双语对齐使用了更新的技术，使对齐的结果更加准确，对齐界面的编辑也更加灵活，自动对齐后还可以进行二次分区域的局部分块对齐。也可以导入时不自动对齐，导入后，分章节、分块对齐，这样则对单次导入的句子数量没有限制，更加方便对大型双语语料的对齐。
6. 加强了对机器翻译的支持，实现机助人译。支持接入 DeepL 翻译、谷歌翻译、百度翻译、阿里翻译、腾讯翻译、有道翻译、同花顺翻译、小牛翻译等多个机器翻译引擎，多个机器翻译引擎的结果可以同时显示，方便用户比较各大机器翻译引擎的译文质量，择优选择最佳译文。
7. 利用用户自定义的本地术语库与机器翻译系统相结合，对术语的翻译，强制机器翻译系统按照用户所定义的词义进行翻译，确保术语翻译的准确性和一致性，同时提升译文的翻译质量。
8. 除了支持传统的记忆库模糊匹配和 100%的精确匹配外，新版还支持上文匹配、下文匹配、上下文匹配的精确匹配。
9. 翻译的项目文件支持多层次的文件目录结构及方便的移动、拖放、复制操作，这个功能使得对翻译后的文档归类、整理也变得十分方便。

10. 支持直接将翻译的项目文件（.SPR ）作为记忆库来引用，这样可以保留原始篇章语境信息，优先匹配选用上下文都能精确匹配的句子，并且保留样式，还可以省去记忆库的导入、导出的工作。
11. 支持全文预翻译。优先选用术语库、记忆库中能精确匹配的句子进行翻译，剩下的交给机器自动翻译，实现快速翻译全文。机器翻译时可以结合用户本地术语库，使得翻译结果更加准确。
12. 支持 Bing、有道等多个在线词典，还支持用户自己添加本地词典。
13. 支持参照翻译，即参考现有已经翻译好的文件翻译另一个文件，同时保留参考文件的样式，即以参考文件为模板翻译新文件，复制参考文件的排版样式。对于同类型的技术文档类文件的升级改版翻译非常实用。
14. 将 CAT 技术和 AI 技术深度融合。支持 AI 大语言模型有：DeepSeek、ERNIE(文心一言)、通义千问、OpenAI(ChatGPT)、智谱、Moonshot(Kimi)、百川、豆包、腾讯混元、讯飞星火、商汤商量、360 智脑。机器翻译模型有：百度翻译、DeepL 翻译。还支持本地部署 AI 大语言模型。支持用 AI 进行全文批量翻译。
15. 支持多种本地机器翻译引擎，为用户提供了额外的翻译解决方案。用户无需额外注册账号及申请 API 密钥，也无需担心翻译数据外泄，因为所有翻译数据均在用户的本地机器上完成。具体有：Facebook 的 M2M-100 本地机翻模型（支持 100 种语言间互译）、谷歌 t5 本地机翻模型、 OPUS 本地机翻模型、Marian 本地机翻模型、以及雪人计算机公司训练的基础翻译模型（“scat_ec_20240812-q8.model”）。
16. 可以为有需要的企业用户或个人用户提供定制本地机器翻译引擎的训练服务。
17. 依据用户术语库，让 AI 译得更准确：让 AI 依据用户术语库中的术语词义，进行语义匹配、术语替换和优化翻译逻辑，确保翻译过程中专业术语的准确无误，可以进一步提升 AI 译文的翻译质量。
18. 使用 AI 时，用户可以自行编写灵活多样的提示语。和普通的机器翻译工具，如谷歌翻译、DeepL 翻译等相比，AI 大语言模型翻译所提供的机译服务超越了普通机器翻译，它可以通过构建合理、清晰的提示语来提升翻译的准确度，实现人机交互式翻译。在雪人中，AI 提示语的编写是非常灵活的，用户可以根据自己具体的翻译文本需求，编写不同的 AI 提示语，使 AI 输出更加符合要求的译文，而不是由 CAT 软件固化几种风格的提示语供用户选择。当然，雪人也会编写一些提示语样例供用户参考及修改。
19. 支持审校模式。支持三种项目文件：翻译项目（*.spr）、审校项目（*.spp）和修订项目（*.spa）。无论采用人工审校方式，还是基于 AI 大语言模型的审校方式，均能确保流程清晰、责任明确，便于追溯。

20. 支持 AI 批量审校。AI 大语言模型对翻译后的译文进行校对具有显著优势，它能够快速识别并纠正语法错误、拼写错误和用词不当，确保译文的准确性和流畅性。
21. 由于 DeepSeek 官方服务器繁忙。雪人支持在阿里平台调用 DeepSeek，还支持在百度智能云千帆平台上调用 DeepSeek。使用成本更低，响应速度更快。
22. 支持云端 AI 通用模型接口，只要是采用与 OpenAI（ChatGPT）兼容的 API 接口，都可以利用通用模型接口来调用它。
23. 支持 Ollama 接口，方便企业部署私有化 AI 大型语言模型。企业可自行部署 DeepSeek 到本地，确保翻译数据安全保密。
24. 原生支持的语种有：简体中文、英语、俄语、日语、法语、西班牙语、韩语、德语；支持这 8 个语种之间的任意互译，雪人会根据不同语种的特点进行优化及支持词形变化。还通过用户自定义语种的扩展方式支持另外 26 个语言，分别是：繁体中文、粤语、阿拉伯语、保加利亚语、捷克语、丹麦语、希腊语、爱沙尼亚语、芬兰语、匈牙利语、意大利语、立陶宛语、拉脱维亚语、挪威语、荷兰语、波兰语、葡萄牙语、罗马尼亚语、印度尼西亚语、斯洛伐克语、斯洛文尼亚语、瑞典语、土耳其语、泰语、乌克兰语、越南语。这些语言都可以使用传统机器翻译引擎和 AI 大语言模型来翻译它们，也可以用本地多语翻译模型（如 m2m100）进行翻译。

如果您觉得软件对您的工作有帮助，请加入雪人 VIP，以支持软件做得更好！

软件持续更新记录

1. 推出雪人 CAT2023 内部测试版本_20230331

新版本支持高清界面

2. 加入对机器翻译的支持_20230412

支持百度、腾讯、阿里、有道、同花顺、谷歌等多个机器翻译引擎同时显示结果，择优选用。

3. 在无 OFFICE 环境下翻译 WORD 文件_20230614

彻底解决了对 OFFICE 的依赖，可以批量导入 WORD 文件，而且导入导出速度也变得飞快!

4. 支持全文快速预翻译及支持翻译项目文件作记忆库被引用_20230705

将用户记忆库、用户术语库、机器翻译三者高效结合，实现全文快速自动翻译！支持直接将翻译的项目文件（.SPR ）作为记忆库来引用，这样可以保留原始篇章语境信息，优先匹配选用上下文都能精确匹配的句子，并且保留样式，还可以省去记忆库的导入、导出的工作。

5. 在无 OFFICE 环境下翻译 EXCEL 文件_20230715

不需要 OFFICE 的支持，也可以预览、翻译 EXCEL 文件，还可以批量导入 EXCEL 文件，而且导入导出速度很快!

6. 在无 OFFICE 环境下翻译 PPT 文件_20230811

不需要 OFFICE 的支持，也可以预览、翻译 PPT 文件，还可以批量导入 PPT 文件，而且导入导出速度很快!

7. 加入对日语版的支持和双语对齐的二次拆分对齐_20230825

在简体中文和英语的基础上，加入对日语版的支持，可以实现这三个语种之间的任意互译，即中英、英中、中日、日中、英日、日英之间的互译，软件会根据不同的语种进行优化和识别词形变化。双语对齐可以先按段落对齐后，再拆分段落里的句子进行二次拆分对齐，提高自动对齐的准确性。

8. 加入对俄语版的支持和在线词典、本地词典、参照翻译_20230908

加入俄语版，可以在中文、英语、日语、俄语之间任意互译。支持 Bing、有道等多个在线词典，还支持用户自己添加本地词典。支持参照翻译，即参考现有已经翻译好的文件翻译另一个文件，同时保留参考文件的样式，即以参考文件为模板翻译新文件，复制参考文件的排版样式。

9. 加入法语版和西班牙语版，支持词频统计_20230922

雪人目前支持的语种有：简体中文、英语、俄语、日语、法语、西班牙语；支持这几个语种之间的任意互译，雪人会根据不同语种的特点进行优化及支持词形变化。

10.支持导出双语对照的译文，支持质量检查_202301027

支持导出段落对照格式的译文。支持对译文进行术语统一、拼写、样式码、漏译、一句多译、数字等非译成分的检查。

11.支持协同翻译_20231228

发布雪人 CAT 协同服务器软件，客户端软件也做了对协同翻译的支持。

12.支持批量查找替换和记忆库搜索_20240126

支持词语的整批查找替换、支持记忆库搜索、自动存盘、导出未翻译的句子、精确匹配时自动填写译文、全角/半角转换、自动填充句尾空格。

13. 增加了 AI 应用的支持_20240423

支持 AI 大语言模型有：ERNIE(文心一言)、通义千问、OpenAI(ChatGPT)、智谱、Moonshot(Kimi)、百川、豆包、腾讯混元、讯飞星火、商汤商量、360 智脑。机器翻译模型有：百度翻译、DeepL 翻译。还支持本地部署 AI 大语言模型。

14. 用 AI 进行全文批量翻译_20240531

AI 应用中增加全文批量翻译功能，实现全文快速预翻译。增加对韩语的支持、导出项目所使用的术语、记忆库匹配分析、文档相似度分析、词频统计中增加导入导出词表功能。

15. 依据用户术语库让 AI 译得更准确及支持本地机器翻译引擎_20240826

一、依据用户术语库，让 AI 译得更准确：让 AI 依据用户术语库中的术语词义，进行语义匹配、术语替换和优化翻译逻辑，确保翻译过程中专业术语的准确无误，可以进一步提升 AI 译文的翻译质量。二、支持多种本地机器翻译引擎，用户无需额外注册账号及申请 API 密钥，也无需担心翻译数据外泄。具体有：Facebook 的 M2M-100 本地机翻模型（支持 100 种语言间互译）、谷歌 t5 本地机翻模型、

OPUS 本地机翻模型、Marian 本地机翻模型、以及雪人计算机公司训练的基础翻译模型（“scat_ec_20240812-q8.model”）。雪人可以为有需要的企业用户或个人用户提供定制本地机器翻译引擎的训练服务。新版本会不断的更新、加入新的功能，使软件越来越完善，为用户带来更加便捷、高效的翻译体验!

16. 审校和 AI 批量审校_20250225

支持三种项目文件：翻译项目（*.spr）、审校项目（*.spp）和修订项目（*.spa）。无论采用人工审校方式，还是基于 AI 大语言模型的审校方式，均能确保流程清晰、责任明确，便于追溯。支持 AI 批量审校。由于 DeepSeek 官方服务器繁忙。支持在阿里平台调用 DeepSeek 和在百度智能云千帆平台上调用 DeepSeek。支持云端 AI 通用模型接口，只要是采用与 OpenAI（ChatGPT）兼容的 API 接口，都可以利用通用模型接口来调用它。支持 Ollama 接口，方便企业部署私有化 AI 大型语言模型。例如，企业可以自行部署 DeepSeek 到本地。

17. 扩展支持更多的语言_20250509

雪人目前原生支持的语种有：简体中文、英语、俄语、日语、法语、西班牙语、韩语、德语；支持这 8 个语种之间的任意互译，雪人会根据不同语种的特点进行优化及支持词形变化。还通过用户自定义语种的扩展方式支持另外 26 个语言，分别是：繁体中文、粤语、阿拉伯语、保加利亚语、捷克语、丹麦语、希腊语、爱沙尼亚语、芬兰语、匈牙利语、意大利语、立陶宛语、拉脱维亚语、挪威语、荷兰语、波兰语、葡萄牙语、罗马尼亚语、印度尼西亚语、斯洛伐克语、斯洛文尼亚语、瑞典语、土耳其语、泰语、乌克兰语、越南语。这些语言都可以使用传统机器翻译引擎和 AI 大语言模型来翻译它们，也可以用本地多语翻译模型（如 m2m100）进行翻译。

更多新的功能的，待续……

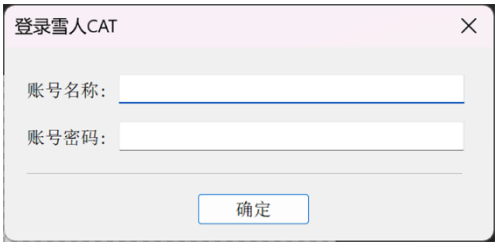
第一章 登录雪人 CAT

第一节 下载软件

登录雪人官网 http://www.gcys.cn 下载新版软件，或者加入雪人 QQ 群（群号 126371112），去 “群共享”中下载软件，下载的文件大小约一百多兆，解压后，操作手册及“雪人 CAT 账号申请表”在其中。填写好“雪人 CAT 账号申请表”后，将它发到 xueren@gcys.cn 邮箱，经雪人公司审核通过后，会以电子邮件方式回复登录密码。

第二节 登录软件

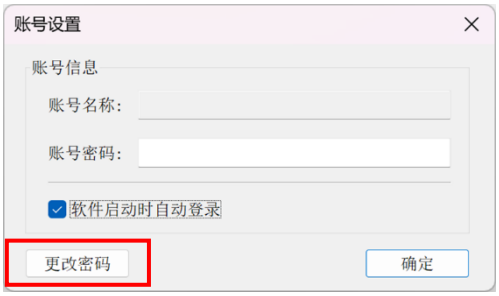
雪人 CAT 是免安装的软件，支持的操作系统有：Win11、Win10，全部为 64 位的操作系统。下载的软件解压后，双击 scat 的应用程序启动软件，将会弹出要求输入“账号名称”及“账号密码”的窗口，输入账号名称及密码后，即可登录雪人 CAT。如果没有连接网络，会出现“无法连接雪人 CAT 服务器”出错窗口，此时请检查网络连接是否正常。如下图：

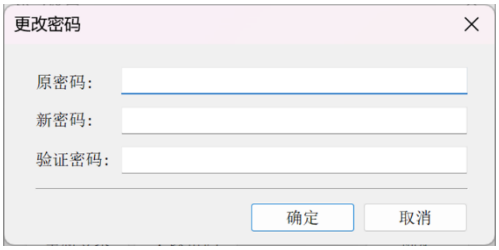




第三节 修改密码

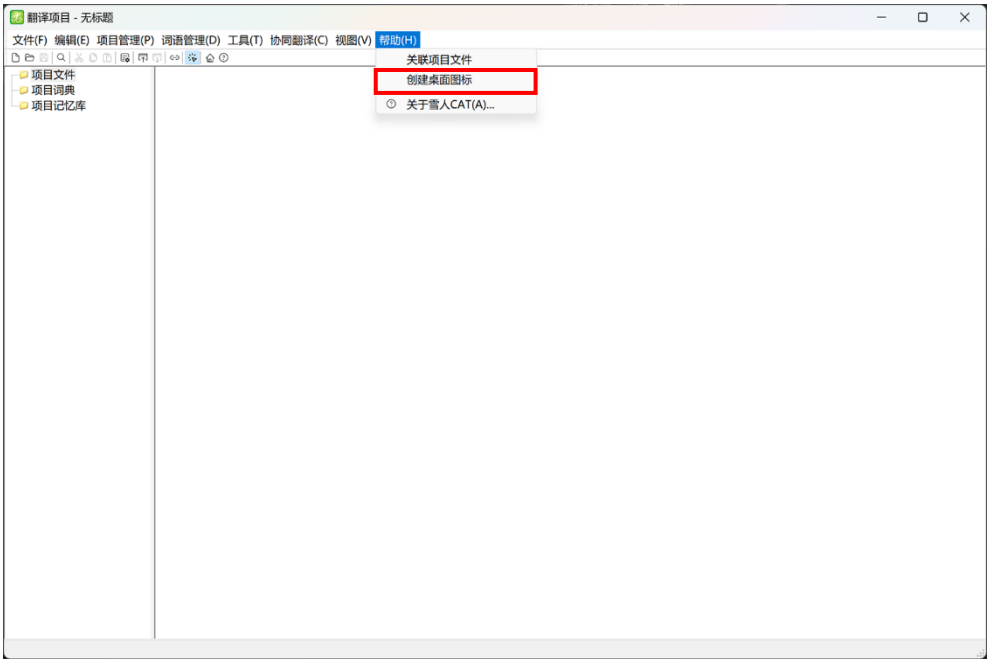
第一次登录雪人软件后，请及时修改密码。修改密码方法：点击**【工具】→【账号设置】**菜单，会弹窗更改密码按钮，如下图：





第四节 创建桌面图标

点击**【帮助】➔【创建桌面图标】**菜单，可以在桌面创建“雪人 CAT”启动图标，方便以后从桌面启动雪人 CAT。如果软件升级，将下载的软件解压到旧的文件夹覆盖旧文件即可。

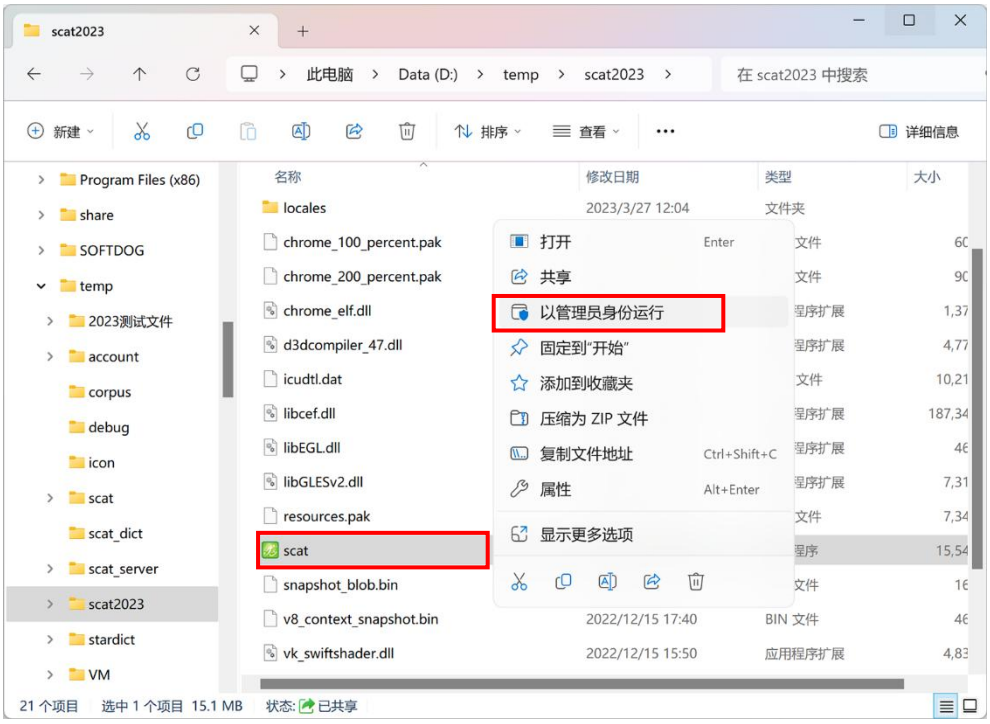


第五节 关联项目文件

设置好关联的项目文件，以后双击某个项目文件，可以直接打开。点击**【帮助】➔【关联项目文件】**菜单，可能会弹出下图的“无法管理项目文件”窗口。进行这个操作，需要 WINDOS 的系统管理身份。



退出软件后，以管理员的身份运行启动雪人 CAT 软件，再次关联。具体操作，打开雪人 CAT 软件的存放的文件夹，如下图：右键点击 scat 的应用程序，会弹出右键菜单，点击**【右键菜单】➔【以管理员身份运行】**功能可以启动雪人 CAT 软件。

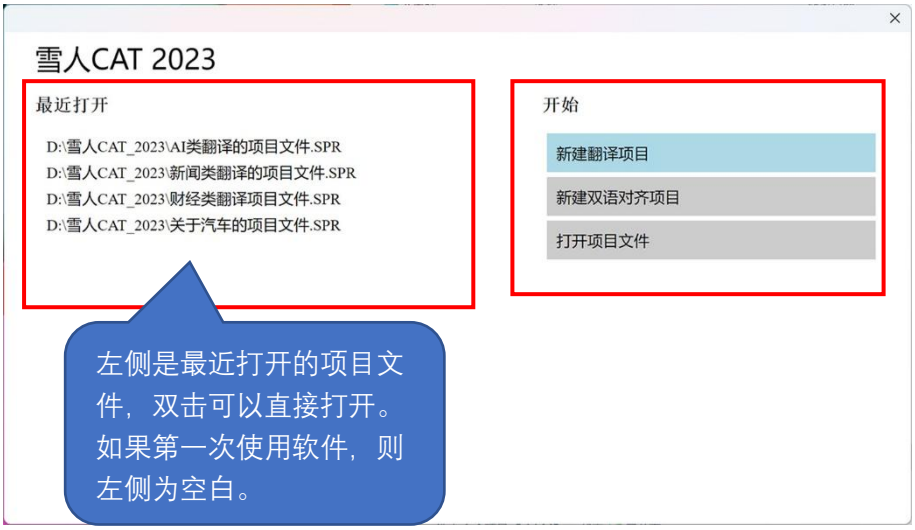


启动软件后，再点击**【帮助】➔【关联项目文件】**菜单，会弹出成功关联项目文件窗口。

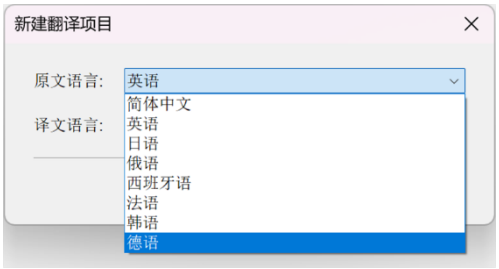
第二章 快速入门

第一节 新建翻译项目文件

正常登录雪人 CAT 后，将会出现如下界面：



左侧是最近打开的项目文件，双击可以直接打开。如果第一次使用软件，则左侧为空白。右侧是新建翻译项目、新建双语对齐项目、打开项目文件三个导航功能，点击新建翻译项目，在弹出的窗口中，选择原文语言和译文语言。新版不再像旧版软件那样按语种划分软件，而是所有语种集中在同一个软件中，用户在新建翻译项目时确定语言方向，点击“原文语言”下拉框和“译文语言”，选择需要的语种，在雪人支持的语种之间，“原文语言”和“译文语言”可以任意选择互译，例如：“原文语言”选择日语，“译文语言”选择简体中文，就是日中互译，反之则是中日互译；或者“原文语言”选择日语，“译文语言”选择英语，则是日英互译，反之则是英日互译。

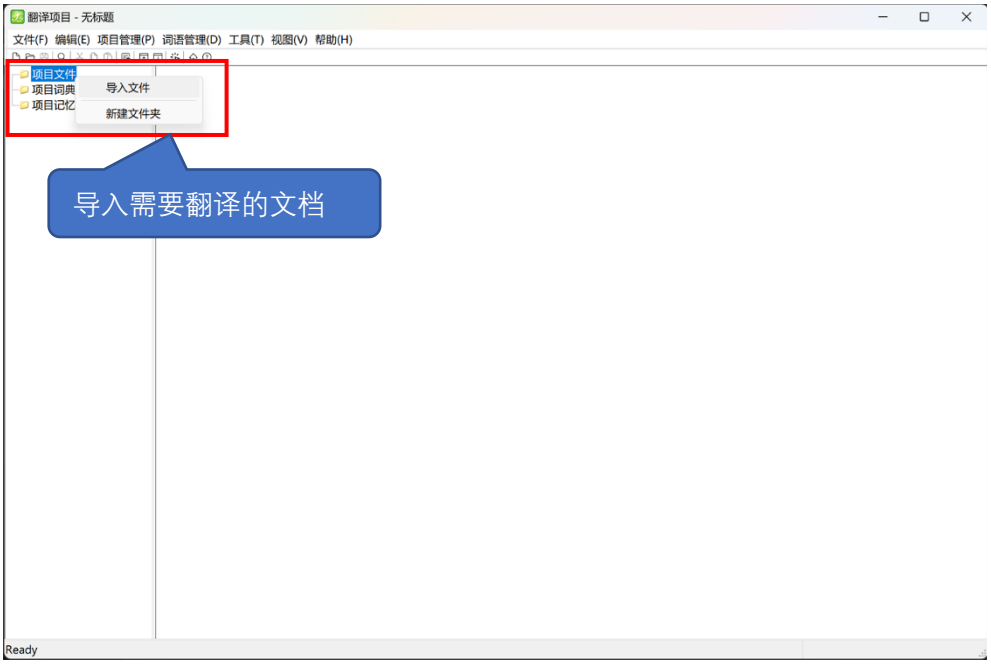


雪人目前原生支持的语种有：简体中文、英语、俄语、日语、法语、西班牙语、韩语、德语；支持这 8 个语种之间的任意互译，雪人会根据不同语种的特点进行优化及支持词形变化。还通过自定义语种的方式扩展支持另外 26 种语言，分别

是：繁体中文、粤语、阿拉伯语、保加利亚语、捷克语、丹麦语、希腊语、爱沙尼亚语、芬兰语、匈牙利语、意大利语、立陶宛语、拉脱维亚语、挪威语、荷兰语、波兰语、葡萄牙语、罗马尼亚语、印度尼西亚语、斯洛伐克语、斯洛文尼亚语、瑞典语、土耳其语、泰语、乌克兰语、越南语（扩展支持的语种操作详见：本章第六节 扩展支持更多的语言）。

下面以英中互译为例说明。

右键点击左侧的项目文件，再点击**【右键菜单】➔【导入文件】**功能，导入一份需要翻译的文档。如下图：



新版支持复杂项目文件结构，在项目文件下面，可以新建多层次的子文件夹。这里为了简单，不新建子文件夹，直接在根目录下导入需要翻译的文档，待译文档是一篇 WORD 的文档，如下图：



将它导入雪人软件中，导入后界面如下图:



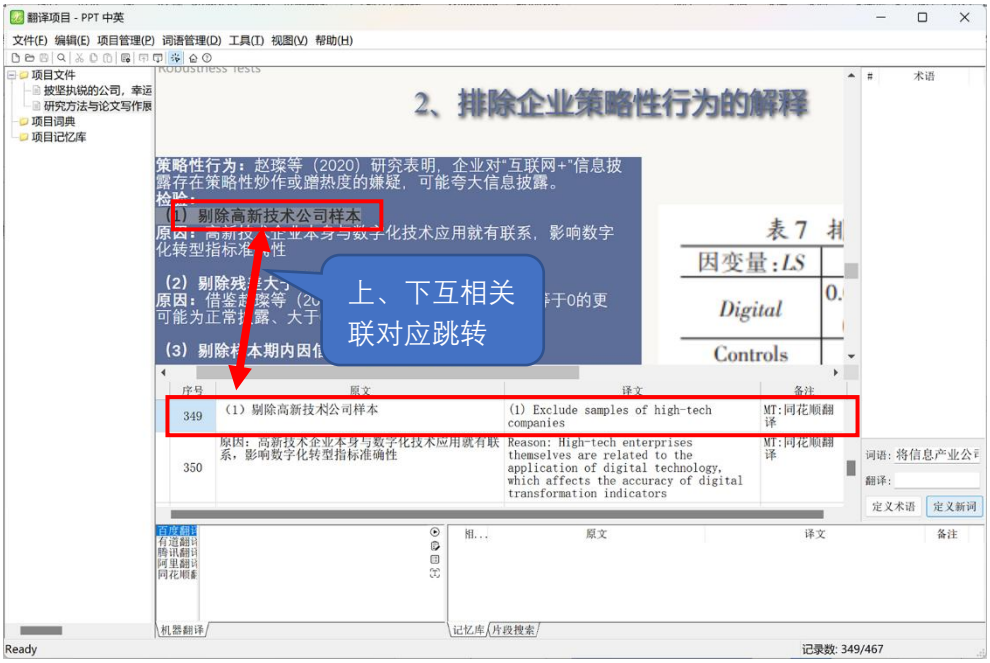
中上部是预览区，可以预览原文、实时预览译文、实时预览双语。中间是表格翻译区，右侧是术语提示区，左下方是机器翻译区，右下方是记忆库和片段搜索

区。各个分区之间的大小可以自由调整，鼠标指向分区边界时会变成双箭头形状，往一侧拖动即可改变分区的大小。

雪人支持 WORD、EXCEL、PPT、TXT 等文件的翻译，在预览窗口，可以预览原文，双击预览窗口的某句原文内容，可以在下方的翻译编辑区中立即定位到该句，非常方便。例如：翻译下图的 PPT 文件，双击预览窗口的 PPT 原文“其他稳键性检验”这一句，在下方的翻译编辑区中，立即跳转定位到该句上，序号为 324，



同理，在下方的翻译编辑区中，点击任意一句，上方的预览窗口立即跳转显示预览该句，如下图：上下窗口的句子，正向、反向都可以联动定位。

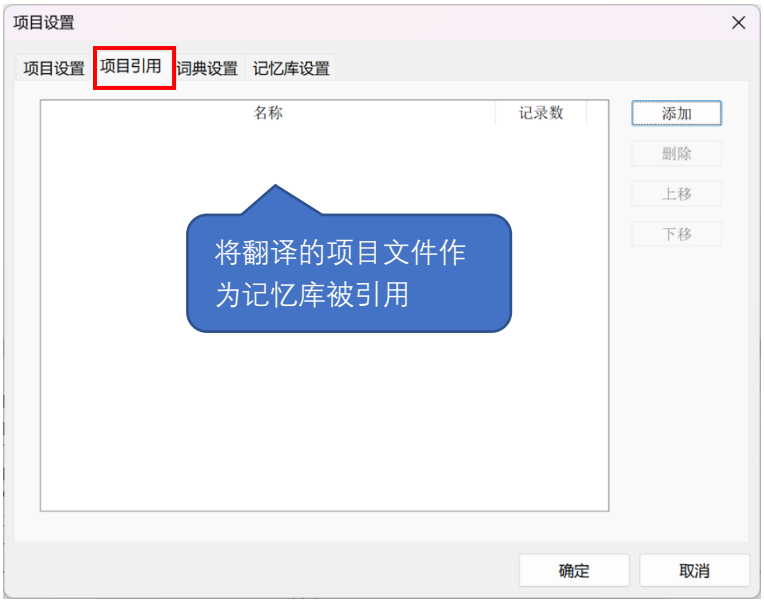


第二节 外挂记忆库和词典

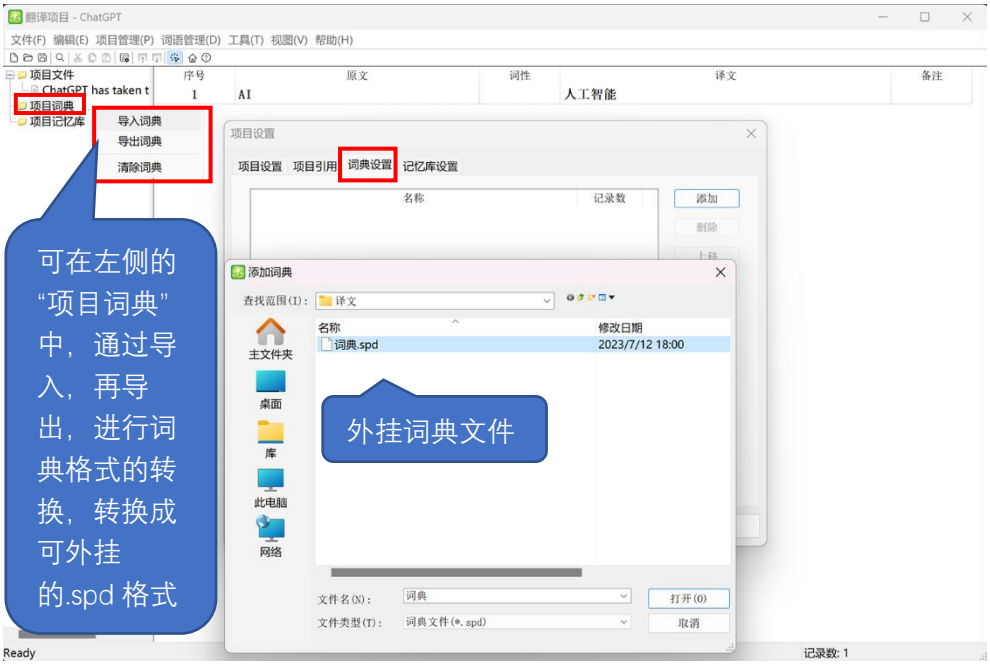
导入需要翻译的文档后，点击【项目管理】➔【项目设置】菜单，会弹出"项目设置"窗口。如下图：



点击“项目引用”页签，可以添加引用多个已经翻译好的项目文件，文件后缀是“.SPR”，即直接引用翻译的项目文件作记忆库，包括被引用的项目文件下的“项目词典”和“项目记忆库”中的内容，都可以被引用。“项目引用”的好处是有完整的篇章信息、可以判断上文是否精确匹配，还可以省去记忆库导入、导出的工作。对于刚开始使用软件的用户，可能还没有已经翻译好的项目文件可以被引用，则此处可留空。



点击“词典设置”页签，可以添加引用词典文件。若暂时没有可以外挂的词典，也可以留空。



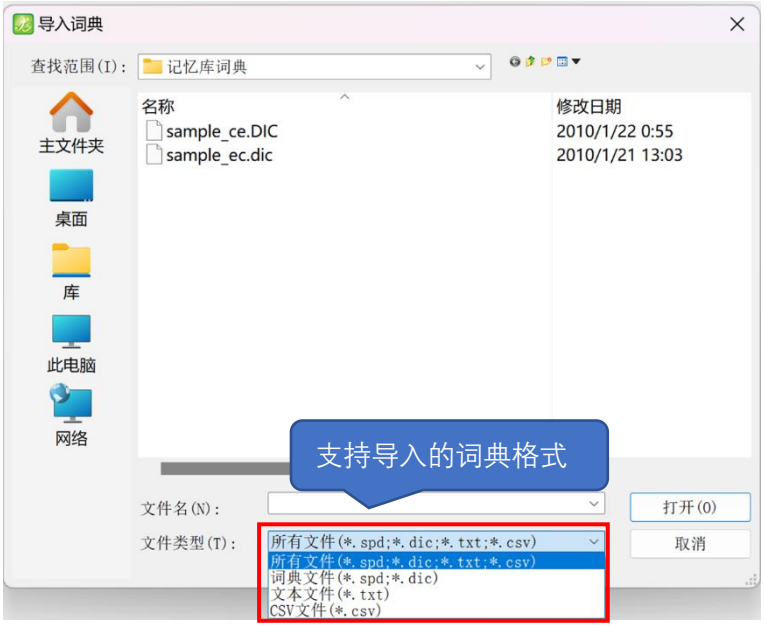
雪人支持导入的词典格式有：如下图:

.spd——新版雪人的词典格式。

.dic——旧版雪人的词典格式。旧版的词典需要在项目词典中进行导入后再导出，才能转换成新版的格式外挂引用。

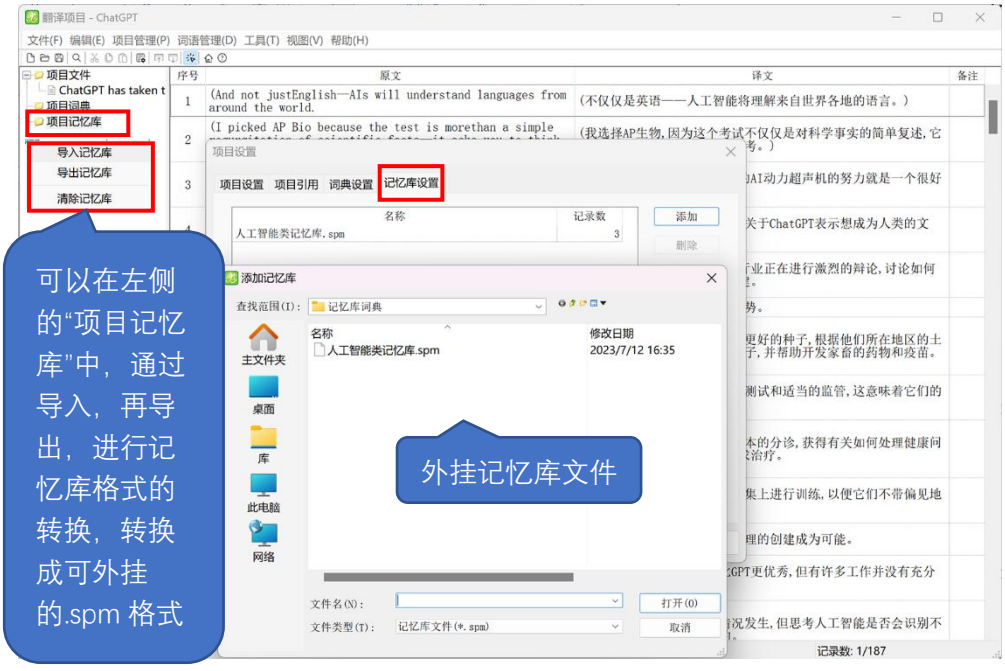
.txt——文本格式，原文、译文之间要有分隔符。

.csv——CSV 格式，EXCEL 文件可以在 EXCEL 软件中另存为.csv 格式，然后导入雪人中。



可以将上述格式的词典文件，在左侧的“项目词典”中，通过右键菜单的导入，再导出功能，进行词典格式的转换，转换成可外挂的.spd 格式。

同理，点击“记忆库设置”页签，可以添加引用记忆库文件。若暂时没有可以外挂的记忆库，也可以留空。



雪人支持导入的记忆库文件格式有：如下图：

.spm——新版雪人的记忆库格式。

.stm——旧版雪人的记忆库格式。旧版的记忆库需要在“项目记忆库”中进行导入后再导出，才能转换成新版的格式外挂引用。

.tmx——各款 CAT 软件记忆库的通用交换格式。

.txt——文本格式，原文、译文之间要有分隔符。

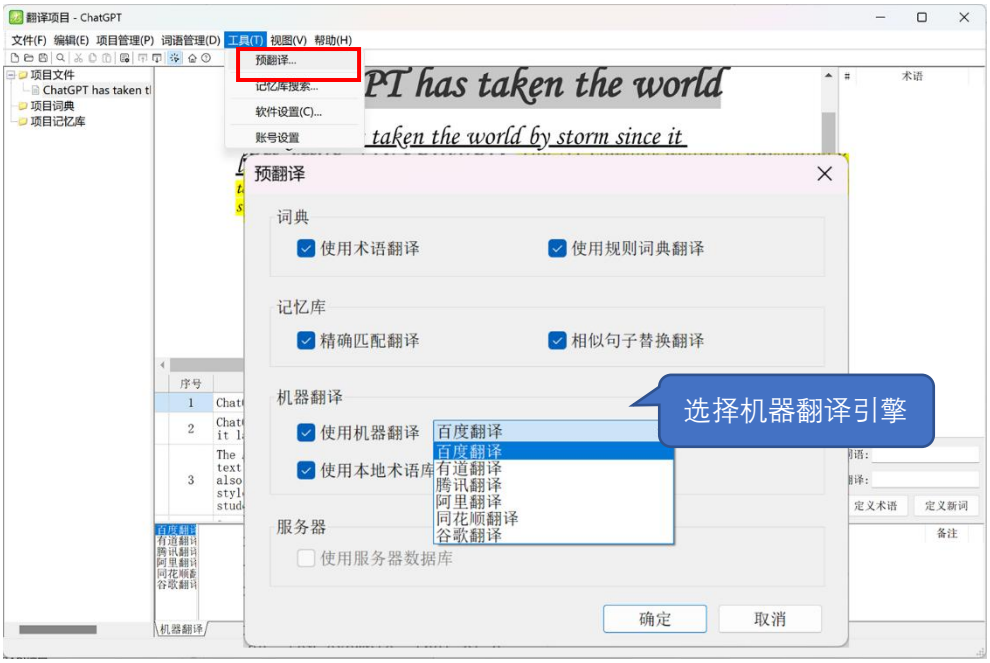
.csv——CSV 格式，EXCEL 文件可以在 EXCEL 软件中另存为.csv 格式，然后导入雪人中。



可以将上述格式的记忆库文件，在左侧的“项目记忆库”中，通过右键菜单的导入，再导出功能，进行记忆库格式的转换，转换成可外挂的.spm 格式。

第三节 快速全文预翻译

在“项目设置”中添加好需要参考引用的“项目文件”或“记忆库”，在“项目词典”中设置好术语，剩下的工作就交给“预翻译”来完成，点击**【工具】➔【预翻译】**菜单，会弹出“预翻译”窗口，如下图:



注意：目前提供百度翻译、有道翻译、腾讯翻译、阿里翻译、同花顺翻译、谷歌翻译引擎供选择。欲使用这些机器翻译引擎，应先去各大引擎站点注册账号，并将 APPID 和 KEY 复制黏贴在雪人的“软件设置”——“机器翻译”窗口中，设置好后，此处才可以出现下拉窗口供选择。具体账号申请及设置方法详见“机器翻译”章节。

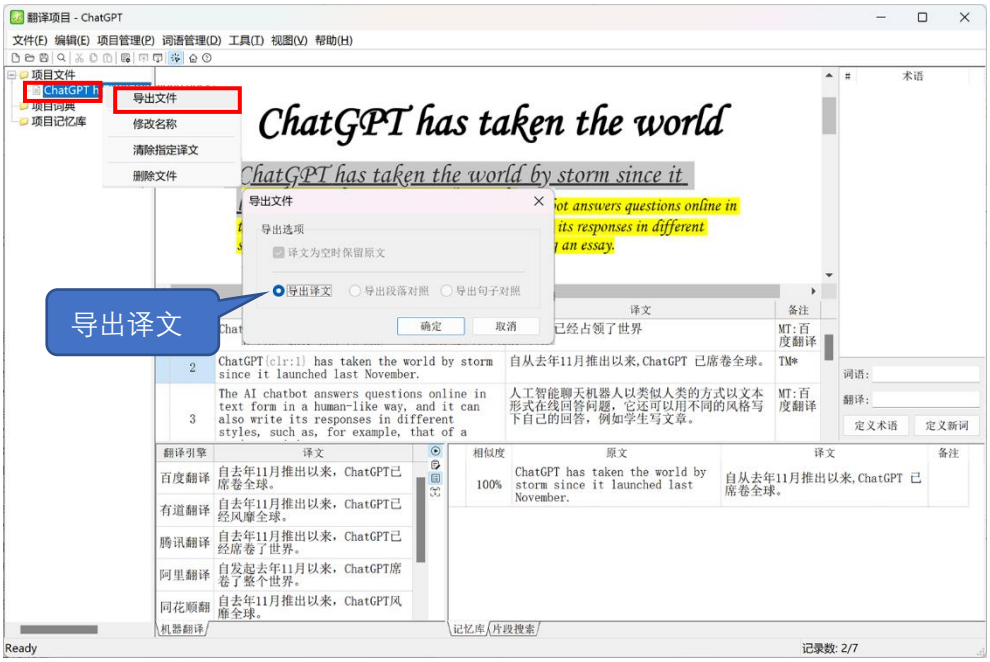
设置好预翻译的相关参数后，点击确定按钮，即进行全文预翻译，如果术语、记忆库中有匹配的，则优先选用这些的结果填写译文，如果找不到匹配的就发送给机器翻译引擎进行翻译，并可以使用本地术语库和机器翻译相结合技术，修正机器翻译的结果。预翻译完成后，在“备注”列中会显示该句译文的来源，如下图:


翻译项目 - ChatGPT
文件(F) 编辑(E) 项目管理(P) 词语管理(D) 工具(T) 视图(V) 帮助(H)
项目文件
ChatGPT has taken t
项目词典
项目记忆库
ChatGPT has taken the world
ChatGPT has taken the world by storm since it launched last November. The AI chatbot answers questions
text form in a human-like way, and it can also write its responses in diff
styles, such as, for example, that of a student writing an essay.
术语
序号
原文
译文
备注
1
ChatGPT has taken the world
ChatGPT已经占领了世界
MT:百度翻译
2
ChatGPT{clr:1} has taken the world by storm since it launched last November.
自从去年11月推出以来,ChatGPT 已席卷全球。
TM*
3
The AI chatbot answers questions online in text form in a human-like way, and it can also write its responses in different styles, such as, for example, that of a
人工智能聊天机器人以类似人类的方式以文本形式在线回答问题，它还可以用不同的风格写下自己的回答，例如学生写文章。
MT:百度翻译
定义术语
定义新词
翻译引擎
译文
百度翻译
自去年11月推出以来，ChatGPT已席卷全球。
有道翻译
自去年11月推出以来，ChatGPT已经风靡全球。
腾讯翻译
自去年11月推出以来，ChatGPT已经席卷了世界。
阿里翻译
自发起去年11月以来，ChatGPT席卷了整个世界。
同花顺翻
自去年11月推出以来，ChatGPT风靡全球。
机器翻译
相似度
原文
译文
备注
100%
ChatGPT has taken the world by storm since it launched last November.
自从去年11月推出以来,ChatGPT 已席卷全球。
记忆库
片段搜索
Ready
记录数: 2/7
来自百度的机器翻译
来自记忆库
同时列出多个机器翻译引擎的结果
来自记忆库匹配结果

第四节 导出译文及双语对照

一、导出译文

预翻译完成后，可以对预翻译的结果进行进一步的修改，修改完成后，右键点击屏幕左侧的某个具体翻译文档后，会弹出右键菜单，点击**【右键菜单】➔【导出文件】**菜单，可以将翻译好的译文导出，如下图：

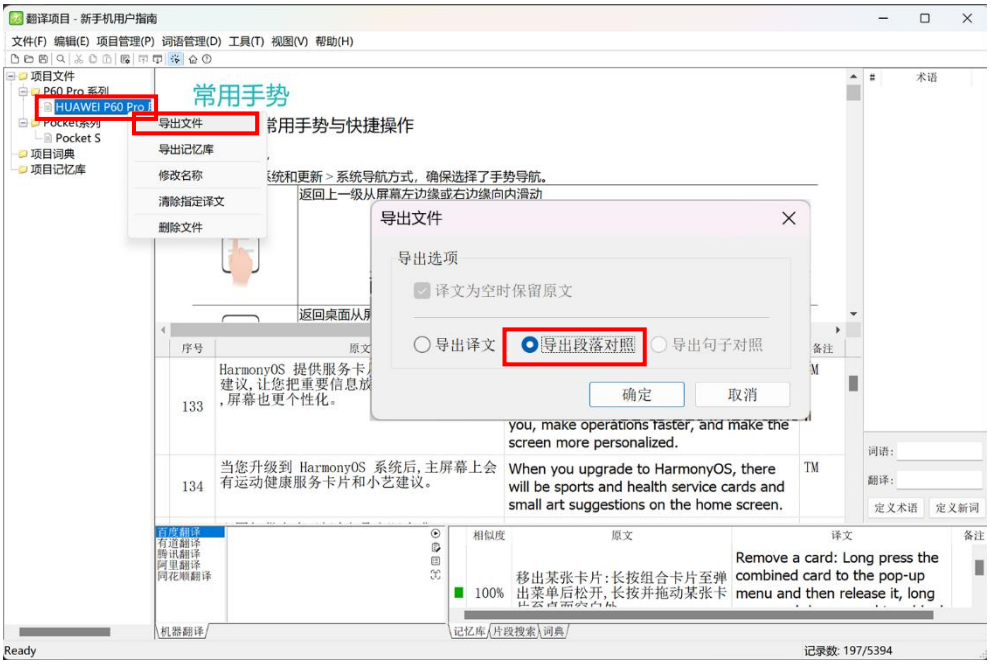


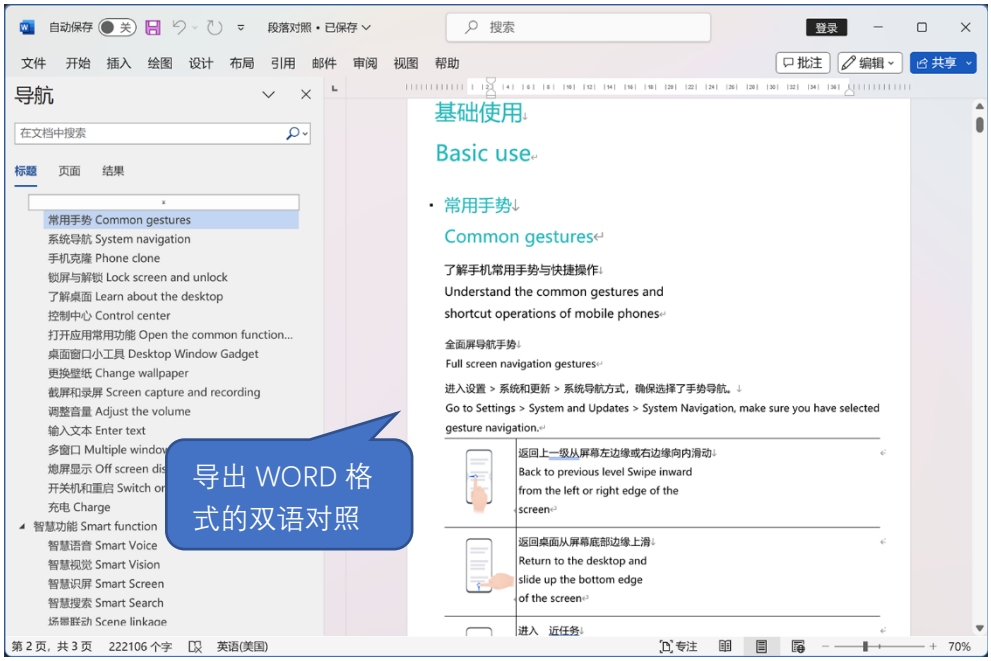
导出的译文如下图：译文可以很好地还原了原文的排版样式。

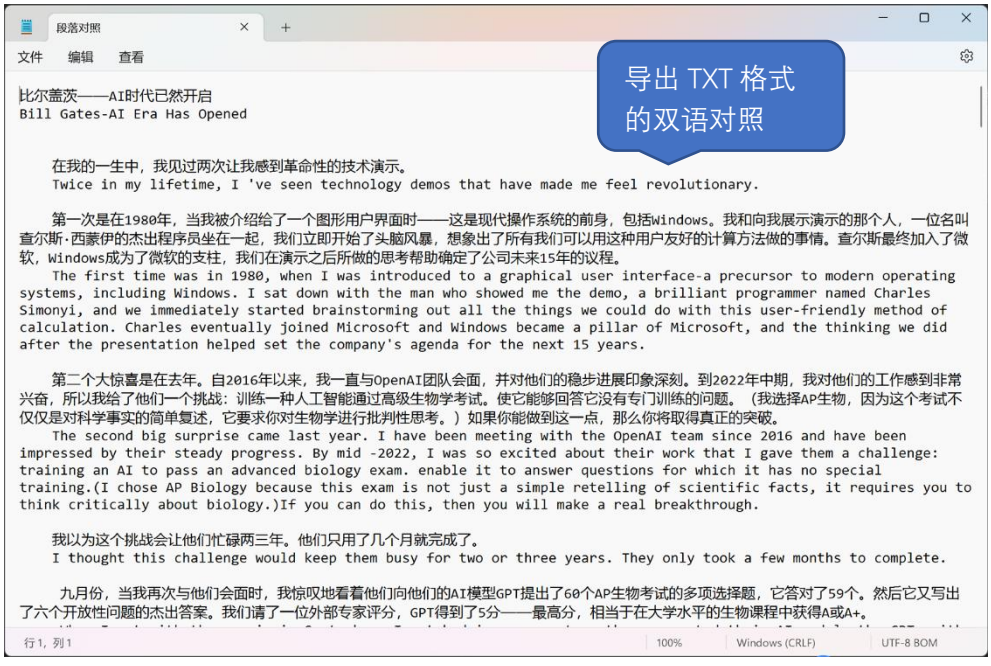


二、导出双语对照格式的译文

点击左侧项目文件下的某个具体翻译文档后，点击【**右键菜单➔导出文件**】功能，在弹出的窗口中选择“导出段落对照”选项，可以导出双语对照格式的译文，如下图：







比尔盖茨——AI时代已然开启
Bill Gates-AI Era Has Opened

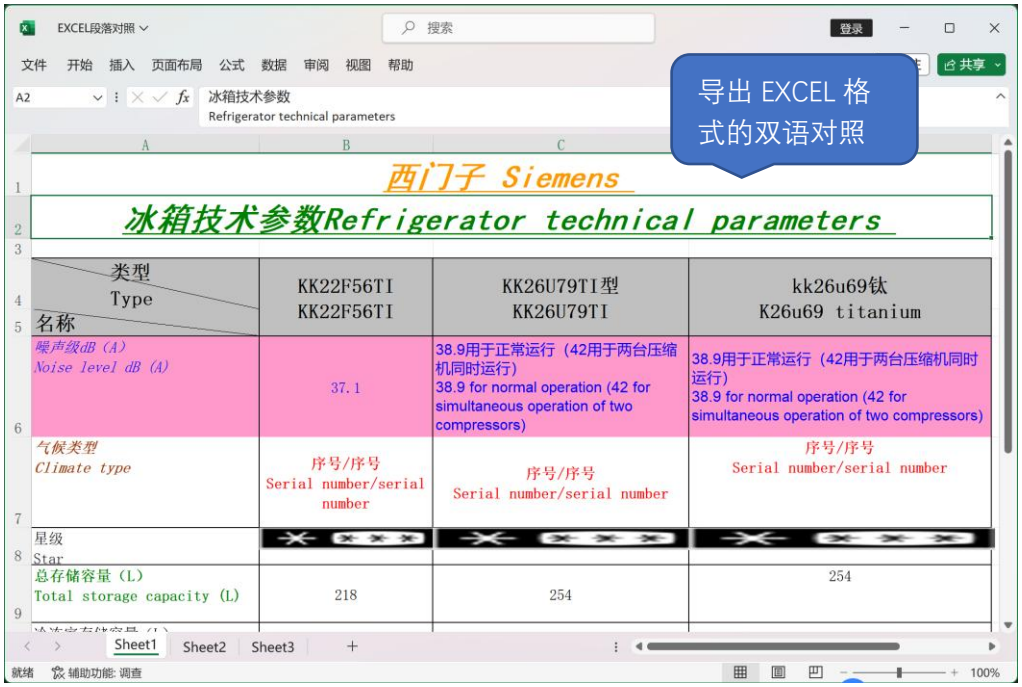
在我的一生中，我见过两次让我感到革命性的技术演示。
Twice in my lifetime, I 've seen technology demos that have made me feel revolutionary.

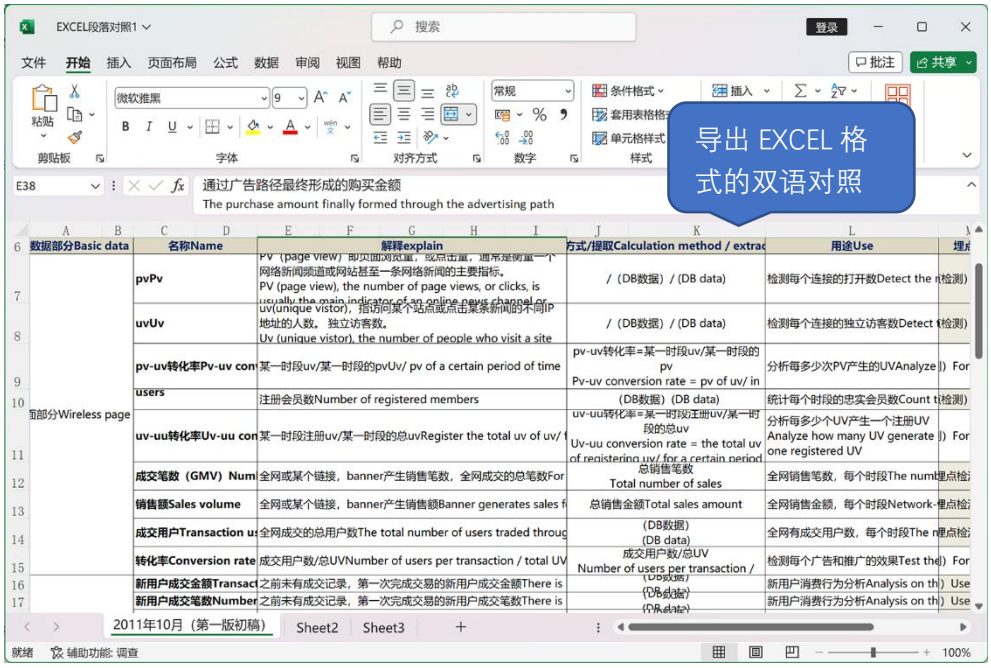
第一次是在1980年，当我被介绍给了一个图形用户界面时——这是现代操作系统的前身，包括Windows。我和向我展示演示的那个人，一位名叫查尔斯·西蒙伊的杰出程序员坐在一起，我们立即开始了头脑风暴，想象出了所有我们可以用这种用户友好的计算方法做的事情。查尔斯最终加入了微软，Windows成为了微软的支柱，我们在演示之后所做的思考帮助确定了公司未来15年的议程。
The first time was in 1980, when I was introduced to a graphical user interface-a precursor to modern operating systems, including Windows. I sat down with the man who showed me the demo, a brilliant programmer named Charles Simonyi, and we immediately started brainstorming out all the things we could do with this user-friendly method of calculation. Charles eventually joined Microsoft and Windows became a pillar of Microsoft, and the thinking we did after the presentation helped set the company's agenda for the next 15 years.

第二个大惊喜是在去年。自2016年以来，我一直与OpenAI团队会面，并对他们的稳步进展印象深刻。到2022年中期，我对他们的工作感到非常兴奋，所以我给了他们一个挑战：训练一种人工智能通过高级生物学考试。使它能够回答它没有专门训练的问题。（我选择AP生物，因为这个考试不仅仅是对科学事实的简单复述，它要求你对生物学进行批判性思考。）如果你能做到这一点，那么你将取得真正的突破。
The second big surprise came last year. I have been meeting with the OpenAI team since 2016 and have been impressed by their steady progress. By mid -2022, I was so excited about their work that I gave them a challenge: training an AI to pass an advanced biology exam. enable it to answer questions for which it has no special training.(I chose AP Biology because this exam is not just a simple retelling of scientific facts, it requires you to think critically about biology.)If you can do this, then you will make a real breakthrough.

我以为这个挑战会让他们忙碌两三年。他们只用了几个月就完成了。
I thought this challenge would keep them busy for two or three years. They only took a few months to complete.

九月份，当我再次与他们会面时，我惊叹地看着他们向他们的AI模型GPT提出了60个AP生物考试的多项选择题，它答对了59个。然后它又写出了六个开放性问题的杰出答案。我们请了一位外部专家评分，GPT得到了5分——最高分，相当于在大学水平的生物课程中获得A或A+。





对于导出的 EXCEL 格式的双语对照，如果希望原文、译文分行显示的，可以在 EXCEL 中，全选后设置“自动换行”，并设置“自动调整行高”，如下图:


自动换行
自动调整行高
调整成分行显示对照
EXCEL段落对照
登录
文件 开始 插入 页面布局 公式 数据 审阅 视图 帮助
批注 共享
微软雅黑
常规
条件格式
套用表格格式
单元格样式
插入
删除
格式
加载项
剪贴板 字体 对齐方式 数字 样式 加载项
J6
计算方式/提取
Calculation method / extraction
单元格大小
行高(H)...
自动调整行高(A)
列宽(W)...
自动调整列宽(I)
默认列宽(D)...
可见性
隐藏和取消隐藏(U)
组织工作表
重命名工作表(R)
移动或复制工作表(M)...
工作表标签颜色(T)
保护
保护工作表(P)...
锁定单元格(L)
设置单元格格式(E)...
基础数据部分 Basic data part
名称 Name
解释 explain
计算方式/提取 Calculation method / extraction
埋点位置
pv
Pv
PV（page view）即页面浏览量，或点击
/（DB数据）
/ (DB data)
uv
Uv
/（DB数据）
/ (DB data)
pv-uv转化率
Pv-uv conversion rate
某一时段uv/某一时段的pv
Uv/ pv of a certain period of time
pv-uv转化率=某一时段uv/某一时段的pv
Pv-uv conversion rate = pv of uv/ in
users
注册会员数
Number of registered members
(DB数据)
(DB data)
uv-uu转化率
Uv-uu conversion rate
某一时段注册uv/某一时段的总uv
Register the total uv of uv/ for a certain period of time
uv-uu转化率=某一时段注册uv/某一时段的总uv
Uv-uu conversion rate = the total uv
无线页面部分
Wireless page section
成交笔数（GMV）
Number of transactions (GMV)
全网或某个链接，banner产生销售笔数，全网成交的总笔数
For the whole network or a link, banner generates the number of sales
总销售笔数
Total number of sales
销售额
Sales volume
全网或某个链接，banner产生销售额
Banner generates sales for the whole network or a link
总销售金额
Total sales amount
sheet1
Sheet2
Sheet3
就绪
辅助功能: 调查
100%

第五节 常用热键

拆分句子

将光标移动到需要拆分的原文处，按**【Alt】+【回车键】**。新版拆分句子需要按【Alt】键，和旧版的操作不一样。

合并句子

将光标移动到句尾，按**【Delete】键**，将下一行的句子合并到上一行。注意：如果上、下句子之间有空行或跨段落的句子，则不能合并句子。

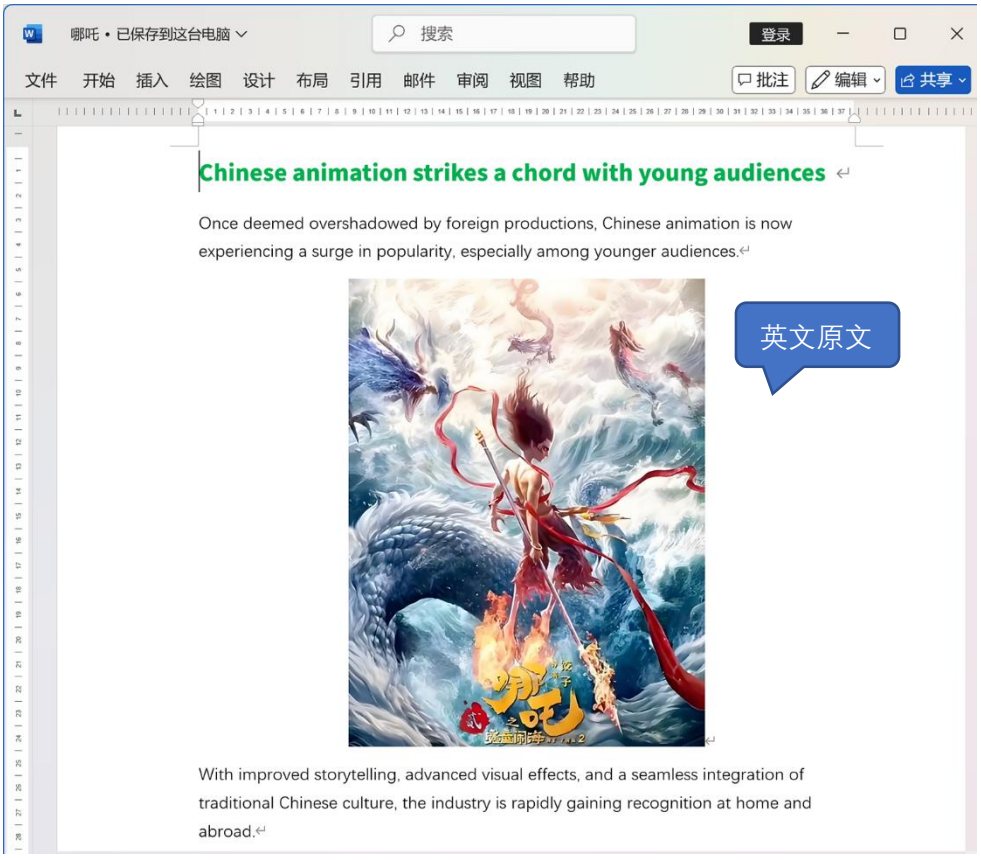
【F2】——浏览与修改状态切换，浏览状态下光标以“整行的表格” 为单位进行上下移动。修改状态下光标以“单元格”为单位进行上下移动，在项目词典或记忆库窗口中，修改状态下可以对词典或记忆库的原文、译文进行编辑、修改。
【F6】——将从记忆库中搜索到的匹配句子的译文复制到当前译文区，并对不同的地方进行了自动替换翻译。【Ctrl+F7】不使用自动翻译替换。
【F7】——将机器翻译结果复制到译文。将大语言模型的翻译结果复制到译文。注意光标在上方的译文区时按此键。
【F8】——将原文直接复制到译文区。
【Ctrl+F】——查找
【Ctrl+R】——查找和替换
【F3】——查找时，按【F3】可以查找下一目标。
【Alt】+【回车】——跳到下一个未翻译的句子，如果记忆库中有精确匹配的句子，则自动填写译文后，继续搜索下一句。

第六节 扩展支持更多的语言

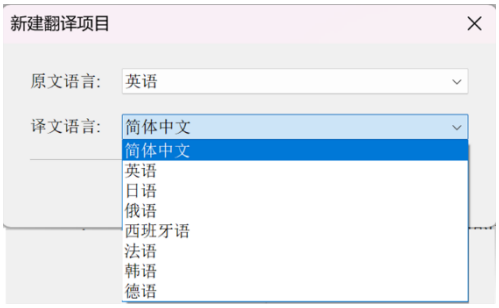
雪人目前原生支持的语种有：简体中文、英语、俄语、日语、法语、西班牙语、韩语、德语；支持这 8 个语种之间的任意互译，雪人会根据不同语种的特点进行优化及支持词形变化。在新建翻译项目时，可以通过下拉框选择这 8 个语种。

还通过自定义语种的方式扩展支持另外 26 种语言，分别是：繁体中文、粤语、阿拉伯语、保加利亚语、捷克语、丹麦语、希腊语、爱沙尼亚语、芬兰语、匈牙利语、意大利语、立陶宛语、拉脱维亚语、挪威语、荷兰语、波兰语、葡萄牙语、罗马尼亚语、印度尼西亚语、斯洛伐克语、斯洛文尼亚语、瑞典语、土耳其语、泰语、乌克兰语、越南语。这些语言都可以使用传统机器翻译引擎和 AI 大语言模型来翻译它们。

下面将举例介绍：例如需要将原文是英语的“哪吒”这篇文档，翻译成简体中文、繁体中文、粤语、阿拉伯语。原文如下：

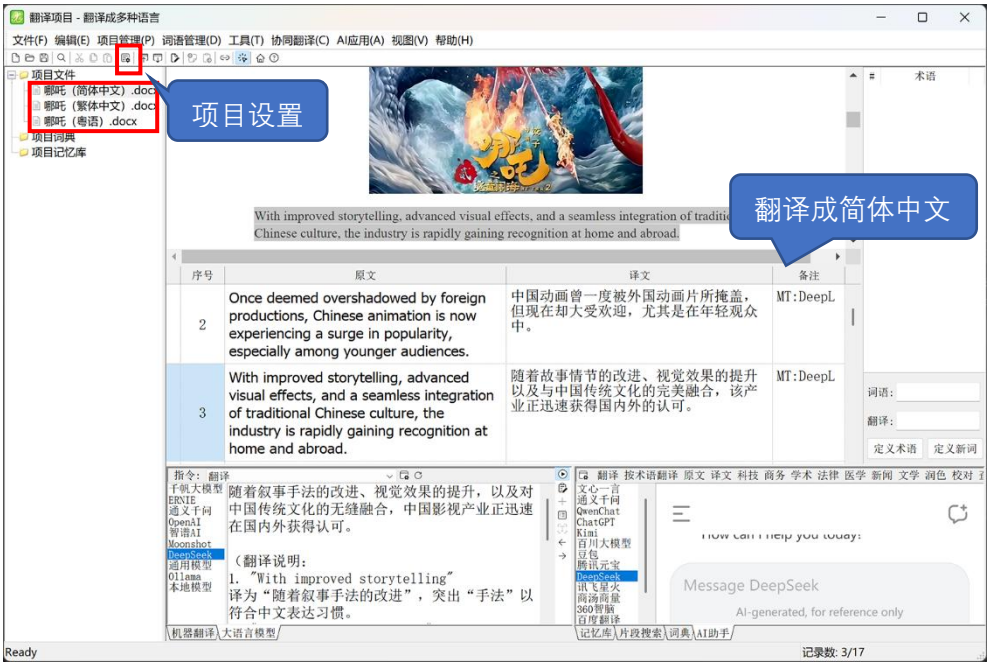


点击**【文件】➔【新建】➔【翻译项目】**菜单，在弹出的“新建翻译项目”窗口中，原文语言选择“英语”，译文语言选择“简体中文”。如下图：



导入需要翻译的“哪吒”这篇文档，导入后将该文档复制两份（复制方法：按住【Ctrl】键+鼠标左键点击左侧的“哪吒”这篇文档并将它拖动到“项目文件”下面，即可实现文件复制），复制后，利用右键菜单分别修改它们的文件名为：哪吒（简体中文）、哪吒（繁体中文）、哪吒（粤语）。

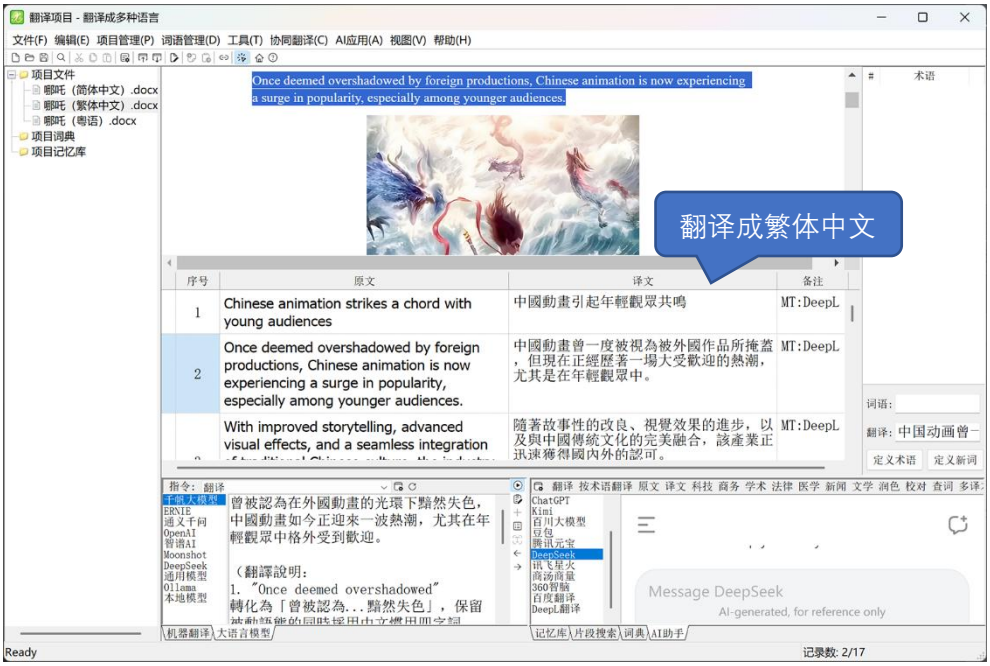
点击左侧的“哪吒（简体中文）”这个文档后，再点击**【工具】➔【预翻译】**菜单，选用“DeepL”机器翻译引擎，将这个文档翻译成简体中文，如下图：



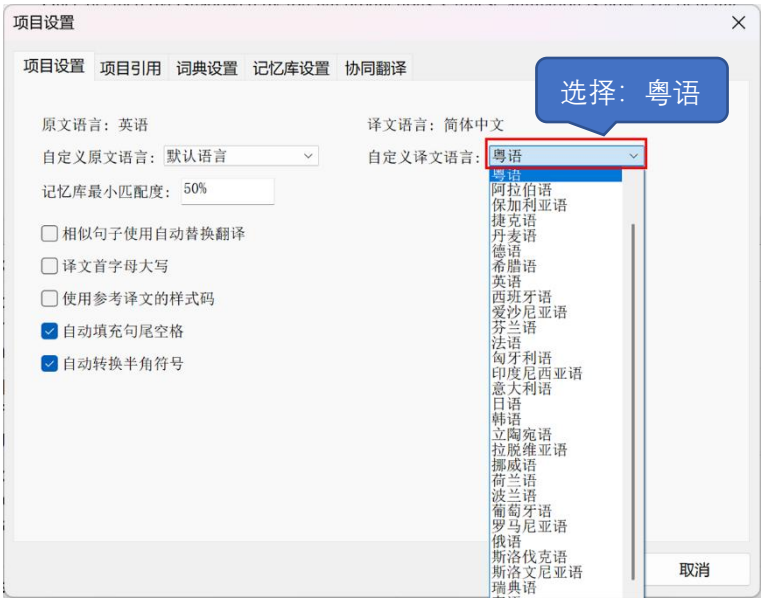
跟着我们再把这个文件翻译为繁体中文版本。首先，我们可以点击工具栏的“项目设置”按钮或点击**【项目管理】➔【项目设置】**菜单，在弹出的“项目设置”窗口中的“自定义译文语言”中，选择“繁体中文”，如下图：



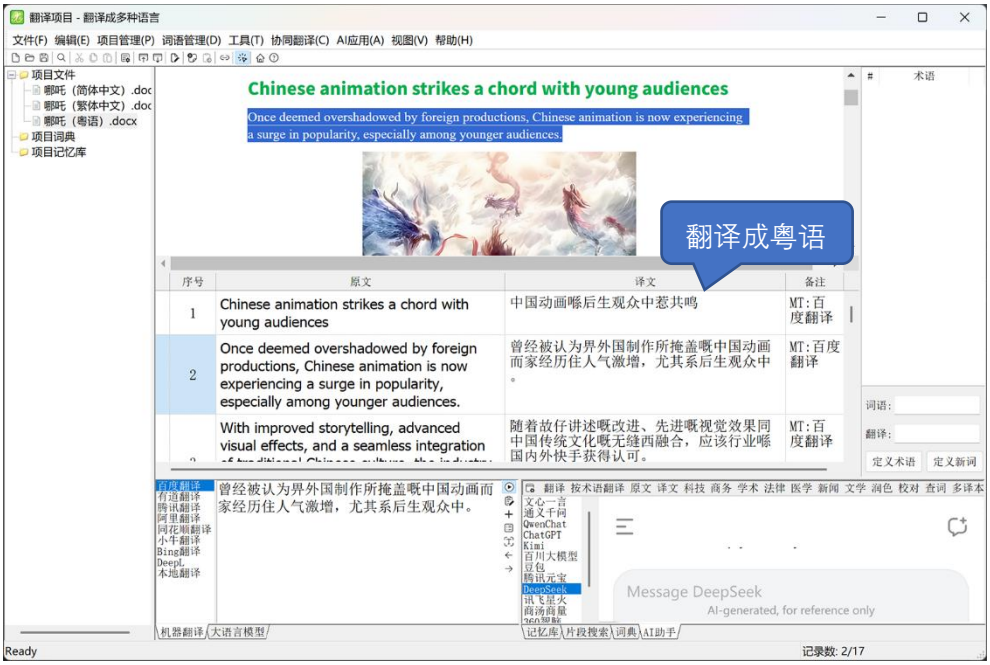
点击左侧的第二个文档即“哪吒（繁体中文）”，再使用【工具】➔【预翻译】菜单，选用“DeepL”机器翻译引擎，就可以将这个文档翻译成繁体中文了。除了用传统的机器翻译引擎来翻译成繁体中文外，我们也可以用 AI 大语言模型来翻译成繁体中文，如下图：



使用同样的方法，我们可以生成粤语版本。通过修改项目设置中的语言，在“自定义译文语言”中选择“粤语”，如下图：



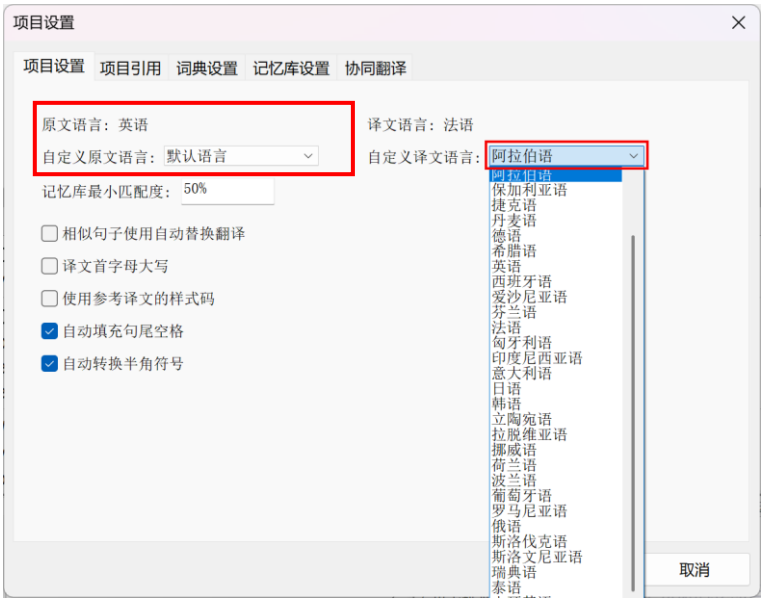
点击左侧的第三个文档即“哪吒（粤语）”，再使用【**工具**】➔【**预翻译**】菜单，选用“百度”机器翻译引擎，就可以将它翻译成粤语。如下图：



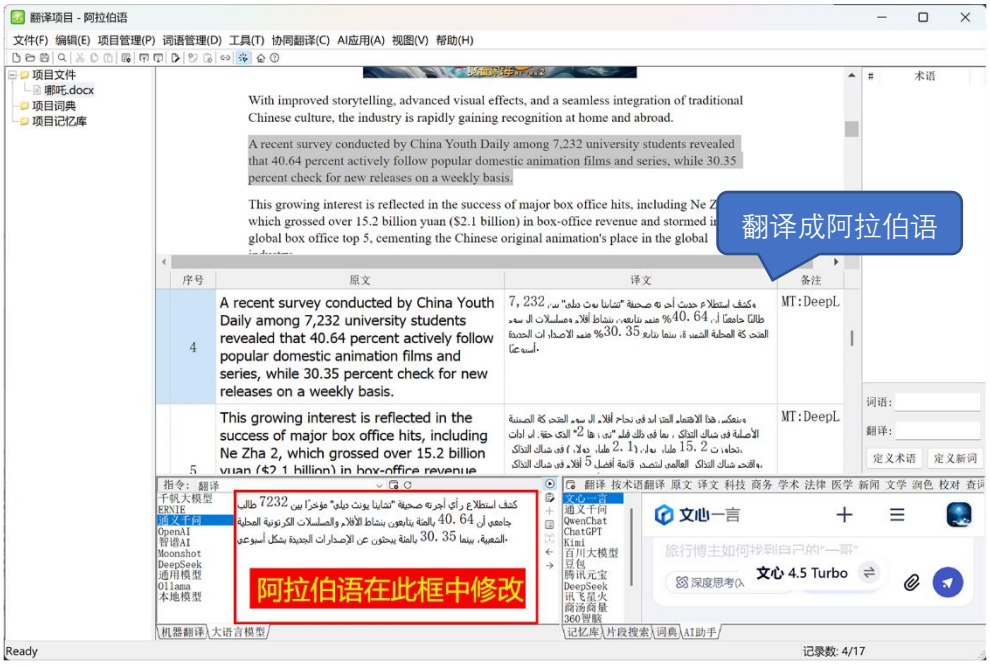
由于雪人目前原生支持的语种是：简体中文、英语、俄语、日语、法语、西班牙

语、韩语、德语这 8 个语种之间的任意互译，更多语种的支持需要在这 8 个语种的基础上进行扩展，并参考这 8 个原生语种的断句规则等信息。如果需要翻译这 8 个语种之外的语言，首先需要根据拟翻译的语言，从这 8 个语种中选取与之相接近的语言作为参考语言，以参考语言创建翻译项目，然后修改“项目设置”中“自定义原文语言”和“自定义译文语言”来选取具体的语言。

由于阿拉伯语和中文相差比较大，我们这里新建一份英译法的项目文件来进行翻译。新建好参考的翻译项目后，导入“哪吒”英文原文，然后点击工具栏的“项目设置”按钮或点击**【项目管理】➔【项目设置】**菜单，在弹出的“项目设置”窗口中的“自定义原文语言”如果填写的是“默认语言”，则表示选用上方的原文语言。在“自定义译文语言”中，选择“阿拉伯语”，如下图：

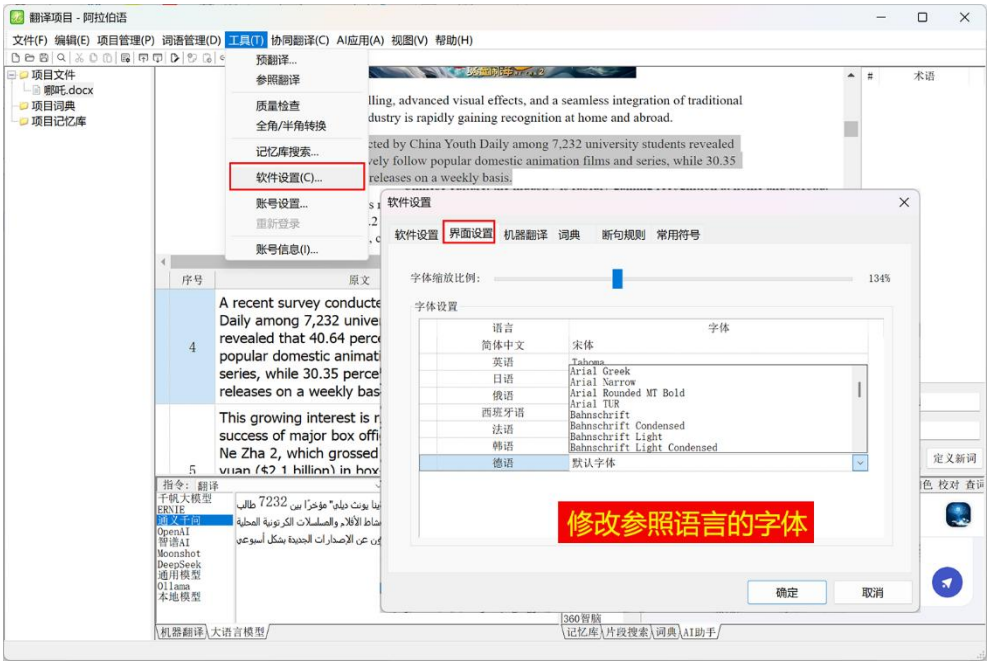


然后点击**【工具】➔【预翻译】**菜单，选用“DeepL”机器翻译引擎，就可以将它翻译成阿拉伯语了，也可以利用 AI 大模型来翻译成阿拉伯语。如下图：



由于目前雪人 CAT 的译文编辑框不能正确地编辑阿拉伯语，如果需要对机器翻译后的阿拉伯语的译文进行修改，建议在屏幕左下方的“机器翻译”框或“大语言模型”框中修改，修改好后再粘贴到上方的表格译文区中。

另外，对于某些语种，可能软件默认的字体不太合适或出现乱码，用户可以自己设置字体获得更佳的显示效果。设置方法是：点击**【工具】➔【软件设置】**菜单，再点击“界面设置”页签，设置各个语言的字体。对于这些扩展支持的语言，建议选择一个自己不常用的语言作为参考语言来翻译其他更多的语种，并设置相应的字体。例如越南语，在 Windows 操作系统中，我们可以选择 Arial，Times New Roman 等字体。

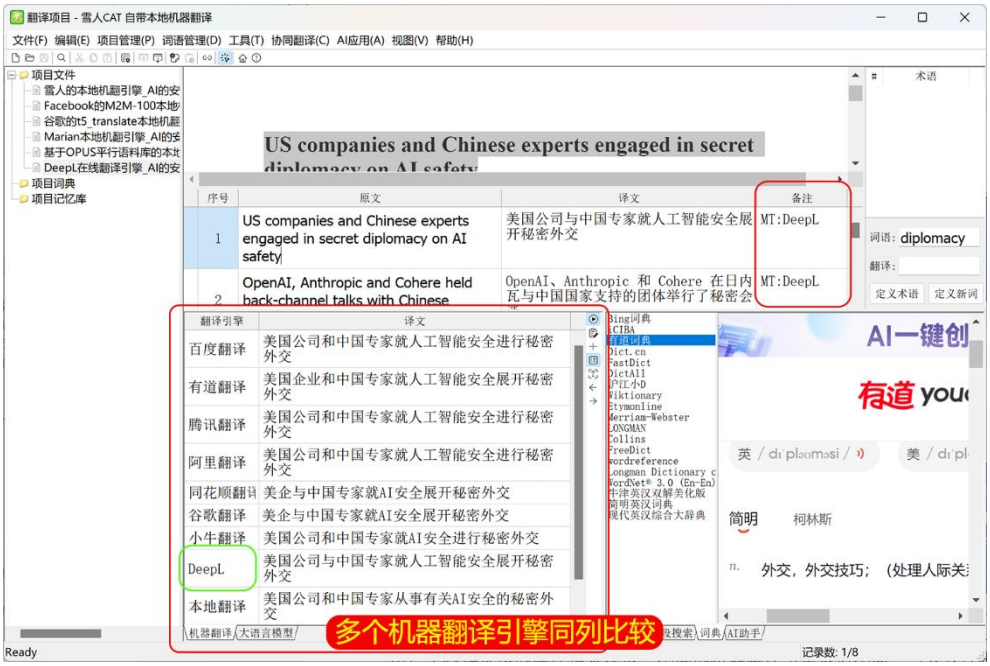


除了使用在线传统机器翻译引擎和 AI 大模型外，本地多语翻译模型（如 m2m100）也可以进行简单的翻译。

第三章 在线机器翻译引擎

第一节 机器翻译引擎的设置

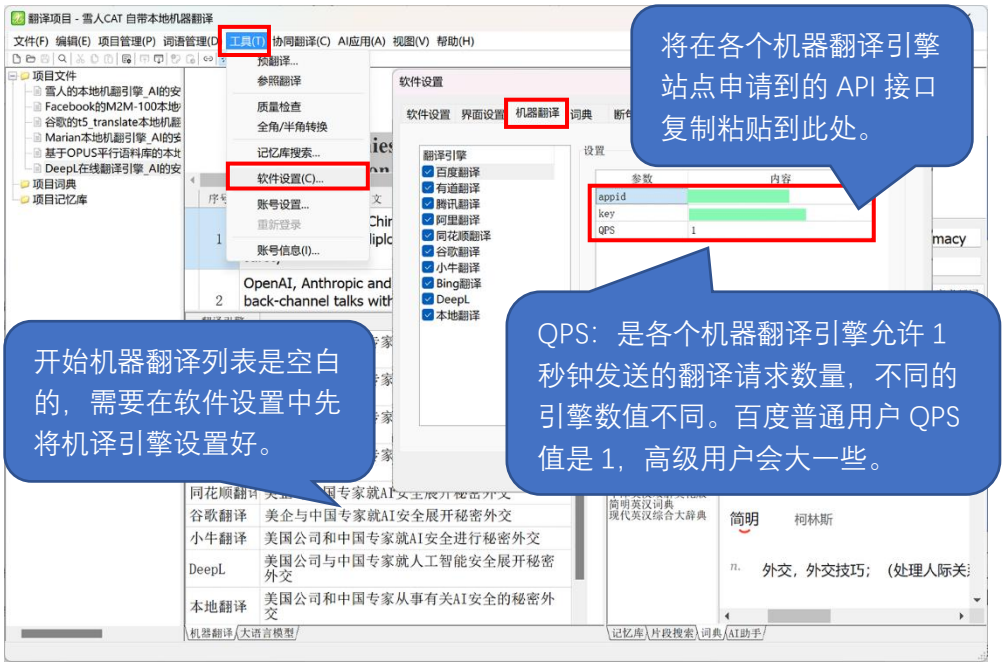
近年来，随着人工智能技术的快速发展，机器翻译质量也得到了显著提升，机器翻译的准确性和速度已经达到了前所未有的水平，用机器翻译辅助人工翻译可以大大提高翻译的效率。雪人新版软件加强了对机器翻译的支持，实现机助人译。目前支持接入的在线机器翻译引擎有：**DeepL 翻译、谷歌翻译、百度云翻译、阿里云翻译、腾讯云翻译、有道翻译、同花顺翻译、小牛翻译**等，多个机器翻译引擎的结果可以同时显示，方便用户比较各大机器翻译引擎的译文质量，择优选择最佳译文，如下图：



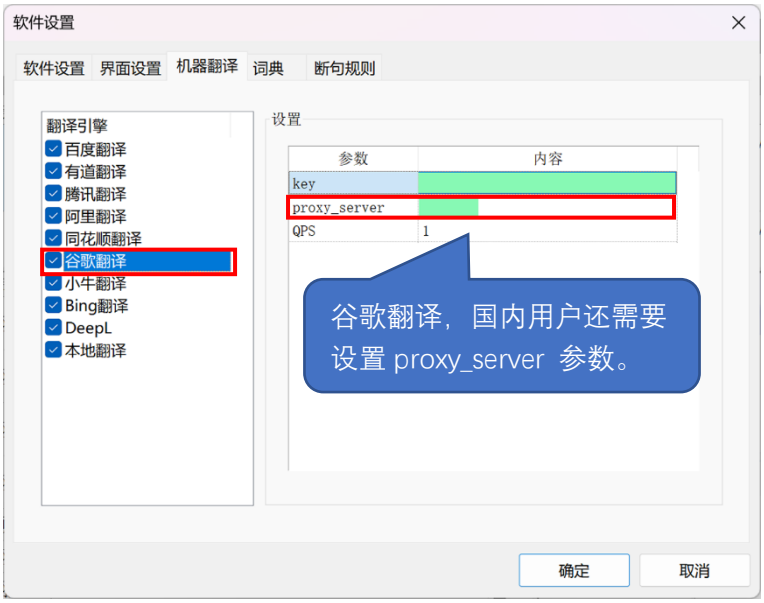
后续还会加入更多的机器翻译引擎，对机器翻译的支持是在雪人服务器中实现的，新增机器翻译引擎用户不需要升级软件即可使用，用户发现好的机器翻译引擎也可以向雪人推荐加入。同时，还可以利用用户自定义的本地术语库与机器翻译系统相结合，对术语的翻译，强制机器翻译系统按照用户所定义的词义进行翻译，确保术语翻译的准确性和一致性，进一步修正机器翻译译文的质量。

设置机器翻译引擎

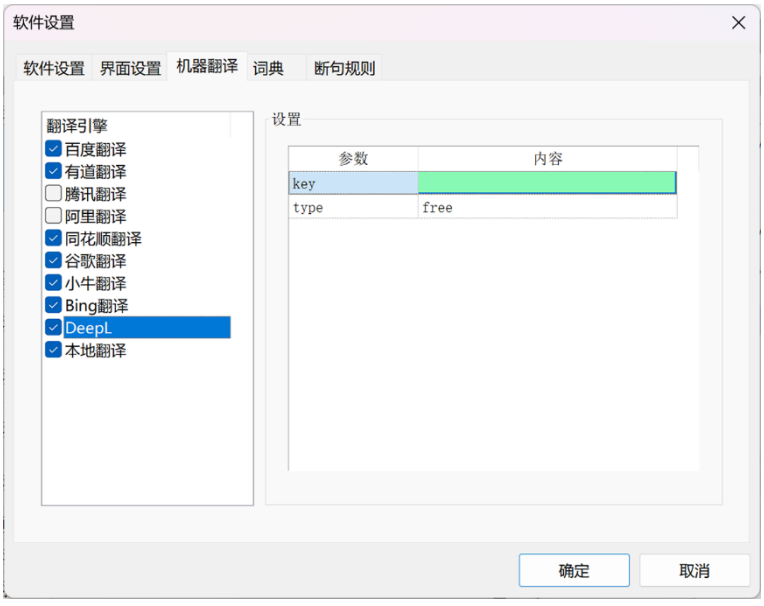
屏幕下方的机器翻译区域默认是空白的，需要先对机器翻译引擎进行设置，才会在此显示出各个机器翻译引擎列表。设置的方法：点击**【工具菜单➔软件设置➔机器翻译】**功能，在“机器翻译”设置窗口中，将百度云翻译、阿里云翻译、腾讯云翻译、有道翻译、同花顺翻译等申请到 API 接口复制粘贴到此处，如下图。如何申请 API 接口？详情请参看 “第六节 申请机器翻译 API 接口”。申请到的 API 接口会有 appid 和 key 两个参数，将它们分别复制粘贴到下图的“内容”列处即可。



在机器翻译设置窗口中，还有一个 QPS 参数，QPS 是各个机器翻译引擎允许用户 1 秒钟发送的翻译请求次数，不同的机器翻译引擎数值不同。雪人软件根据各个机器翻译引擎的限制设置了不同的数值， 一般情况下，用户不需要修改 QPS 参数值。例如：百度普通用户 QPS 值设置为 1，是因为百度允许普通用户 1 秒钟，只能向它发送一个翻译请求，QPS 值设置大了，百度会报出错，错误信息返回的结果会以**红色字体**显示在机器翻译区中。如果是百度的高级用户，则允许 QPS=10，具体要看您申请到的 API 接口而定，软件默认是按普通用户设定的，如果您是百度的高级用户，则可以自行修改 QPS 值 。 阿里的 QPS 默认设置为 50，即阿里允许每秒钟向它发送 50 个翻译请求。国内用户，由于不能直接访问谷歌站点，如果需要使用谷歌翻译，还需要在“proxy_server”参数中，设置代理服务器的地址。如下图：

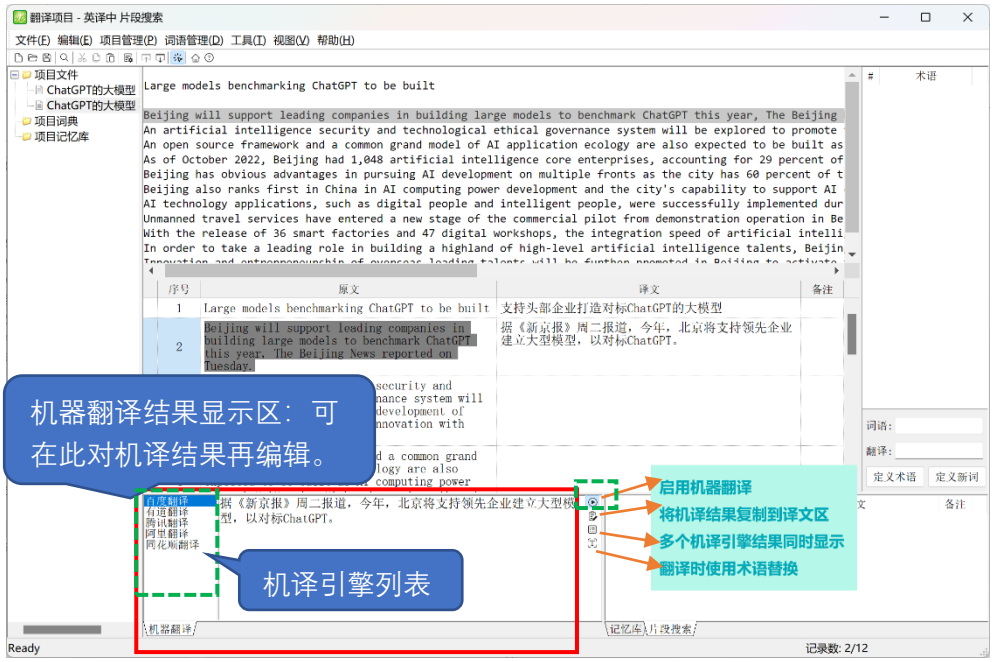


同时能显示多少个机器翻译引擎的结果？取决于用户在**【工具菜单➔软件设置➔机器翻译】**中添加了多少个机器翻译引擎，如下图：即使添加了的机器翻译引擎，若暂时不需要使用它，也可以在“机器翻译引擎”设置窗口中，不勾选它，则该引擎就不会出现在屏幕下方的“机器翻译”引擎列表中，也不会出现在“预翻译”的机器翻译引擎下拉选项中。



第二节 逐句机器翻译提示

机器翻译引擎设置好后，在屏幕左下方的“机器翻译”区中，会出现刚才设置的多个机器翻译引擎列表，**为了节省机器翻译的流量和避免意外发送数据，默认状态下机器翻译功能是关闭的。**点击 ▶ 按钮启动机器翻译，该按钮变成蓝色 ▶，表示开启了机器翻译的功能，会对外发送数据，**注意：如果翻译有保密要求的文档，请不要点击该按钮。**



机器翻译区中，右侧按钮解释如下：

▶ ——启用机器翻译，再次点击可关闭机器翻译。▶ 关闭状态，▶ 启用状态。

——将机器翻译结果复制到译文区中，或按【F6】键。

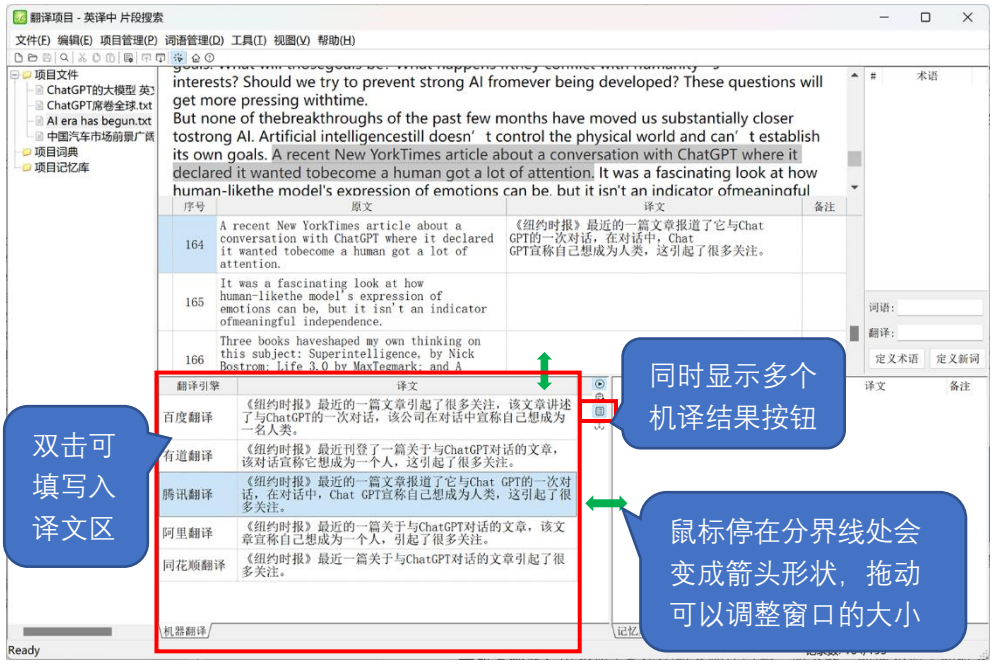
——将多个机器翻译结果同时显示，便于比较、选用最佳结果。

——翻译时使用术语库中的术语进行替换翻译，确保对机器翻译对术语翻译结果的准确性和一致性。（此次更新暂未加入）

如果对机器翻译的结果不满意，可以在机器翻译结果显示区中进行编辑，修改好后，再填入译文区中。

第三节 同时显示多个机译引擎结果

不同的机器翻译引擎各有长短，雪人 CAT 将多个机器翻译引擎的翻译结果并列显示出来，方便用户比较各个机器翻译引擎译文的质量，择优选用最佳译文，在屏幕下方的“机器翻译”区中，点击旁边的 显示模式按钮，切换到多个机器翻译引擎结果同时显示界面，如下图。双击某个机器翻译引擎的翻译结果，可以将它填写到译文区中。



例如：对于这一句：A recent New York Times article about a conversation with ChatGPT where it declared it wanted to become a human got a lot of attention.

百度翻译：《纽约时报》最近的一篇文章引起了很多关注，该文章讲述了与 ChatGPT 的一次对话，该公司在对话中宣称自己想成为一名人类。

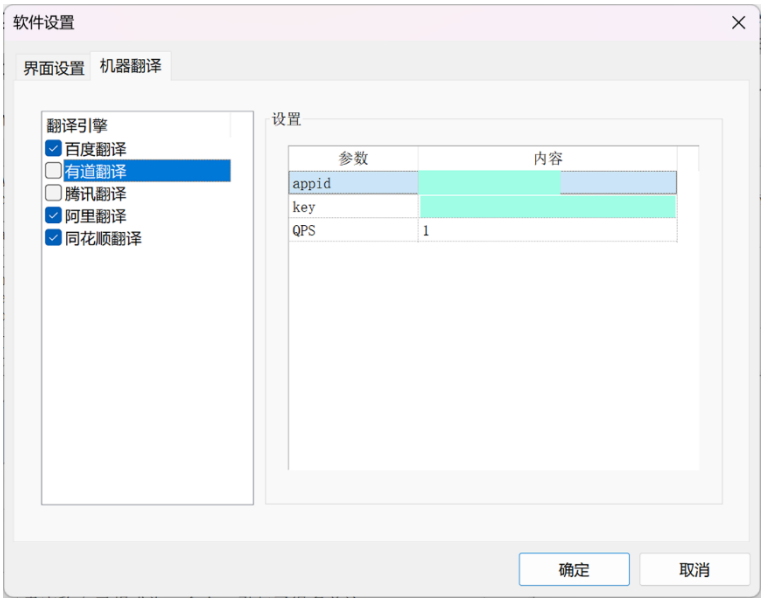
有道翻译：《纽约时报》最近刊登了一篇关于与 ChatGPT 对话的文章，该对话宣称它想成为一个人，这引起了很多关注。

腾讯翻译：《纽约时报》最近的一篇文章报道了它与 Chat GPT 的一次对话，在对话中，Chat GPT 宣称自己想成为人类，这引起了很多关注。

经过对比后发现：针对这句的翻译，百度和有道都将“it”的指代翻译错了，**腾讯**的译文质量最好，鼠标双击**腾讯翻译**的结果，将它填写入译文区。

如果觉得“机器翻译”区的窗口不够大，可以将鼠标指向窗口的边界位置（如上图的绿色箭头处），此时鼠标会变成双向箭头状态，按下鼠标左键并拖动，可以调整窗口的大小。

同时能显示多少个机器翻译引擎的结果？取决于用户在**【工具菜单➔软件设置➔机器翻译】**中添加了多少个机器翻译引擎，如下图：即使添加了的机器翻译引擎，若暂时不需要使用它，也可以在“机器翻译引擎”设置窗口中，不勾选它，则该引擎就不会出现在屏幕下方的“机器翻译”引擎列表中，也不会出现在“预翻译”的机器翻译引擎下拉选项中。

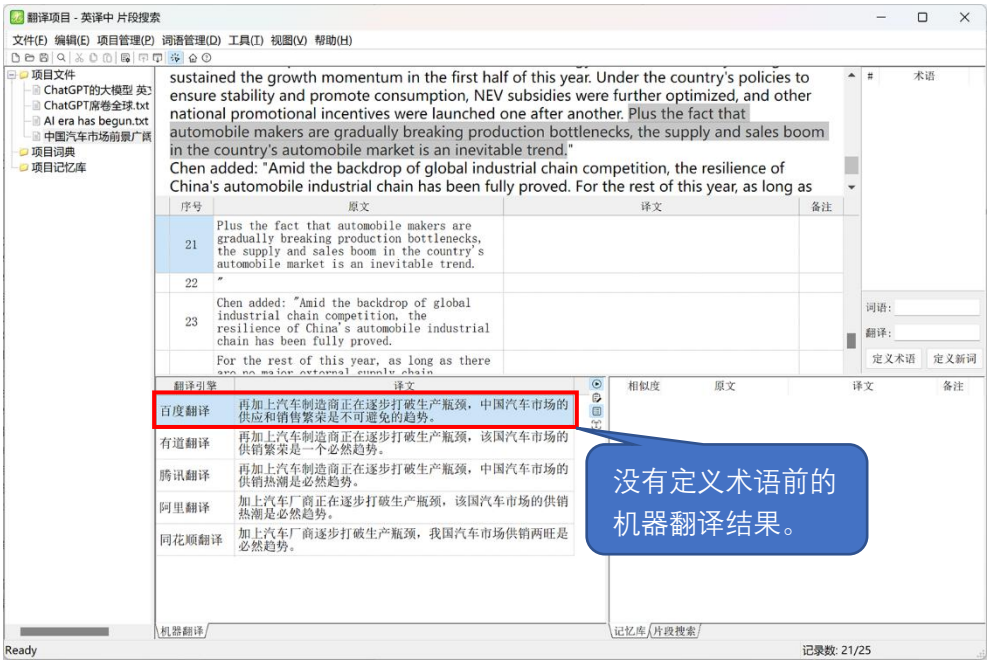


第四节 机器翻译时进行术语替换

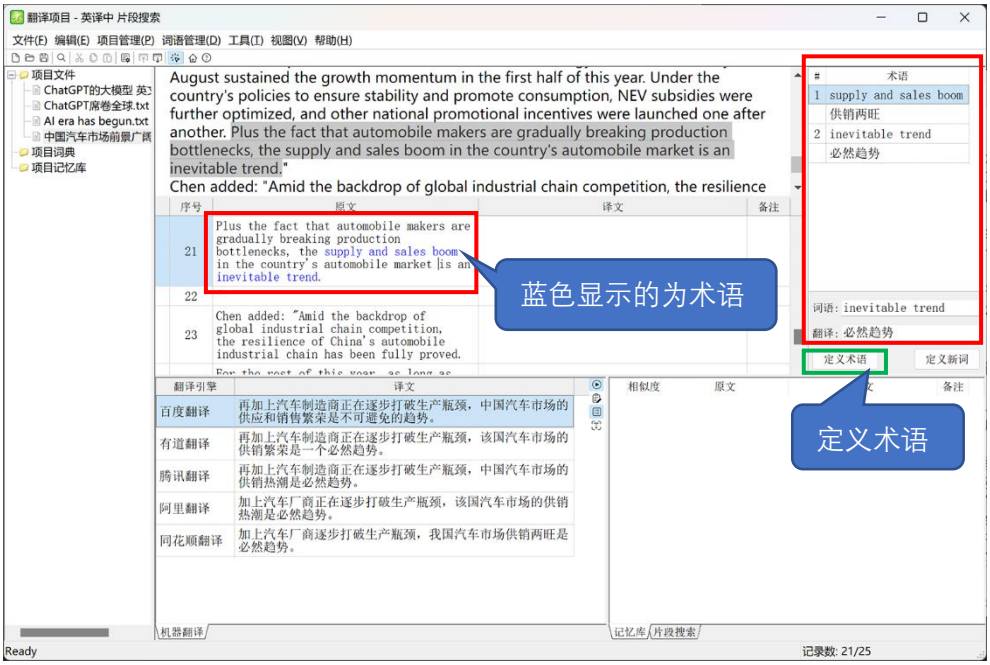
将用户自定义的本地术语库与机器翻译系统相结合，可以实现对术语的精准翻译。软件会强制机器翻译系统按照用户所定义的词义进行翻译，以确保术语翻译的准确性和一致性，从而提高译文的质量。

例如对以下这句：Plus the fact that automobile makers are gradually breaking production bottlenecks, the **supply and sales boom** in the country's automobile market is an **inevitable trend**.

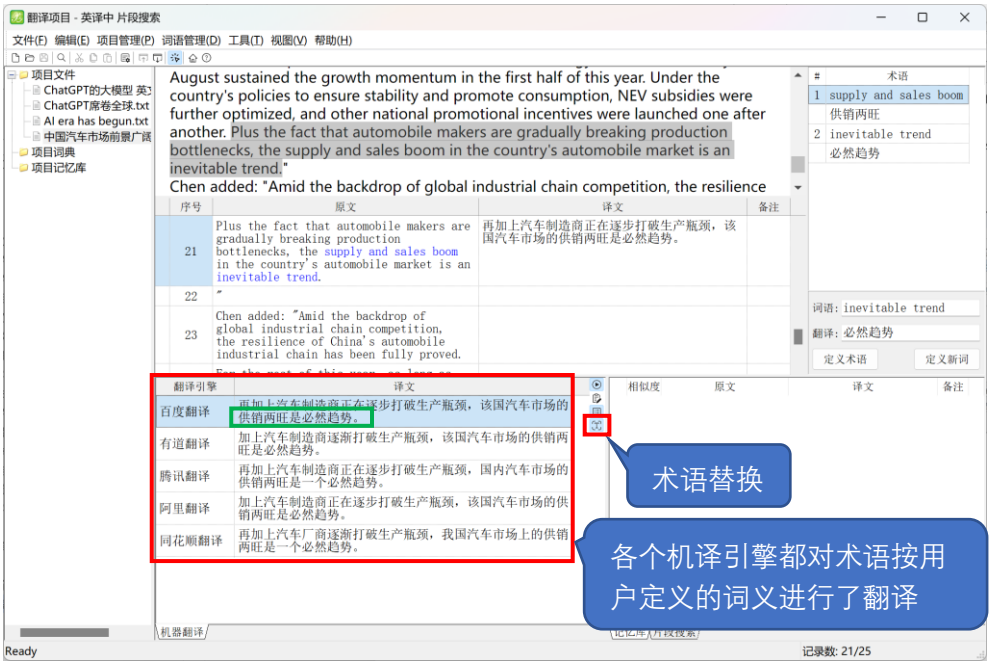
百度机器翻译结果是：再加上汽车制造商正在逐步打破生产瓶颈，中国汽车市场的供应和销售繁荣是不可避免的趋势。如下图：



如果将该句的 supply and sales boom 定义为术语：供销两旺。将 inevitable trend 定义为术语：必然趋势。



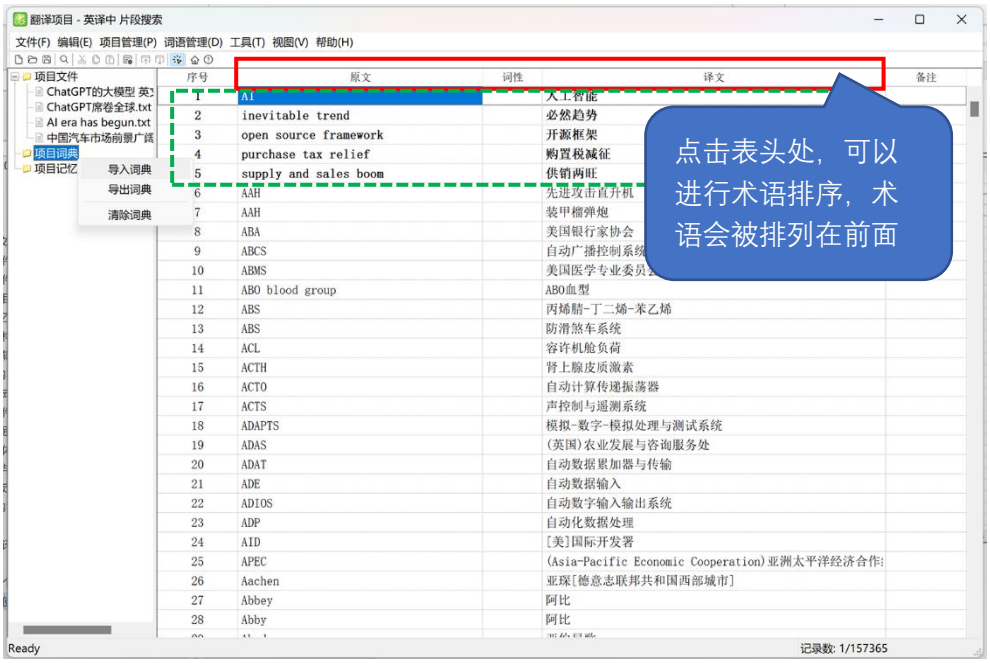
定义好术语后，点击机器翻译窗口右侧的 术语替换按钮，**百度机器翻译**的结果变成：再加上汽车制造商正在逐步打破生产瓶颈，该国汽车市场的供销两旺是必然趋势。同时，有道、腾讯、阿里等其他几个机器翻译引擎也对术语进行了替换翻译。如下图：



用户定义的术语，会自动保存在屏幕左侧的项目词典中并**加粗显示**，如下图：在项目词典中，按住【Shift】+【↑】或【↓】光标键，可以批量选择多条记录，然后按【**右键菜单➔设为术语**】功能将它们设为术语。

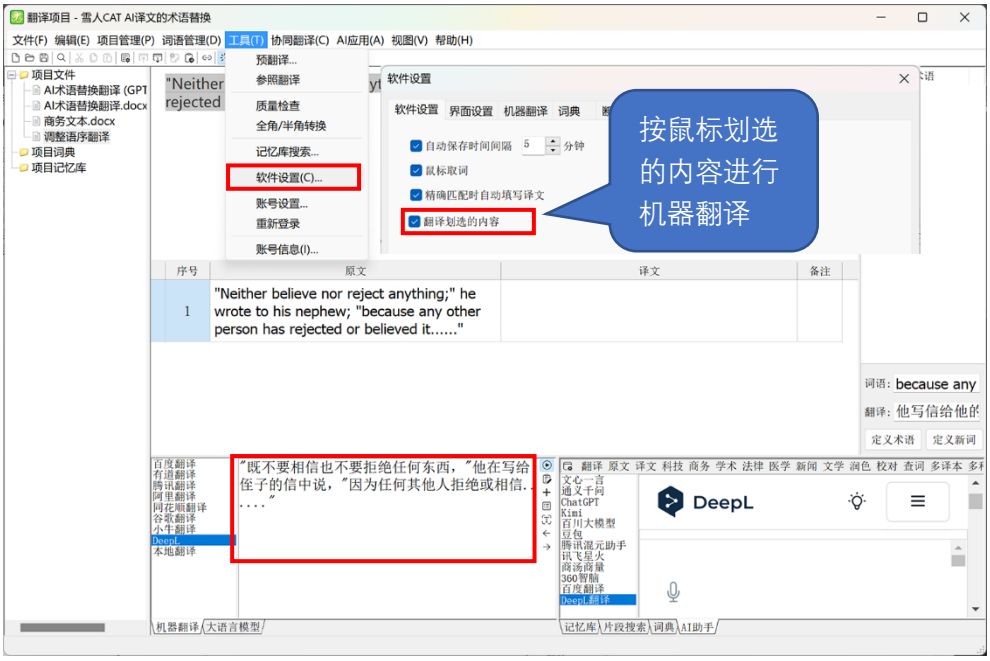


点击下图表头的“原文”处或“译文”处，可以排序（项目记忆库中点击表头也可以实现排序），术语会被排列在最前面，其他的词汇也会按字母顺序进行排序。


翻译项目 - 英译中 片段搜索
文件(F) 编辑(E) 项目管理(P) 词语管理(D) 工具(T) 视图(V) 帮助(H)
项目文件
ChatGPT的大模型 英
ChatGPT席卷全球.txt
AI era has begun.txt
中国汽车市场前景广
项目词典
项目记忆
导入词典
导出词典
清除词典
序号 原文 词性 译文 备注
1 AI 人工智能
2 inevitable trend 必然趋势
3 open source framework 开源框架
4 purchase tax relief 购置税减征
5 supply and sales boom 供销两旺
6 AAH 先进攻击直升机
7 AAH 装甲榴弹炮
8 ABA 美国银行家协会
9 ABCS 自动广播控制系统
10 ABMS 美国医学专业委员
11 ABO blood group ABO血型
12 ABS 丙烯腈-丁二烯-苯乙烯
13 ABS 防滑煞车系统
14 ACL 容许机舱负荷
15 ACTH 肾上腺皮质激素
16 ACTO 自动计算传递振荡器
17 ACTS 声控制与遥测系统
18 ADAPTS 模拟-数字-模拟处理与测试系统
19 ADAS (英国)农业发展与咨询服务处
20 ADAT 自动数据累加器与传输
21 ADE 自动数据输入
22 ADIOS 自动数字输入输出系统
23 ADP 自动化数据处理
24 AID [美]国际开发署
25 APEC (Asia-Pacific Economic Cooperation)亚洲太平洋经济合作
26 Aachen 亚琛[德意志联邦共和国西部城市]
27 Abbey 阿比
28 Abby 阿比
点击表头处，可以进行术语排序，术语会被排列在前面
Ready
记录数: 1/157365

第五节 按鼠标划选的内容进行机译

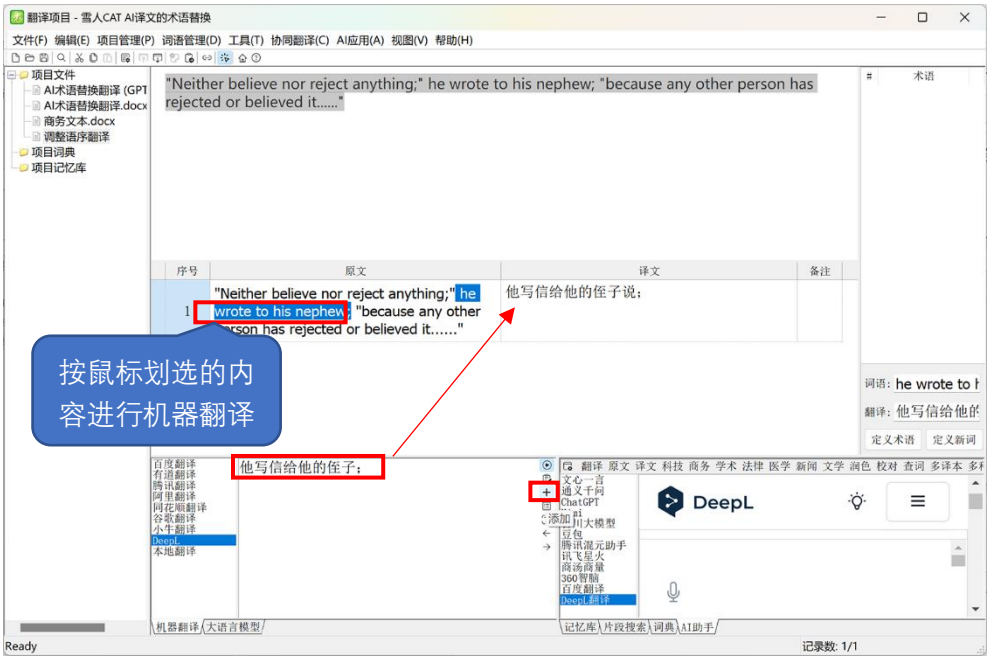
有时候，机器翻译整句话可能不太准确，通过鼠标划选，将整句拆分成多个小句段来翻译，结果会准确很多。点击【**工具**】➔【**软件设置**】菜单，在“软件设置”窗口中，有“翻译划选的内容”选项，软件默认打开了这个选项，在机器翻译时可以根据鼠标划选的内容进行翻译，如果不需要此功能，可以关闭该选项。如下图：



例如上图中的这一句：“Neither believe nor reject anything;" he wrote to his nephew; "because any other person has rejected or believed it.....”

DeepL 机器翻译的译文：“既不要相信也不要拒绝任何东西，"他在写给侄子的信中说，"因为任何其他人拒绝或相信.....”

机器翻译的译文语序不太好，我们可以通过鼠标划选，先翻译哪个部分，再翻译哪个部分？这样调整就可以得到更加通顺的译文。先用鼠标划选“he wrote to his nephew;”下方的机器翻译窗口就仅出现所划选句子的译文，点击旁边的 ＋ “添加”译文按钮，将译文粘贴上去，如下图：



接着划选原文的"Neither believe nor reject anything;"这部分，再点击旁边的＋“添加”译文按钮，将译文粘贴在句尾。＋“添加”译文按钮是将下方机器翻译的译文添加到当前光标定位处，也可以在下方的机器翻译窗口，使用鼠标划选部分译文来粘贴。



第六节 申请机器翻译 API 接口

雪人允许用户自己去各大机器翻译引擎站点申请 API 接口，这样可以降低用户的翻译浏览成本。机器翻译需要消耗服务器的算力，各大机器翻译引擎站点是需要按翻译的流量收费的。但是，一般的机器翻译引擎，每月都会给用户一定额度的免费流量，如果不允许用户自己去申请 API 接口，只能使用 CAT 软件内置的 API 接口，用户就无法使用上这些免费的流量，表面看操作是简单了，但翻译的流量成本确增加了。

一、申请阿里云翻译 API

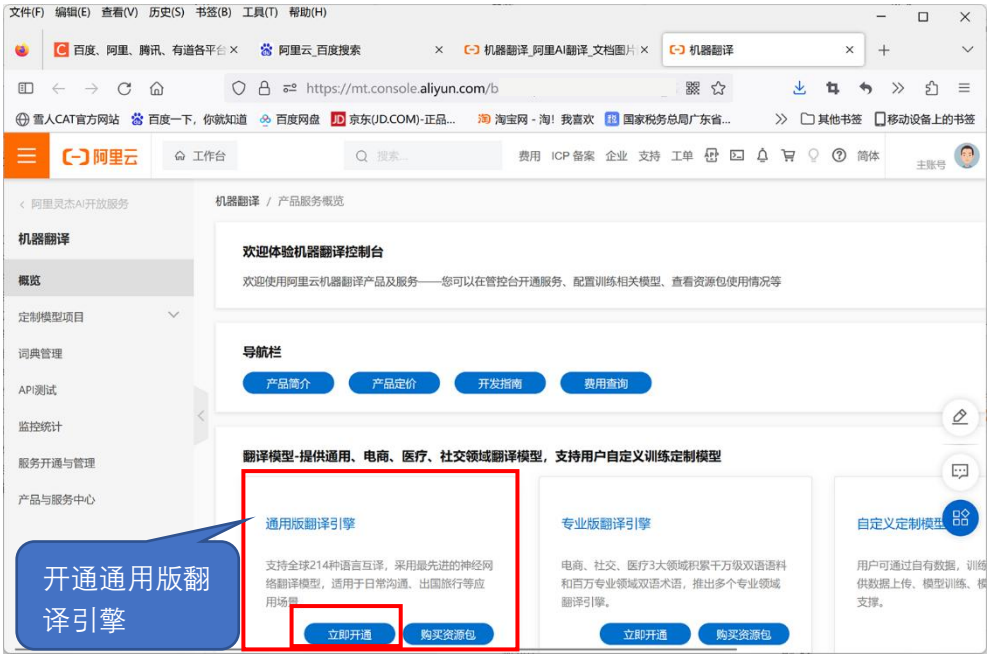
（一）、注册并登陆阿里云：https://account.aliyun.com，在云产品中，找到机器翻译。



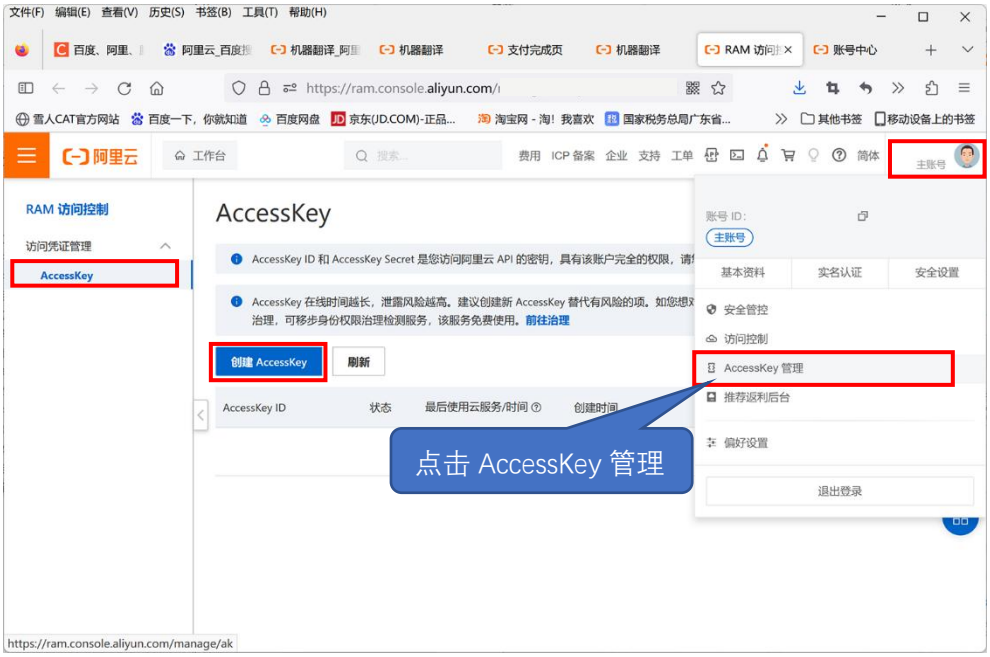
（二）、点击立即开通按钮。



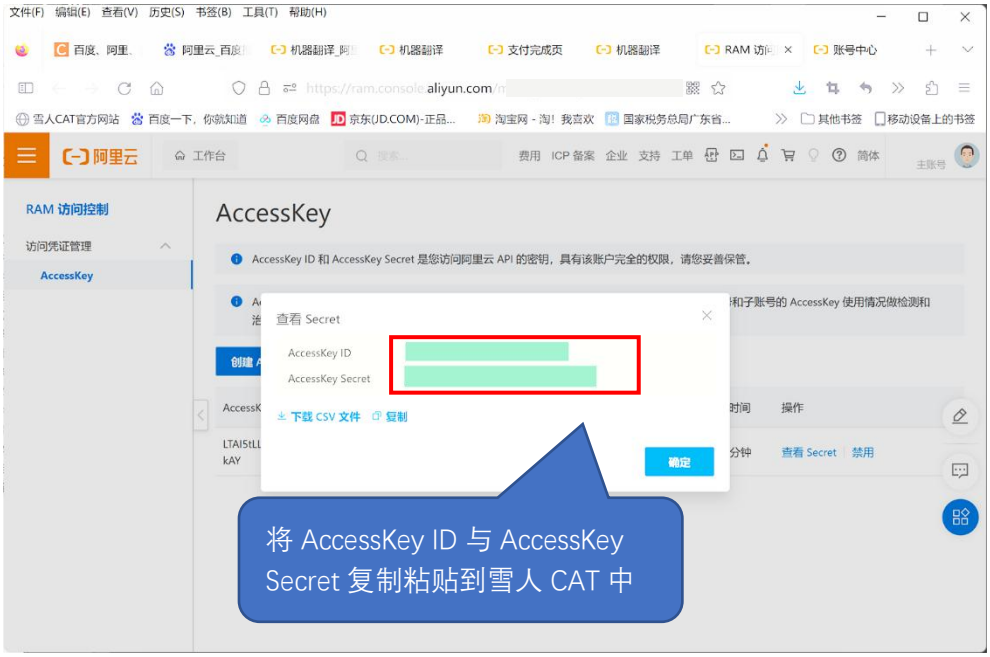
（三）、申请通用版翻译引擎



（四）、开通后，鼠标悬停在右上角用户头像，点击 AccessKey 管理进入控制台，然后点击创建 AccessKey 按钮。

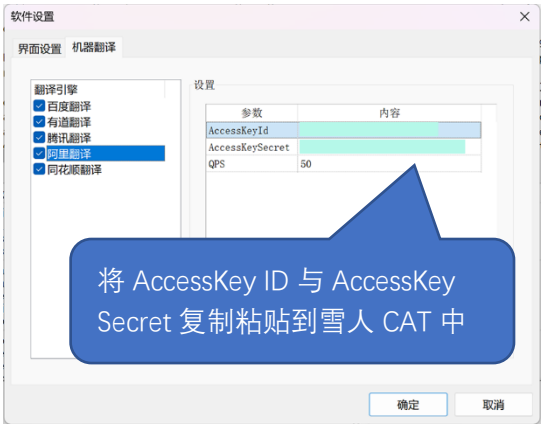


（五）、通过短信验证后，提示 AccessKey 创建成功，将 AccessKey ID 与 AccessKey Secret 复制粘贴到雪人 CAT 中。



将申请的 API 接入雪人 CAT

在雪人软件中，点击【**工具菜单➔软件设置➔机器翻译**】功能，在“机器翻译”设置窗口中，将阿里云翻译申请到 API 接口复制粘贴到此处，如下图。

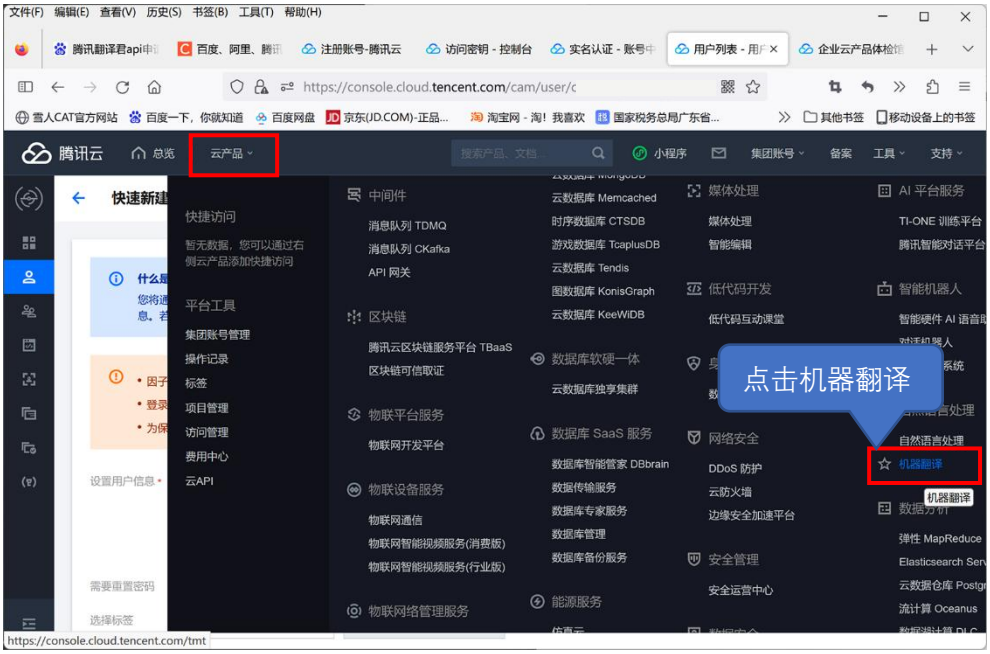


二、申请腾讯云翻译 API

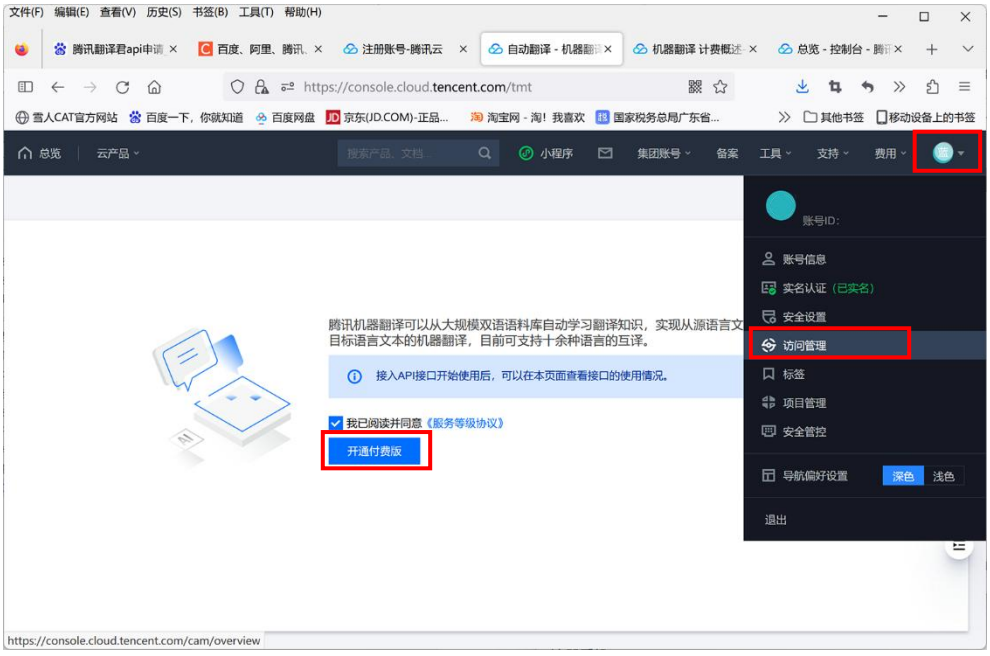
（一）、注册腾讯云账号 https://cloud.tencent.com/register



（二）、在云产品中，找到“机器翻译”。



（三）、点击开通付费版



（四）、开通后 鼠标悬停在右上角头像上，点击上图的“访问管理”。



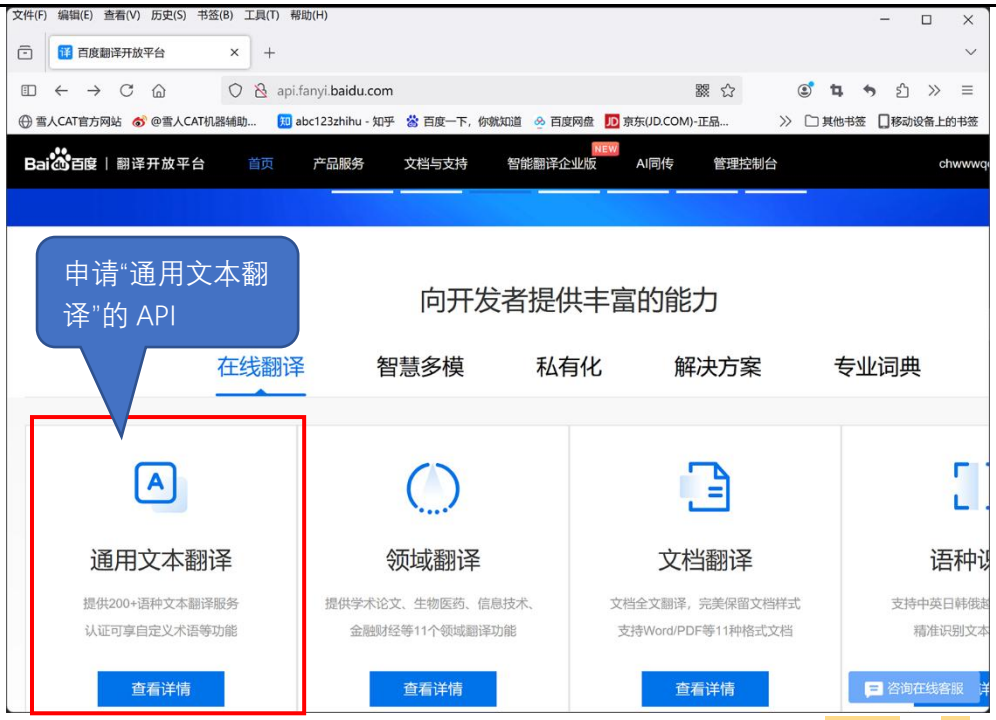
（五）点击上图左侧的"API 密钥管理"，再点击新建密钥按钮创建密钥。然后将上图显示出的 SecretId 和 SecretKey 复制粘贴到雪人 CAT 中。

三、申请百度云翻译 API

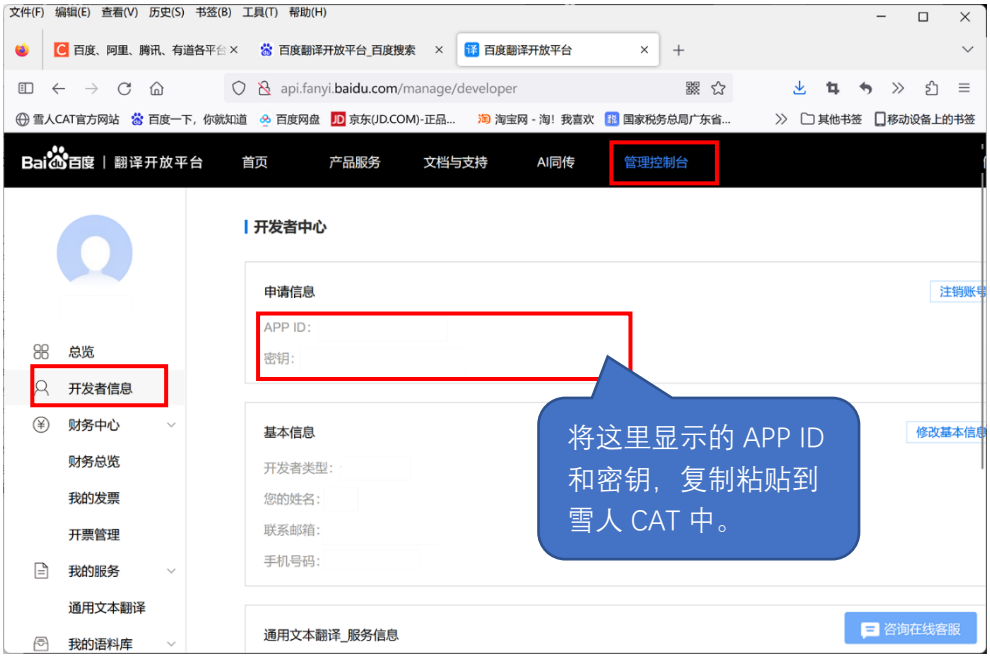
（一）、登陆百度翻译开发平台（http://api.fanyi.baidu.com），注册账号，可以注册为个人开发者或企业开发者。



（二）、申请“通用文本翻译”的 API

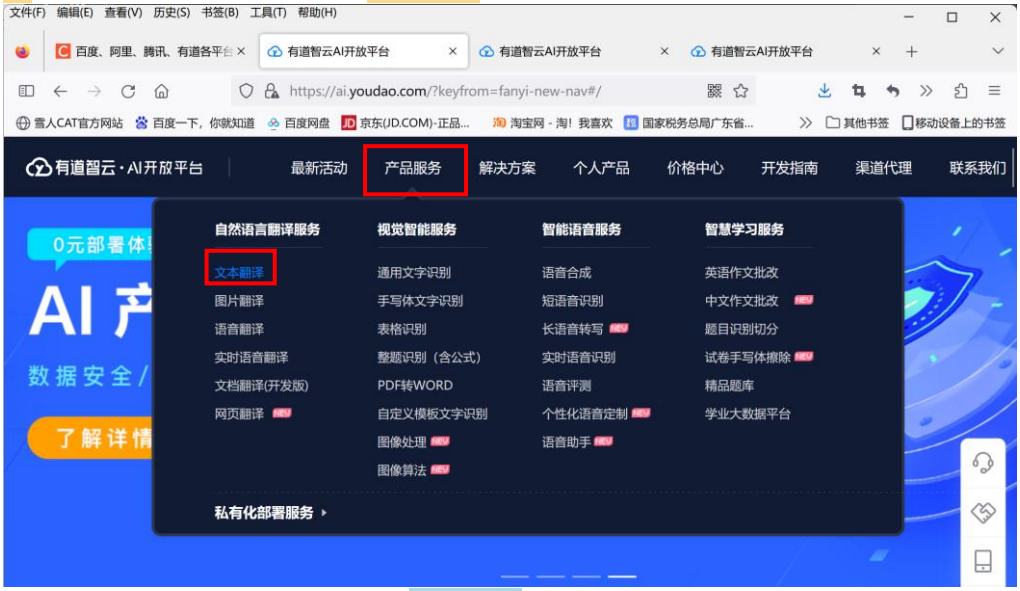


（三）、进入管理控制台，点击屏幕左侧的“开发者信息”，查看接口 APPID 与秘钥。将 APP ID 和密钥，复制粘贴到雪人 CAT 中。



四、申请有道翻译 API

（一）、注册并登陆有道智云 https://ai.youdao.com，在“产品服务”中，找到文本翻译。注册时，开发者名称填写文本翻译。



（二）、开通文本翻译服务。点击创建应用按钮，创建一个应用。

应用名称：写文本翻译，应用类别：选择实用工具，选择服务：选择文本翻译。开通后进入控制台，选择自然语言翻译服务下的文本翻译即可查看 APP ID 和秘钥，将 APP ID 和秘钥复制粘贴到雪人 CAT 中。



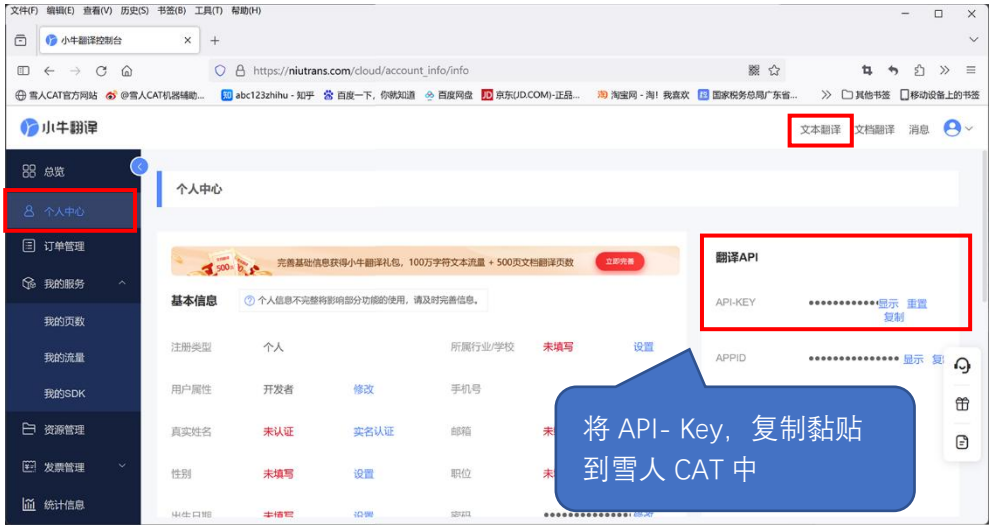
五、申请同花顺翻译 API

注册并登陆同花顺翻译网站，https://translate.10jqka.com.cn/translate，同花顺的注册比较简单，对经济、财经类的翻译有其长处，所以加入了它。点击导航栏“API”，创建一个文本翻译的应用，然后将 App ID 和 App Key 复制粘贴到雪人 CAT 中。



六、申请小牛翻译 API

登陆小牛翻译网站，https://niutrans.com/register ，注册账号并申请文本翻译 API-KEY，在控制台的“个人中心”中，可以查看申请的 API-KEY，将 API-KEY 复制粘贴到到雪人 CAT 中。



七、申请谷歌翻译 API

Google 翻译也同样要去谷歌站点申请 API，国内用户一般不能直接连上谷歌站点，需要使用代理服务器才能访问谷歌站点，请自行想办法申请到 API？除了需要谷歌的将 API key 复制粘贴到雪人中，国内用户还需要在“机器翻译”设置窗口的“proxy_server”栏中，输入代理服务器的地址及端口，例如：203.129.209.8:1080。

八、申请 DeepL 翻译 API

登录 DeepL 官方网站 https://www.deepl.com ，注册账号及申请 API key，申请 API 时需要绑定 DeepL 官方认可的国家的信用卡，如果没有，则只能考虑购买其他第三方注册好的 API key，或者在雪人的“AI 助手”中，使用网页界面的方式调用 DeepL 翻译，网页界面调用的方式不需要注册 DeepL 账号及申请 API key。



第四章 本地机器翻译引擎

第一节 支持本地机器翻译

DeepL 翻译、谷歌翻译、Bing 翻译及百度翻译等，均属于业界领先的在线机器翻译服务平台。使用这些在线翻译引擎，用户需访问其官方网站完成账号注册流程和获取 API 密钥（API Key）。使用在线翻译服务时，翻译数据需上传至它们的服务器，这对有严格保密需求的用户而言不太适用。

雪人 CAT 支持多种本地机器翻译引擎，为用户提供了额外的翻译解决方案。用户无需额外注册账号及申请 API 密钥，也无需担心翻译数据外泄，因为所有翻译数据均在用户的本地机器上完成。有条件的用户也可以使用自己的数据训练出来的翻译模型，在专门领域中获得更高质量的译文。

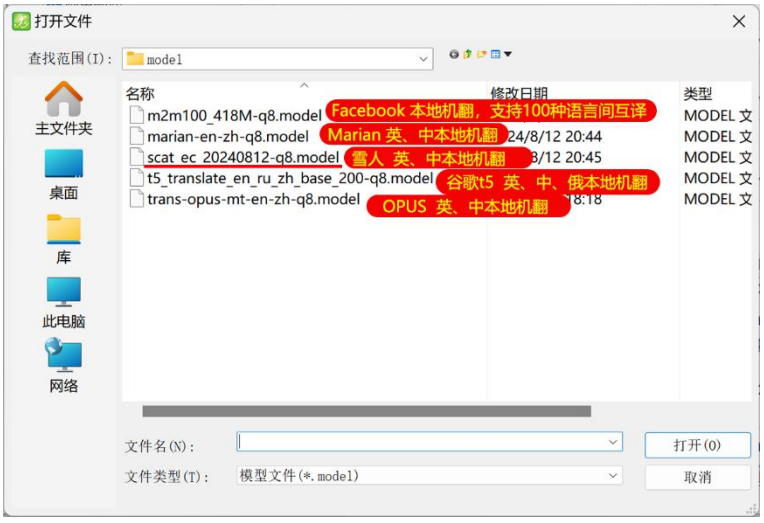
下面将介绍本地机器翻译引擎的配置及使用。点击**【工具】➔【软件设置】**菜单，如下图：



点击“软件设置”窗口中的“机器翻译”页签，如下图：勾选窗口左侧的“本地翻译”。



在窗口右侧的的“path”行处，双击鼠标，此时会出现下拉按钮，点击该按钮 ..., 选择本地机器翻译模型，如下图：



上图中，是目前雪人支持的本地机器翻译模型，用户可以去雪人的官网或者 QQ 群中下载这些模型，（群号：526362494 雪人翻译软件 AI 交流群、180692245 雪人翻译软件 6 群、126371112 雪人翻译软件 4 群）。以下是这些模型的介绍：

M2M-100 模型是 Facebook 推出的一个多语言机器翻译（MMT）模型，该模型具有在 100 种语言之间进行直接翻译的能力，而无需依赖英语作为中转语言。Facebook 构建了一个包含 75 亿个句子的多对多翻译数据集，该数据集涵盖了 100 种不同语言，为模型提供了丰富的训练资源。M2M-100 模型是 Facebook 在机器

翻译领域的一项重要成果，其多语言直接翻译的能力、高性能表现以及广泛的应用前景都使其成为该领域的佼佼者。“m2m100_418M-q8.model”是基于 M2M-100 微调的模型，支持 100 种语言之间的互译。

t5_translate：T5（Text-To-Text Transfer Transformer）模型是由谷歌研究团队提出的一种用于自然语言处理的预训练语言模型，在处理长文本和复杂语言结构时具有更高的效率和准确性。强大的预训练与微调机制，T5 模型能够高效、准确地完成各种翻译任务。“t5_translate_en_ru_zh_base_200-q8.model”是基于 t5_translate 微调的模型，支持英语、俄语、中文三个语言之间的互译。

trans-opus-mt 模型通常被认为是基于 OPUS 数据集训练的 Transformer 架构的多语言机器翻译模型。可以在不同的翻译任务和数据集上进行微调，以适应特定的翻译需求。“trans-opus-mt-en-zh-q8.model”是基于 trans-opus-mt 微调的模型，支持英语、中文的互译。

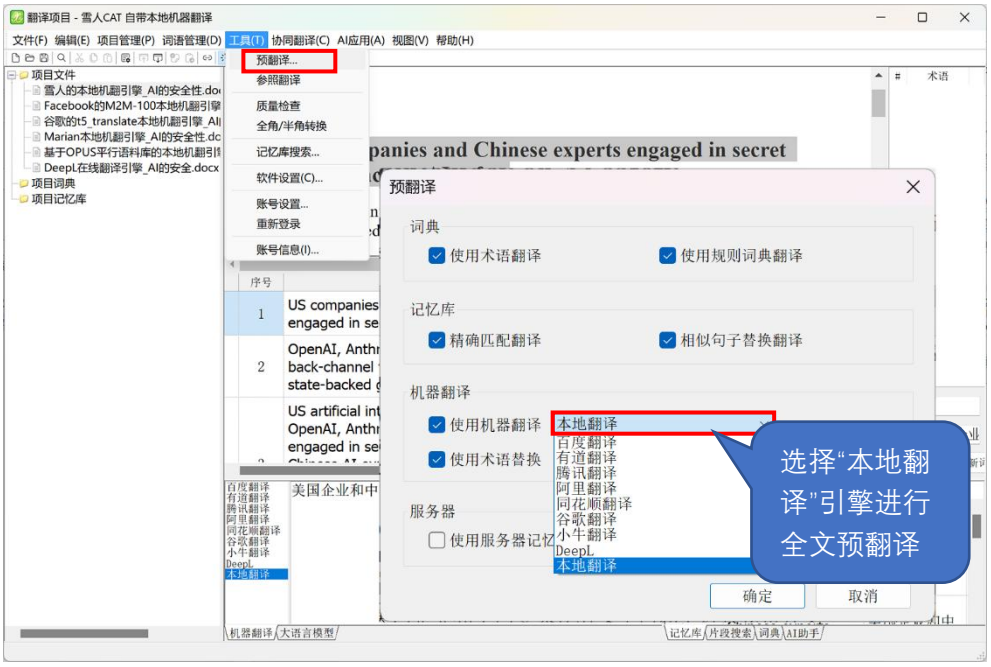
Marian（也称为 Marian-NMT）是一种专注于机器翻译（Machine Translation, MT）的神经网络模型。它基于深度学习技术，特别是 Transformer 架构，为自然语言处理（NLP）领域中的翻译任务提供了高效和强大的解决方案。“marian-en-zh-q8.model”是基于 Marian 微调的模型，支持英语、中文的互译。

“scat_ec_20240812-q8.model”是雪人计算机公司训练的基础翻译模型，雪人团队将持续对该模型进行迭代升级与性能优化。可根据用户提供的特定领域的记忆库，定制化训练出更加贴合该用户翻译习惯、风格、领域的本地翻译模型。特定领域翻译记忆库是指专注于某一特定领域（如法律、医学、科技、专利、军事、航天、船舶、汽车、金融财经、机械、建筑、客户的技术手册等）的翻译记忆库，通过深度学习技术，机器翻译模型能够从记忆库中学习特定领域的语言模式和专业术语，从而实现更精准、更符合行业标准的翻译输出。根据用户的记忆库进行特定训练，不仅充分挖掘了用户记忆库的价值，使得训练后的本地机器翻译的质量在某个特定的领域表现得更好，还可以使翻译结果更好地模拟用户的个性化文风，从而为用户提供更加精准和个性化的翻译体验。**雪人团队可以为有需要的企业用户或个人用户提供定制本地机器翻译引擎的训练服务。**

选择好具体的本地翻译模型后，在屏幕左下方的“机器翻译”窗口，点击“本地翻译”，就会调用所选择的本地翻译模型逐句进行翻译。如下图：



还可以利用“本地翻译引擎”进行全文的快速预翻译，即使遇上长文翻译也可快速完成。点击【**工具**】➔【**预翻译**】菜单，在“使用机器翻译”栏中，点击旁边的下拉按钮，选择“本地翻译”，如下图：



下图是一篇“关于 AI 安全性”的英文原文：



我们分别采用的了雪人的本地机翻引擎、facebook 的 m2m100 本地模型、谷歌 t5、Marian、OPUS 这五个本地机翻引擎进行翻译，并将各个模型的机翻译文导出并截图。同时，为了比较，也通过 API 的方式，调用 DeepL 来翻译。本地机翻和 DeepL 在线机翻的译文截图如下：



第二节 本地化机器翻译模型定制与训练服务

雪人公司致力于在自然语言处理（NLP）这一前沿领域的深化研发，积极引入并整合行业的先进技术，以全面赋能翻译工作，引领翻译服务迈向新高度。"scat_ec_20240812-q8.model"作为雪人计算机公司训练的基础翻译模型，该模型由雪人专业团队全程护航，雪人团队会不断地对它进行优化与性能迭代，使得模型在每一次迭代中都能实现翻译质量与效率的双重提升。

针对用户独特的翻译需求与特定行业领域（包括但不限于法律、医学、科技、专利文献、军事策略、航空航天、船舶工程、汽车制造、金融财经、机械工程、建筑设计，乃至客户专有技术手册等），雪人团队提供高度定制化的翻译模型训练服务。这一过程依托于先进的深度学习算法，能够精准捕捉并学习用户提供的特定领域翻译记忆库中的语言模式、专业术语及行文风格，进而构建出高度贴合用户偏好与行业标准的本地化翻译模型。

特定领域翻译记忆库，作为这一定制化流程的核心资源，其价值在于汇聚了领域内丰富的翻译经验与专业知识，为机器翻译模型提供了宝贵的学习素材。通过深度挖掘与智能分析，模型能够有效吸收并内化这些专业知识，从而在翻译过程中展现出更高的准确性、专业性和行业适应性。

更为重要的是，基于用户记忆库的定制化训练，不仅实现了对记忆库价值的最大化利用，提升了翻译结果的领域专属性与精准度，还进一步增强了翻译输出的个性化特征，使其能够模拟用户的独特文风与表达习惯。这一特性对于追求品牌一致性、信息传递精准度及用户体验优化的企业或个人用户而言，无疑具有不可估量的价值。AI 大语言模型，对中英、英中的效果普通不错，但对于小语种，效果则差一些。对于小语种，利用自己的记忆库，在基础模型的基础上训练的专有模型，效果可能会比普通的大模型好。

雪人团队可以为有需求的企业用户及个人用户定制本地机器翻译引擎训练服务。记忆库的使用价值不仅仅是作为句子匹配之用，还可以利用记忆库来定制训练自己的机器翻译模型，可以充分挖掘出用户记忆库的更大价值！我们坚信，通过双方的紧密合作与不懈努力，定能共创翻译技术的新篇章。

第五章 预翻译

第一节 全文预翻译

新版中有功能强大的全文预翻译，可以实现快速翻译。一篇待译文档，导入后在“项目设置”中添加好需要参考引用的“项目文件”或“记忆库”，在新版中支持直接将翻译的项目文件（.SPR ）作为记忆库来引用（具体请参看“直接引用翻译的项目文件作记忆库”），设置这些的目的是：在进行全文预翻译时，优先选用用户之前的翻译资源找到精确匹配的内容填写，找不到匹配的内容，则交给机器自动翻译完成。并在“项目词典”中设置好术语，设置术语的好处是：若整句可以匹配术语的，则用术语的结果填写，若句子中存在术语的，则利用术语的词义去修正机器自动翻译的结果，提升机器自动翻译的译文质量。点击**【工具】➔【预翻译】**菜单，会弹出“预翻译”窗口，如下图：



各个选项解释如下：

词典中，有两个选项，分别是：使用术语翻译、使用规则词典翻译

使用术语翻译——如果存在整句是术语的句子，则直接用术语的翻译作译文填写，并在译文旁边的“备注列”中，写明来源是“TERM”。

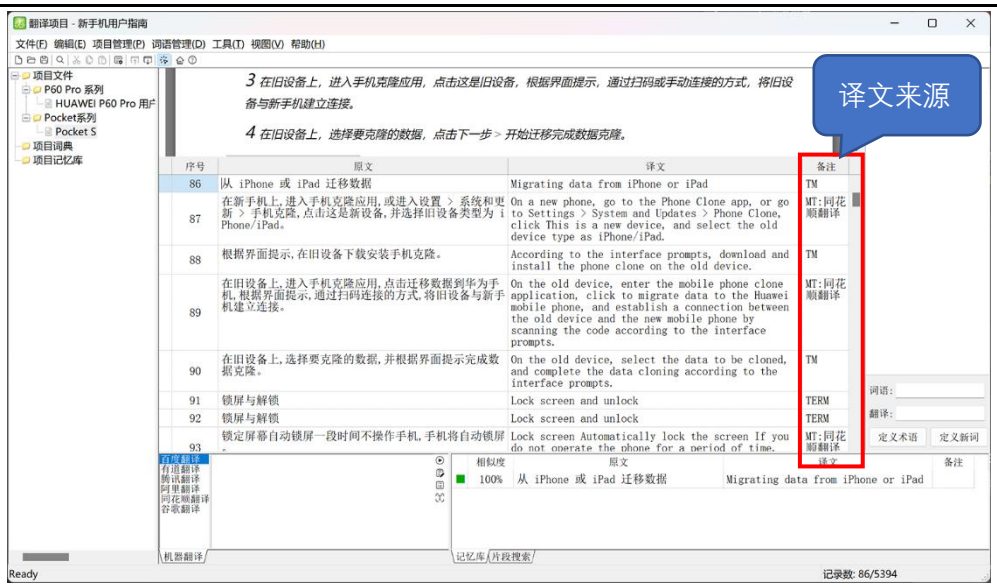
使用规则词典翻译——对于能套用用户在正则表达式词典中定义的翻译，则调用正则表达式词典的翻译结果填写译文，并在译文旁边的“备注列”中，写明来源是“RULL”。例如对于以下非译成分的定义，非译成份的翻译等等。

记忆库中有两个选项，分别是：精确匹配翻译、相似句子替换翻译。
精确匹配翻译——和记忆库中 100%相似的句子，自动填写译文。
相似句子替换翻译——和记忆库中结构相类似的句子，根据可以替换的部分进行替换翻译后自动填写译文。
记忆库来源包括：当前项目文件下面的所有已经翻译过的句子、当前“项目记忆库”中保存的句子、在项目设置中添加引用的.SPR 的项目文件、添加引用的记忆库文件。其优先级是：100%精确匹配且上下文都匹配的句子最优先被选用，次之选用 100%匹配且上文匹配或下文匹配的句子。当前项目文件夹下面同一篇翻译文档中的句子最优先，其次是同层次的文件夹下面的其他翻译文档，再次是同项目文件下面的其他文件夹里的翻译文档，再次是当前项目记忆库中的句子、再次是添加引用的项目文件，再次是添加引用的记忆库文件。

机器翻译中有 3 个选项，分别是：使用机器引擎选项、使用本地术语库、模拟批量翻译选项。
使用机器翻译——目前提供百度翻译、有道翻译、腾讯翻译、阿里翻译、同花顺翻译、谷歌翻译引擎供选择。欲使用这些机器翻译引擎，应先去各大引擎站点注册账号，并将 APPID 和 KEY 复制黏贴在雪人的“软件设置”——“机器翻译”窗口中，设置好后，此处才可以出现下拉窗口供选择。具体账号申请及设置方法详见“机器翻译”章节。
使用本地术语库——利用本地术语库中定义的词义替换机器翻译中相应术语的翻译，修正机器翻译的结果。
模拟批量机器翻译——由于有些机器翻译引擎的 QPS 允许值限制的比较低，QPS 是机器翻译引擎允许用户 1 秒钟发送的翻译请求次数，有些 QPS=1，即 1 秒只允许发送 1 个翻译请求，发送多了就会报出错，这样完成全文预翻译就需要等待较长的时间，此时勾选“模拟机器翻译”选项，可以加快机器翻译的速度，目前如果选用“百度翻译”和“腾讯翻译”可以勾选这个选项。

服务器中有：使用服务器数据库选项。这个是对接雪人协同翻译平台服务器中的数据接口，目前没有开放。

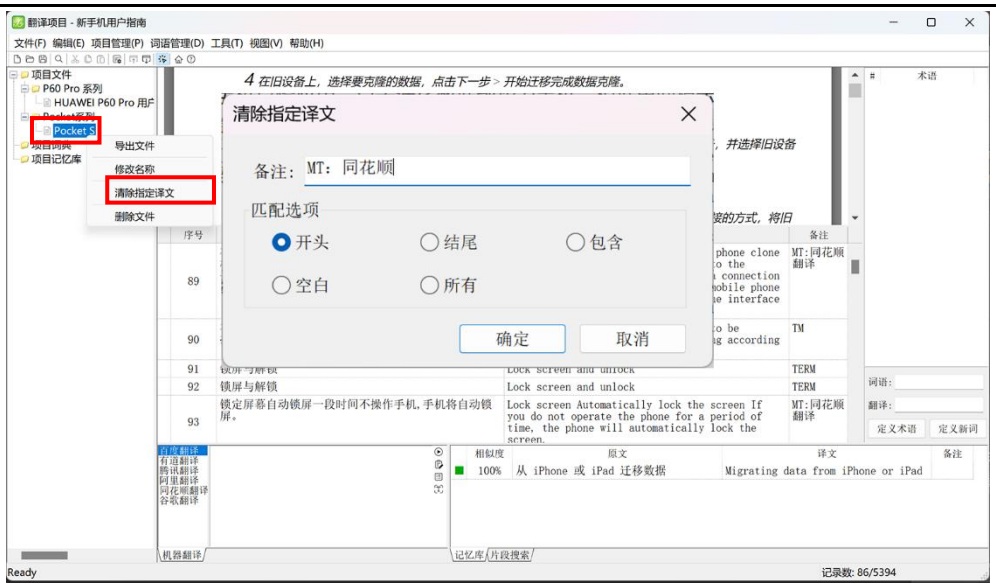
设置好预翻译的相关参数后，点击确定按钮，即进行全文预翻译，如果术语、记忆库中有匹配的，则优先选用这些的结果填写译文，如果找不到匹配的就发送给机器翻译引擎进行翻译，在调用机器翻译引擎时，可能需要等待一些时间，根据选用的机器翻译引擎响应速度不同，等待的时间有所不同。预翻译完成后，在“备注”列中会显示该句译文的来源，如下图:



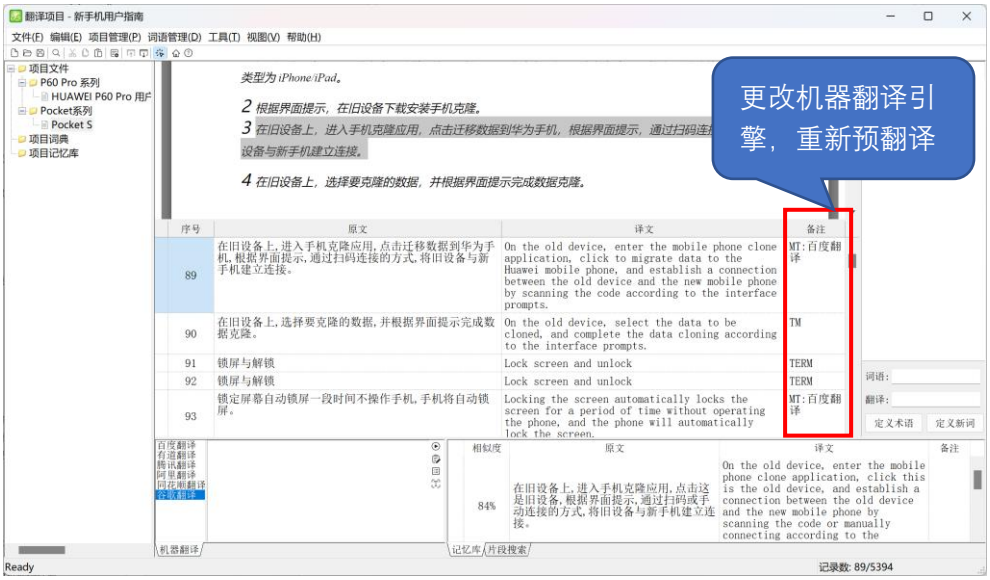
译文来源解释如下：

MT:机器翻译引擎——MT 表示机器翻译的结果，如 MT:百度翻译，表示是使用百度翻译引擎的机器翻译译文。
TM——来自记忆库 100%精确匹配的译文。
TM+——来自记忆库 100%精确匹配的译文，并且提示记忆库中存在多个精确匹配的译文版本，点击该句后，在下方的“记忆库”提示区中查看其他几条译文。
TM*——参考记忆库中结构相似的句子进行替换翻译后填写的译文。
TERM——直接使用术语翻译出来的译文。
TERM+——直接使用术语翻译出来的译文且存在多个译文版本。
RULE——使用正则表达式词典匹配翻译出来的译文。
RULE+——使用正则表达式词典匹配翻译出来的译文且存在多个译文版本。

如果对某个机器翻译引擎的翻译结果不满意，可以将其译文清除后，再次选择其他的机器翻译引擎重新预翻译。例如下图中，如果对所选用的“同花顺翻译”引擎的翻译结果不满意，可以将“同花顺翻译”的译文删除后，重新调用其他机器翻译引擎进行预翻译。具体方法：用鼠标右键点击屏幕左侧的某翻译文档后，点击**【右键菜单】➔【清除指定译文】**菜单，会弹出“清除指定译文”窗口，如下图：

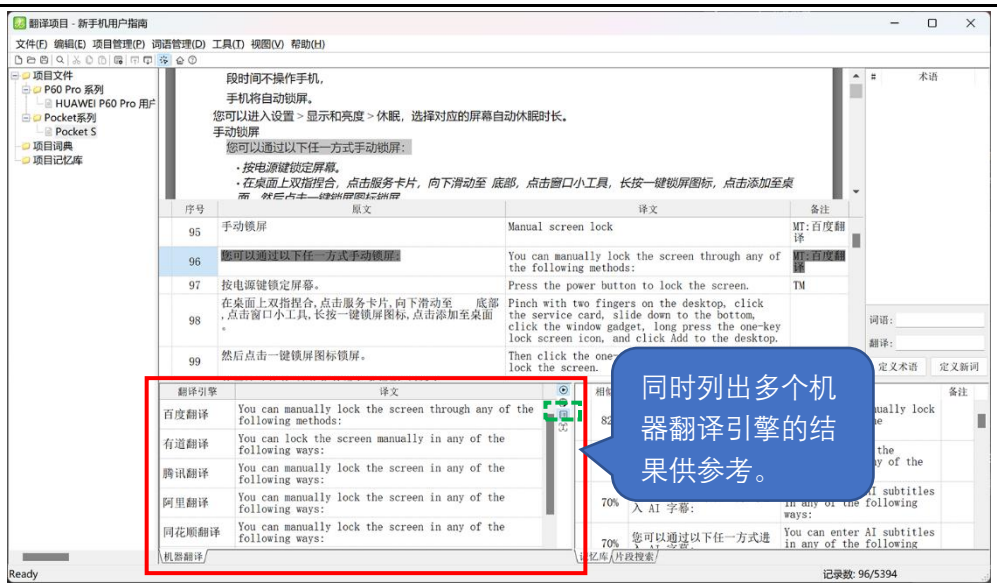


在匹配选项中，点选“开头”选项，在备注栏中输入“MT：同花顺”，表示删除备注以“MT：同花顺”开头的句子的译文。当然，也可以利用其他匹配选项，删除不想要的译文，或者删除所有。删除指定译文后，再点击**【工具】➔【预翻译】**菜单，选择其他翻译引擎重新进行预翻译，例如：选择百度翻译重新预翻译，结果如下：



预翻译或重新预翻译，只影响译文为空的句子，如果已经填写了译文的句子，预翻译不会影响它。

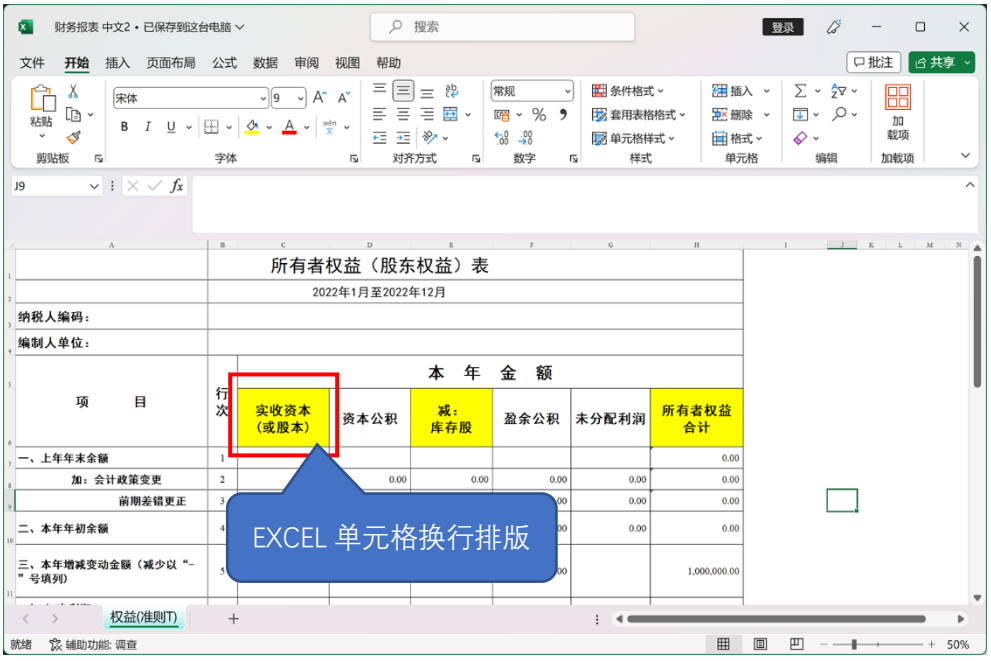
预翻译时，只能选择其中的一个翻译引擎进行，如果还想比较一下其他翻译引擎的结果，则可以在屏幕下方的“机器翻译”区中，将其他的多个机器翻译引擎的结果同时列出来供参考选用。如下图：


翻译项目 - 新手机用户指南
文件(F) 编辑(E) 项目管理(P) 词语管理(D) 工具(T) 视图(V) 帮助(H)
项目文件
P60 Pro 系列
HUAWEI P60 Pro 用户
Pocket系列
Pocket S
项目词典
项目记忆库
段时间不操作手机，
手机将自动锁屏。
您可以进入设置 > 显示和亮度 > 休眠，选择对应的屏幕自动休眠时长。
手动锁屏
您可以通过以下任一方式手动锁屏：
·按电源键锁定屏幕。
·在桌面上双指捏合，点击服务卡片，向下滑动至 底部，点击窗口小工具，长按一键锁屏图标，点击添加至桌
序号
原文
译文
备注
95
手动锁屏
Manual screen lock
MT:百度翻译
96
您可以通过以下任一方式手动锁屏：
You can manually lock the screen through any of the following methods:
MT:百度翻译
97
按电源键锁定屏幕。
Press the power button to lock the screen.
TM
98
在桌面上双指捏合，点击服务卡片，向下滑动至 底部，点击窗口小工具，长按一键锁屏图标，点击添加至桌面。
Pinch with two fingers on the desktop, click the service card, slide down to the bottom, click the window gadget, long press the one-key lock screen icon, and click Add to the desktop.
99
然后点击一键锁屏图标锁屏。
Then click the one
lock the screen.
术语
词语:
翻译:
定义术语
定义新词
翻译引擎
译文
百度翻译
You can manually lock the screen through any of the following methods:
有道翻译
You can lock the screen manually in any of the following ways:
腾讯翻译
You can manually lock the screen in any of the following ways:
阿里翻译
You can manually lock the screen in any of the following ways:
同花顺翻译
You can manually lock the screen in any of the following ways:
机器翻译
记忆库
片段搜索
70%
入 AI 字幕:
70%
您可以通过以下任一方式进
You can enter AI subtitles in any of the following
同时列出多个机器翻译引擎的结果供参考。
Ready
记录数: 96/5394

第二节 参照翻译

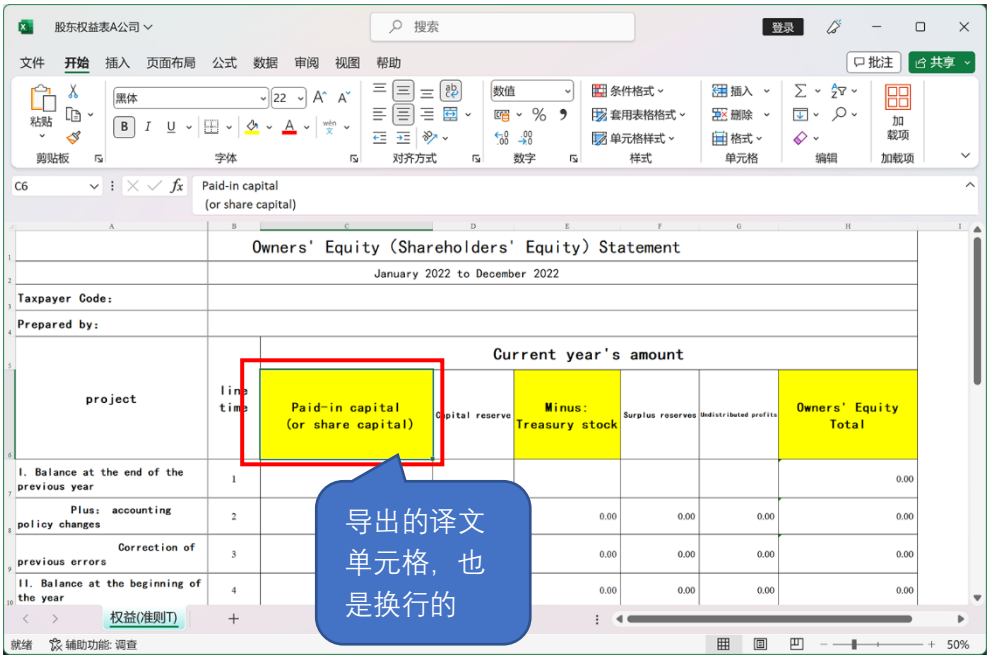
参照翻译：参考现有已经翻译好的文件翻译另一个文件，同时保留参考文件的样式，即以参考文件为模板翻译新文件，复制参考文件的排版样式，不是使用记忆库翻译，参照翻译和记忆库无关，因为记忆库的句子是没有样式且是完整的一句或几句，中间没有换行符等信息的，记忆库翻译实现不了模板样式的套用式翻译的需求。参照翻译对于翻译格式要求严格一致的 EXCEL 表格，或多版本技术文档的更新升级等，会比记忆库翻译更有优势，它是直接套用旧的文档为模板翻译新文档，可以非常省事地保持新旧文档风格和排版的一致。

例如：下图的 EXCEL 原文（股东权益表 A 公司）中，“实收资本（或股本）”这个单元格是进行了换行排版的，即在“实收资本“和“（或股本）”之间换行。要求译文也需要这样换行排版。





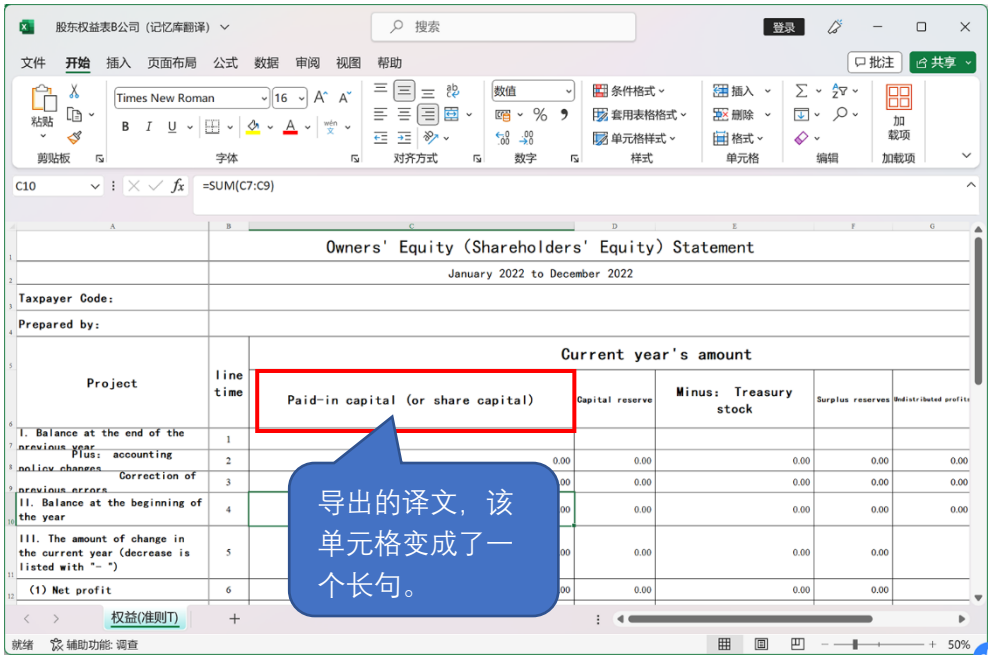
导出的 EXCEL 译文，这个单元格也是和原文一样进行了换行。



现在，以“股东权益表 A 公司”这个 EXCEL 文件，来翻译“股东权益表 B 公司”文件，将它导入雪人后，当光标去到“实收资本（或股本）”这个单元格时，下方的“记忆库”窗口，显示出 100%精确匹配的句子，双击即可应用在译文中，但是丢失了样式。如下图：

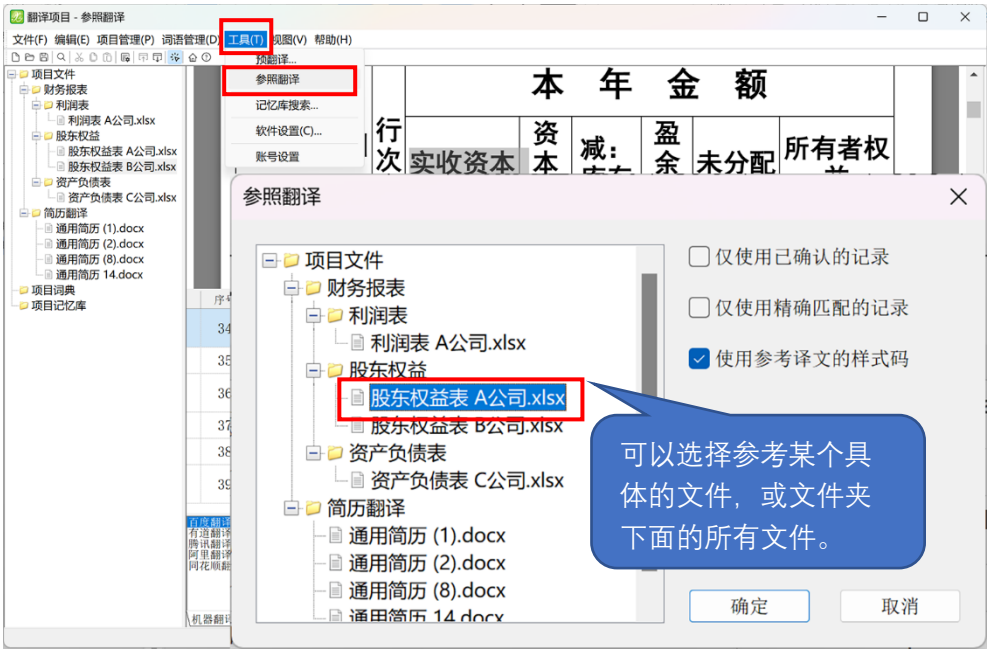


导出的译文，该单元格变成了一个长句，这对版式要求严格的表格就达不到要求了。

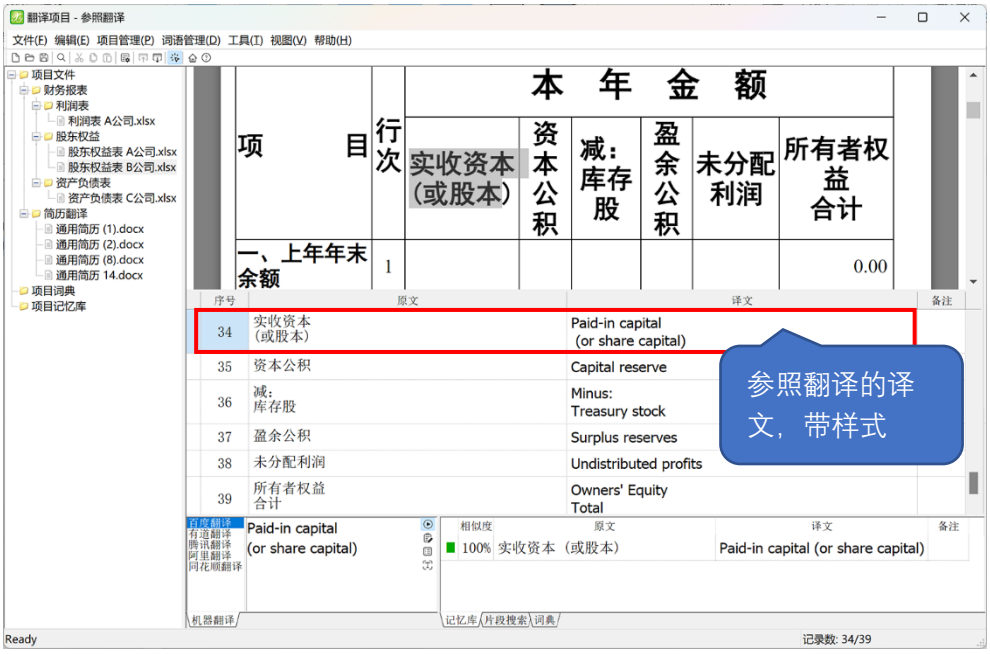


如果不使用传统的记忆库翻译模式，而是使用“参照翻译”，即以“股东权益表 A 公司”这个 EXCEL 文件为模板，来翻译“股东权益表 B 公司”文件，导入“股东权益

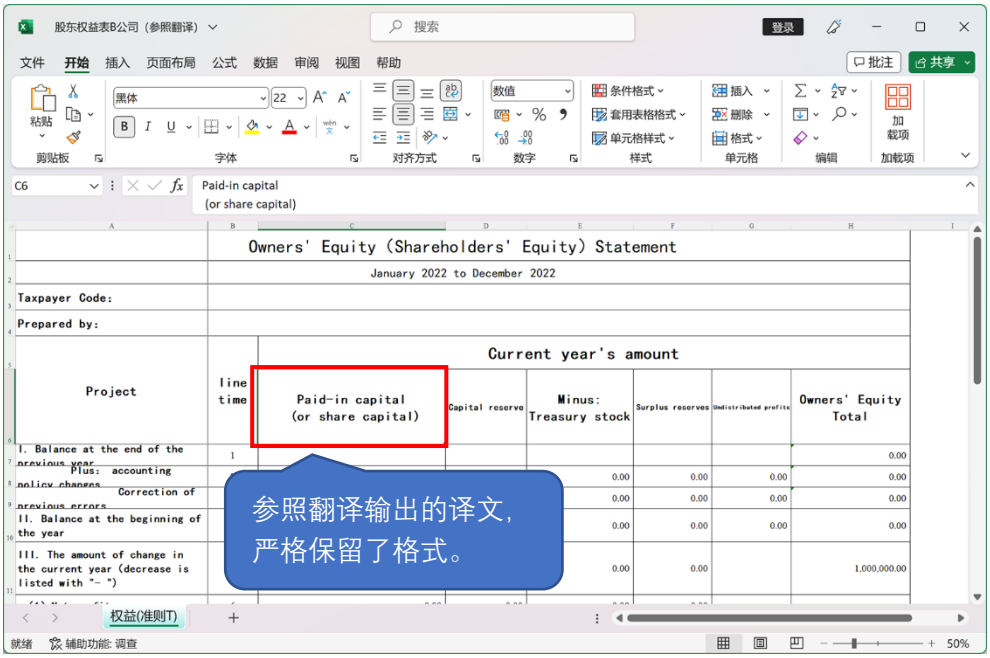
表 B 公司”文件后，点击【**工具菜单➔参照翻译**】功能，在弹出的“参照翻译”窗口中，选择“股东权益表 A 公司”这个文件。



就会以“股东权益表 A 公司”这个 EXCEL 文件为模板来翻译“股东权益表 B 公司”这个文件，对“实收资本（或股本）”这个单元格，参照翻译的译文也自动进行了换行处理，即复制了参照翻译的译文格式过来。



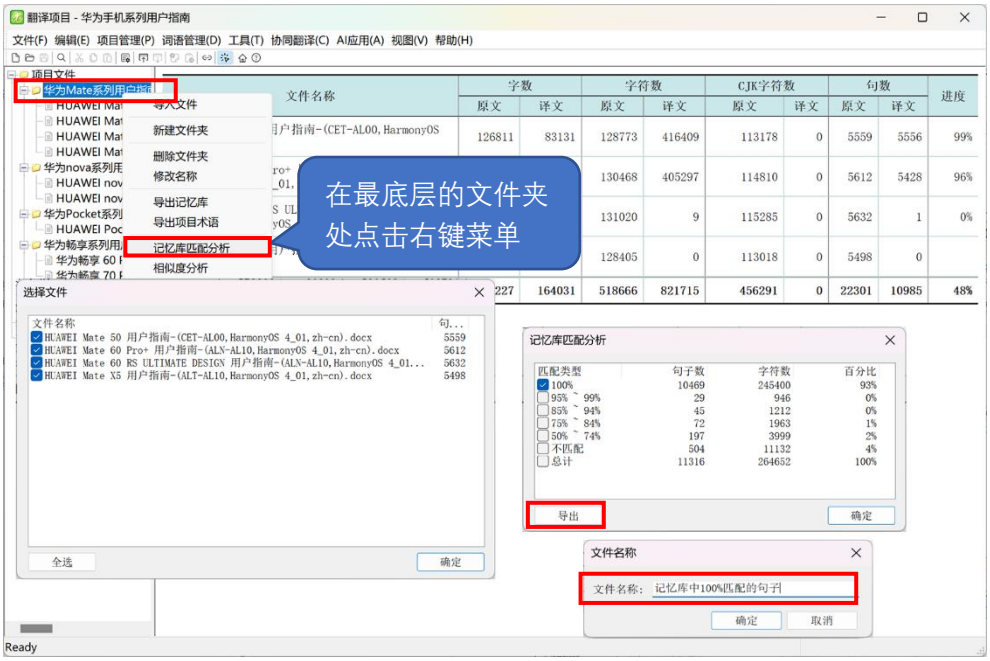
导出的译文，如下图：参照翻译输出的译文，严格保留了格式。



第三节 记忆库匹配分析

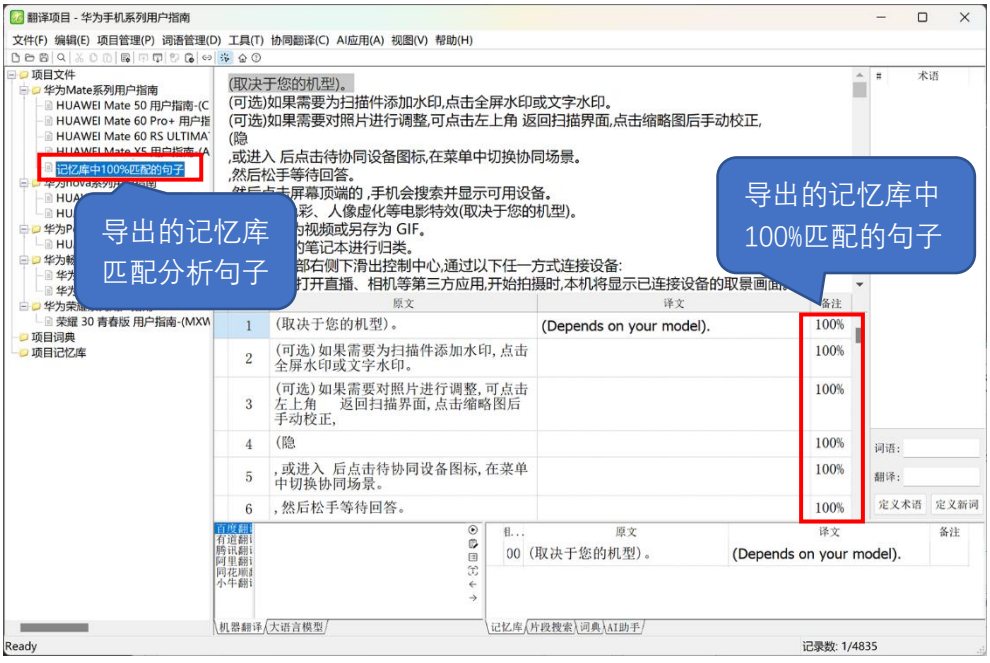
“记忆库匹配分析”是将选择的文档与记忆库中的句子进行比较，统计出待译文档中与记忆库的匹配程度，按 100%、95%、50%、不匹配等多个匹配程度给出计算结果，并可以导出其相应的句子。

选择屏幕左侧项目文件下的相应的文件夹，再点击**【右键菜单】➔【记忆库匹配分析】**功能，会弹出“选择文件”窗口。注意：如果左侧是多层次的复杂的项目文件结构，先点击最底层的文件后，再点击右键菜单，不支持在中间层次的“子文件”位置上点击右键菜单“记忆库匹配分析”。如果点击某个具体的翻译文档后再点击**【右键菜单】➔【记忆库匹配分析】**功能，则是对当前文档进行记忆库匹配分析。如下图：



选择需要进行记忆库匹配分析的文档（可以单选或多选），如上图：点击“确定”按钮，就可看到分析结果。

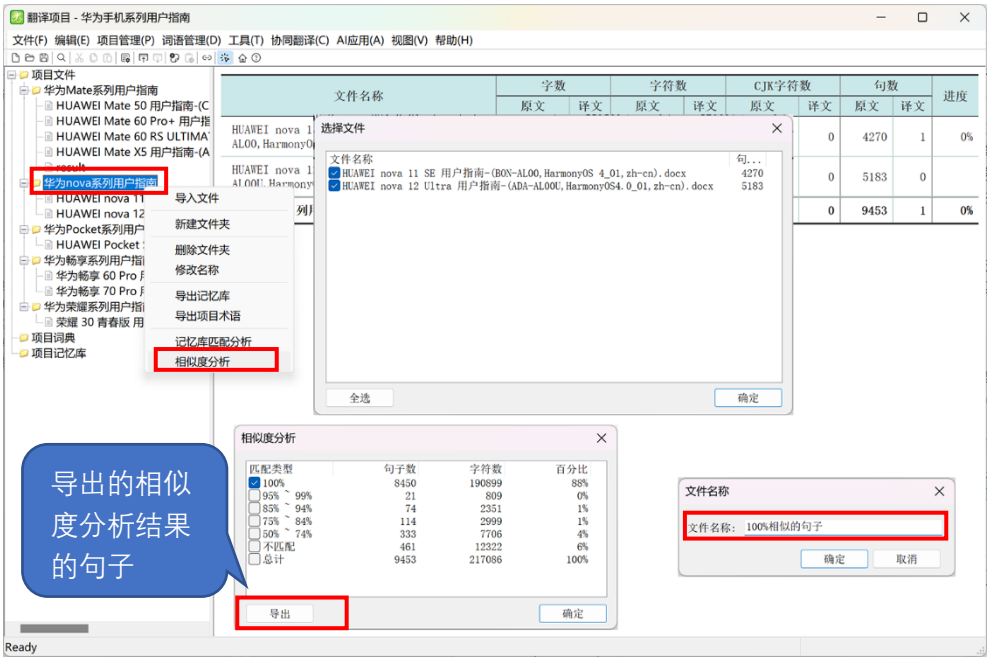
选择具体的匹配率后，点击“导出”按钮，在屏幕左侧的该文件夹下面将会生成一个新的文件。可以根据需要导出不同匹配率的句子，也可以导出不匹配的句子。如下图：



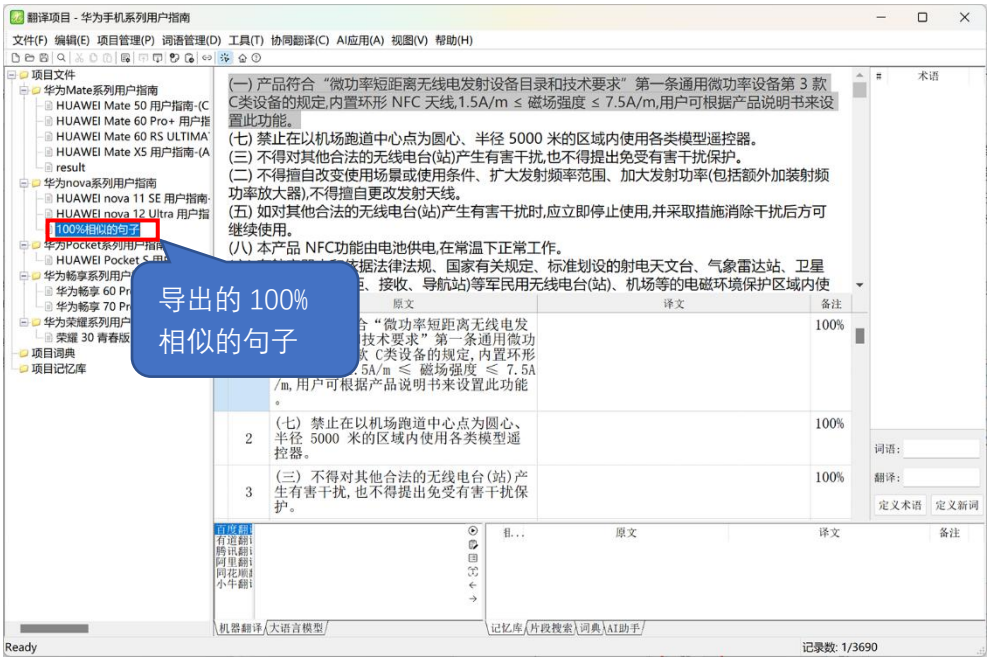
第四节 文档相似度分析

“相似度分析”是对文档内部或多份文档之间存在的相似句子进行比较、分析。例如对技术文档的翻译，可以将它与同类型的其他技术文档进行比较，统计出完全相同的句子、50%以上相同的句子、不相同的句子，将这些句子按需导出以便做进一步的处理。

选择屏幕左侧项目文件下的相应的文件夹，再点击**【右键菜单】➔【相似度分析】**功能，会弹出“选择文件”窗口。注意：如果左侧是多层次的复杂的项目文件结构，要先点击最底层的文件后，再点击右键菜单，不支持在中间层次的“子文件”位置上点击右键菜单“记忆库匹配分析”。如果点击某个具体的翻译文档后再点击**【右键菜单】➔【相似度分析】**功能，则是对当前单个文档内部的相似情况进行分析。如下图：



勾选具体的匹配率后，点击“导出”按钮，可以将它导出到屏幕左侧的该文件夹下面。如下图：

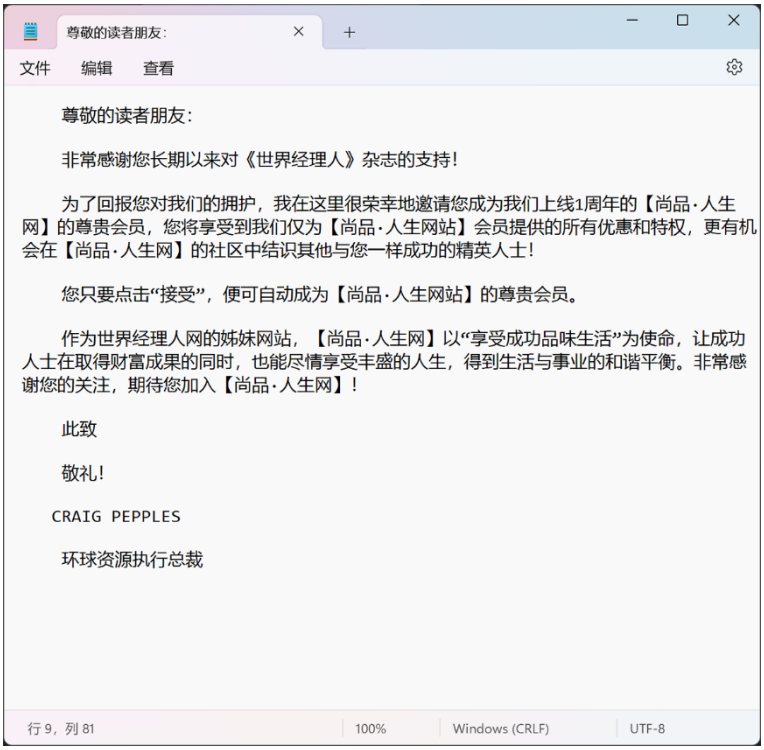


通过“相似度分析”和“记忆库匹配分析”，译员可以很好地了解待译文档的难易程度，给翻译项目报价和估计完成的时间等方面提供了很好的参考。

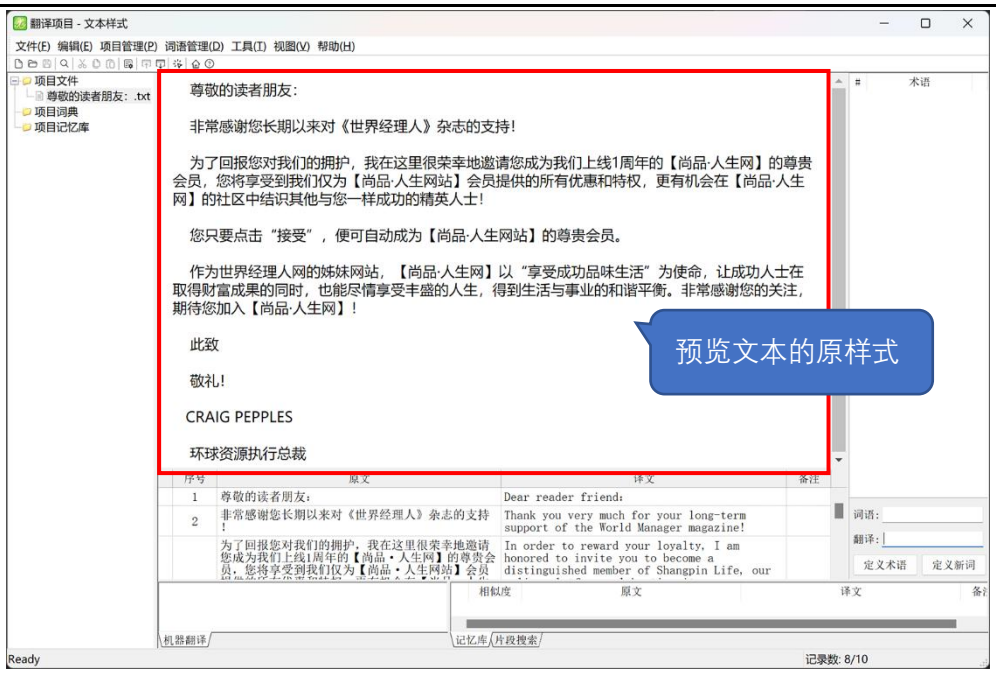
第六章 翻译各种格式的文档

第一节 翻译 TXT 格式的文档

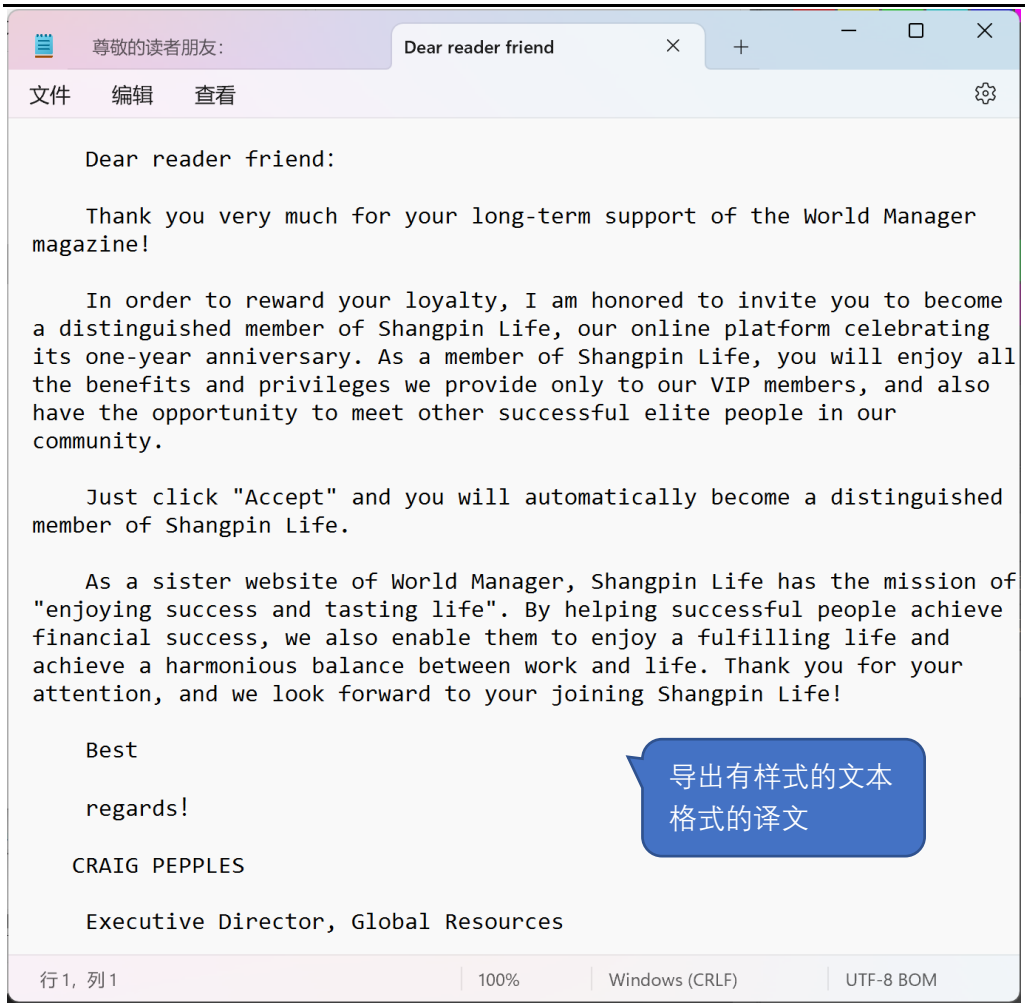
新版加强了对文本格式文件的支持，旧版的文本文件是以句子为单位，不支持段落、排版缩进等格式的。新版的文本文件也可以有样式，例如下图的文本格式的电子邮件：



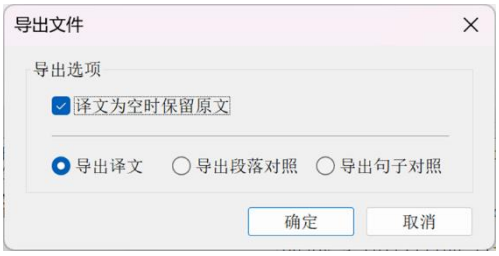
将它导入雪人 CAT 中，在预览区中也保持了原来的文本样式，如下图：



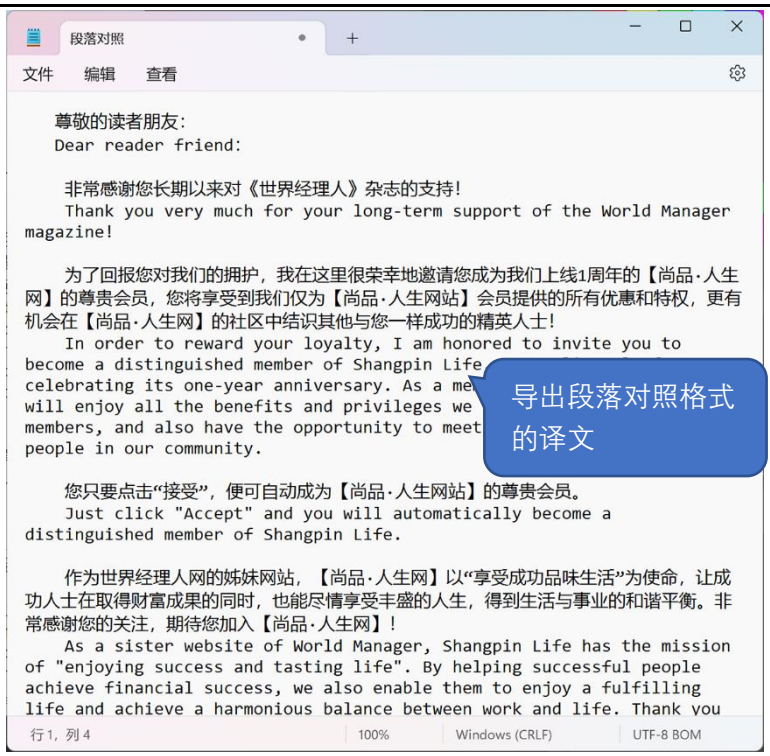
翻译完成后，右键点击屏幕左侧的该项目文件，点击**【右键菜单】➔【导出译文】**，导出的译文格式可以恢复原文的文本格式。如下图:



导出译文时，会弹出如下对话框，若勾选“译文为空时保留原文”选项，对没有翻译的句子，导出时将会用原文填充；如果不勾选，则译文将为空。

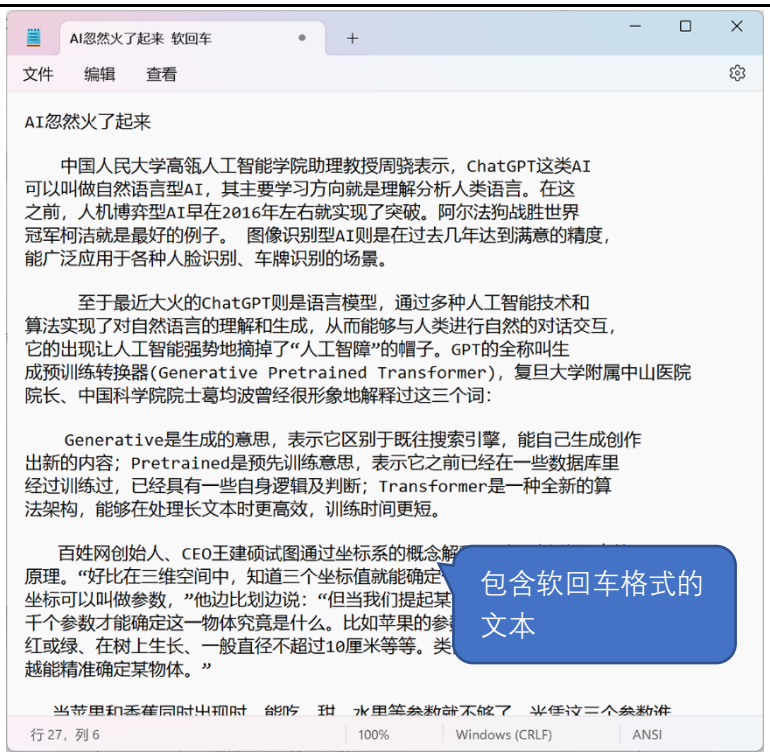


勾选“导出段落对照”选项，可以导出一段原文、一段译文的对照格式。如下图：勾选“导出句子对照”选项，可以导出一句原文、一句译文的对照格式。

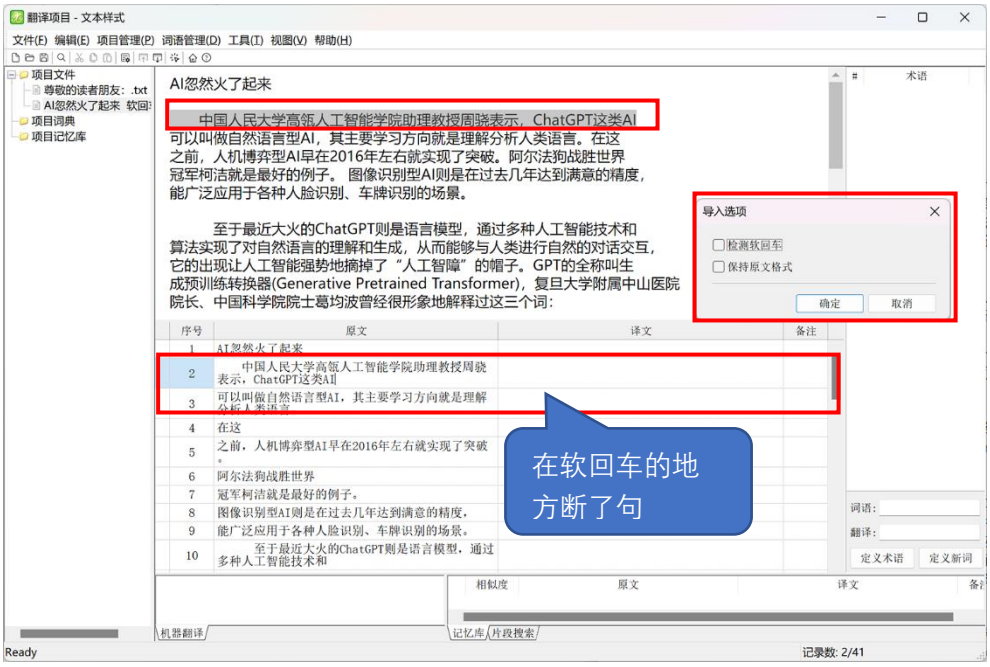


对软回车的支持

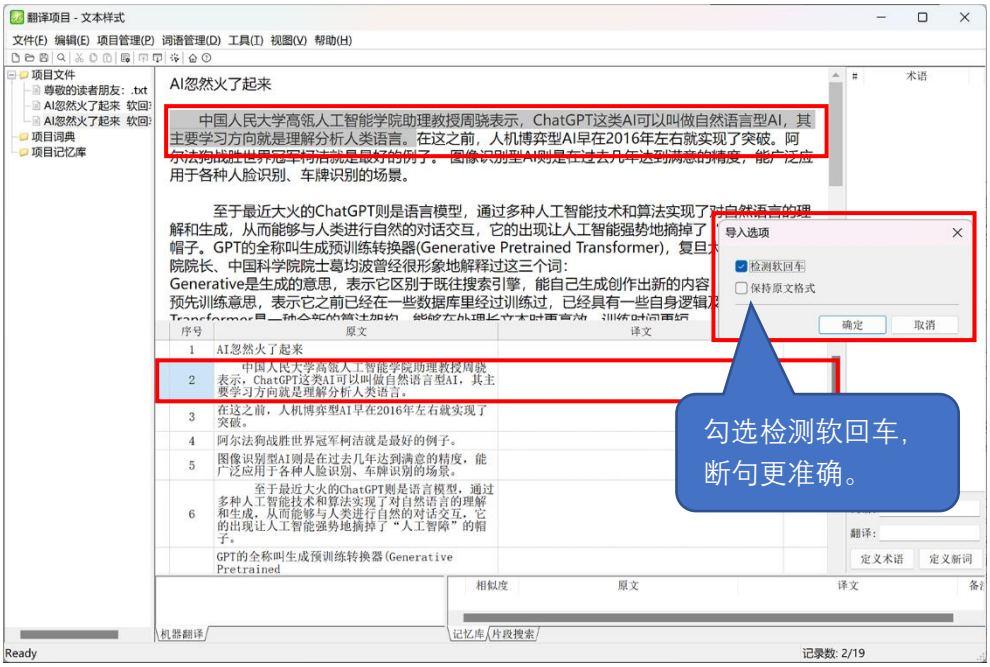
有时遇到的文本文件，包含许多的软回车，例如下图:



如果导入时，在弹出的“导入选项”窗口中，不勾选“检测软回车”选项，在软回车的地方会作断句处理，由于软回车影响了断句的准确性，导入结果如下图：



如果导入时，勾选“检查软回车”选项，在软回车的地方不作断句处理，断句就更准确，导入结果如下图：

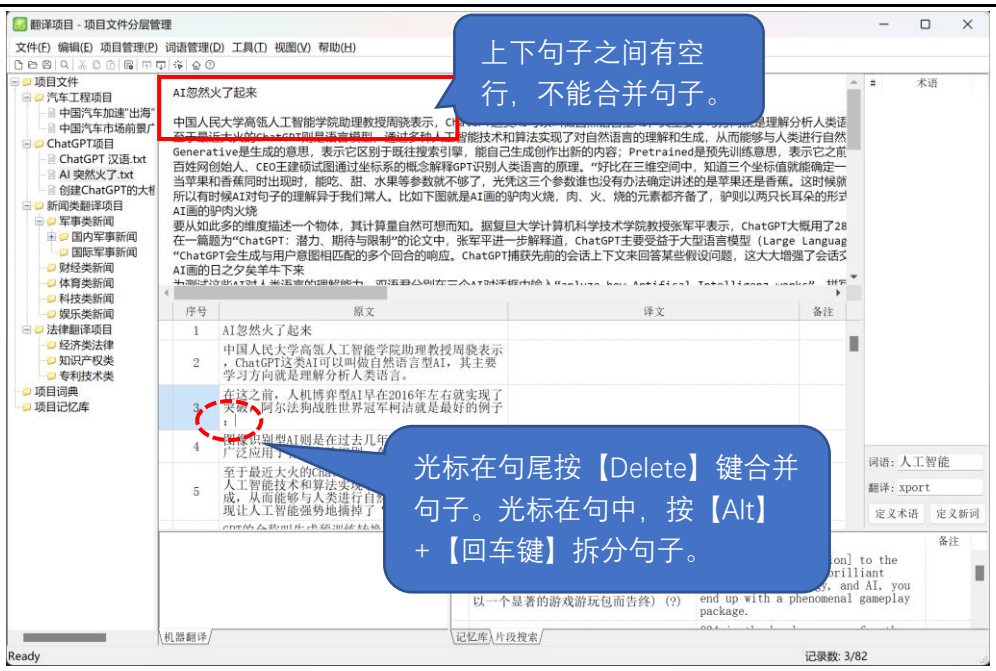


句子的合并与拆分

如果默认断句不准确时，可能需要对原文进行拆分句子或合并句子的操作。

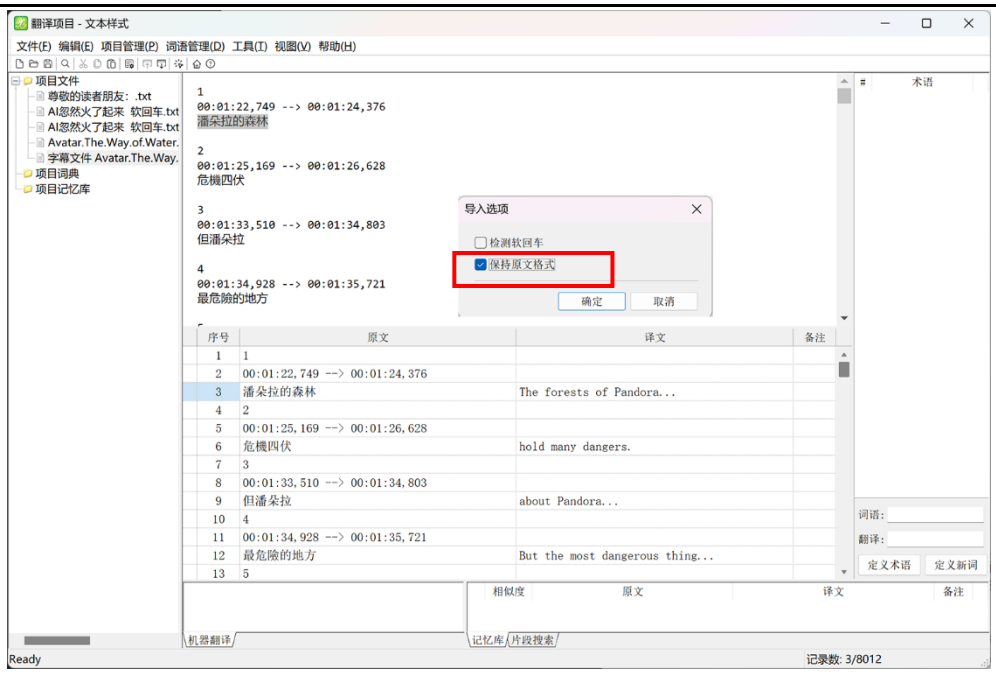
拆分句子：将光标移动到需要拆分的原文处，按**【Alt】+【回车键】**。新版拆分句子需要按【Alt】键，和旧版的操作不一样。

合并句子：将光标移动到句尾，按**【Delete】键**，将下一行的句子合并到上一行。注意：如果上、下句子之间有空行，则不能合并句子。例如下图中，第一句“AI 忽然火了起来”是标题，和第二句之间存在一个空行，此时在第一句的句尾按按按**【Delete】键**，是不能将第二句合并上来的。



字幕文件的翻译

Srt 格式，全称是“SubRip Text”格式；因其制作规范简单，一句时间代码+一句字幕，是文本字幕最为流行格式。可以将 Srt 格式的字幕文件另存为.txt 后，导入雪人 CAT 中，导入时选择“保持原文格式”选项，导入后时间轴代码全部留空不翻译，仅翻译字幕部分，如下图：译好导出时选择“译文为空时保留原文”选项，就可以严格恢复原来的字幕格式。

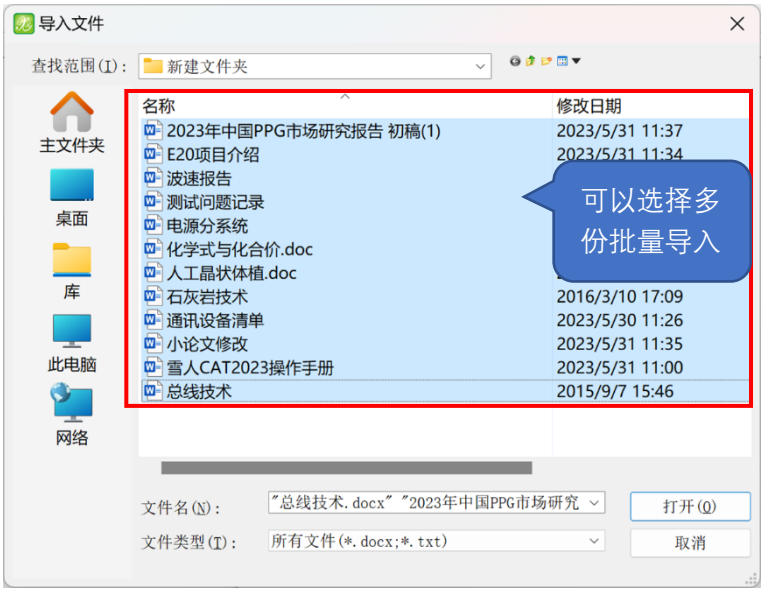


第二节 翻译 WORD 格式的文档

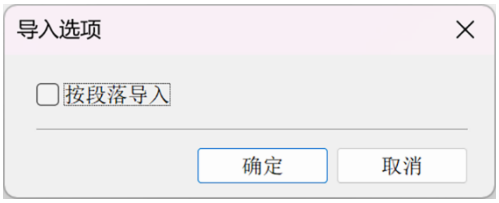
由于雪人 CAT 旧版软件采用的是嵌入 OFFICE 的方式，这种方式虽然能够实现与 OFFICE 的无缝集成，但是由于 OFFICE 的版本众多且更新频繁，也带来了 OFFICE 版本兼容性的问题。而且由于对 OFFICE 的依赖和受限于 OFFICE 处理速度，导致雪人旧版本在原文导入、译文的导出速度方面表现不佳。

新版软件彻底解决了对 OFFICE 版本的依赖，用户不再需要担心 OFFICE 版本的问题，以及和 WPS 的兼容性问题，用户机器上即使没有安装 OFFICE 版本，也不会影响导入、导出和预览，而且导入、导出速度变得飞快。此外，新版本还支持批量导入，一次性可以导入多个需要翻译的文档，大大提高了工作效率。

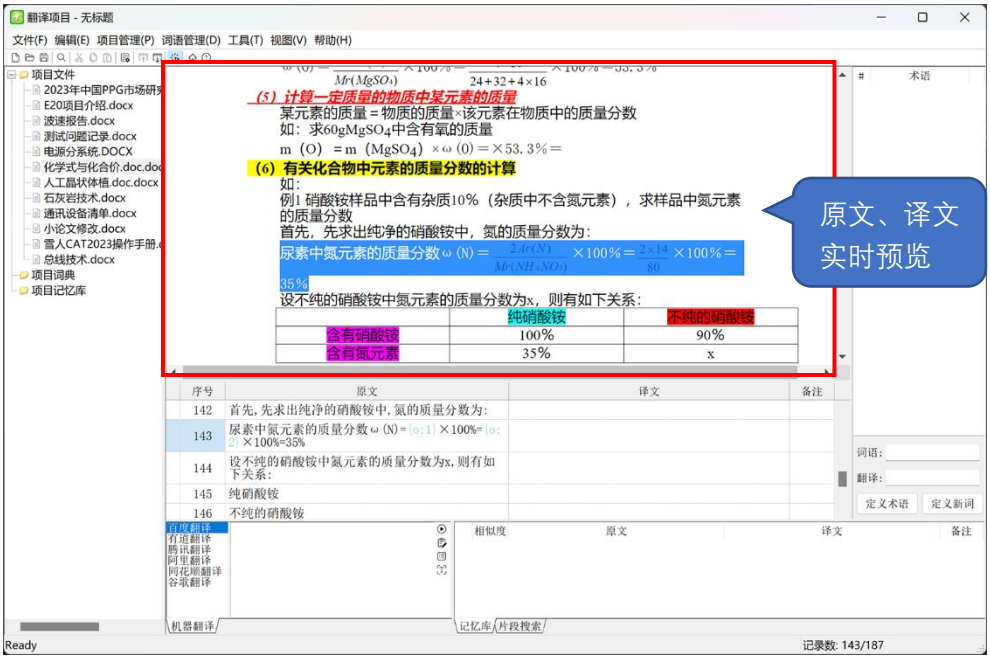
右键点击屏幕左侧的项目文件，再点击**【右键菜单】➔【导入文件】**功能，在弹出的“导入文件”窗口中，可以选择多份 WORD 文档批量导入，如下图：



导入文件时，会弹出如下图的“导入选项”窗口，若勾选“按段落导入”这个选项，则导入后会以段落为单位进行断句划分。以段落为单位断句，可以将整段发送给机器翻译引擎，由于有更多的上下文信息，可以提高机器翻译的译文质量。如果不需要使用机器翻译，可以不勾选“按段落导入”这个选项，则是按句子为单位进行断句。

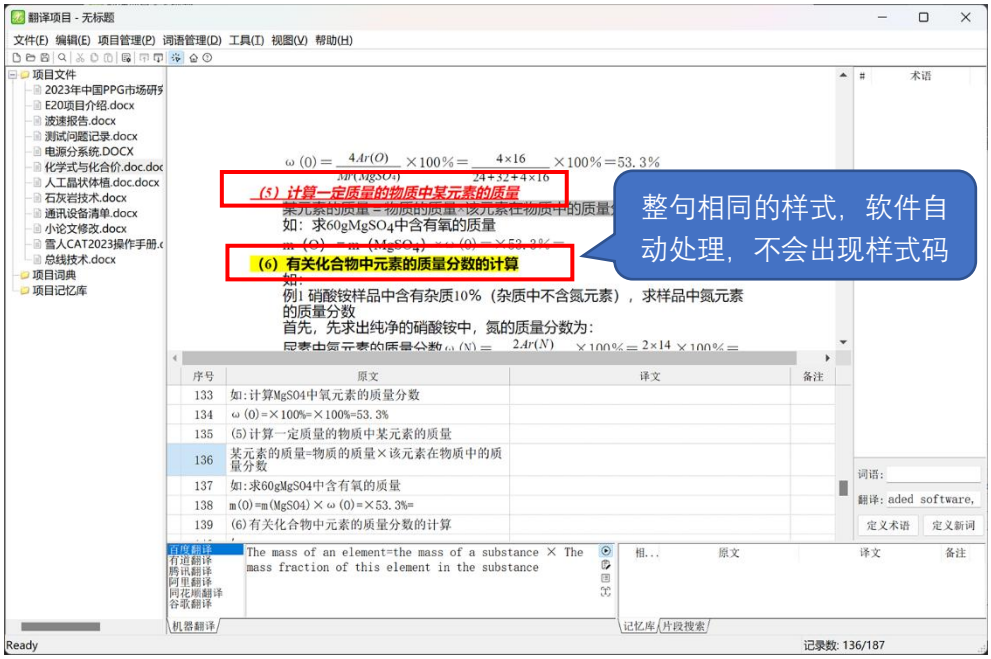


新版中，即使是一次性批量导入数十篇原文，导入的速度也很快，导入后，预览如下图：



第三节 样式码

为了能使译文保持和原文一致的排版风格，样式码是不可避免的。雪人尽可能地自动处理大多数的样式码，例如：当一句中样式相同时，不出现样式码，当一句中存在多种样式时，会出现样式码。如下图的斜体红色字、黄色底色等样式的句子，其样式自动处理了，不会出现样式码。



样式码用“{ }”表示，分为三类，点击菜单**【视图】➔【样式码】**可以根据要显示。

第一类 基本样式代码：需要在译文的相应位置插入，不可忽略，蓝色字体显示。

{o:序号} ——以文字的形式插入的图片或 WORD 的域代码。如：“{o:1 }”

{ref:序号}——脚注的参考位置。如：“{ref:1 }”

第二类 常用样式代码：主要是字体样式、上下标等此类样式，以灰色字体显示。

{b} {/b}——字体加粗/取消加粗

{i} {/i}——斜体字/取消斜体

{u} {/u}——下划线/取消下划线

{h} {/h}——隐藏字符/取消隐藏字符

{sup} {/sup}——上标/取消上标

{sub} {/sub}——下标/取消下标

{str} {/str} ——删除线/取消删除线

{clr:序号} ——字体颜色

{lnk:序号}{/ lnk:序号} ——超链接开始/超链接结束，一般在句中成对出现

第三类 所有样式码：包括第一类、第二类、以及其他样式码。

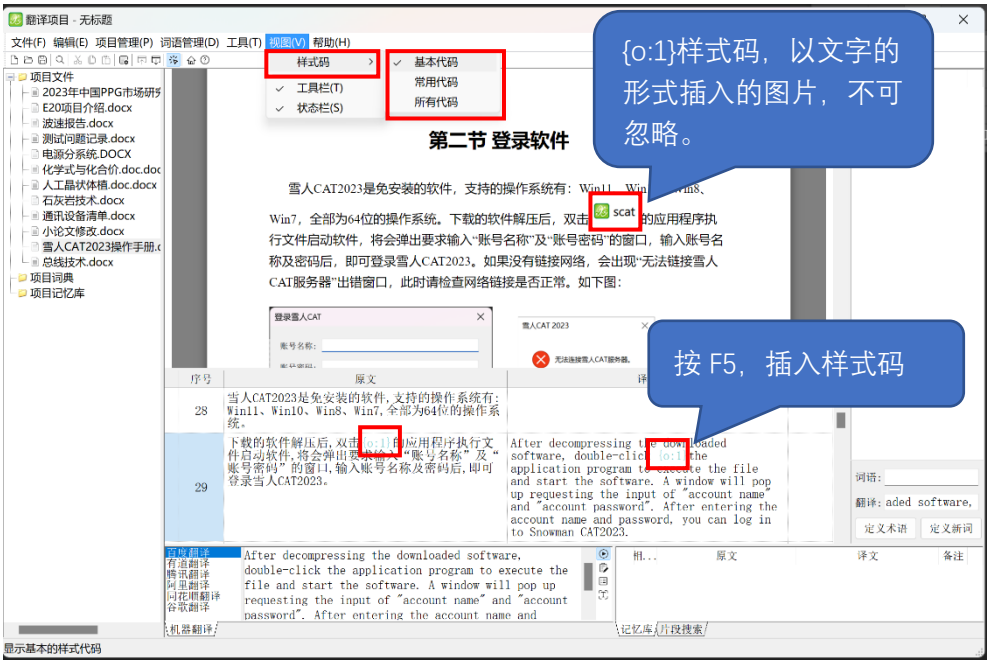
处理样式码的常用热键：

【F5】——在译文处插入样式码。或鼠标划选原文后拖到译文处，可以复制原文，再按【Ctr+T】将普通文字转成样式码，此时该文字会变成灰色显示。

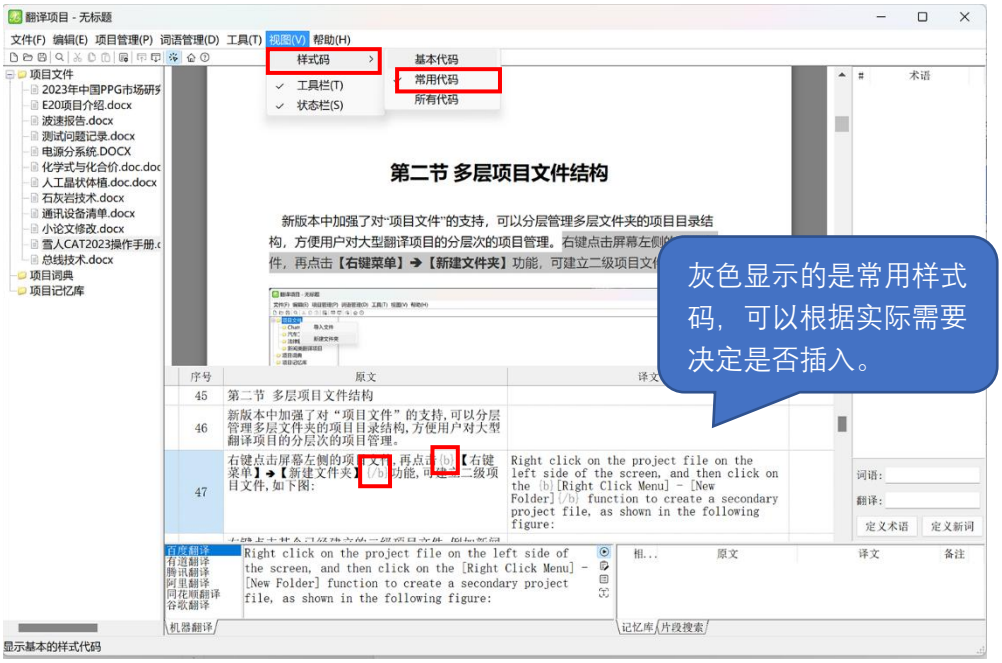
【Ctr+S】——删除样式码，或按工具栏的剪切按钮。

【Ctr+T】——将普通文字变为样式码，再按，将样式码变为普通文字。

例如：下图中，原文这个“ scat ”图片是以文字的形式插入在句子中的，所以在原文中会出现蓝色“{o:1 }”字体显示的样式码，这是属于第一类的基本样式码，在译文的相应位置必须按【F5】键插入它，如果没有插入，导出译文时就会丢失该图片。不是以文字形式插入在句子中的图片，软件会自动处理，不会出现图片的样式码。



例如：下图中，由于原文句子中“**【右键菜单】➔【新建文件夹】**”这几个字是加粗的，所以，在原文中出现了“{b}**【右键菜单】➔【新建文件夹】**{/b}”样式码，表示“{b}　{/b}”范围内的文字需要加粗处理，这是常用样式码，用户可以根据需要决定是否在译文中插入它。若整句都加粗的，则不出现样式码。



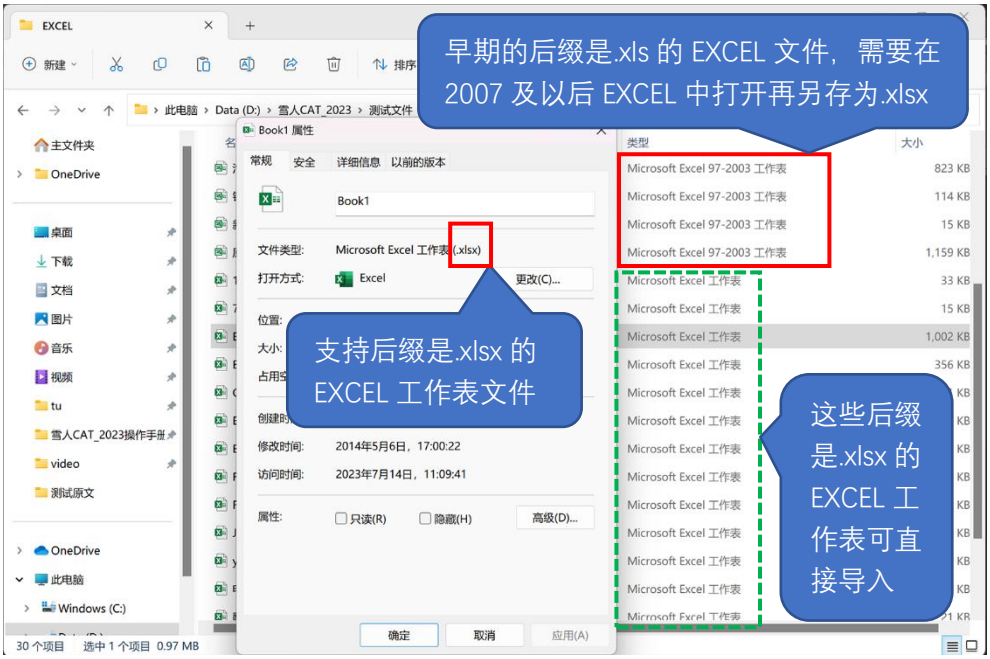
如果不想在翻译中处理太多的样式，可以仅处理蓝色显示的基本样式码，这样也可以导出一个样式基本正确的译文，导出后再进行样式排版处理。

进行了合并句子、分拆句子操作的句子，对于如{ ：序号}的样式码，可能需要调整序号。

对于某些复杂排版的 WORD 文件，预览时可能存在某些样式没有精确还原的问题，但这仅仅影响屏幕预览，不会影响导出译文，导出的译文是可以很好还原的。

第四节 翻译 EXCEL 格式的文档

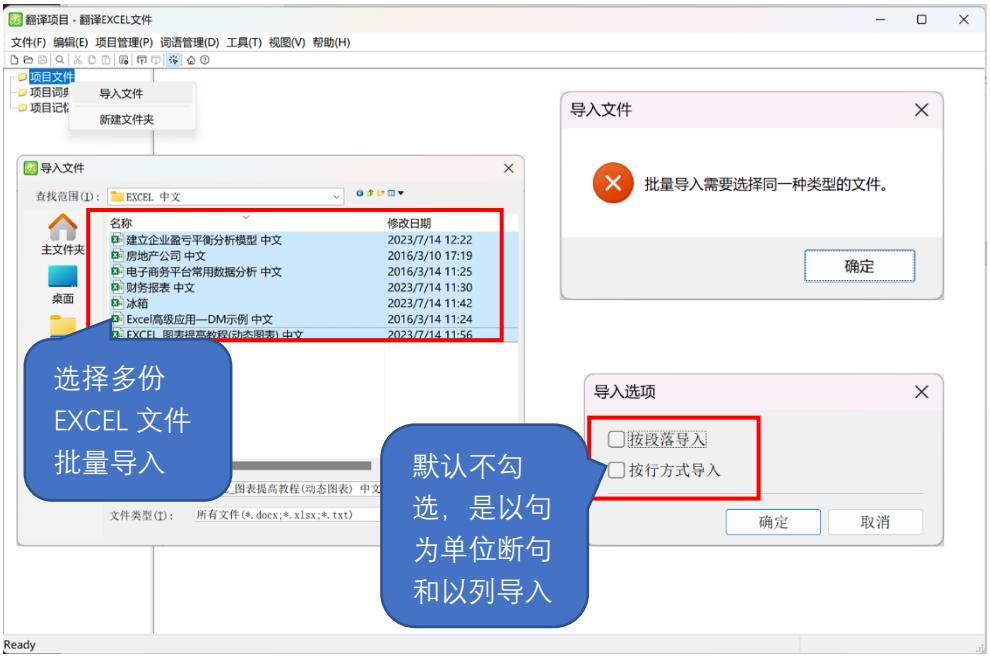
雪人新版支持翻译后缀是“.xlsx”的 EXCEL 文件格式，如果是 Microsoft Excel97-2003 版本生成的 EXCEL 文件，其后缀是“.xls”，需要用 Microsoft Excel2007 或以上的版本打开并另存为“.xlsx”，然后再导入雪人软件中。查看文件后缀的方法：右键点击某个 EXCEL 文件后，点击**【右键菜单】➔【属性】**，在“文件类型”中可以查看到该文件的后缀。如下图：



雪人新版支持一次性批量导入多份 EXCEL 文件，**注意，批量导入时需要选择同一种类型的文件，**例如，选择的全部是 EXCEL 文件才可以进行批量导入，不能选择几个 EXCEL 文件和几个 WORD 文件，这样会报出错窗口，如下图：

导入时，会弹出“导入选项”窗口，如下图：其中“按段落导入”选项如果勾选，则表示以段落为单位进行断句，如果不勾选，则是默认的以句子为单位进行断句。“按行方式导入”，若勾选，翻译时光标是以 EXCEL 单元格的行方式进行跳转，默认

是不勾选，表示翻译时光标是以 EXCEL 单元格的列方式进行跳转。



例如：下图是需要翻译的 EXCEL 文件的原文。



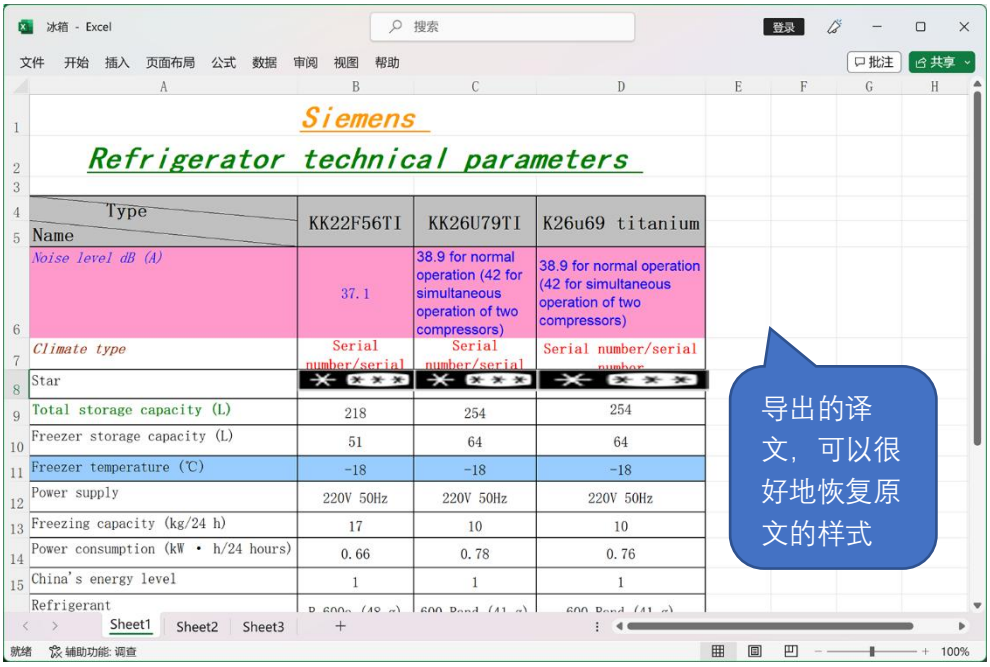
类型 / 名称	KK22F56TI	KK26U79TI型	kk26u69钛
噪声级dB（A）	37.1	38.9用于正常运行（42用于两台压缩机同时运行）	38.9用于正常运行（42用于两台压缩机同时运行）
气候类型	序号/序号	序号/序号	序号/序号
星级	[redacted]	[redacted]	[redacted]
总存储容量（L）	218	254	254
冷冻室存储容量（L）	51	64	64
冷冻室温度（℃）	-18	-18	-18
电源	220V 50Hz	220V 50Hz	220V 50Hz
冷冻能力（kg/24h）	17	10	10
功耗（kW·h/24小时）	0.66	0.78	0.76
中国能源水平	1	1	1
制冷剂	R600a（48克）	600兰特（41克）	600兰特（41克）
外形尺寸（mm）	590x649x1750	640×633×1850	640×641×1900

将 EXCEL 文件导入雪人中的，点击**【工具菜单】➔【预翻译】**，选择一个机器自动翻译引擎进行全文自动翻译，如下图：然后对自动翻译不满意的译文进行修

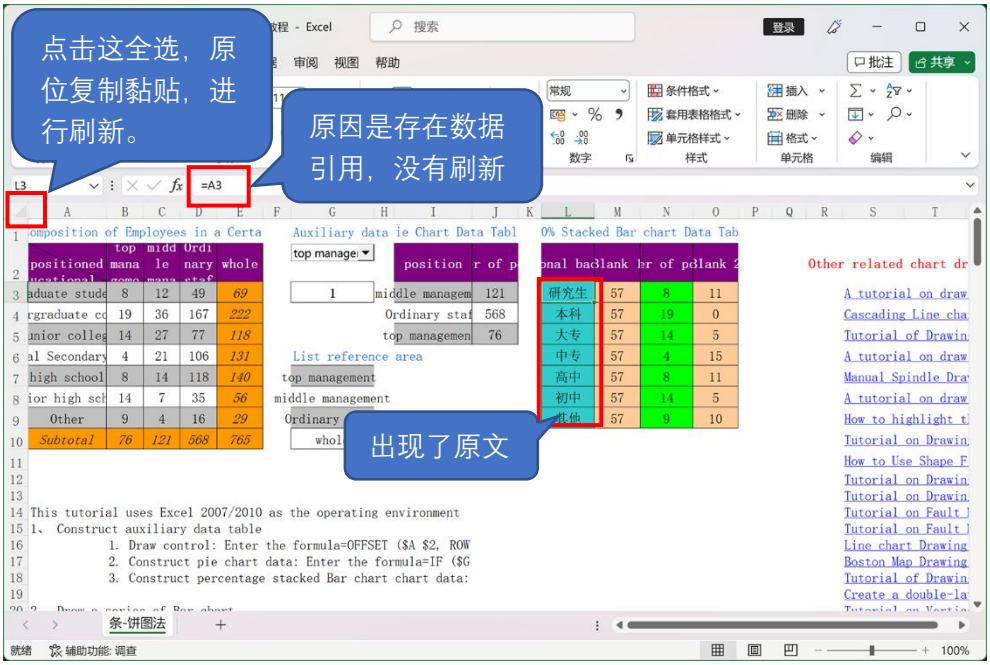
改，修改完成后，点击屏幕左侧的这个具体 EXCEL 文件，再点击**【右键菜单】➔【导出文件】**导出译文。



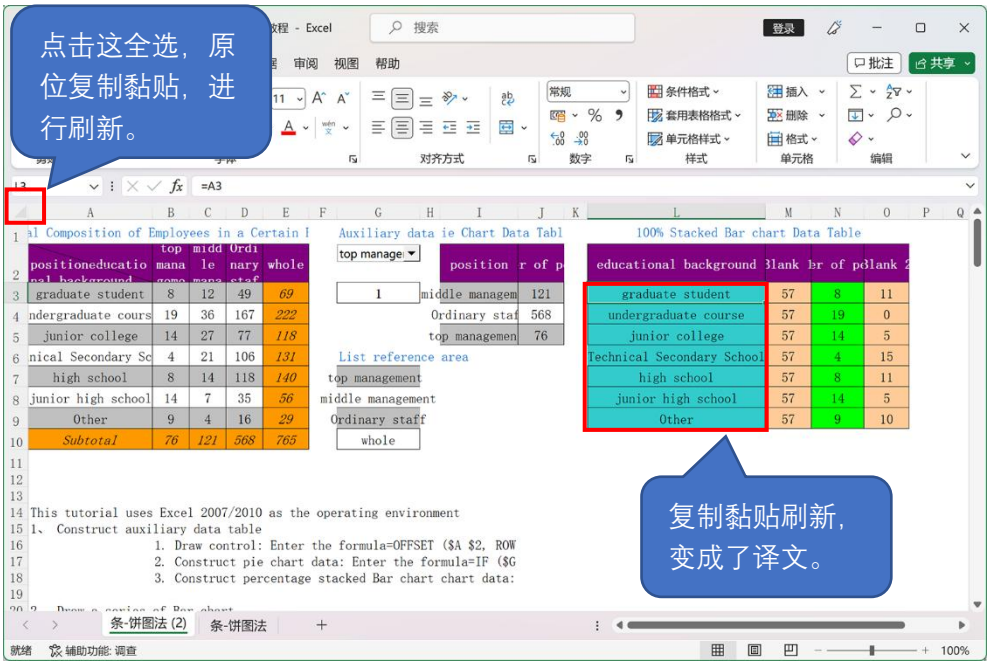
雪人可以自动处理绝大多数的样式，导出的译文可以很好地还原原文的样式。如下图:



如果 EXCEL 中，存在某些公式、数据的引用，导出的译文可能由于刷新的问题，会出现原文，如下图：



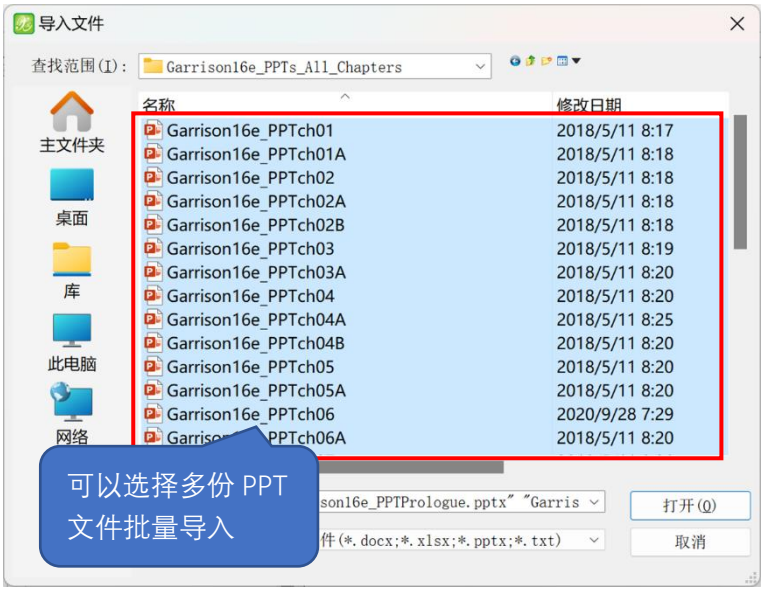
解决办法：在 EXCEL 中，点击左上角选择整个工作表，然后按【Ctrl】+【C】复制，再按【Ctrl】+【V】在原位进行黏贴，这样就可以刷新引用的数据。如下图：

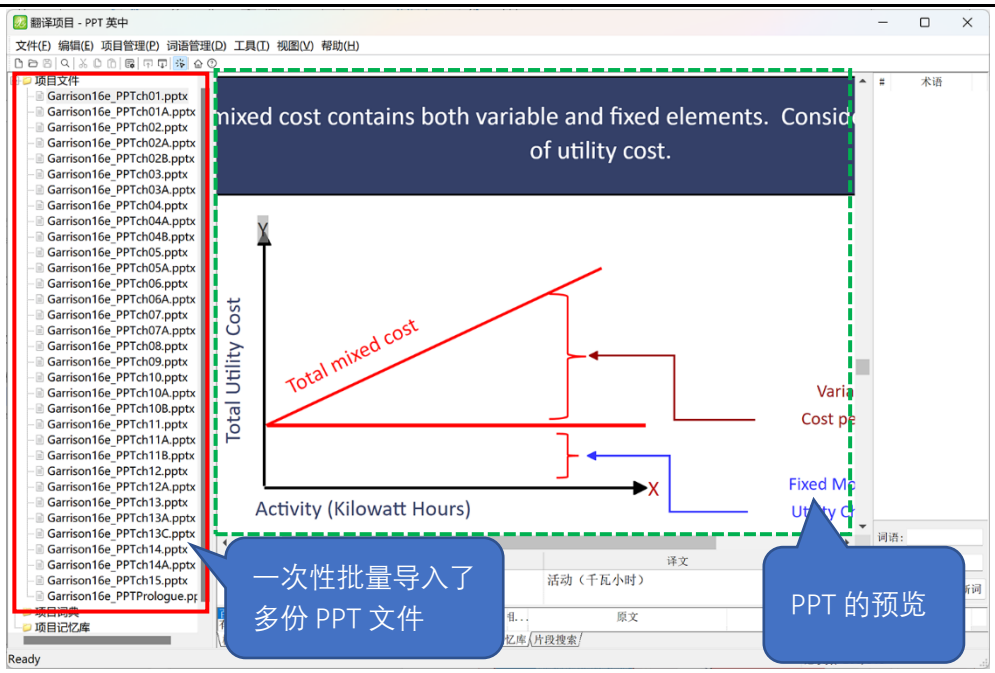


第五节 翻译 PPT 格式的文档

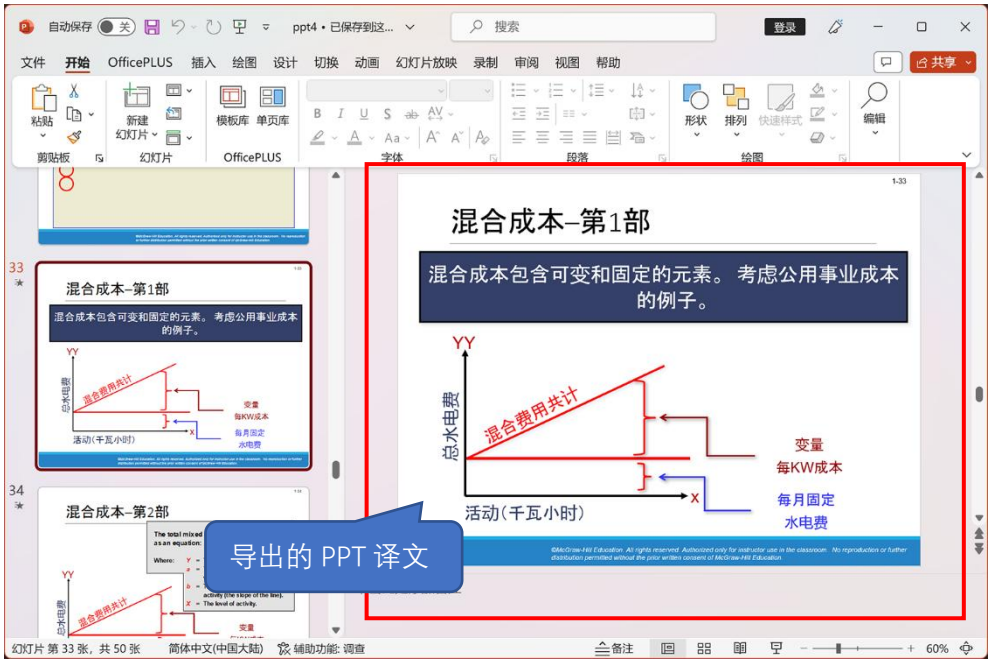
雪人新版支持翻译后缀是“.pptx”的 PPT 文件格式，如果是 PowerPoint97-2003 版本生成的 PPT 文件，其后缀是“.ppt”，需要用 PowerPoint2007 或以上的版本打开并另存为“.pptx”，然后再导入雪人软件中。

右键点击屏幕左侧的项目文件，再点击**【右键菜单】➔【导入文件】**功能，在弹出的“导入文件”窗口中，可以选择多份 PPT 文档批量导入，如下图：**注意，批量导入时需要选择同一种类型的文件，**即全部是 PPT 文件才可以进行批量导入。



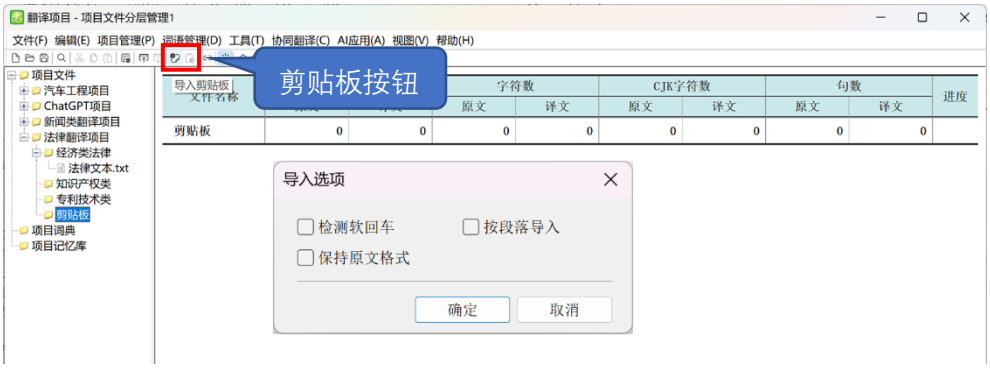


翻译完成后，导出的 PPT 译文如下图:



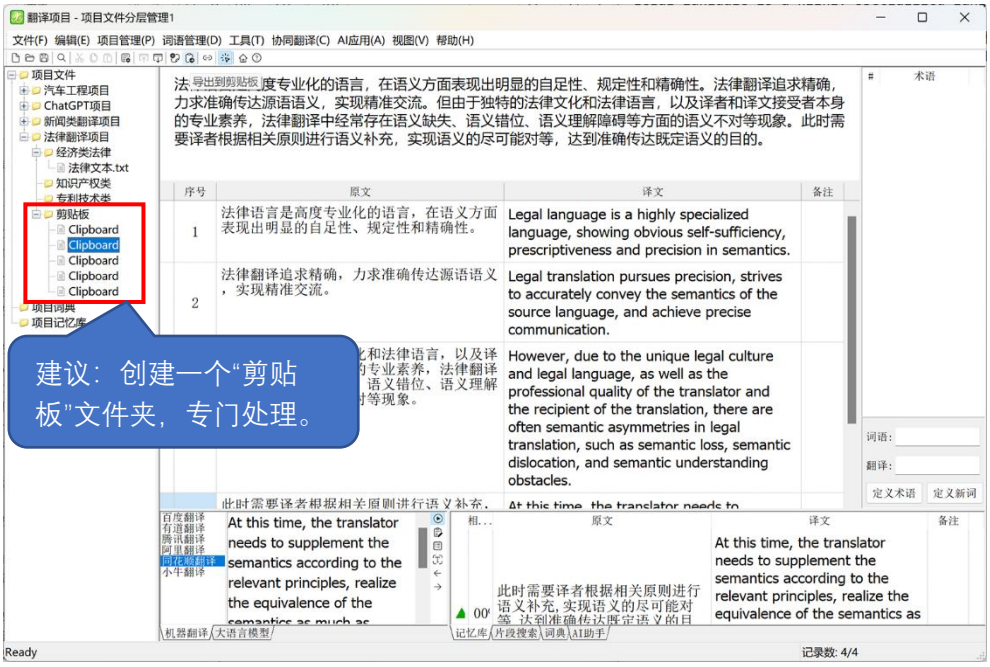
第六节 剪贴板翻译

利用“剪贴板”功能，可以方便、快捷对一些文本进行翻译，将需要翻译的内容按【Ctrl+C】键复制到剪切板上后，点击工具栏的“导入剪贴板”按钮，可立即将它粘贴到雪人 CAT 中进行翻译，如下图：



建议：创建一个“剪贴板”文件夹，专门处理这些来自剪贴板的内容，好处是：可以以文件夹为单位来管理这些内容，例如删除、导出等等。

翻译完成后，按“导出到剪切板”按钮，将译文导出到剪切板上，再按【Ctrl+V】键可以粘贴到目标位置上。



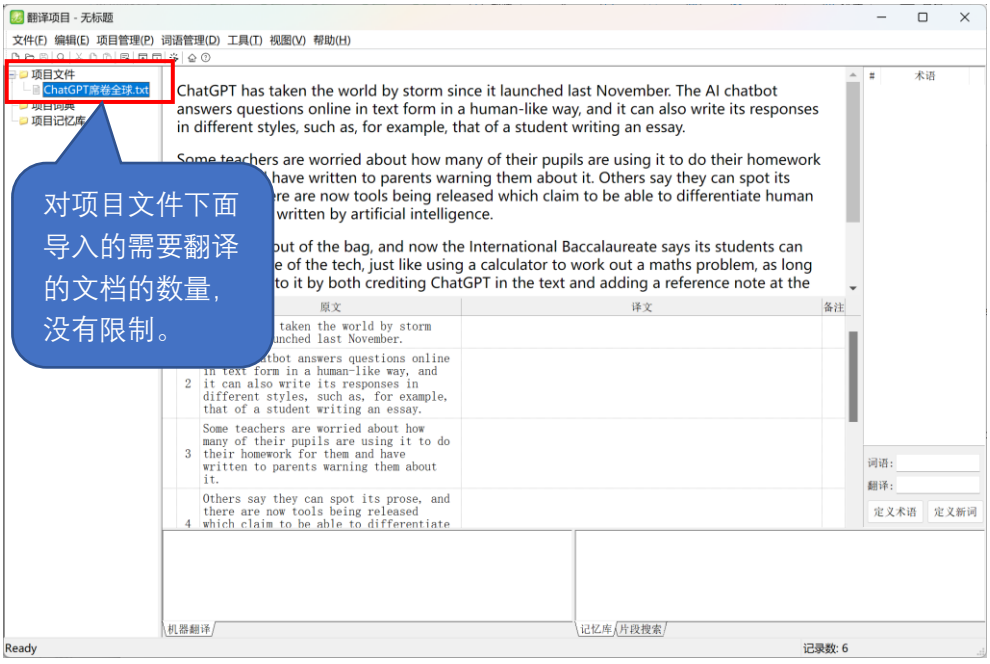
第七章 支持复杂的项目文件结构

第一节 单层项目文件结构

单层项目文件结构就是指在屏幕左侧的项目文件根目录下面，导入需要翻译的文档，不新建子文件夹进行分类。即右键点击屏幕左侧的项目文件，再点击**【右键菜单】➔【导入文件】**，导入一份或批量导入多份需要翻译的文档。



所导入的文档会放在项目文件根目录下面，其显示结果和雪人旧版一样，不同的是旧版项目文件下面最多支持 200 份需要翻译的文档，新版则没有限制。如下图：



第二节 多层项目文件结构

新版本中加强了对“项目文件”的支持，支持分层管理的多层复杂的项目目录结构，项目文件分多少层子文件夹不作限制，每个文件夹下面存放多少个翻译文档不做限制，对导入的每一份翻译文件的大小亦不作限制，非常方便用户对大型翻译项目的分层次的项目管理。

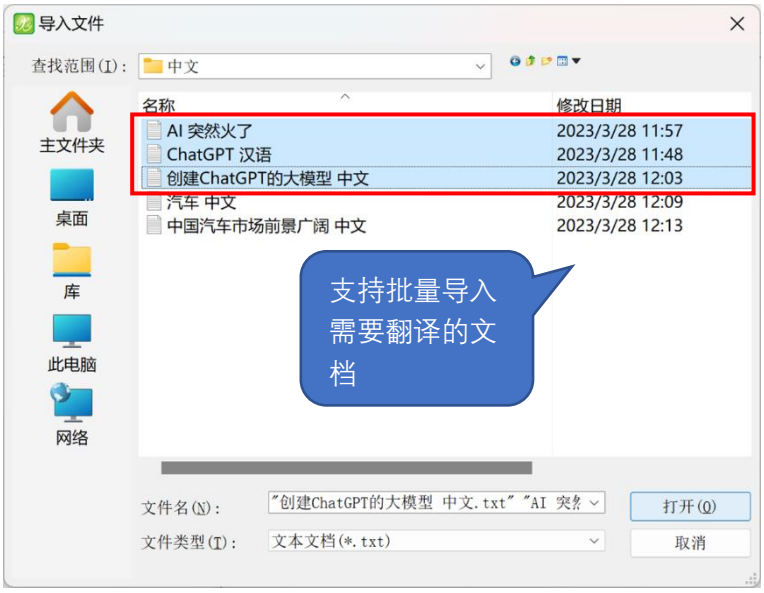
右键点击屏幕左侧的项目文件，再点击**【右键菜单】➔【新建文件夹】**功能，可建立二级项目文件，如下图：



右键点击某个已经建立的二级项目文件，例如新闻类翻译项目，再点击**【右键菜单】➔【新建文件夹】**功能，可在该二级项目文件下面建立三级项目文件，同理，可以建立四级项目文件、五级项目文件，具体分多少层不作限制，用户根据实际使用需要而定。如下图：



鼠标点击某层次的项目文件，再点击**【右键菜单】➔【导入文件】**功能，则可在该级项目文件下面导入待译文档，导入文档时支持批量导入，可选择多份文档一次性导入。新版本对导入的文件大小和文件数量都没有限制。如下图：



例如：选择了“AI 突然火了”、”ChatGPT”、“创建 ChatGPT 的大模型”这三个文档，将它们一次性导入“ChatGPT 项目”中，如下图：



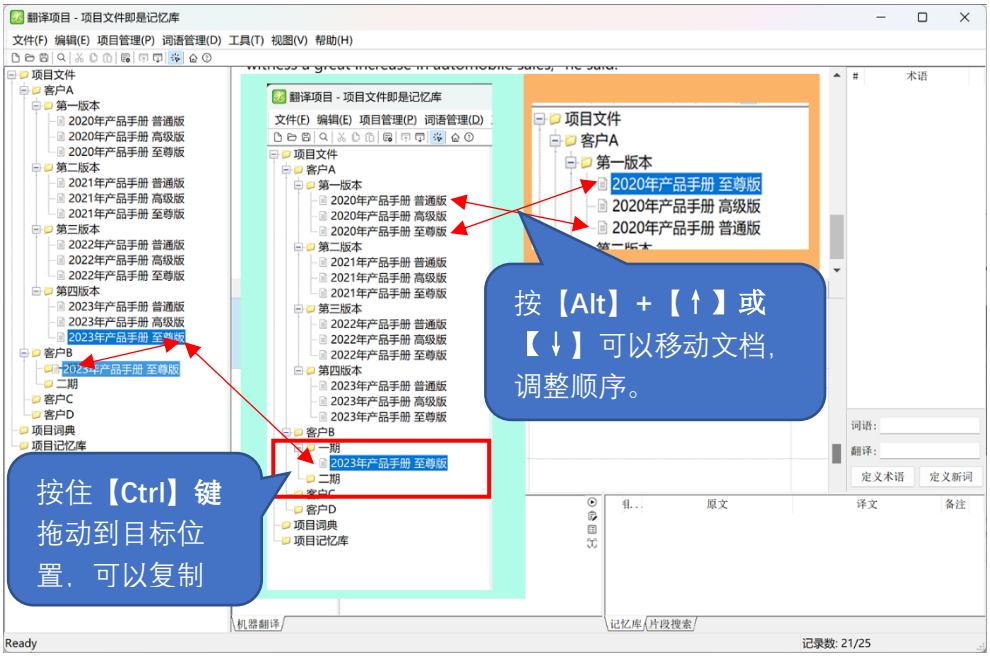
第三节 项目文件的整理

在复杂的项目文件结构中，会需要对项目文件下面的各个子目录进行调整，归类等操作。调整各级项目文件的顺序及项目文件的移动、复制也非常灵活方便。

同文件夹下面移动文档——按【Alt】+【↑】或【↓】光标键，可以调整它在此文件夹下面的顺序。

移动文档到另一个文件夹——按住鼠标左键单击并拖动可以将它移动到其他文件夹下面。

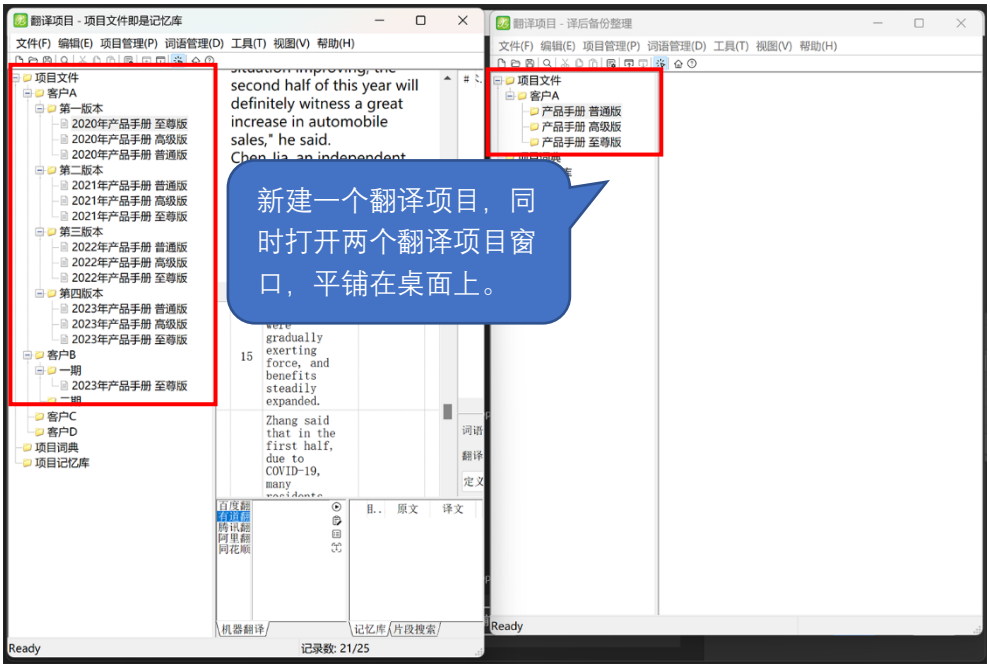
复制文档——按住鼠标左键单击拖动，并同时按住【Ctrl】键，此时鼠标会变成一个“+”号，则可复制一份到其他文件夹下面。



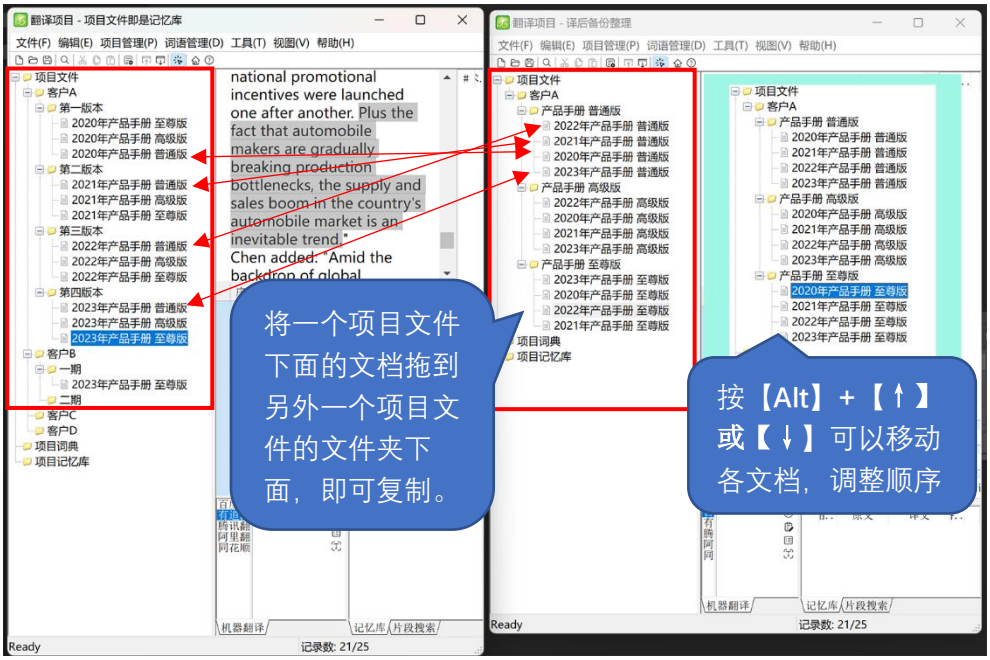


同理，文件夹的顺序也可以调整，鼠标点击某个文件夹后，按【Alt】+【↑】或【↓】光标键，可以调整同层次的文件夹顺序。如上图：

新版本中，由于翻译的项目文件支持多层次的文件目录结构，这个功能对翻译后的文档归类、整理也变得十分方便。例如：给某客户 A 翻译了多年不同版本的产品手册，现在想将它们归类整理，则新建一个翻译项目，将两个翻译项目窗口同时平铺在桌面上，如下图：



例如：下图中，将客户 A 的“普通版各年的产品手册“”归类到“产品手册 普通版"这个文件夹中，方法是：点击某个项目文件下面的文档拖到另外一个项目文件的文件夹下面，即可将它复制到新的文件夹下面，这一就可以很方便地进行归类、整理，集中起来，按文件夹导出记忆库。


将一个项目文件下面的文档拖到另外一个项目文件的文件夹下面，即可复制。
按【Alt】+【↑】或【↓】可以移动各文档，调整顺序

第八章 词典和术语

第一节 定义术语和定义新词

术语一般是指专业的词汇，术语是在整个翻译项目中需要统一译法的词，一般情况下一个术语只对应于一种解释。

鼠标划选原文和划选对应的译文，它们就会分别出现在右下角的新词定义区中，点击**定义术语**按钮或按**【Ctrl+D】键**即可添加为术语。若按点击**定义新词**按钮或按**【Ctrl+W】键**则是定义为新词。定义为术语的词在原文以蓝色提示，同时会自动出现在右侧的术语提示区。定义新词则是定义为普通的词汇。如下图:

在上方的工具栏中，有“鼠标取词”按钮，按下此按钮，当鼠标指向某词或划选某些词语时，会自动弹出屏幕取词窗口，点击取词窗口中的提示词义，可以直接输入到译文中。如果不希望自动弹出屏幕取词窗口，可以再按一次该按钮切换状态，关闭自动弹出的取词窗口。



刚才定义的新词和术语，会保存在屏幕左侧的“项目词典”中，如下图：有些翻译项目，客户会提供术语库，可以将它导入到“项目词典”中，利用**【右键菜单➔设为术语】**功能将它们设定为术语。

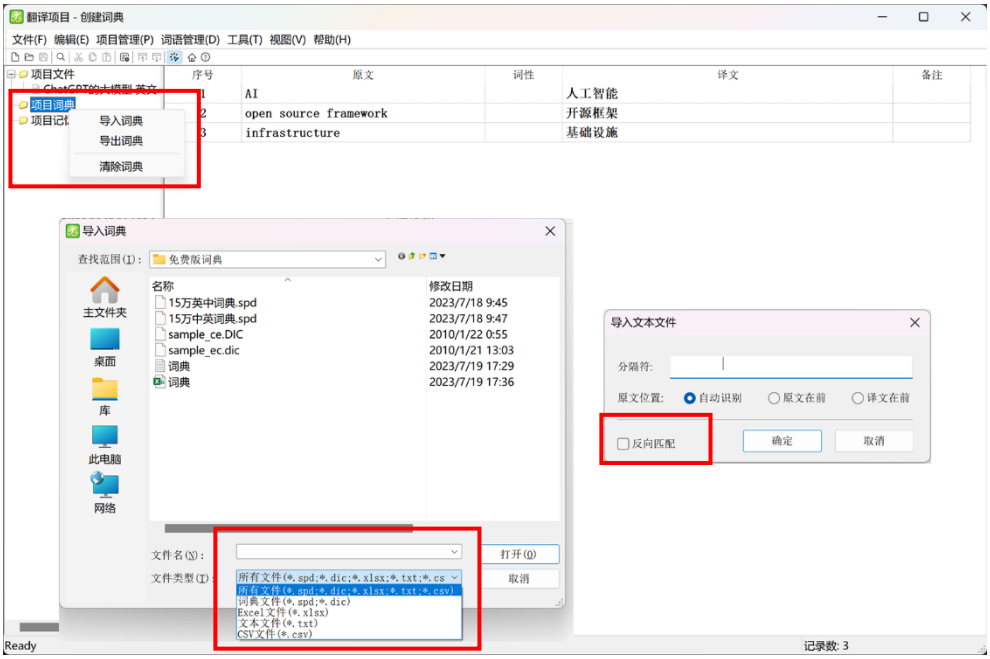

翻译项目 - 英译中 片段搜索
文件(F) 编辑(E) 项目管理(P) 词语管理(D) 工具(T) 视图(V) 帮助(H)
项目文件
项目词典
项目记忆
导入词典
导出词典
清除词典
原文
译文
备注
1 open source framework 开源框架
2 AI 人工智能
3 infrastructure 基础设施
增加词语
设为术语
设为词语
删除词语
清除词典
术语加粗显示，普通词汇不加粗

第二节 词典的导入导出

“项目词典”用于存放于用户在翻译过程中增加的新词、术语，用户也可以在“项目词典”中导入自己的词汇或术语表。

词典的导入

右键点击左侧的项目词典，会弹出右键菜单，点击**【右键菜单】➔【导入词典】**功能，可以将词汇导入雪人软件中。



雪人支持导入的词典格式有：

.spd——新版雪人的词典格式。

.dic——旧版雪人的词典格式。旧版的词典需要在项目词典中进行导入后再导出，才能转换成新版的格式外挂引用。

.txt——文本格式，原文、译文之间要有分隔符。

.csv——CSV 格式，EXCEL 文件可以在 EXCEL 软件中另存为.csv 格式，然后导入雪人中。

雪人导入词典的速度非常快，而且**会自动删除重复的词汇**，有时会发现词汇导入前的数量和导入雪人后的数量不一致，例如导入前 EXCEL 记录显示是 2000 行，导入雪人后显示 1980 记录数，数量上相差了 20 个，是因为存在重复的词汇，导入时雪人自动去重了。

若是导入 EXCEL 格式的词库，则只需在 EXCEL 中分 A、B 两列，一列原文、一列译文。若是文本格式的词库，**则在原文和译文之间需要设置分隔符**，分隔符建议采用某些特殊的符号，如：Tab 制表符、||等。设置好分割符后，就可导入。

```
B to B (B2B) ||business to business
2024 夏季奥运会  ||2024 summer Olympics
甲 A 足球队|| Division A soccer team
360 度环幕电影|| Circamara
申办 2024 年奥运会|| bid for 2024 Olympics
主办 2024 年奥运会  ||host the 2024 Olympic Games
```

反向匹配

在导入 TXT 格式的词典或者记忆库时，有时会遇到下图所示的 TXT 原文件（这里以词典为例说明，记忆库同理）：（图一）的中英文间是以空格分割的，且中文在前，英文在后面；这种以空格为分隔符的格式，可以直接导入，软件会自动正确区分原文、译文。



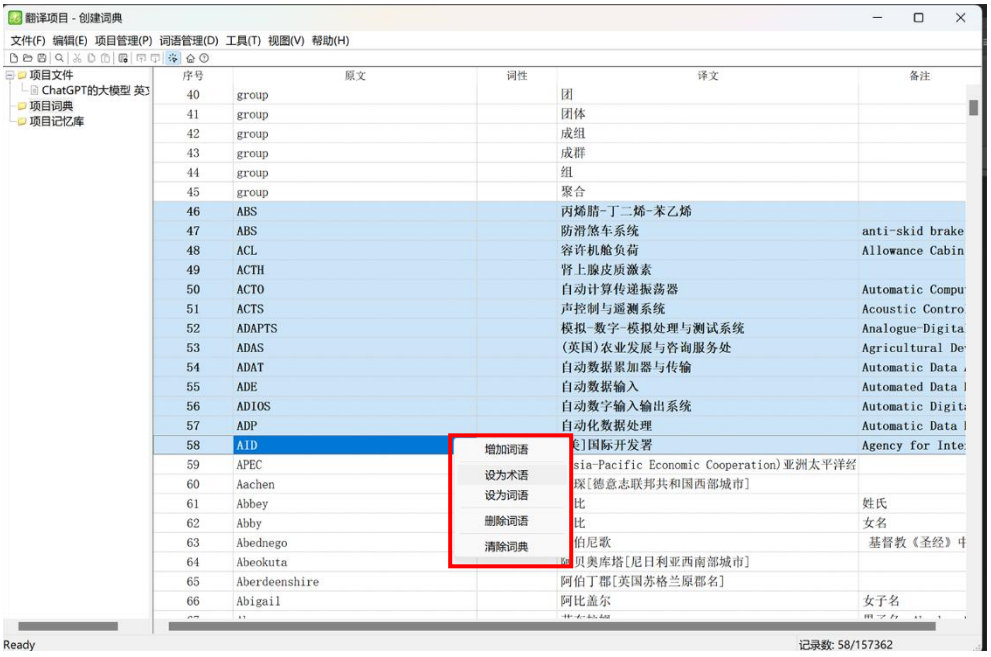
而（图二）也是以空格分割的词语，不同的是英文在前，中文在后面；由于英文的单词之间也是以空格分割的，直接导入就不正确了，勾选导入窗口的“反向匹配”功能，则能正确导入。

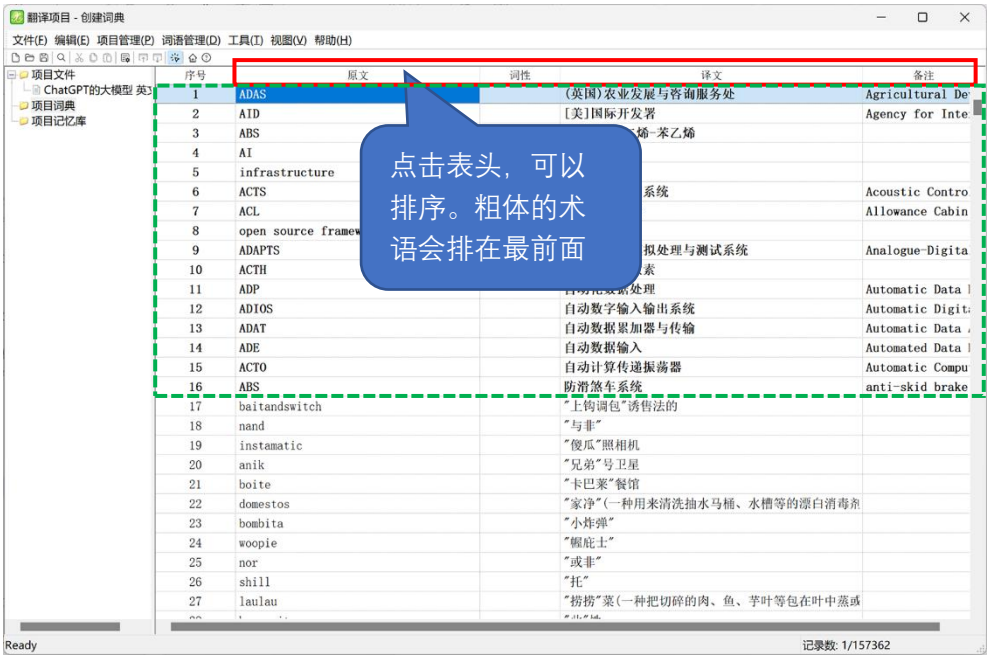
词典的编辑

将词汇导入“项目词典”后，可以进行编辑。点击右键，会弹出右键菜单，有【增加词语】、【设为术语】、【设为词语】、【删除词语】、【清除词典】功能。

整批设为术语——按住【Shift】键+【鼠标单击前后记录】选择多条记录，再点击**【右键菜单】➔【设为术语】**功能，可以整批将它们设定为术语，术语会以粗体显示，机器翻译时若勾选“使用术语替换”选项，就会以“项目词典”中术语的词义进行翻译。

修改——双击某词条或按【F2】键，进入编辑修改状态。
删除——点击某条词语后，点击**【右键菜单】➔【删除词语】**功能，则删除一条记录。若按住【Shift】键+【鼠标单击前后记录】，可选择多条连续的记录进行批量删除；按住【Ctrl】键+鼠标点选，可以选择多条不连续的记录进行批量删除。
排序——点击表头的原文，译文、词性等，可以进行排序。粗体的术语会排在最前面。如下图：





词典的导出

右键点击左侧的“项目词典”，会弹出右键菜单，点击**【右键菜单】➔【导出词典】**功能，可以将词汇导出。支持导出的格式有:

.psd——雪人新版的词典格式，外挂词典需要导出成.psd 格式，雪人旧版的词典为.dic 格式，需要在新版中导入后在导出转换成新版的.psd 格式来外挂。

.txt——文本文件格式，原文、译文之间自动用 Tab 制表符分隔，方便其他软件读取。也可以用 EXCEL 软件打开再另存为 EXCEL 格式。

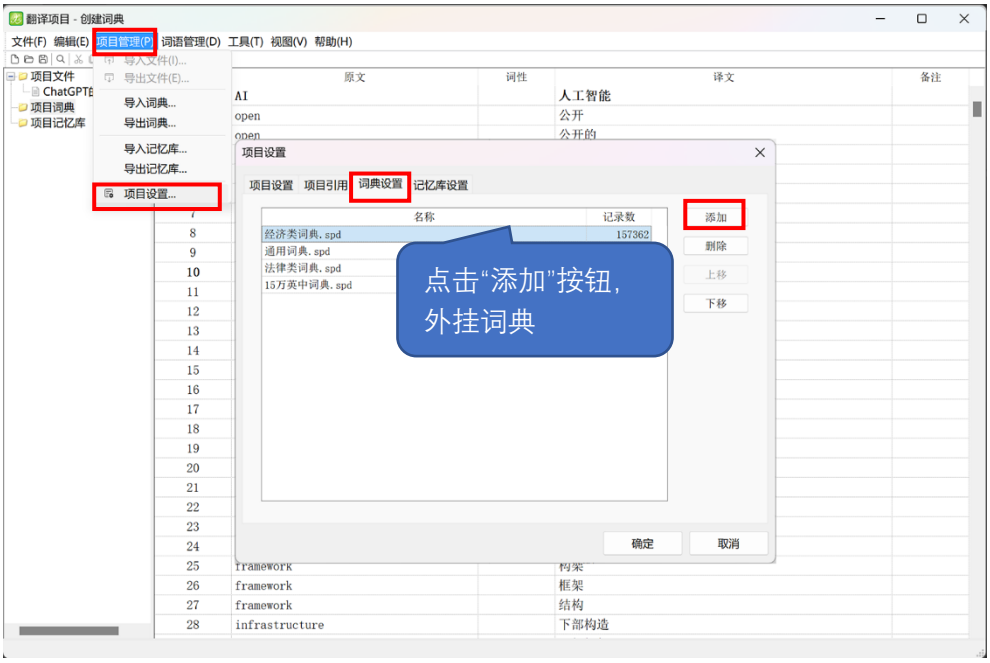
.csv——CSV（Comma-Separated Values）是一种常用的数据交换格式，也可以用 EXCEL 软件打开再另存为 EXCEL 格式。

词典的合并

如果有多个词典文件，希望将它们合并成一个词典文件，则可以在同一个项目文件的“项目词典”中导入这些词典文件，然后点击**【右键菜单】➔【导出词典】**功能即可生成一个合并了的词典文件，而且会自动删除记忆库中的重复词汇。

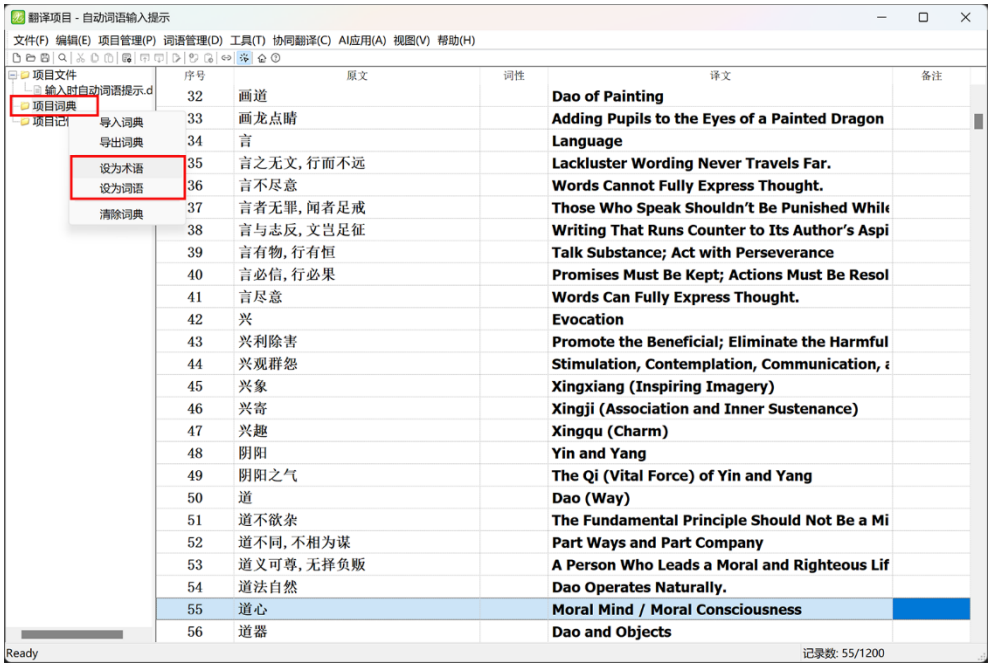
第三节 词典的外挂

词典可以在“项目词典”中直接导入使用，也可以通过外挂词典文件的方式来使用。外挂词典的好处是减少内存的消耗。在“项目词典”中将词典导出成“.psd”格式文件，就可以被以后的翻译项目添加引用。方法是：点击菜单**【项目管理】➔【项目设置】**（或点击工具栏的 “项目设置”图标），再点击“词典设置”标签页，点击“添加”按钮，可以加载多个“.psd”格式的词典文件，根据需要调整各个词典的优先级顺序，放在上面的比下面的优先。如下图：



第四节 整批设置术语/词语

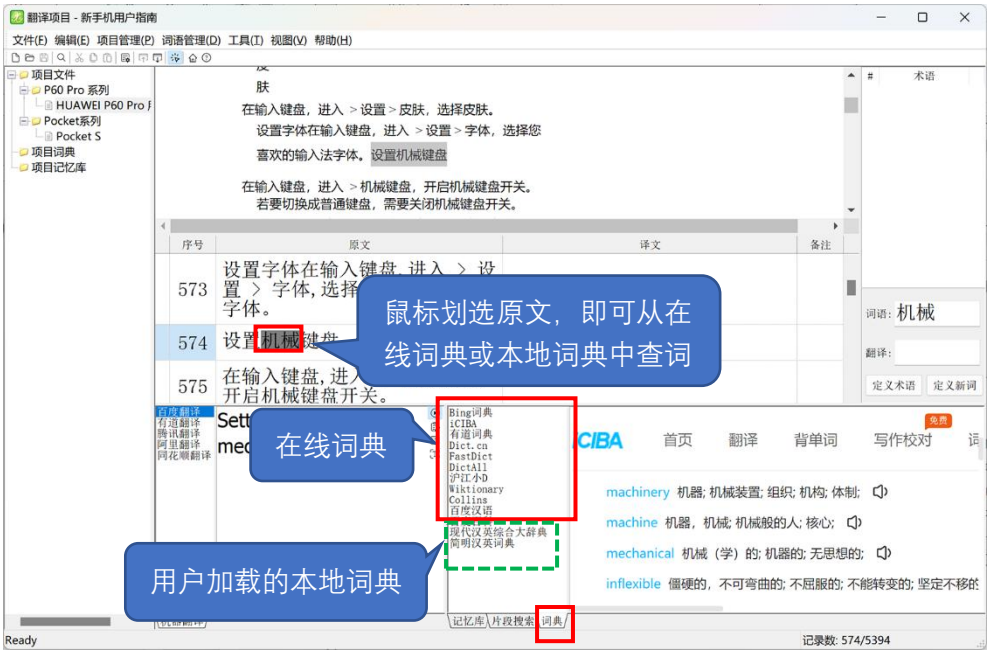
点击屏幕左侧的“项目词典”，然后点击**【右键菜单】➔【设为术语】**功能，可以将“项目词典”中所有的内容设为术语，设为术语后会加粗显示。点击**【右键菜单】➔【设为词语】**功能，则全部设为普通词汇，普通词汇不会加粗显示。



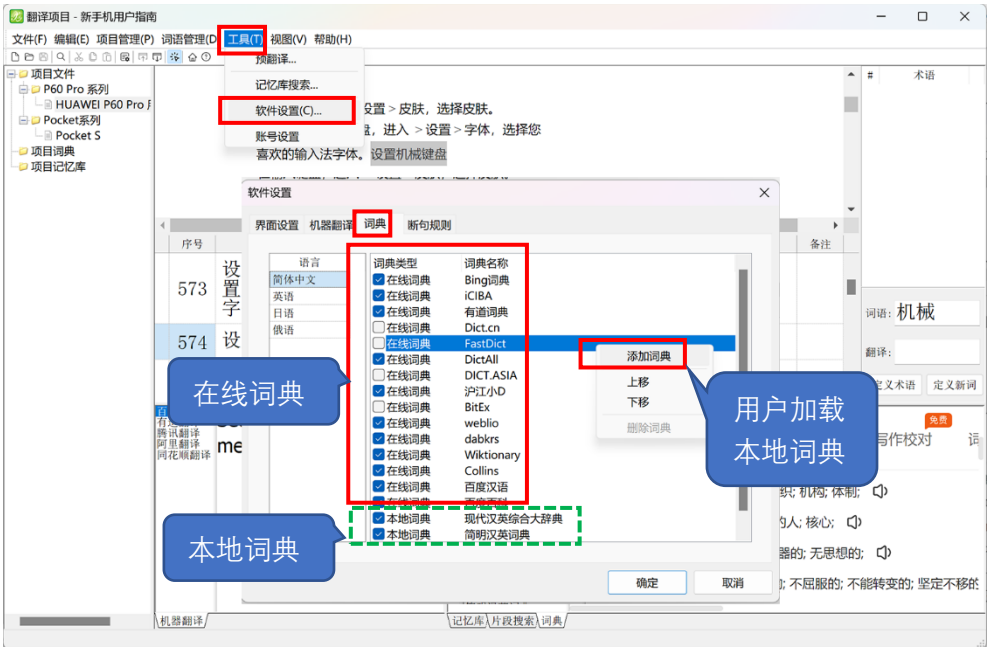
第五节 在线词典和本地词典

点击翻译界面右下方的“词典”页签，可以在此显示 Bing、有道等多个在线词典和用户自己添加的本地词典，鼠标划选原文的词语，即可从词典中查询该词的解释及例句等信息。如下图：

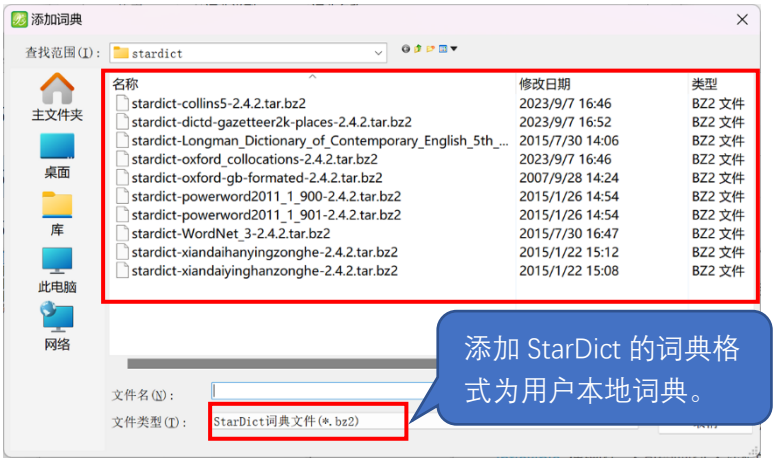
雪人内置的在线词典，用户可以在“词典设置”中，勾选显示与不显示，但是不能删除，也不能自行添加在线词典，如果您发现好的在线词典，欢迎发邮件到 xueren@gcys.cn 推荐，雪人会将它加入到在线词典列表中。



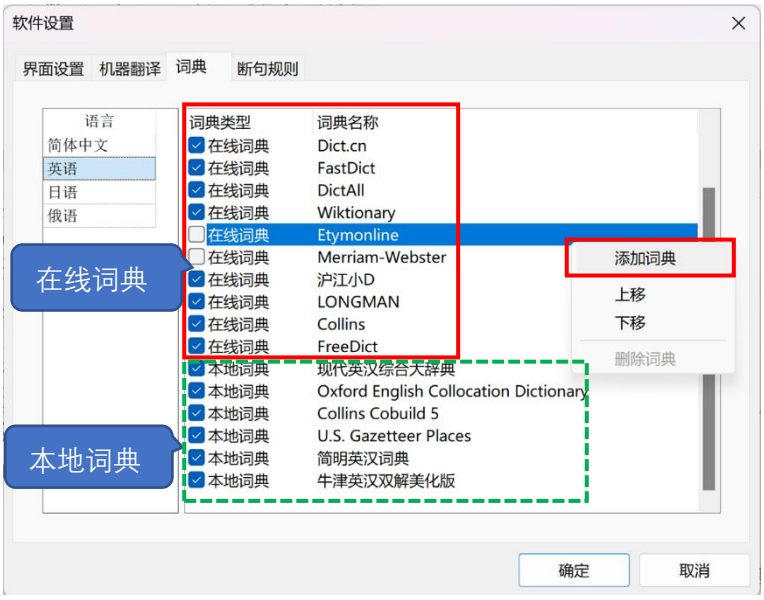
词典的显示与否及各个词典的排列顺序，用户可以自行设置。点击**【工具菜单➔软件设置➔词典】**功能，在“词典设置”窗口，可以为各个语种设置显示哪些在线词典?如果不想要显示某个词典，可以不勾选它。点击**【右键菜单➔上移/下移】**功能（或者按**【Alt】+【↑】或【↓】**光标键），可以调整各个词典的显示顺序。如下图：



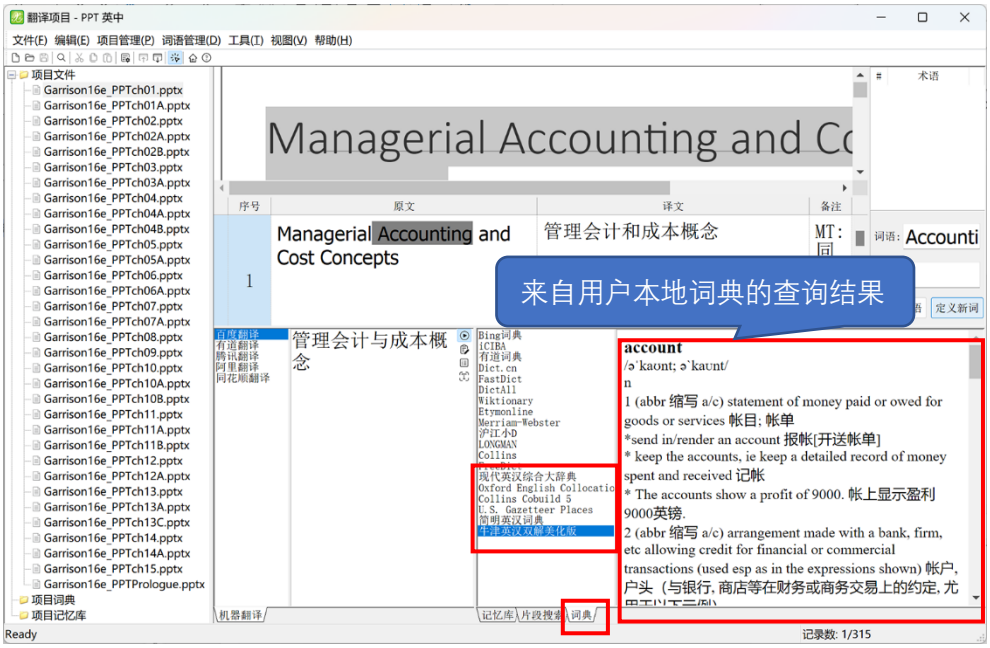
点击【**右键菜单➔添加词典**】功能，用户可以添加自己的本地词典，雪人支持直接添加“**starDict** ”2.42 版本的词典，文件名后缀是 BZ2”的词典文件作为用户的本地词典，即“星际译王词库”，如下图：



Stardict 是开源的词典项目，用户可以自己制作或使用现成制作好的词典。在“百度”中搜索，输入“星际译王词库”或“stardict”等，可以找到一些可用的词典，然后根据自己的需要下载各个语言方向的词典，下载后不用解压，直接添加到雪人中作为本地词典使用，也可以下载一些单语的词典，加载到雪人中作本地词典使用。

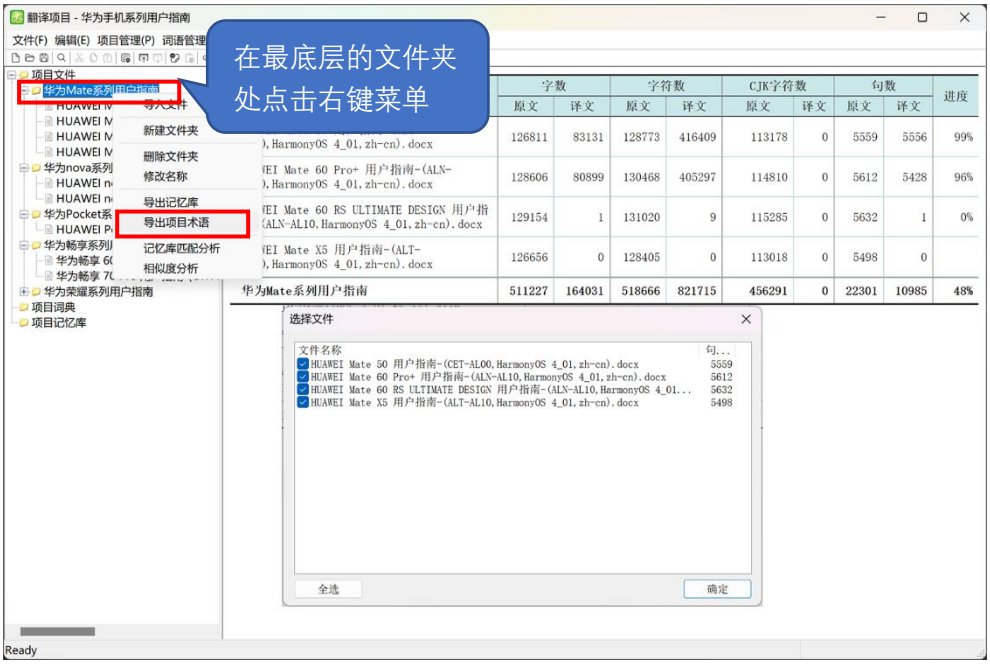


例如，下图鼠标划选原文的词语，点击屏幕下方的“词典”列表中的本地词典，即可显示本地词典的查询结果，本地词典的优势是响应快，没有广告，但是也会占用一些内存，用户根据需要添加，也不是越多越好，宜选择一些高质量的词典下载。

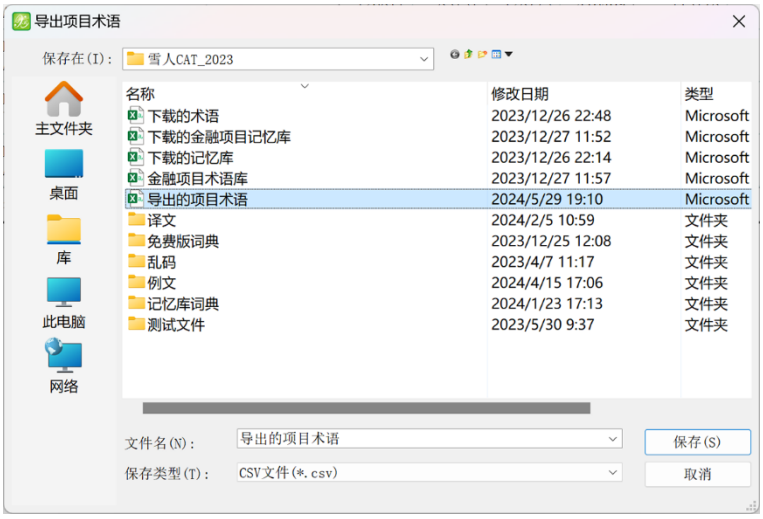


第六节 导出项目所使用的术语

加载术语库后，如果需要了解当前的翻译项目使用了哪些术语，或者需要将这部分术语导出发给其他人，就可以用“导出项目术语”功能来解决。选择屏幕左侧 项目文件 下的相应的文件夹，再点击**【右键菜单】➔【导出项目术语】**功能后，选择所需的文件后，可以导出这些文件中使用到的术语。（注意：如果左侧是多层次的复杂的项目文件结构，要先点击最底层的文件后，再点击右键菜单，不支持在中间层次的“子文件”位置上点击右键菜单“导出项目术语”。）如下图：



选择具体的文件后，点击“确定”按钮。在弹出的窗口输入文件名后导出。



会将当前项目所使用的术语导出成一个“.csv”格式的文件，这个文件可以用 EXCEL 打开它。如下图：

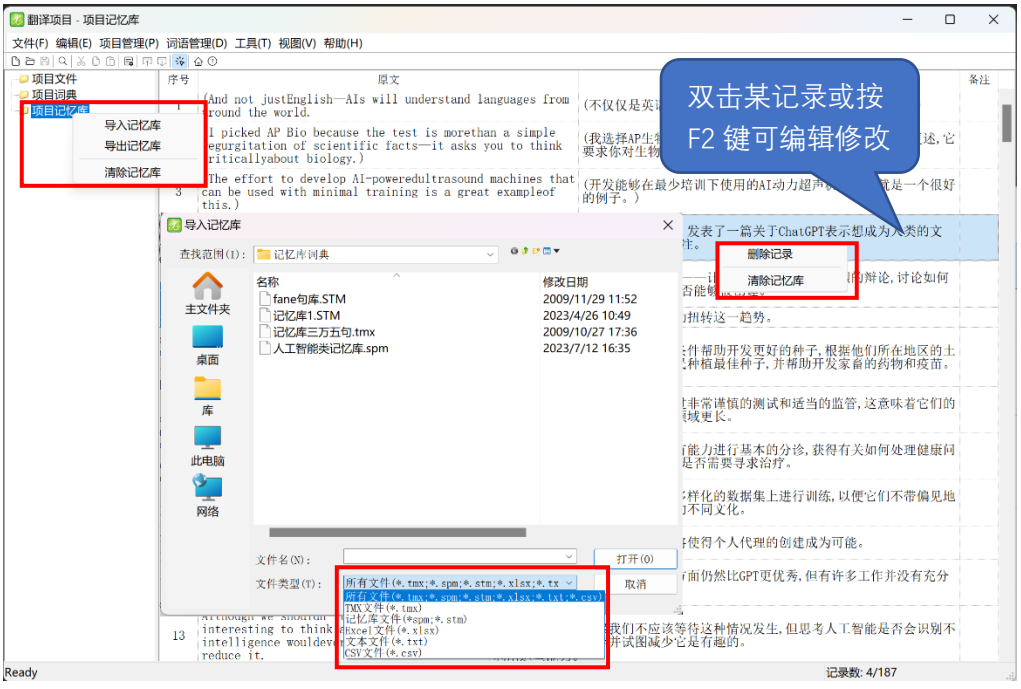
第九章 记忆库

第一节 记忆库的导入导出

新版由于支持直接引用项目文件做记忆库（详见本章第三节），这种方式既保留了原始篇章信息，还可以作上下文匹配分析，而且不必担心若原句修改了，记忆库中的那句有没有修改的问题？所以，对已经翻译过的句子不再进“项目记忆库”（旧版是进“项目记忆库”的），已经翻译的资源若在雪人软件中使用则直接引用项目文件做记忆库，若需要用在其他 CAT 软件中，可以在“项目文件”中导出记忆库。新版的“项目记忆库”，主要用于记忆库格式的转换、维护、合并。

记忆库的导入

右键点击左侧的“项目记忆库”，会弹出右键菜单，点击**【右键菜单】➔【导入记忆库】**功能，可以将多种格式的语料导入雪人软件中。



雪人支持导入的记忆库文件格式有：

.spm——新版雪人的记忆库格式，外挂记忆库需要转换成这种格式。

.stm——旧版雪人的记忆库格式。旧版的记忆库需要在“项目记忆库”中进行导入后再导出，才能转换成新版的格式外挂引用。

.tmx——各款 CAT 软件记忆库的通用交换格式。

.txt——文本格式，原文、译文之间要有分隔符。

.csv——CSV 格式，EXCEL 文件可以在 EXCEL 软件中另存为.csv 格式，然后导入雪人中。

记忆库的编辑

将语料导入“项目记忆库”后，可以进行编辑。点击右键，会弹出右键菜单，有【删除记录】、【清除记忆库】功能。

修改——双击某词条或按【F2】键，进入编辑修改状态。

删除——点击某条记录后，点击**【右键菜单】➔【删除记录】**功能，则删除一条记录。若按住【Shift】键+【鼠标单击前后记录】，可选择多条连续的记录进行批量删除；按住【Ctrl】键+鼠标点选，可以选择多条不连续的记录进行批量删除。

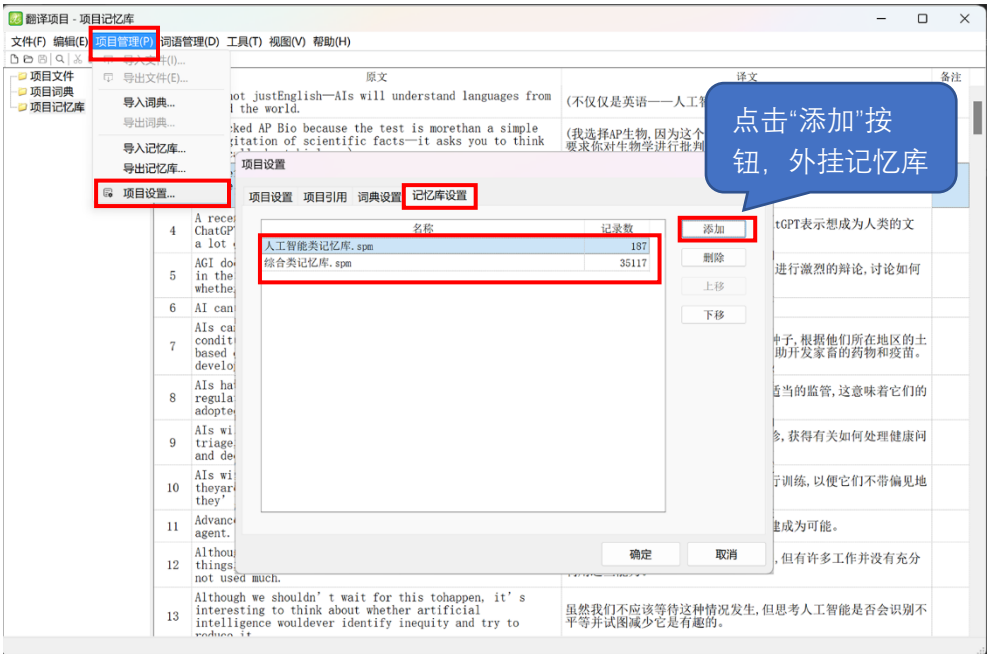
排序——点击表头的原文，译文、词性等，可以进行排序。粗体的术语会排在最前面。

记忆库的合并

如果有多个记忆库文件，希望将它们合并成一个库，则可以在同一个项目文件的“项目记忆库”中导入这些文件，然后点击**【右键菜单】➔【导出记忆库】**功能即可生成一个合并了的记忆库文件，而且会自动删除记忆库中的重复句子。

第二节 记忆库的外挂

积累的语料可以在“项目记忆库”中直接导入使用，也可以通过外挂记忆库文件的方式来使用。外挂记忆库的好处是减少内存的消耗。在“项目记忆库”中将句库导出成“.spm”格式文件，就可以被以后的翻译项目添加引用。方法是：点击菜单**【项目管理】➔【项目设置】**（或点击工具栏的“项目设置”图标），再点击“记忆库设置”标签页，点击“添加”按钮，可以加载多个“.spm”格式的记忆库文件，根据需要调整各个库的优先级顺序，放在上面的比下面的优先。如下图：



新版支持直接将翻译的项目文件（.SPR ）作为记忆库来引用，也是属于记忆库外挂的调用方式，具体请参看本章的第四节。

第三节 记忆库的上下文精确匹配

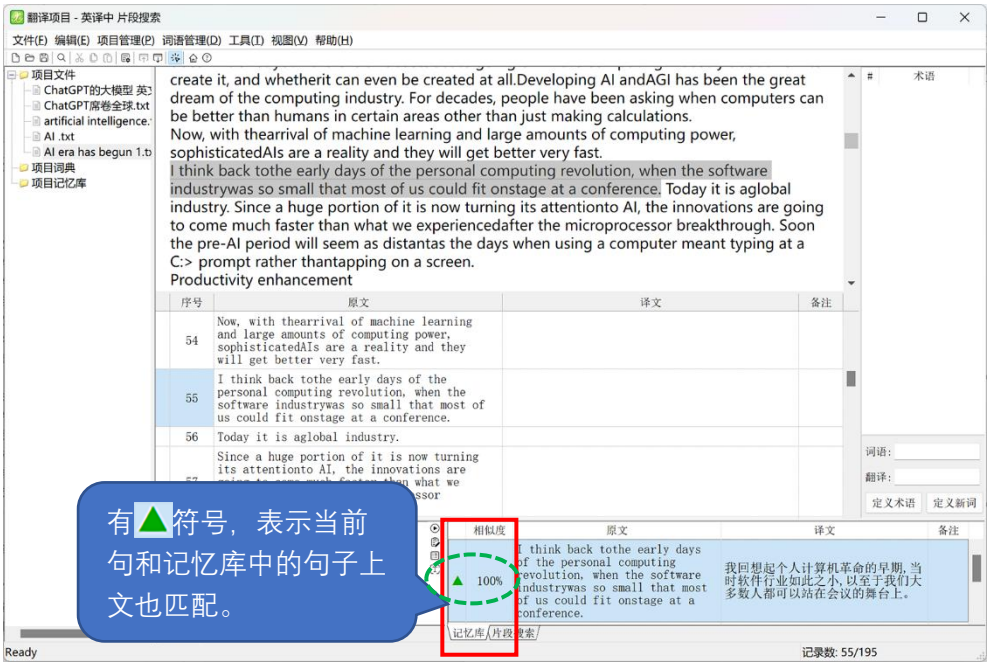
CAT 的一个重要功能是“记忆翻译”，即对已有语料进行搜索和匹配，找出相同或相似的译文，在翻译过程中使用翻译记忆库，可以减少重复劳动，帮助译者更高效地完成翻译工作。计算出待翻译文档中的句段和记忆库中的句段相似程度，搜索出来供译者翻译时参考，相似度 50%~99%为模糊匹配；相似度为 100%为精确匹配；100% 匹配就是完全相同了，吗？当然不是，同样一句话，语境不同，翻译就不一定相同。在精确匹配基础上，和前一句段也精确匹配，为上文匹配；在精确匹配基础上，和后一句段也精确匹配，为下文匹配；在精确匹配基础上，不但和前一句段精确匹配，而且和后一句段也精确匹配，为上下文匹配。

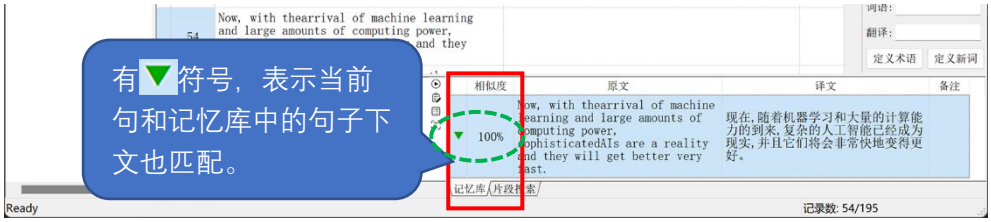
除了支持传统的记忆库模糊匹配和 100%的精确匹配外，还支持上文匹配、下文匹配、上下文匹配的精确匹配。在相似度前面，分别用以下**绿色**符号标注，以示区分不同的匹配精度。

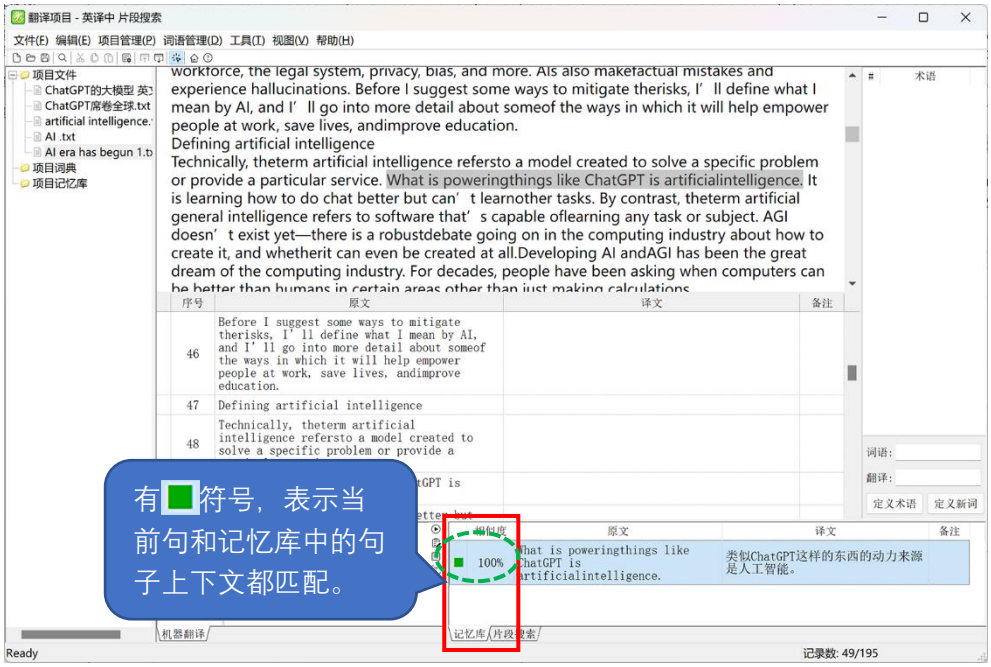
▲——**上文匹配**：在相似度 100%的精确匹配基础上，和前一句段也精确匹配。

▼——**下文匹配**：在相似度 100%的精确匹配基础上，和后一句段也精确匹配。

■——**上下文匹配**：在相似度 100%的精确匹配基础上，不但和前一句段精确匹配，而且和后一句段也精确匹配。









如果在相似度前面，没有显示任何标志符号，则和旧版约定一样，没有上下文匹配的信息。

例如对产品手册、技术文档、政府工作报告、学术论文等文档的翻译，由于这些文档通常在不断地更新，新版本与旧版本之间可能存在较多的相似之处，充分利用记忆库的上下文匹配技术，可以减少出错的几率和提高翻译的工作效率。

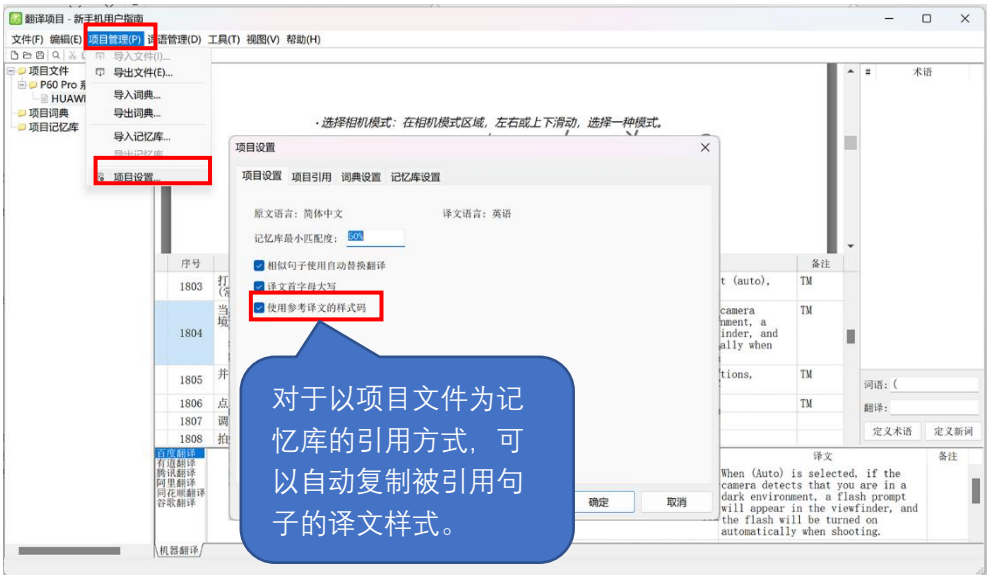
第四节 直接引用翻译的项目文件作记忆库

在新版中，支持直接将翻译的项目文件（.SPR ）作为记忆库来引用，包括这个项目文件下面的“项目词典”和“项目记忆库”中的内容，都可以一并打包被其他项目引用。这样做的好处是：可以保留被引用句子的上、下文信息，记忆库搜索匹配时，除了判断本句 100%匹配外，还可以判断是否和被引用的句子上文一致？下文一致？及上下文都相同？而且还可以直接套用被引用句子的样式，在全文预翻译时，会优先选用这些上、下文都 100%精确匹配的句子。传统的记忆库，是拆成以句子为单位保存的，没有办法保留原文的整体篇章信息、语境及上下文的关系。在新版中，用户可以不必太过关注记忆库的导入、导出等问题，直接引用翻译的项目文件即可实现资源的再利用，这对于翻译一些产品说明、手册更新等重复率较高的技术文档，更显优势。

例如下图中：已经翻译好了“多个不同型号的手机用户指南”。



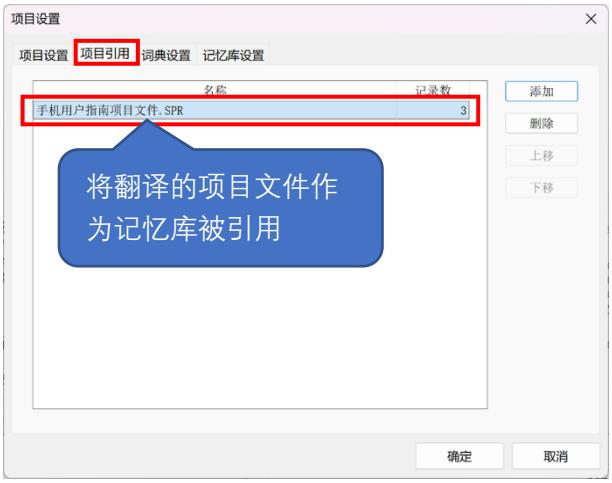
现在又有一个“新的型号的手机用户指南”需要翻译，由于这类技术文档，存在较高的相似性，之前翻译过的“手机用户指南”肯定能给现在要翻译的文档带来帮助，所以直接将之前翻译过的项目文件作记忆库使用，将它引用到新的翻译项目中。具体操作：新建一个新的翻译项目后，导入需要翻译的文档，点击**【项目管理】➔【项目设置】**菜单，会弹出“项目设置”窗口。如下图：



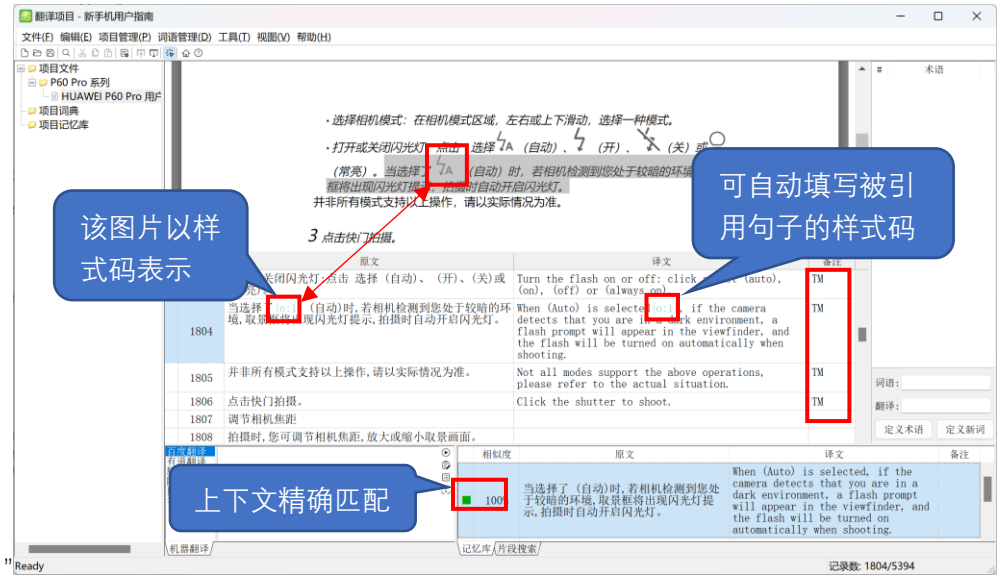
在“项目设置”窗口，会显示当前项目的语言方向，被引用的项目文件，语言方向应该与当前的项目文件一致。“相似句子使用自动替换翻译”选项，勾选它后，对于和记忆库中结构相似的句子，会作自动替换翻译。“使用参考译文的样式码”选项，对于以项目文件为记忆库的引用方式，可以自动复制被引用句子的译文样式，对于引用的是普通的记忆库，该选项则无效，因为普通的记忆库保存时不保存样式数据。

点击“项目引用”页签，可以添加引用多个已经翻译好的项目文件，文件后缀是“.SPR”，即直接引用翻译的项目文件作记忆库，包括被引用的项目文件下的“项目词典”和“项目记忆库”中的内容，都可以被引用。“项目引用”的好处是有完整的篇章信息、可以判断上文是否精确匹配，还可以省去记忆库导入、导出的工作。坏处是会占用比单纯的记忆库更多的内存，搜索速度也会比单纯的记忆库搜索速度慢一些。注意：如果当前翻译的项目文件为“A”，在“A”中引用了“B”项目文件，现在又同时打开了“B”项目文件并进行了修改，此时保存“B”项目文件时会出现无法保存的情况，解决办法是:暂时关闭“A”后，让“B”保存，或者将“B”另存。

点击“词典设置”页签，可以添加引用词典文件。点击“记忆库设置”页签，可以添加引用记忆库文件，这些和旧版操作一样。



设置好引用项目后，点击【工具】➔【预翻译】菜单，对两个项目文件中精确匹配的句子，会自动填写译文，包括译文的样式码都一并填写。例如：下图的那句，在下方的“记忆库”提示区中，相似度显示 100%，而且前面还有一个“ ”图标，表示该句上下文都精确匹配，自动填写的译文还包括{o:1}这个样式码，它代表在译文处要插入图片，即这个插入的图片也一并自动处理了。

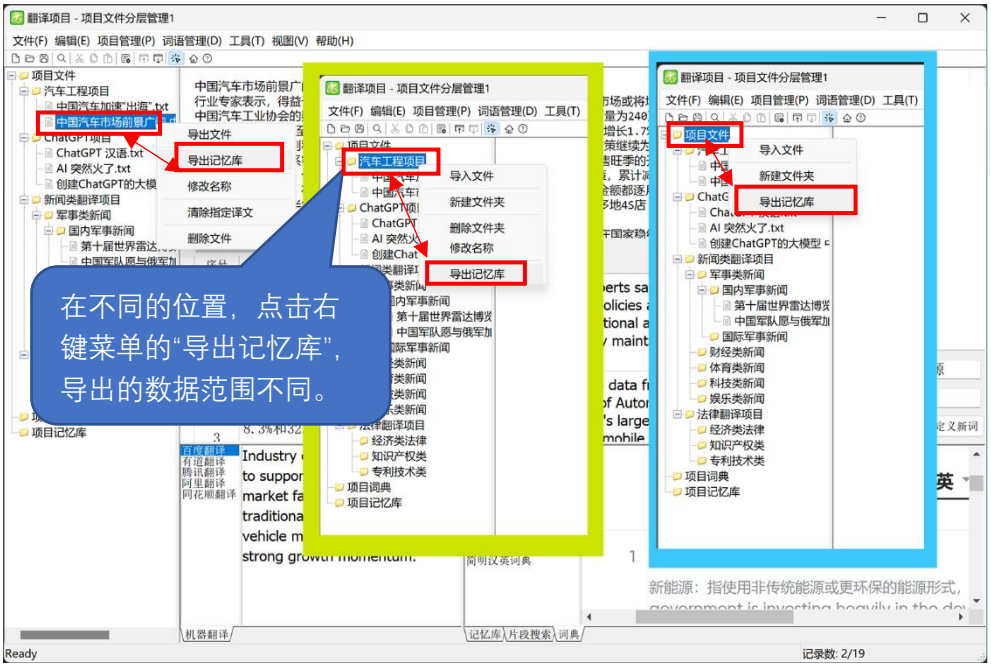


第五节 项目文件的记忆库导出

雪人新版支持将翻译的项目文件作为记忆库被其他项目引用，而且保留被引用句子的上下文信息及样式码等信息，所以，翻译的项目文件的句子不再自动进屏幕左侧的“项目记忆库”（旧版是自动进项目记忆库）。用户可以在“项目文件”中直接导出这些记忆库，在不同的文件夹层次上点击**【右键菜单➔导出记忆库】**功能，可以导出该文件夹下面的内容，导出的记忆库是合并的数据，并且重新排序。例如：下图中，若先点击屏幕左侧的“中国汽车市场前景”这个具体的翻译文档，再点击**【右键菜单➔导出记忆库】**功能，则仅仅导出“中国汽车市场前景”这个文档的数据。

若点击“汽车工程项目”这个文件夹，再点击**【右键菜单➔导出记忆库】**功能，则导出的是属于“汽车工程项目”这个文件夹下面的数据。

若点击最上面的根目录——“项目文件”，再点击**【右键菜单➔导出记忆库】**功能，则导出根目录下面的所有数据，即所有数据合并、重新排序导出。



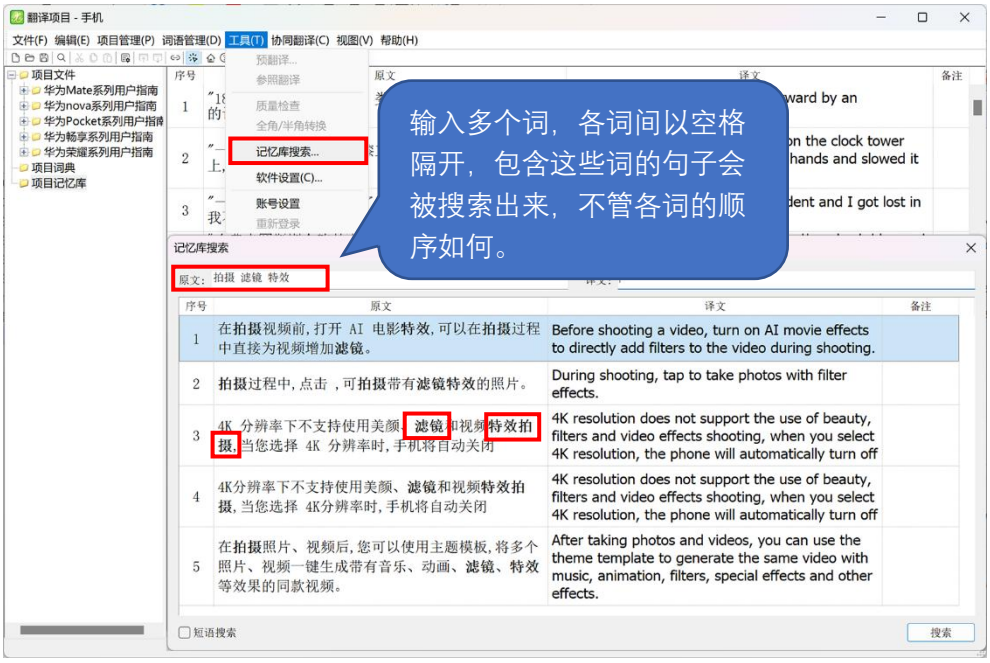
第六节 记忆库搜索

支持独立的“记忆库”搜索，搜索条件可以自由设定，搜索的范围包括：当前翻译的项目文件、项目记忆库以及添加引用的外部记忆库。

搜索条件灵活多样，可单独搜索原文或单独搜索译文，当原文和译文搜索框中同时输入内容时，表示搜索的结果需要同时满足条件才显示出来。可以按顺序或不按顺序搜索多个词，各词之间以空格隔开。亦可搜索带…… 的短语，如：a？of ；a * of 等。

下面举例说明：

点击【工具】➔【记忆库搜索】菜单，会弹出记忆库搜索窗口。在原文搜索框中输入：拍摄 滤镜 特效 这三个词，中间以空格隔开，可以搜索出原文包含这三个词语的句子，各词出现的顺序不限。如下图：

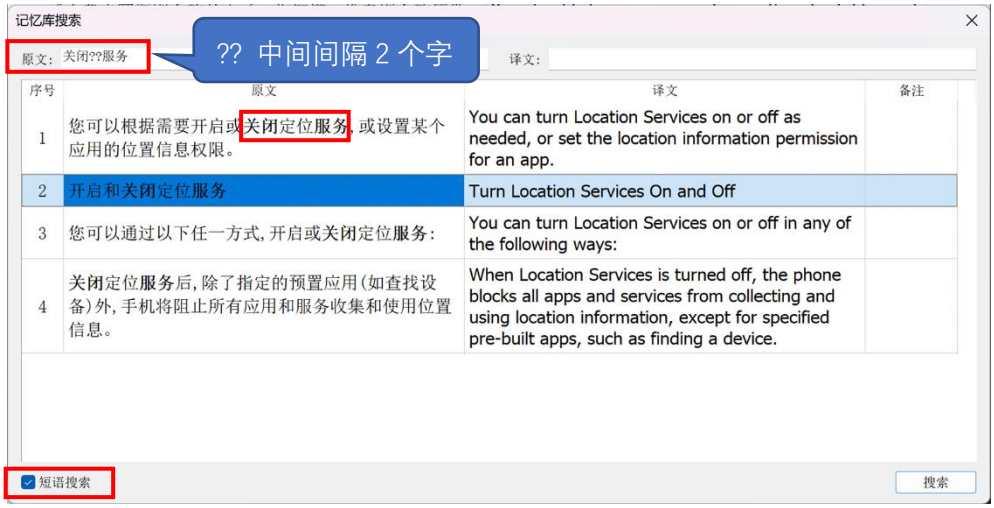


若在原文搜索框中输入：关闭服务　这四个字，则完全匹配这四个字的句子才会被搜索出来。如下图：



点击搜索窗口左下角的“短语搜索”选项，将它设为“√”。在短语搜索中，有两个常用的特殊符号：**“?”和“*”**，“?”表示匹配一个汉字或一个单词，“*”表示匹配任意个词，注意：**“?”和“*”都要在英文状态下输入半角的**。使用**“?”和“*”**进行中文的搜索时，注意不要留空格，因为空格也是作为一个匹配单位来计算的；搜索英文则需要输入空格。

例：在原文搜索框中输入：关闭??服务，一个?号表示一个字，两个??号表示中间有两个字，搜索的结果是：关闭和服务之间间隔 2 个汉字的句子会被搜索出来。如下图：



若在原文搜索框中输入：关闭*服务 则表示：关闭和服务之间包含任意个字的句子会被搜索出来。如下图：



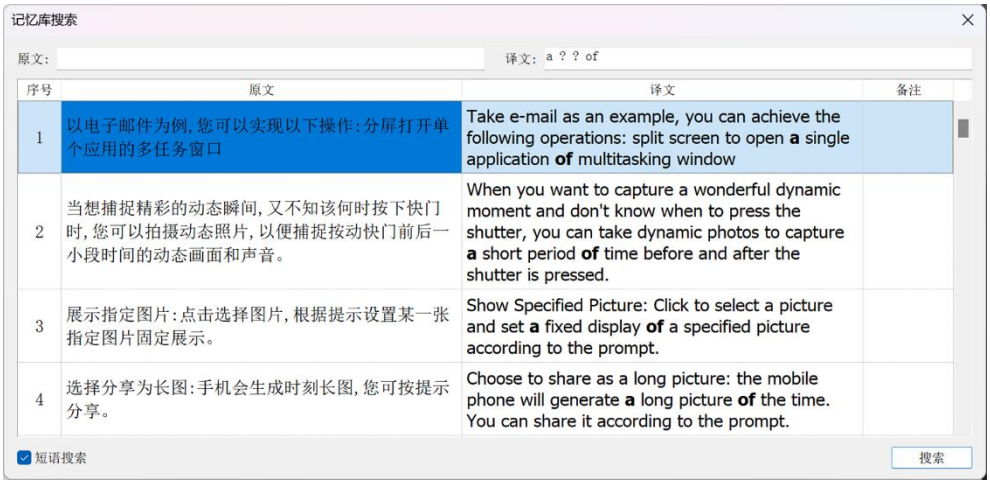
对英文的搜索操作也类似，例如：在译文搜索框中输入：a * of 会搜索出 a 和 of 之间有任意个单词的句子，如下图：



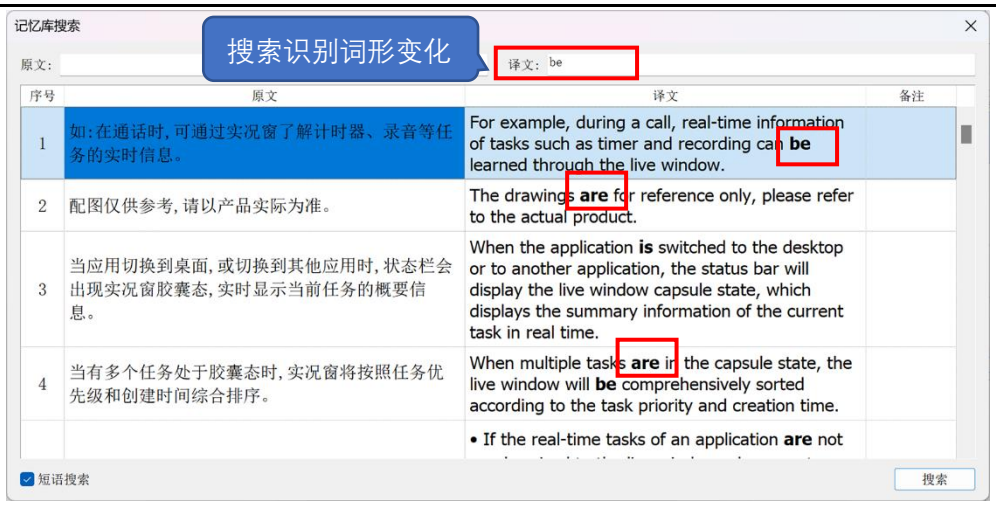
若在译文搜索框中输入：a ? of 会搜索出 a 和 of 之间仅一个单词的句子，例如：搜索出 a period of、a group of 等这些句子。如下图：



若在译文搜索框中输入：a ? ? of 会搜索出 a 和 of 之间间隔两个单词的句子，例如：搜索出 a short period of、a fixed display of 等这些句子。如下图：



搜索时会识别词形变化，例如搜索 be，则 am、is、are、be 的句子都会被搜索出来。如下图：



第七节 片段搜索

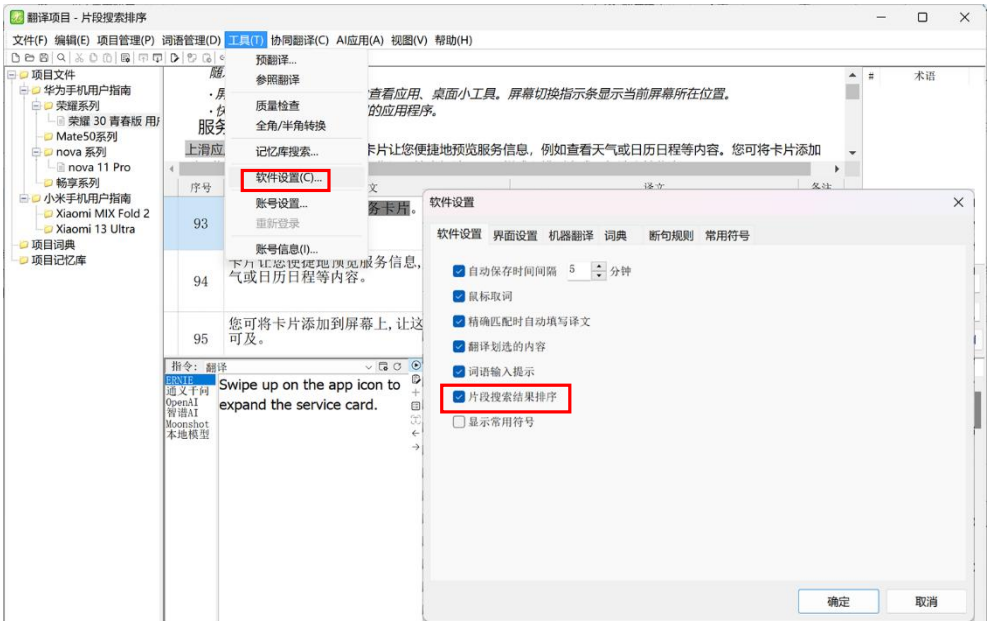
点击屏幕下方的“片段搜索”页签，鼠标划选原文某个词，可以从记忆库中查找该词在记忆库中的译法。片段搜索支持识别词形变化，例如：鼠标划选原文的 factories,可以从记忆库中搜索出含有 factory 和 factories 的所有句子。片段搜索划选的词语最长支持 30 个汉字或单词。

旧版的片段搜索对长句不能自动换行，新版的片段搜索的结果以表格形式显示，对于长句可以自动转行，浏览起来更加方便些。如下图：

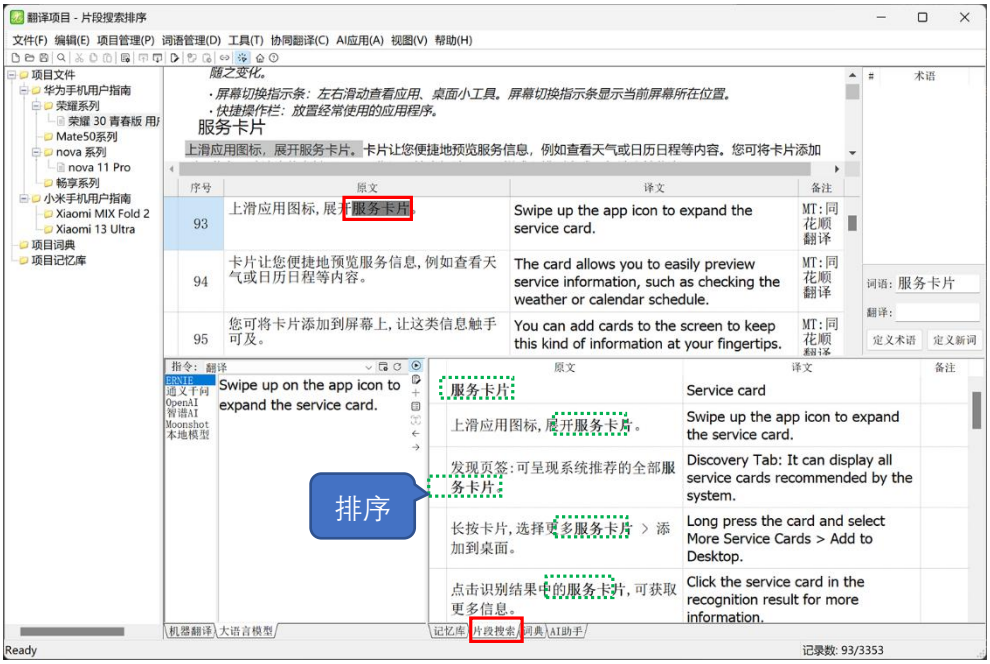


第八节 将片段搜索的结果排序

点击【工具】➔【软件设置】菜单，在“软件设置”窗口中，选择“片段搜索结果排序”这个选项可以开启这一功能。如下图：



鼠标划选原文后，在屏幕下方的“片段搜索”窗口，会按匹配句子的长度进行排序，简短的句子排在最前面。如下图：



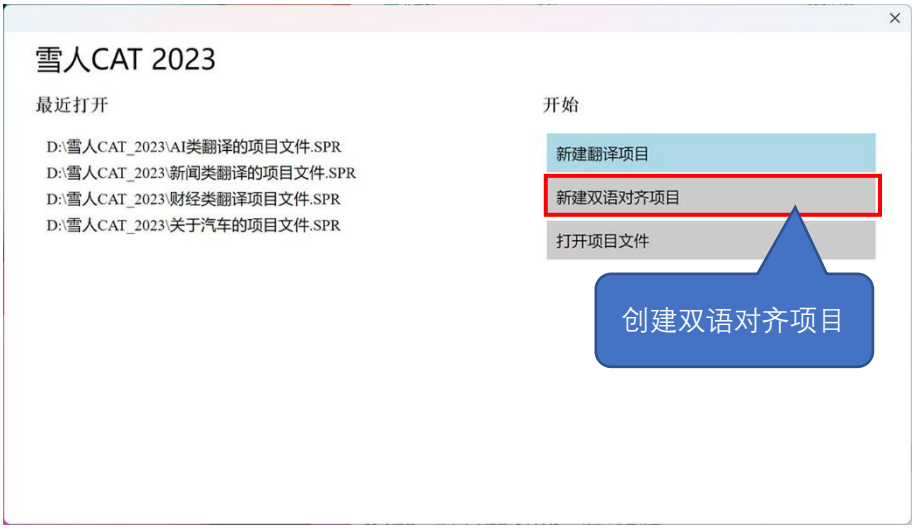
第十章 双语对齐建记忆库

第一节 创建多层结构的双语项目

CAT 软件要充分发挥作用离不开专业化、高质量、大型翻译记忆库的支持。记忆库是译者的宝贵财富，拥有大型、高质量的记忆库，可以成倍地提高翻译的效率。翻译记忆库既可以通过译员个人翻译逐渐积累，也可以直接将自己已翻译过的或是收集到的他人的双语文本制作成翻译记忆库。雪人旧版软件集成了功能强大的双语对齐工具，自动对齐效果业界一流，多年来得到大量用户的好评，利用它可以很轻松的创建大型记忆库。

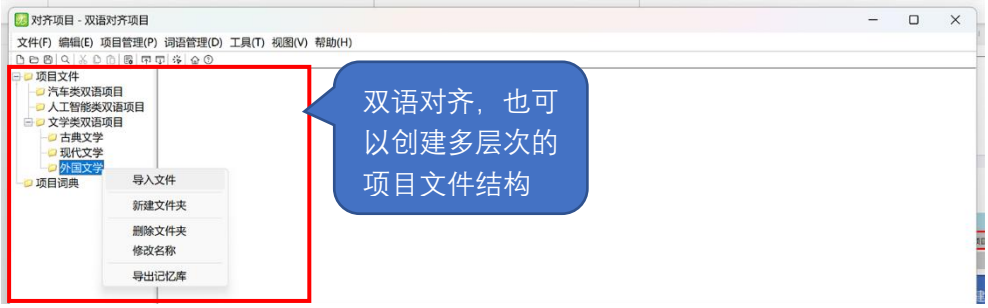
雪人新版在双语对齐中使用了更先进的技术，使对齐的结果更加准确，对齐界面的编辑也更加灵活，自动对齐后还可以进行二次分区域的局部分块对齐。也可以导入时先不自动对齐，导入后分章节、分块对齐，这样则对单次导入的句子数量没有限制，更加方便对大型双语语料的对齐。新版的双语对齐项目，还支持创建多层次、复杂的项目文件结构，方便对语料的归类管理。

在下方的开始窗口中，点击新建双语对齐项目（或点击**【文件菜单】➔【新建】➔【双语项目】**功能）。



创建一个双语对齐项目，新版本的双语对齐，支持创建多层次的项目文件结构，以方便归类管理。如果不需要创建多层次的分类目录结构，则右键点击左侧的**【项目文件】➔【导入文件】，**可以导入需要对齐双语资料。如果需要分类整理对

齐的双语资料，则可点击**【右键菜单】➔【新建文件夹】**，创建好分类目录文件结构，然后在相应的目录文件夹中导入双语文件。如下图：



右键点击左侧的某级分类的项目文件夹后，点击**【导入文件】**，弹出“导入文件”窗口，如下图：

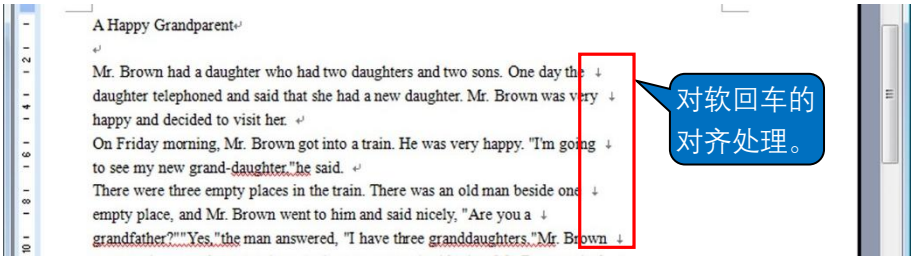


窗口中的选项解释如下：

使用词典辅助对齐——勾选这个选项，在对齐处理时会暂用一定的处理时间，比不选用它会稍微慢一些。如果预先对整篇对齐的文档进行词频统计分析，将文章中的高频词、术语、人名、地名等固定译法先定义在项目词典中，对齐时再勾选“使用词典辅助对齐”选项，可以进一步提高对齐的准确度。

导入时对齐——勾选这个选项，导入时就会自动进行对齐，但对单次导入对齐的句子数量有限制，单次对齐大约 2 万句，如果文件太大，需要分拆多次导入。如果不勾选，导入时则不自动对齐，导入后可以指定范围，分区域进行对齐，具体方法：导入后，按**住【Shift】键+ 鼠标左键**单击选择前后范围，然后点击**【右键菜单】➔【部分对齐】**功能，可以就所选择区域范围内进行对齐。若导入时不自动对齐，对单次导入的句子数量没有限制，更加方便对大型双语语料的对齐。

检测软回车——有时会遇到一些文章中含有许多软回车，它是由于排版等其他原因造成的，在不该断句的地方转行了，如下图：对这种文档的对齐，就需要勾选“检测软回车"选项才能有好的对齐效果。



对齐时可合并句子——不勾选，对齐时将不会自动合并句子，对句子是一一对应的文章的对齐效果会比较好，可以避免因为合并句子而造成的对齐问题，提升对齐的准确性。对一句原文对应多句译文的情况，对齐时由于限制自动合并句子，对齐效果不一定会更好，用户可以两种方式都测试一下，哪种对齐效果好就选用哪种。

拆分段落中的句子——勾选，则导入对齐时，会将段落中的句子拆分成一句一句对齐。若不勾选，则以段落为单位进行段与段的对齐，进行段落对齐后，利用锁定区域后，选择局部内容，进行二次拆分对齐。对于大型语料，可以使用先按段落为单位的初对齐，然后再以句子为单位进行第二次对齐。

第二节 各种文档格式的对齐

一、直接导入文件对齐

如果需要对齐的语料是一篇原文、一篇译文，分属在不同的文件中，则点选文件存放路径直接导入文件进行对齐，如下图：



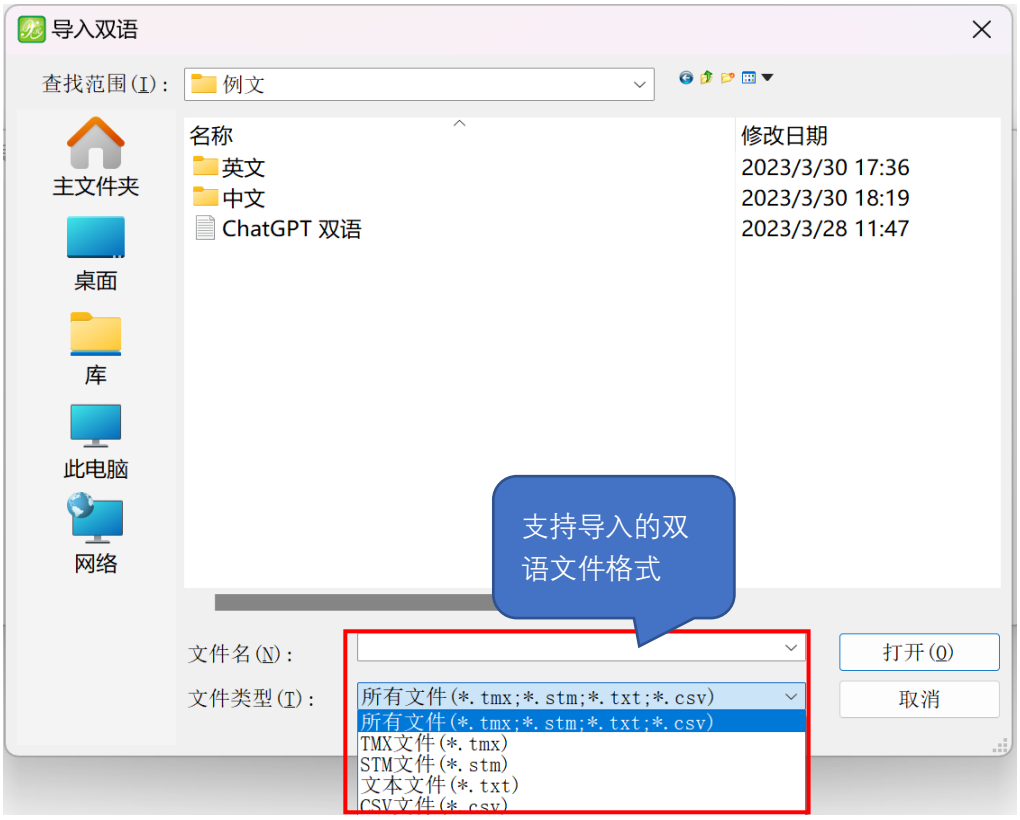
分别选择原文、译文相对应的文件并根据具体需要勾选相应的选项后，点击“确定”按钮，自动对齐结果如下：左侧“项目文件”下面的默认文件名为“Untitled”，可以右键点击它后，利用右键菜单修改文件名。



二、导入双语对齐

如果同一个文件中包含原文、译文的的双语语料，则点击导入双语按钮，如下图：





雪人支持的导入双语格式有：.txm、.stm、.txt、.csv 等格式。其中 TMX 格式是各种 CAT 软件的记忆库通用的交换格式，STM 是雪人旧版的记忆库文件，如果是 EXCEL 格式的文件，可以在 EXCEL 中另存为.csv 格式再导入。

三、从编辑器导入

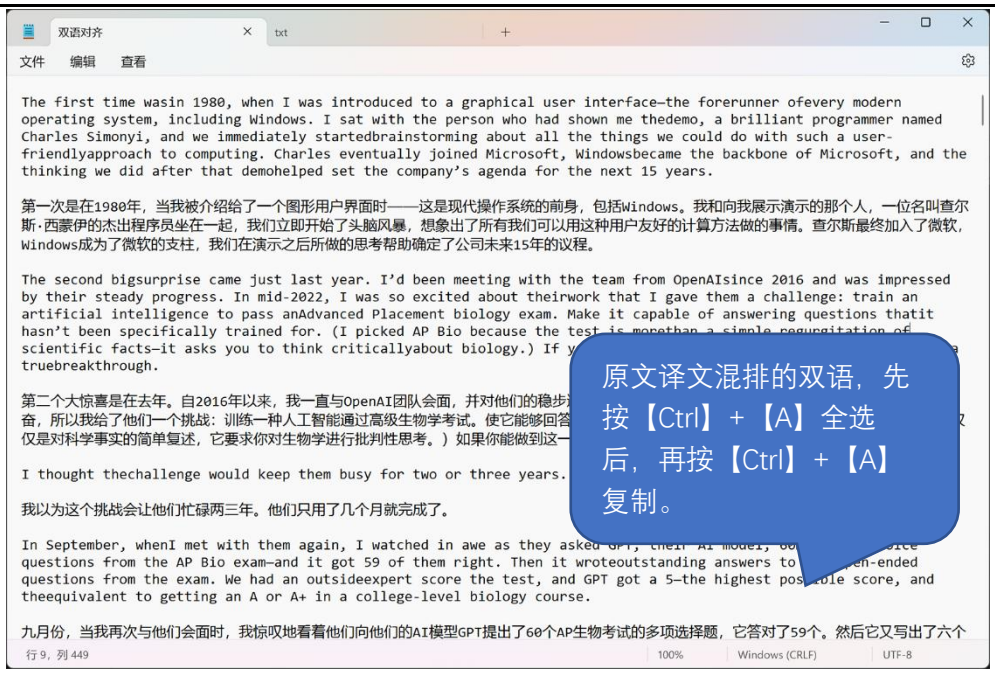
对于原文、译文可以分开进行复制、粘贴的，则点击从编辑器导入按钮，如下图：将原文、译文分别复制粘贴在以下窗口进行对齐，也可在此窗口进行编辑、删除、修改等操作，适合从网上或其他地方收集双语资料。





三、从剪切板导入的对齐

如果是原文、译文混在一起的双语资料，一段原文、一段译文，原文和译文不好分开，如下图：



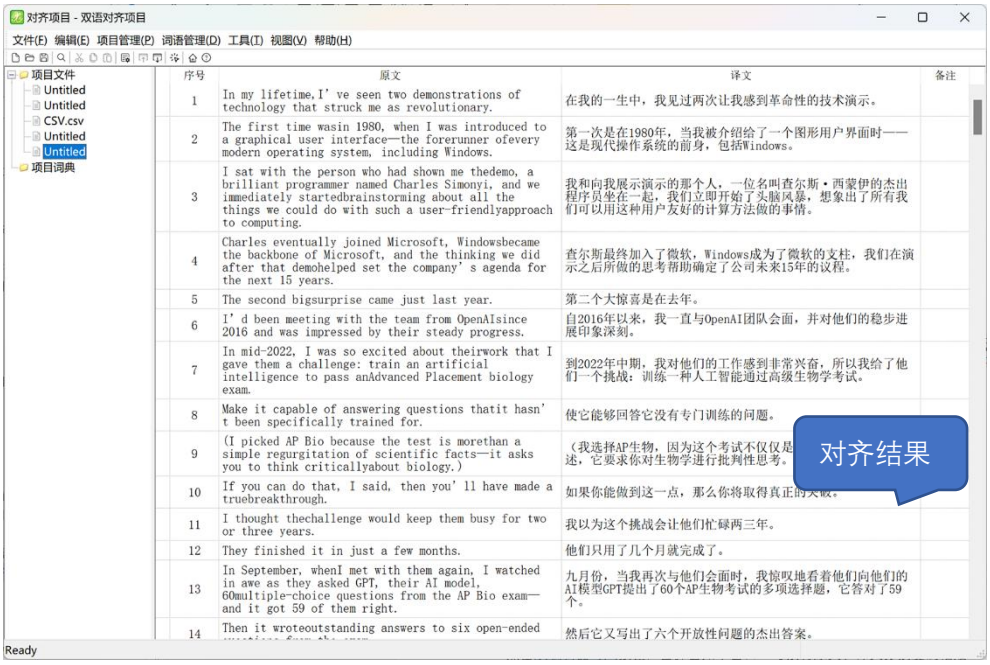
处理这种语料的对齐，应先按【Ctrl】+【A】全选后，再按【Ctrl】+【C】复制到剪切板。



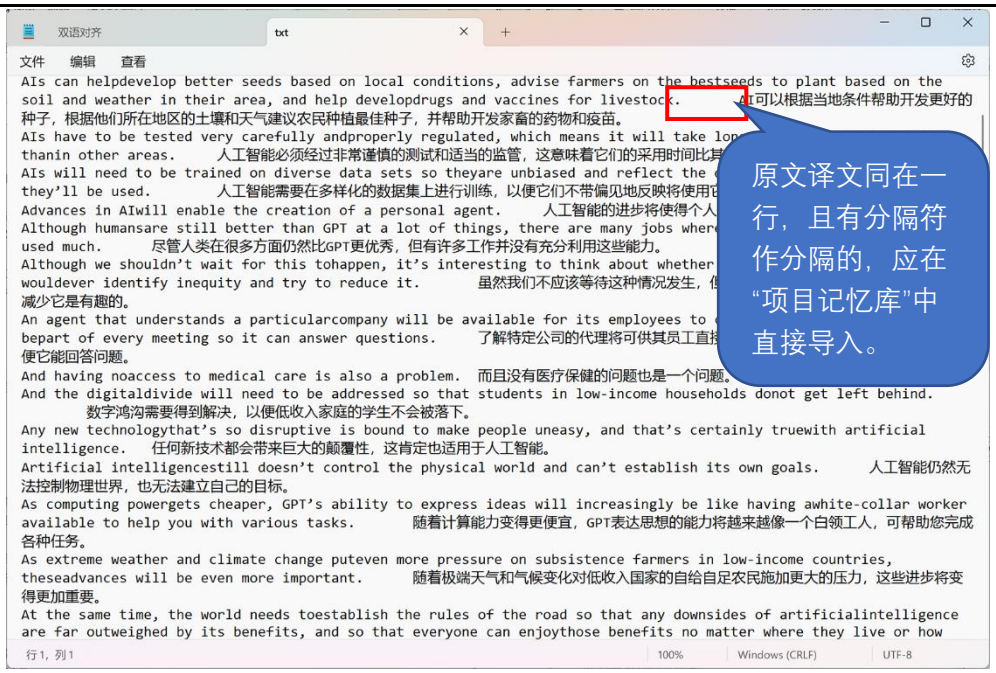
然后点击从剪贴板导入按钮，雪人会自动区分原文、译文。如下图：



点击"确定"按钮，对齐结果如下图：



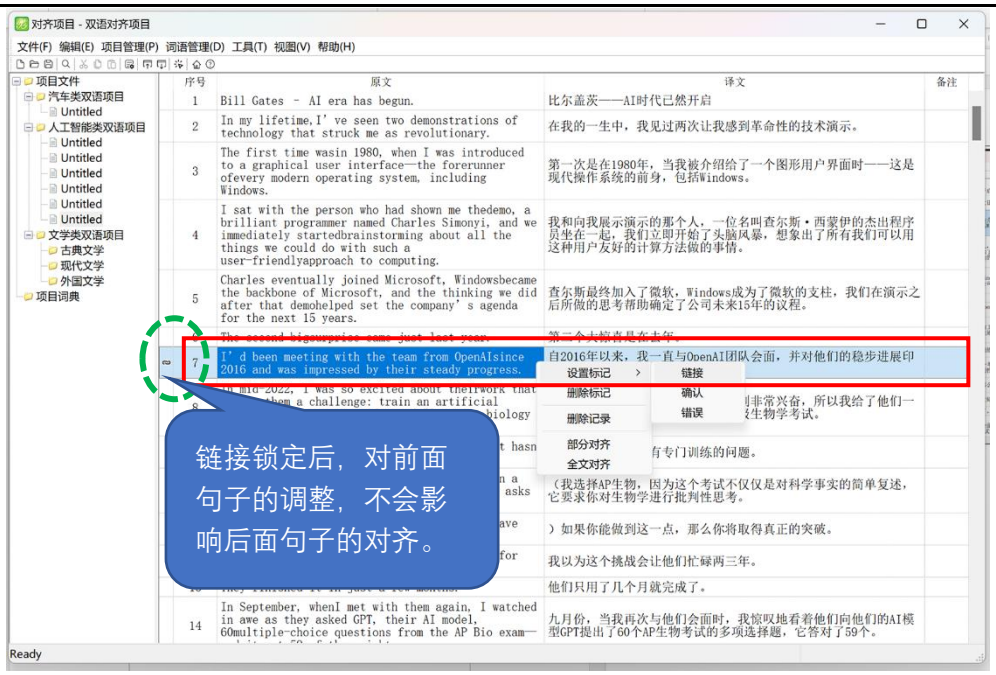
如果原文、译文同在一行，且原文译文之间是以制表符作分隔的，原文译文同在一行，且有分隔符作分隔的，应在"项目记忆库"中直接导入。如下图：



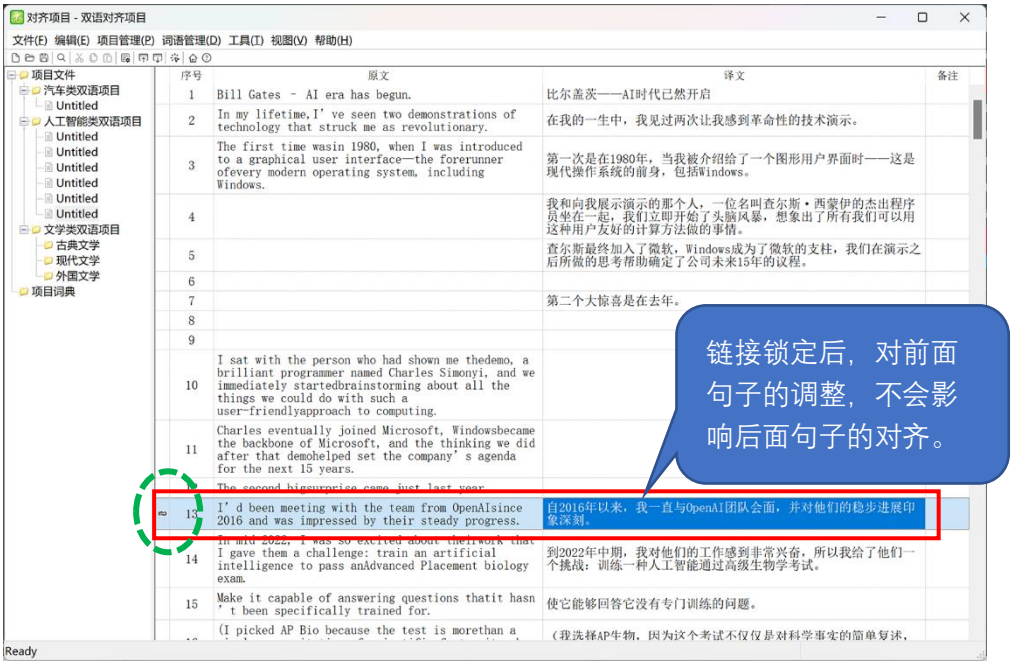
在双语对齐项目进行“词频统计”，统计出语料中的高频词语并将它们加入到“项目词典”中，特别是一些地名、人名、机构名、专业术语等，先将这些词进行定义，可以提升双语对齐的准确度，具体请参看“词频统计”章节。

第三节 对齐后的编辑

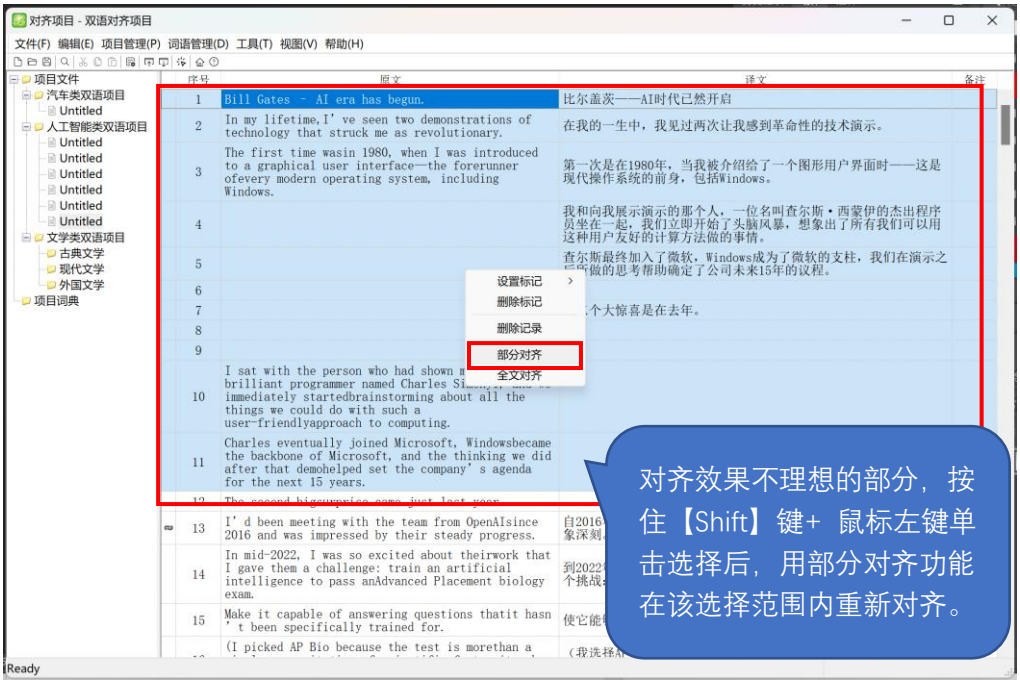
在对齐界面中，点击**【右键菜单】➔【设置标记】➔【链接】**功能，会在该句的最左列中添加一个“锁定链接”标记，例如下图的第 7 句：调整时，表示这句及其后面的句子始终会一起移动。



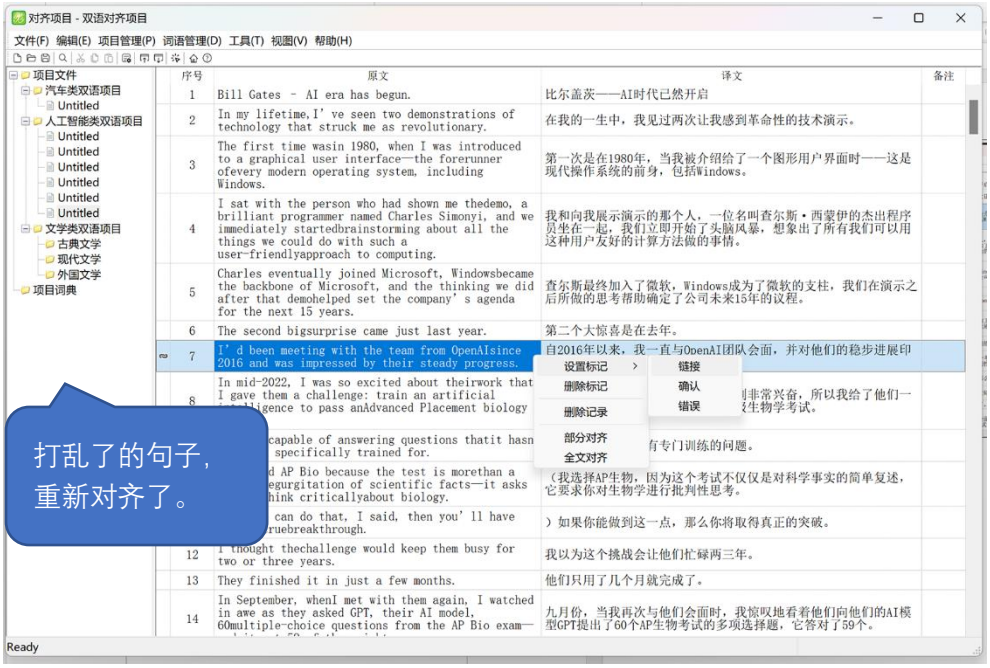
此时，对第 7 句之前的句子，进行修改、例如插入行、删除等，不会影响它及之后句子的对齐。如下图：在第 7 句之前插入了很多空行，但并不会影响其后面句子的对齐（按【**Insert**】**键**插入空行）。



对齐效果不理想的部分，按**住【Shift】键+ 鼠标左键**单击选择后，用**【右键菜单】➔【部分对齐】**在该选择范围内重新对齐。



对齐结果如下图：打乱的句子，又重新对齐了。



对齐后编辑的常用操作键：

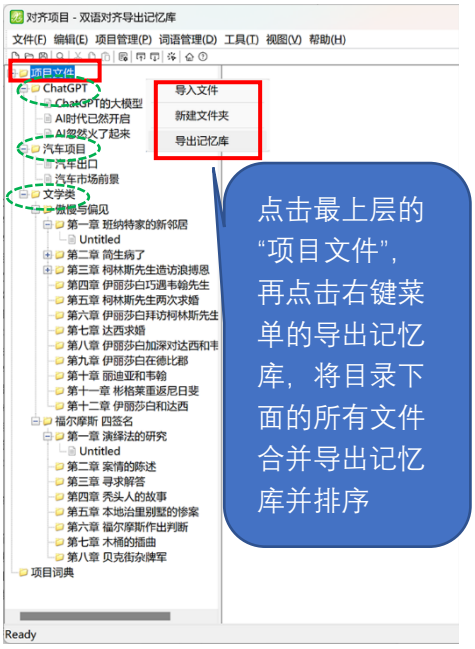
句子拆分：双击鼠标左键或按【F2】键进入编辑状态后，将光标放在需要拆分的地方按**【Alt】+回车键**将一句拆成两句。
句子合并：双击鼠标左键或按【F2】键进入编辑状态后，将光标放在句尾，按**【Delete】**键可以将上下两句合并。
插入空行：按**【Insert】**键插入空行。
移动句子：双击鼠标左键或按【F2】键进入编辑状态后，按**【Alt】+【↑】或【↑】键，**可以将当前句子移动到空白行处。注意：长光标状态下或没有插入空白行状态下，按**【Alt】+【↑】或【↑】键**是没有反应的。

二次对齐

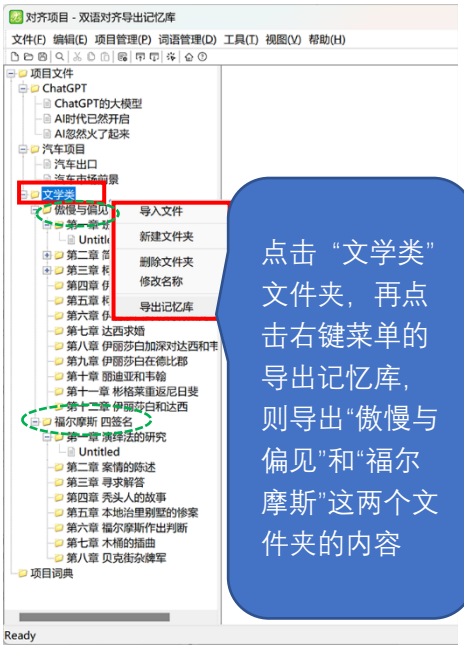
对于某些对齐效果不理想的句子，可以利用“锁定链接”标记功能进行区域划分，即对在某区域的前一句进行“锁定链接”，在区域的末尾也进行“锁定链接”，锁定后，可以在该区域中按**【Insert】**键插入空行，然后按住【Shift】键+ 鼠标左键单击选择前、后句子后，针对这个区域用**【右键菜单】➔【部分对齐】**再次进行局部的对齐。

第四节 双语对齐记忆库的导出

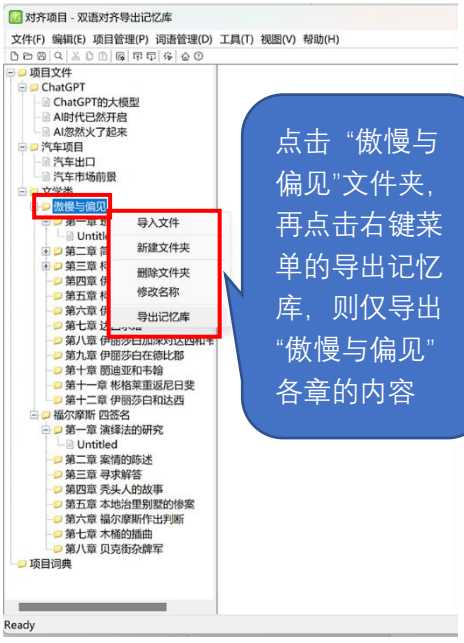
双语对齐项目中支持复杂的多层项目文件结构，对齐后，在屏幕左侧的目录结构中，点击不同的位置，再点击**【右键菜单】➔【导出记忆库】**功能，则导出不同范围的记忆库内容。例如下图中，划分了“ChapGPT”、“汽车项目”、“文学类”三个子项目，“文学类”中又划分了“傲慢与偏见”和“福尔摩斯”两个子项目，点击最上层的“项目文件”，再点击**【右键菜单】➔【导出记忆库】**功能，则将目录下面的所有文件合并导出记忆库，导出时会将这三个子项目的内容合并成一个记忆库导出并进行了排序。新版的记忆库文件格式后缀为“.SPM”，雪人旧版的记忆库文件格式后缀为“.STM”，旧版的记忆库文件可以在新版的“项目记忆库”中导入后再导出，转换成新版的“.SPM”记忆库文件使用。点击其他层次的文件夹，再点击**【右键菜单】➔【导出记忆库】**功能，则将导出属于该文件夹下面的内容，具体参看下面的截图：



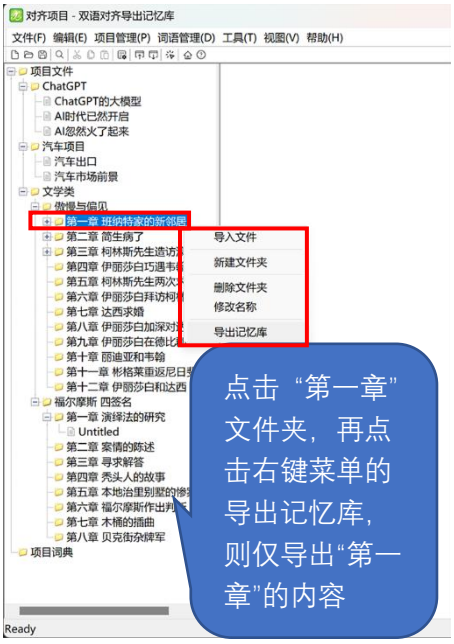
点击最上层的“项目文件”，再点击右键菜单的导出记忆库，将目录下面的所有文件合并导出记忆库并排序



点击“文学类”文件夹，再点击右键菜单的导出记忆库，则导出“傲慢与偏见”和“福尔摩斯”这两个文件夹的内容



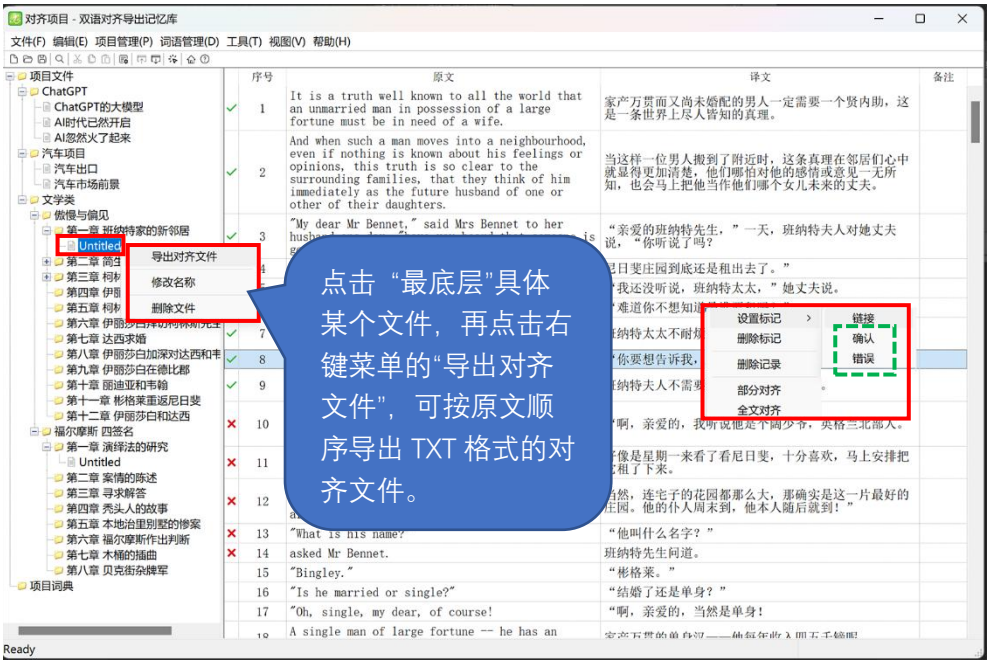
点击“傲慢与偏见”文件夹，再点击右键菜单的导出记忆库，则仅导出“傲慢与偏见”各章的内容



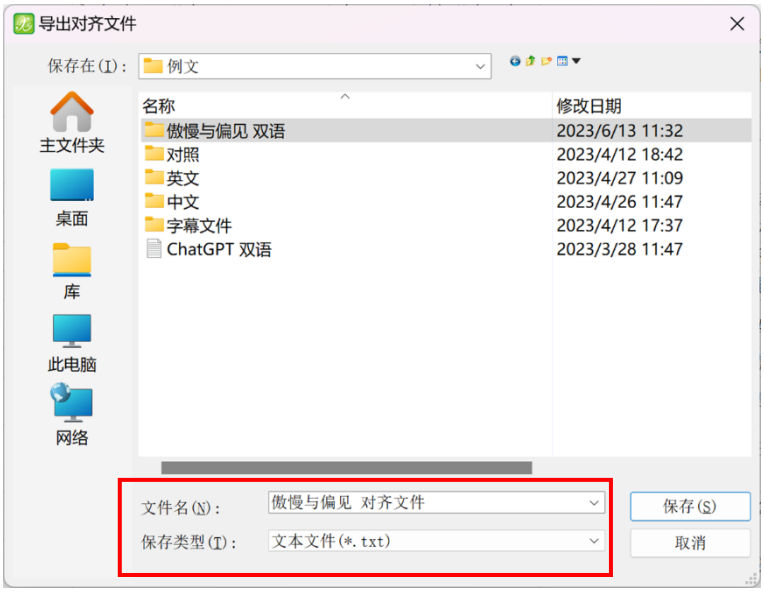
点击“第一章”文件夹，再点击右键菜单的导出记忆库，则仅导出“第一章”的内容

导出保持原文顺序的对齐文件

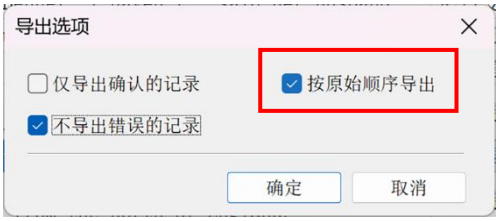
点击“最底层”具体某个文件，再点击【**右键菜单**】➔【**导出对齐文件**】功能，可按原文顺序导出 TXT 格式的对齐文件。如下图：



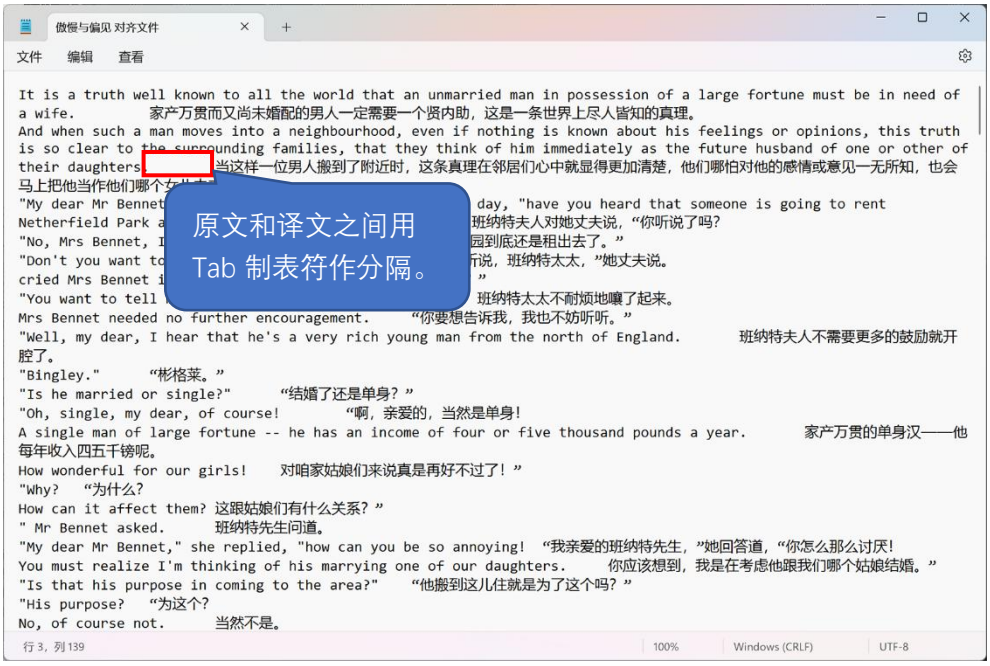
导出时，输入导出的文件名，如下图：



此时会弹出如下的“导出选项”窗口，其中“仅导出确认的记录”是指校对打了✓的正确句子，“不导出错误的记录”是指不导出打了 X 的句子。选择某些句子后点击**【右键菜单】➔【设置标志】➔【正确】**功能可以打✓，点击**【右键菜单】➔【设置标志】➔【错误】**功能可以打 X。勾选“按原始顺序导出”，则保持原文的顺序导出对齐文件。



按原文顺序导出的 TXT 对齐文件，打开后如下图：在原文和译文之间，用 Tab 制表符作了分隔。



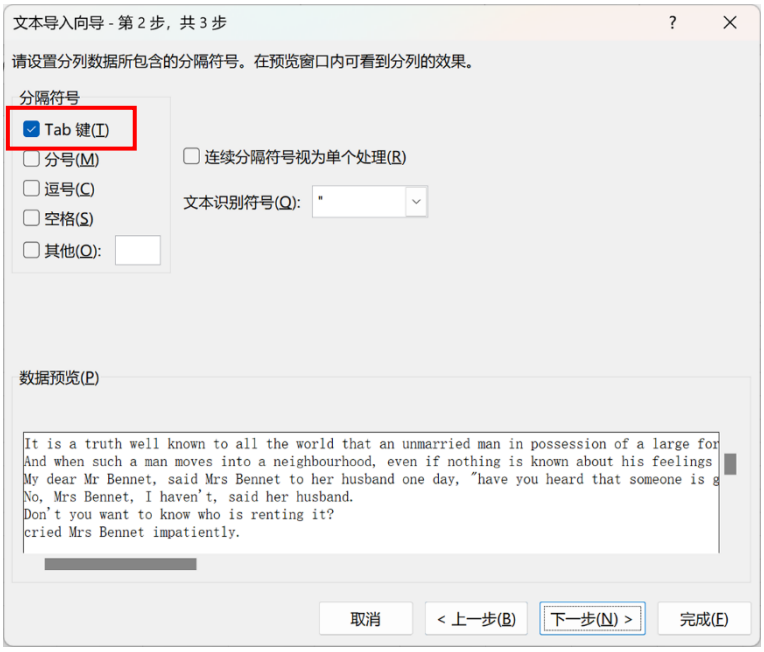
导出的这个 TXT 对齐文件，可以在 EXCEL 打开后，再另存为 EXCEL 文件。方法是：在 EXCEL 中点击“打开”文件按钮，在弹出的“打开文件”窗口的文件类型中，点击下拉按钮，选择“文本文件”类型，则可以看到 TXT 文件，如下图：



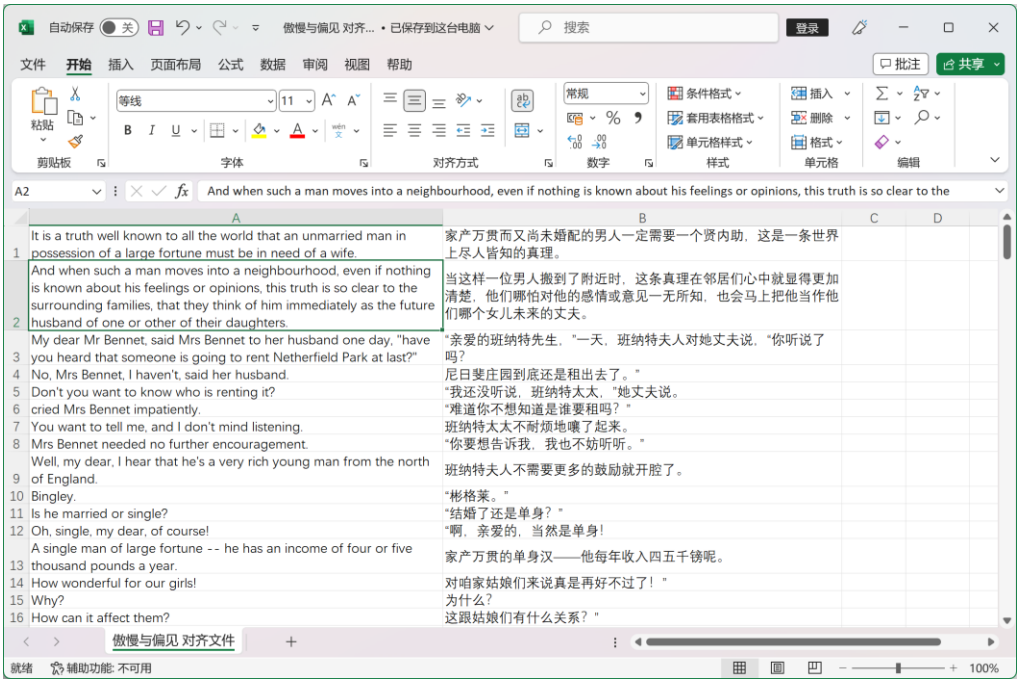
双击 TXT 对齐文件，会弹出导入向导，选择“分隔符号”选项。



点击“下一步”按钮，选择“Tab 键”，便可以打开这个 TXT 文件了。



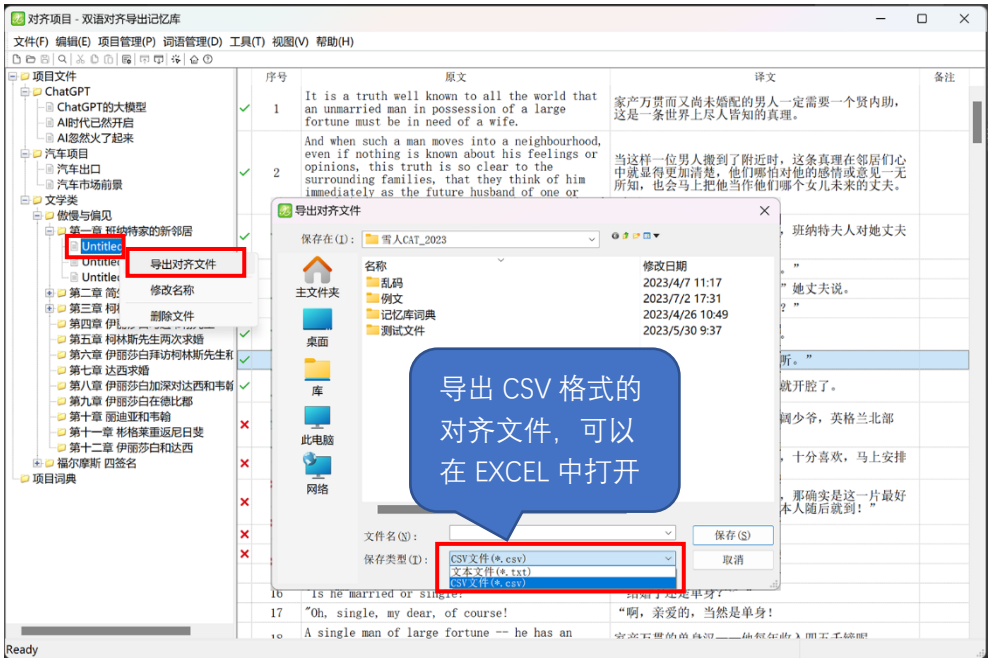
打开后，调整 A、B 列的宽度及设置自动换行，再另存为 EXCEL 文件即可。如下图：



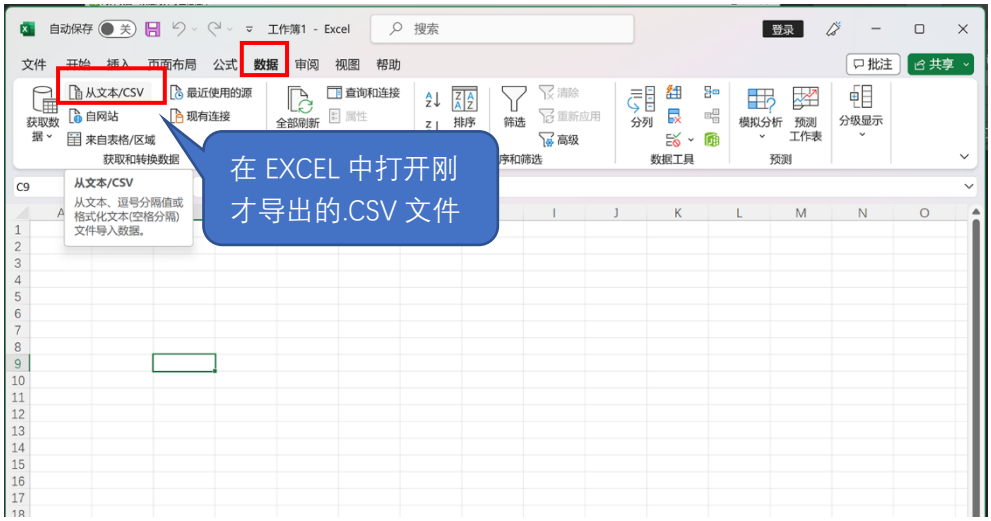
	A	B
1	It is a truth well known to all the world that an unmarried man in possession of a large fortune must be in need of a wife.	家产万贯而又尚未婚配的男人一定需要一个贤内助，这是一条世界上尽人皆知的真理。
2	And when such a man moves into a neighbourhood, even if nothing is known about his feelings or opinions, this truth is so clear to the surrounding families, that they think of him immediately as the future husband of one or other of their daughters.	当这样一位男人搬到了附近时，这条真理在邻居们心中就显得更加清楚，他们哪怕对他的感情或意见一无所知，也会马上把他当作他们哪个女儿未来的丈夫。
3	My dear Mr Bennet, said Mrs Bennet to her husband one day, "have you heard that someone is going to rent Netherfield Park at last?"	"亲爱的班纳特先生，"一天，班纳特夫人对她丈夫说，"你听说了吗？
4	No, Mrs Bennet, I haven't, said her husband.	尼日斐庄园到底还是租出去了。"
5	Don't you want to know who is renting it?	"我还没听说，班纳特太太，"她丈夫说。
6	cried Mrs Bennet impatiently.	"难道你不想知道是谁要租吗？"
7	You want to tell me, and I don't mind listening.	班纳特太太不耐烦地嚷了起来。
8	Mrs Bennet needed no further encouragement.	"你要想告诉我，我也不妨听听。"
9	Well, my dear, I hear that he's a very rich young man from the north of England.	班纳特夫人不需要更多的鼓励就开腔了。
10	Bingley.	"彬格莱。"
11	Is he married or single?	"结婚了还是单身？"
12	Oh, single, my dear, of course!	"啊，亲爱的，当然是单身！
13	A single man of large fortune -- he has an income of four or five thousand pounds a year.	家产万贯的单身汉——他每年收入四五千镑呢。
14	How wonderful for our girls!	对咱家姑娘们来说真是再好不过了！"
15	Why?	为什么？
16	How can it affect them?	这跟姑娘们有什么关系？"

导出 CSV 格式的对齐文件

对双语对齐文件，支持导出 CSV 格式，这个文件可以在 EXCEL 中打开并另外为 EXCEL 文件。具体方法：点击“最底层”某个具体文件，再点击**【右键菜单】➔【导出对齐文件】**功能，在弹出窗口的“保存类型”中，点击旁边的下拉框，选择 CSV 文件（*.csv），如下图：



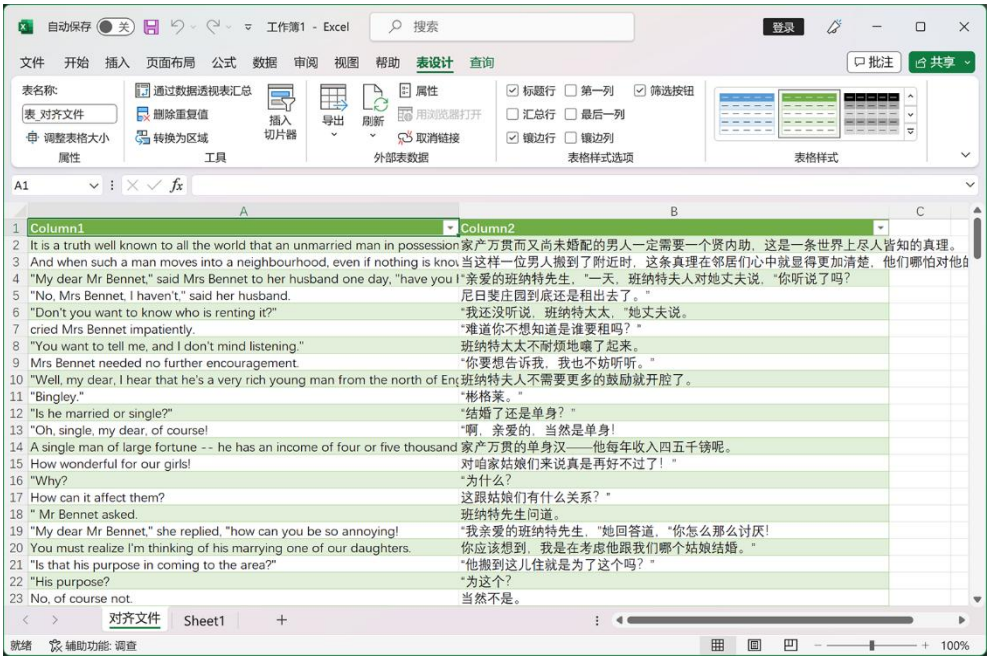
导出后，启动 EXCEL 软件，点击**【数据】➔【从文本/CSV】**，打开刚才导出的.CSV 文件。如下图：



点击“加载”按钮，如下图： 如果 Excel 自动识别的编码有错，会出现乱码，此时请选择下拉框中的 Unicode（UTF-8）编码格式。



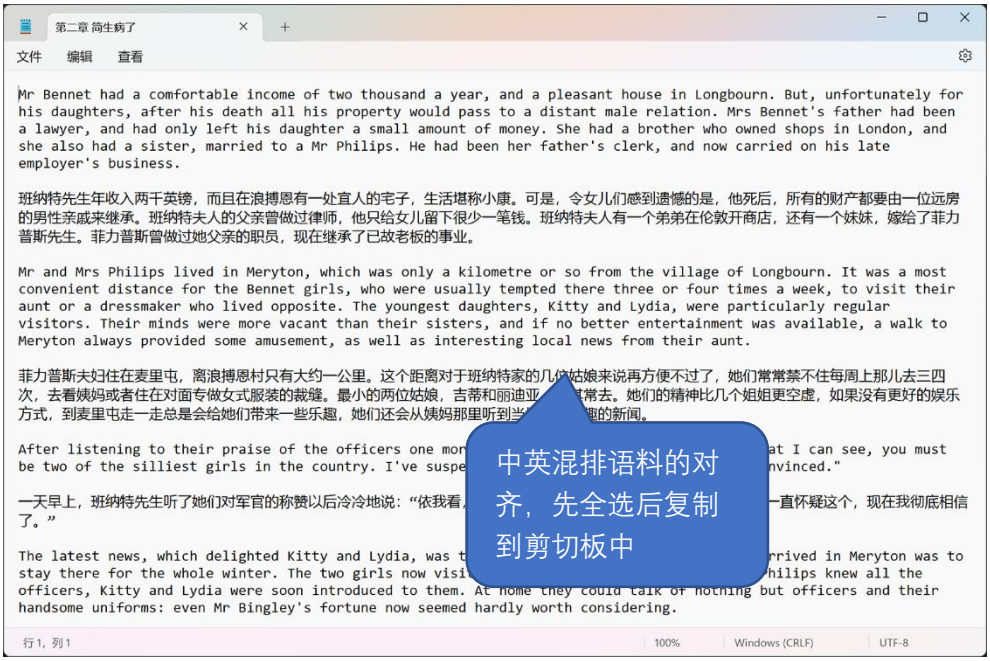
在 EXCEL 中打开了这个导出的 CSV 文件，再另外为 EXCEL 文件。如下图：



第五节 双语对齐二次拆分对齐

二次拆分对齐可以提升对齐的准确度，对于大型语料的对齐，可以先按章节、段落进行第一次的对齐，然后再按句子进行第二次对齐。下面举例说明：

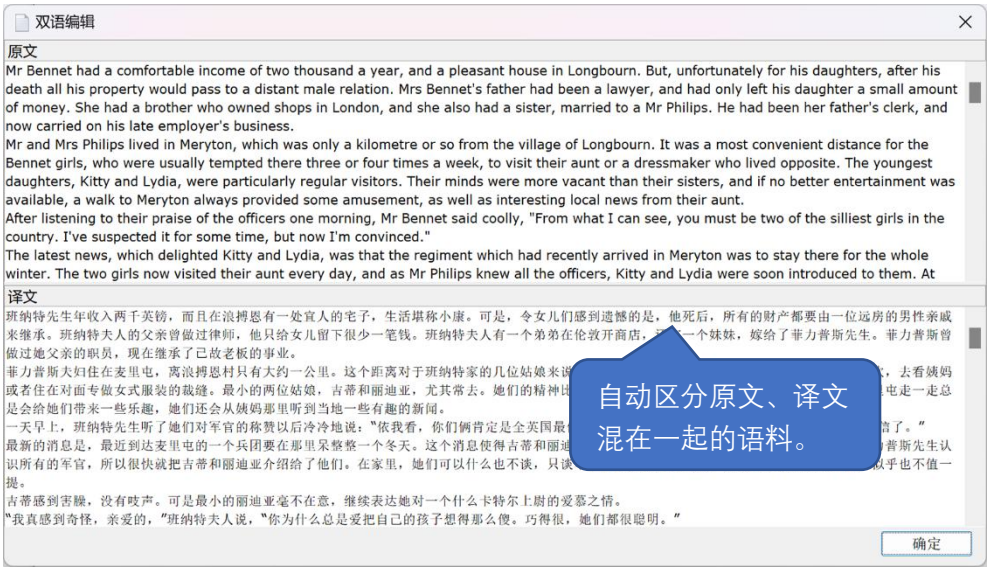
例如：下图中欲对齐的双语语料是中英混排的，先按【Ctrl+A】全选、【Ctrl+C】复制，将需要对齐的语料复制到剪切板中。



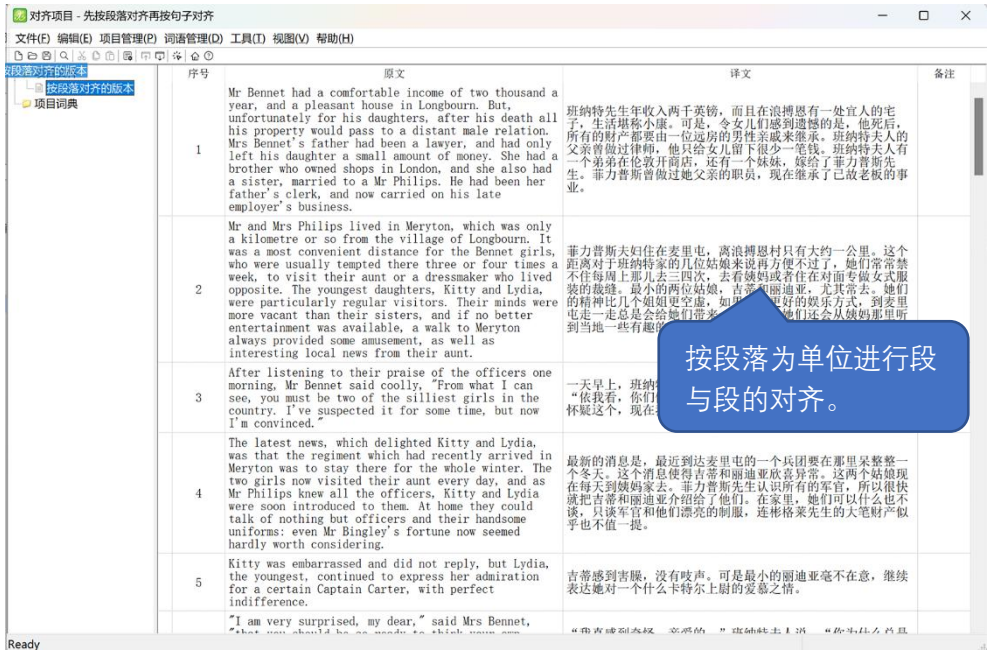
回到雪人软件中，点击**【文件】➔【新建】➔【对齐项目】**菜单，新建一个双语对齐项目，然后右键点击屏幕左侧的"项目文件"，点击**【右键菜单】➔【导入文件】**功能，弹出窗口如下图：



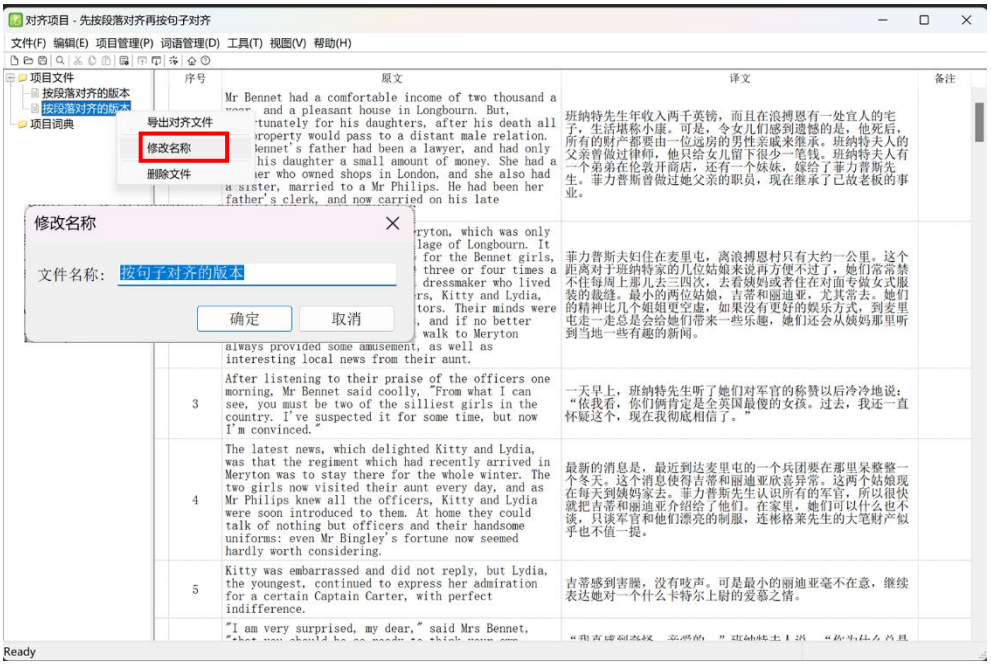
不勾选“拆分段落中的句子”这个选项，表示第一次对齐时将按段落进行对齐，不拆成一句一句进行对齐。然后点击“从剪切板导入”按钮，刚才原文、译文混在一起的语料，雪人已经自动分开了，如下图：



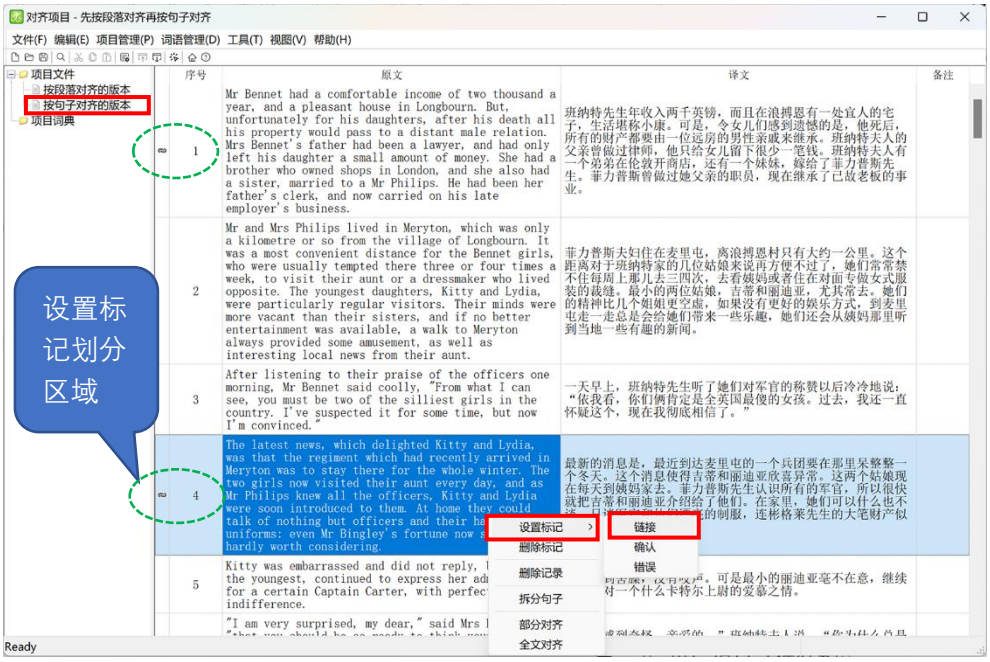
点击“确定”按钮，将按段落为单位进行段与段的对齐，段落对齐一般会比较准确，如下图：



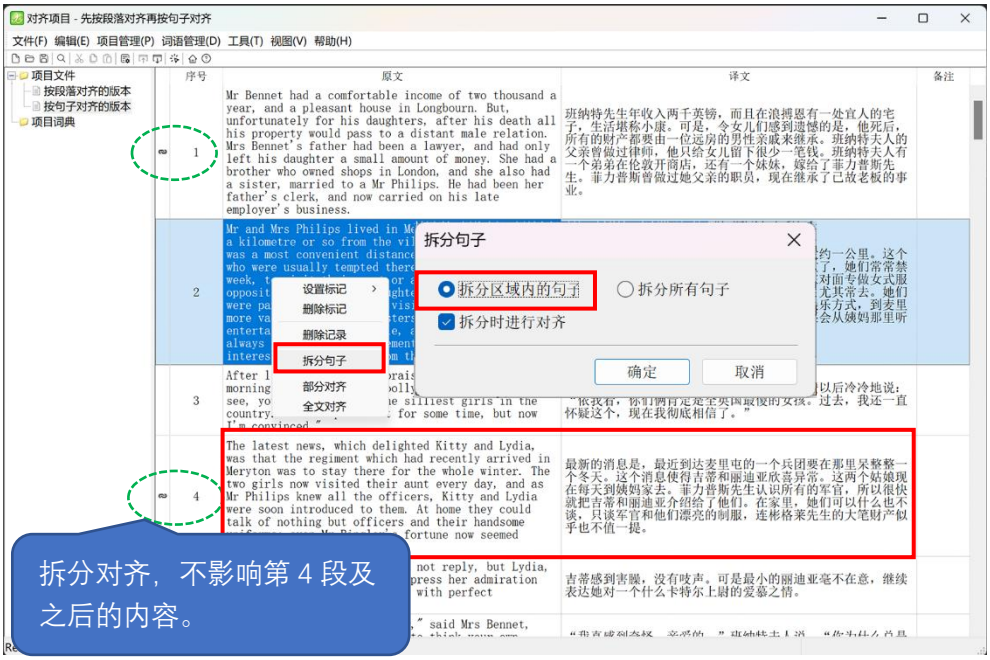
段落对齐后，接下来对段落里的句子拆分对齐。如果希望保留段落对齐的版本，可以复制一份来进行句子的拆分对齐。在雪人中，可以非常方便地进行文件的复制，方法是：在屏幕左侧的项目文件夹下面，点击某个文件后，按住鼠标左键单击拖动，并同时按住**【Ctrl】键**，此时鼠标会变成一个“+”号，则可复制一份。复制后可以点击**【右键菜单】➔【修改名称】**功能将文件名修改为“按句子断句的版本”，如下图：



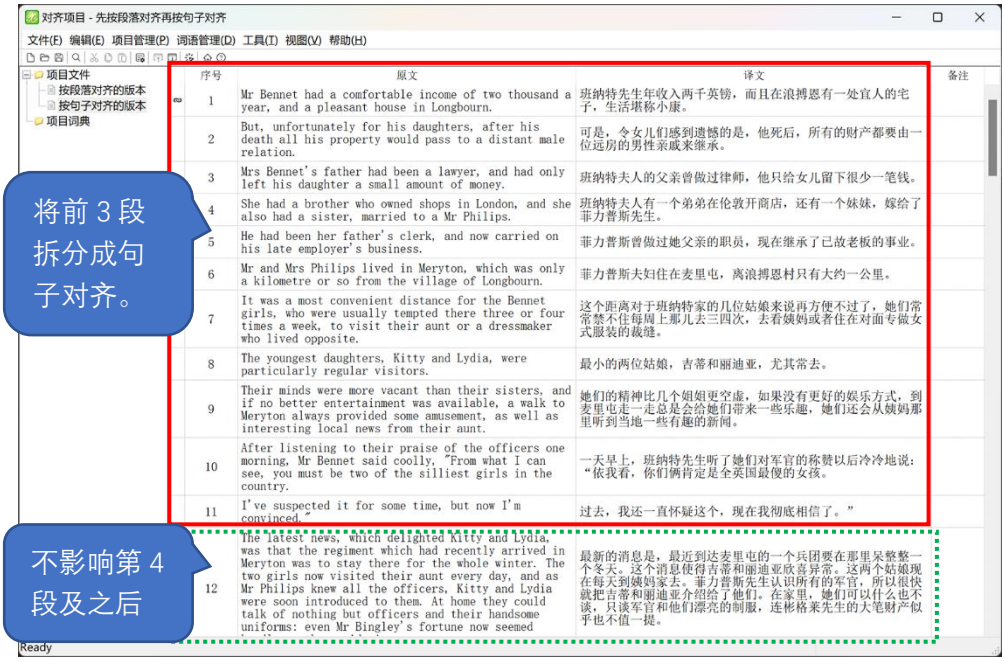
点击屏幕左侧复制出来的“按句子断句的版本”文件，在这个文件中，设置标记进行区域划分，例如，下图中，利用**【右键菜单】➔【设置标记】➔【链接】**功能，在第 1 段和第 4 段设置了两个标记。



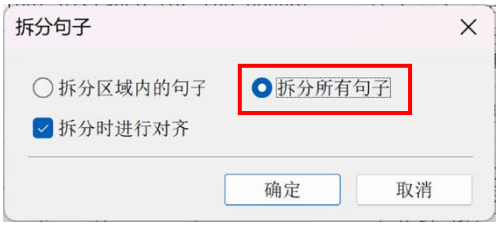
然后，点击【**右键菜单**】➔【**拆分句子**】功能，在弹出的“拆分句子”窗口中，选择“拆分区域内的句子”选项，则表示仅在区域内进行拆分对齐，不影响区域外的段落。例如下图中，由于在第 1 段和第 4 段设置了两个标记，那么这个区域范围就是第 1 段、第 2 段、第 3 段，即针对这 3 段的内容进行拆分对齐，不包括第 4 段的内容。



点击“确定”按钮后，则第 1 段、第 2 段、第 3 段这个区域范围内的句子进行了拆分对齐，而第 4 段及之后的段落则没有拆分。如下图：



用户可以根据实际需要灵活选用“拆分区域内的句子”这个选项。当然，也可以不进行区域划分，直接针对所有的段落进行拆分对齐，选择下图的“拆分所有句子”选项，则在各自段落对齐的前提下，将各个段落中的句子再拆分出来，以句子为单位在本段落中进行对齐。



第十一章 查找和替换

第一节 词语的整批查找和替换

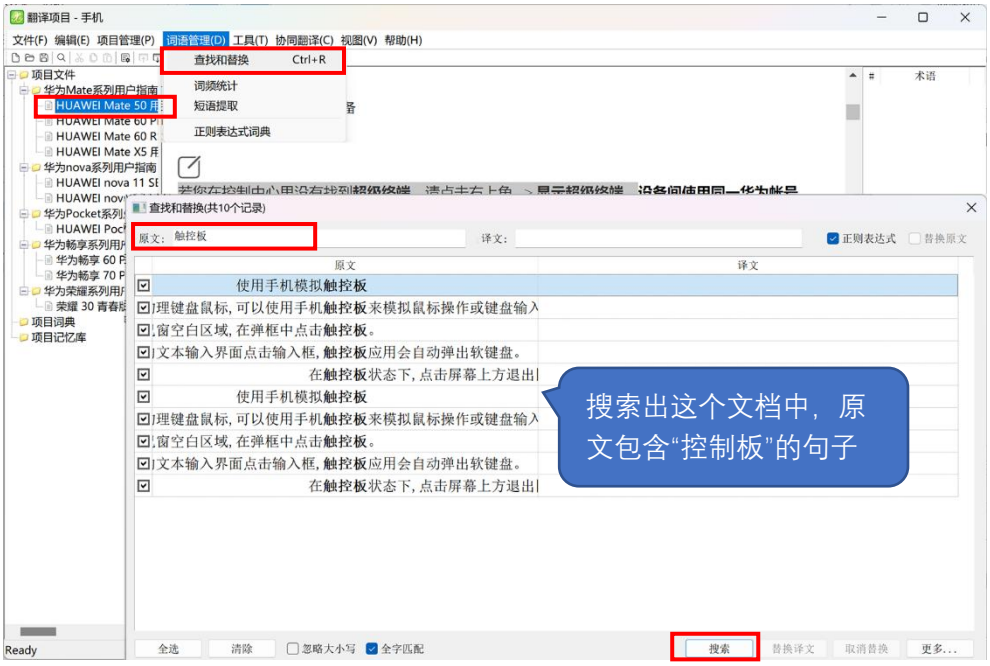
新版进一步加强了“查找和替换”的功能，可以很方便地对译文、原文、项目词典、项目记忆库、双语对齐项目进行整批的修改替换。查找时支持用“**正则表达式**”，替换时也支持用“**正则表达式**”进行复杂的替换。查找替换方法灵活，目标结果准确。对翻译的项目文件和双语对齐的项目文件，支持单一文件内的替换或多文件夹下的文件统一替换。双击查找替换窗口的结果，可以直接跳转定位到该句上。

（一）、查找替换在翻译项目文件中的应用

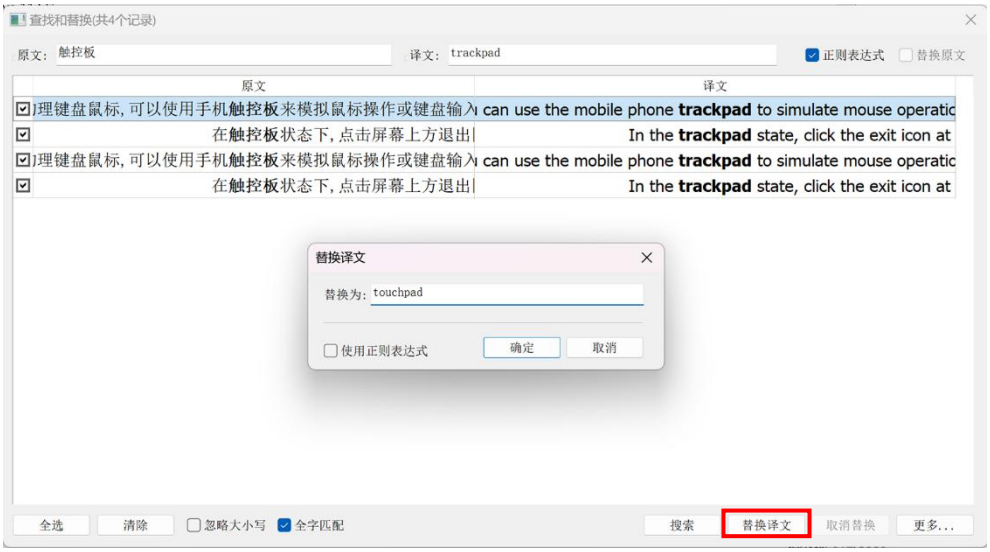
例如：下图的“手机用户指南”翻译项目文件是一个复杂结构的项目文件，其中包含：“华为 Mate 系列”、“华为 nova 系列”、“华为 Pocket 系列”、“华为畅享系列”、“华为荣耀系列”五个子文件夹，每个子文件夹下面有多份翻译文档。翻译完成后，发现“触控板”这个词有些翻译成“trackpad”，有些翻译成“touchpad”，现在希望对它进行统一的修改。

(1)、仅在单个文档内替换

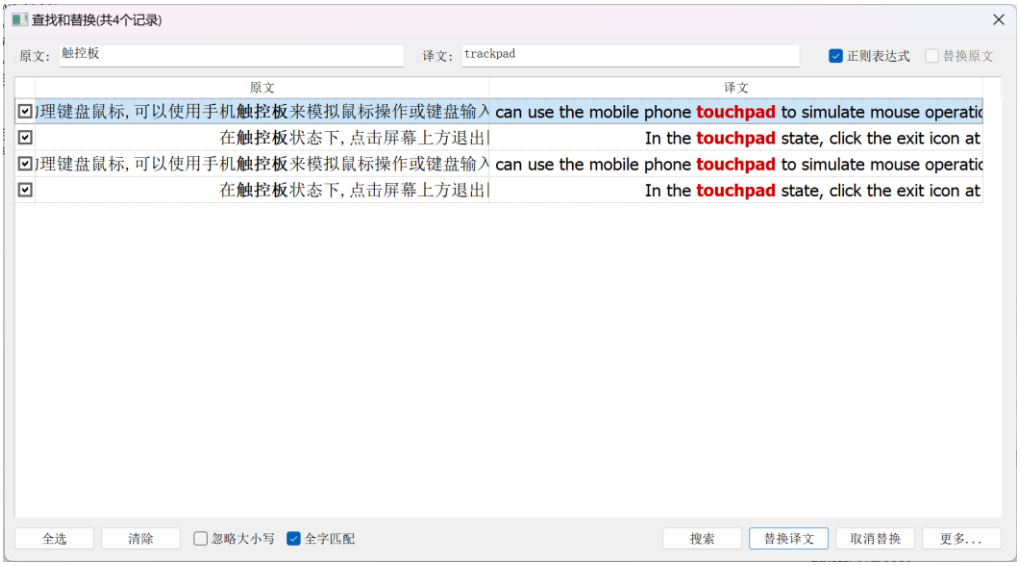
如果仅希望在单个文档内进行替换，鼠标先点击屏幕左侧“项目文件”下面的某个具体的翻译文档，例如点击“HUAWEI Mate 50 用户指南.docx”这个文档，然后点击**【词语管理】➔【查找和替换】**菜单或按 Ctrl+R】快捷键，会弹出“查找和替换”窗口，在原文搜索框中输入“触控板”，点击搜索按钮，会搜索出这个文档中原文包含“触控板”的 10 个句子，如下图：



在译文框中输入“trackpad”，点击搜索按钮，可在这个文档中搜索出原文包含“触控板”，且译文包含“trackpad”的 4 条记录，如下图：

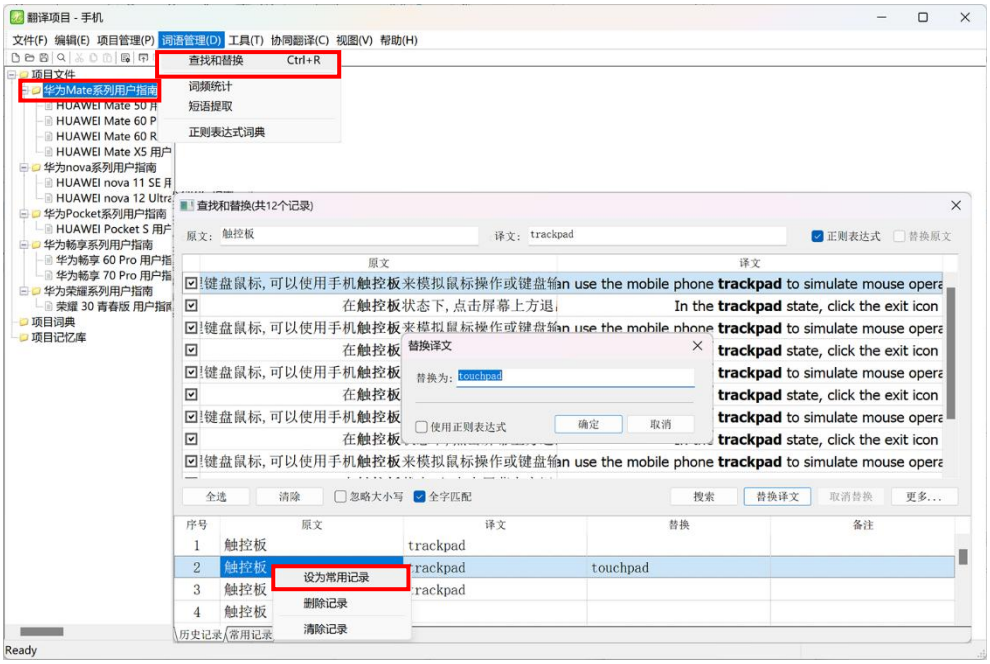


点击替换译文按钮，将这 4 条记录替换为“touchpad”，点击确定按钮后，进行替换，如下图：

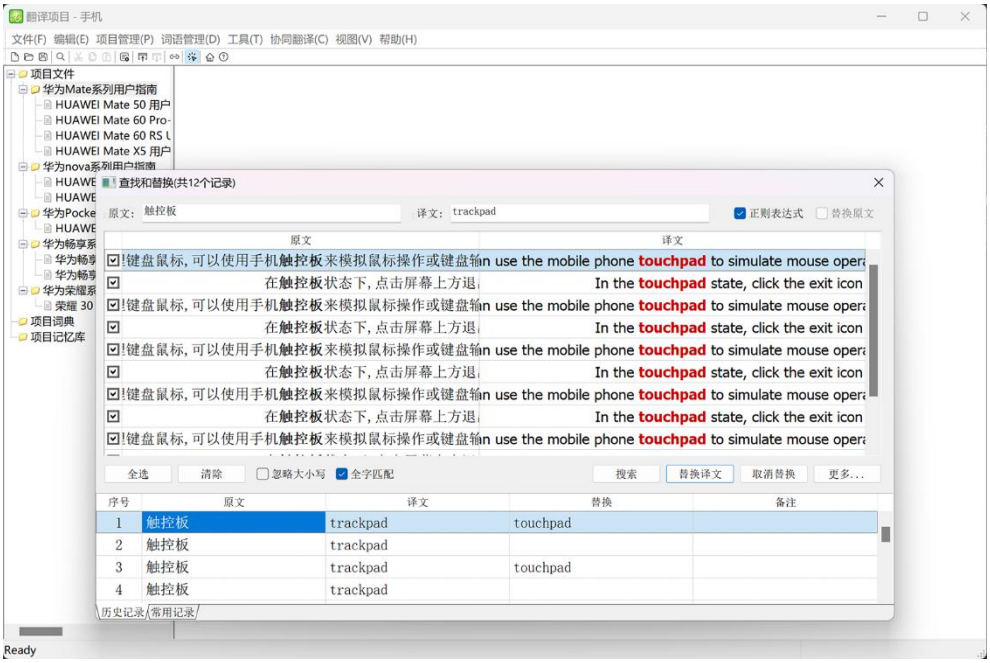


(2)、在某个子文件夹内替换

鼠标先点击屏幕左侧项目文件下面的某个子文件夹，例如点击“华为 Mate 系列用户指南”这个子文件夹，然后点击【词语管理】➔【查找和替换】菜单，会对这个子文件夹下面的 4 篇文档进行查找替换操作，例如下图中，在这个子文件夹下面的 4 篇文档中搜索出原文包含“触控板”，且译文包含“trackpad”的 12 条记录：

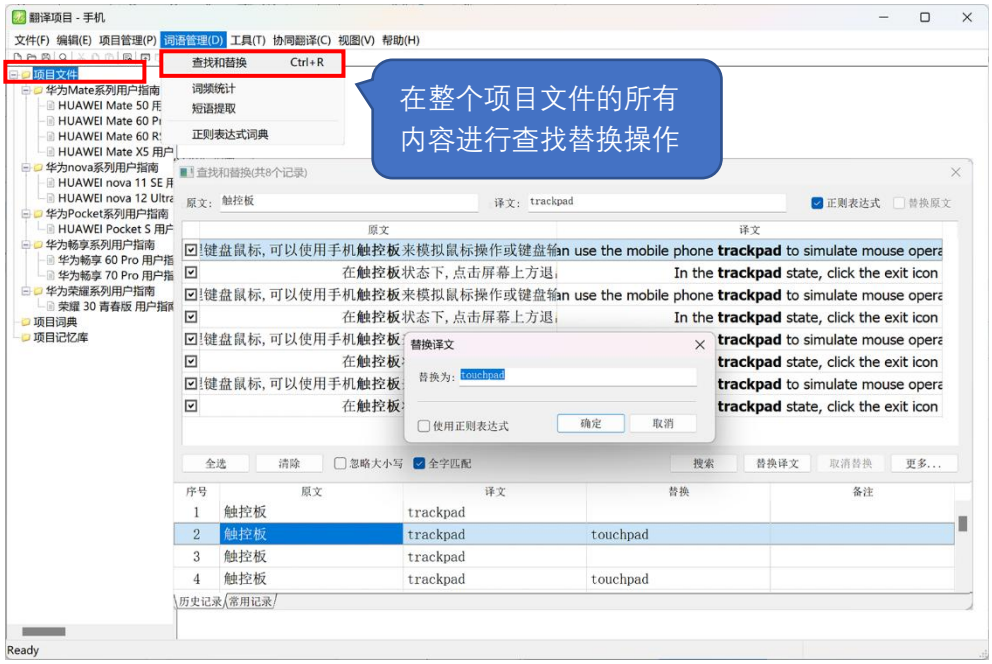


点击替换译文按钮，将这 12 条记录替换为“touchpad”，如下图：对于一些可能经常使用的查找替换条件，可以点击更多按钮，将它设置为常用记录，下次双击就可以调用这些预设好的查找替换条件。

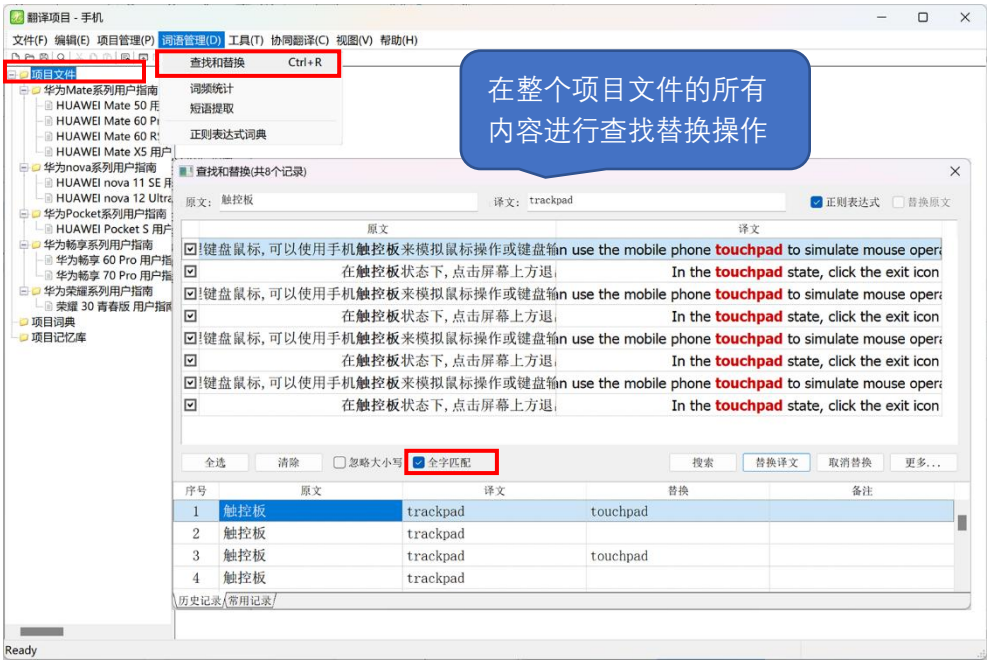


(3)、在整个项目文件内替换

鼠标先点击屏幕左侧最上面的项目文件，再点击**【词语管理】➔【查找和替换】**菜单，会对整个项目文件下面的所有文件进行查找替换操作，即：对该项目文件下的“华为 Mate 系列”、“华为 nova 系列”、“华为 Pocket 系列”、“华为畅享系列”、“华为荣耀系列”五个子文件夹下的所有翻译文档执行查找替换操作。如下图:



例如：下图中，将整个项目文件下的“触控板”的译文替换成“touchpad”。

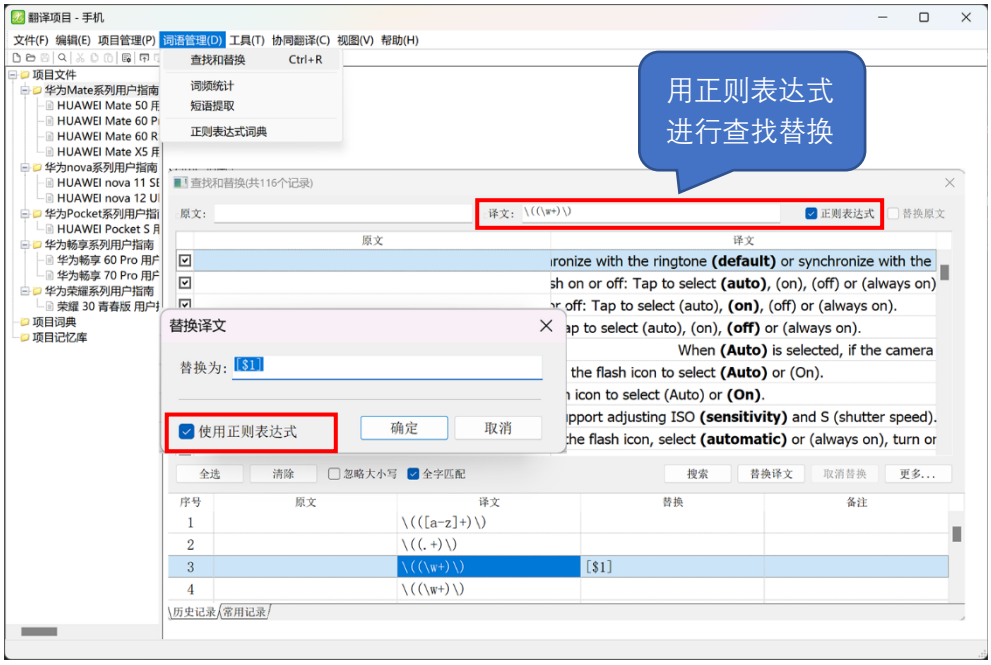


全字匹配

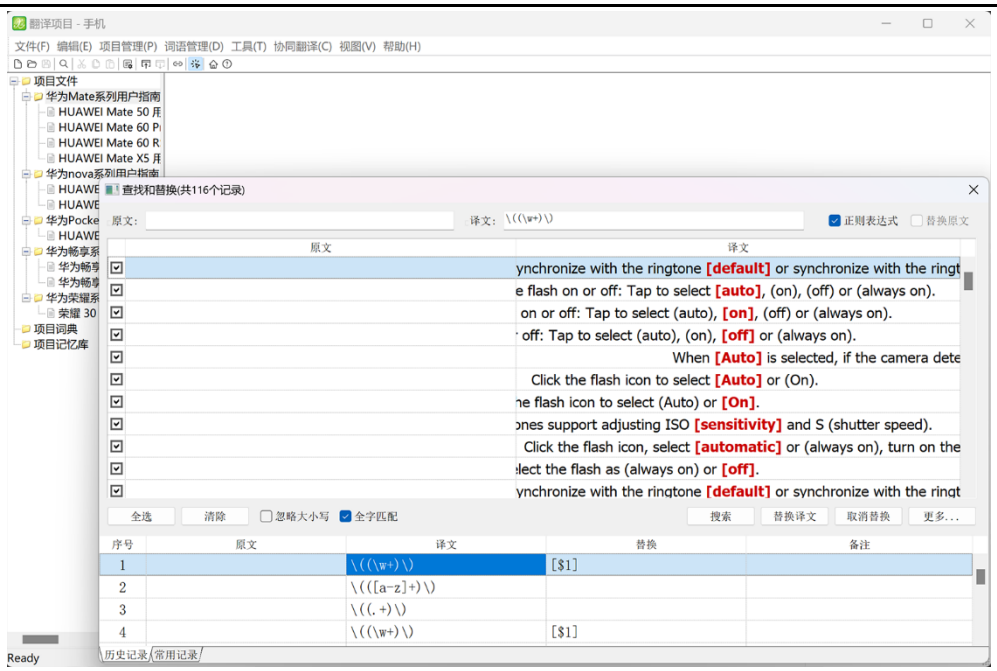
例如搜索“on”，只要包含有“on”的单词也会显示出来，如 song、online。勾选“全字匹配”后重新搜索，仅出现是以“on”为单词的句子。

(4)、用正则表达式进行查找和替换

查找条件可以输入正则表达式，替换时也可以输入正则表达式。如下图中：在译文搜索框中输入：\((\w+)\)，表示查找（）圆括号内只有一个单词的词语



替换时输入：［$1］ 表示将这些词语的圆括号替换为[]方括号，并勾选“使用正则表达式”这个选项，替换的结果如下图：



(I)、正则表达式理论知识

正则表达式就是用于描述查找符合某些复杂规则的字符串的代码。类似于 Windows/Dos 下用于文件查找的通配符，如*.docx 可以查找某个目录下的所有的 Word 文档，只不过正则表达式更能精确实现你的查找目标，同时也会更加复杂些，这里只简单介绍几个常用的元字符，若需更加深入地学习正则表达式的使用，请参考“正则表达式”的相关书籍。

\b——是正则表达式的一个特殊代码，代表单词的开头或结尾。有了它，当你搜索“is”这个单词时，this、his、miss 等就不会出现，写法：\b 单词\b，如：\bis\b，表示仅仅搜索 is 这个单词，其他含有 is 这两个字母的单词不要出现。

.——是一个很有用的元字符，它可以匹配除了换行符以外的任意字符。

*——同样也是一个很常用的元字符，不过它代表的不是字符，也不是位置，而是数量；它指定*前边的内容可以连续重复使用任意次以使整个表达式得到匹配。

.*——.*连在一起就意味着任意数量的不包含换行的字符。例如：\bis\b.*\bhappy\b 的意思是:搜一个单词 is，然后是任意个任意字符(但不能是换行)，最后是 happy 这个单词。

^——匹配字符串的开始。(^是和数字 6 在同一个键位上的符号)

$——匹配字符串的结束。

\d——匹配数字。
\s——匹配任意的空白符。
\w——匹配字母或数字或下划线或汉字

反义：

有时需要查找不属于某个能简单定义的字符类的字符。比如想查找除了数字以外，其它任意字符都行的情况，这时需要用到反义:

\W ——匹配任意不是字母，数字，下划线，汉字的字符
\S ——匹配任意不是空白符的字符
\D ——匹配任意非数字的字符
\B ——匹配不是单词开头或结束的位置
[^x]——匹配除了 x 以外的任意字符
[^nes]——匹配除了 nes 这几个字母以外的任意字符

(II)、正在表达式实际应用

①将句尾的“.”号全部替换成“。”号

在英译中时，经常出现用英文的“.”号来做句子的句号。如何将它们全部找出一次替换成“。”号呢?

\.$　　表示查找以“.”号结尾的句子
。$　表示查找以“。”号结尾的句子

②搜索多个词

输入 go|went　　表示搜索含 go 或 went 的单词，例如:
go|went|goes|gone|going，可以搜索出 go 的各种变形词。

③精确查找单词

例如想查找“go”这个单词，在原文搜索框中输入“go”时，查找到的结果除了有“go”这个单词外，还出现了许多其他含 go 这两个字母的单词，例如 good、got 等等，这些当然不是我们想要的结果。

但若输入的条件写成：“\bgo\b”表示只查找 go 这个单词，勾选了屏幕下方的“忽略大小写”选项，则搜索的结果就只是出现 go 这个单词了，还包括 Go、GO 大小

\b\w+ing\b　　表示搜索以 ing 结尾的词。
\bw+(ing|ed)\b　　表示搜索以 ing 或 ed 结尾的单词。

④对间隔短语的搜索

利用“正在表达式”的强大搜索功能，可以搜索出中间带省略号的复杂短语，例如希望搜出所有含 Not only…, but…的句子，只需要在搜索框中输入：\bnot only\b.+\bbut\b 即可实现这一需求。\bnot only\b.+\bbut \b 搜索固定空格数的 Not only…, but…短语。若 not 和 only 之间是不固定空格数，如 not … only…, but…also…则应该写成：\bnot +only\b.+\bbut +also\b

⑤中文搜索

“正则表达式”对中文的搜索也是非常方便的，例如输入：赚|亏 可以搜索包含“赚”字或“亏”字的句子。

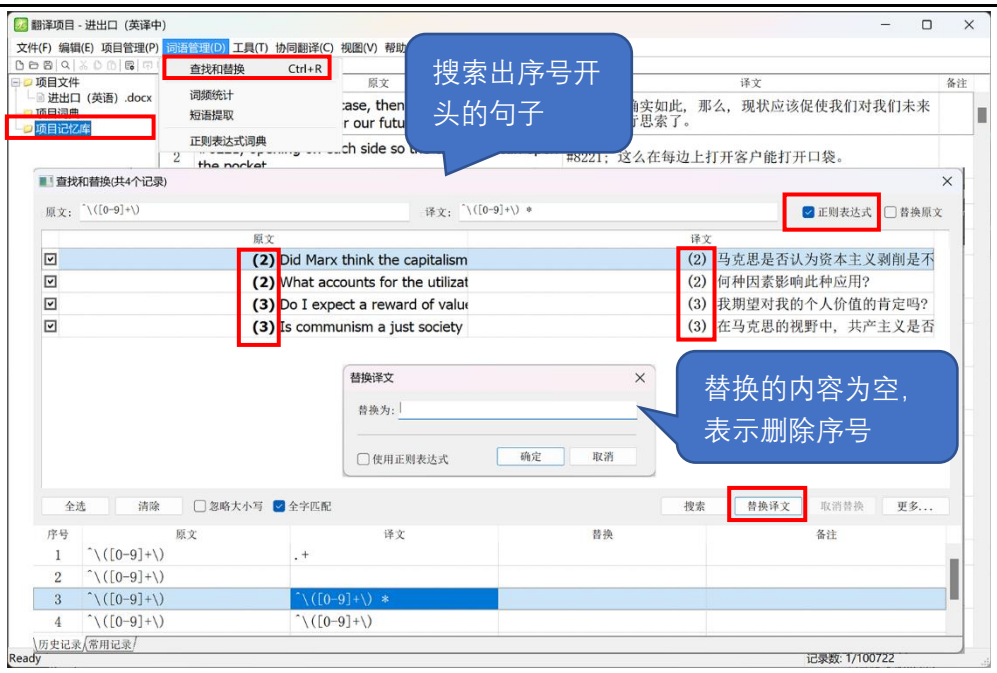
若输入：赚.*万 可以搜索含“赚……万”的句子，如赚进数百万、赚一百万、赚进十一二万等等都会被搜索出来。

若输入：足球.*之后.*会议 表示搜索句子中含有：“足球…….之后……会议”字样的句子。例如：在观看足球比赛之后，他还有一个重要的会议要参加 ，这样的句子就会被搜索出来。

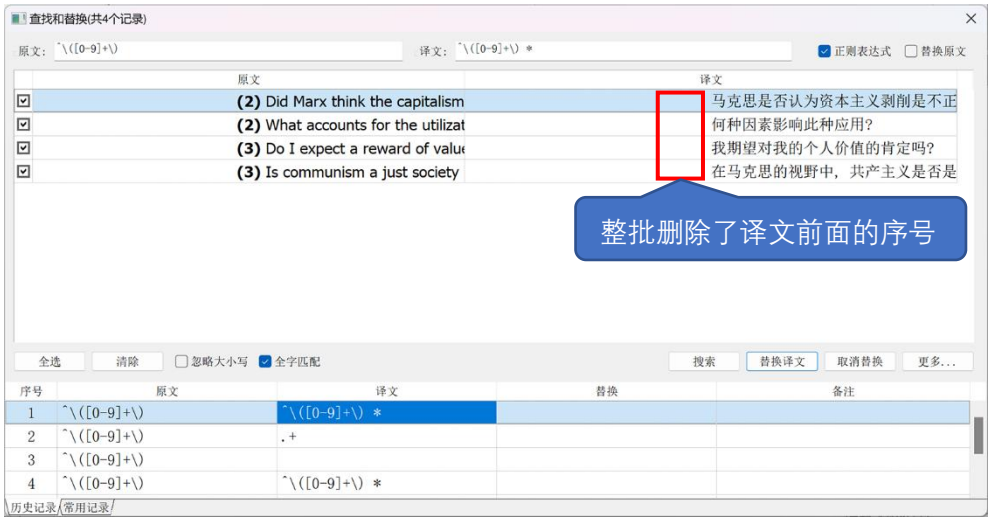
正则表达式的使用有标准的语法，更多关于正则表达式的高级运用，请用户自行查阅相关的正则表达式的书籍。

（二）、查找替换在项目记忆库中的应用

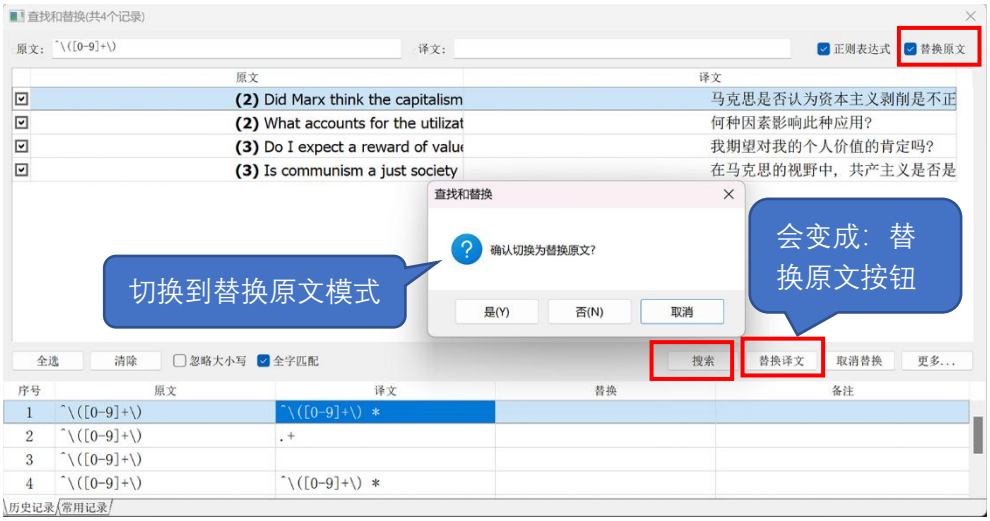
鼠标先点击屏幕左侧的 项目记忆库，然后点击**【词语管理】➔【查找和替换】**菜单，可以对记忆库的内容进行整理、维护。例如：想搜索出来记忆库中那些序号开头的句子，将序号整批删除掉。利用正则表达式可以实现，在下图的原文搜索框中输入 ^\([0-9]+\) 译文搜索框中输入 ^\([0-9]+\) * 勾选右侧的“正则表达式”选项，点击搜索按钮，可以搜索出记忆库中序号开头的句子，如下图：



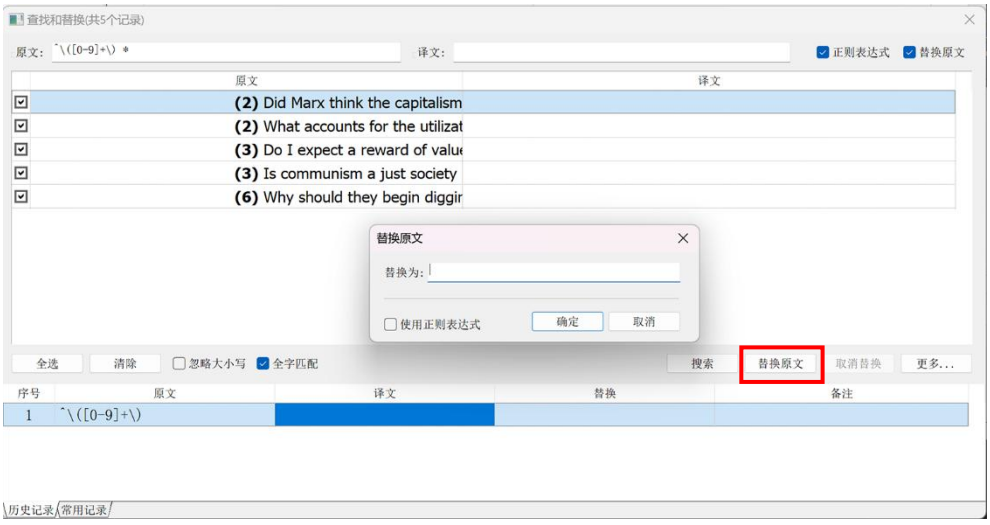
点击下方的替换译文按钮，替换的内容为空，可整批删除序号，如下图：



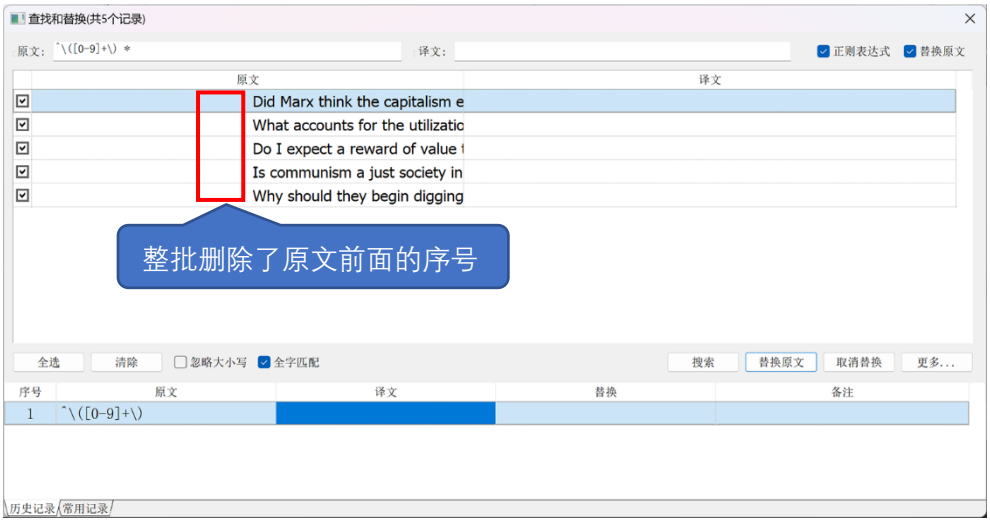
接着处理原文的序号，删除刚才译文搜索框中输入的条件，然后勾选右侧的替换原文选项，切换到替换原文的模式，下方的按钮会变成替换原文按钮。



点击搜索按钮，搜索出原文开头带序号的句子，如下图:



然后点击替换原文按钮，替换内容为空，原文序号开头的句子就被批量删除了。如下图:



（三）、查找替换在项目词典中的应用

鼠标先点击屏幕左侧的项目词典，然后点击**【词语管理】➔【查找和替换】**菜单，可以对项目词典中的内容进行整理、维护。具体方法相类似，不累述。

（四）、查找替换在双语对齐项目中的应用

新建双语对齐项目，导入或黏贴双语语料进行对齐。双语对齐项目和翻译项目一样，也支持复杂的多文件夹的目录结构，鼠标先点击屏幕左侧的项目文件下面的哪个子文件夹或哪个具体的文档，然后点击**【词语管理】➔【查找和替换】**菜单，可以对双语对齐中的内容进行整理、维护。具体方法相类似，不累述。

（五）、错误替换的撤销

如果发现替换错误，可以点击取消替换进行撤销此次操作，如果没有进行下一个内容的搜索操作，可以多次取消替换。如果已经进行了其他内容的搜索操作或者退出了查找替换窗口，则之前的替换已经生效，不可以撤销，这时只能通过放弃存盘操作来恢复到替换前的状态，方法：点击**【文件】➔【关闭】**菜单，关闭文件时会询问您是否保存文件？点击否按钮，放弃存盘。建议在进行批量查找替换前，手动存一次盘，这样不管进行了多少次的错误替换，都可以通过放弃存盘操作来恢复到替换前的状态。

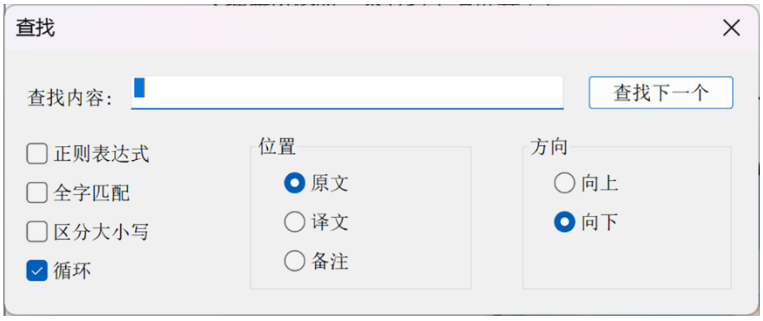


（六）、自动记录查找替换条件

软件会自动记录所输入的“查找条件”，避免了一些常用查找条件的重复输入。点击“查找与替换”窗口右下角的“更多”按钮，可展开窗口，其中有“历史记录”和“常用记录”两个页签，“历史记录”页签用于自动记录用户曾经输入过的查找条件，它仅对当前的翻译项目有用，“常用记录”中的查找条件则可适应于所有的翻译项目。所以，若是一些经常使用的查找条件，可以点击**【右键菜单】➔【设为常用记录】**，以后只要点击“替换译文”按钮就可以瞬间完成，省去了再次输入查找条件、输入替换内容等的操作。例如：经常需要查找译文中不正确的标点符号，将“点号”替换成“句号”。 “历史记录”和“常用记录”最多支持记录 100 个条件记录，超过 100 个后会覆盖之前的内容。

第二节 灵活准确的查找

点击【编辑】➔【查找】菜单、或点击工具栏查找按钮按、或按【Ctrl】+【F】快捷键，可以调出查找窗口，如下图:



在翻译的项目文件、项目记忆库、项目词典、双语对齐项目中，都可以调出这个查找窗口。查找功能很强大，可以对译文、原文、备注分别进行查找，还支持“正则表达式”，使查找的条件更加灵活，目标结果准确，想找什么都能找到。查找时，按【F3】可以查找下一目标。

第十二章 协同翻译

“雪人 CAT 协同服务器”，是全新研发的新版本协同翻译服务器软件。**服务器架设在用户自己的机器上**，可以在局域网或互联网架设服务器，也可去服务器云提供商租用虚拟服务器主机（例如：去阿里云、华为云等租用云主机）；使翻译企业、翻译团队**拥有私有化自主的协同翻译平台**，切实保障翻译资料的安全、保密性。

“雪人 CAT 协同服务器”简化了服务器的架设和对服务器系统的要求。无论是拥有专门的技术人员、高性能的服务器的大型企业，还是小型的翻译公司和小型翻译团队，甚至是个人译者，都可以轻松地架设“雪人 CAT 协同服务器”。“雪人 CAT 协同服务器”对操作系统的要求：Windows 7 及以上的 64 位操作系统都可以。

“雪人 CAT 协同服务器”软件可以登录雪人官网 http://www.gcys.cn 下载和申请使用。即使是“个人译者”，也可以去下载“雪人 CAT 协同服务器”，有了“协同服务器”可以实现译者之间的资源共享，或者把“协同服务器”软件安装在本地机器，将翻译的内容自动保存在到本地的服务器上。

如果觉得自己架设服务器比较麻烦，或觉得自己租用“虚拟云主机”成本较高，但又希望和他人在互联网上进行协同翻译、数据共享。可以申请加入雪人公司提供的协同服务器。

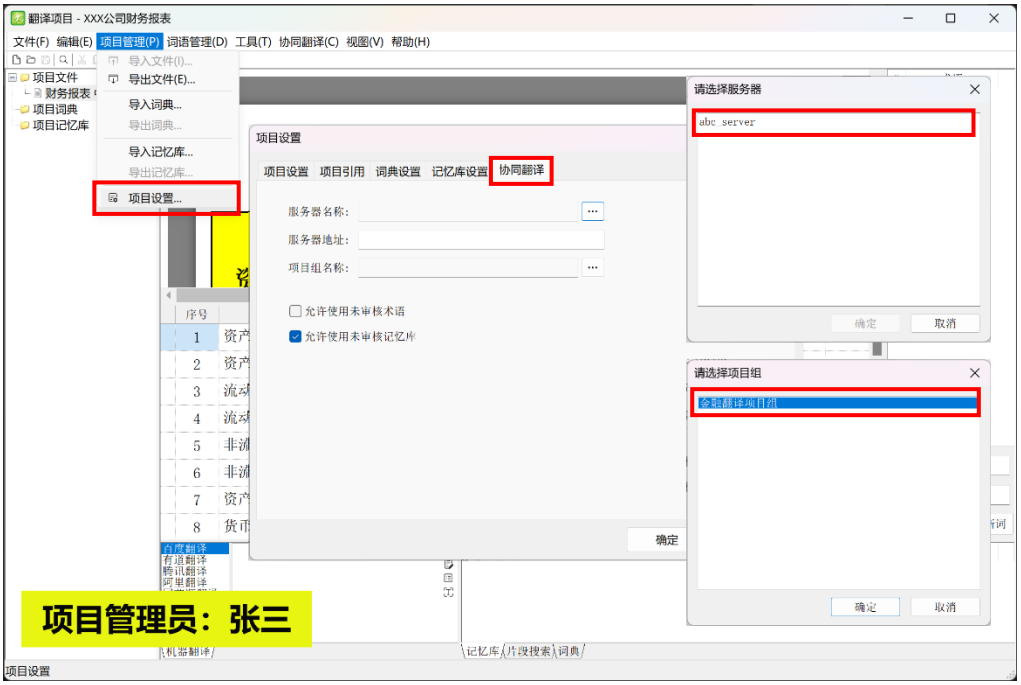
更多关于“协同服务器”的架设、授权允许哪些成员访问这个服务器、在服务器中外挂资源供团队使用等相关问题，请参看“雪人 CAT 协同服务器”的操作手册。

接“雪人 CAT 协同服务器”操作手册中的介绍：“李明”申请了一个名称为：“abc_server”的协同服务器”。“系统管理员”为 test/李明（test 是他注册雪人 CAT 客户端账号时填写的账号名），还授权了 test2 张三、test3 李四、test4 王五、test5 赵六 这几个人可以访问“abc_server”协同翻译服务器。除此之外的其他任何人，是无权访问这个“abc_server”协同翻译服务器的。“系统管理员 test/李明”还在“abc_server”协同翻译服务器中外挂了多个记忆库和词典，创建了“金融翻译项目组”，并指定“test2 张三”为“金融翻译项目组”的“项目管理员”。

接下来我们将介绍“金融翻译项目组”的“项目管理员”——“test2 张三”如何在雪人 CAT 客户端中接入“abc_server”协同翻译服务器？如何分配其他成员角色？及给“金融翻译项目组”添加上可以使用的记忆库、术语库。译员、审校如何参与团队工作？

第一节 客户端如何连接服务器

“项目管理员”——“test2 张三”启动雪人 CAT 客户端软件后，根据实际情况选择原文语种和译文语种并新建翻译项目。然后点击【**项目管理**】➔【**项目设置**】菜单，在“项目设置”窗口中，点击“协同翻译”，如下图：



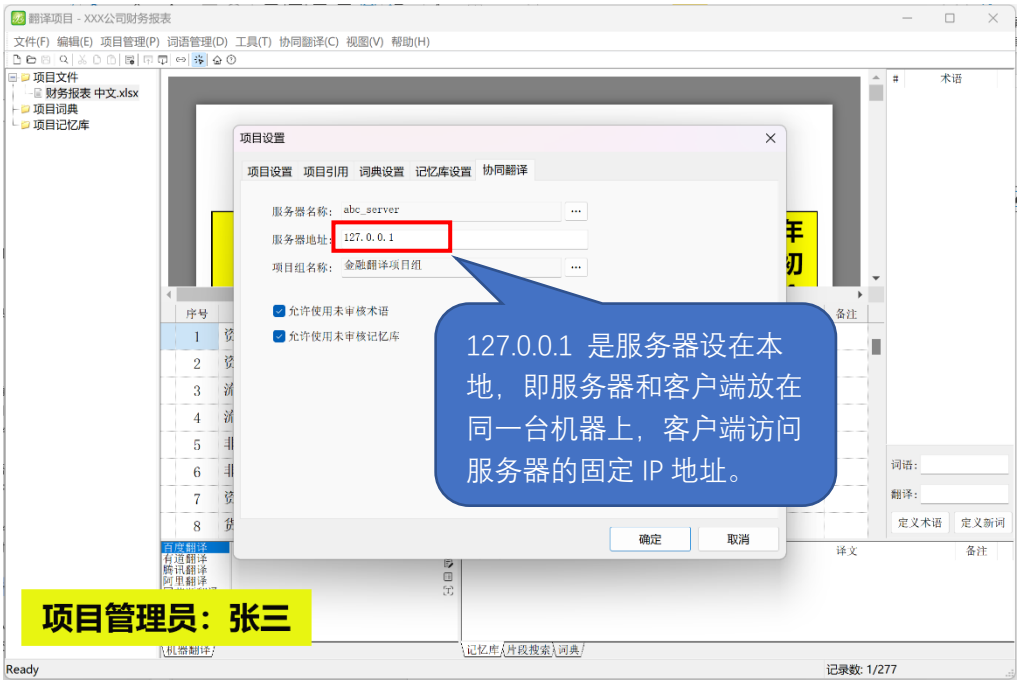
一、选择被授权加入的服务器名称

点击“服务器名称”右侧的…按钮，会弹出允许当前用户加入的具体服务器名称，双击某服务器名称即可选择加入该服务器。（例：目前仅有“abc_server”服务器的“系统管理员 test 李明”授权允许“test2 张三”加入，所以张三只能在该窗口中看见“abc_server”这一个服务器。假如有多个服务器的“系统管理员”授权允许“test2 张三”加入，张三在这个窗口就可以看见多个服务器的名称供他选择，如果没有任何服务器的“系统管理员”授权允许“test2 张三”加入，则张三看到的这个窗口是空白的）。利用“雪人 CAT 协同服务器”，在保障数据安全的前提下，可以实现企业与企业之间的数据共享，或公司与团队之间的数据共享，或个人译者与个人译者之间的数据共享。

二、服务器地址及端口设置

如果这个“abc_server”协同翻译服务器是放在用户自己租用的“虚拟云主机上”（例如：“系统管理员”李明将“abc_server”服务器放在他租用的“虚拟云主机上”），给外网（互联网）的团队成员提供服务的，若“系统管理员”李明在“服务器设置”中选择“公开服务器地址”。则经“系统管理员”授权允许访问“abc_server”协同翻译服务器的成员，在客户端中连接“abc_server”协同翻译服务器时，不需要再输入服务器的地址，避免因地址输入错误而产生连不上的问题；而且若“abc_server”协同翻译服务器的地址发生变更，也不需要通知团队成员更新。如果选择“不公开服务器地址”，则“系统管理员”李明需要将服务器的地址告诉团队成员，由团队成员在客户端的“服务器地址”中手工输入所连接的服务器地址，若服务器的地址发生变更，也需要手动修改。服务器的端口默认为：9001，用户可以修改端口，如果使用默认的 9001 端口，则团队成员在客户端中可以不输入端口（即：在服务器的地址栏中仅需输入地址：例如输入：192.168.1.6 即可）。如果“系统管理员”修改了服务器的默认端口（例如端口修改为：8020，则团队成员在客户端的“服务器地址”栏中，需要输入地址及端口，输入格式如：192.168.1.6:8020 **注意：冒号必须在英文状态下输入**）

如果这个“abc_server”协同翻译服务器软件是放在本地机器上运行的（即服务器和客户端都在同一台机器，则输入本地机器访问服务器的**固定地址 IP 地址：127.0.0.1** 即可，如下图：



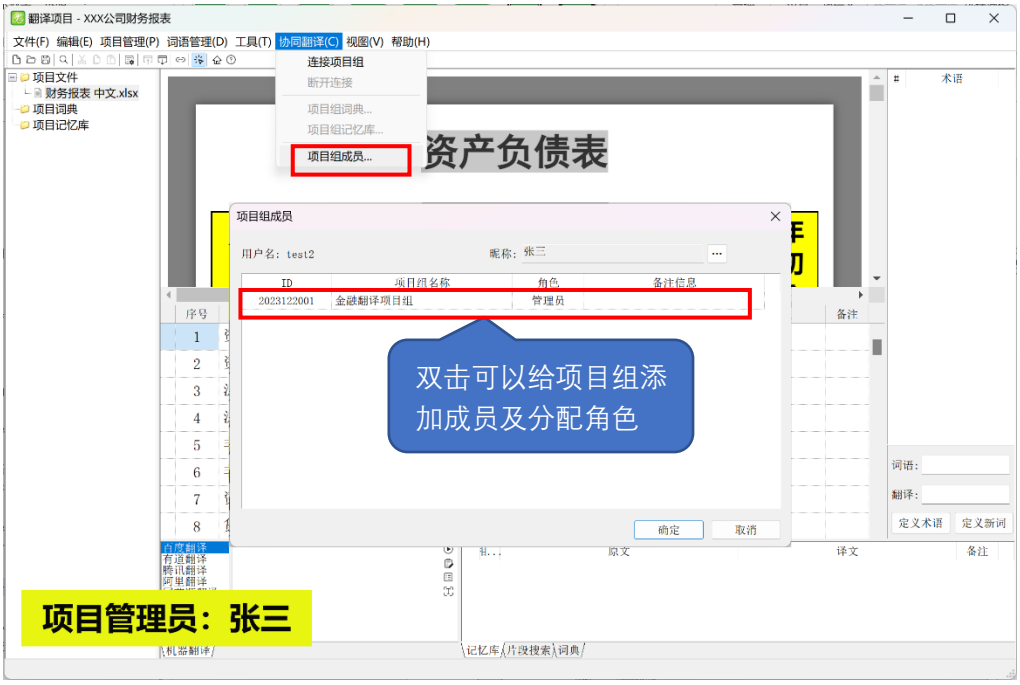
三、选择被授权加入的项目组名称

点击上图的“项目组名称”右侧的…按钮，会弹出允许当前用户加入的具体项目组名称供选择，双击某“项目组名称”即可选择加入该项目组。（例：由于“abc_server”服务器的“系统管理员 test 李明”在服务器中创建了一个“金融翻译项目组”，并指定了“test2 张三”为“项目管理员”，所以张三在该窗口中能看见“金融翻译项目组”并可以选择加入它。假如 “系统管理员”创建了多个“项目组”，并授权“test2 张三”为多个项目的“项目管理员”，则张三在这个窗口就可以看见多个项目组的名称供他选择。如果没有任何服务器的“系统管理员”授权允许“test2 张三”加入任何“项目组”，则张三看到的这个窗口是空白的）。

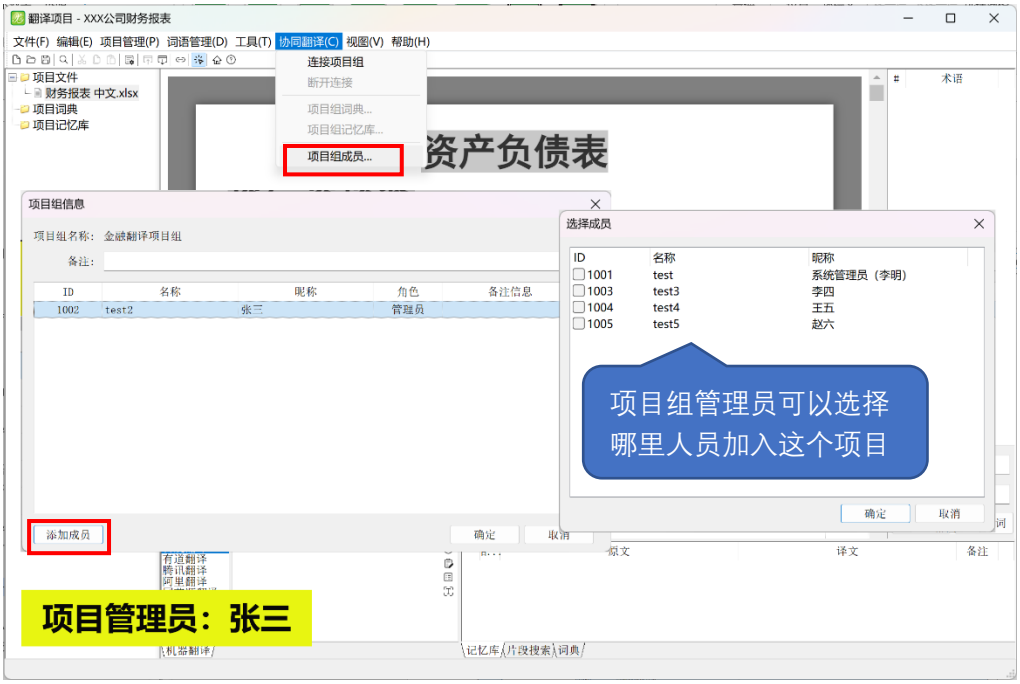
“项目管理员”在“协同翻译”窗口，还可以设置是否“允许使用未审核的术语”和是否“允许使用未审核的记忆库”。

四、添加项目组成员

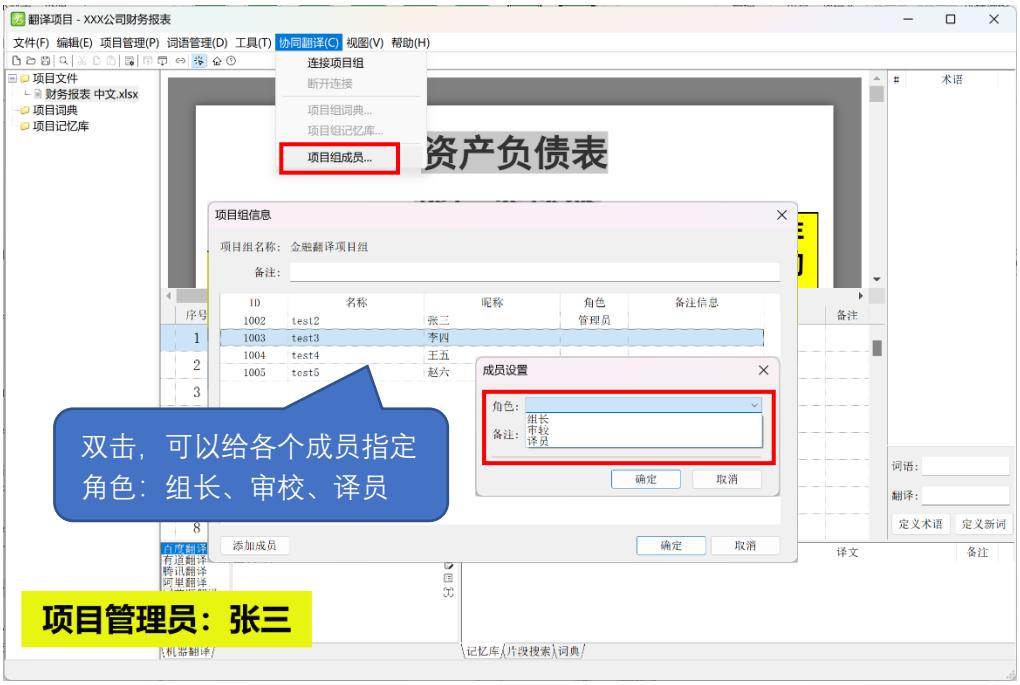
完成了前面的一、二、三项设置后，“项目管理员”可以在“协同翻译平台”中，挑选哪些人员加入这个项目组？及给这些人员指派具体的角色，角色分：项目组长、译员、审校。点击**【协同翻译】➔【项目组成员】**菜单，会弹出“项目组成员”窗口，如下图：在“项目组成员”窗口的上方，可以修改自己的昵称。



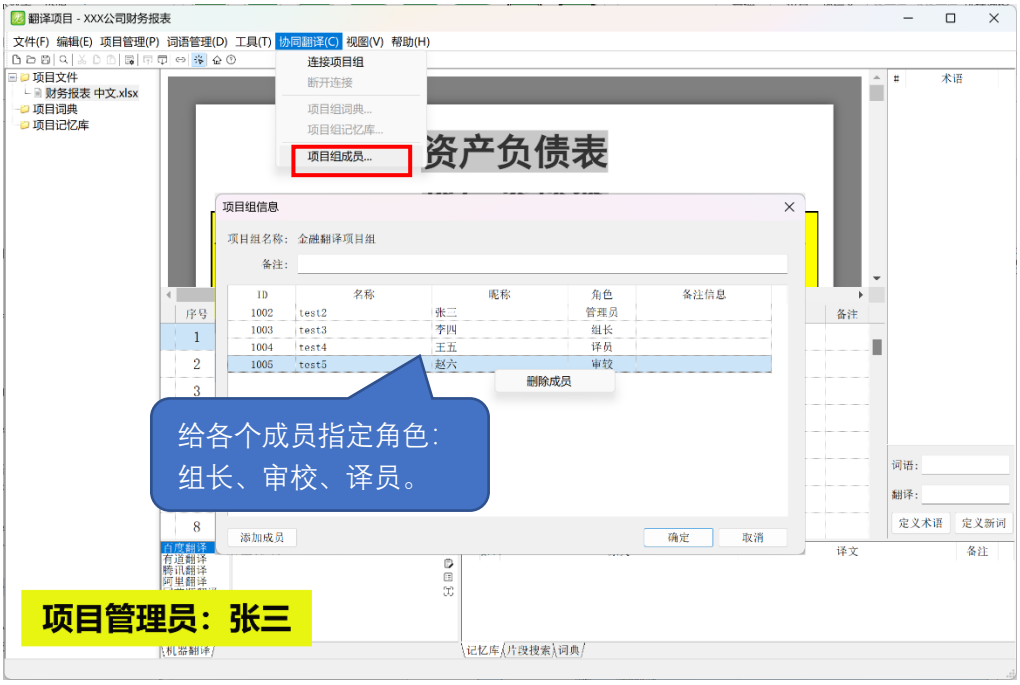
双击某项目组会弹出“项目组信息”窗口，如下图：点击“项目组信息”窗口下方的添加成员按钮，会弹出“选择成员”窗口，这些成员是经“协同服务器系统管理员”授权允许访问此服务器的成员。例如：“abc_server”协同服务器”的“系统管理员”——test/李明，授权了 test2 张三、test3 李四、test4 王五、test5 赵六 这几个人可以访问“abc_server”协同翻译服务器，所以，“项目管理员 张三”可以在下图的“选择成员”窗口中看见、选用这些人到他负责的项目中来。决定哪些人能访问服务器的权限在“系统管理员”，在有资格访问这个服务器的人员中，挑选哪些人加入某个“项目组”的权限在“项目管理员”。



挑选好项目组的成员后，点击确定按钮。例如：挑选了 test3 李四、test4 王五、test5 赵六 这几个人加入到“金融翻译项目组”中。如下图：



双击某个成员，可以指定他在该项目组担任的角色，有项目组长、译员、审校三个角色。“项目管理员”也可指定自己为“组长”，或指派其他成员担任“组长”角色。如下图：“项目管理员” test2 张三——指定 test3 李四为“组长”、test4 王五为“译员”、test5 赵六为“审校”。点击**【右键菜单】➔【删除成员】**可以删除某个成员。



五、连接项目组

点击【协同翻译】➔【连接项目组】菜单或点击工具栏的 “连接项目组”图标（该图标变成蓝色表示已经连接上了），可以将当前的翻译项目连接到服务器的某个项目组中，如下图：连接好后，翻译的内容会自动上传到服务器相应的项目的记忆库中，定义的术语也会自动上传到服务器的相应项目的术语库中。团队其他成员，可以实时共享、审核。



第二节 组长给项目组添加词典及记忆库

一、给项目组添加词典

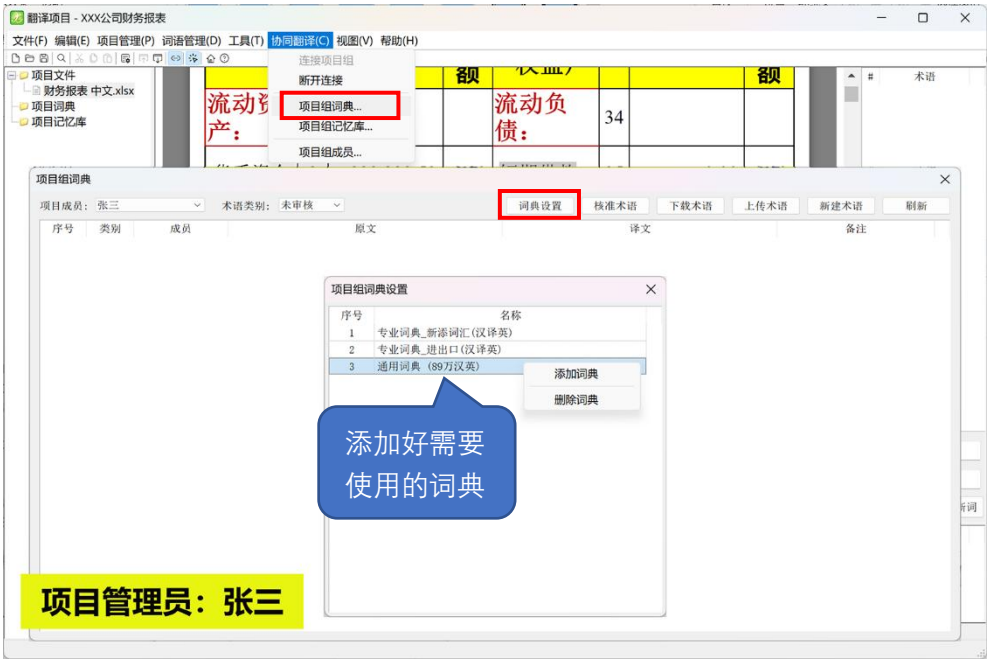
如果没有连接上项目组，**【协同翻译】➔【项目组词典】**菜单是灰色的、不可用。此时，应该先连接项目组，点击**【协同翻译】➔【连接项目组】**菜单或点击工具栏的 “连接项目组”图标。

连接上项目组后，点击**【协同翻译】➔【项目组词典】**菜单，会弹出“项目组词典”窗口，如下图：



如果是项目管理员或项目组长的界面，在窗口的上方会出现“词典设置”按钮，其他身份则没有这个按钮。点击词典设置按钮，会弹出“项目组词典设置”窗口，在窗口中点击**【右键菜单】➔【添加词典】**功能，系统会根据当前项目的语言方向给出可以选用的词典，前提是“系统管理员 test/李明”在“abc_server”协同翻译服务器中已经外挂这些词典。勾选需要的词典后，点击确定按钮即可，如下图：

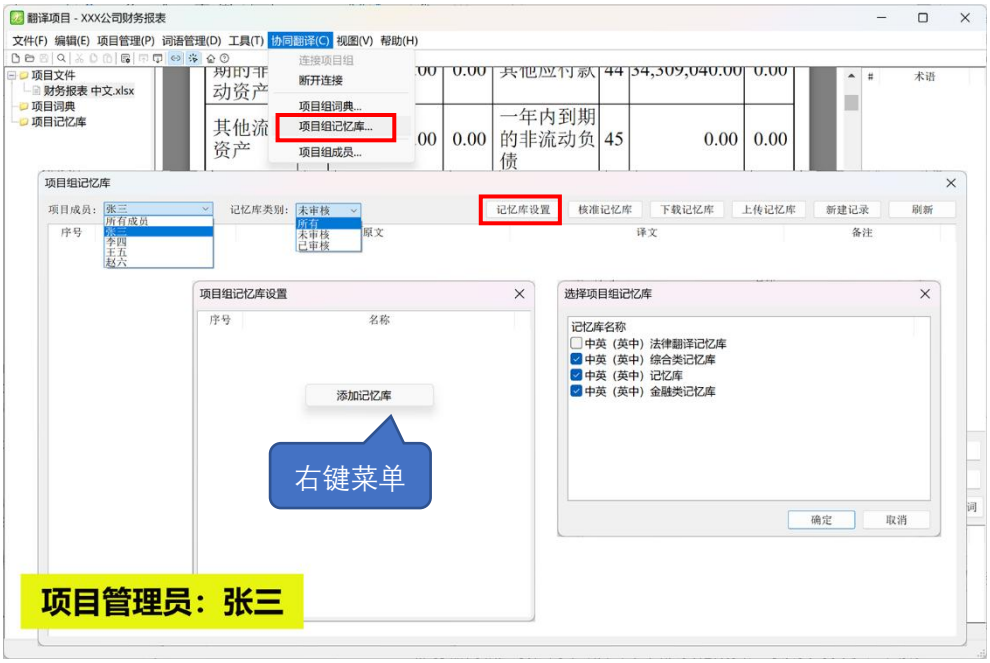
“系统管理员”在“abc_server”协同翻译服务器中外挂的 60 本词典，用户如果需要，可以加入“雪人 QQ 群”（群号 126371112），去 “群共享”中下载。教师可以创建自己的教学“协同翻译服务器”，在服务器中外挂词典供学生使用。



二、给项目组添加记忆库

如果没有连接上项目组，**【协同翻译】➔【项目组记忆库】**菜单是灰色的、不可用。此时，应该先连接项目组，点击**【协同翻译】➔【连接项目组】**菜单或点击工具栏的 🔗 “连接项目组”图标。

连接上项目组后，点击**【协同翻译】➔【项目组记忆库】**菜单，会弹出“项目组记忆库 ”窗口，如下图：



如果是项目管理员或项目组长的界面，在窗口的上方会出现“记忆库设置”按钮，其他身份则没有这个按钮。点击记忆库设置按钮，会弹出“项目组记忆库设置”窗口，在窗口中点击**【右键菜单】➔【添加记忆库】**功能，系统会根据当前项目的语言方向给出可以选用的记忆库，前提是“系统管理员 test/李明”在“abc_server”协同翻译服务器中已经外挂这些记忆库。勾选需要的记忆库后，点击确定按钮即可。



第三节 译员连接项目组进行翻译

下面以普通“译员”身份介绍：译员如何连接项目组进行翻译？

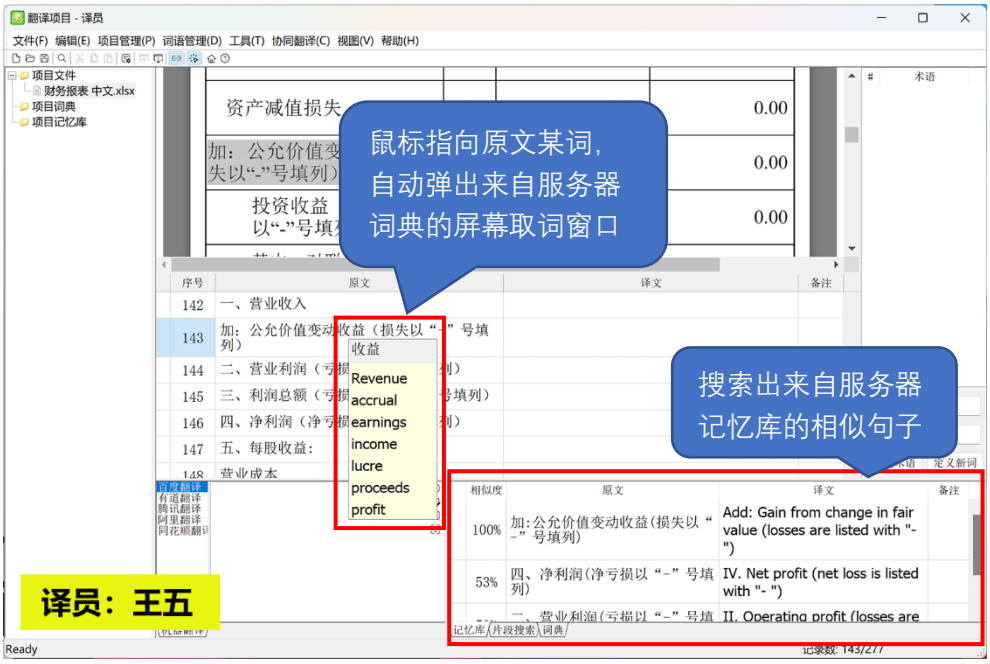
“项目管理员 test2/张三” 指定了“test/4 王五”为译员角色。“译员”王五在客户端界面中，选择语言方向新建翻译项目，导入需要的翻译的文档，然后点击**【项目管理】➔【项目设置】**菜单，在“项目设置”窗口中，点击“协同翻译”，选择能加入的服务器及项目组，如下图：



如果服务器的“系统管理员”选择了公开服务器地址，则团队成员不需要输入服务器地址。否则，要输入地址。更详细的介绍请参看本章的 “第一节 客户端如何连接服务器”。然后点击**【协同翻译】➔【连接项目组】**菜单或点击工具栏的 “连接项目组”图标（该图标变成蓝色表示已经连接上了）。如下图：



译员王五连接上项目组后，由于“项目管理员”张三在这个“金融翻译项目组”中外挂了词典和记忆库，团队成员就可以使用上服务器的这些资源，例如：鼠标指向原文的某个词停留片刻，会自动弹出屏幕取词窗口显示来自服务器词典的词义，点击即可输入。如果服务器的记忆库中有相似的句子，也会搜索出来现在屏幕的右下方，双击或按【F6】可以填入译文区。如下图：



还可以利用服务器的资源进行快速预翻译，点击【**工具**】➔【**预翻译**】菜单，会弹出“预翻译”窗口，如下图：勾选“使用服务器记忆库”和“使用服务器属于”这两个选项，如果服务器中有的句子，就会自动填写上，并在备注栏中显示“GTM”，表示该译文是来自服务器的。



服务器记忆库中能精确匹配的句子可能不多，点击下方的“片段搜索”，然后鼠标划选原文的某个词语，在“片段搜索”窗口立即从服务器的记忆库中搜索出该词的曾经译法供参考。如下图：鼠标划选原文的“递延”二字，立即从服务器的记忆库中搜索该词的曾经译法。

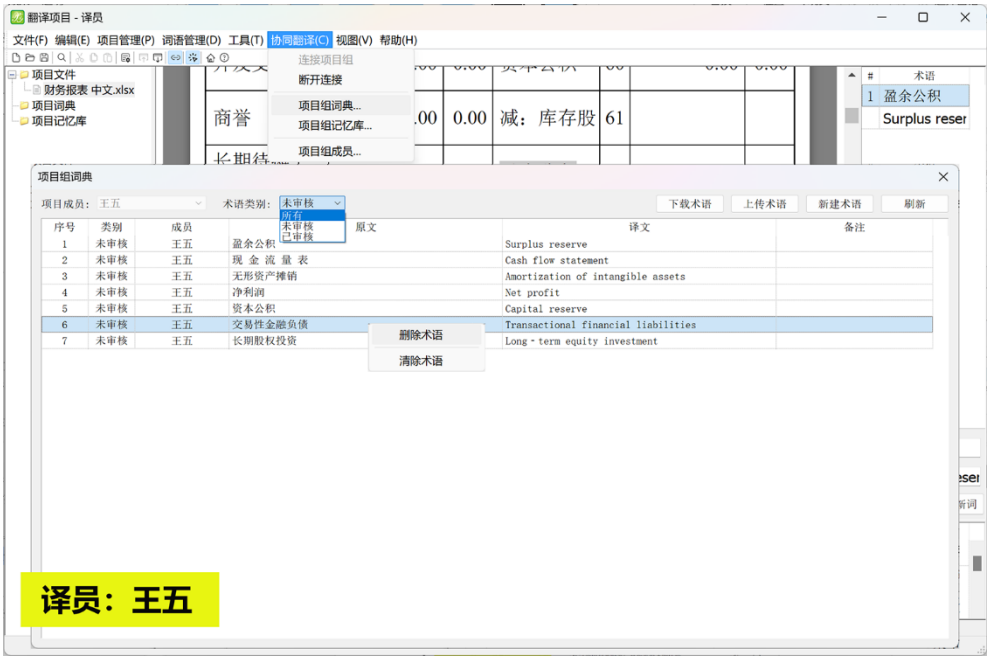


译员王五翻译的句子，会自动上传到协同翻译的“项目组记忆库”中，团队其他成员可实时看见。点击**【协同翻译】➔【项目组记忆库】**菜单，可以看见上传的这些句子。如下图：



普通译员身份，在“项目组记忆库”窗口，只能看见自己翻译的句子；如果是“项目组管理员”或“项目组长”或“审校”，可以看见该项目组所有成员翻译的句子。

译员王五翻译时定义的术语，会自动上传到协同翻译的“项目组词典”中，团队其他成员可实时看见。点击**【协同翻译】➔【项目组词典】**菜单，可以看见上传的这些术语。如下图：

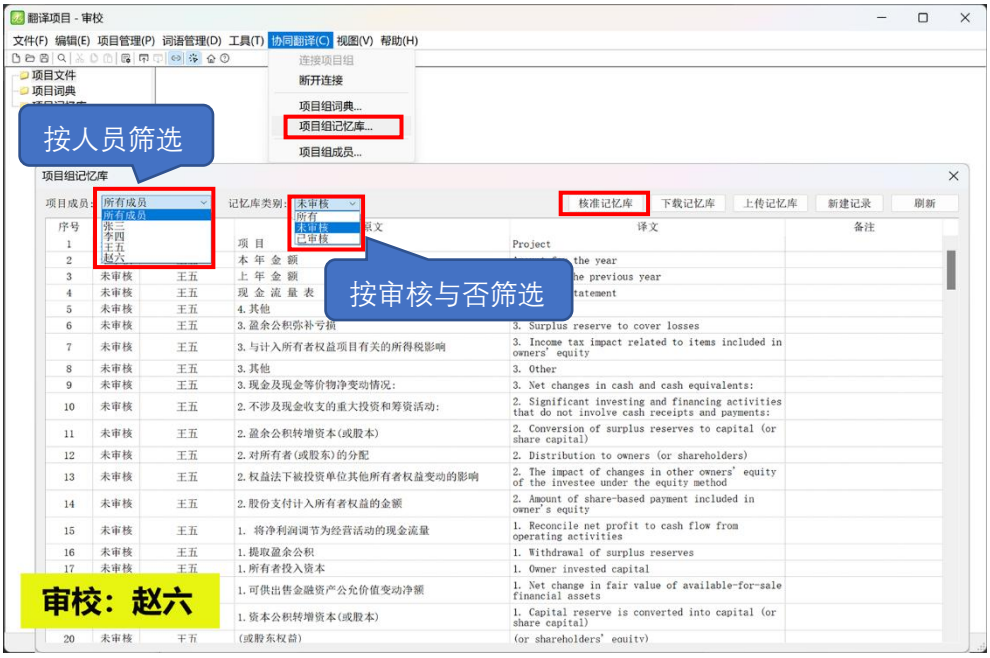


普通译员身份，在“项目组词典”窗口，只能看见自己定义的术语；如果是“项目组管理员”或“项目组长”或“审校”，可以看见该项目组所有成员定义的术语。

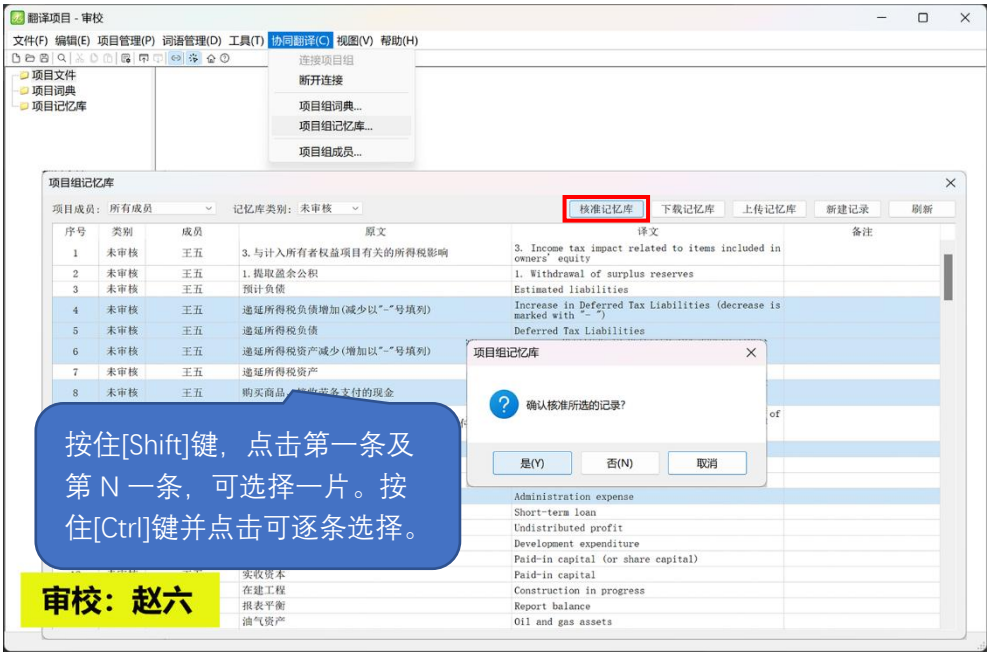
第四节 审校连接项目组进行审核

“审校”赵六在客户端界面中，新建翻译项目的语言方向要与译员的语言方向一致，然后点击**【项目管理】➔【项目设置】**菜单，在“项目设置”窗口中，点击“协同翻译”中选择能加入的服务器及项目组，再点击**【协同翻译】➔【连接项目组】**菜单或点击工具栏的 “连接项目组”图标（该图标变成蓝色表示已经连接上了）。

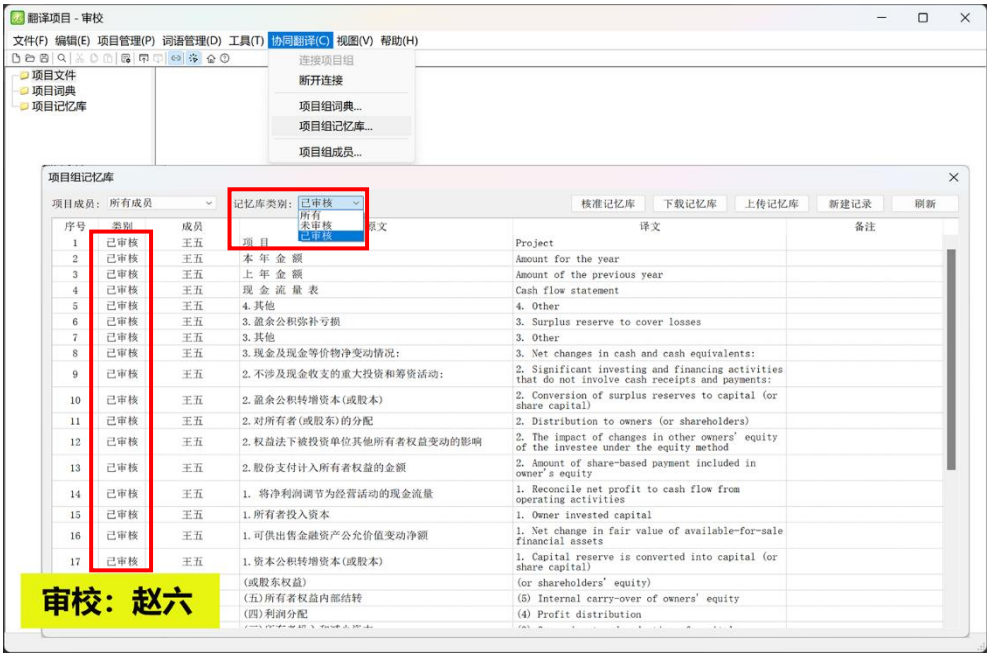
连接上项目组后，如下图：



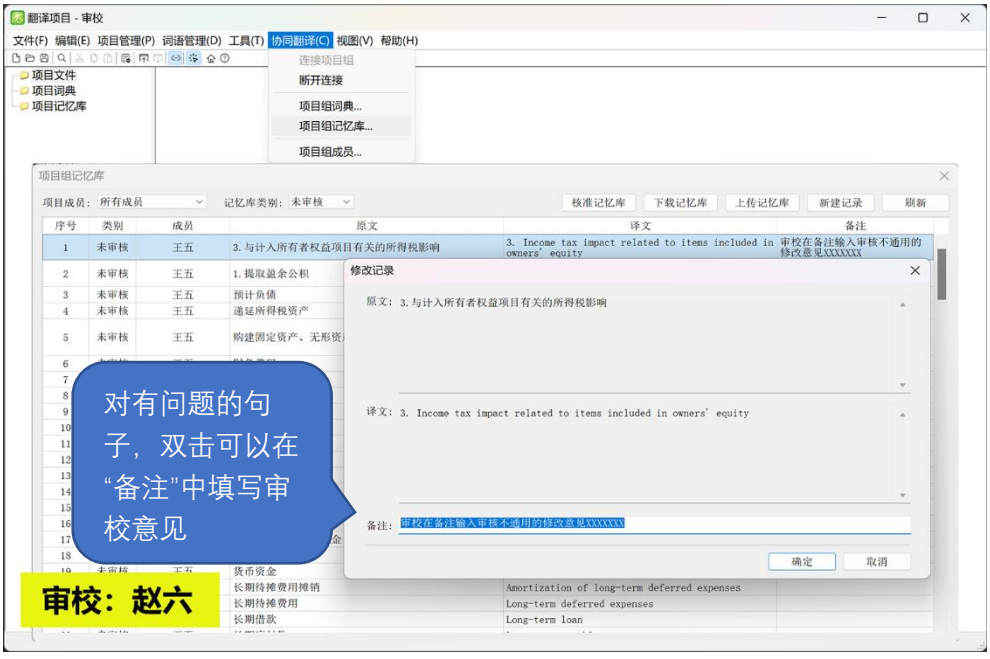
选择认为翻译正确的句子后，点击上方的“核准记忆库”按钮核准。按住[Shift]键，点击第一条和第 N 一条，可选择一片进行审核；按住[Ctrl]键并点击可不连续选择多条进行审核。如下图：



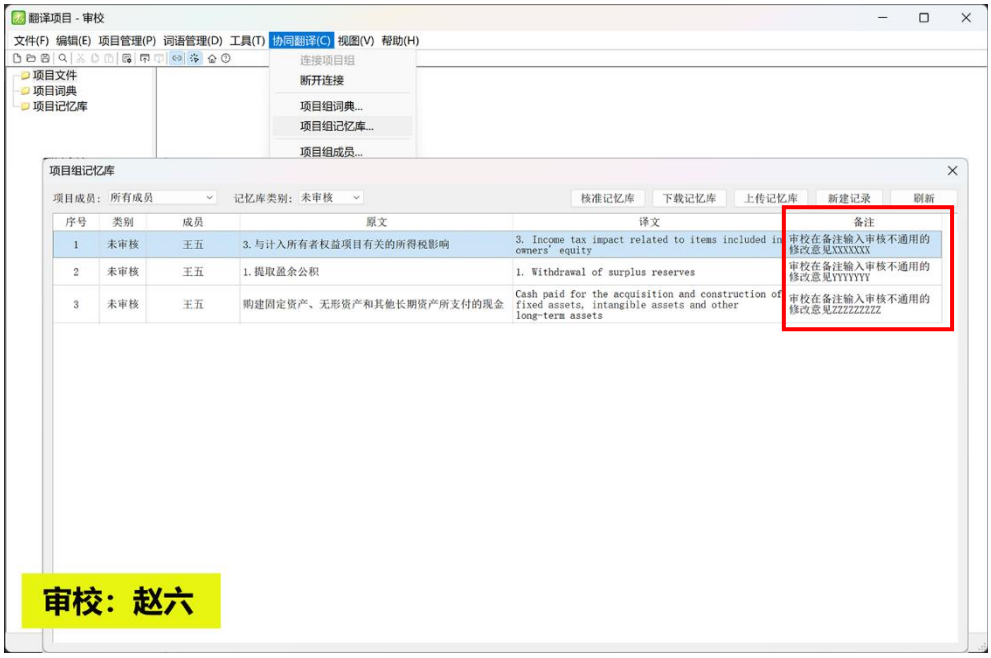
审核后的句子会从“未审核的库”放到“已审核的库”。点击“记忆库类别”旁边的下拉按钮，可以筛选出“已审核”的句子。如下图：



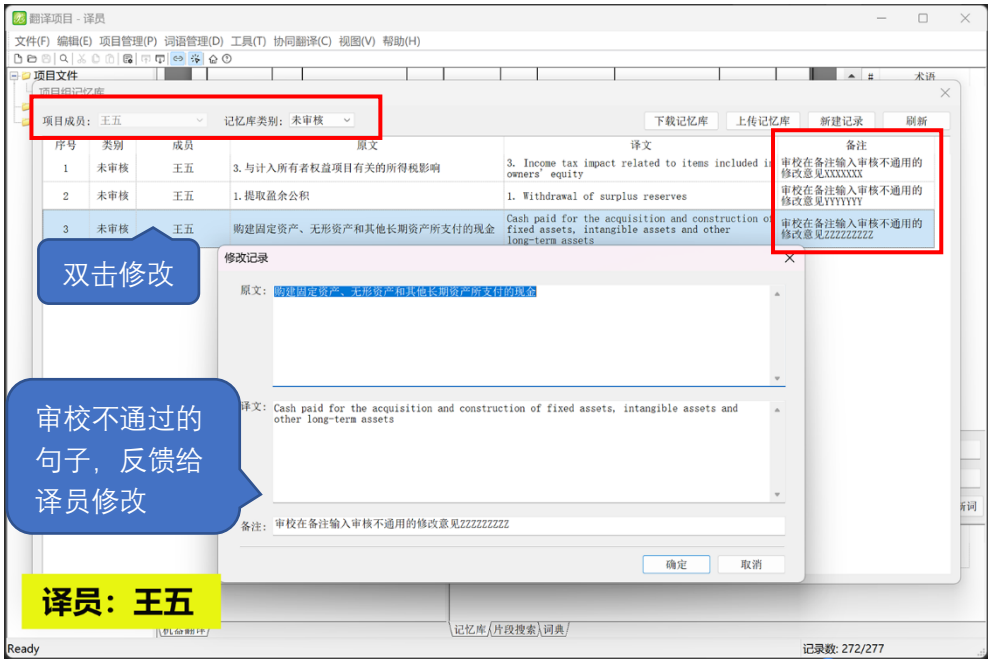
对有问题的句子，双击该句，会弹出“修改记录”窗口，审校可以在“备注”中填写修改意见。如下图：



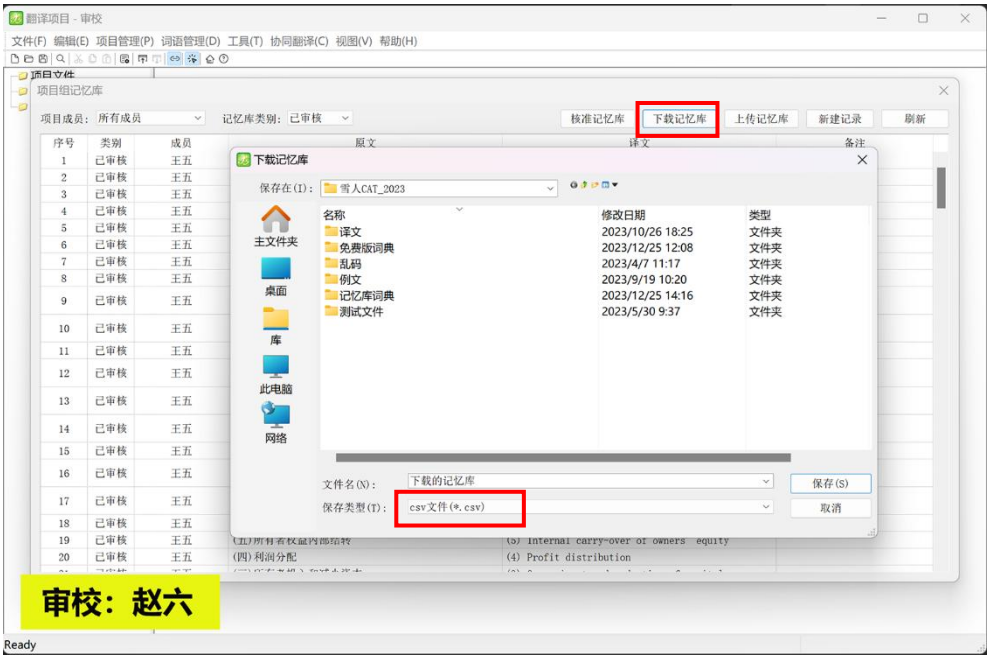
例如：下图中，有 3 个来自“译员”王五的句子审校没有通过，并在备注中填写了意见。



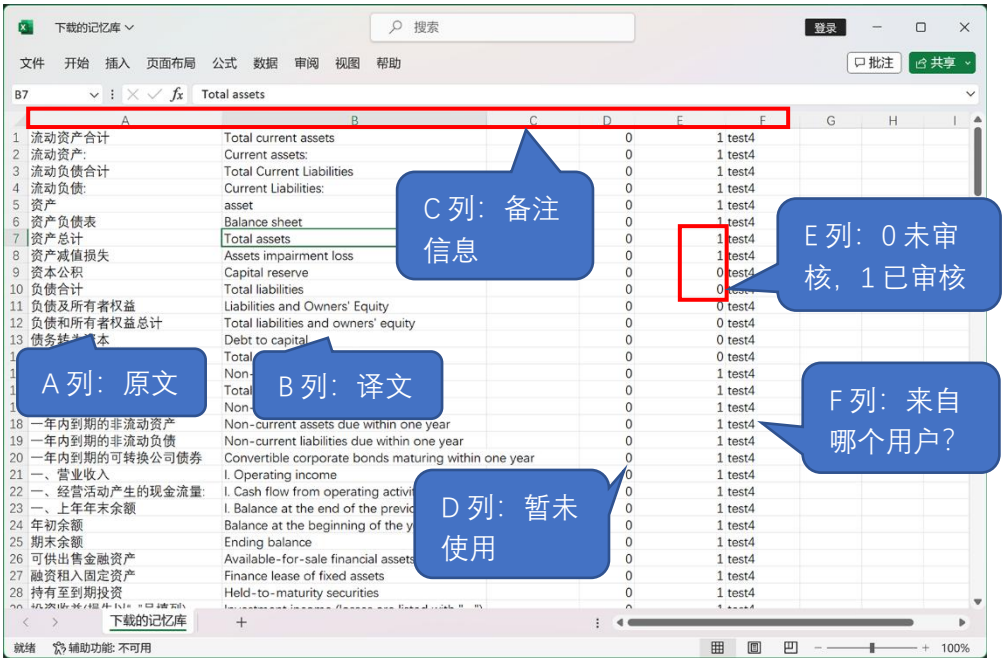
此时，“译员”王五的的界面可以马上看到审校的备注意见，如下图：双击某句子，可以进行修改。



也可以将数据导出后进行外部审校。点击“下载记忆库”按钮，可以将数据保存为一个“.csv”格式的文件，如下图：

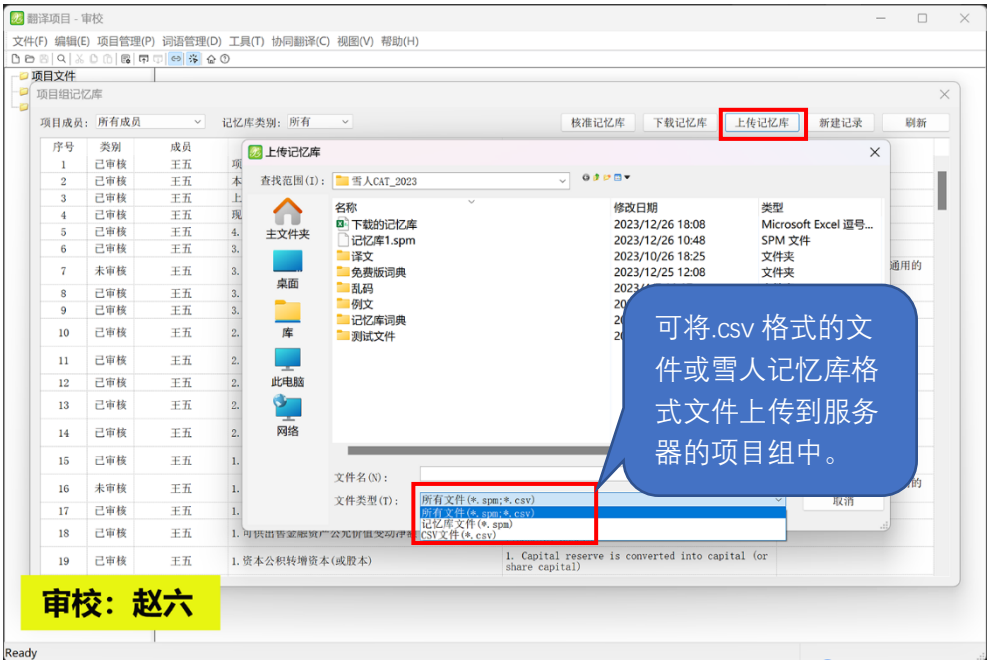


双击这个“.csv”格式的文件，可以在 EXCEL 打开它，如下图：A 列原文、B 列译文、C 列备注信息、D 列暂未使用该列的信息、E 列的 0 表示未审核　1 表示已审核、F 列显示的是用户的账号名，表示来自哪个用户翻译的？

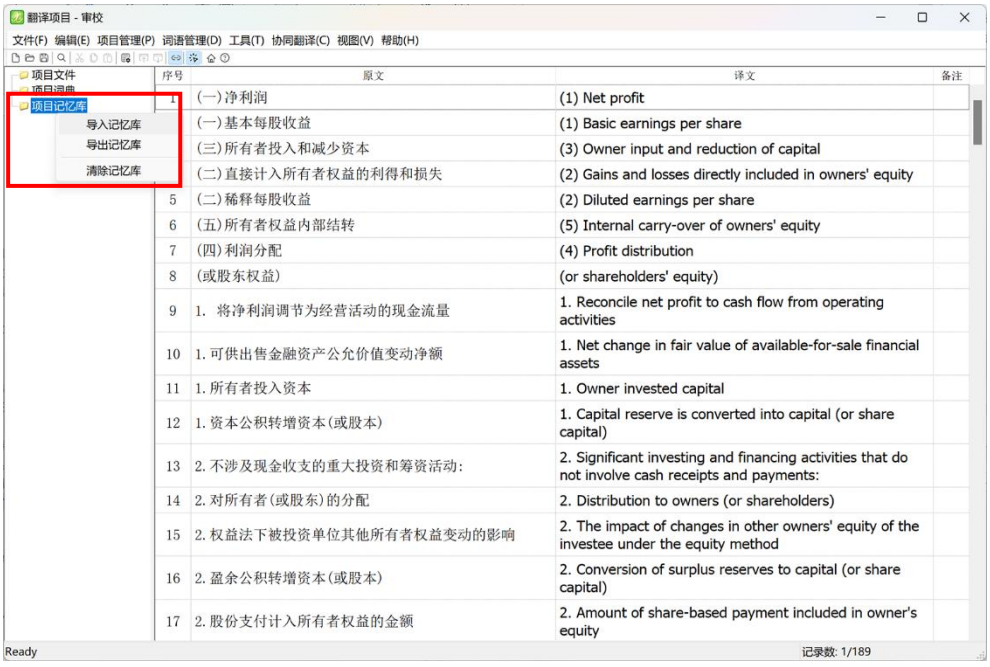


如果进行外部审核的数据，在 EXCEL 中按照上面的 A、B、C、D、E、F 列的约定填写好后，再另存为“.csv”格式的文件，点击“上传记忆库”按钮，就可以直接上传

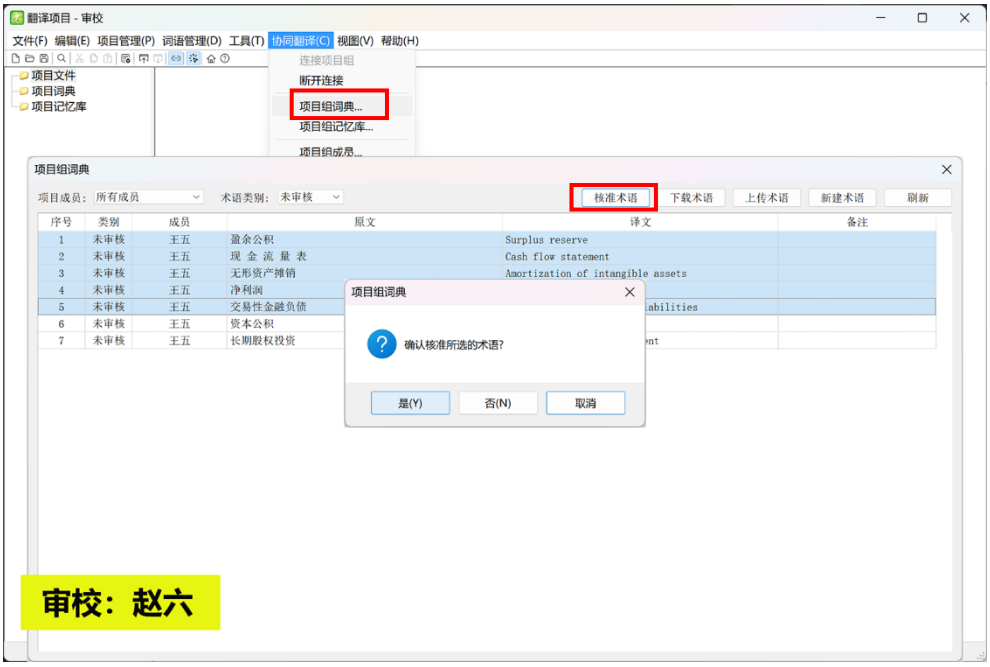
到服务器的项目组中。雪人格式的记忆库文件（.spm）也可以直接上传到服务器的项目组中。



下载的这个.csv 格式的记忆库文件，还可以直接在“项目记忆库”中导入，如下图：导入后，再导出转换成雪人的.spm 记忆库格式，在服务器中外挂。



点击【协同翻译】➔【项目组词典】菜单，会弹出“项目组词典”窗口，如下图：



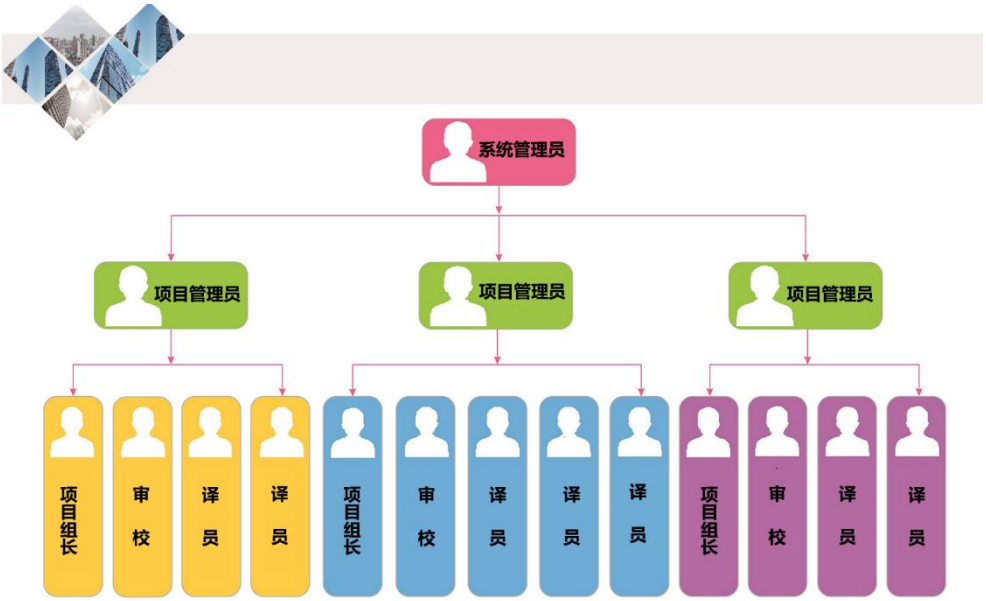
审校可以看到团队成员定义的术语，核准术语、上传术语、下载术语的操作和记忆库相类似，不累述。如果上传的是雪人(.spd)格式的词典文件，里面包含有普通词汇和术语，点击“上传术语”按钮，仅上传术语那部分的词语，普通词汇不会被上传。

在“协同翻译”设置窗口中，如果勾选了“允许使用未审核的记忆库”，翻译时如果遇到有相似的句子，会实时自动出现在翻译界面右下方的的记忆库中供参看，并提示“未审核”字样。如果勾选了“允许使用未审核的术语”，翻译时如果遇到该术语，会实时自动出现在翻译界面右侧的“术语提示区”中，并提示“未审核”记忆库中供参看。如果不勾选，则需要审核后，才可以使用。如下图：


翻译项目 - 译员
文件(F) 编辑(E) 项目管理(P) 词语管理(D) 工具(T) 协同翻译(C) 视图(V) 帮助(H)
导入文件(I)...
导出文件(E)...
导入词典...
导出词典...
导入记忆库...
导出记忆库...
项目设置...
项目文件
财务报表
项目词典
项目记忆库
资产负债表
项目设置
项目设置 项目引用 词典设置 记忆库设置 协同翻译
服务器名称: abc_server
服务器地址:
项目组名称: 金融翻译项目组
允许使用未审核术语
允许使用未审核记忆库
序号
术语
定义新词
备注
如果勾选使用未审核的术语、记忆库。一人翻译后他人立即可以使用，但会有“未审核”字样提示。如果不勾选，则需要审核后，才可以使用。

第五节 团队成员的权限

协同翻译平台由服务器+客户端构成，如果不需要与他人共享和团队协作，客户端软件就可以作为单机版使用。团队成员角色分为：系统管理员、项目管理员、项目组长、审校、译员。如下图：



团队成员权限表

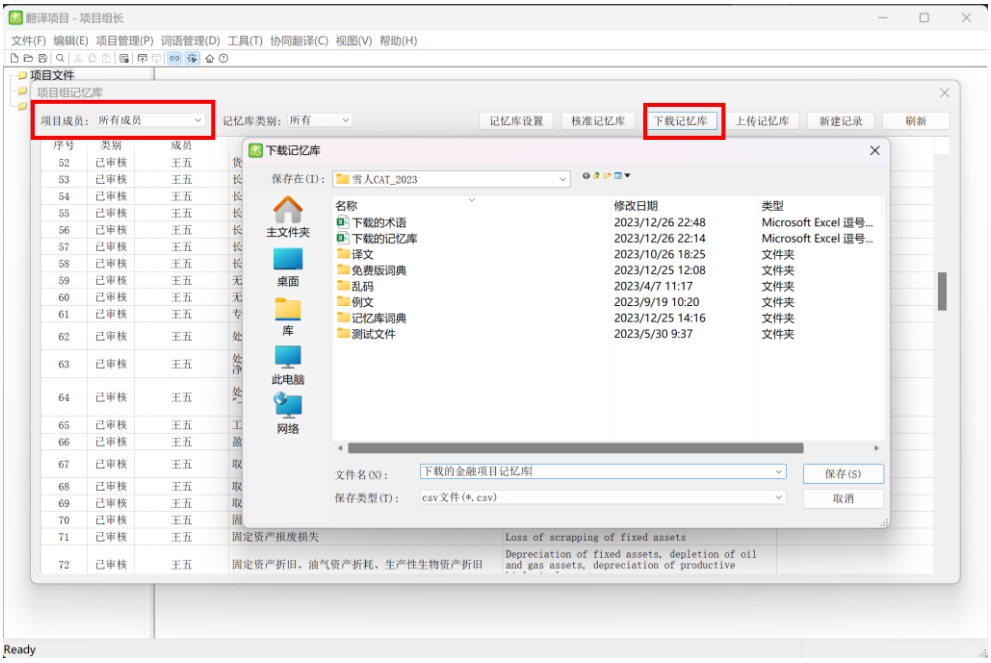
序号	权限内容	系统管理员	项目管理员	项目组长	审校	译员
1	启动、关闭服务器	√	X	X	X	X
2	服务器数据的备份、导出、迁移	√	X	X	X	X
3	授权谁可以访问服务器	√	X	X	X	X
4	在服务器中创建项目组	√	X	X	X	X
5	指定项目组的项目管理员	√	X	X	X	X
6	在服务器中添加项目组可引用的记忆库和词典	√	X	X	X	X
7	决定本项目选用哪些记忆库和词典	X	√	X	X	X
8	指定本项组的组长、审校、译员角色	X	√	X	X	X
9	下载本项目组积累的记忆库、词典	X	√	√	√	X
10	上传记忆库、词典到本项目组	X	√	√	√	√
11	下载本人创建的记忆库、词典	X	√	√	√	√
12	清除项目组的记忆库、词典	X	√	√	X	X

第六节 服务器项目组中数据的导出

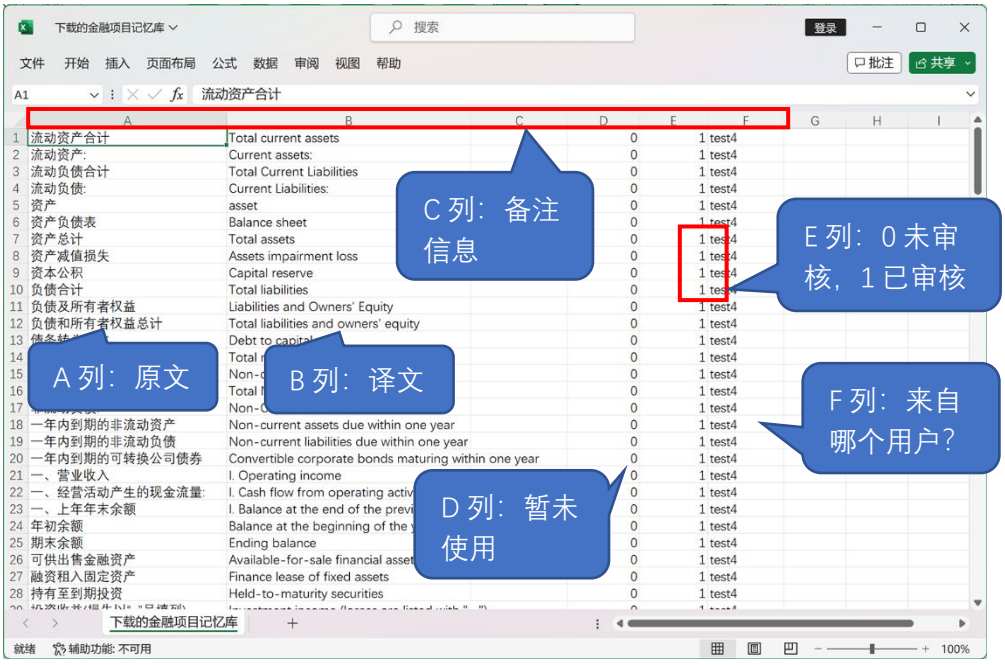
为了提高服务器的运行效率，服务器项目组中积累的数据应该不定期进行数据导出，导出后转成服务器外挂的方式供团队使用。单个项目组内的动态记忆库建议在 10 万条左右，在下图的“项目记忆库”窗口的动态记忆库中最多显示 1 万记录，（“项目词典”窗口显示的术语记录也是 1 万条），超过需要下载成“.csv”文件在 EXCEL 打开浏览。

一、下载项目组记忆库转外挂

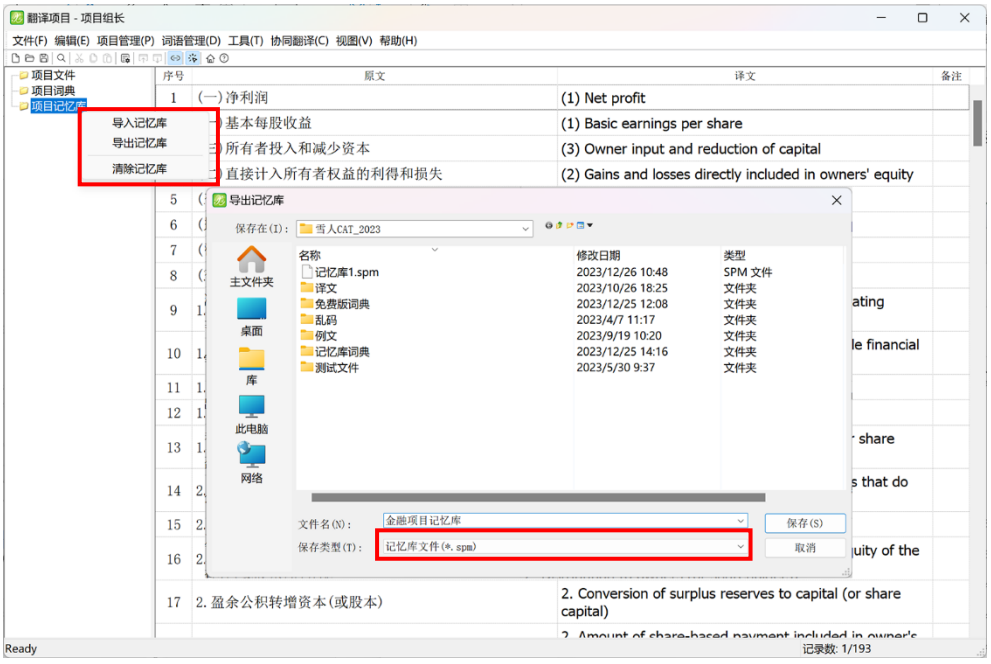
“项目管理”、“项目组长”、“审校”都有权限下载“项目组”积累的记忆库。点击**【协同翻译】➔【项目组记忆库】**菜单，会弹出“项目组记忆库”窗口，这个是项目组的动态记忆库。筛选所有成员、已经审核的记忆库，然后点击“下载记忆库”按钮。如下图：



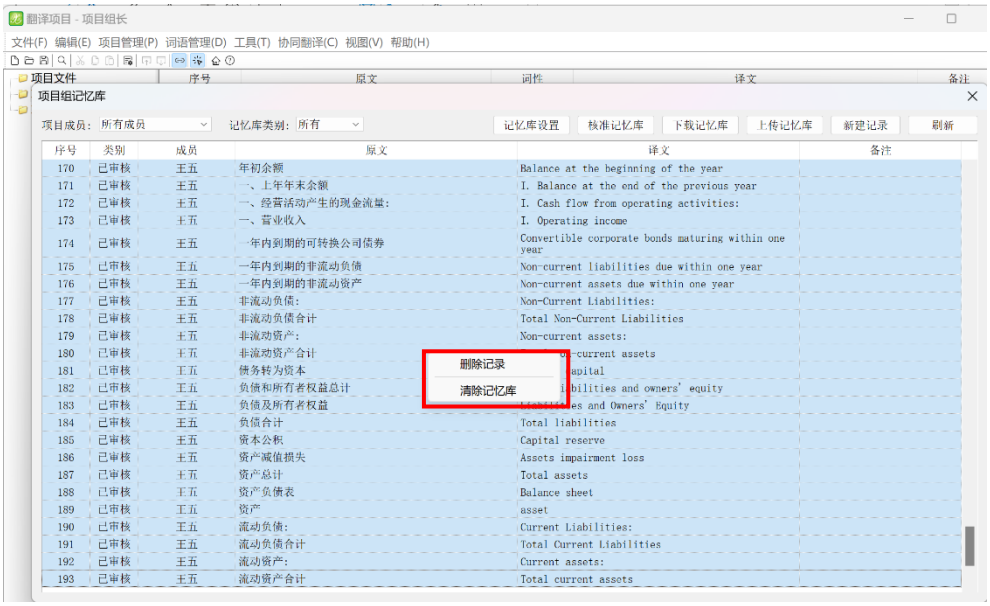
下载的是一个“.csv”格式的文件，双击这个“.csv”格式的文件，可以在 EXCEL 打开它，如下图：A 列原文、B 列译文、C 列备注信息、D 列暂未使用该列的信息、E 列的 0 表示未审核 1 表示已审核、F 列显示的是用户的账号名，表示来自哪个用户翻译的？



这个“.csv”格式的文件，可以在屏幕左侧的“项目记忆库”中导入，如下图：点击**【项目记忆库】➔【导入记忆库】**右键菜单进行导入，导入后再点击**【项目记忆库】➔【导出记忆库】**右键菜单导出，导出时选择“记忆库文件（*.spm）”的格式，转换成可以外挂的（*.spm）文件格式。将这个（*.spm）文件发给服务器的“系统管理员”外挂。

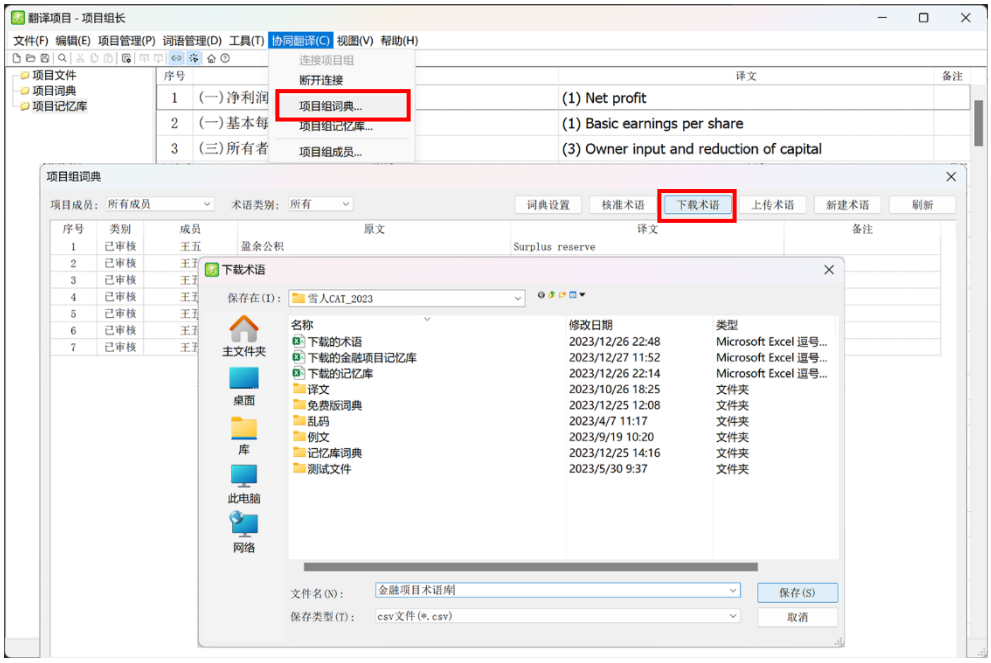


然后，可以将“项目组记忆库”中这些动态的记忆库记录删除，释放所占用的服务器资源。如下图，点击**【右键菜单】➔【清除记忆库】**功能可以全部删除。

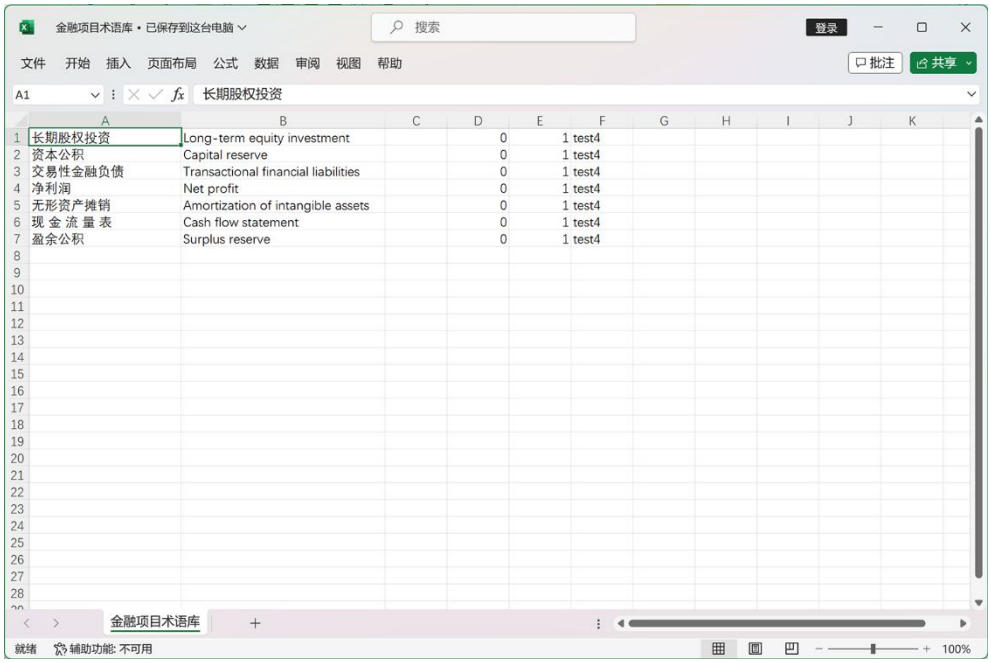


二、下载项目组词典转外挂

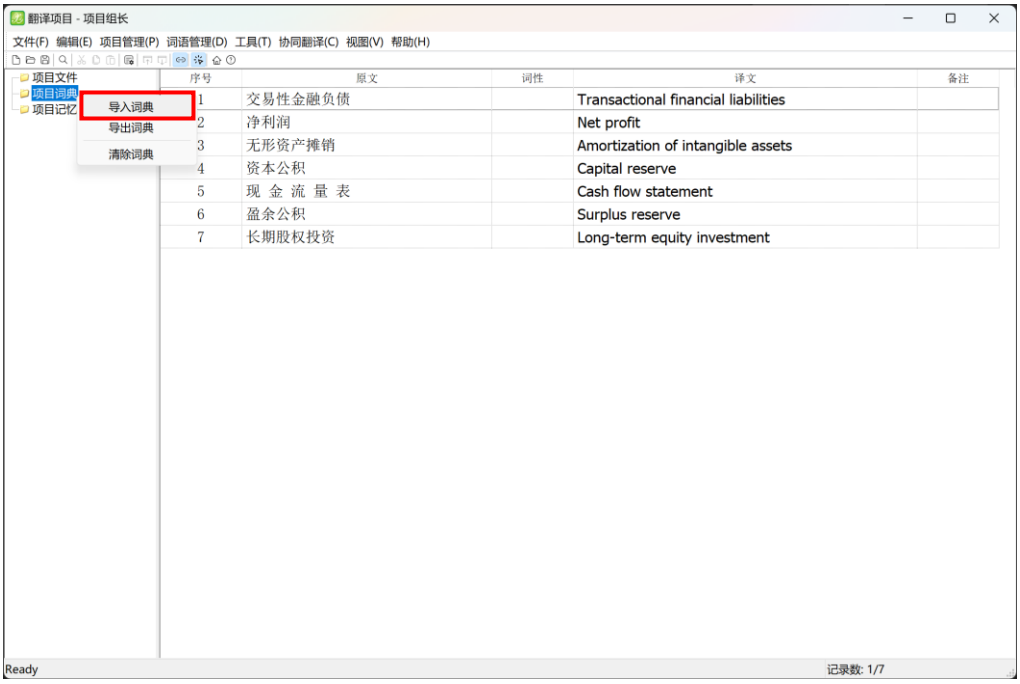
“项目管理”、“项目组长”、“审校”都有权限下载“项目组”积累的词典。点击**【协同翻译】➔【项目组词典】**菜单，会弹出“项目组词典”窗口，这个是项目组的动态词典库。筛选所有成员、已经审核的记录，然后点击“下载术语”按钮。如下图：



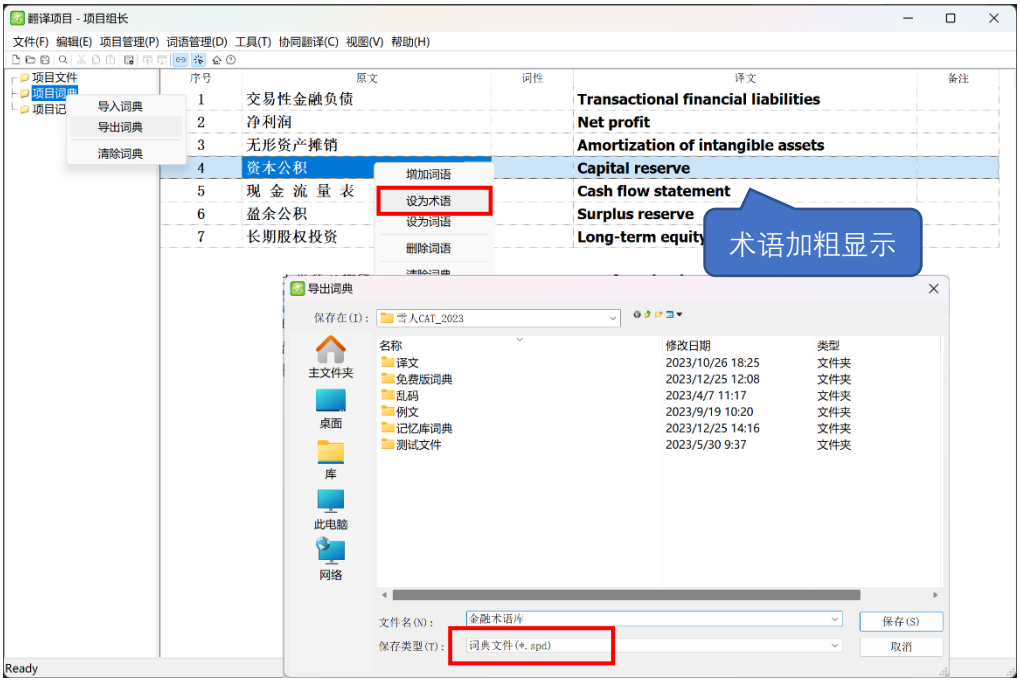
下载的是一个“.csv”格式的文件，双击这个“.csv”格式的文件，可以在 EXCEL 打开它，如下图： A 列原文、B 列译文、C 列备注信息、D 列暂未使用该列的信息、E 列的 0 表示未审核 1 表示已审核、F 列显示的是用户的账号名，表示来自哪个用户创建的?



这个“.csv”格式的文件，可以在屏幕左侧的“项目词典”中导入，点击**【项目词典】➔【导入词典】**右键菜单进行导入，如下图：

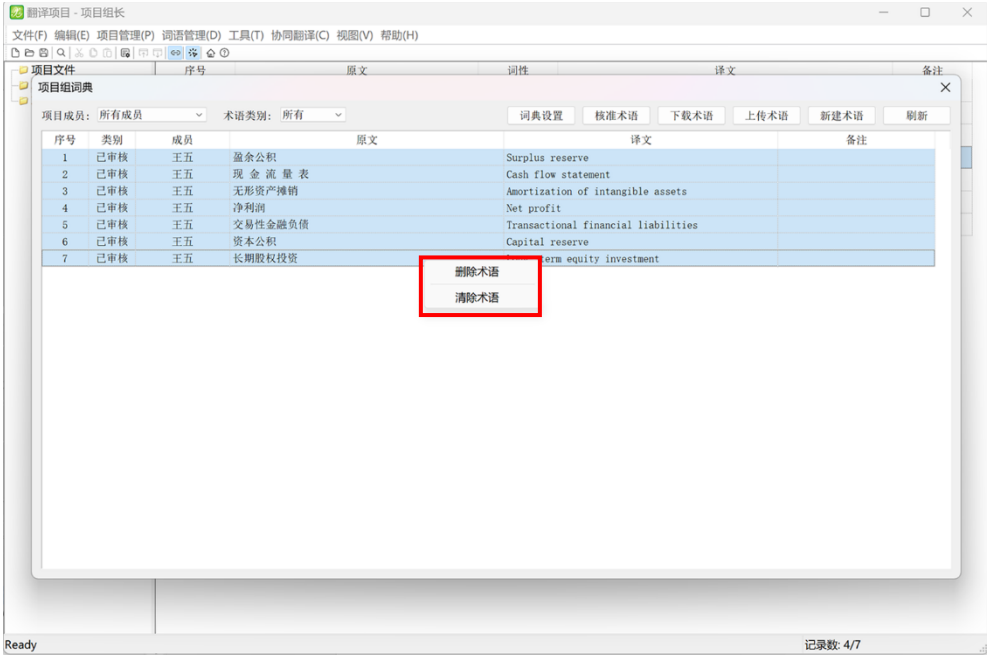


导入后默认是普通词汇，即未加粗显示。需要全选后再点击**【右键菜单】➔【设为术语】**功能将这些词汇设为术语。如下图：



再点击**【项目词典】➔【导出词典】**右键菜单导出，导出时选择“词典文件（*.spd）”的格式，转换成可以外挂的（*.spd）文件格式的词典。将这个（*.spd）词典文件发给服务器的“系统管理员”外挂。

然后，可以将“项目组词典”中这些动态的术语记录删除，释放所占用的服务器资源。如下图，点击**【右键菜单】➔【清除术语】**功能可以全部删除。



第十三章 在雪人 CAT 中接入 AI

第一节 AI+MTPE 人机交互翻译模式

ChatGPT 是由 OpenAI 公司于 2022 年 11 月发布的一种基于自然语言处理和深度学习技术的语言模型，全称为“Generative Pre-Trained Transformer”，ChatGPT 一经问世，迅速火爆全球，让更多人看到 AIGC(人工智能技术) 的发展前景。面对这一浪潮，国内多家企业纷纷响应，竞相推出了各自的人工智能大模型，其中包括：百度的“文心一言”、阿里的“通义千问”、智谱科技的同名产品“智谱”、腾讯的“混元”、科大讯飞的“讯飞星火”、百川智能的“百川”以及月之暗面的“Kimi”、抖音的“豆包”、商汤科技的“商量”、360 的“智脑”等等。

生成式人工智能语言模型的多轮智能对话、文本内容生成、自然语言理解等“全知全能”的强大性能在翻译领域展现出了超越想象的潜力，也给翻译行业带来了震撼性的冲击。AI 以其超强的知识集成和整合生成能力，在翻译的准确度和流畅度等方面取得了长足进展。AI 不仅具有严密的语言逻辑和强大的文本生成能力；还能根据译者的指令指出并修正翻译错误、调整句法结构、增强篇章连贯、调整语域，从多个方面对译本进行修改、校对、润色等译后编辑，其强大的功能集成性已经超越了普通的机器翻译。

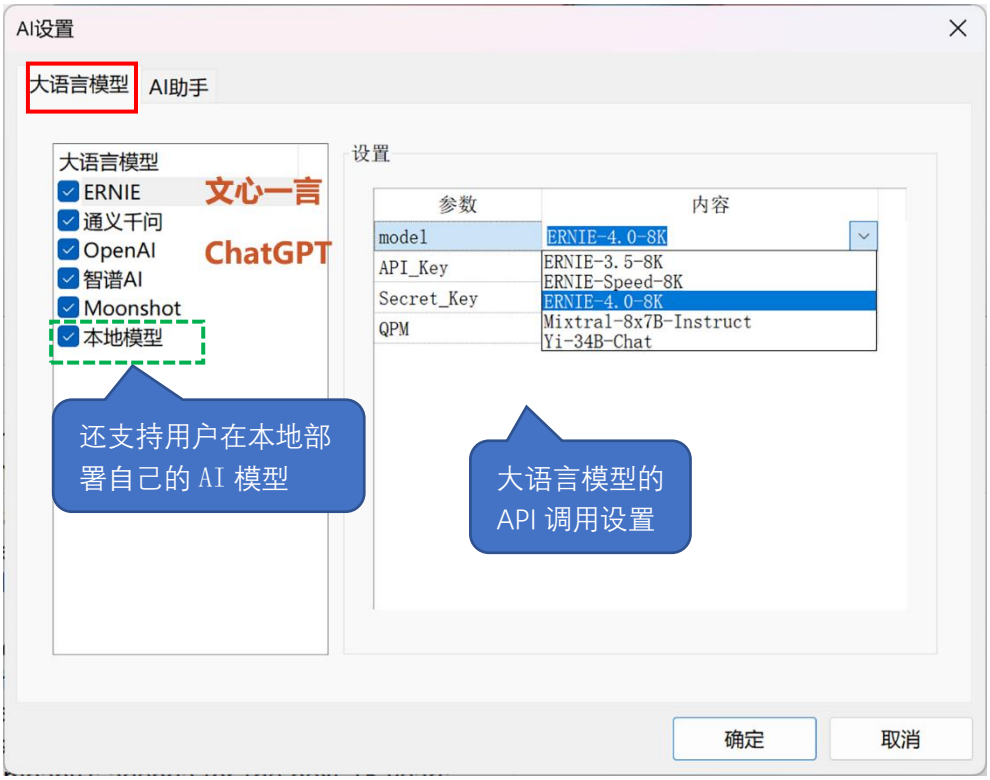
翻译技术的快速发展引发了翻译生产模式的变革，“AI+MTPE”(Machine Translation Post－editing)交互式译后编辑必将成为翻译产业的主流工作模式，译者可以实时与 AI 进行交互，针对特定翻译难点或特定语境需求给出 Prompt(提示语)，AI 随即生成对应译文供译者参考或直接采纳。这种“人机交互模式”是“人的智慧”与“机器智能”的高度交融。融合人工智能技术的机器翻译，以及 AI 技术的飞速发展和迭代，将对翻译行业、翻译教育带来了颠覆性的影响。

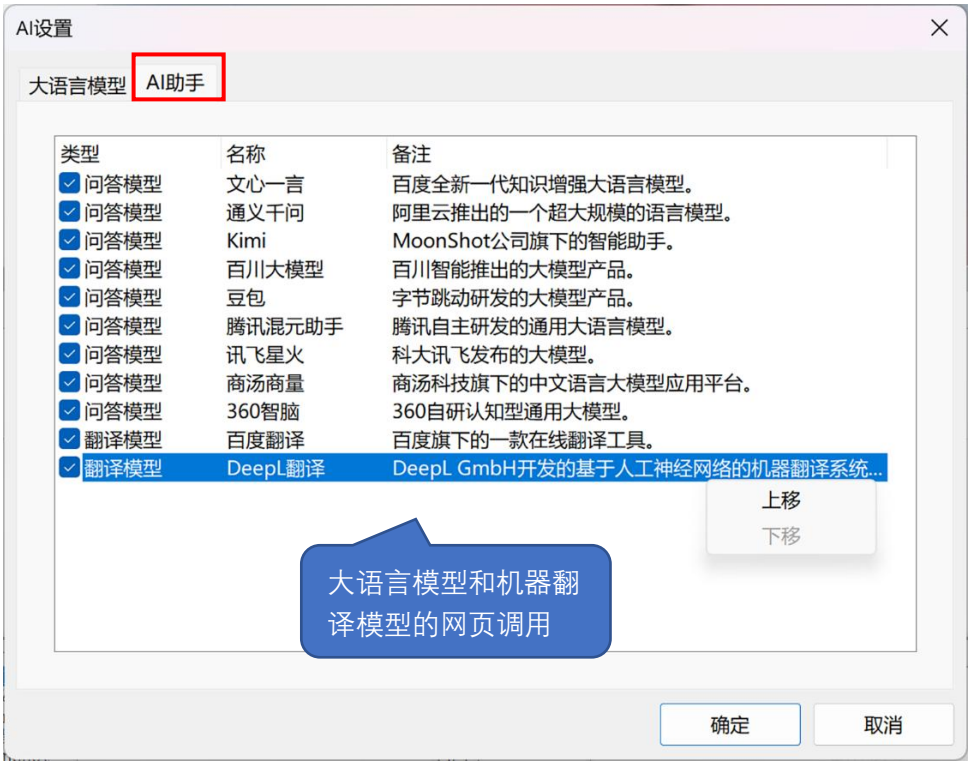
第二节 注册 AI 账号

雪人将 AI 技术和 CAT 技术深度融合，让 AI 技术更加高效地服务于翻译工作。同时开设了“雪人翻译软件 AI 交流群”（群号为：526352494），欢迎加入讨论交流！点击软件主菜单**【AI 应用】➔【AI 设置】**功能（注：AI 应用需要 VIP 权限），如下图：



在【AI 设置】窗口中，“大语言模型”页面是多个主流的“大语言模型”的 API 调用设置界面，其中包括:ERNIE(文心一言)、通义千问、OpenAI(ChatGPT)、智谱 AI、Moonshot。还支持用户在本地部署自己的 AI 模型。如下图:


AI设置
大语言模型
AI助手
大语言模型
ERNIE
文心一言
通义千问
OpenAI
ChatGPT
智谱AI
Moonshot
本地模型
还支持用户在本地部署自己的 AI 模型
设置
参数
内容
model
ERNIE-4.0-8K
API_Key
Secret_Key
QPM
ERNIE-3.5-8K
ERNIE-Speed-8K
ERNIE-4.0-8K
Mixtral-8x7B-Instruct
Yi-34B-Chat
大语言模型的 API 调用设置
确定
取消

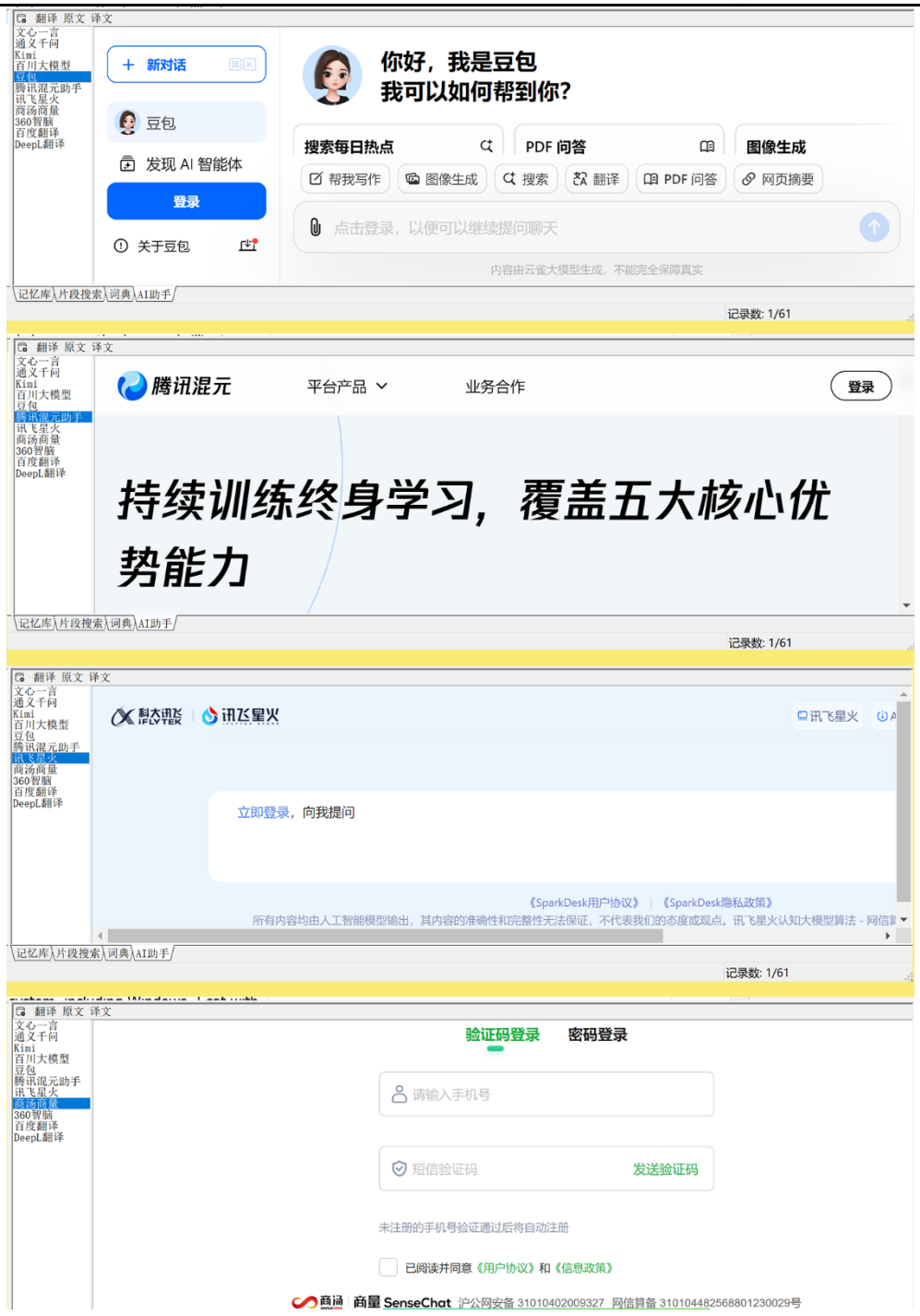

AI设置
大语言模型
AI助手
类型
名称
备注
问答模型 文心一言 百度全新一代知识增强大语言模型。
问答模型 通义千问 阿里云推出的一个超大规模的语言模型。
问答模型 Kimi MoonShot公司旗下的智能助手。
问答模型 百川大模型 百川智能推出的大模型产品。
问答模型 豆包 字节跳动研发的大模型产品。
问答模型 腾讯混元助手 腾讯自主研发的通用大语言模型。
问答模型 讯飞星火 科大讯飞发布的大模型。
问答模型 商汤商量 商汤科技旗下的中文语言大模型应用平台。
问答模型 360智脑 360自研认知型通用大模型。
翻译模型 百度翻译 百度旗下的一款在线翻译工具。
翻译模型 DeepL翻译 DeepL GmbH开发的基于人工神经网络的机器翻译系统...
上移
下移
大语言模型和机器翻译模型的网页调用
确定
取消

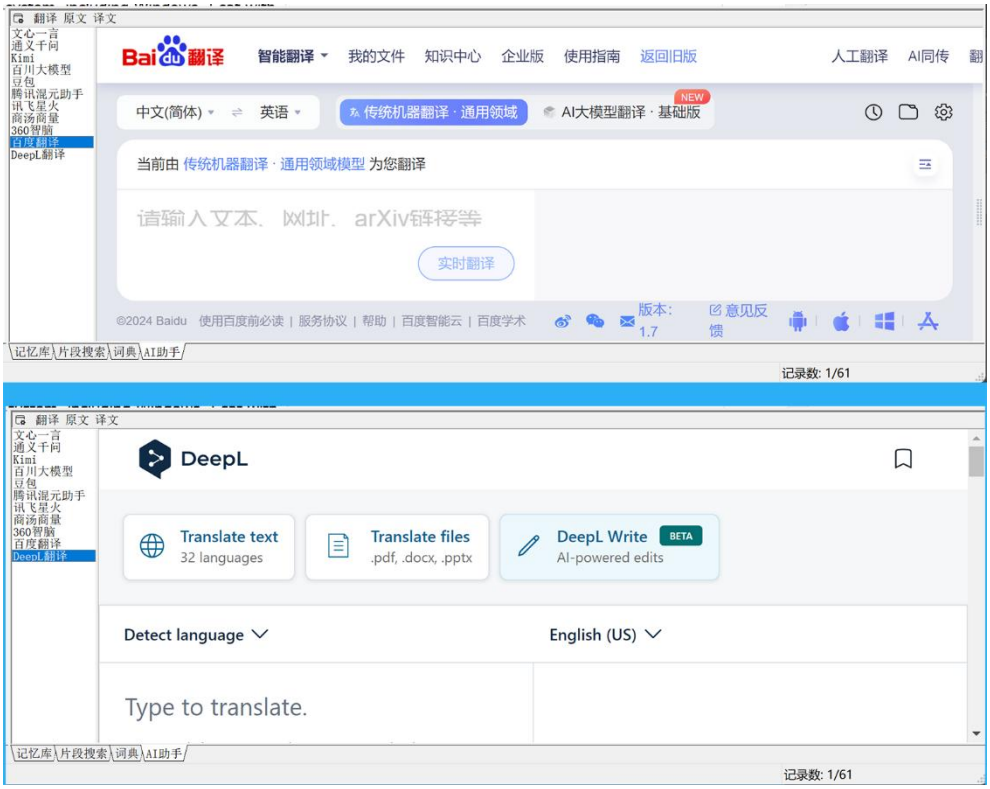
在“AI 助手”页面，是用网页方式调用的 AI 问答模型和翻译模型，目前支持的问答类模型有：文心一言、通义千问、Kimi、百川、豆包、腾讯混元、讯飞星火、商汤商量、360 智脑。机器翻译模型有：百度翻译、DeepL 翻译。

网页方式调用的好处是注册简单，无需复杂的设置。问答模型不仅可以翻译，还可以对译本进行修改、校对、润色等译后编辑，也可以向问答模型咨询某词的译法、解析，翻译时遇到各种问题，可以随时向它提问，AI 会从它的海量预训练知识库中帮你找到问题的答案供参考。在上图的【AI 助手】窗口中，点击**【右键菜单】➔【上移】或【下移】**功能，可以调整它们的排列顺序，如果不需要某个模型，可以点击最左边的“√”取消它。

欲使用这些网页版的 AI 问答模型，首先需要注册 AI 账号。如下图：




翻译 原文 译文
文心一言
通义千问
Kimi
百川大模型
豆包
腾讯混元助手
讯飞星火
商汤商量
360智脑
百度翻译
DeepL翻译
新对话
豆包
发现 AI 智能体
登录
关于豆包
你好，我是豆包
我可以如何帮到你？
搜索每日热点
PDF 问答
图像生成
帮我写作
图像生成
搜索
翻译
PDF 问答
网页摘要
点击登录，以便可以继续提问聊天
内容由云雀大模型生成，不能完全保障真实
记忆库
片段搜索
词典
AI助手
记录数: 1/61
腾讯混元
平台产品
业务合作
登录
持续训练终身学习，覆盖五大核心优势能力
讯飞星火
立即登录，向我提问
《SparkDesk用户协议》 | 《SparkDesk隐私政策》
验证码登录
密码登录
请输入手机号
短信验证码
发送验证码
未注册的手机号验证通过后将自动注册
已阅读并同意《用户协议》和《信息政策》
商量 SenseChat 沪公网安备 31010402009327 网信算备 310104482568801230029号



注册账号的方法很简单，点击雪人软件界面下方的“AI 助手”页面，然后点击左侧的各个 AI 模型，输入手机号即可以注册账号，各个问答模型的注册方法都类似，需要使用哪几个就注册它们。目前软件还提供了一些网页版本的在线翻译，例如 DeepL 翻译和百度翻译。

第三节 面向翻译的提示语(Prompt)工程

为了使人机交互真正成功，我们需要清晰、有效地对人工智能系统表达我们的需求，精心制作的指令是人工智能的指路明灯，指导它的行为并确保生成的输出内容与预期目标保持一致。提示语(Prompt)工程提供了人机对话更加自然和直观的方式。通过设计恰当、精妙的提示语，引导 ChatGPT 类的大语言模型生成我们期望的回答或完成特定任务。好的提示语可以让 AI 大语言模型在各种应用场景中表现出更高的准确性和智能性。

面向翻译的提示语工程

通过与 AI 大语言模型进行人机交互的"问答"形式，可以完成查阅、翻译、校对、润色等译前、译中、译后等各项翻译任务。基于后续提问不断完善译文是 AI 大语言模型翻译的特色。

和普通的机器翻译工具，如谷歌翻译、DeepL 翻译等相比，AI 大语言模型翻译所提供的机译服务超越了普通机器翻译，它可以通过构建合理、清晰的提示语来提升翻译的准确度，实现人机交互式翻译。提示语工程（Prompt Engineering）已经成为使用 ChatGPT 类 AI 工具所必须的技能。清晰而准确的提示语有助于 AI 理解你的请求和输出高质量的译文。例如：在提示语中提供一些背景信息，可以引导 AI 去调用与所提供的背景信息更加相关的数据，使翻译结果更加贴合需求;在提示语中也可以指定自己想要的翻译风格，如正式、学术、口语、诗意、幽默、商务、庄重等等风格；给 AI 大语言模型指定要翻译的领域也会让它的精准度提高，比如指定要翻译的领域是法律文书、商业计划书、新闻、医学报告、政府公文、学术论文等等；如果给 AI 一些实例的话，它会更懂你的需求，告诉 AI 去参考某某翻译家的风格，它生成的内容会有此翻译家的味道，如参考翻译家严复的翻译风格：（严复主张翻译应遵循"信、达、雅"的原则，忠实原文（信）、表达流畅（达）、文字优美（雅））。

也可以把 AI 大语言模型翻译和其他的机器翻译工具结合起来。先用专门的翻译工具进行预翻译，如果翻译的效果不理想，可以用 AI 大语言模型进行优化。比如：很多时候专门的翻译工具翻译出来的译文不地道，显得比较生硬，有明显的机翻痕迹，这时候可以让 AI 大语言模型对翻译的文本进行修饰，优化翻译文本的质量，通过提示语让最终的译文更加自然、地道和生成自己想要的风格。

AI 大语言模型可校对或润色译文，例如可以通过提示语要求译文句式复杂化，选用更为地道的用词表达。ChatGPT 类的大语言模型集成了 Grammarly、Phrase

Bank、Collocaid 等软件的校对、替换、改述功能；可以准确地识别拼写错误、语法错误、词汇错误、简单逻辑错误等，且可针对错误予以一定的分析与评估，并可以在评价中给出相关错误的原因及修改建议，其对错误识别比较精准，给出的修改意见也较为合理，通常能够为译员提供较为明确的反馈意见。

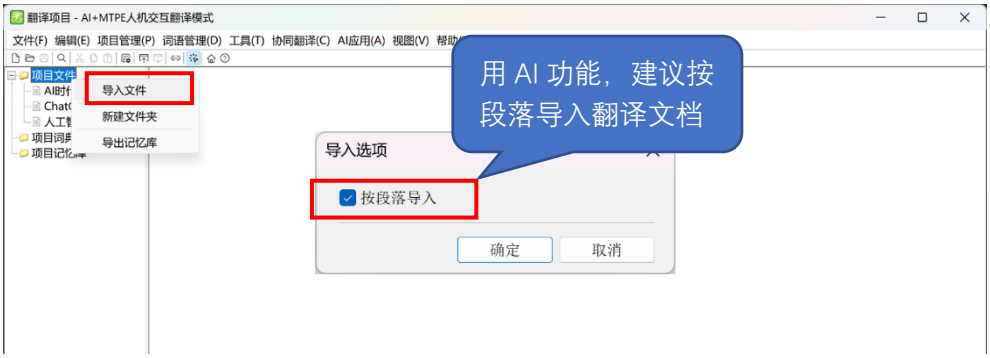
通过使用高质量的提示语（Prompt），可以得到更好的结果。一个好的提示语（Prompt）的结构:

- 角色、技能、个性
- 目标、任务
- 生成规则
- 输入输出格式
- 示例

可以在输出格式中要求 AI 输出多种不同类型的翻译风格：意译、直译、口语化翻译、创意翻译、文学翻译。这样就能得到多个不同角度的翻译结果供参考。

第四节 AI 助手

AI 助手是使用网页方式调用各个 AI 大模型，目前支持有：文心一言、通义千问、Kimi、百川、豆包、腾讯混元、讯飞星火、商汤商量、360 智脑，这些模型可以执行翻译、问答等各种任务。机器翻译模型有：百度翻译、DeepL 翻译。网页方式调用的优势是注册简单，各个模型目前可以免费使用。使用 AI 进行翻译时，建议导入翻译的文档时，勾选“按段落导入”选项（如下图），以段落为单位发送给 AI 大语言模型进行翻译，以段落为单位有更多上下文的信息，有利于 AI 提高翻译质量和降低向模型发送请求的频率。



为了方便用户编写提示语时引用原文或译文，软件提供了如下变量：

{{SourceLanguage}}——原文语言

{{TargetLanguage}}——译文语言（目标语言）

{{SourceSentence}}——原文整句/整段的内容

{{TargetSentence}}——译文整句/整段的内容

{{TermList}} ——是术语变量，可以从用户术语库中检索到当前句子所使用的术语，并将它提交给 AI 大语言模型

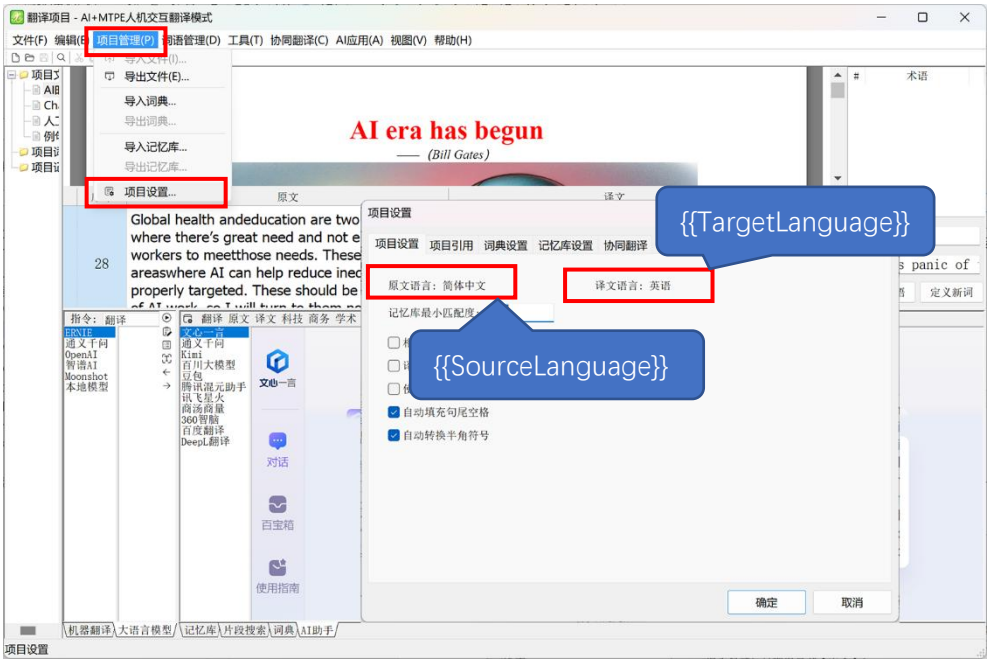
{{SourceSelection}}——原文中鼠标划选的内容

{{TargetSelection}}——译文中鼠标划选的内容

{{SourceLanguage:1}}——出现在提示语时，“原文语言”所指定的语种用中文显示，1 代表中文，不写序号默认用英文表示“原文语言”的语种。

{{TargetLanguage:1}}——出现在提示语时，“译文语言”所指定的语种用中文显示，1 代表中文，不写序号默认用英文表示“译文语言”的语种。用中文写提示语时，可以用带序号的变量，用英文写提示语时，不带序号。

{{SourceLanguage}}、{{TargetLanguage}}变量所指代的语言方向，在新建翻译项目时决定，点击**【项目管理】➔【项目设置】**菜单，在“项目设置”窗口中可以查看语言方向。如下图：



点击屏幕下方的“AI 助手”页签，在屏幕的中部是“自定义 AI 工具条”，默认有设置、翻译、原文、译文四个按钮，用户可以修改它或添加更多的按钮。如下图：



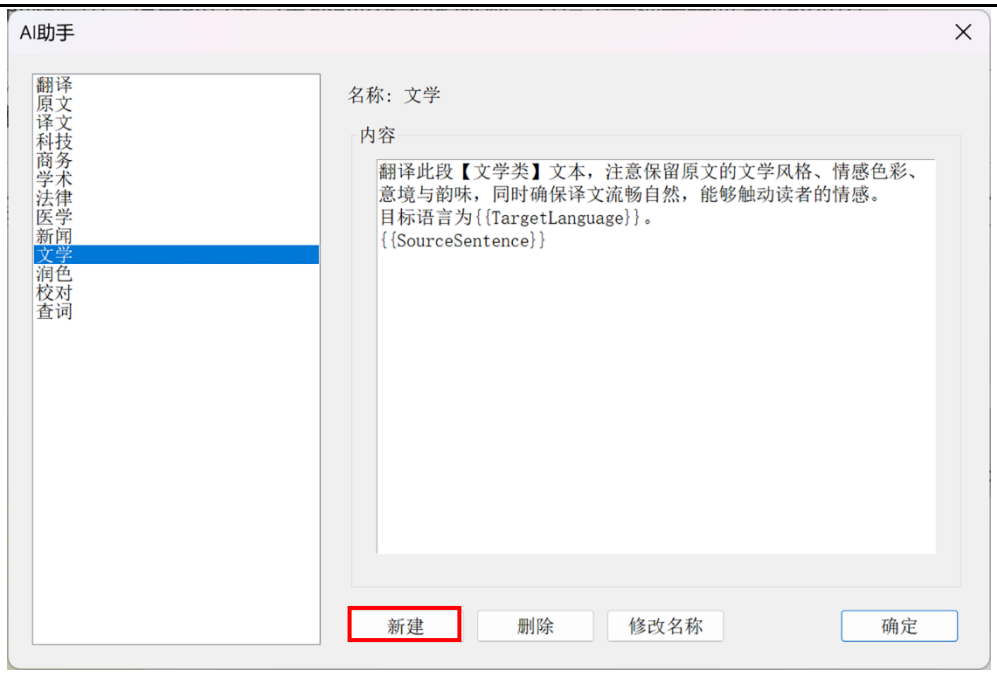
点击工具条最前面的设置按钮，会弹出“AI 助手”窗口，如下图：



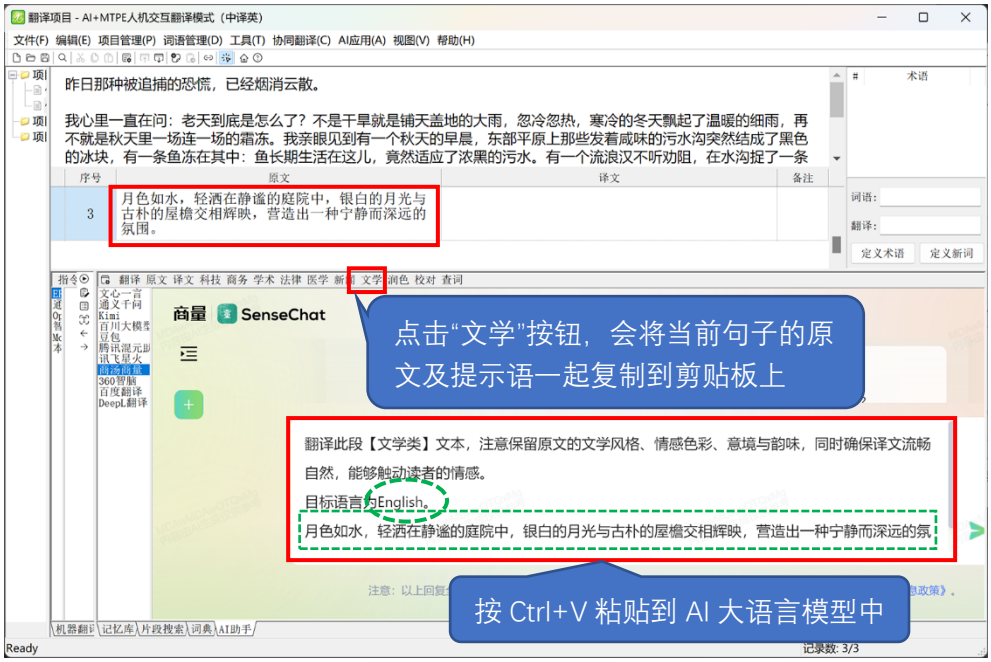
在 AI 助手窗口中，可以修改默认的按钮执行的功能或新建按钮，在窗口右侧点击**【右键菜单】➔【上移】**或**【下移】**功能，可以调整各个按钮的排列顺序。在右侧的内容输入框中，输入提示语（指令），要求 AI 做什么？提示语可以用中文写，也可以用英文写，如果是用 ChatGPT，用英文的提示语效果会比中文的好。

下面以“新建”文学翻译按钮为例，说明其具体运用。点击“AI 助手”下方的“新建”按钮（如下图），输入按钮名称“文学”二字，在右侧的“内容”中输入提示语：

翻译此段【文学类】文本，注意保留原文的文学风格、情感色彩、意境与韵味，同时确保译文流畅自然，能够触动读者的情感。
目标语言为{{TargetLanguage}}。
{{SourceSentence}}



设置好后，点击“确定”按钮，退出窗口。然后点击屏幕中部“工具条”的“文学”按钮，会将当前句子的原文及提示语一起复制到剪贴板上，按【Ctrl+V】直接粘贴到 AI 大预言模型输入框中，如下图:



若提示语变量用{{TargetLanguage:1}},则生成的提示语是：“目标语言为英语”。
若提示语变量用{{TargetLanguage}},则生成的提示语是：“目标语言为 English”

用变量的好处是一条提示语可以用在多个语言方向，若直接写明目标语种，则另一个语言方向就不合适用了。

点击大语言模型的“复制”按钮，再按【Ctrl+V】将译文粘贴到雪人 CAT 的译文区中，或者点击**【右键菜单】➔【Paste】**粘贴，或者选择译文后按住鼠标左键不松手，直接将大语言模型的译文拖到译文区中。如下图：

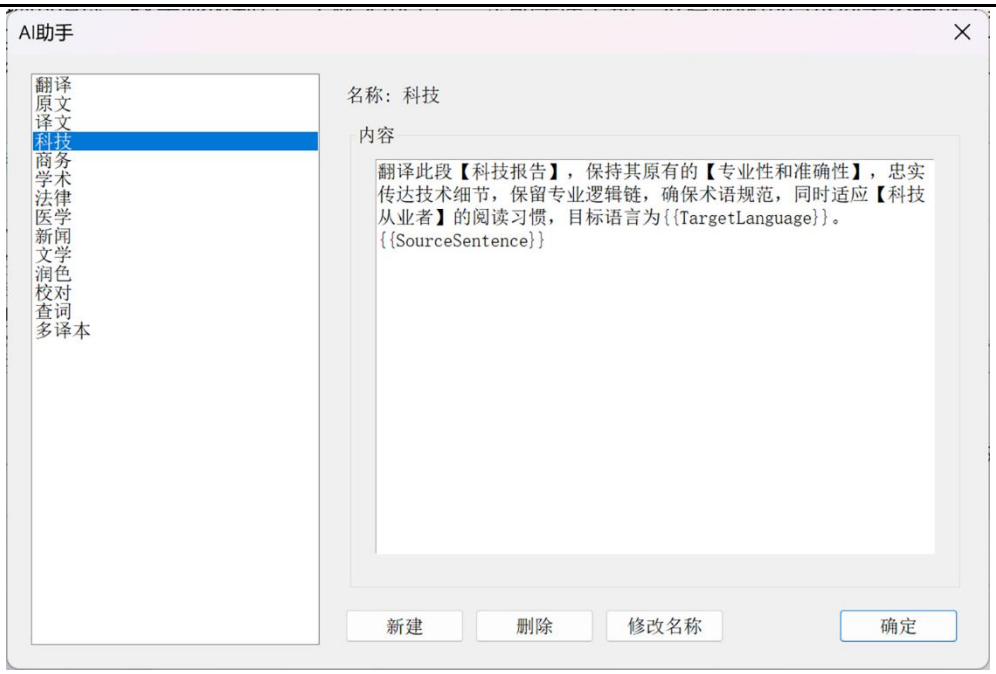


还可以将剪切板中的提示语及原文，粘贴到其他的几个大语言模型中，比较看看哪个模型翻译的好？

以下给出多种文本翻译的提示语供参考：

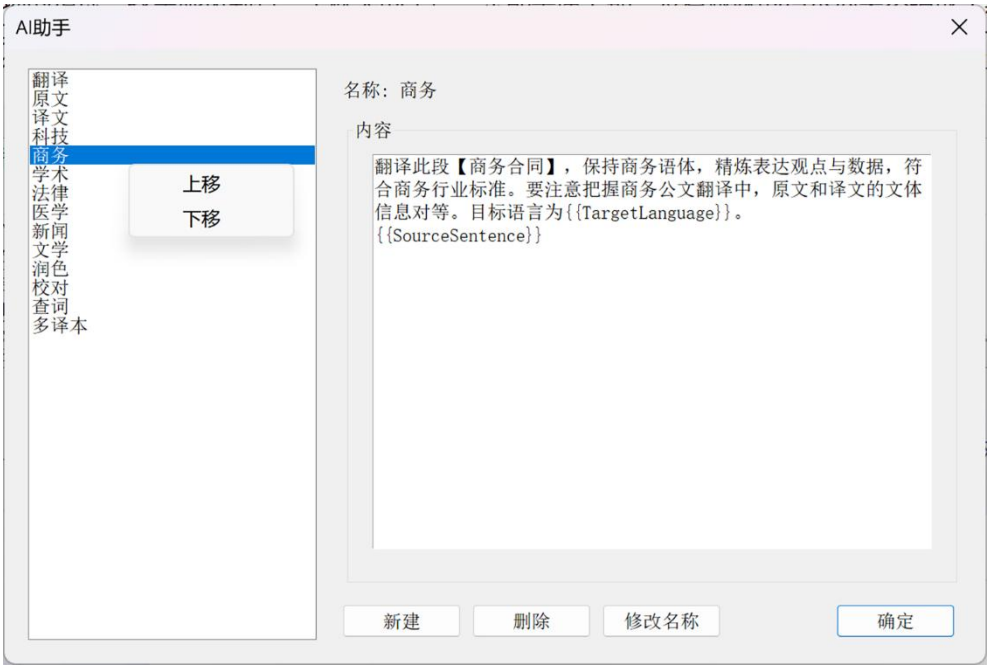
一、【科技文本】参考提示语：

翻译此段【科技报告】，保持其原有的【专业性和准确性】，忠实传达技术细节，保留专业逻辑链，确保术语规范，同时适应【科技从业者】的阅读习惯，目标语言为{{TargetLanguage}}。
{{SourceSentence}}



二、【商务文本】参考提示语：

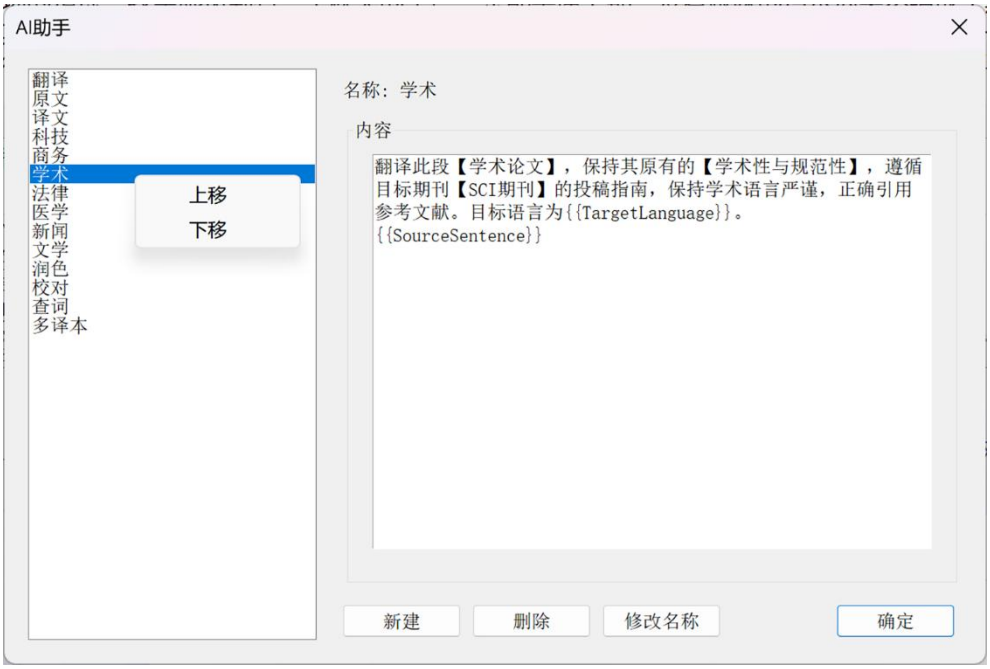
翻译此段【商务合同】，保持商务语体，精炼表达观点与数据，符合商务行业标准。要注意把握商务公文翻译中，原文和译文的文体信息对等。目标语言为{{TargetLanguage}}。
{{SourceSentence}}



三、【学术论文】参考提示语：

翻译此段【学术论文】，保持其原有的【学术性与规范性】，遵循目标期刊【SCI 期刊】的投稿指南，保持学术语言严谨，正确引用参考文献。目标语言为{{TargetLanguage}}。

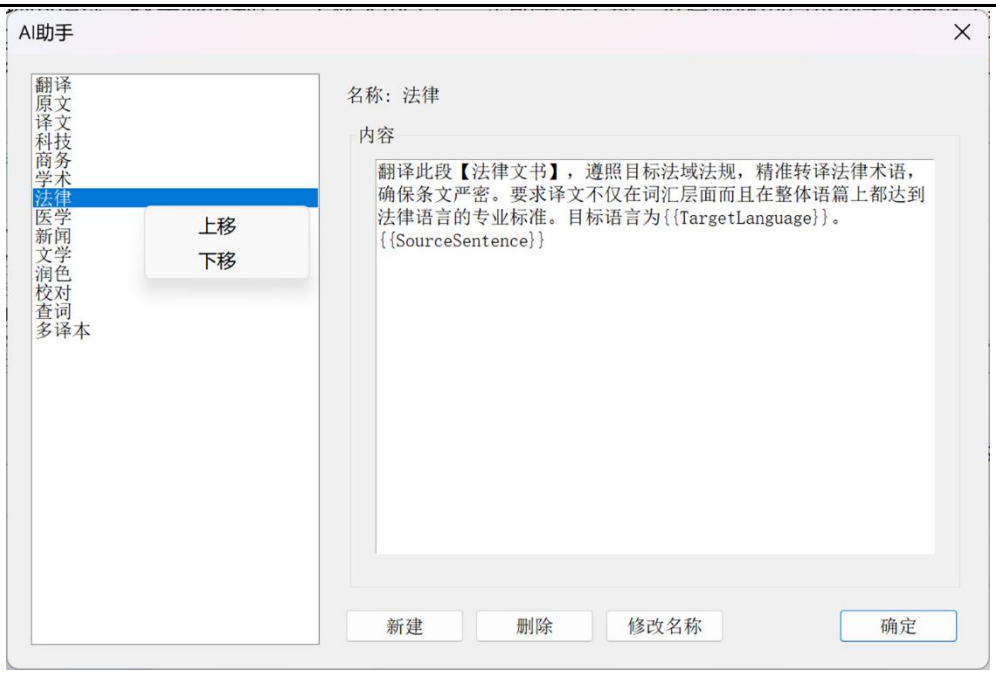
{{SourceSentence}}



四、【法律文本】参考提示语：

翻译此段【法律文书】，遵照目标法域法规，精准转译法律术语，确保条文严密。要求译文不仅在词汇层面而且在整体语篇上都达到法律语言的专业标准。目标语言为{{TargetLanguage}}。

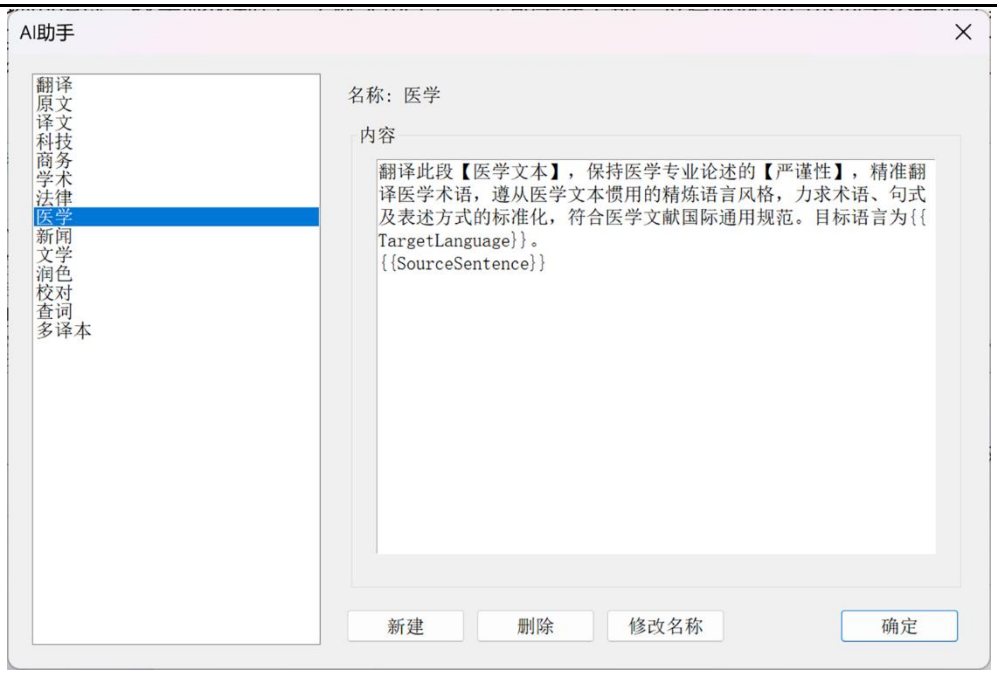
{{SourceSentence}}



五、【医学文本】参考提示语：

翻译此段【医学文本】，保持医学专业论述的【严谨性】，精准翻译医学术语，遵从医学文本惯用的精炼语言风格，力求术语、句式及表述方式的标准化，符合医学文献国际通用规范。目标语言为{{TargetLanguage}}。

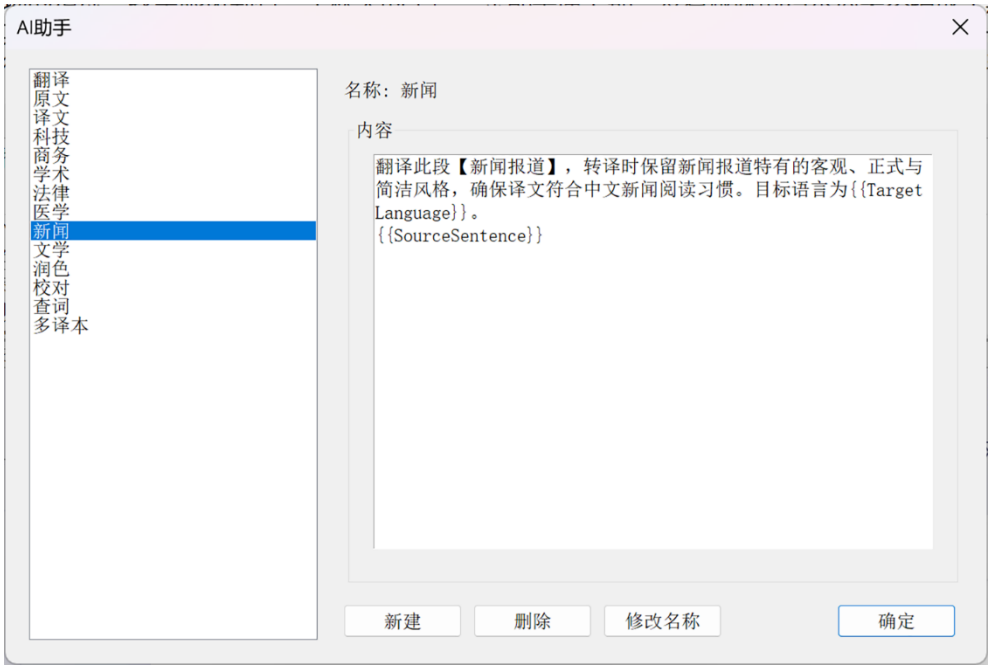
{{SourceSentence}}



六、【新闻报道】参考提示语：

翻译此段【新闻报道】，转译时保留新闻报道特有的客观、正式与简洁风格，确保译文符合中文新闻阅读习惯。目标语言为{{TargetLanguage}}。

{{SourceSentence}}

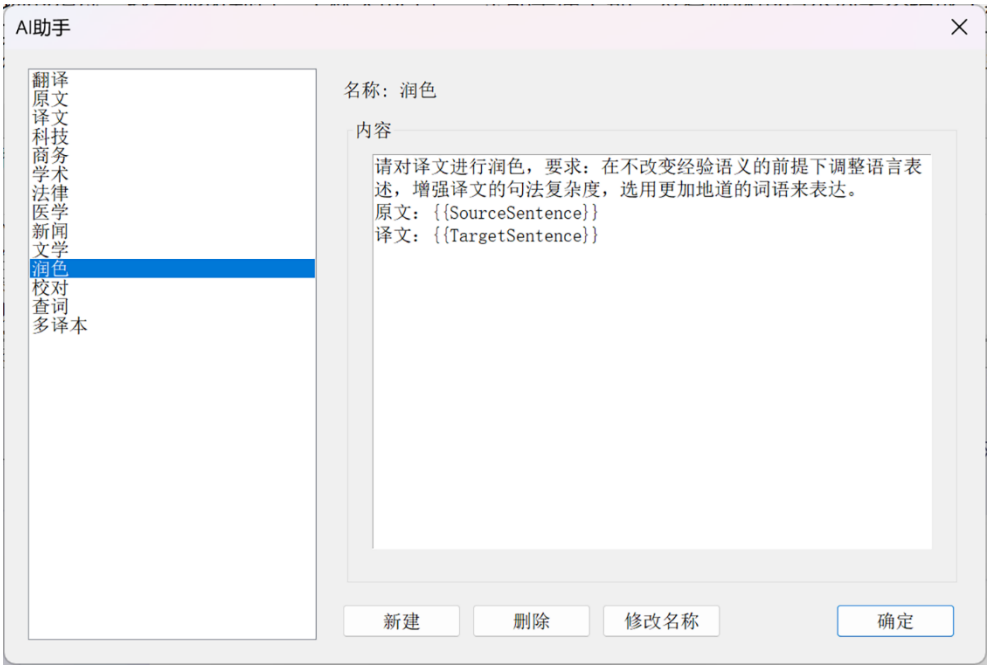


七、【文本润色】参考提示语：

请对译文进行润色，要求：在不改变经验语义的前提下调整语言表述，增强译文的句法复杂度，选用更加地道的词语来表达。

原文：{{SourceSentence}}

译文：{{TargetSentence}}

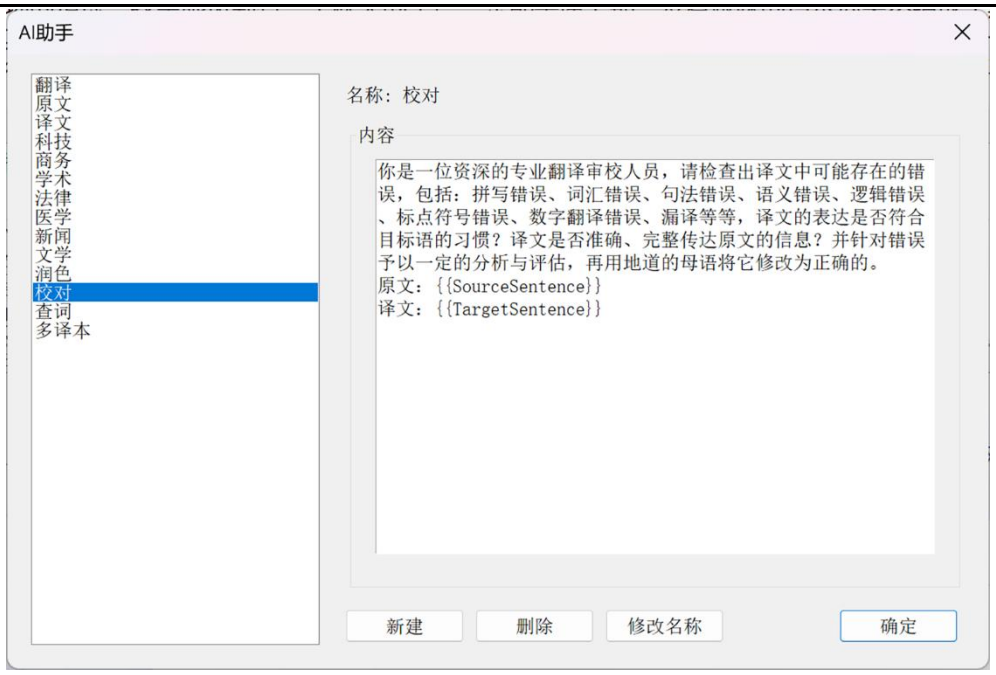


八、【译文校对】参考提示语：

你是一位资深的专业翻译审校人员，请检查出译文中可能存在的错误，包括：拼写错误、词汇错误、句法错误、语义错误、逻辑错误、标点符号错误、数字翻译错误、漏译等等，译文的表达是否符合目标语的习惯？译文是否准确、完整传达原文的信息？并针对错误予以一定的分析与评估，再用地道的母语将它修改为正确的。

原文：{{SourceSentence}}

译文：{{TargetSentence}}



写好校对的提示语后，退出“AI 助手”窗口。



然后点击工具条的“校对”按钮，会将当前句的原文、译文及提示语一起复制到剪贴板上，具体内容如下图：

```
你是一位资深的专业翻译审校人员，请检查出译文中可能存在的错误，包括：拼写错误、词汇错误、句法错误、语义错误、逻辑错误、标点符号错误、数字翻译错误、漏译等等，译文的表达是否符合目标语的习惯？译文是否准确、完整传达原文的信息？并针对错误予以一定的分析与评估，再用地道的母语将它修改为正确的。
原文：我心里一直在问：老天到底是怎么了？不是干旱就是铺天盖地的大雨，忽冷忽热，寒冷的冬天飘起了温暖的细雨，再不就是秋天里一场连一场的霜冻。我亲眼见到有一个秋天的早晨，东部平原上那些发着咸味的污水沟突然结成了黑色的冰块，有一条鱼冻在其中：鱼长期生活在这儿，竟然适应了浓黑的污水。有一个流浪汉不听劝阻，在水沟捉了一条鱼烧了吃，结果肚子疼得打滚。不仅沟渠里的鱼不能吃，就连大河里的鱼吃了也要出事。不知多少人因为吃了有毛病的鱼给拉到医院里抢救，几乎每年都有人死于受污染的鱼。
译文：I always ask myself: what's wrong with the weather? It's either be in drought or rains cats and dogs. It sometimes cold and turns into hot suddenly. Warm rain happens in the cold winter or frosts happen one after another. I once saw it with my own eyes in a autumn morning when salty smelling water in the drainage suddenly frozed and became dark with a fish inside the ice. Those fish have lived for long periods has adapted the black water unexpectedly. A tramp had been deaf to advice caught a fish and cooked, resulting from a dreadful stomach. Not only the fish in the drainage cannot be eaten, but also the fish in the wide river. There's countless people got into the emergency room because of the polluted fish or died from them almost every year.
```

然后按【Ctrl+V】将剪贴板的内容粘贴到 AI 大语言模型中，就会得到 AI 对这段译文的校对结果。如下图：下图的校对结果来自“通义千问”大模型。

译文修订：

I constantly ask myself: What on earth is happening with the weather? It oscillates between prolonged droughts and relentless downpours, with sudden fluctuations from chilly to warm. In the midst of a cold winter, there may be unexpected spells of gentle warmth accompanied by fine rain, or during autumn, a series of consecutive frosts may occur. I vividly recall one autumn morning when the briny water in the drainage ditches on the eastern plains abruptly froze into thick, black ice, encapsulating a fish within. Astonishingly, these fish, having long inhabited such conditions, have adapted to the pitch-black effluent. Despite warnings, a vagrant caught one of these fish, cooked it, and suffered excruciating stomach pain as a result. Not only are the fish in these drainage channels inedible, but consumption of those from even large rivers can prove hazardous. Countless individuals have been rushed to emergency rooms due to consuming contaminated fish, with fatalities occurring almost annually.

来自“通义千问”大模型的校对结果

错误分析与评估：

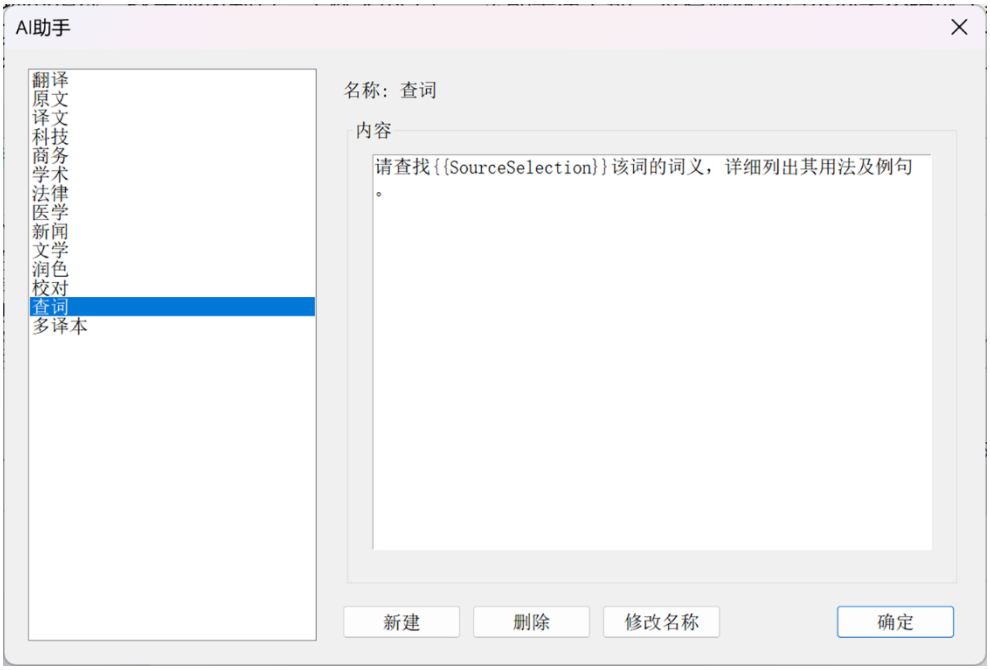
1. **词汇错误**：原文中的“忽冷忽热”被译为“sometimes cold and turns into hot suddenly”，虽然大致传达了意思，但不够地道。修订为“with sudden fluctuations from chilly to warm”更符合英语表达习惯。
2. **句法错误**：译文中“Warm rain happens in the cold winter or frosts happen one after another.” 的表述不够流畅。修订为 “In the midst of a cold winter, there may be unexpected spells of gentle warmth accompanied by fine rain, or during autumn, a series of consecutive frosts may occur.” 使句式更为紧凑且清晰。
3. **语义错误**：“It's either be in drought or rains cats and dogs” 中的 “be” 应删除，修正为 “It's either drought or rains cats and dogs”。
4. **逻辑错误**：“Not only the fish in the drainage cannot be eaten, but also the fish in the wide river.” 逻辑连接词位置不当，应改为 “Not only are the fish in these drainage channels inedible, but consumption of those from even large rivers can prove hazardous.”
5. **漏译**：“鱼长期生活在这儿，竟然适应了浓黑的污水。” 中的 “竟然” 表达出意外之意，修订为 “astonishingly, these fish, having long inhabited such conditions, have adapted to the pitch-black effluent” 更贴切地传达原文意思。
6. **标点符号错误**：原文中的冒号”：“在译文中未得到恰当处理。修订后的译文使用逗号和句号来正确断句，确保语义连贯。

综上所述，通过对译文的审校与修订，已纠正了其中的拼写、词汇、句法、语义、逻辑、标点符号错误，以及漏译问题，确保译文准确、完整地传达原文信息，且符合英语表达习惯。

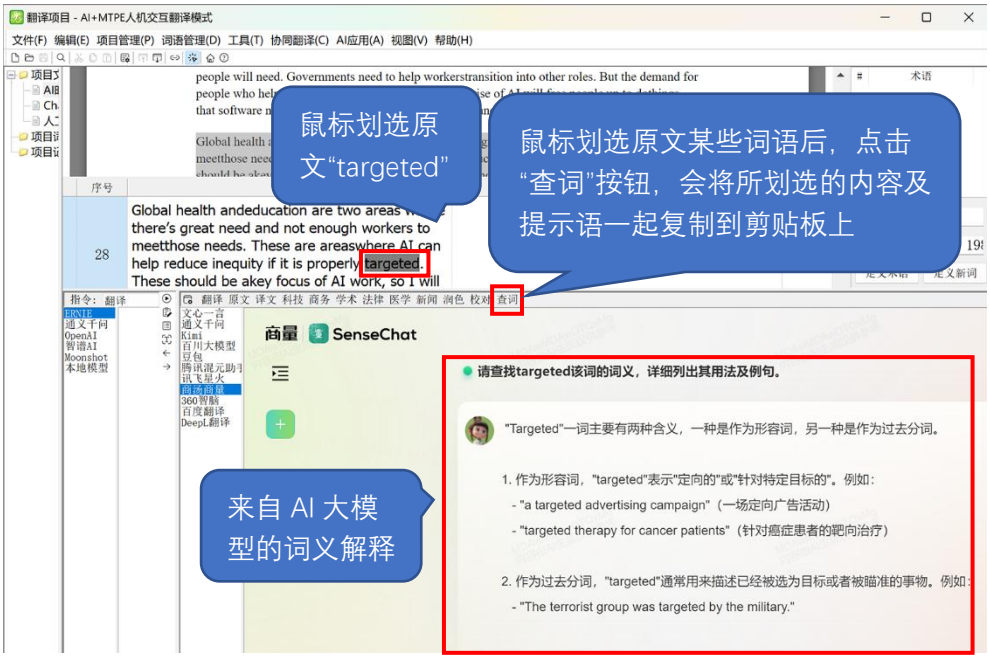
如果觉得 AI 校对的不错，可以将其修订的内容复制、粘贴回雪人 CAT 软件中。为了保留自己原始的版本，建议校对前，在屏幕左侧的“项目文件”下面，点击某文件后，按住鼠标左键拖动，将文件复制一份。

九、【查词】参考提示语：

请查找{{SourceSelection}}该词的词义，详细列出其用法及例句。



鼠标划选原文某些词语后，点击“查词”按钮，会将所划选的内容及提示语一起复制到剪贴板上，然后按【Ctrl+V】将剪贴板的内容粘贴到 AI 大语言模型中查询。如下图:



十、【多译本】参考提示语：

翻译此段文本。输出以下多种译文风格：【正式风格】、【诗意风格】、【幽默风格】、【庄重风格】、【讽刺风格】。

目标语言为{{TargetLanguage}}。

{{SourceSentence}}。



可以通过提示语，让 AI 输出多个不同风格译本的译文供参考，如下图：



翻译此段文本。输出以下多种译文风格：【正式风格】、【诗意风格】、【幽默风格】、【庄重风格】、【讽刺风格】。

目标语言为English。

昨日那种被追捕的恐慌，已经烟消云散。

让 AI 生成多个不同风格的译文版本。

【正式风格】

The panic of being pursued yesterday has vanished into thin air.

【诗意风格】

Yesterday's panic of pursuit has dissolved into the ether, leaving only tranquility in its wake.

【幽默风格】

The panic of being chased yesterday? Nah, it's like a bad dream that's already fading from memory!

【庄重风格】

The dread of yesterday's pursuit has dissipated, leaving behind a sense of peace and tranquility.

【讽刺风格】

Yesterday's panic of being chased? Hah, that's just a thing of the past now. Guess you're not as wanted as you thought, eh?

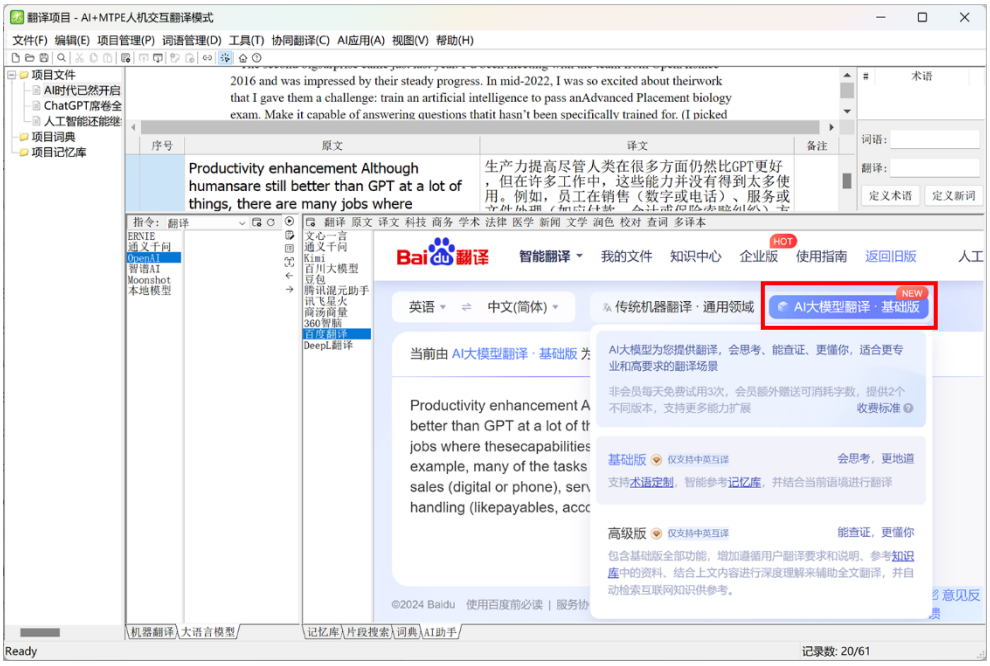
由文心大模型 3.5 生成

十、【调用百度翻译】：

支持百度翻译网页调用，网页上可以用传统的机器翻译，根据所翻译的内容选择所属领域，其中有：通用领域、学术论文、生物医药、信息技术、金融财经、新闻资讯、航空航天、法律法规、机械制造、人文社科多个领域供选择。网页传统的机器翻译仅提供翻译结果，不能交互、按要求翻译、问答等。操作方法：点击“原文”按钮，将原文复制到剪切板，再按【Ctrl+V】粘贴到百度中，然后再将译文复制粘贴回来。如下图：

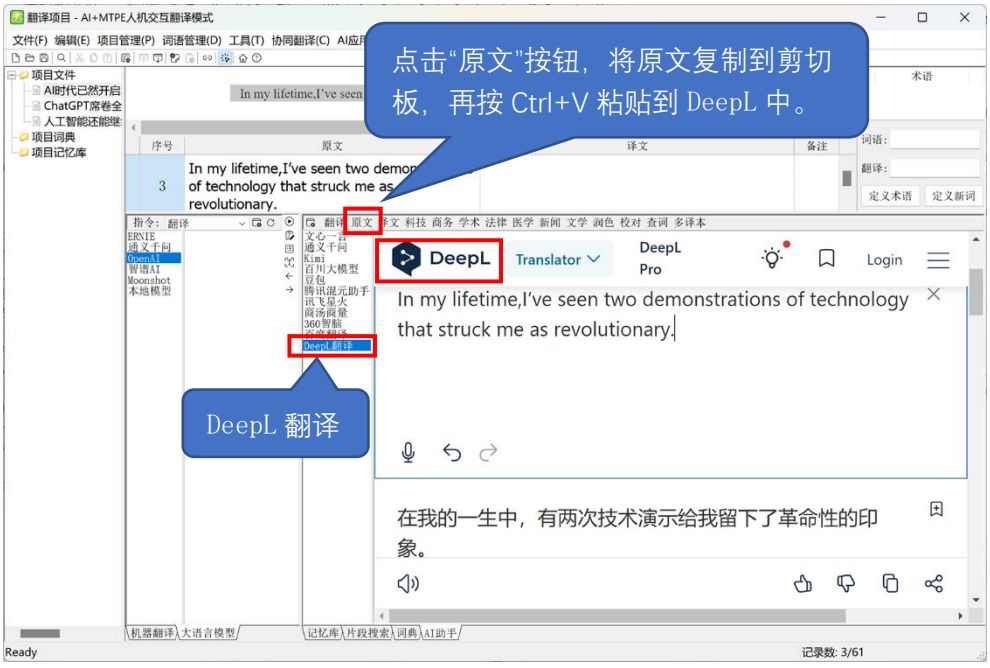


也可以使用百度的 AI 大模型翻译，如下图：用 AI 大模型翻译则需要注册。



十、【调用 DeepL 翻译】：

支持使用 DeepL 网页进行翻译。操作方法：点击“原文”按钮，将原文复制到剪切板，再按【Ctrl+V】粘贴到 DeepL 网页中，然后再将译文复制粘贴回来。如下图：

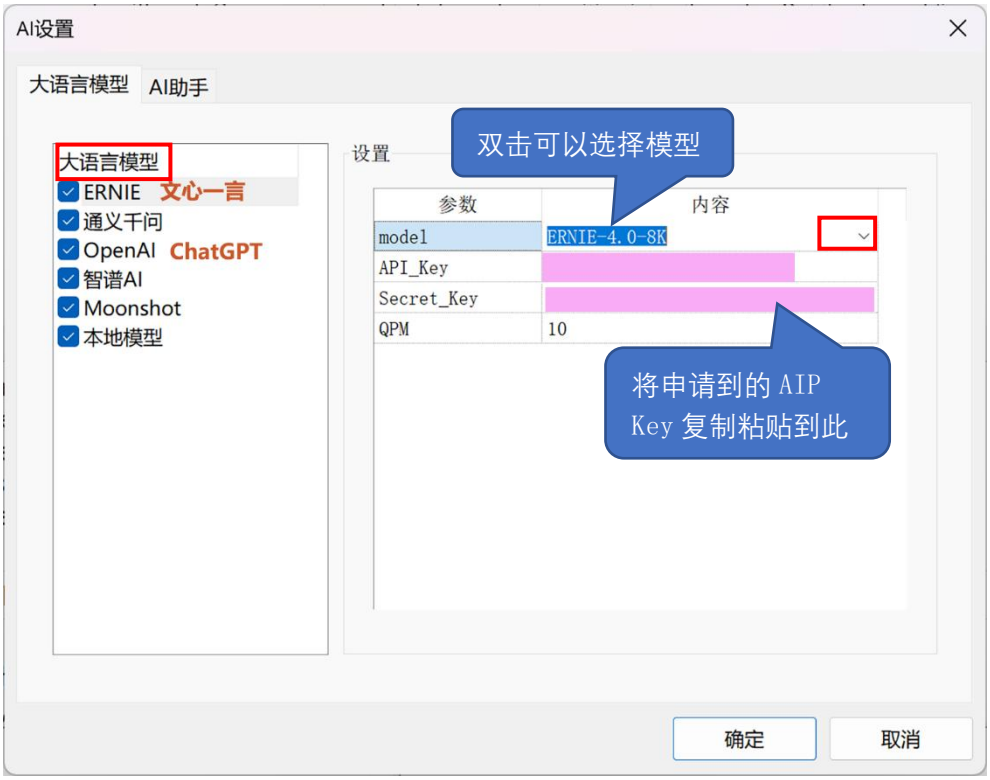


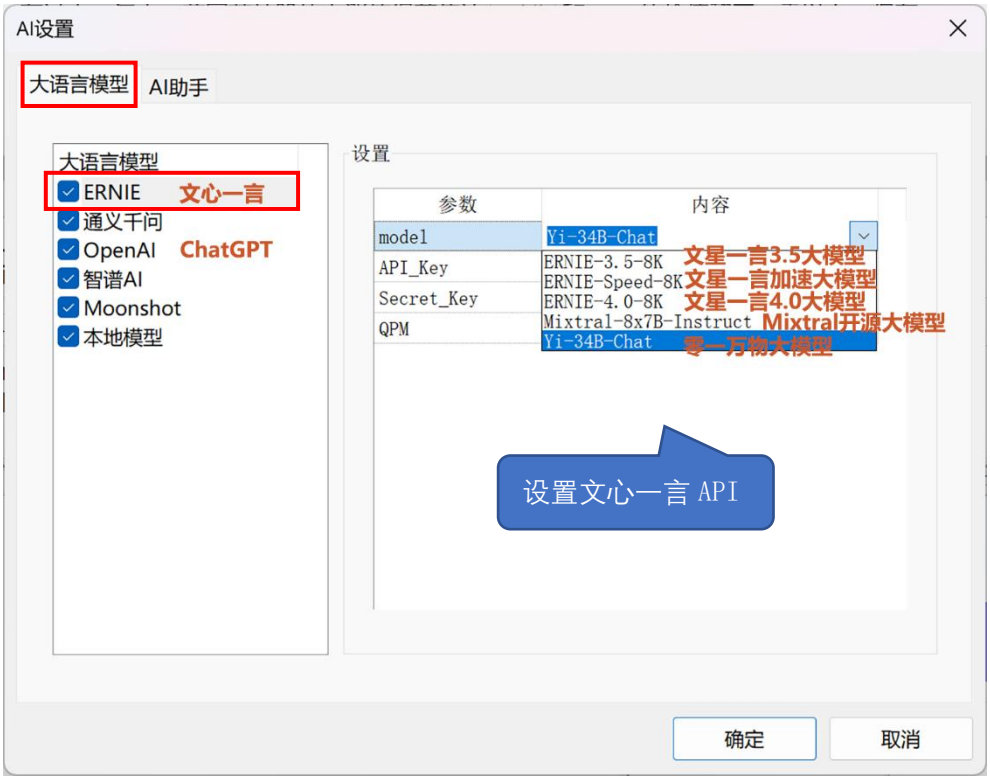
第五节 AI 大模型的 API 调用

上节讲述的是 AI 大模型的网页调用方法，本节将讲述 AI 大模型 API 接口调用方法，API 方式会比网页方式高效，且界面整洁。

点击软件主菜单**【AI 应用】➔【AI 设置】**功能，在“AI 设置”窗口中，点击“大语言模型”页签，可以看到软件支持的多个 AI 大语言模型，有 ERNIE(文心一言)、通义千问、OpenAI(ChatGPT)、智谱 AI、Moonshot (Kimi)。还支持用户在本地机器上部署 AI 模型。

要想用 API 方式使用各个 AI 大语言模型，首先要去各个大语言模型网站申请 API，将申请到的 API Key 复制粘贴到雪人 CAT 的**【AI 应用】➔【AI 设置】➔【大语言模型】**窗口中，如下图：如何申请 API？请参看本章的“第六节 AI 大模型 API 的申请”。



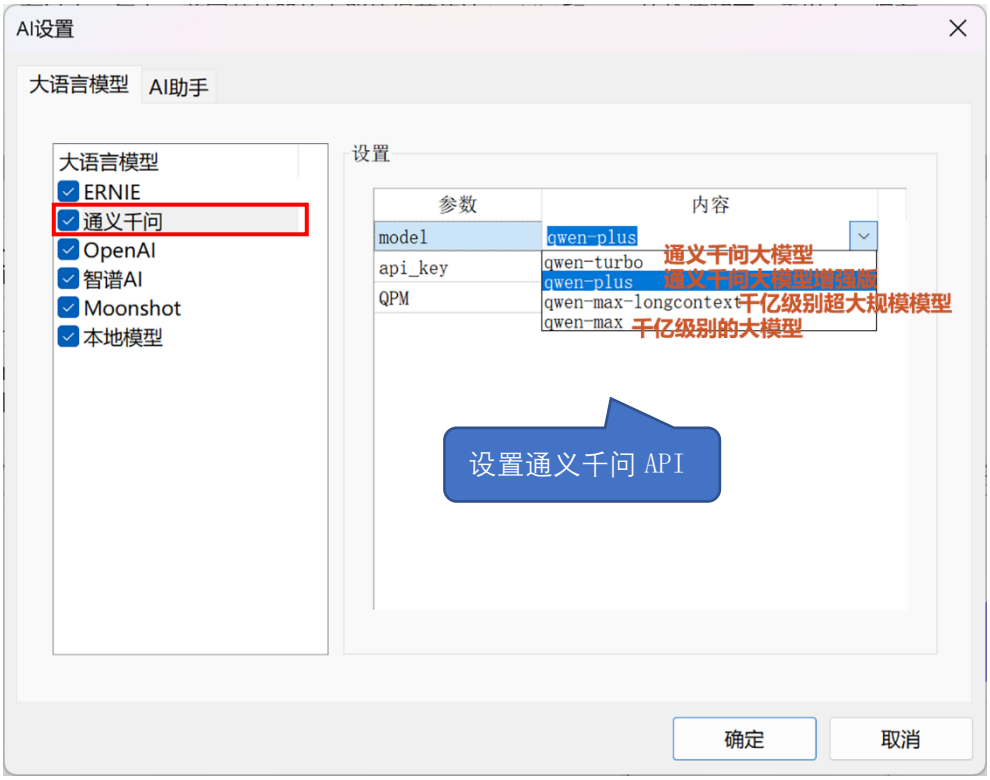


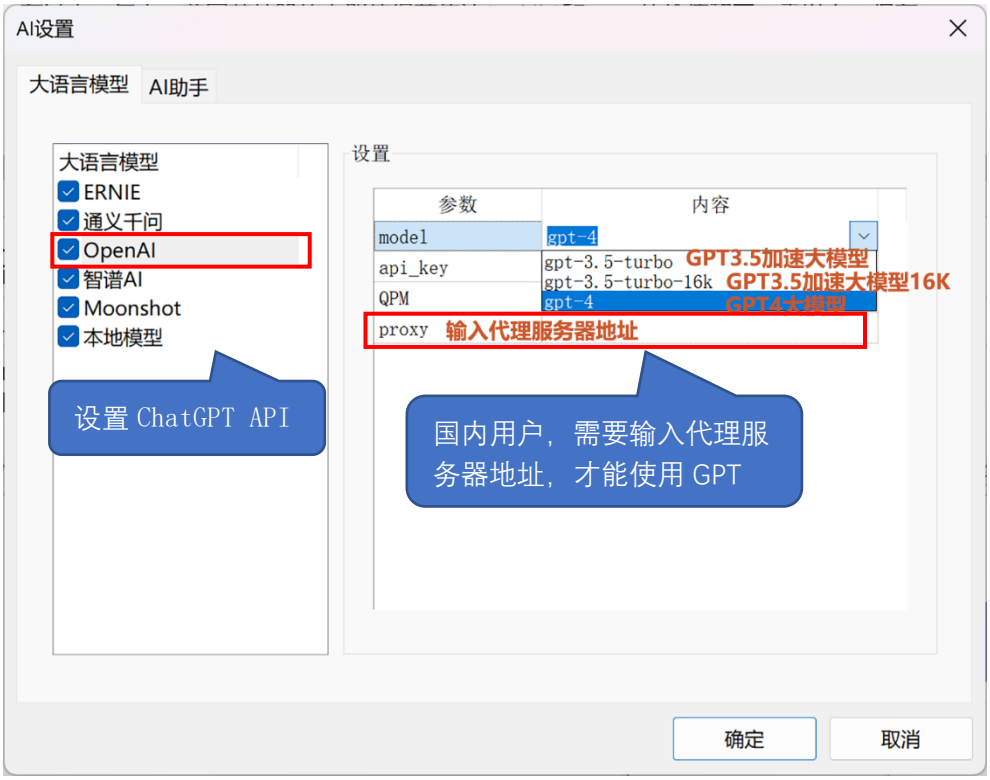
双击“model”那一行，光标会变成可编辑状态，点击旁边的下拉按钮，可以选择不同的 AI 大语言模型。上图中，列出了“文心一言”API 支持的多个大语言模型，不同的模型，收费的标准是不一样的，请仔细查看各个模型的相关收费说明，根据需要选择合适自己使用的模型。其中：Yi-34B-Chat （零一万物）模型的 API 目前是可以免费使用的。

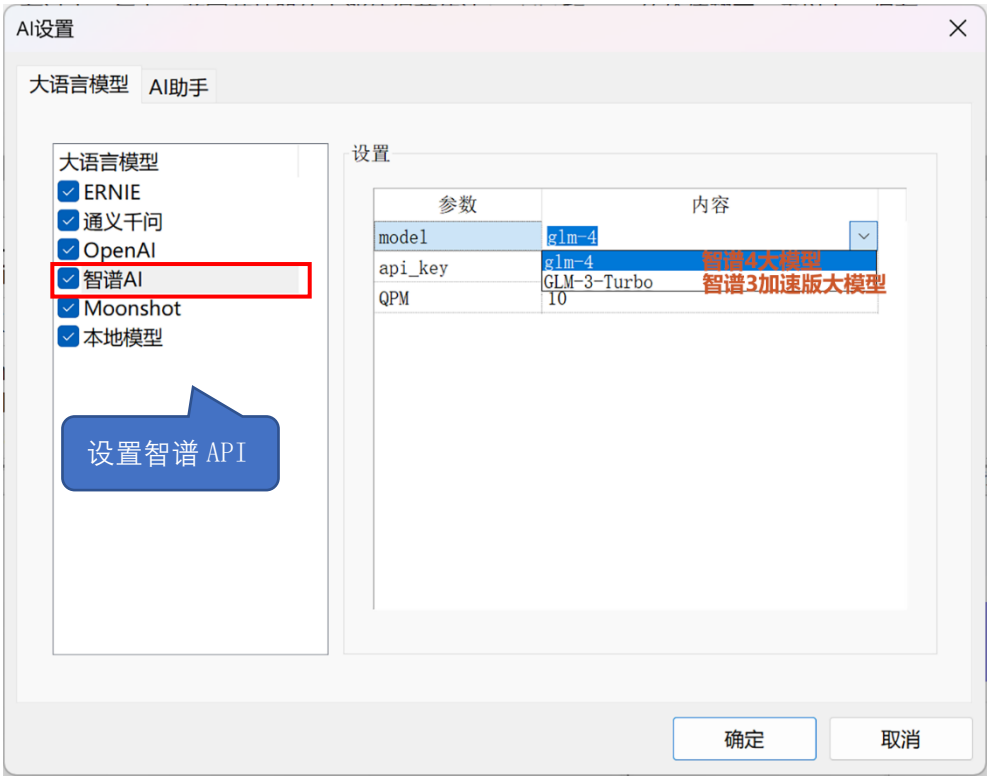
QPM——是允许每分钟访问大模型的次数，软件设置了默认的数值，用户可以修改，但是如果修改过大，繁忙时可能会造成出错，无结果返回。

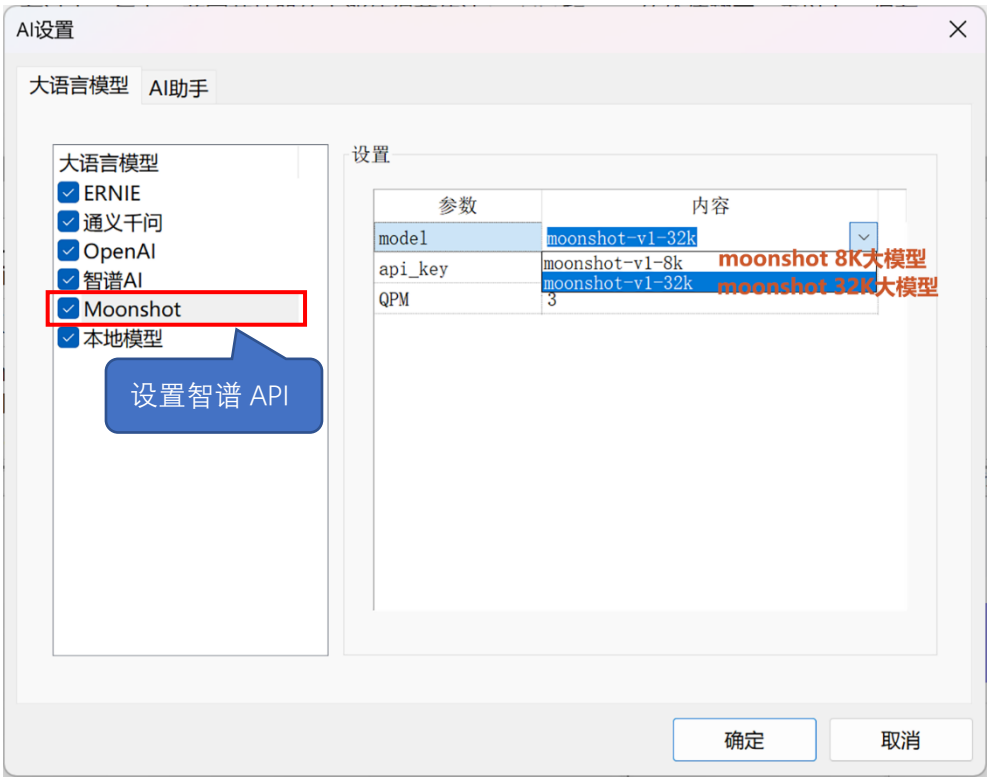
同理，设置好通义千问、OpenAI(ChatGPT)、智谱 AI、Moonshot 几个 AI 大语言模型，将申请到的 API Key 复制粘贴到各个模型相应的位置上。繁忙时段，由于算力有限，有些大模型调用时可能存在一定时间的响应等待。

注意：国内用户需要输入代理服务器地址，才能使用 ChatGPT。ChatGPT-4 的收费标准比 ChatGPT-3.5 贵不少，请根据需要选用。

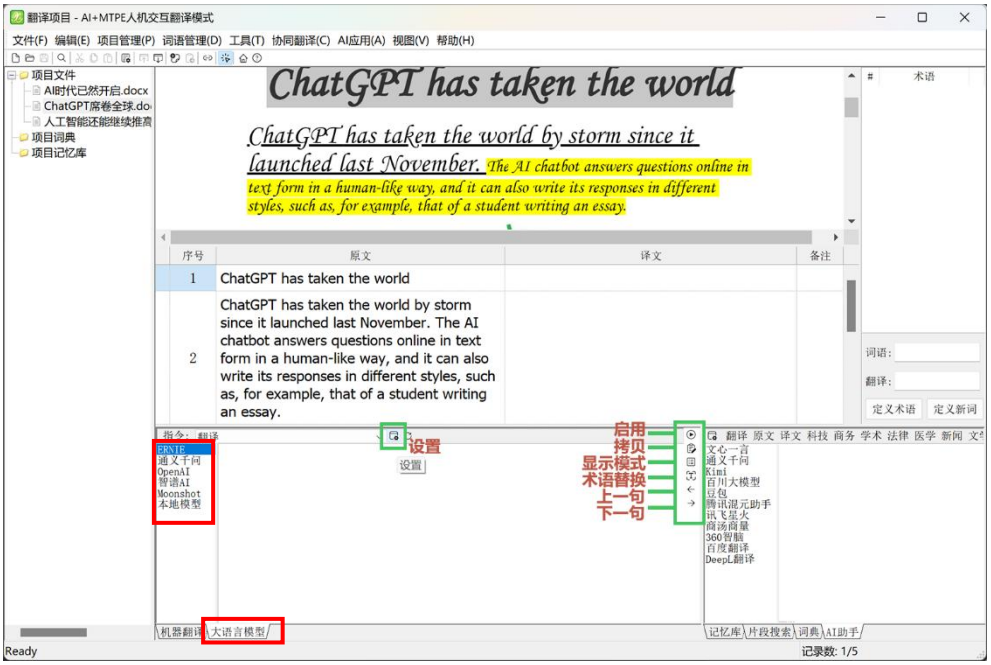

AI设置
大语言模型
AI助手
大语言模型
ERNIE
通义千问
OpenAI
智谱AI
Moonshot
本地模型
设置
参数
内容
model
qwen-plus
api_key
QPM
qwen-turbo
qwen-plus
qwen-max-longcontext
qwen-max
通义千问大模型
通义千问大模型增强版
千亿级别超大规模模型
千亿级别的大模型
设置通义千问 API
确定
取消


AI设置
大语言模型
AI助手
大语言模型
ERNIE
通义千问
OpenAI
智谱AI
Moonshot
本地模型
设置
参数
内容
model
gpt-4
api_key
QPM
proxy
gpt-3.5-turbo
gpt-3.5-turbo-16k
gpt-4
GPT3.5加速大模型
GPT3.5加速大模型16K
GPT4大模型
输入代理服务器地址
设置 ChatGPT API
国内用户，需要输入代理服务器地址，才能使用 GPT
确定
取消


AI设置
大语言模型
AI助手
大语言模型
ERNIE
通义千问
OpenAI
智谱AI
Moonshot
本地模型
设置智谱 API
设置
参数
内容
model
glm-4
api_key
QPM
glm-4
GLM-3-Turbo
10
智谱4大模型
智谱3加速版大模型
确定
取消


AI设置
大语言模型
AI助手
大语言模型
ERNIE
通义千问
OpenAI
智谱AI
Moonshot
本地模型
设置智谱 API
设置
参数
内容
model
moonshot-v1-32k
api_key
QPM
moonshot-v1-8k
moonshot-v1-32k
3
moonshot 8K大模型
moonshot 32K大模型
确定
取消

设置好 AI 大语言模型的 API 后，点击屏幕下方的“大语言模型”页签，在“大语言模型”窗口的左侧是雪人支持的各个大语言模型，有：ERNIE(文心一言)、通义千问、OpenAI(ChatGPT)、智谱 AI、Moonshot 和本地模型，点击它即可以实现切换，右侧是工具图标，如下图：



各图标解释如下：

——设置 AI 的提示语和执行操作的指令。

——启用 AI 大语言模型

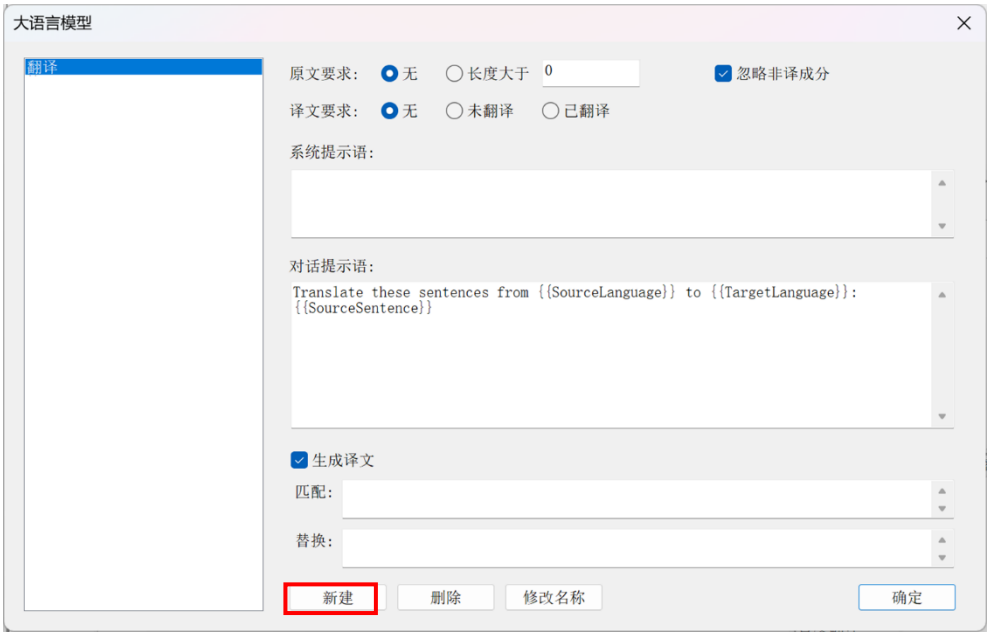
——将译文拷贝到上方的译文输入框中

——显示模式，可以显示单个 AI 大语言模型的结果，也可以将多个模型结果并列显示，方便进行比较，点击该按钮可以在这两种显示模式之间进行切换。

——翻译/校对/润色 上一句

——翻译/校对/润色 下一句

点击“大语言模型”窗口上方的“设置”图标，弹出“大语言模型”窗口，默认仅有一项简单的“翻译”指令，如下图：



用户可以自己“新建”各种 AI 指令。下面以“商务文本翻译”为例，说明其使用方法：

点击“新建”按钮，新建“商务文本翻译”指令，在“系统提示语”和“对话提示语”中分别输入以下内容：

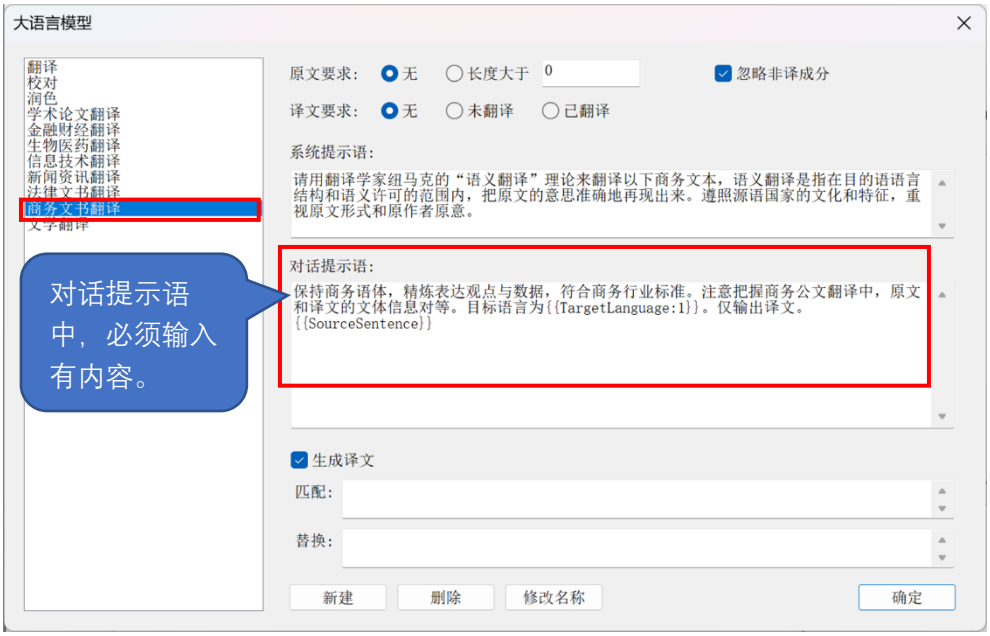
系统提示语：

请用翻译学家纽马克的“语义翻译”理论来翻译以下商务文本，语义翻译是指在目的语语言结构和语义许可的范围内，把原文的意思准确地再现出来。遵照源语国家的文化和特征，重视原文形式和原作者原意。

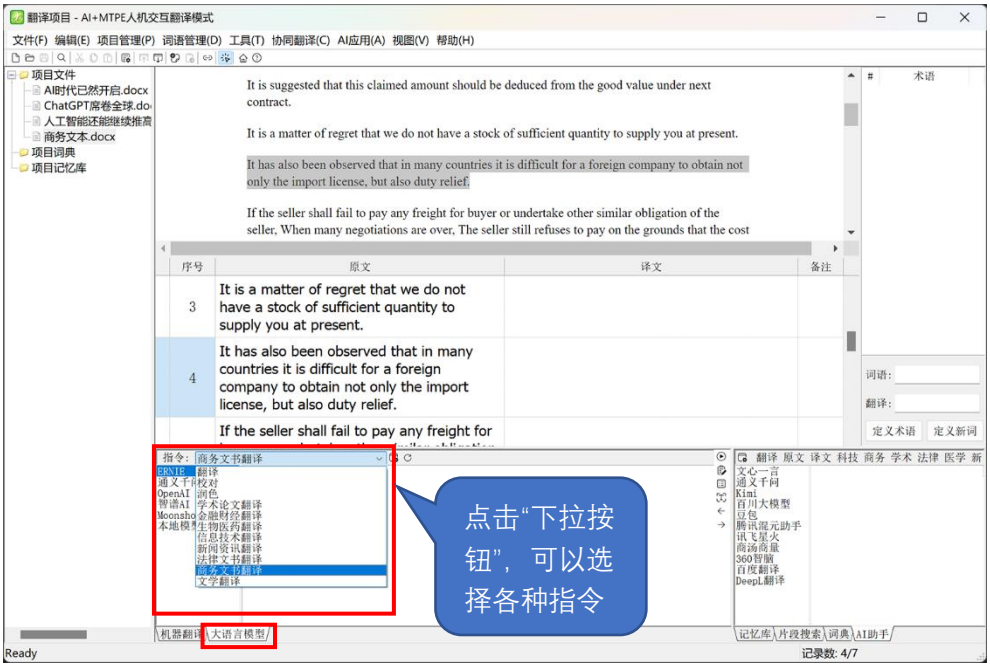
对话提示语：

保持商务语体，精炼表达观点与数据，符合商务行业标准。注意把握商务公文翻译中，原文和译文的文体信息对等。目标语言为{{TargetLanguage:1}}。仅输出译文。

{{SourceSentence}}

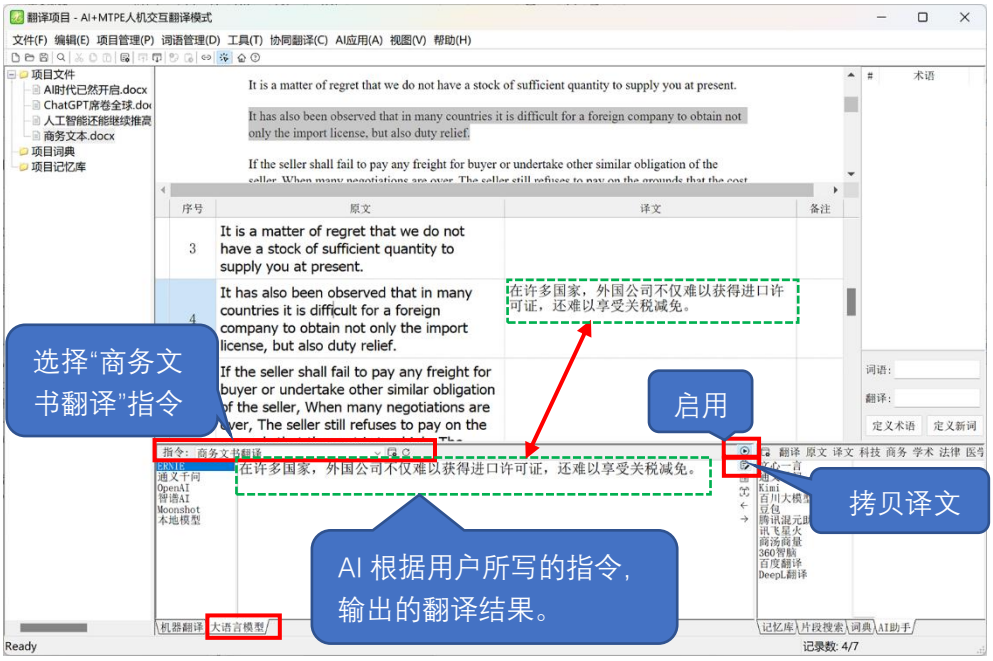


对话提示语中，必须输入有内容，不能为空，否则该指令新建失败。有些大模型区分“系统提示语”和“对话提示语”，有些模型不区分，可以全部在“对话提示语”中输入。“原文要求”、“译文要求”是设置是否发送给大模型的条件。“忽略非译成分”选项若勾选了，对于整句都是“非译成分”的内容，则不发送给大模型。



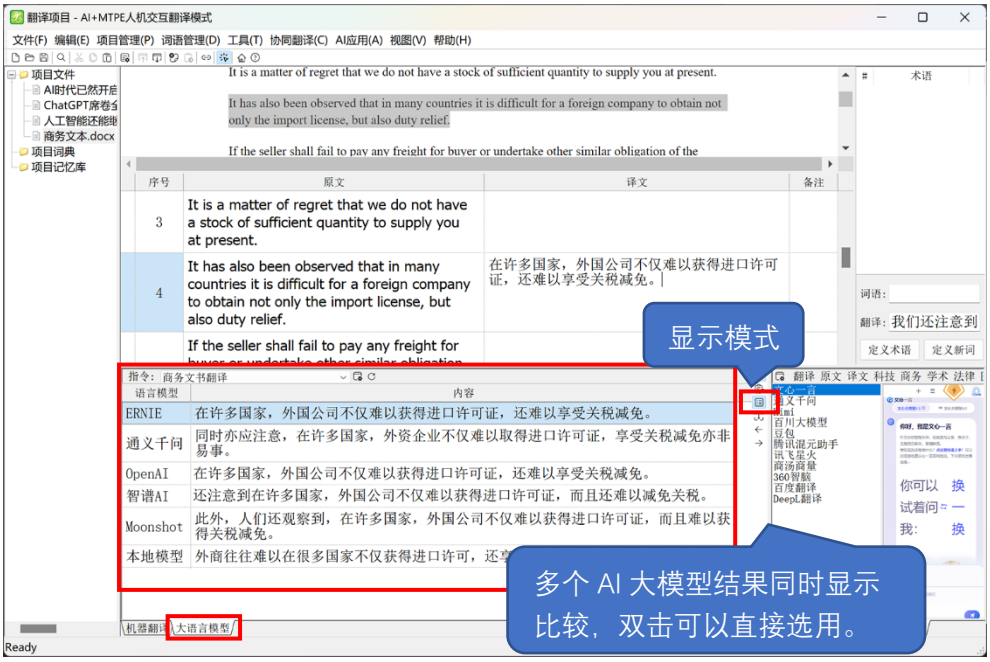
点击上图的“指令”右侧的下拉按钮，可以选择自己创建的各种 AI 指令，例如有：校对、润色、学术论文翻译、金融财经翻译、生物医药翻译、信息技术翻译、新闻资讯翻译、法律文书翻译、商务文书翻译、文学翻译等多种 AI 指令。

这里选择“商务文书翻译”指令，然后点击右侧的启用按钮，AI 就会根据用户所写的指令输出翻译结果，如果用户在指令中的提示语写得好，AI 输出的结果就会更符合预期。



点击右侧的拷贝按钮，可以将翻译结果拷贝到上方的译文区中。当光标在上方的译文区时，按【F7】也可以将翻译结果拷贝到译文区中。

点击窗口左侧的其他大语言模型，可以切换查看其他模型的翻译结果。点击右侧的“显示模式”按钮，可以将多个模型结果并列显示，方便进行比较，如下图：同时列出了：ERNIE(文心一言)、通义千问、OpenAI(ChatGPT)、智谱 AI、Moonshot、本地模型的结果以便比较，择优选用，双击某模型的结果，可以直接将它粘贴到上方的译文区中。



再举一例“法律文书翻译”，输入如下提示语：

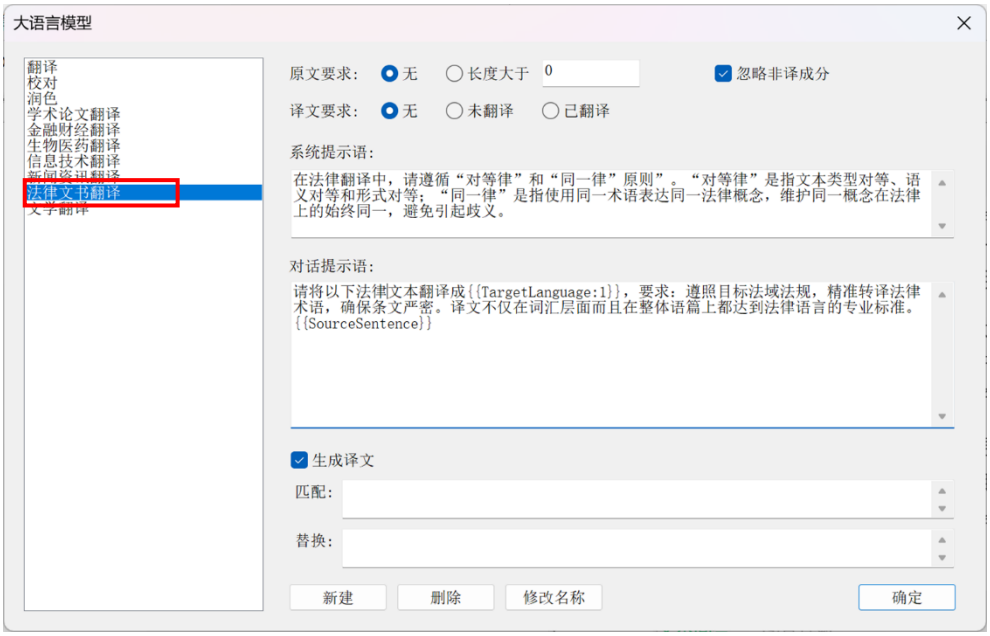
系统提示语：

在法律翻译中，请遵循“对等律”和“同一律”原则”。“对等律”是指文本类型对等、语义对等和形式对等；“同一律”是指使用同一术语表达同一法律概念，维护同一概念在法律上的始终同一，避免引起歧义。

对话提示语：

请将以下法律文本翻译成{{TargetLanguage:1}}，要求：遵照目标法域法规，精准转译法律术语，确保条文严密。译文不仅在词汇层面而且在整体语篇上都达到法律语言的专业标准。

{{SourceSentence}}



设置好提示语后，点击“确定”按钮退出窗口，选择“法律文书翻译”指令，“通义千问”大模型翻译结果，如下图：



点击右侧的拷贝按钮，可以将翻译结果拷贝到上方的译文区中。然后点击“下一句”按钮，AI 大模型马上显示出下一句的翻译结果，后续翻译可以不去理会提示语，由雪人软件自动处理，非常方便快捷，API 方式调用会比网页方式调用

高效，不需要像网页调用那样句句需要复制粘贴原文及提示语到 AI 大模型中。如下图：

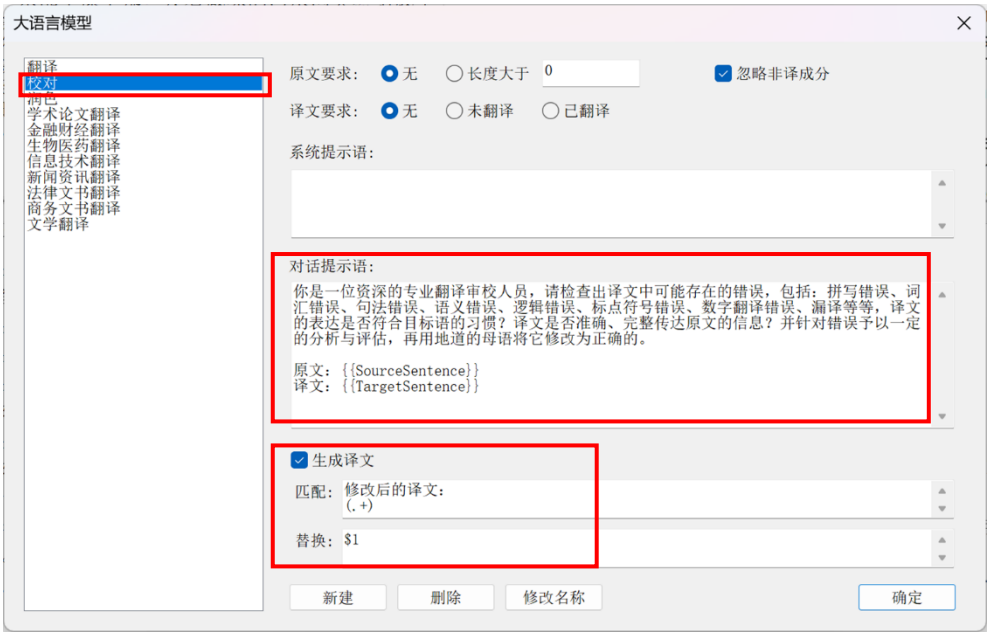


校对

创建“校对”指令，输入如下的提示语：

你是一位资深的专业翻译审校人员，请检查出译文中可能存在的错误，包括：拼写错误、词汇错误、句法错误、语义错误、逻辑错误、标点符号错误、数字翻译错误、漏译等等，译文的表达是否符合目标语的习惯？译文是否准确、完整传达原文的信息？并针对错误予以一定的分析与评估，再用地道的母语将它修改为正确的。

原文：{{SourceSentence}}
译文：{{TargetSentence}}

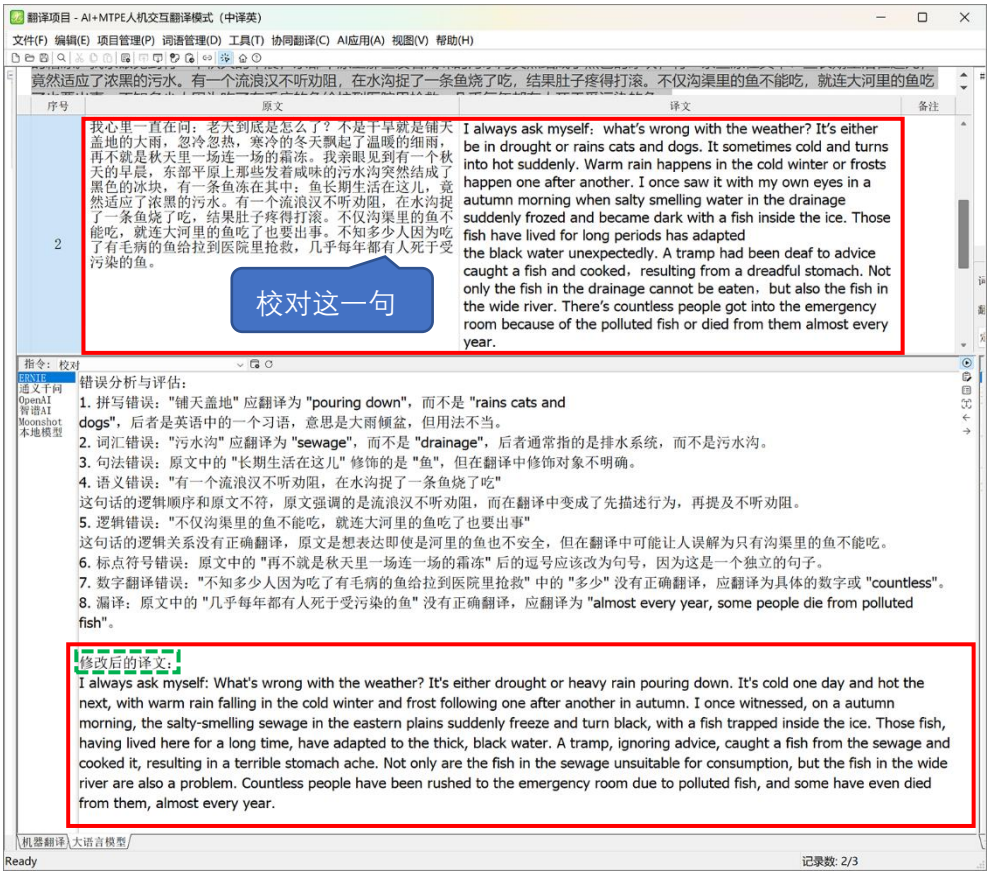


上图中“生成译文”选项，若勾选它，生成的译文可以按拷贝按钮粘贴到上方的译文区中，若不勾选它，也会生成译文，但仅用于浏览，按拷贝按钮不会粘贴到上方的的译文区中。“匹配”框、“替换”框中可以输入“正则表达式”，有时大模型输出的内容不仅仅是译文，还有许多其他内容，在“匹配”框、“替换”中可以输入“正则表达式”，以方便提取需要的内容到上方的译文输入框中。例如：校对时，AI会输出许多分析错误的原因内容供我们参考，这部分内容我们仅是浏览，并不进入译文。

点击下图的“校对”指令，选择“ERNIE(文心一言)”模型进行校对，当光标停在“昨日那种被追捕的恐慌，已经烟消云散。His panic of the previous evening was quite gone now.”这一句时，下方的“大语言模型”窗口立即显示来自“ERNIE(文心一言)”模型针对这一句的分析、校对结果：“译文在保持原意的基础上，更加突出了恐慌的个体性和消失的彻底性”。如下图：

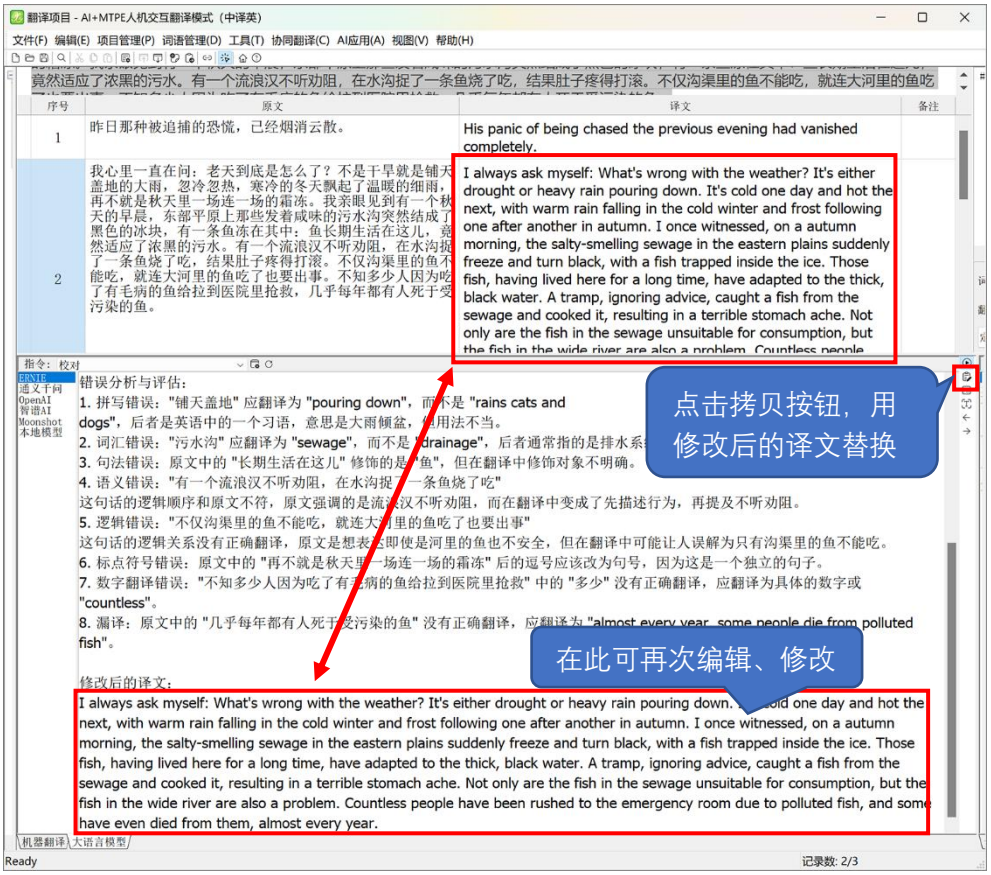


然后点击 → “下一句”按钮，会显示下一句的校对结果，如下图：



这一句是一个长段落，AI 逐一指出了存在“拼写错误”、“词汇错误”、“标点符号错误”等各种问题，并给出了“修改后的译文”。用户可以在下方的“大语言模型”窗口，对 AI“修改后的译文”再进行编辑、修改。修改完成后，点击右侧的拷贝按钮，可以直接提取“修改后的译文”此部分内容去替换旧译文，而不是粘贴窗口中所有的内容为译文，能这样有选择地从 AI 大模型反馈的信息中截取所需要的内容，是因为在创建“校对”指令时，在“生成译文”框的“匹配”、“替换”框中，“用正则表达式”输入了“匹配”和“替换”的条件，如下图：**如果在创建“指令”时，规范大模型的输出结果，使其“格式化”，在此写“匹配”和“替换”的条件时会比较容易。**如果此处不填写“匹配”和“替换”条件，则点击拷贝按钮，会将窗口中显示的所有内容粘贴到上方的译文区，显然不合适，此时只能手动选取需要的内容进行复制粘贴了。





第六节 用 AI 大模型进行全文批量翻译

在 AI 应用中增加批量翻译功能，可以利用 AI 进行全文的快速翻译。点击【**AI 应用**】➔【**大语言模型批量翻译**】菜单，如下图：注意：先点击左侧的具体文档，再点击“大语言模型批量翻译”菜单，否则，该菜单是灰色的。



在弹出的“大语言模型批量翻译”窗口中，点击“大语言模型”旁边的下拉按钮，选择需要使用的 AI 大语言模型，目前支持的模型有：ERNIE(文心一言)、通义千问、OpenAI(ChatGPT)、智谱 AI、Moonshot (Kimi)、本地模型。点击“指令”旁边的下拉按钮，可以选择自己编写好的各种翻译指令。如下图：



如果你还没有编写自己的提示语指令，请点击左下方的“ ”设置按钮，在弹出的“大语言模型”窗口中，点击下方的“新建”按钮，根据实际需要，创建提示语指令。例如：下图中，创建了一个“新闻资讯翻译”提示语指令。

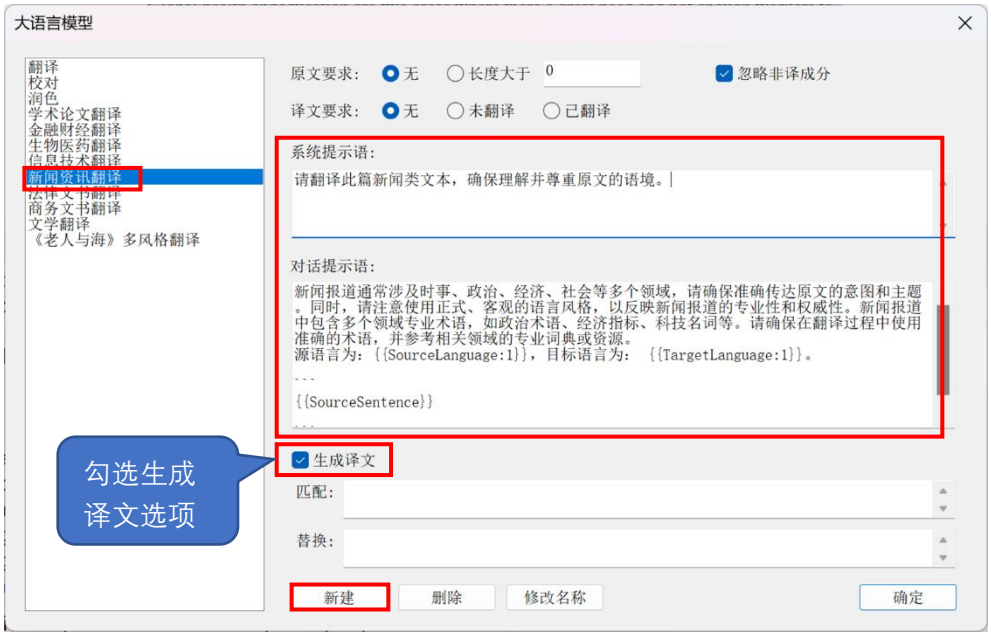
请翻译此篇新闻类文本，确保理解并尊重原文的语境。新闻报道通常涉及时事、政治、经济、社会等多个领域，请确保准确传达原文的意图和主题。同时，请注意使用正式、客观的语言风格，以反映新闻报道的专业性和权威性。新闻报道中包含多个领域专业术语，如政治术语、经济指标、科技名词等。请确保在翻译过程中使用准确的术语，并参考相关领域的专业词典或资源。

源语言为：{{SourceLanguage:1}}，目标语言为：{{TargetLanguage:1}}。

```

{{SourceSentence}}

```

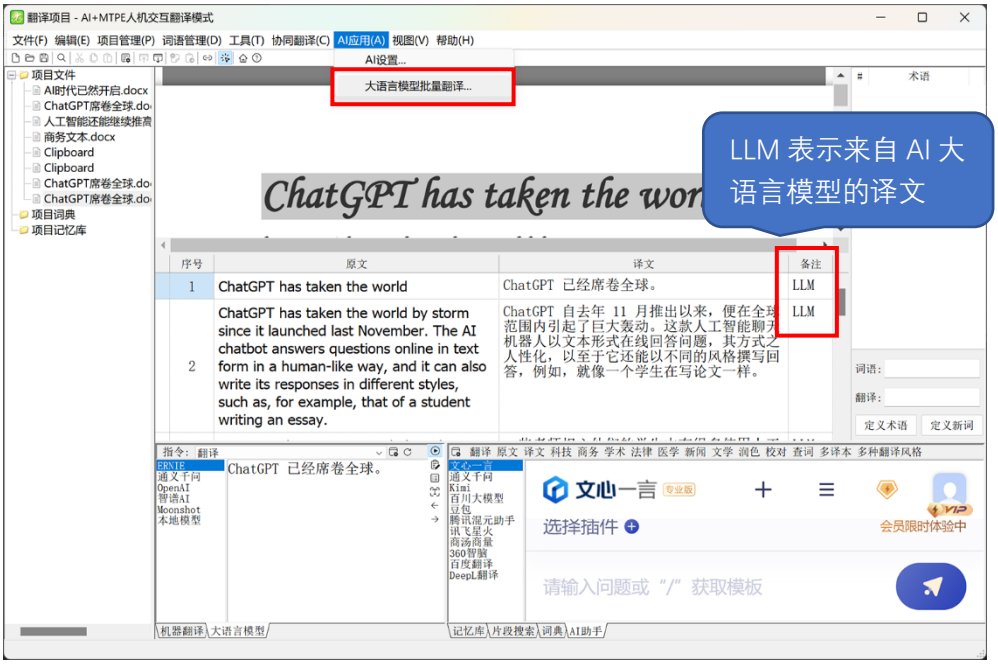


创建好后，在“大语言模型批量翻译”窗口中就可以选择这个“新闻资讯翻译”提示语指令。点击“确定”按钮进行全文翻译。此时会弹出翻译的进度条，如下图：



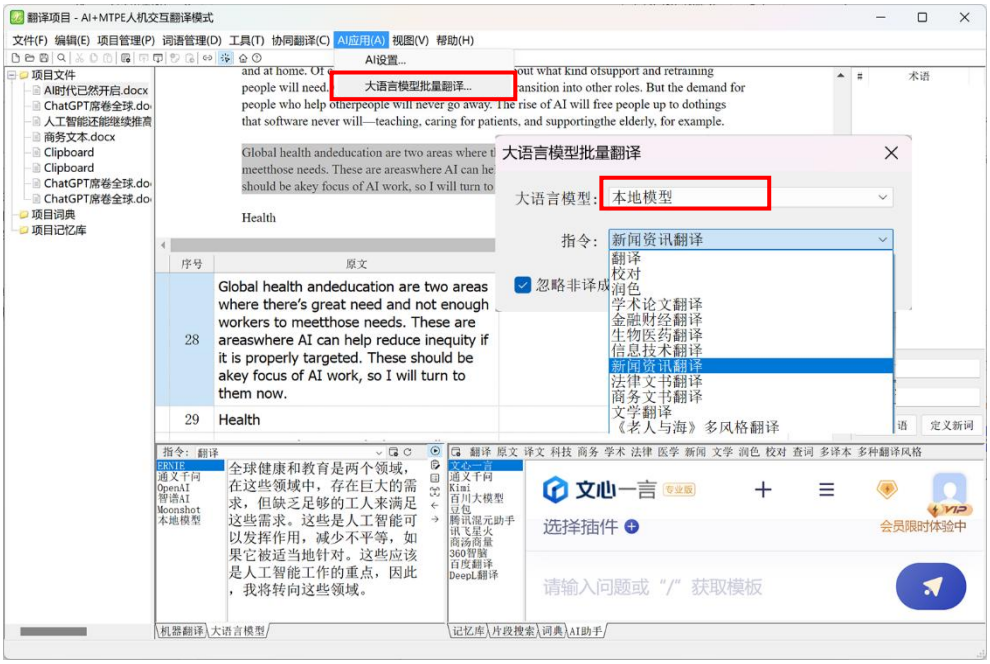
不同的 AI 模型，响应的速度不一样，有些 AI 模型 1 分钟内对发送的请求有限制，所以进行全文翻译时，需要耐心等待。如果等待的时间较长，可以点击进行条窗口的“取消”按钮，中断当前操作。再改用其他 AI 模型进行翻译，之前已经翻译过

的内容，将不再发送。翻译结束后，结果如下图：备注栏中会标注“LLM”，表示译文来自于 AI 大语言模型模型。



如果机器的性能比较好，对翻译质量要求不太高，可以先用“本地模型”进行全文翻译，使用“本地模型”进行翻译可以节省使用大语言模型的成本。对于“本地模型”译得不满意的句子，可以在参考在线 AI 大语言模型的结果。

如果有记忆库，也可以点击**【工具】➔【预翻译】**菜单，先将记忆库中有匹配的句子填上，然后再用“大语言模型批量翻译”，这样，已经有译文的句子，将不会再发送给 AI 大语言模型。


翻译项目 - AI+MTPE人机交互翻译模式
文件(F) 编辑(E) 项目管理(P) 词语管理(D) 工具(T) 协同翻译(C) AI应用(A) 视图(V) 帮助(H)
AI设置...
大语言模型批量翻译...
项目文件
AI时代已然开启.docx
ChatGPT席卷全球.do
人工智能还能继续推
商务文本.docx
Clipboard
Clipboard
ChatGPT席卷全球.do
ChatGPT席卷全球.do
项目词典
项目记忆库
大语言模型批量翻译
大语言模型：
本地模型
指令：
新闻资讯翻译
翻译
校对
润色
学术论文翻译
金融财经翻译
生物医药翻译
信息技术翻译
新闻资讯翻译
法律文书翻译
商务文书翻译
文学翻译
《老人与海》多风格翻译
忽略非译成
序号
原文
28
Global health andeducation are two areas where there's great need and not enough workers to meetthose needs. These are areaswhere AI can help reduce inequity if it is properly targeted. These should be akey focus of AI work, so I will turn to them now.
29
Health
术语
定义新词
指令：翻译
ERNIE
通义千问
OpenAI
智谱AI
Moonshot
本地模型
全球健康和教育是两个领域，在这些领域中，存在巨大的需求，但缺乏足够的工人来满足这些需求。这些是人工智能可以发挥作用，减少不平等，如果它被适当地针对。这些应该是人工智能工作的重点，因此，我将转向这些领域。
机器翻译
大语言模型
翻译 原文 译文 科技 商务 学术 法律 医学 新闻 文学 润色 校对 查词 多译本 多种翻译风格
文心一言
专业版
选择插件
会员限时体验中
请输入问题或"/"获取模板
记忆库
片段搜索
词典
AI助手

综述

人工智能大模型在与人类进行交互时，展现出卓越的适应性与表达优化能力，能够根据对话情境精准调整输出，从而在译后编辑领域展现显著优势。如何设计高效且精准的提示语指令，引导 AI 大模型生成与预期高度吻合的语言内容，成为了一项极具价值的研究课题。AI 大模型在翻译实践与教学场景中的辅助潜能不容忽视，其多元化的应用为提升翻译质量、效率及教育效果提供了有力支持。

对于专业译者，AI 大模型能够在译后编辑阶段扮演重要角色。通过精准识别并修正译文中的错误，以及对句法结构进行精细化调整与润色，AI 大模型有效提升了译文的专业度与流畅度。同时，它还能提供对译文的整体评估与针对性反馈，助力译者持续优化翻译技艺，提高译文品质。

对于翻译学习者，AI 大模型如同一位智能导师，能够提供基础且普适的翻译表达建议，帮助他们及时规避常见翻译误区，逐步掌握翻译规律与技巧。这种即时、个性化的辅导方式，极大地增强了学习者的自主学习能力和实践效果。

对于翻译教育工作者，AI 大模型在批改作业与教学资源建设方面发挥了显著作用。它能高效批阅学生译文，精准指出问题所在，减轻教师的工作负担。此外，AI 大模型还能生成高质量的示范译文，为学生提供丰富、规范的学习素材，有力推动翻译教学质量的提升。

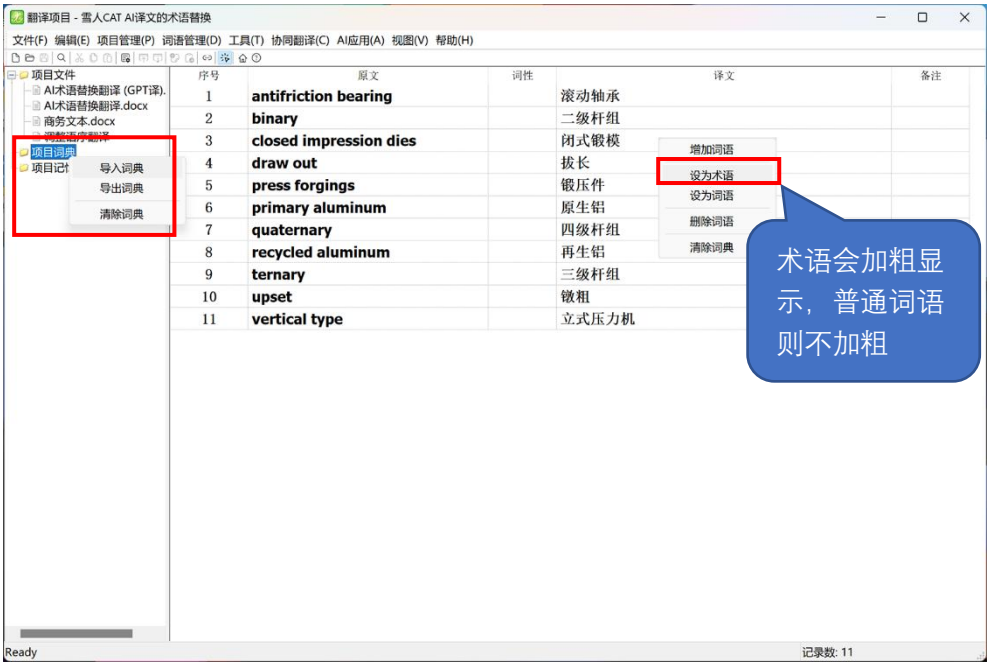
面对人工智能技术尤其是 AI 大语言模型技术的迅猛发展与迭代，人机交互在为翻译行业带来颠覆性挑战的同时，也开启了人机协作的新纪元。这促使翻译领域的产学研各界积极应对变革，积极探索将人工智能深度融入翻译实践，以实现翻译效率的显著提升，推动翻译行业的创新与进步。

第七节 依据用户术语库，让 AI 译得更准确

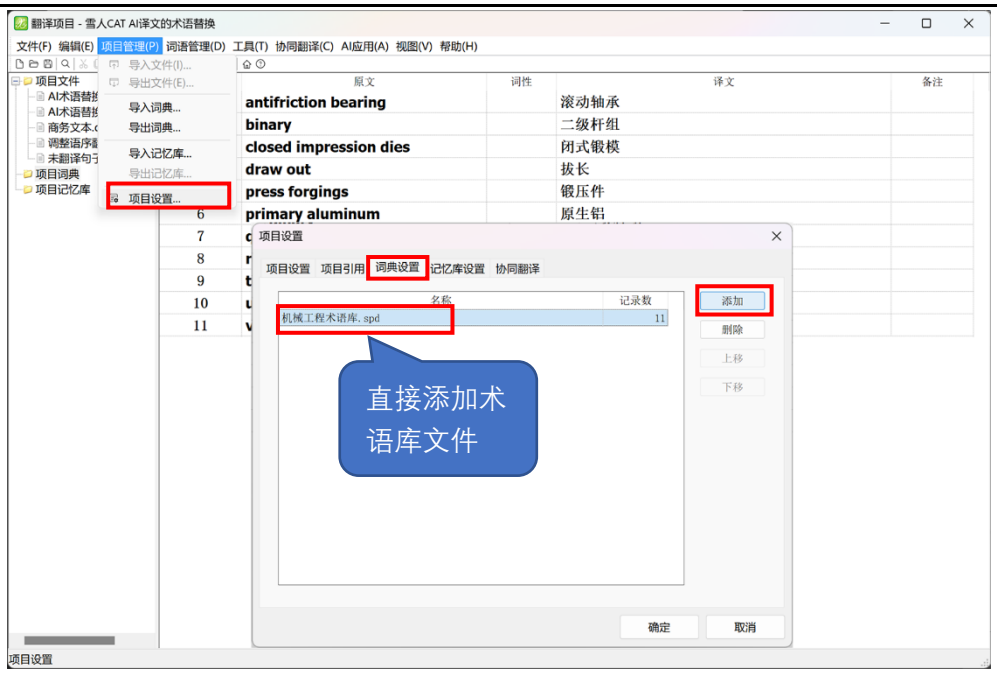
AI 大语言模型在翻译通用领域的文本时，通常表现出较高的准确性；但在涉及专业术语和技术细节时往往存在局限性。为解决这一问题，引入用户术语库成为了提高专业领域翻译准确度的有效手段。在翻译过程中，让 AI 依据用户术语库中的术语词义，进行语义匹配、术语替换和优化翻译逻辑，确保翻译过程中专业术语的准确无误，可以进一步提升 AI 译文的翻译质量。

下面介绍在雪人 CAT 中如何进行此项操作?

首先，将客户给的术语表或自己专业领域的术语在屏幕左侧的项目词典中导入，导入后默认这些词汇是普通词语。在右侧窗口中点击**【右键菜单】➔【设为术语】**功能将它们设定为术语，**设为术语后会加粗显示**，如下图：按住**【Shift】**键+鼠标点选可以选择多条记录进行设置，按住**【Shift】**键+鼠标点选第一条和最后一条记录，可以选择所有记录进行设置。注意：不是术语的词汇，不要盲目地将它设为术语，术语库太多、太乱，反而会降低翻译质量。

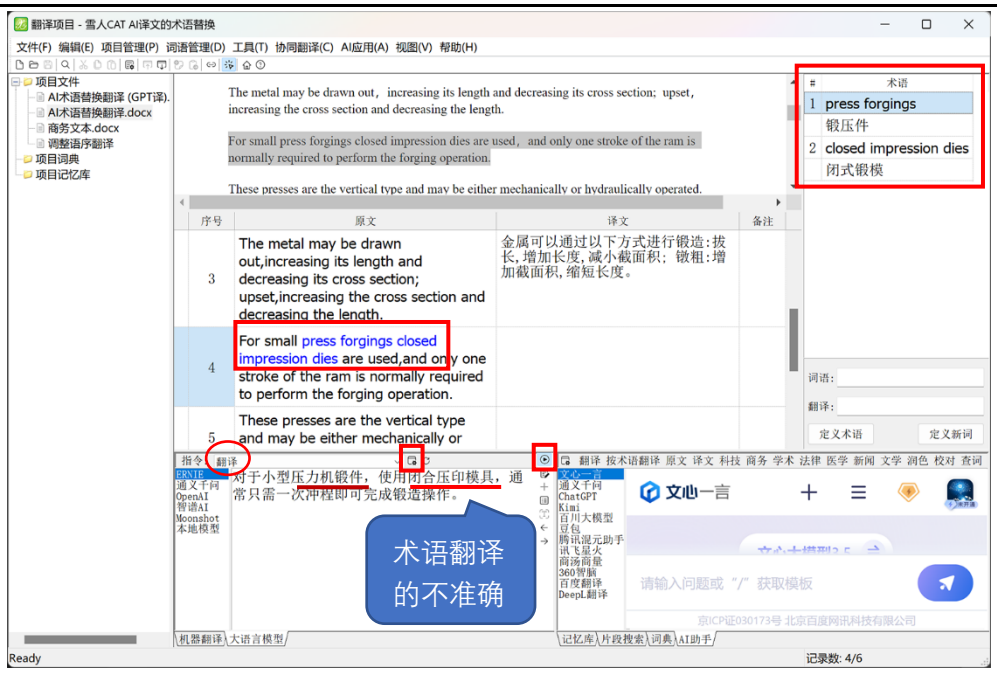


设好的术语，也可将它们导出成“术语库文件”，以后，可以通过点击**【项目管理】➔【项目设置】**菜单，在“词典设置”窗口中，直接添加这个“术语库文件”来使用。如下图：

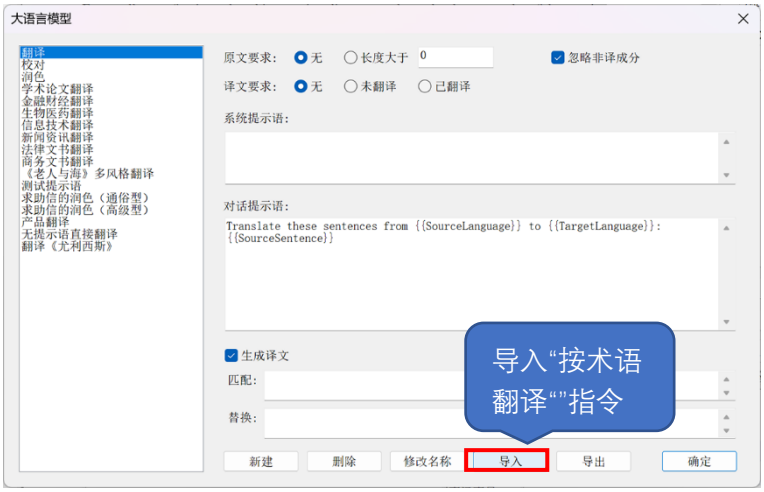


例如下图中这一句：For small **press forgings closed impression dies** are used, and only one stroke of the ram is normally required to perform the forging operation. 该句包含 2 个术语，在窗口右侧的术语区已经提示出来了，在原文中涉及的术语也用**蓝色字体**标注出来。但是，如果我们选择的是普通“翻译”指令（该指令仅是要求 AI 翻译成中文），AI 将“**press forgings**”译为“压力机锻件”，“**closed impression dies**”译为“闭合压印模具”，显然不够专业，经测试各个传统的机器翻译引擎也是一样，都不能准确翻译这两个术语。

我们希望 AI 能按我们术语库中定义的词义来翻译，在术语库中“**press forgings**”的词义是“锻压件”，“**closed impression dies**”的词义是“闭式锻模”。



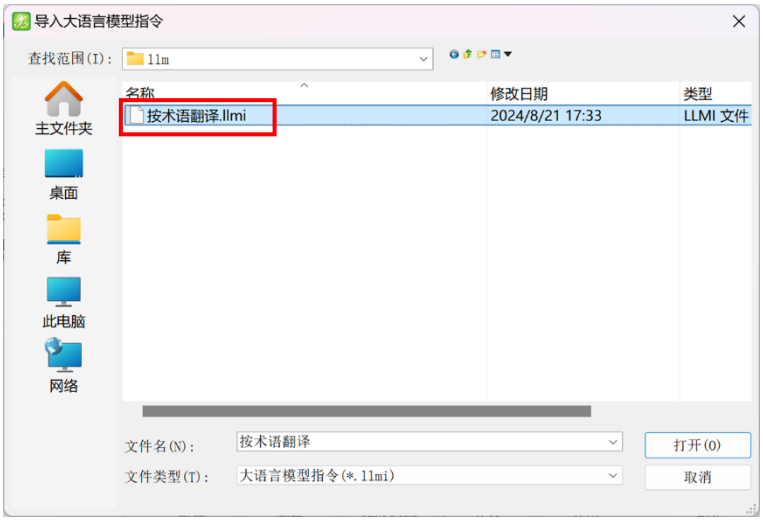
如果点击指令旁边的下拉按钮，在弹出的菜单栏里没有“按术语翻译”这一指令，可以使用以下方法加入：点击指令旁边的设置按钮，弹出“大语言模型”指令窗口，如下图：



点击窗口下方的“导入”按钮，找到雪人 CAT 软件存放数据的“SCAT”文件夹，再双击“llm”文件夹（该文件夹用于存放 AI 的提示语指令），如下图：



双击“按术语翻译.llmi”这个文件，如下图:



就可以导入这条“按术语翻译”指令了。如下图:



这条提示语指令是：

You are a helpful assistant.

Your task is provide a professional translation from {{SourceLanguage}} to {{TargetLanguage}}. The source text is below, delimited by <SOURCE> and </SOURCE>. Translate only the part within the source text, with specific attention given to accurately translating these terms delimited by <TERMS> and </TERMS>. Only the translation of the text needs to be output, not other extraneous content.

<TERMS>

{{TermList}}

</TERMS>

<SOURCE>

{{SourceSentence}}

</SOURCE>

其中：{{TermList}}是术语变量，可以从用户术语库中检索到当前句子所使用的术语，并将它提交给 AI 大语言模型。{{SourceSentence}}是原文变量。转换后的表达如下：

<TERMS>

press forgings(锻压件)

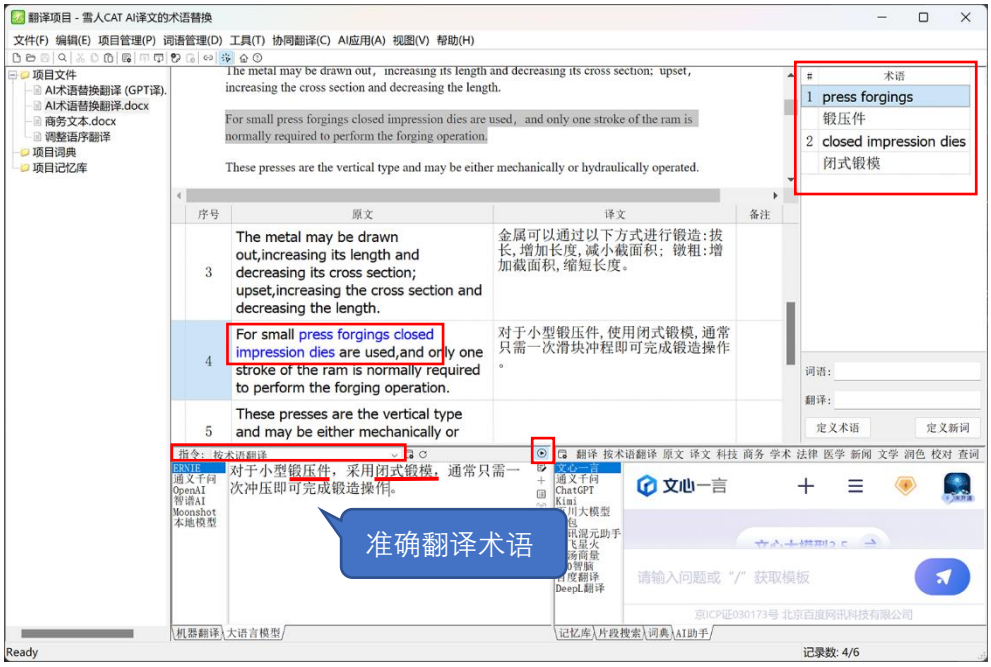
closed impression dies(闭式锻模)

</TERMS>

<SOURCE>
For small press forgings closed impression dies are used, and only one stroke of the ram is normally required to perform the forging operation.
</SOURCE>

用户可以在此指令的基础上修改，例如增加背景知识、翻译文风等要求。通过导入、导出按钮，可以方便与他人共享、交流提示语指令。“按术语翻译.llmi”这条指令，如果您是第一次安装运行雪人 CAT 软件，默认会自动出现在下列按钮的选择框中，如果您已经编写了许多自己的提示语指令，则需要按上述的方法导入这条指令，如果导入时出现和您已有的指令同名，则会弹出窗口要求输入新的指令名称。

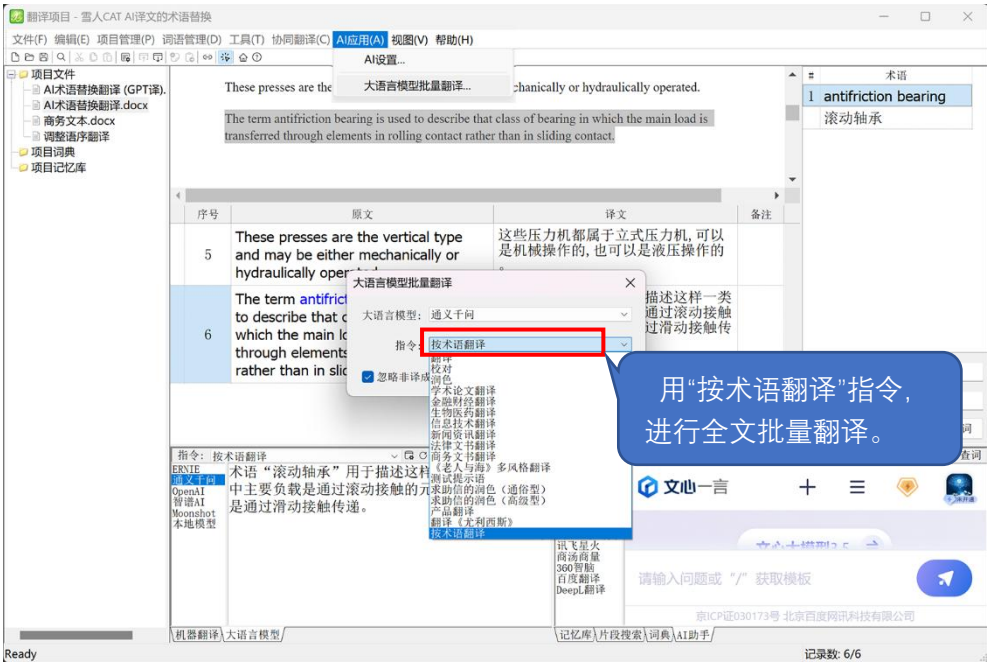
导入“按术语翻译.llmi”指令后，退出窗口。点击屏幕左下方的“指令“旁边的下拉按钮，选择这个“按术语翻译”指令，并点击右侧的“启用”按钮， AI 大语言模型就会按这条指令，进行精确的术语翻译，我们可以看到：“**press forgings**”已经被准确翻译成“**锻压件**”，“**closed impression dies**”也被准确翻译成“**闭式锻模**”。如下图：



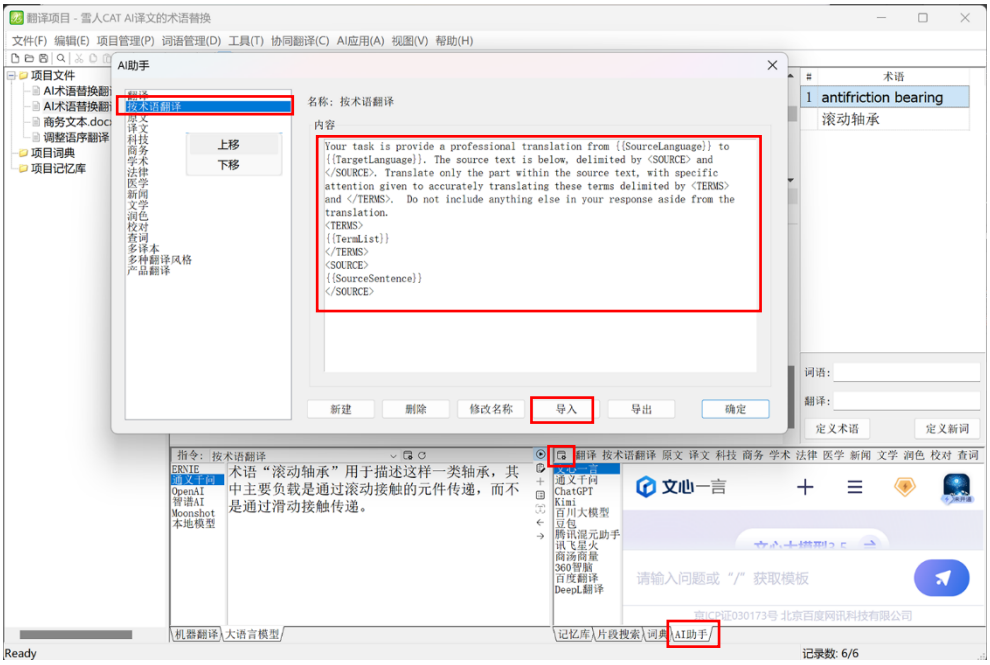
接着，光标移动到下一句，“**antifriction bearin**”也被准确翻译成“**滚动轴承**”了。如下图：



如果希望 AI 对全文的术语都能快速、精确地翻译，可以点击【**AI 应用**】➔【**大语言模型批量翻译**】菜单，并在“指令”中，选择“按术语翻译”指令，则可以快速完成全文的翻译，并且使术语翻译准确。如下图：



“按术语翻译.llmi”这个指令文件，也可以在右下方的“AI 助手”中导入，点击“AI 助手”窗口中的设置按钮，再点击下方的“导入”按钮，可以将它导入，“AI 助手”窗口的各个 AI 大语言模型，可以在无 API KEY 的情况下，通过网页界面来使用。在右侧窗口，也可点击**【右键菜单】➔【上移/下移】**来调整各个指令按钮在屏幕中部的排布位置。如下图:



利用术语库使 AI 更精准翻译，对大型的 AI 模型执行的效果会比较好。一些小型的 AI 模型，可能不能正确理解这种提示语指令，所以执行效果会不太理想。对于中英、英中，由于大语言模型的训练数据丰富，效果普遍不错，但对于小语种，由于训练的数据相对少一些，效果也会差一些。对于小语种，利用自己的记忆库，在基础模型的基础上训练的专有模型，效果会比普通的大模型好。

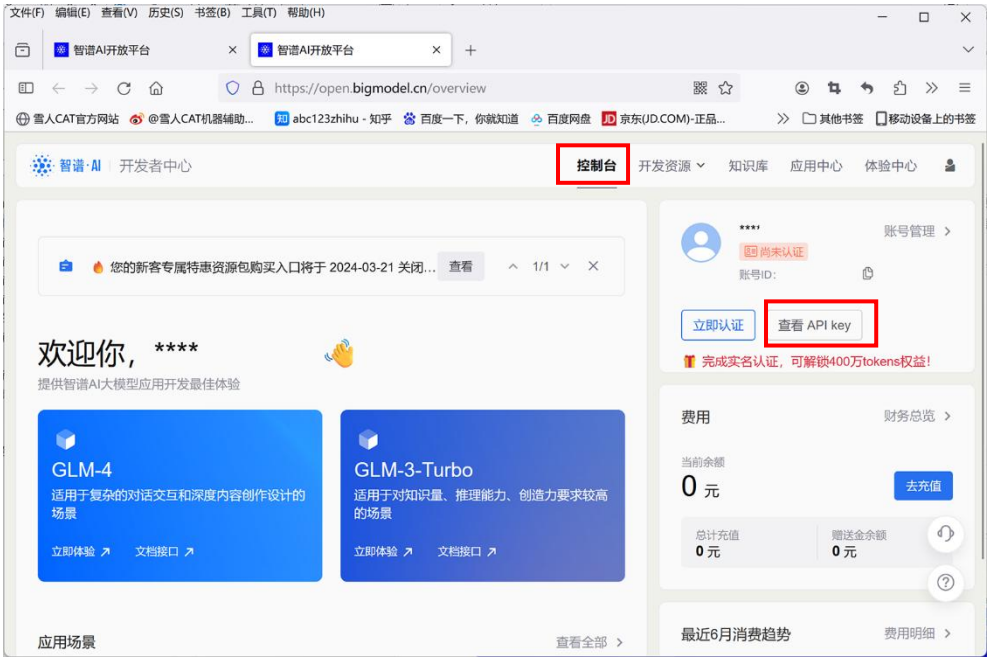
第八节 AI 大模型 API 的申请

一、智谱 AI 大模型 API 申请

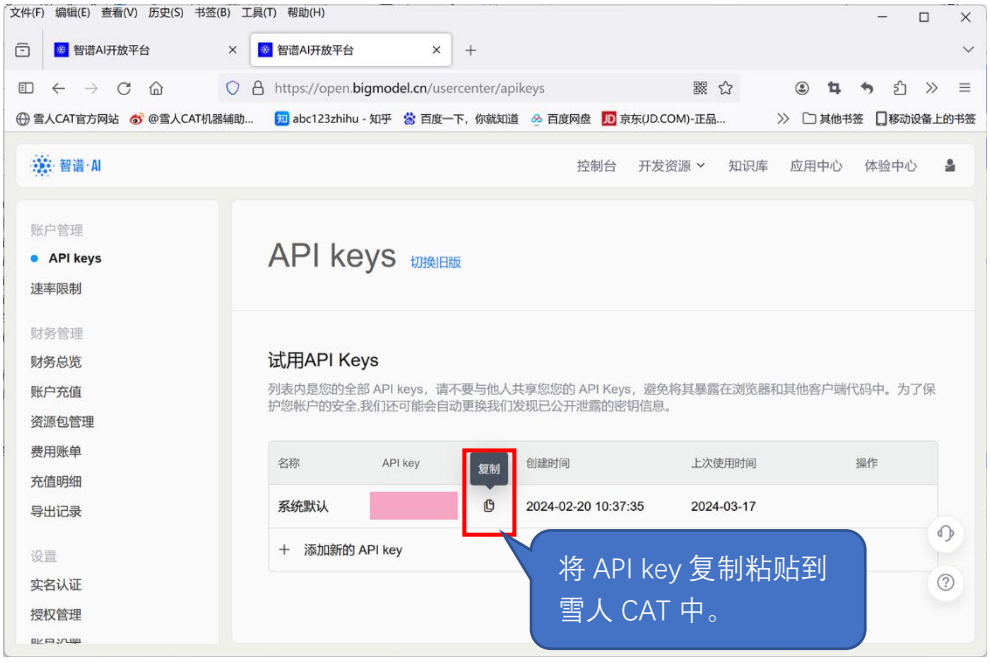
智谱 AI 大模型 API 申请比较简单，新用户有一定的免费额度，首先登录智谱官网 https://open.bigmodel.cn ，注册账号，如下图：



点击右上角的开发工作台，在控制台页面中点击查看 API key 按钮，如下图

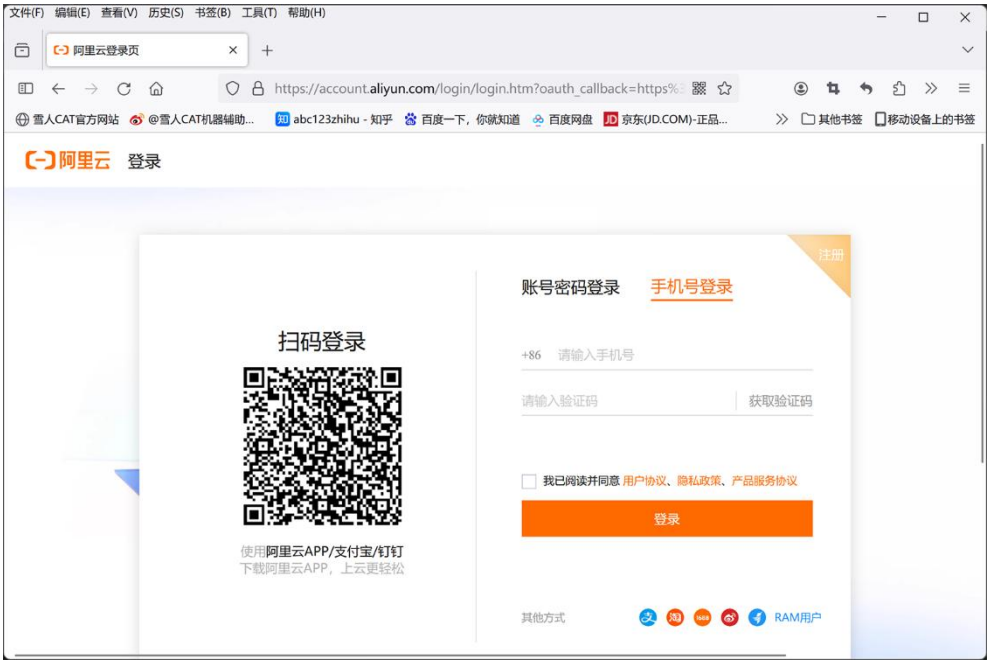


可以看见自己申请的 API key，点击下方的“复制”按钮，将申请到的 API Key 复制粘贴到雪人 CAT 的【AI 应用】➔【AI 设置】➔【大语言模型】窗口中。

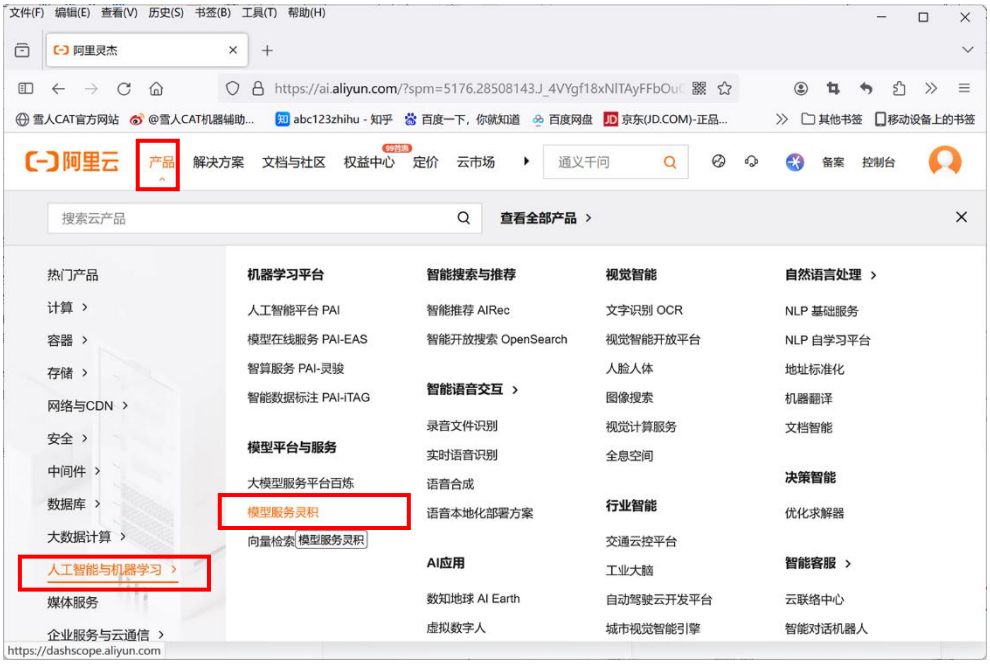


二、阿里通义千问　AI 大模型 API 申请

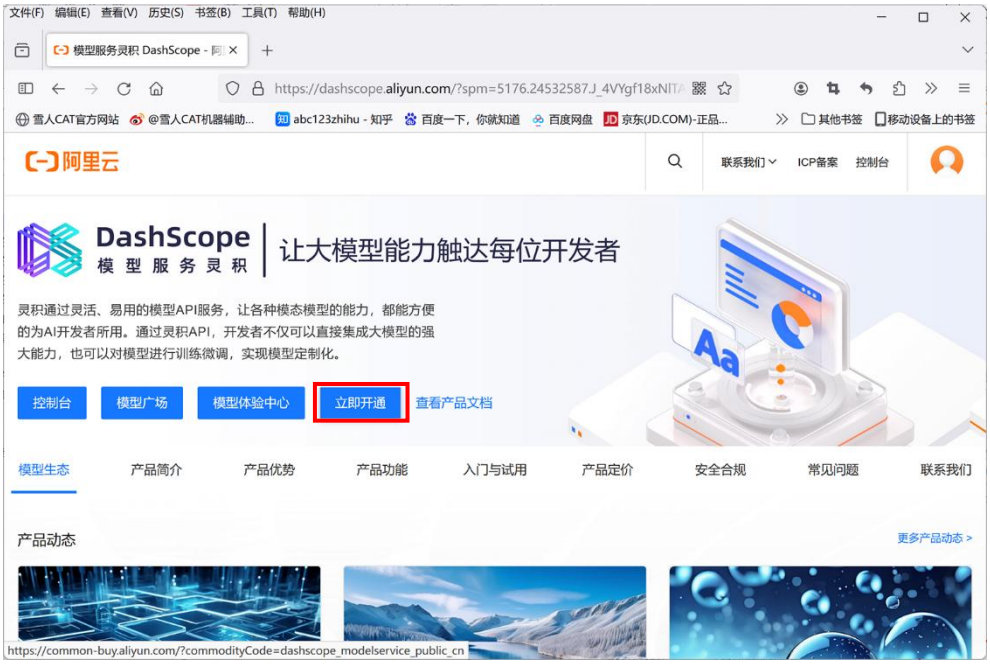
登录阿里云官网：https://account.aliyun.com ，注册阿里云账号，并完成实名认证。



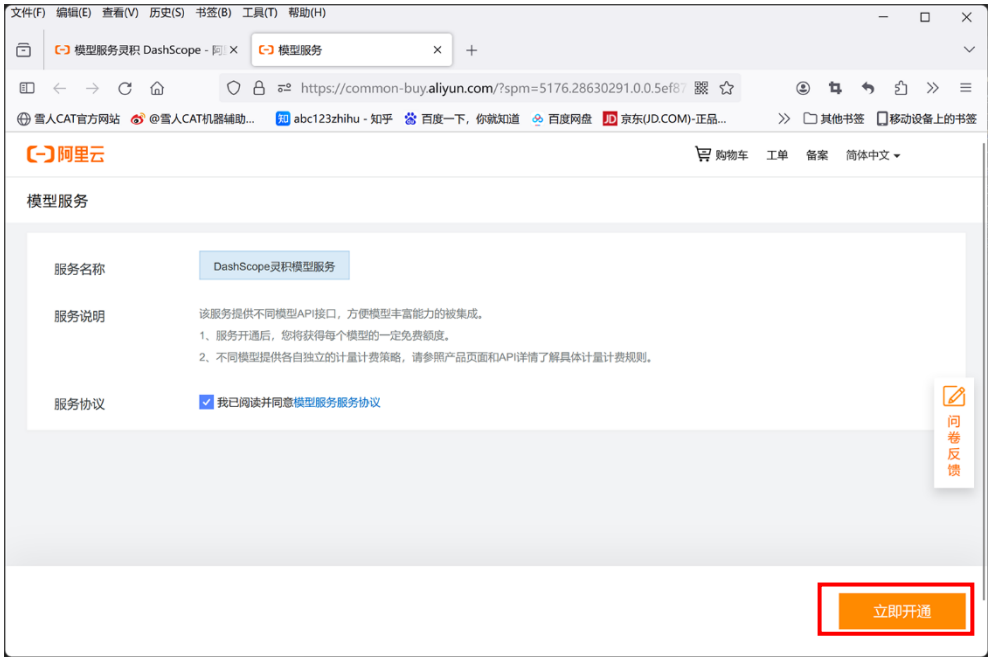
点击“产品”导航栏，在“产品中”找到“人工智能与机器学习”，在该分类中找到“模型服务灵积”，如下图：

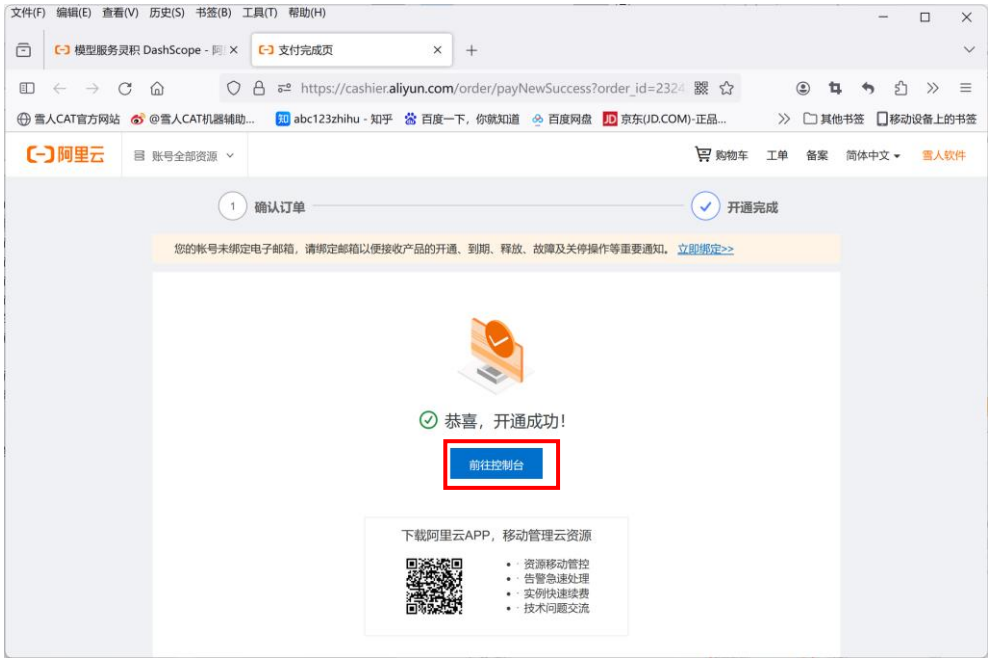


点击“模型服务灵积”，弹出窗口如下图：

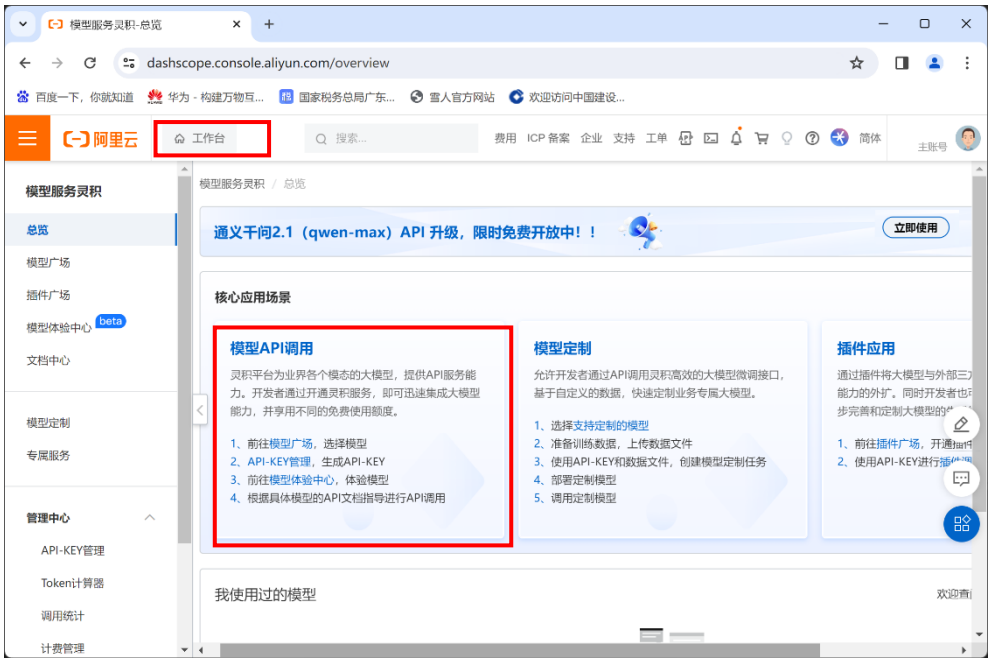


点击“立即开通”按钮。如果账号没有进行实名认证，开通时会弹出窗口，要求先进行实名认证。

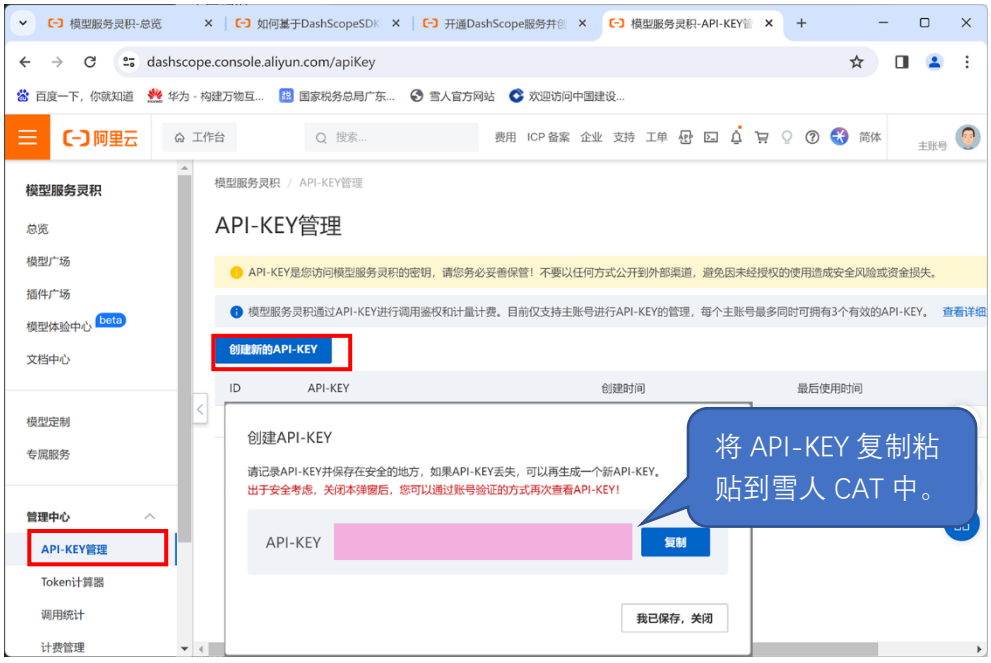




成功开通后，前往控制台，如下图:



点击左侧的“API-KEY 管理”，再点击“创建新的 API-KEY”按钮，将申请到的 API Key 复制粘贴到雪人 CAT 的**【AI 应用】➔【AI 设置】➔【大语言模型】**窗口中。

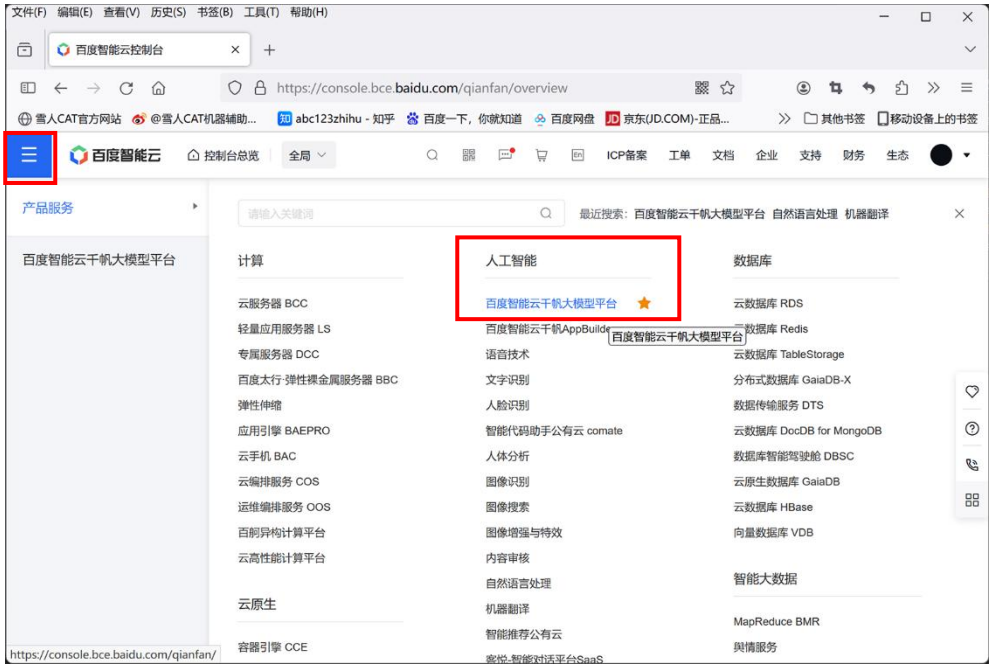


三、百度文心一言　AI 大模型 API 申请

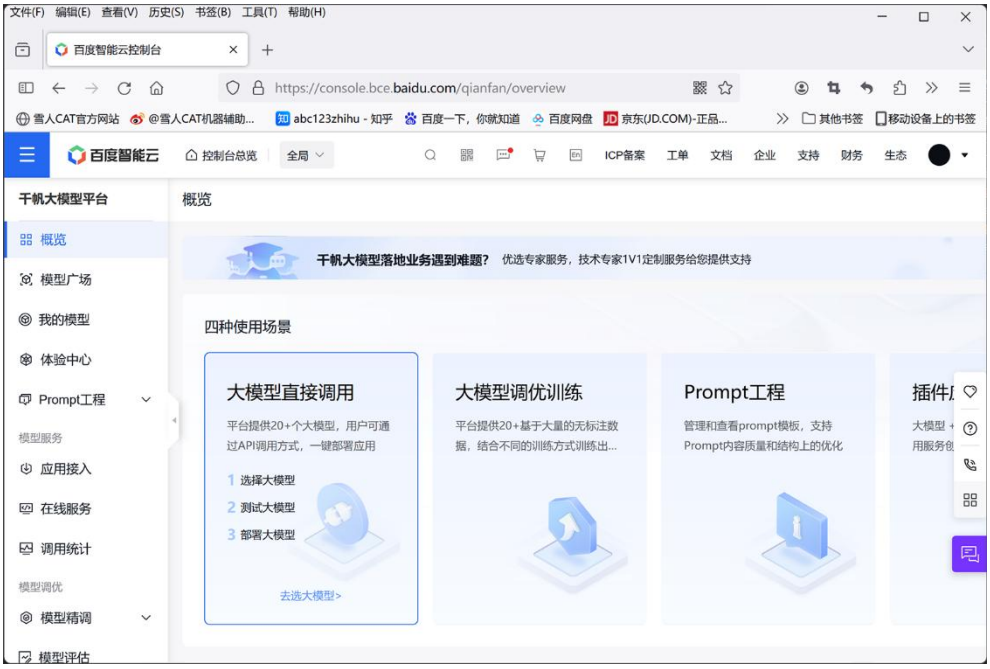
登录百度智能云网站 https://login.bce.baidu.com 注册百度账号。



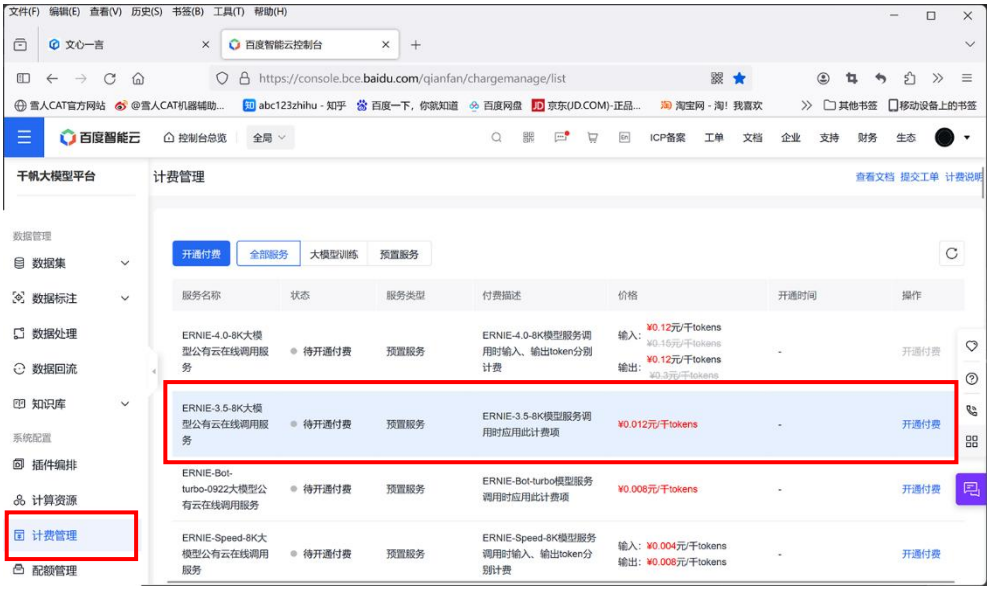
点击左上角的蓝色图标，在“人工智能”分类中，找到“百度智能云千帆大模型平台”，如下图：



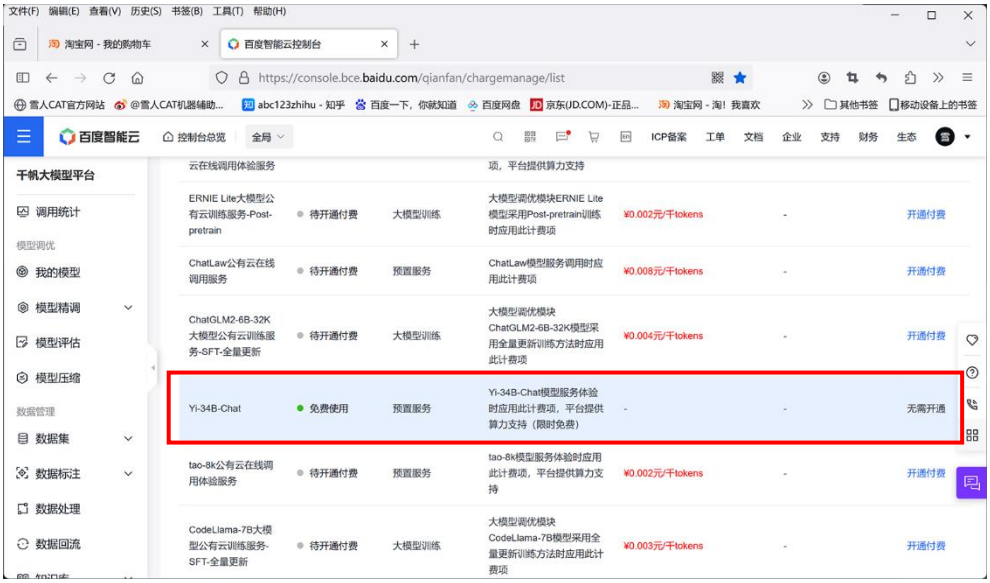
点击“百度智能云千帆大模型平台”后，如下图：



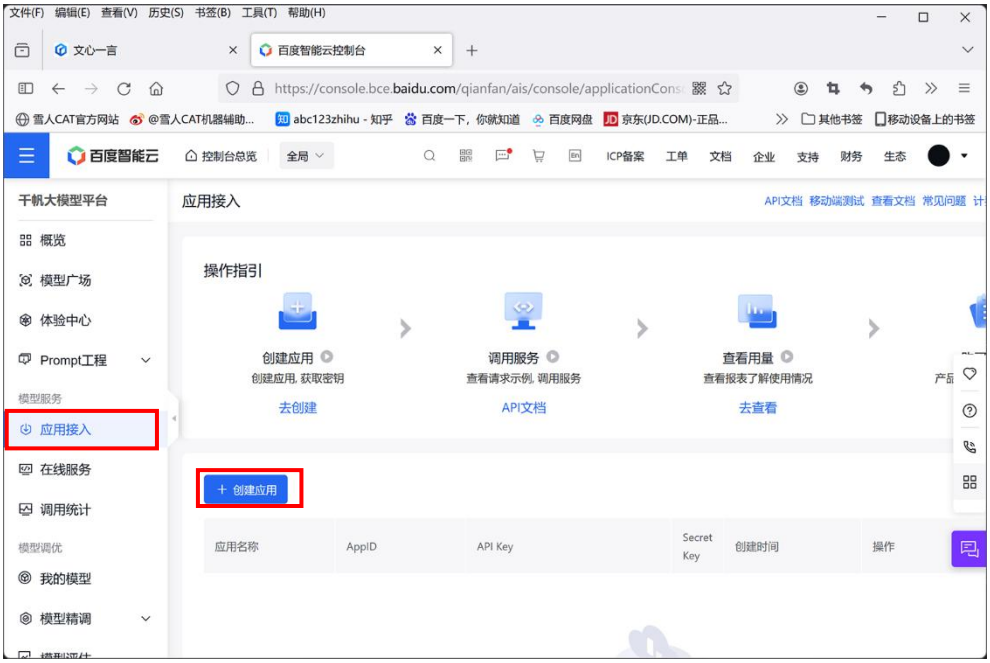
点击左侧导航栏的“计费管理”，根据需要选择开通的大模型，不同模型的收费标准不同，用户自行决定开通哪些模型。这里选择了开通“ERNIE-3.5-8K 大模型公有云在线调用服务”这个模型。



目前“Yi-34B-Chat 模型”无需开通，可以免费使用。如下图：



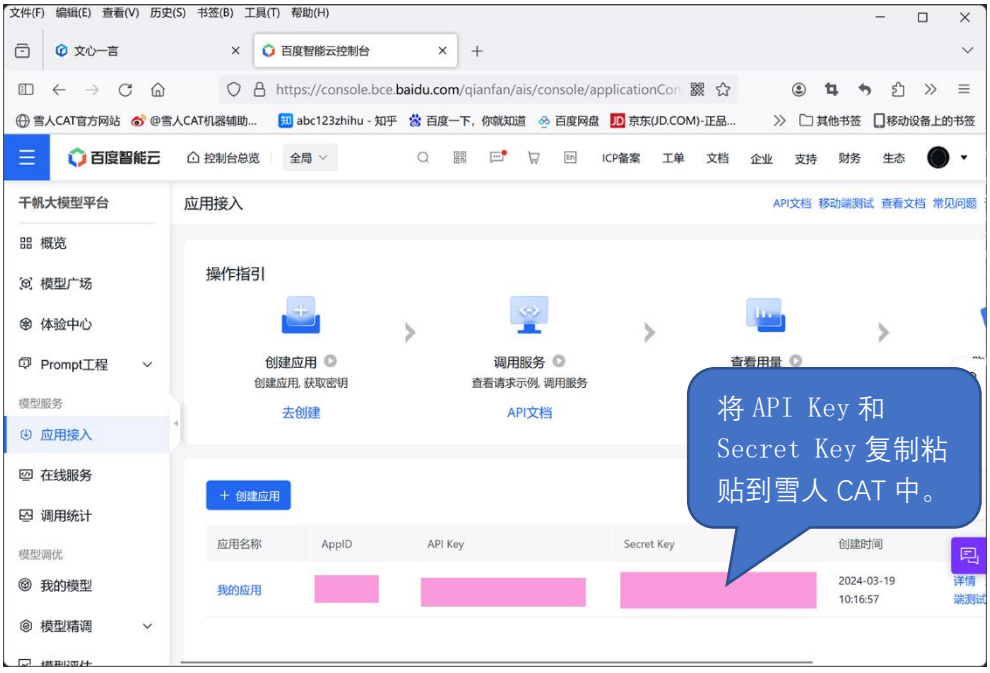
点击左侧导航栏的“应用接入”，再点击“创建应用”按钮，如下图：



创建一个自己的应用，其中填写“应用名称”和“应用描述”。如下图：

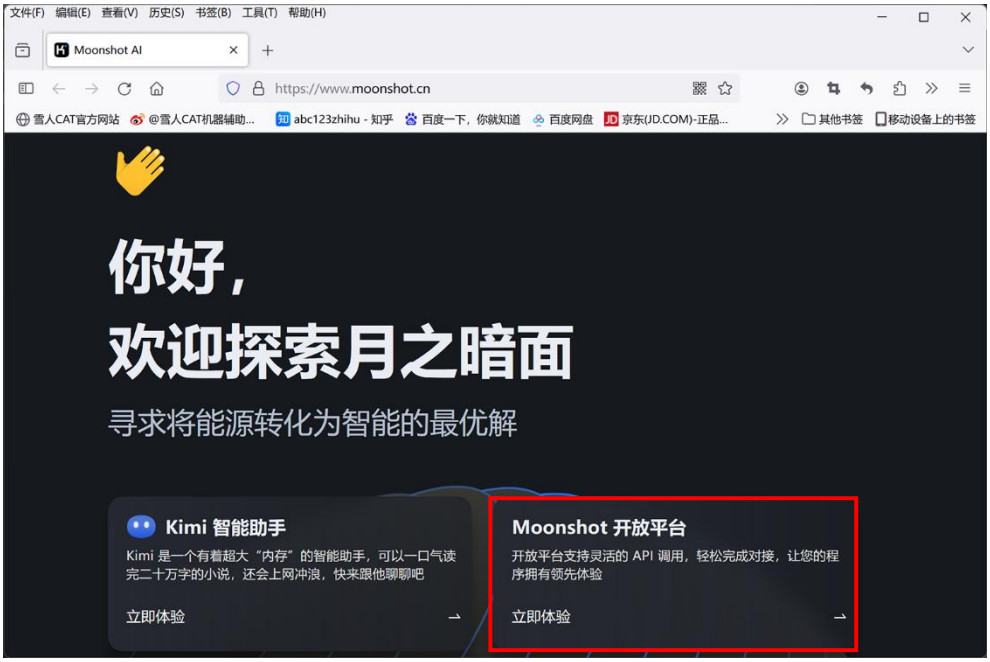


点击“确定”按钮后，将所创建应用的 API Key 和 Secret Key 复制粘贴到雪人 CAT 的**【AI 应用】➔【AI 设置】➔【大语言模型】**窗口中。



四、Kimi　AI 大模型 API 申请

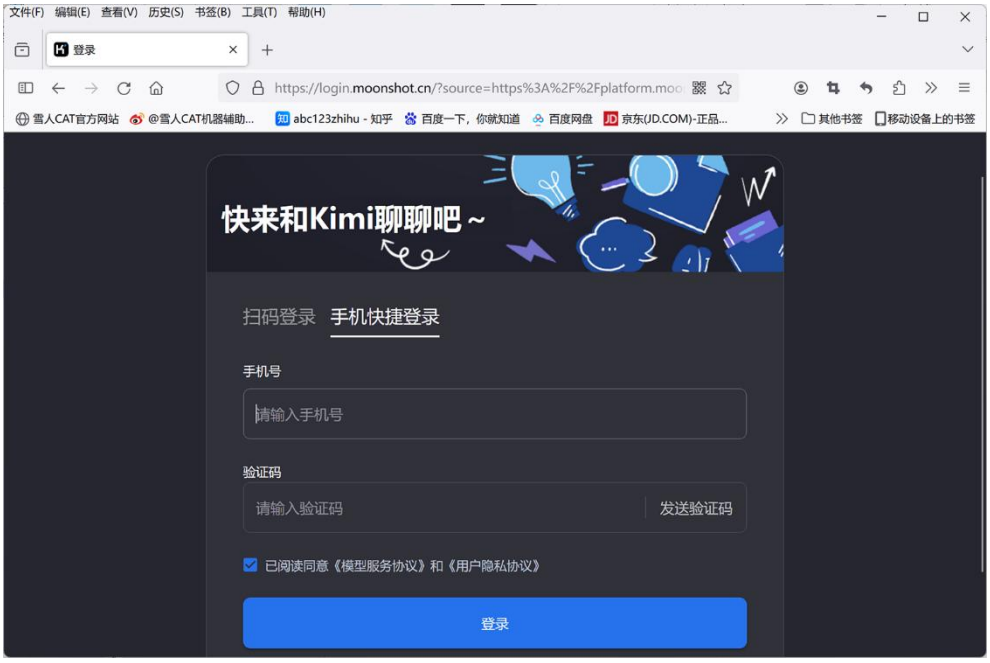
登陆 Kimi 的官网 https://www.moonshot.cn ，如下图:



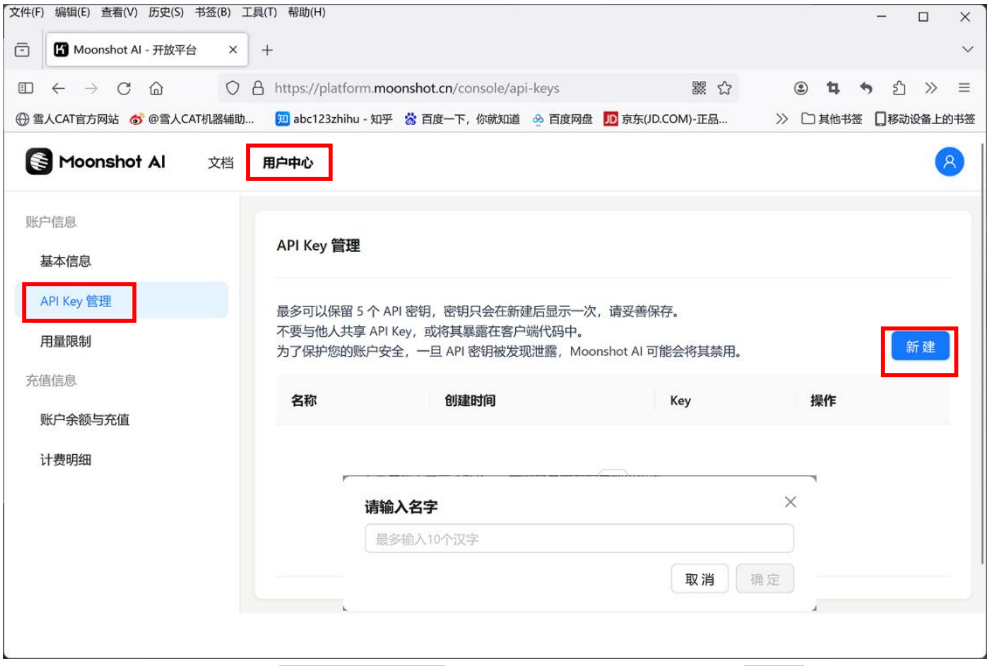
点击 Moonshot 开放平台，会进入如下界面:



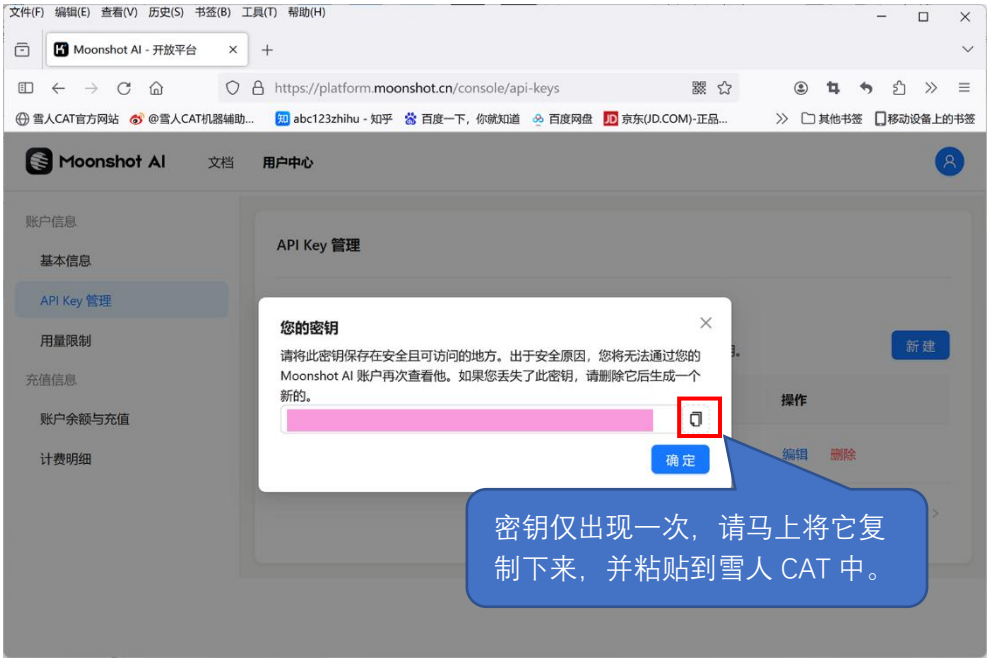
点击上方导航栏的“用户中心”，会弹出要求注册/登录的界面，如下图：



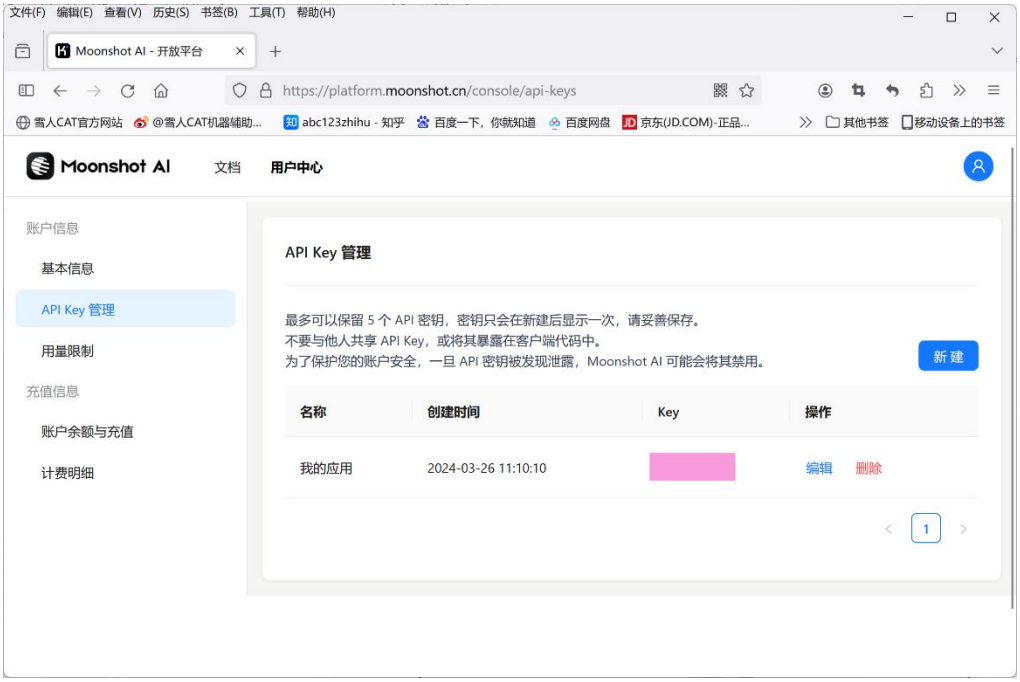
输入手机号注册后，点击登录按钮，进入如下界面：



点击左侧导航栏的“API Key 管理”，然后点击屏幕右侧的“新建”按钮，输入应用名称，创建一个应用。此时会弹出密钥窗口，如下图：

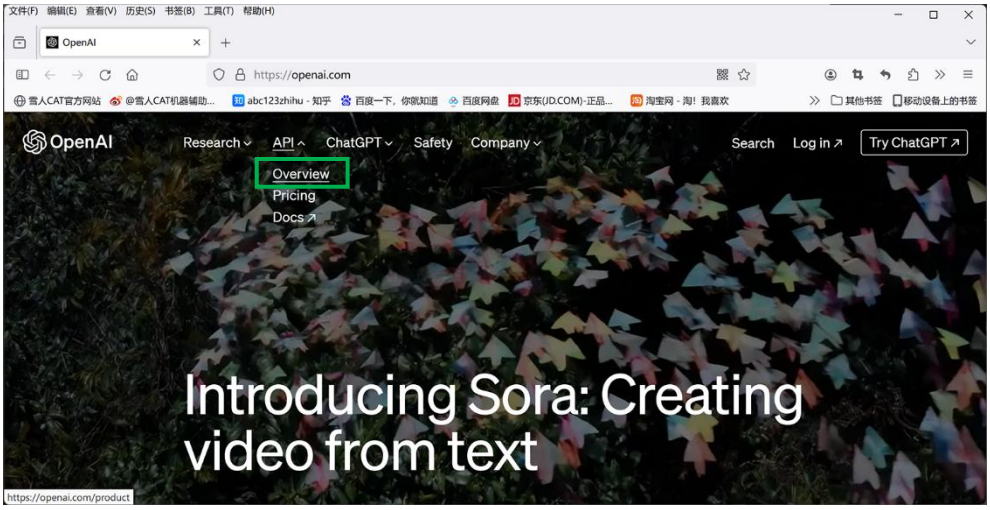


注意：密钥仅出现一次，请马上将它复制下来，并将申请到的 API Key 复制粘贴到雪人 CAT 的**【AI 应用】➔【AI 设置】➔【大语言模型】**窗口中。

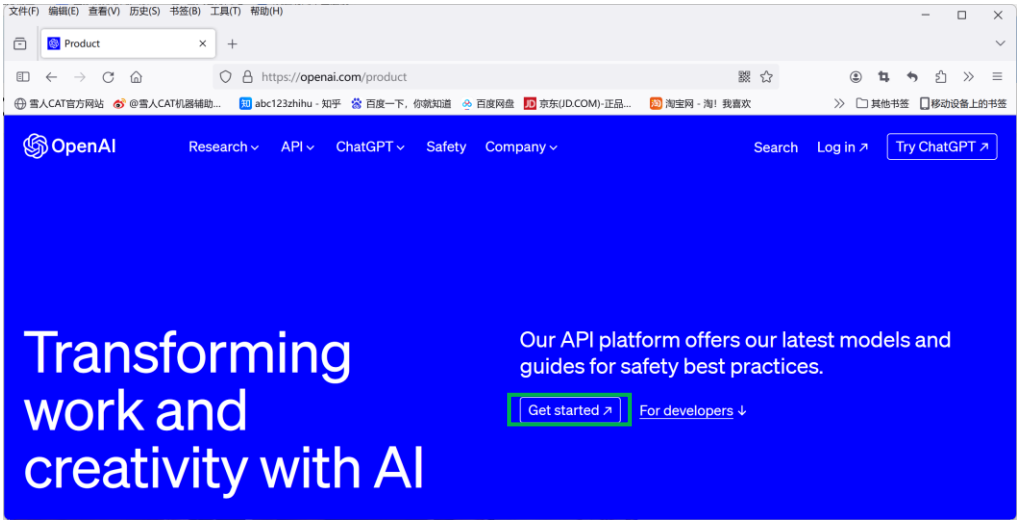


五、ChatGPT AI 大模型 API 申请

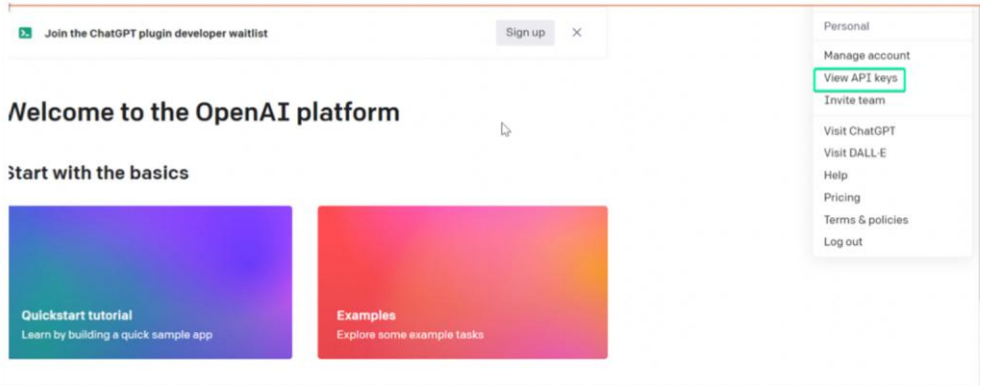
注册 OpenAI 账号:登录 OpenAI 的官方网站（https://www.openai.com）注册一个账号。在注册过程中，需要提供一些个人信息，如邮箱地址、用户名和密码等。还要先设置一下扣费账户，点页面右上角的【personal】，然后选择【Manage account】，之后在跳转的页面选择【Billing】，然后点击【Set up paid account】，把信用卡信息依次填写保存即可。



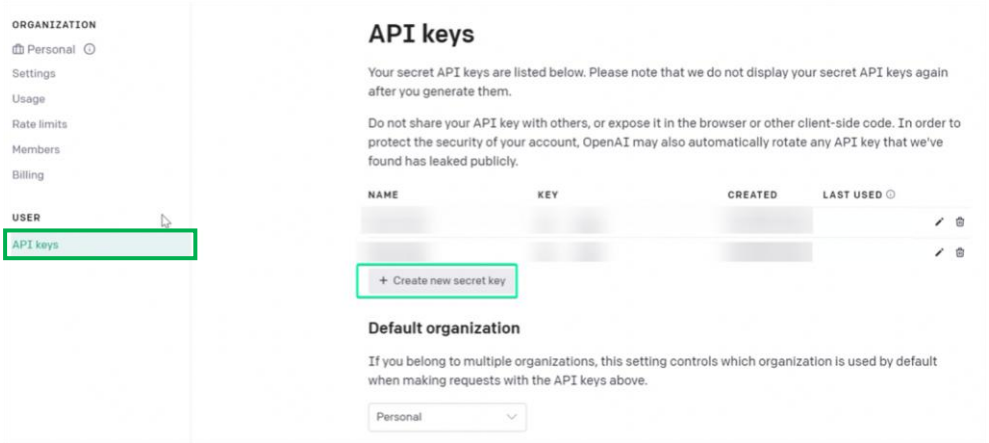
点击导航栏的“API”下面的“Overview”菜单，再点击【get started】，如下图:



然后点击右上角【personal】，点击【view API key】，如下图:



点击左侧窗口的【API keys】导航，再点击【Create new secret key】，新建一个 key，如下图：



在弹出的创建框中填写一个名字，方便备注，然后点击【Create secret key】，创建之后出现的一串字符就是 API key 了，点击旁边的绿色按钮即可复制，**注意：复制之后要自己妥善保管好，关闭这个窗口即无法再看到这个 key，也无法再行复制这个 key。如果没有复制下来就关闭了该窗口，只能将这个 key 删除，再重新创建一个。**将申请到的 API Key 复制粘贴到雪人 CAT 的**【AI 应用】➔【AI 设置】➔【大语言模型】**窗口中。



由于 ChatGPT 不对中国大陆用户开放注册及国内的限制使用，所以，国内用户仅有 ChatGPT 的 API Key 是无法使用的，还需要设置合适的代理服务器。另外，GPT

的使用费用比较高，尤其是 GPT-4，以下是从 OpenAI 网站摘录的费用供参考，Turbo 是提速、降质版。

GPT-4	: Input $30.00/1M tokens	Output	$60.00/1M tokens
GPT-4 Turbo	: Input $10.00/1M tokens	Output	$30.00/1M tokens
GPT-3.5 Turbo:	Input $ 0.50/1M tokens	Output	$ 1.50/1M tokens

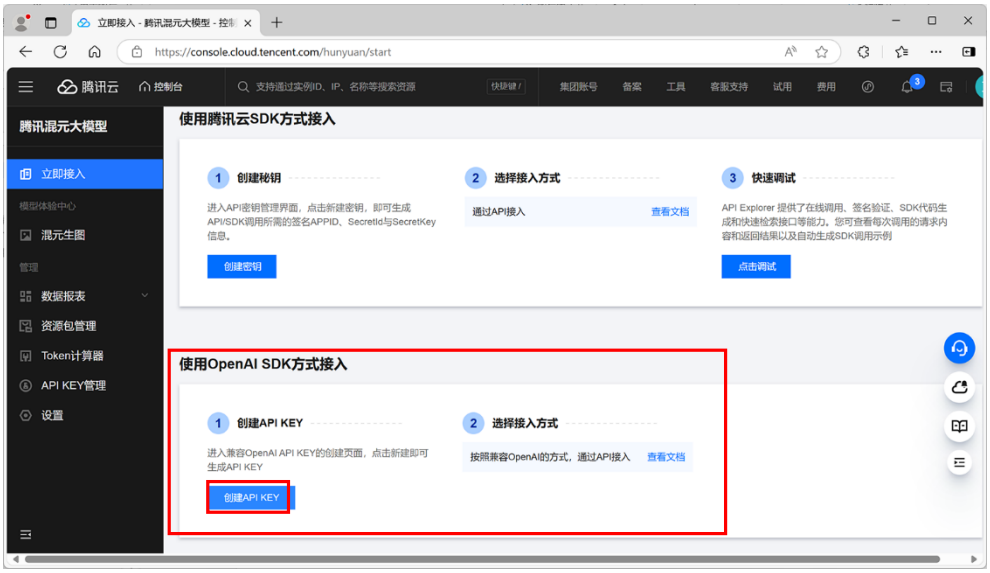
六、接入腾讯混元 AI 模型的 API

登录腾讯云网站 https://console.cloud.tencent.com/hunyuan/start 注册账号并申请开通“腾讯混元大模型”的 API，方法如下：

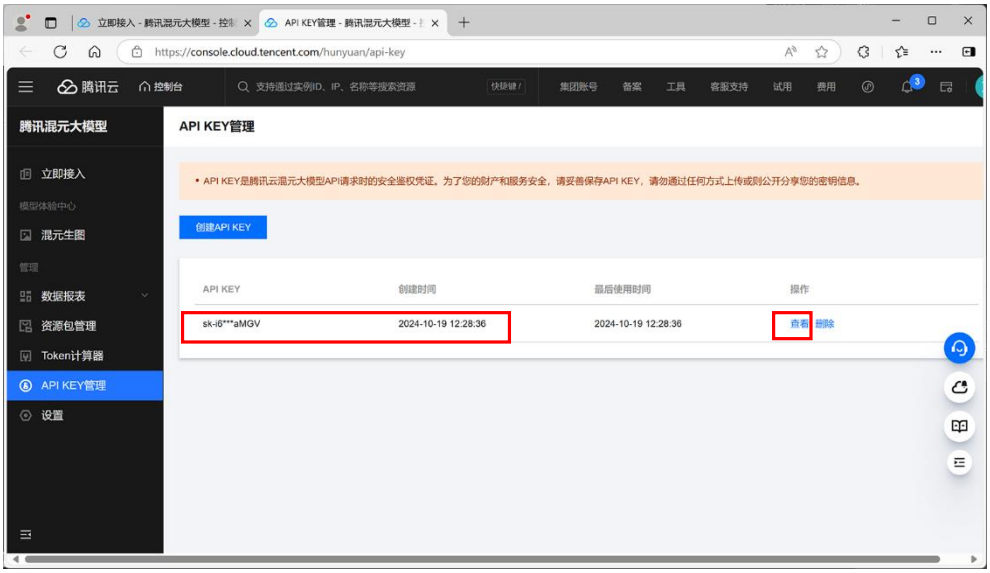
点击左侧导航栏的“立即接入”，然后点击“立即开通”按钮。如下图：



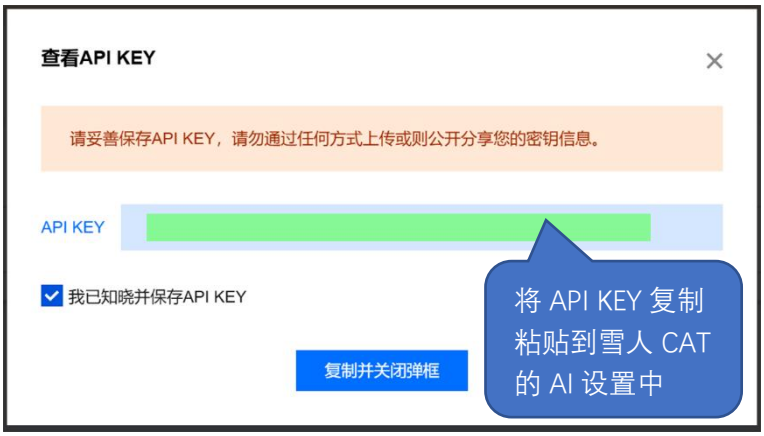
在下面窗口中，选择“使用 OpenAI SDK 方式接入”，点击“创建 API KEY”按钮。



将会创建一个 API KEY，是一串字符。如下图：



点击“查看”，可以查看到完整的 API KEY，将 API KEY 复制粘贴到雪人 CAT 的 AI 设置中，即可实现腾讯混元的 API 调用。如下图：

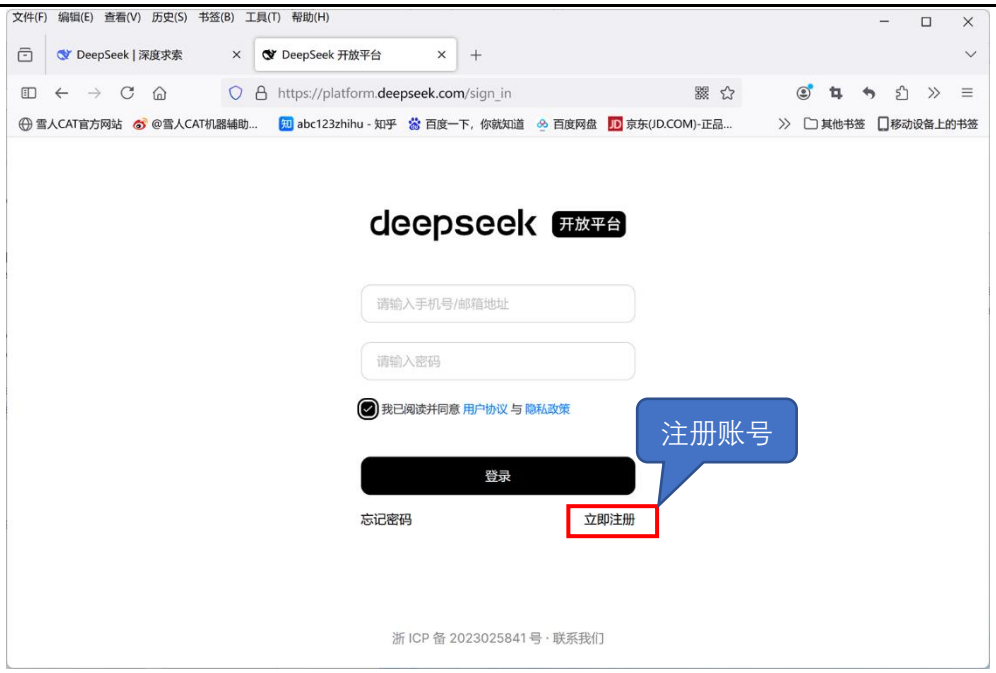


七、DeepSeek 深度求索 AI 大模型 API 申请

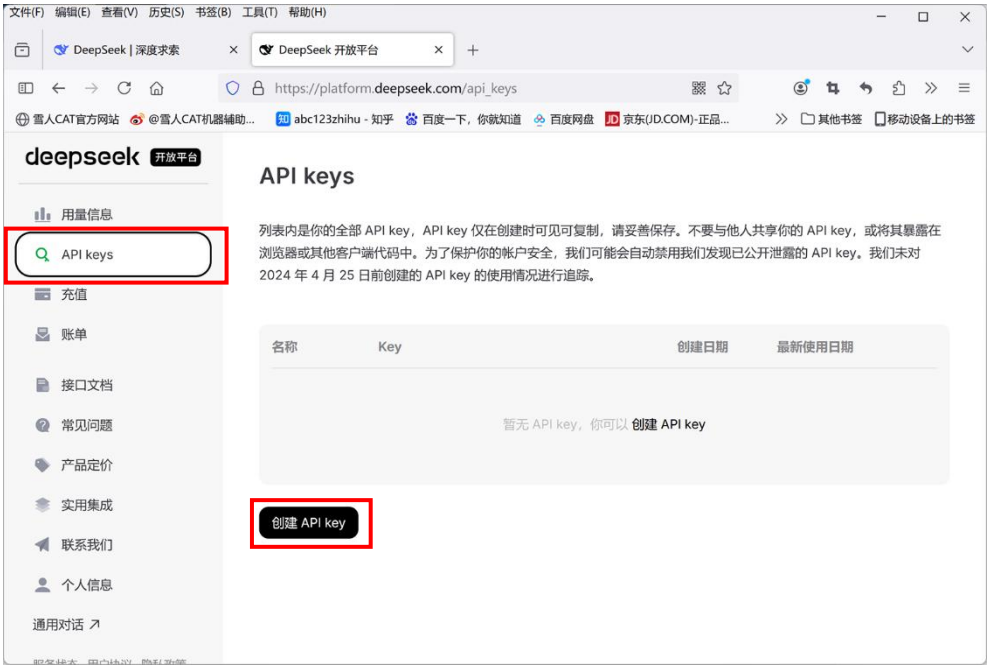
登录深度求索的网站 https://www.deepseek.com 注册账号。如下图:



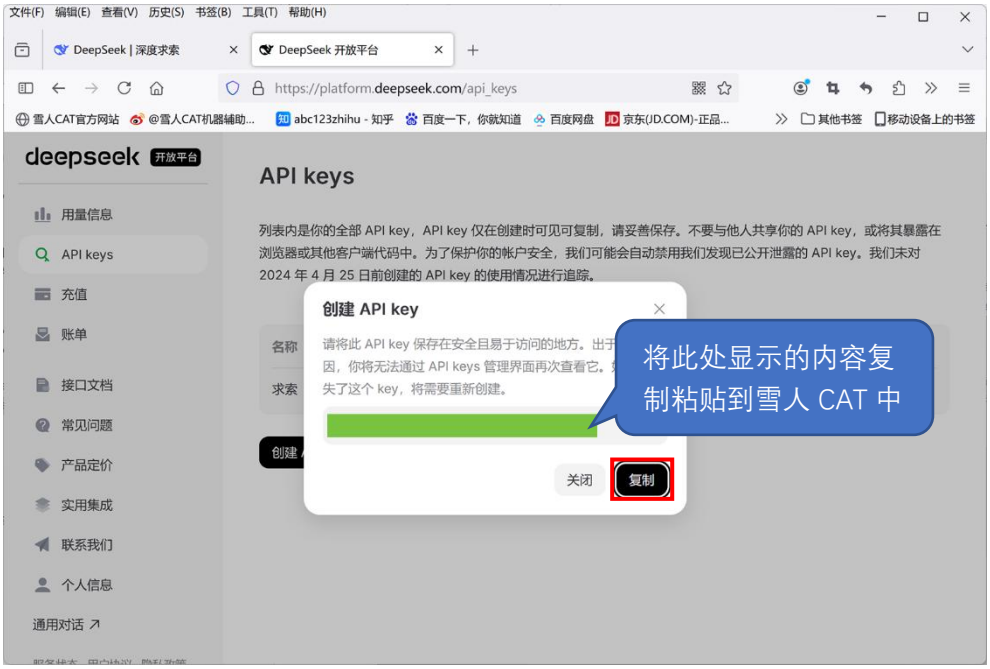
点击上图的“接入 API”，弹出窗口如下:



点击“立即注册”按钮，输入手机等信息后，将会跳转到以下页面：



点击左侧导航栏的“API keys”，在右侧窗口中，点击下方的“创建 API key”按钮，输入名称后，即可获得一个“API keys”，如下图：

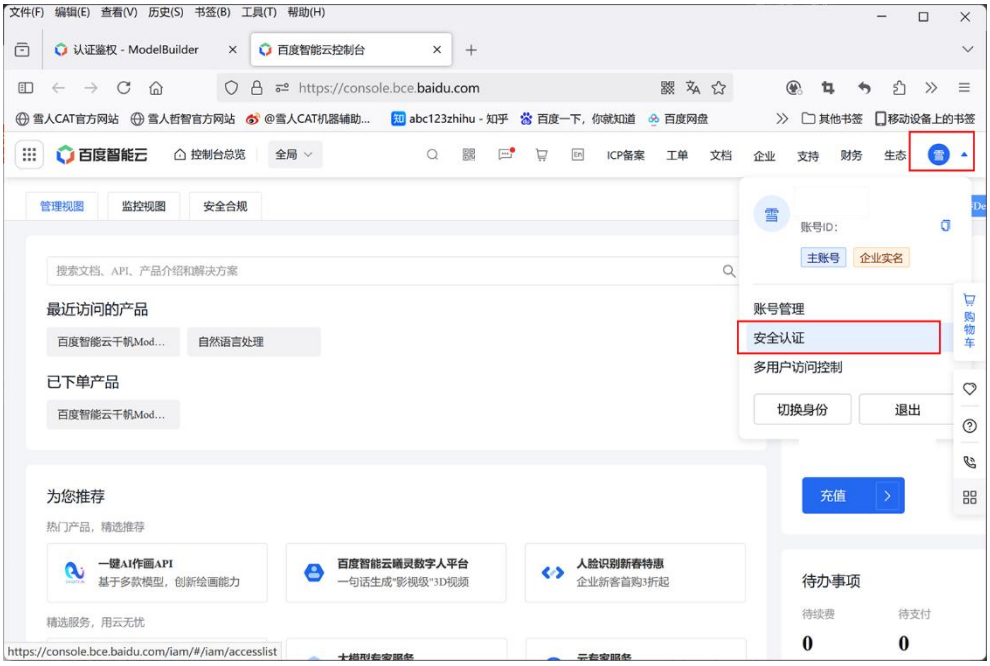


将申请到的 API Key 复制粘贴到雪人 CAT 的**【AI 应用】➔【AI 设置】➔【大语言模型】**窗口中。

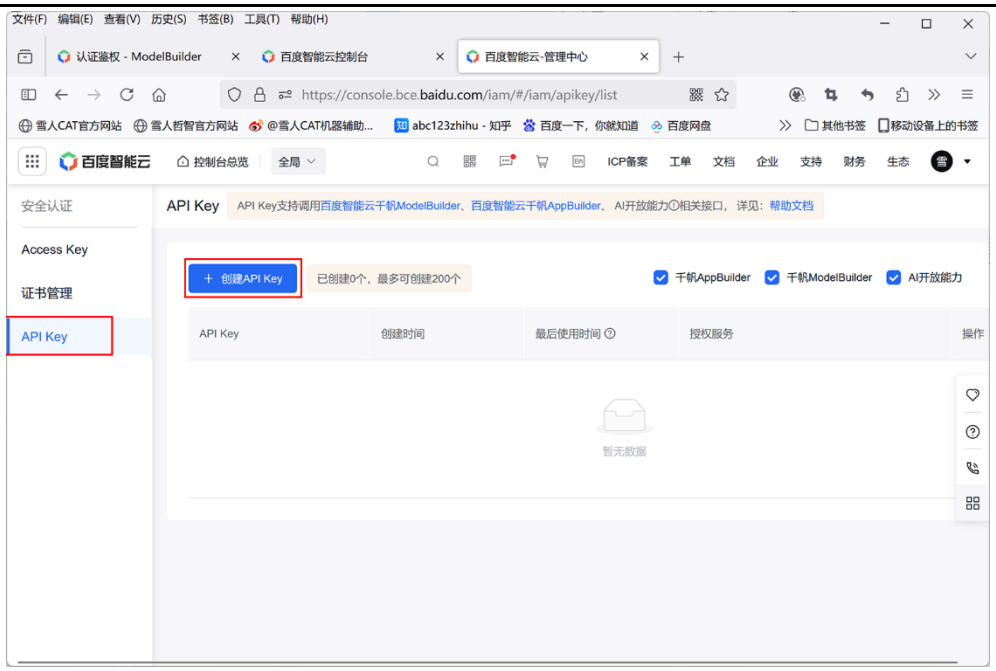
第九节 在百度平台调用 DeepSeek

由于使用 DeepSeek 的人太多，DeepSeek 官方服务器繁忙，响应速度比较慢，建议在其他平台上调 DeepSeek。百度智能云千帆 ModelBuilder，全面支持 DeepSeek-R1/V3 调用。但不能直接使用之前申请的“文心一言”的 API Key 来调用 DeepSeek，需要另外创建百度 V2 API Key。方法是：

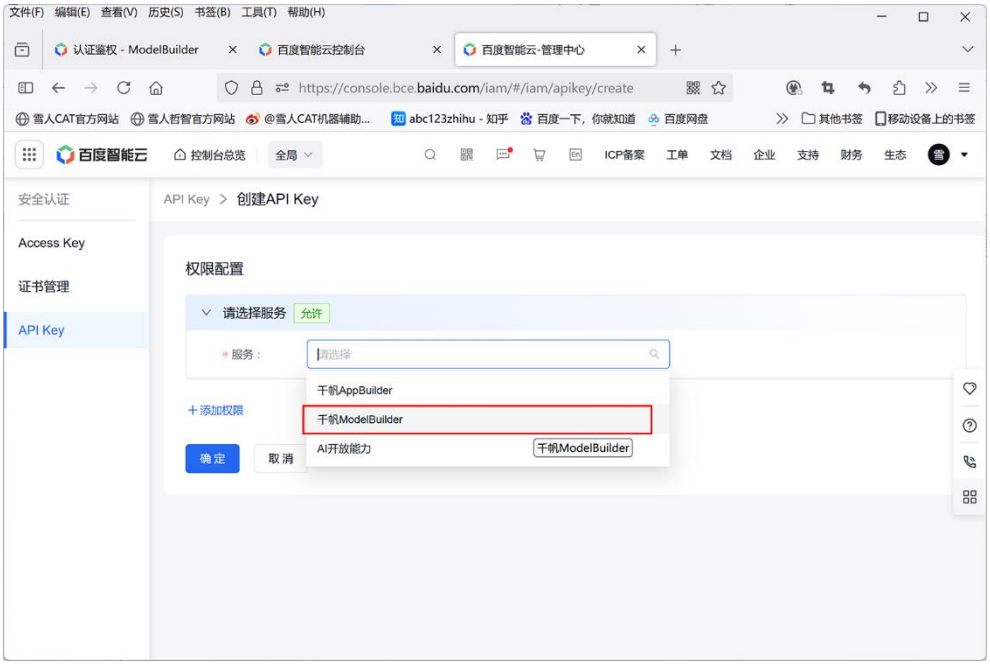
1. 登录百度智能云控制台，https://login.bce.baidu.com
2. 点击右上角的账号图标，会弹出菜单，选择“安全认证”，如下图：



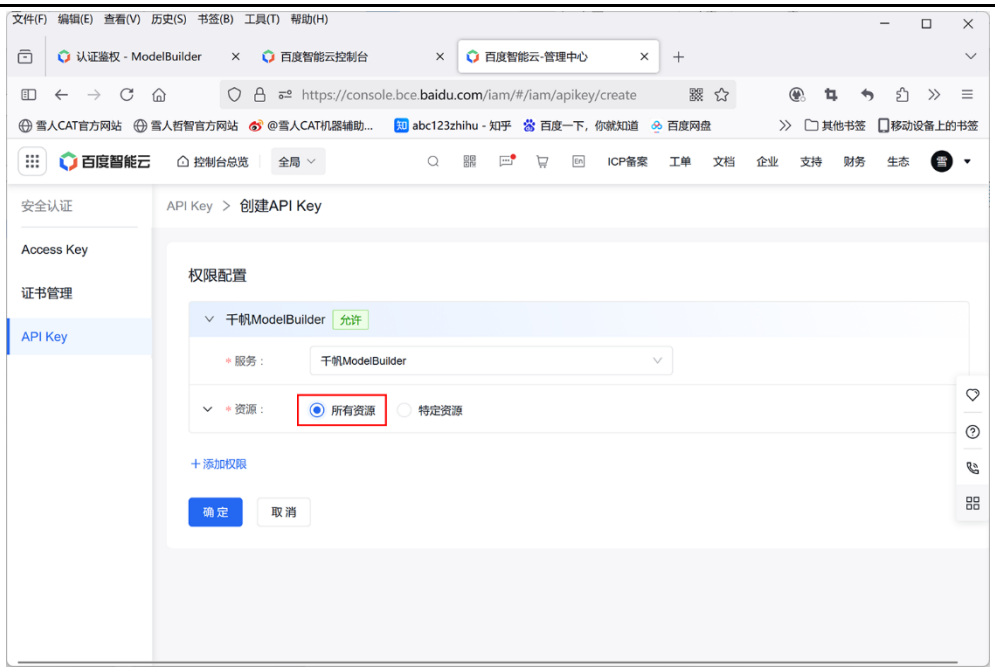
3. 会跳转至如下窗口，点击屏幕左侧的“API Key”,然后点击“创建 API Key”按钮。



4. 在“服务”中选择“千帆 ModelBuilder”，如下图：



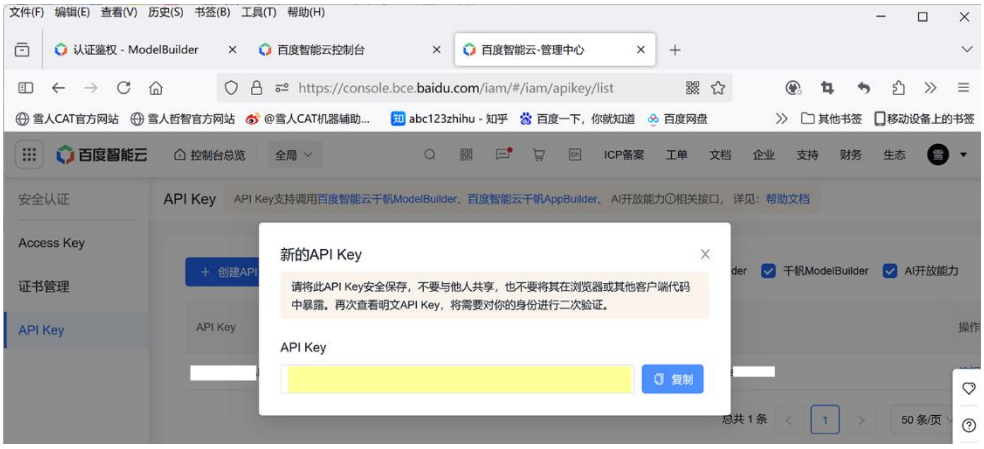
5. 在“资源”中选择“所有资源”这个选项。如下图：



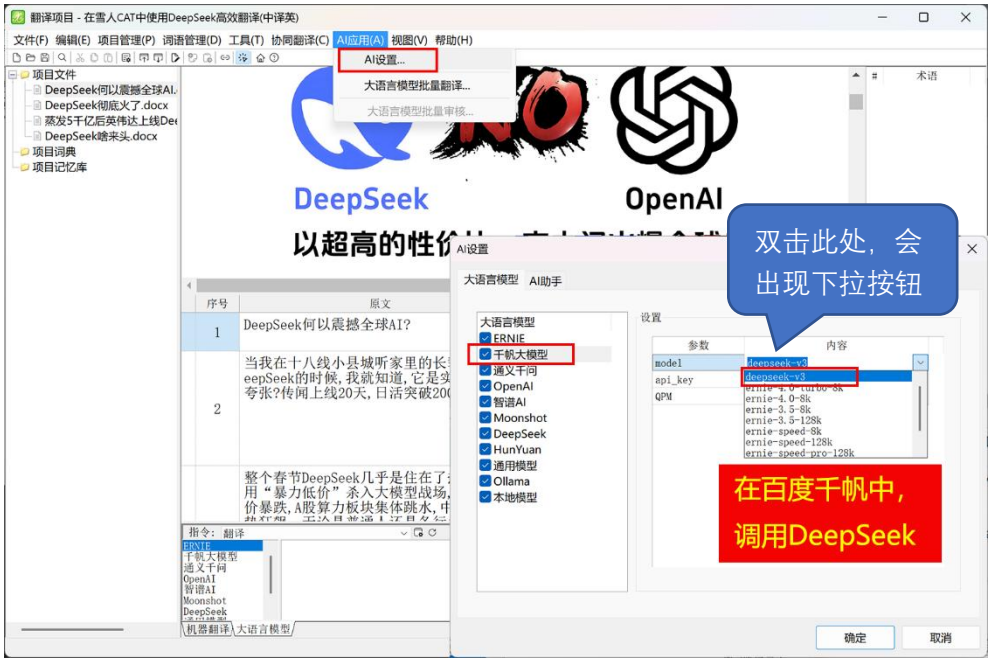
- 所有资源，表示 appid 对应的 V2 版本应用，都可以使用该 API Key。
- 特定资源，表示特定 V2 应用对应的 appid，可以使用该 API Key。

选择“所有资源”这个选项。

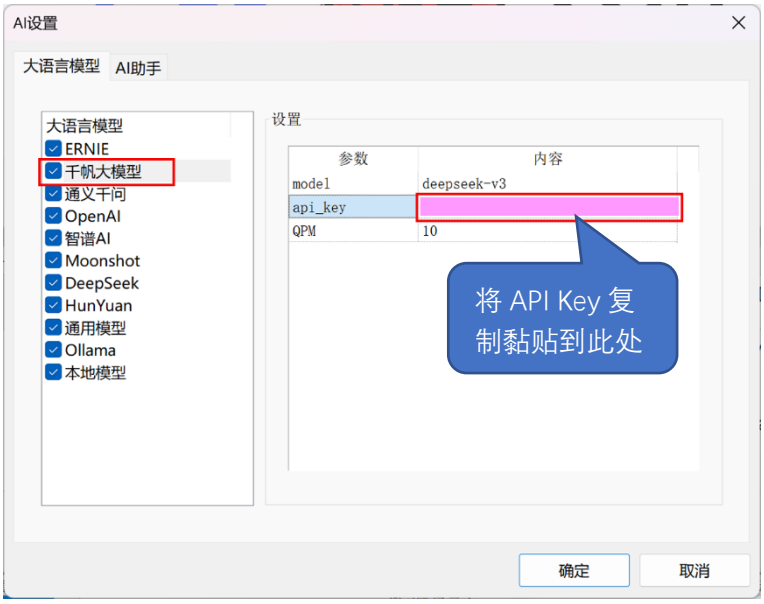
6. 最后，点击“确定”按钮，将会创建一个 API Key，如下图：



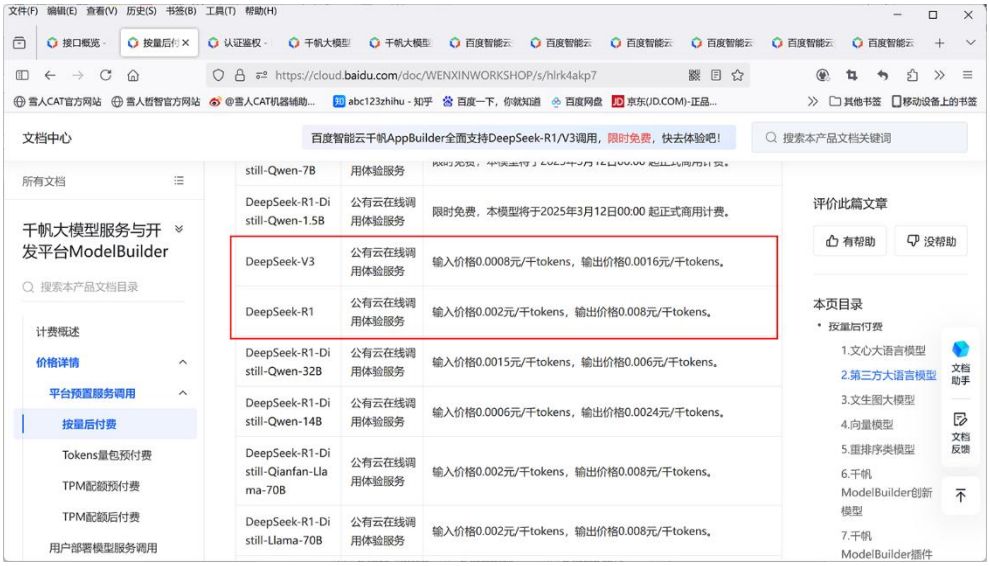
7. 将创建的这个 API Key 复制下来。
8. 切换到雪人 CAT 界面，点击【AI **应用**】➔【**AI 设置**】菜单，在弹出的“AI 设置窗口”中，点击左侧的“千帆大模型”，在右侧的“model”栏处，双击鼠标左键，此时会出现下拉按钮，点击下拉按钮，选择 DeepSeek-V3 即可。



9. 将刚才创建的 API Key 复制粘贴到此处，即可实现在百度“千帆大模型”中调用 DeepSeek-V3。如下图：



在百度中千帆中调用 DeepSeek-V3，速度比较快，而且价格比较便宜。下图是百度的价格，仅供参考。



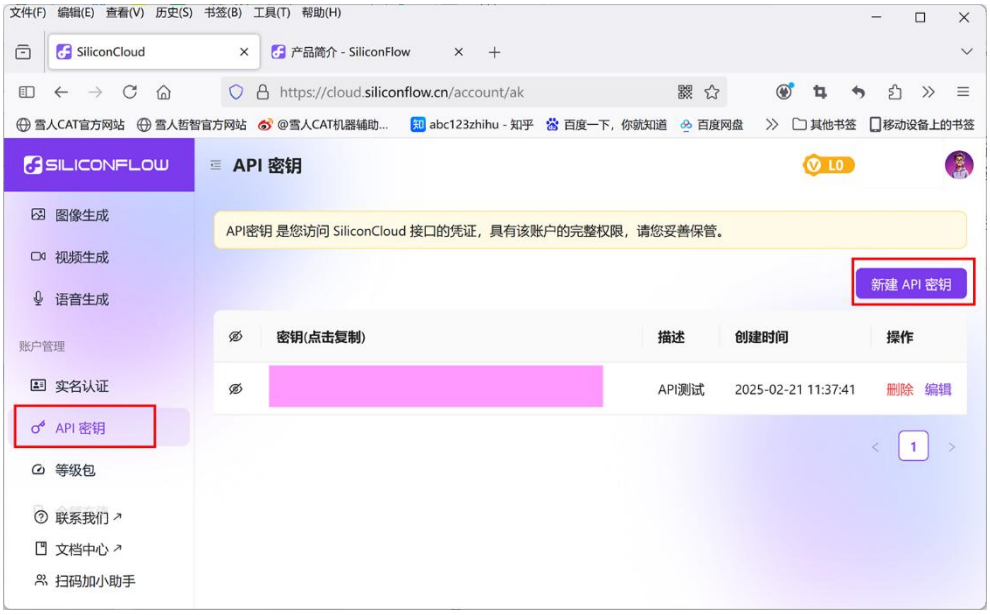
DeepSeek-V3 是通用型模型，专注于自然语言处理、知识问答、内容创作等通用任务。DeepSeek-R1 是推理专用模型，专为数学、代码生成和复杂逻辑推理任务设计，通过大规模强化学习（RL）提升推理能力，对标 OpenAI o1 系列。DeepSeek-R1 会输出许多推理的过程，耗费的 token 数会比较多，而且响应速度也比 V3 慢，作为翻译使用，建议使用 DeepSeek-V3 模型。目前雪人 CAT 默认支持的是 DeepSeek-V3 模型，如果需要使用 DeepSeek-R1 模型，可以用“通用的模型接口”自己设置（后面会讲述）。

第十节 云端 AI 通用模型接口

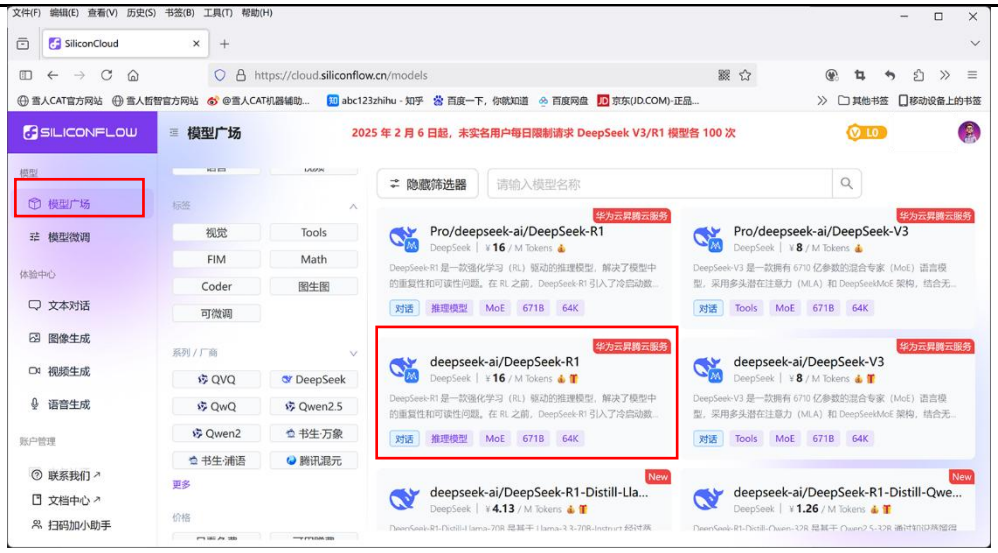
为了更加方便用户使用各种 AI 大模型，雪人支持 AI 通用模型接口，用户可以根据需要，自己设置和选用各种 AI 大模型。只要是采用与 OpenAI（ChatGPT）兼容的 API 接口，都可以利用通用模型接口来调用它。

DeepSeek 的开放让众多云服务厂商迅速行动，纷纷部署了相关模型。让普通人也能轻松接入 DeepSeek。下面以“硅基流动”第三方 AI 模型云服务平台为例介绍在雪人 CAT 中，如何用通用模型接口接入各种 AI 模型？

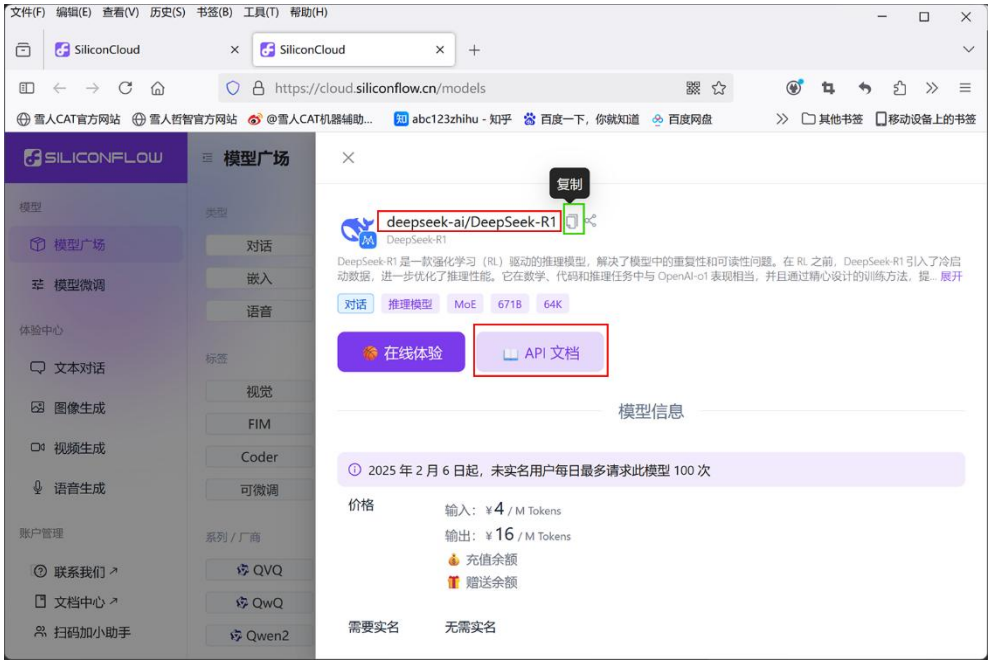
登录 https://cloud.siliconflow.cn/models “硅基流动”模型广场，注册之后，账户里默认会赠送 2000 万 tokens。点击左侧的“API 密钥”，再点击“新建 API 密钥”按钮，创建一个 API 密钥，创建好后将它复制粘贴下来。如下图：



点击左侧的“模型广场”，有许多模型供选择，有包括 DeepSeek-V3、DeepSeek-R1、Qwen 系列等各种模型，有收费的，也有免费的。如下图：

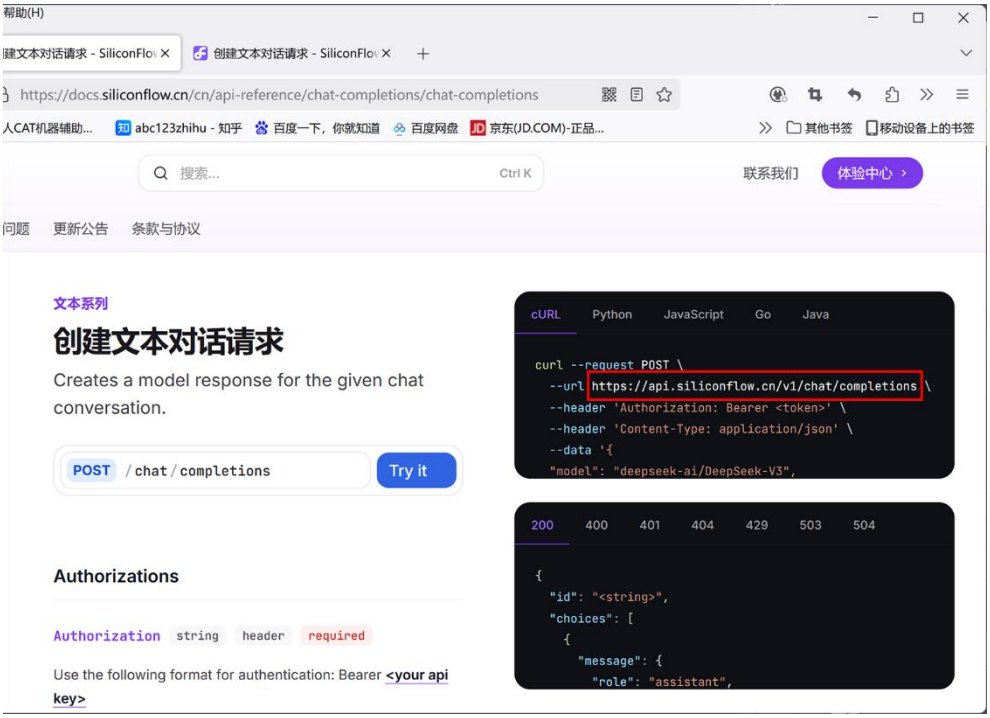


这里以选择“deepseek-ai/DeepSeek-R1”这个模型为例来说明。Pro 版速度会快一下，但是不能使用赠送的流量。点击“deepseek-ai/DeepSeek-R1”这个模型，页面跳转如下图：



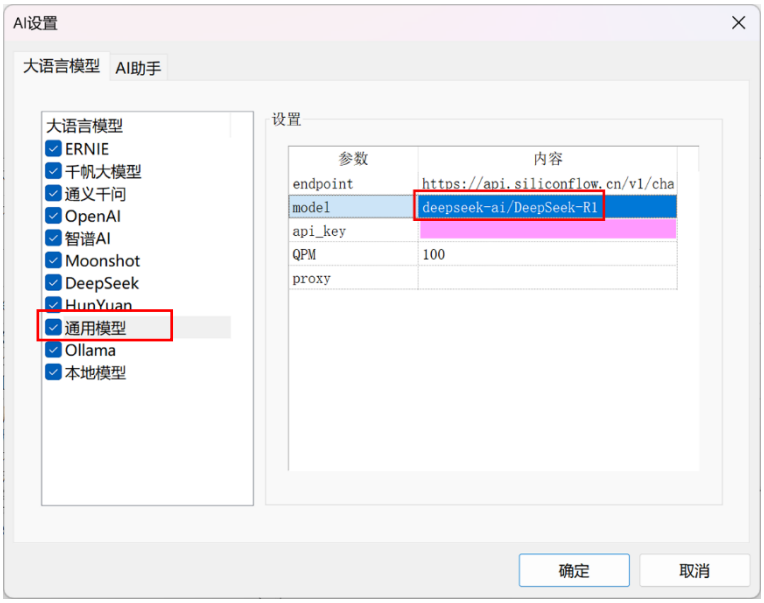
点击该模型名称旁边的“复制”按钮，将模型名称完整复制粘贴下来。例如名称粘贴出来是：deepseek-ai/DeepSeek-R1

点击上图的“API 文档按钮”，页面跳转如下图：



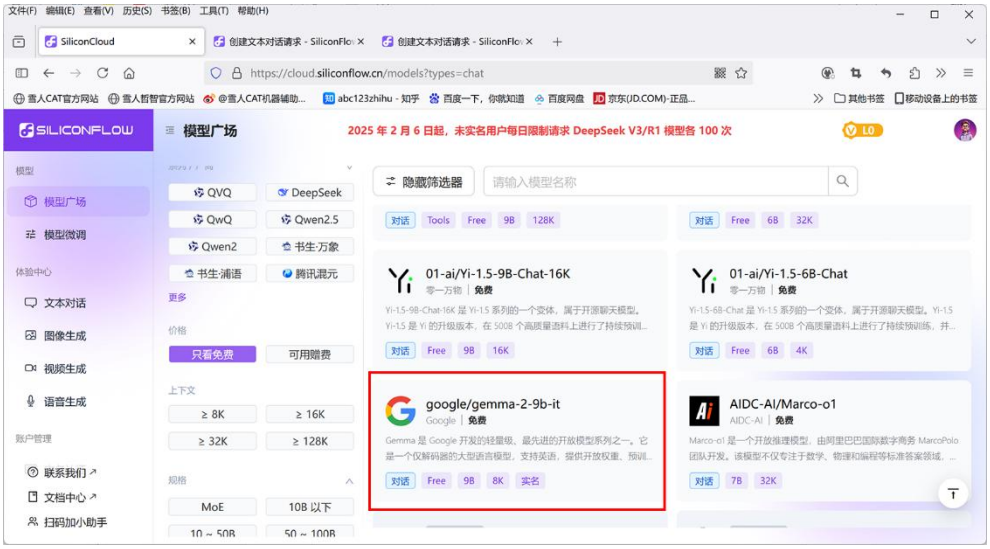
找到模型接入的 endpoint(端点地址，即服务接入地址)，实际上是一个 url 链接地址，要求 http 或 https 开头的地址。“硅基流动”平台模型调用的 endpoint(端点地址)是：https://api.siliconflow.cn/v1/chat/completions 将这个地址复制下来。

切换回雪人 CAT 软件界面，点击【AI **应用**】➔【**AI 设置**】菜单，在弹出的“AI 设置窗口”中，点击左侧的“通用模型”，如下图：

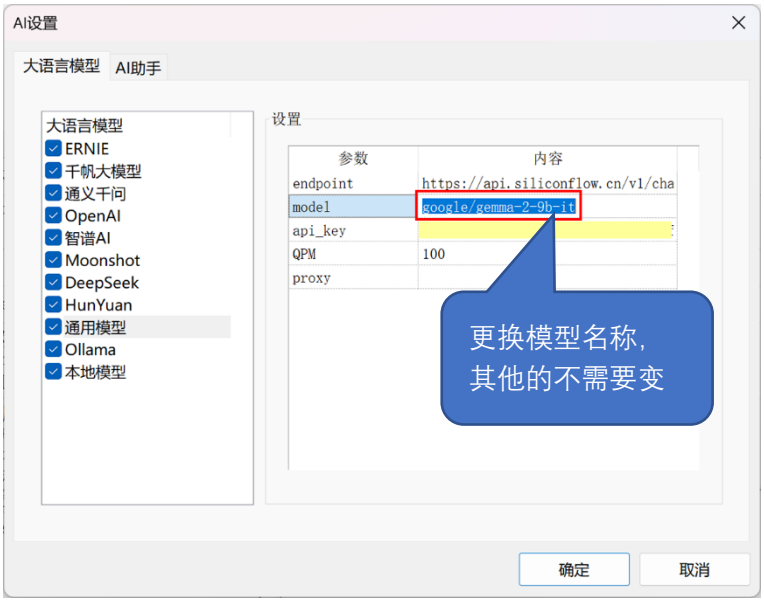


在“endpoint”(端点地址，即服务接入地址)处粘贴刚才复制的 url 地址，即 https://api.siliconflow.cn/v1/chat/completions。在“model”处，粘贴复制下来的模型的名称，即 deepseek-ai/DeepSeek-R1 。在“api_key”处，粘贴申请创建的“api_key”。“proxy”是用来填写代理服务器地址，如果有些平台提供的模型，是需要使用代理服务器才可以使用的，则需要填写其地址，否则不需要填写。

按照上述方法，填写好后就可以在“通用模型”中，调用自己设置的 AI 模型了。如果要更换成“硅基流动”模型广场中的其他模型，方法也非常简单，只要复制该模型的名称填写中雪人 CAT 中即可。例如，下图中，想更换成 google/gemma-2-9b-it 这个免费模型，将其模型名称 google/gemma-2-9b-it 复制下来。



将新的模型名称粘贴到雪人中，如下图：其他的无需变化即可使用新的模型。



注意：“model”处填写的模型名称，必须与接入点的服务提供商给定的模型名称严格一致，包括名称的字母大小写、全角、半角、下划线等等都必须一致，否则可能影响调用不成功。

使用通用模型接口可以实现：在阿里平台调用 DeepSeek，具体操作方法是：在上图的 endpoint 中输入： https://dashscope.aliyuncs.com/compatible-mode/v1/chat/completions

model 中输入：deepseek-v3

api_key 中复制粘贴从阿里申请到的 api_key。

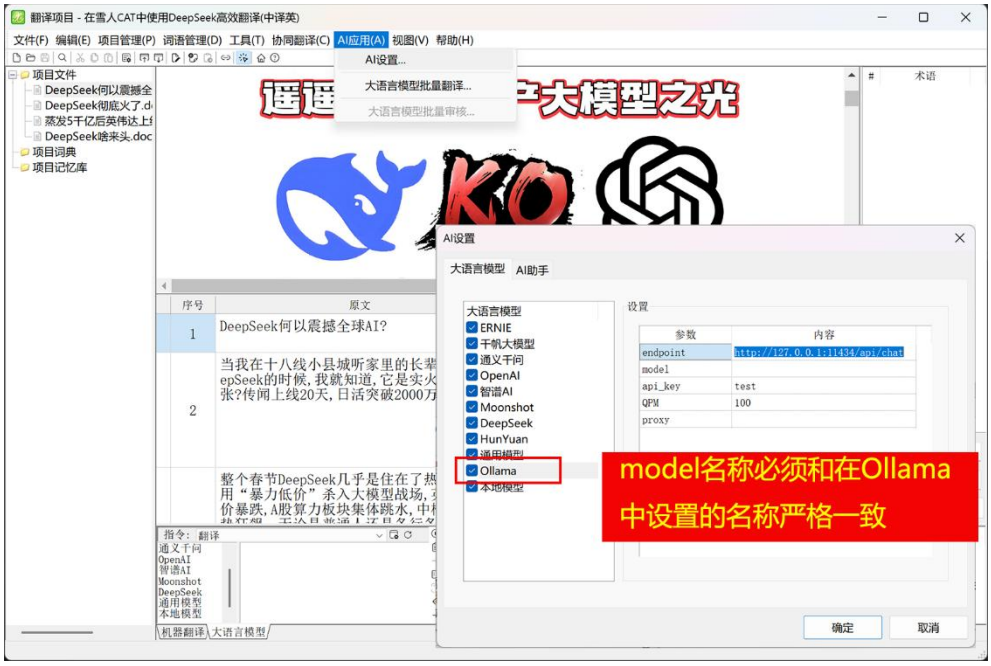
在“阿里”调用 DeepSeek，响应速度比 DeepSeek 官方快，而且价格还比官方的便宜，目前阿里有赠送 100 万 token 的免费额度。

第十一节 Ollama 接口

Ollama 是一个开源的本地化大模型部署工具，旨在简化大型语言模型（LLM）的安装、运行和管理。它支持多种模型架构，适合开发者和企业快速搭建私有化 AI 服务。利用 Ollama 可以在本地部署 AI 大模型，也可也向云提供商租用虚拟云主机，在云主机上部署私有化的 AI 模型。Ollama 的主要特点包括： 轻量化部署：支持在本地设备上运行模型和云端服务器运行。多模型支持：兼容多种开源模型，如 LLaMA、DeepSeek 等。高效管理：提供命令行工具，方便用户下载、加载和切换模型，也支持图形界面。跨平台支持：支持 Windows、macOS 和 Linux 系统。

如果您的业务对数据隐私、定制化需求、实时性或长期成本控制有较高要求，且具备相应的技术能力和硬件资源，那么自行部署 DeepSeek 到本地是一个值得考虑的选择。

在雪人 CAT 软件界面，点击【AI 应用】➔【AI 设置】菜单，在弹出的“AI 设置窗口”中，点击左侧的“Ollama”，如下图：

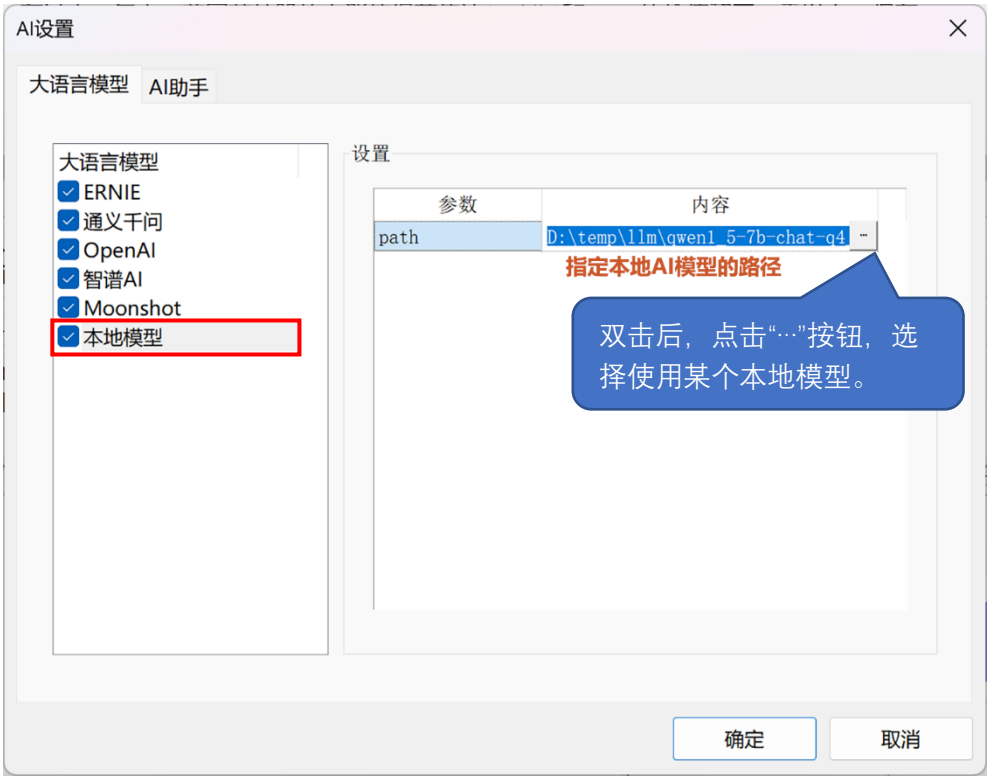


在 endpoint(端点地址)中输入“Ollama”部署的地址，如果是本地部署的，默认的地址是：http://127.0.0.1:11434/api/chat 前面 http://127.0.0.1 是部署在本机的访问地址，如果是部署在单位的局域网内，则是局域网内的访问地址（例如：http://168.0.1.2）。后面:11434/api/chat 保持不变，其中 11434 是“Ollama”的默认端口，如果修改了端口，则更换为新的端口。api key 处，必须填写内容，如果没有，则可随便填写，但是不能为空白。

第十二节 本地 AI 模型的部署

对于有保密需求的用户，不能将翻译的数据对外发送，又希望应用 AI 技术，可以部署本地 AI 模型。本地 AI 模型，省去注册的麻烦，不需要额外的流量费用，如果用户的机器性能可以，本地模型的响应速度比较快，即使对于普通用户，本地模型也是不错的选择。考虑一般普通用户的机器性能，雪人支持接入的本地模型有“千问 1.8B 模型”、“千问 7B 模型”和“零一万物模型”，这些模型是公开的、可以免费使用的，用户可以从网上下载，例如去“通义千问”网站下载或去“雪人翻译软件 AI 交流群”（群号为：526352494）的群共享中下载。模型越大，对机器的性能要求越高。有条件的用户，建议架设专门的服务器来运行本地 AI 模型。如果企业用户，有更好的机器算力，希望使用更大的本地 AI 模型，雪人可以为企业用户定制部署接入更大的本地 AI 模型。

将下载的本地 AI 模型解压后放在一个文件夹内，点击雪人软件主菜单**【AI 应用】➔【AI 设置】**功能，在“AI 设置”窗口中，点击“大语言模型”页签，然后点击窗口左侧的“本地模型”，如下图：



在窗口右侧“path”行处双击鼠标，此时会出现 ··· 按钮，点击该按钮，选择使用某个本地模型，这样就配置好了本地模型了，如下图：普通的机器如果用“7B 模型”觉得慢，就使用“1.8B 模型”。



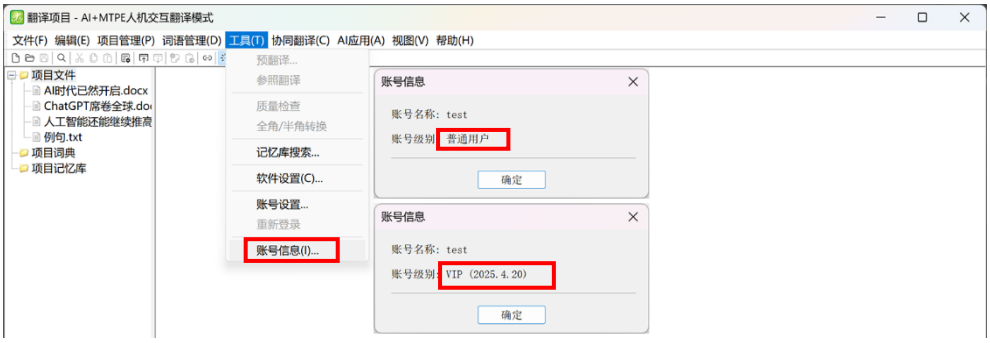
设置好“本地模型”后，点击屏幕下方的“大语言模型”页签，点击该窗口左侧的“本地模型”，就可以用“本地模型”进行翻译、校对等操作了，使用方法同前面的各个大语言模型一样。下图是来自“本地模型（1.8B 模型）”对“昨日那种被追捕的恐慌，已经烟消云散。His panic of the previous evening was quite gone now.”这一句的校对结果，本地 1.8B 模型对这一句的校对结果是：“生动且准确，没有发现明显的错误，保持原译”。



使用“本地模型”第一次加载时需要等待一些时间，可能几秒、十几秒，不同机器由于性能不同，第一次加载需要的时间也不同。

第十三节 升级为 VIP 用户

AI 应用功能需要 VIP 权限，普通用户使用不了这部分的功能。点击主菜单**【工具】➔【账号信息】**功能，可以查看当前用户的账号级别，升级为“VIP 用户“”后，在“账号信息”窗口中，可以看到 VIP 的使用时长。如下图：



如果是普通用户，点击主菜单**【AI 应用】➔【AI 设置】**功能时，会弹出“AI 支持需要 VIP 权限”窗口，如下图：点击窗口中的“帮助”按钮，会链接到雪人官网(http：//www.gcys.cn)开通 VIP。



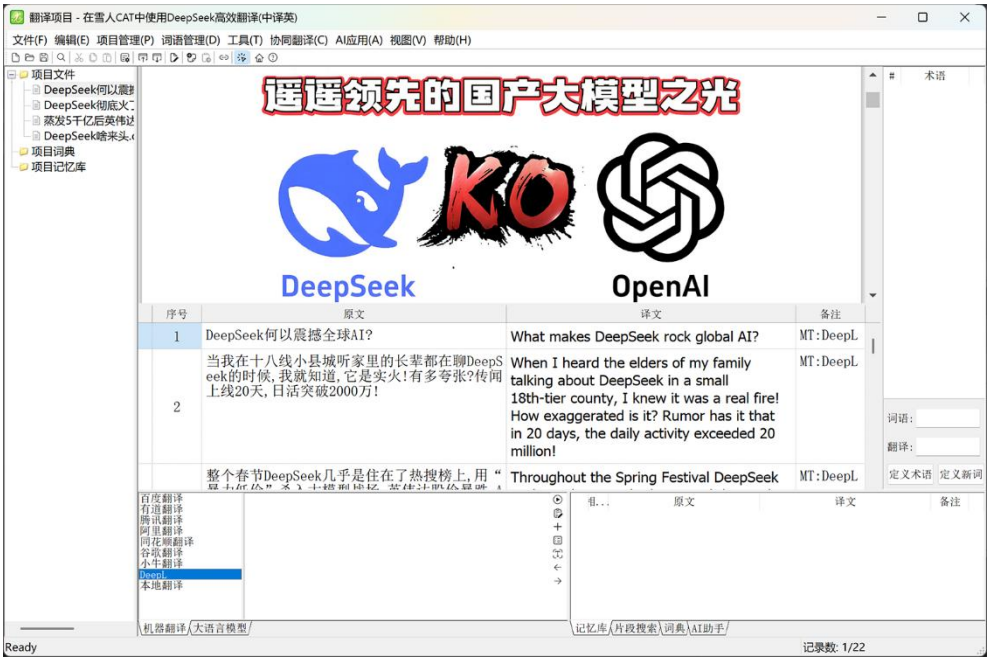
第十五章 审校与 AI 批量审校

第一节 审校

支持三种项目文件格式，分别是：翻译项目（*.spr）、审校项目（*.spp）和修订项目（*.spa）。无论采用人工审校方式，还是基于 AI 大语言模型的审校方式，均能确保流程清晰、责任明确，便于追溯。

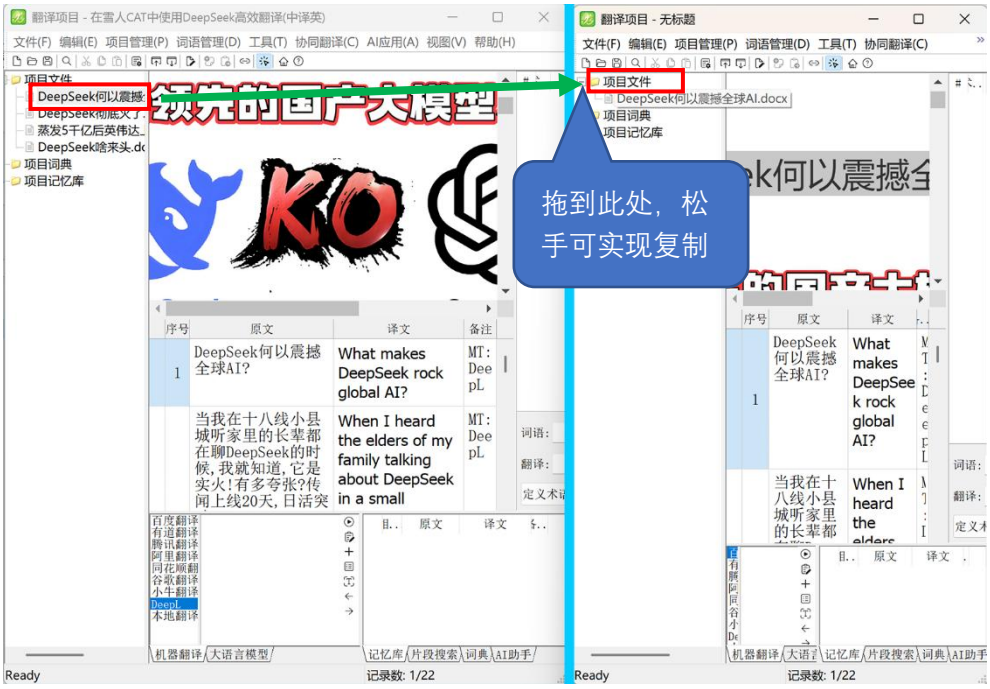
下面以“DeepSeek 何以震撼全球 AI?”的文档为例，介绍审校、修订的各项功能及流程。

由于雪人的一个翻译项目（*.spr）文件中，支持导入数十份甚至上百份需要翻译的文档，而当前需要审校的可能是其中的某一个文档或几个文档，这种情况下将翻译项目（*.spr）中的部分文件发给审校人员就可以了。例如：下图中的翻译项目（*.spr）包含 4 个翻译文档，而当前需要审校的是“DeepSeek 何以震撼全球 AI?”这个文档。如下图：

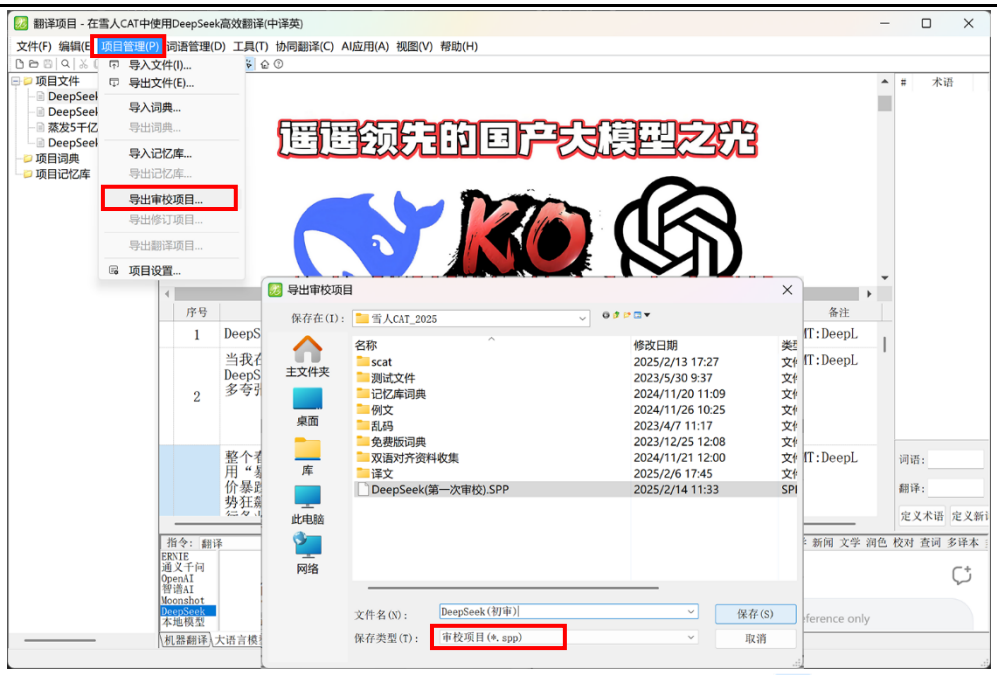


解决办法：新建一个翻译项目（*.spr），将新建的翻译项目窗口和原来的窗口缩小后平铺在桌面，用鼠标左键点击“DeepSeek 何以震撼全球 AI?”这个文档，拖动到新窗口的“项目文件”夹下面，此时，鼠标会变成带有“+”号的图标，松手后即可实现复制（如下图），然后将这个新的项目文件取名保存起来，接下来就针对这个新项目文件进行操作。当然，如果整个项目文件都需要发给他人审校的，则不需要

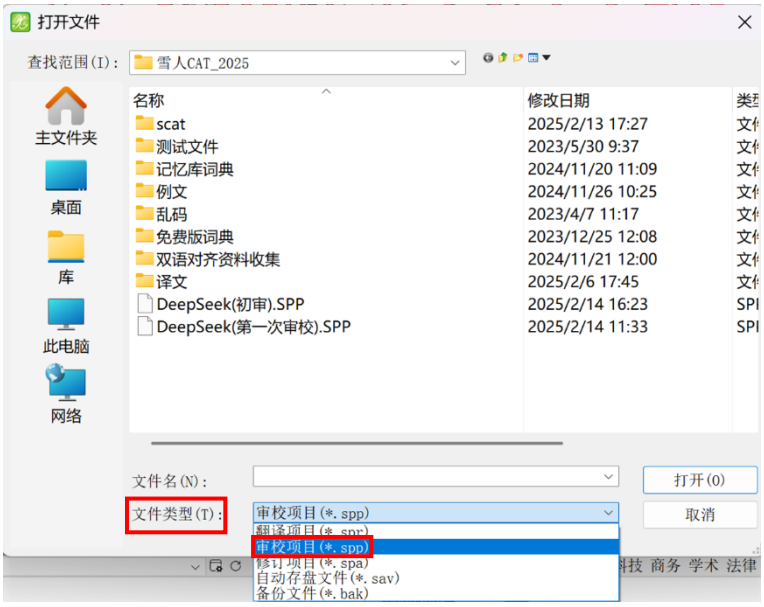
如此操作。



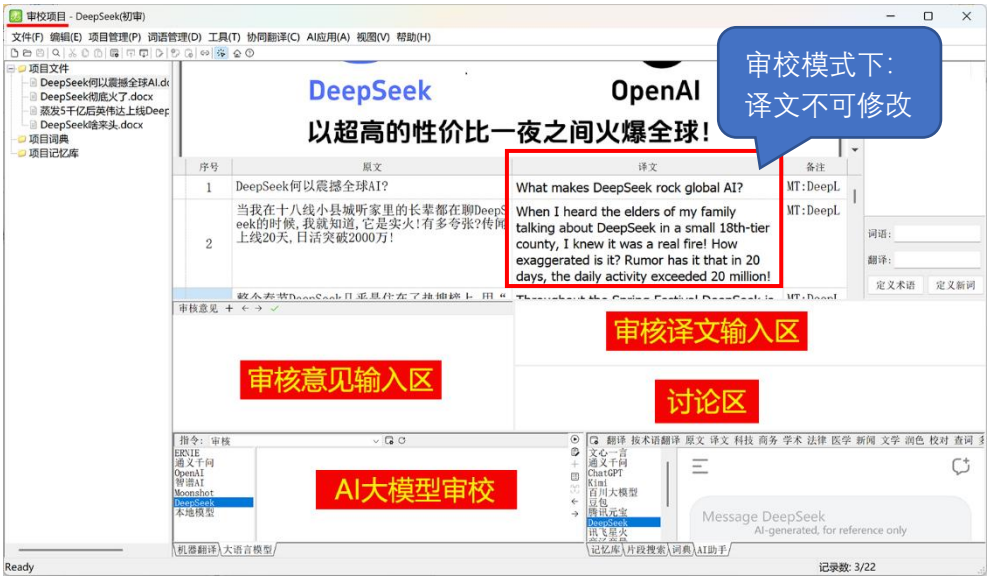
译员将需要审校的翻译项目文件（*.spr）发送给项目组长，组长收到文件后打开文件，点击**【项目管理】➔【导出审校项目】**菜单（如下图），将当前的翻译项目（*. spr）文件导出生成一个审校项目文件（*. spp），将这个“审校项目文件（*. spp）”发送给具体的审校人员。



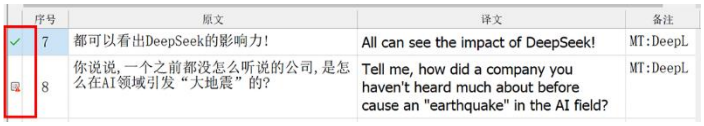
审校人员收到这个“审校项目文件（*.spp）”后，点击工具栏打开按钮，在弹出的“打开文件”窗口中，点击“文件类型”右边的下拉框，选择“审校项目（*.spp）”类型，就可以看到这个文件，如下图。



双击将它打开，审校模式下的软件界面如下图：在屏幕的中部，增加了三个区域，分别是：审核意见输入区、审核译文输入区和讨论区。



在审校模式下，不可以修改原来的译文，可以在下方的“审核译文输入区”中输入修改后的译文。审校认为正确的句子，可以点击绿色的✓核准按钮，将会在该句的序号前面打上一个✓，对于打了✓的句子，译员不能再对其译文进行修改。对于有审核意见及审核译文的句子，在该句的序号前面会显示图标，如下图:



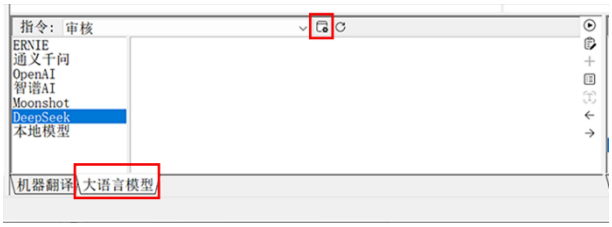
点击＋“添加消息”按钮，可以在讨论区中输入讨论信息。

第二节 AI 逐句审校

利用 AI 大语言模型对翻译后的译文进行校对具有显著优势，它能够快速识别并纠正语法错误、拼写错误和用词不当，确保译文的准确性和流畅性，极大地提升校对效率，节省时间和精力。尤其是 DeepSeek 的横空出世，引爆了全球 AI 市场，也大幅降低了 AI 的使用成本，DeepSeek 的开源策略，使得企业本地部署也成为可行。

下面将讲述如何利用 AI 大语言模型审校来?

使用 AI 大语言模型审校的前提是：需要先设置好这些 AI（点击**【AI 应用】➔【AI 设置】**菜单，具体见 AI 设置章节），然后点击屏幕左下方的“大语言模型”页签（如下图）：



点击窗口上方的设置按钮，会弹出“大语言模型”提示语输入窗口，如下图：



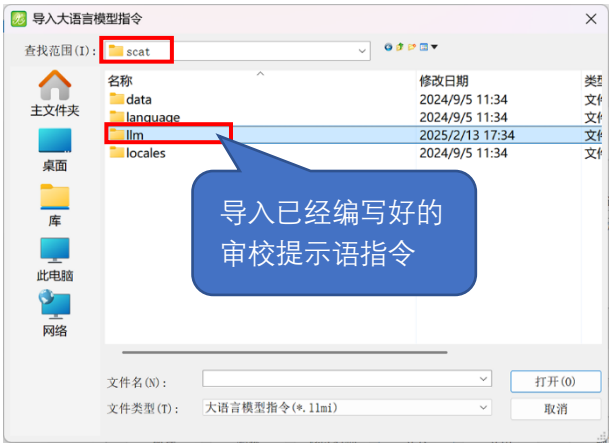
点击“新建”按钮，创建一个“审校”提示语指令，具体指令的编写参考如下：

你是一位资深的专业翻译审校人员，请检查出译文中可能存在的错误，包括：拼写错误、词汇错误、句法错误、语义错误、逻辑错误、标点符号错误、数字翻译错误、漏译等等，译文的表达是否符合目标语的习惯？译文是否准确、完整传达原文的信息？并针对错误予以一定的分析与评估，再用地道的母语将它修改为正确的译法。修改后的译文使用<Result>括起来。

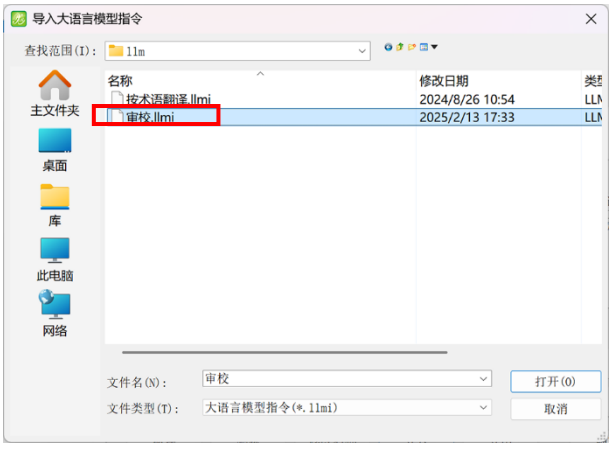
{{SourceLanguage}}原文：{{SourceSentence}}
{{TargetLanguage}}译文：{{TargetSentence}}

勾选“生成译文”选项，并在“匹配”输入框中输入正则表达式：<Result>\s*(.+)\s*</Result> ，“替换”输入框中输入：$1 ，目的是为了能自动抽取 AI 修改后的译文，尤其是当使用 AI 进行全文批量校对时，希望能自动抽取 AI 修改后的译文，并能自动填写在“审核译文输入区”中。

如果觉得自己编写审校提示语指令比较麻烦，可以导入我们已经编写好的指令。点击上图中的“导入”按钮，找到“scat”文件夹，双击“llm”这个存放提示语指令的的文件夹，如下图：



双击“审校.llmi”这个指令文件，导入它。如下图：



如果是第一次用 AI 大语言模型功能的用户，系统会默认内置这个审校指令，不需要导入。其他用户，为了避免和用户已有的指令冲突，则需导入或自己创建。

在“大语言模型”窗口左侧选择具体使用的“大语言模型”，例如点击“DeepSeek”，选择“审核”指令，点击右侧的▶启动按钮，就会从“DeepSeek”返回该

句的审校结果。如下图:



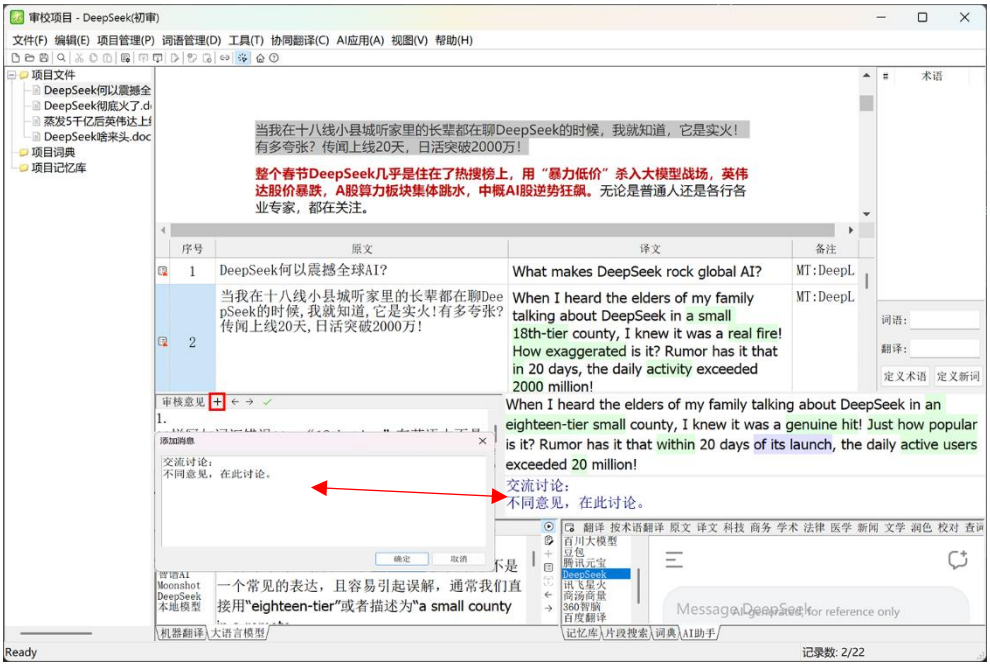
点击右侧的拷贝按钮，可以将 AI 的审校意见复制到上方的“审核意见输入区”，并自动提取 AI 修改后的译文填写在上方的“审核译文输入区”中。

点击“下一句”按钮，可以审核下一句，如下图:

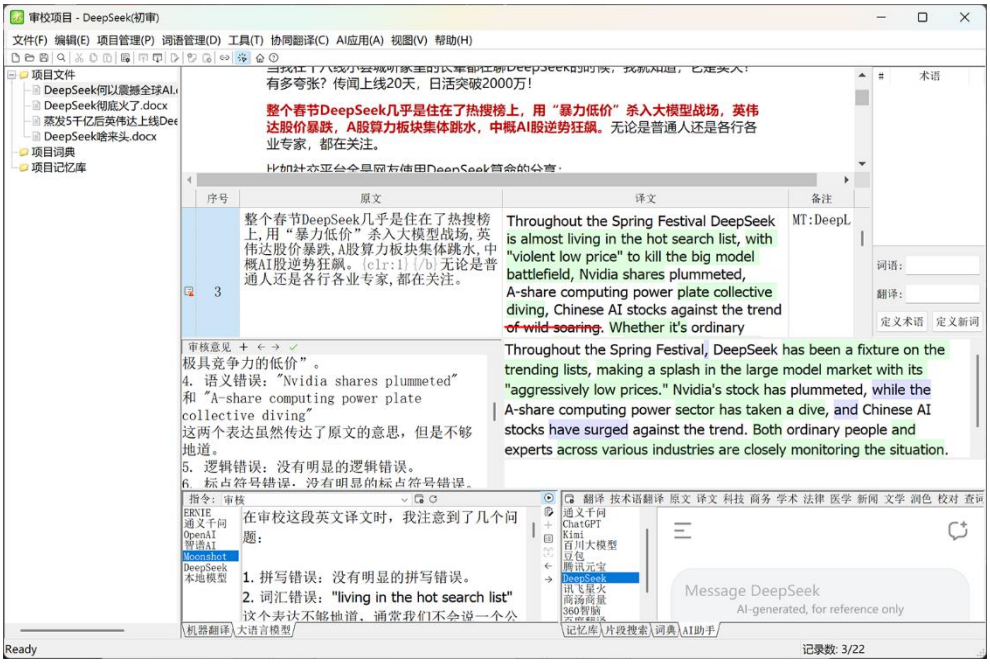


点击 + “添加消息”按钮（如下图），可以在“讨论区”中输入讨论的信息，会不

断滚动记录关于当前句子的讨论内容，已经发送的讨论信息，不可以删除。



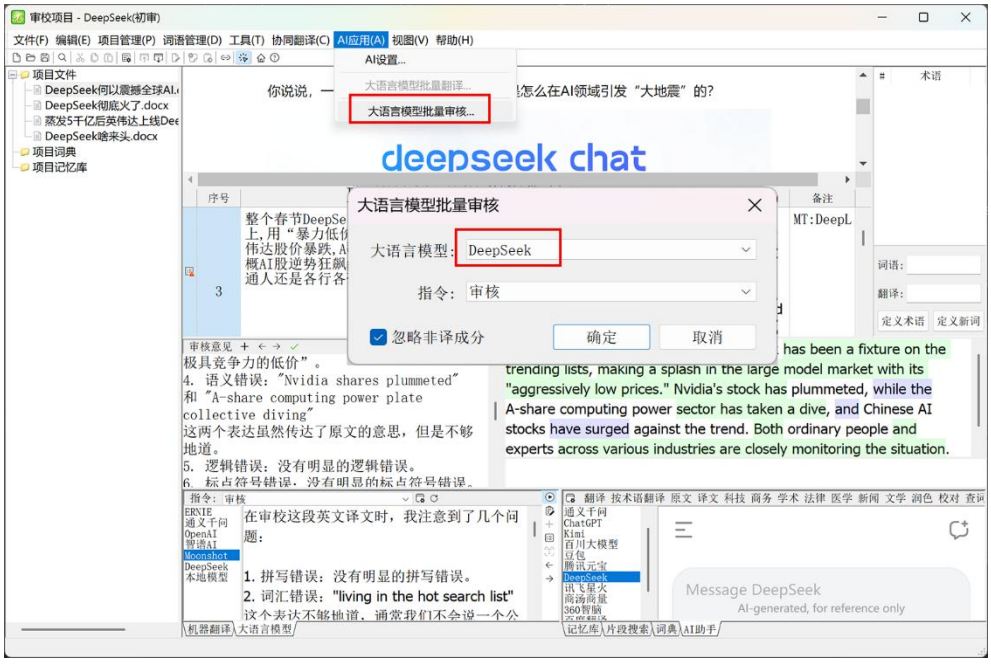
审核译文和原始译文的不同之处会用不同的颜色、删除线等自动标识，绿色表示修改为不同的词汇，紫色表示增加的词汇，删除的词汇会划上红色的删除线。如下图：



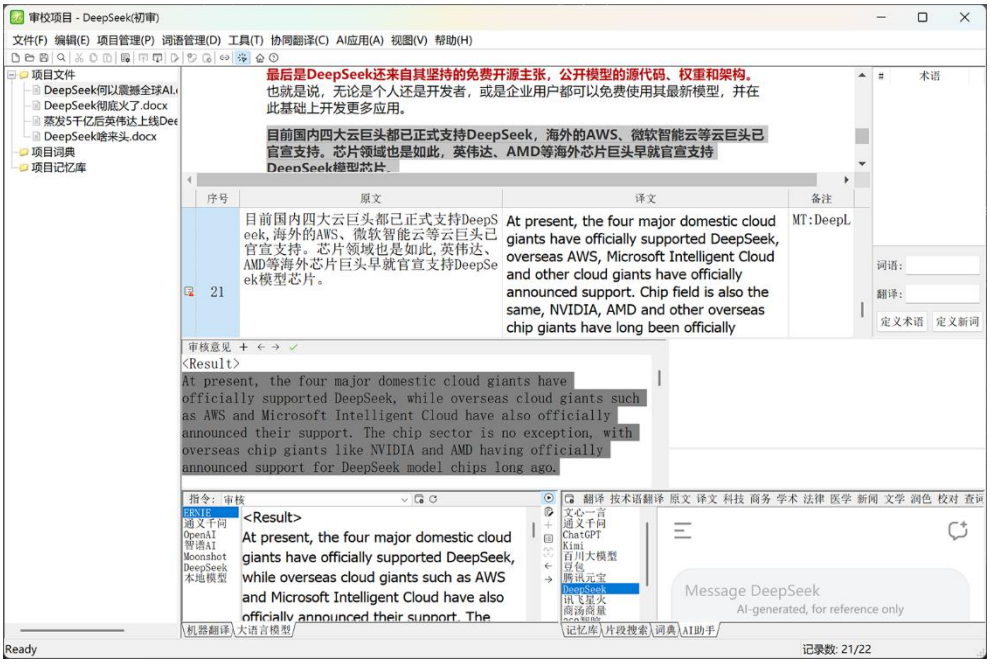
第三节 AI 批量审校

有些 AI 大语言模型，由于服务器比较繁忙，有时逐句审校可能需要等待，此时，可以使用 AI 批量审校功能，执行该功能后，耐心等待 AI 审校完成。点击**【AI 应用】➔【大语言模型批量审校】**菜单（如下图）：

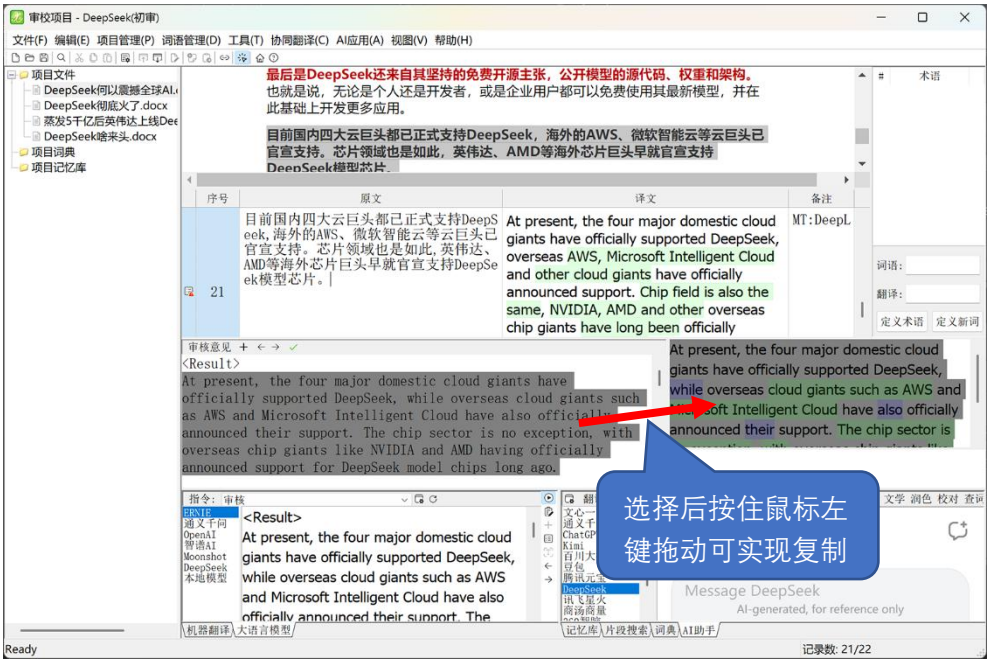
点击“大语言模型”下拉框，选择具体使用哪个大语言模型来审校？目前支持的大语言模型有：DeepSeek、ERNIE(文心一言)、通义千问、OpenAI(ChatGPT)、智谱、Moonshot(Kimi)、本地部署的 AI 等等。这里选择 DeepSeek，再选择“审校”指令，点击“确定”按钮后，就可以进行批量审校了。软件会自动将各句的 AI 审校意见填写在“审核意见输入区”中，同时抽取 AI 修改后的译文自动填写在“审核译文输入区”中，AI 译文和原始译文不同之处，会以不同的颜色及红色删除线表示。



有些 AI 大模型，对指令的遵循不太好，可能会造成 AI 的审核译文不能正确抽取的问题，如下图：右侧的“审核译文输入区”中没有自动填写译文。

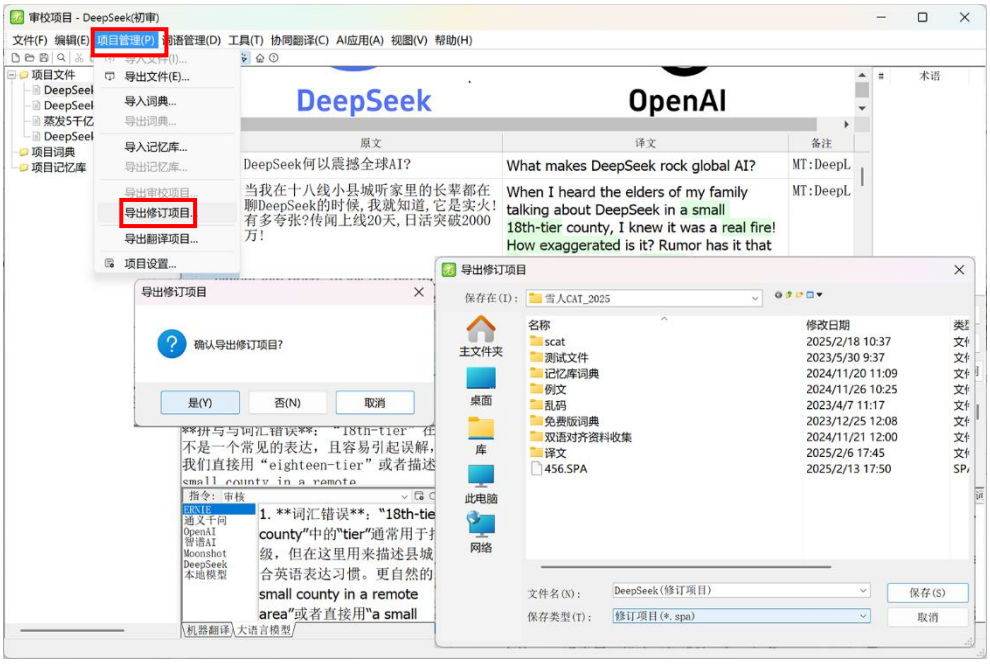


此时，可以用鼠标在左侧的“审核意见输入区”中，选择 AI 修改后的译文，拖到右侧的“审核译文输入区”中。

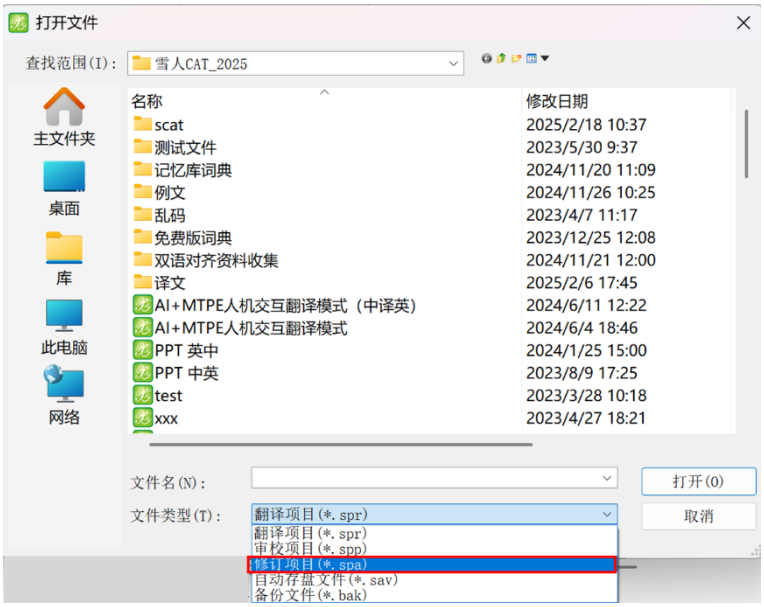


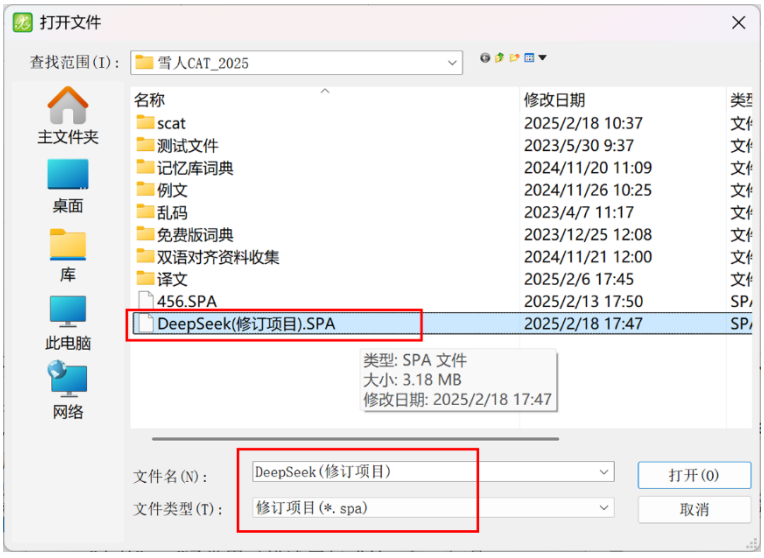
第四节 导出修订项目

审校完成后，将这个审校项目（*. spp）发送给项目组长，组长收到审校项目文件后打开它，点击**【项目管理】➔【导出修订项目】**菜单（如下图）：

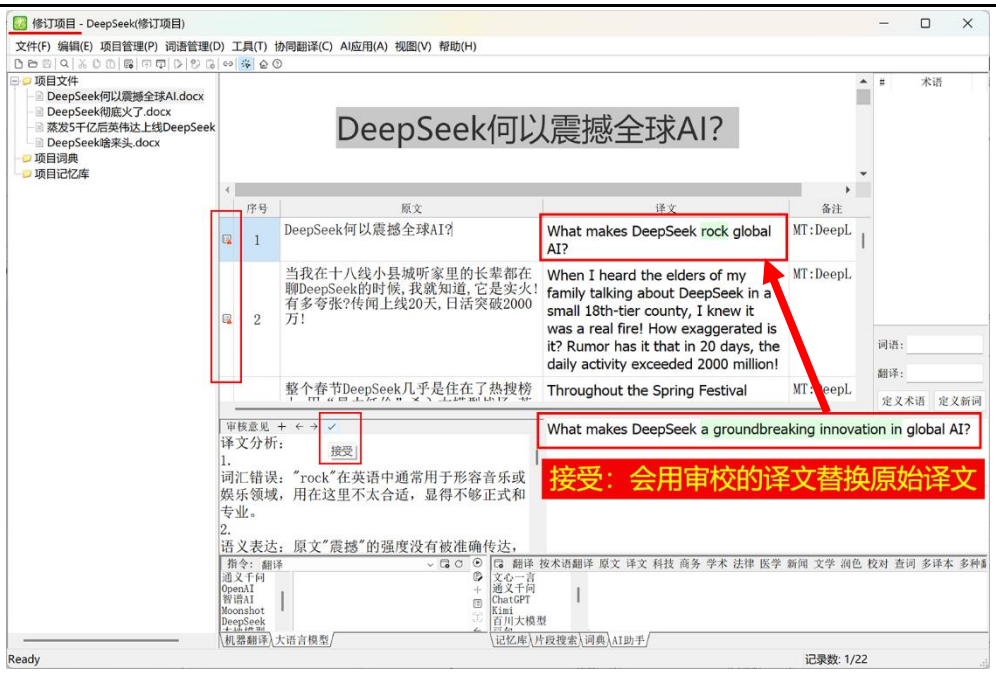


“导出修订项目”将生成一份新的项目文件，后缀为（*. spa），然后将这个修订项目文件发送给译员。译员收到这个修订项目文件后，点击工具栏打开按钮，在弹出的“打开文件”窗口中，点击“文件类型”右边的下拉框，选择“修订项目（*.spp）”类型，就可以看到这个文件，如下图。

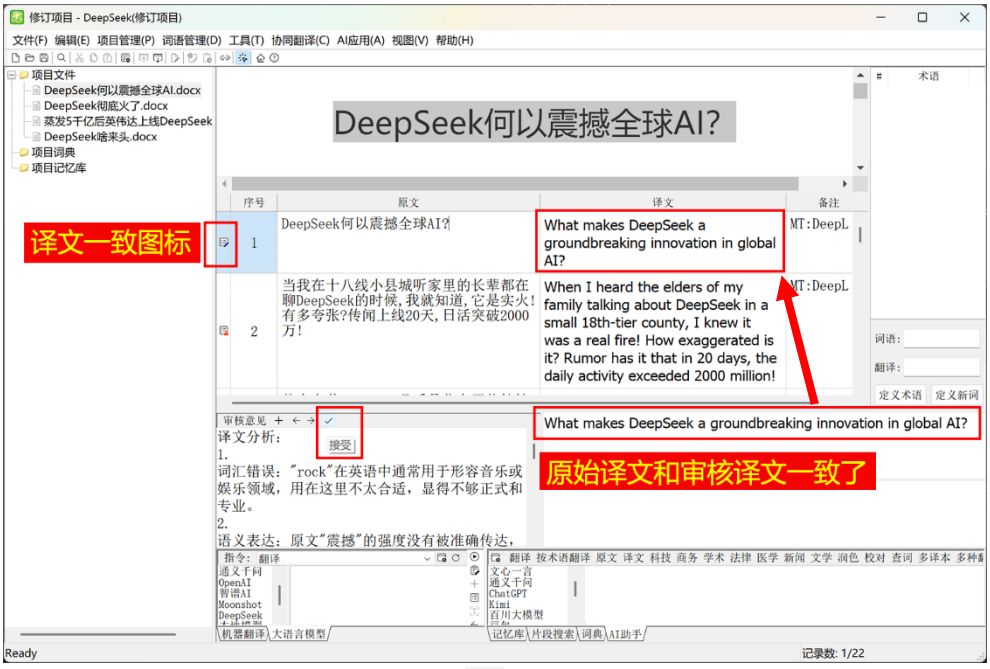




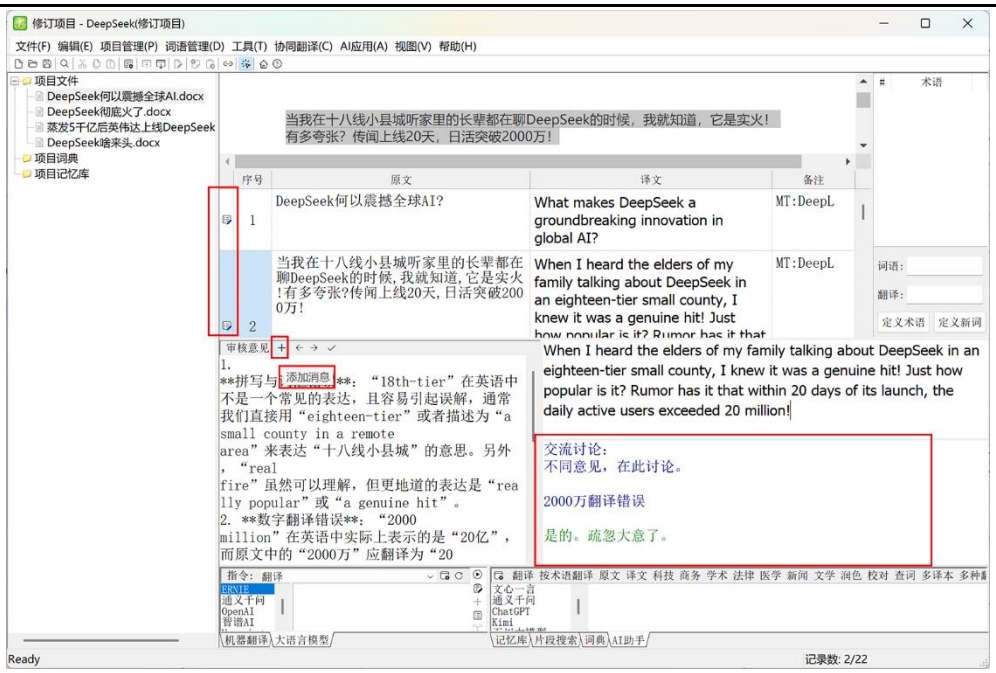
双击打开这个修订项目，在序号前面有图标的，表示有审校信息，点击这些句子，可以看到原始译文和审校译文的差异，绿色表示修改为不同的词汇，紫色表示增加的词汇，删除的词汇会划上红色的删除线。点击✓接受按钮，表示同意用审校译文替换原始译文。注意：在审校项目模式下，这个✓是绿色的（✓），表示审核通过，一旦打了✓的句子，译员不可以再次修改其译文。在修订项目模式下，这个✓是黑色的（✓），表示接受审校的译文，点击它就会用审校的译文替换原始的译文。



接受审核译文后，该句序号前面的图标会发生变化，由 图标变成 ， 图标表示原始译文和审核的译文一致。



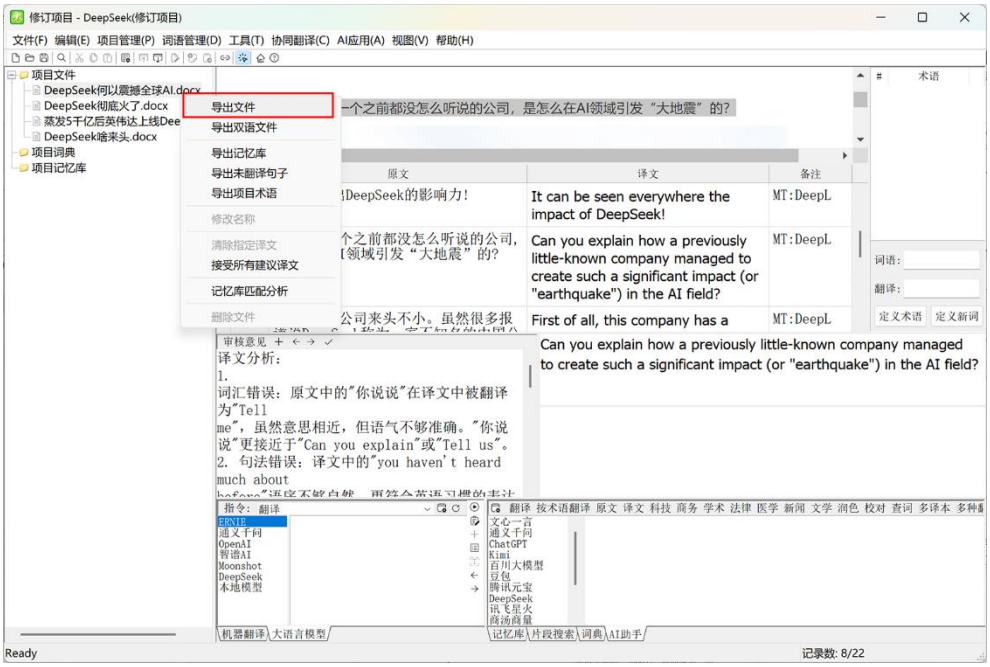
译员在修订项目界面，也可以点击 “添加消息”按钮（如下图），在“讨论区”中输入讨论的信息与审校进行交流。



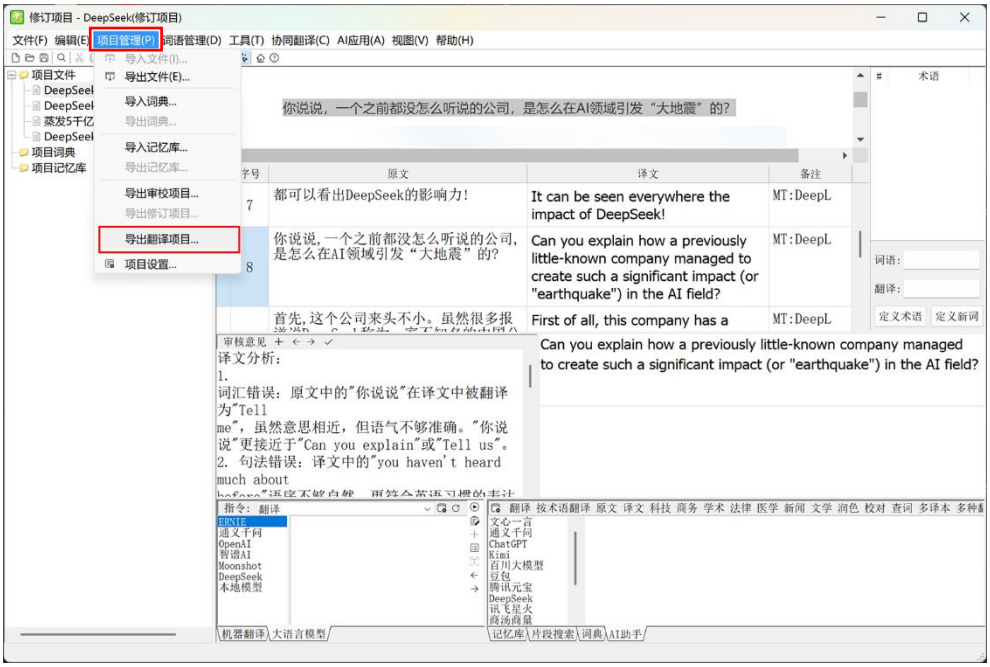
若觉得逐句接受审校译文太慢，可以在屏幕左侧点击当前文档后，点击**【右键菜单】➔【接受所有建议译文】**菜单（如下图）：一次性地接受所有的审核译文，即用审核译文替换所有的原始译文。稳妥起见，可以将当前文件另存一份后再进行“接受所有译文”的操作；如果接受后又后悔了，还可以通过放弃存盘来达到回退到之前的状态（点击点击**【文件】➔【关闭】**菜单，关闭文件时选择“不保存”）。



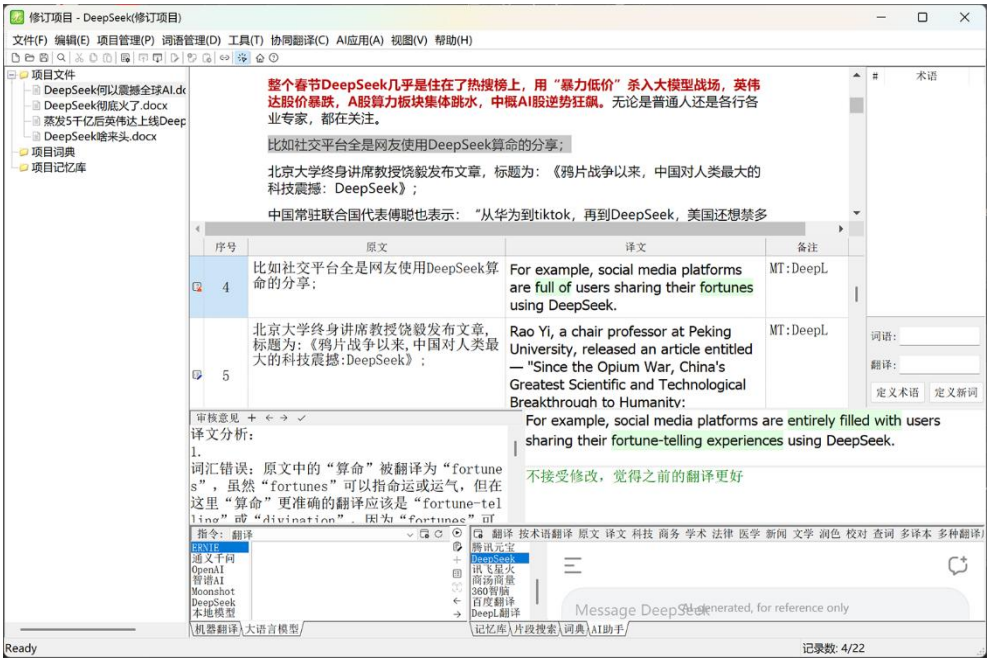
修订完成后，点击**【右键菜单】➔【导出译文】**菜单，可以将译文导出。译员也可以将这个“修订项目”（*. spa）文件发送给组长，由组长导出译文。如下图：



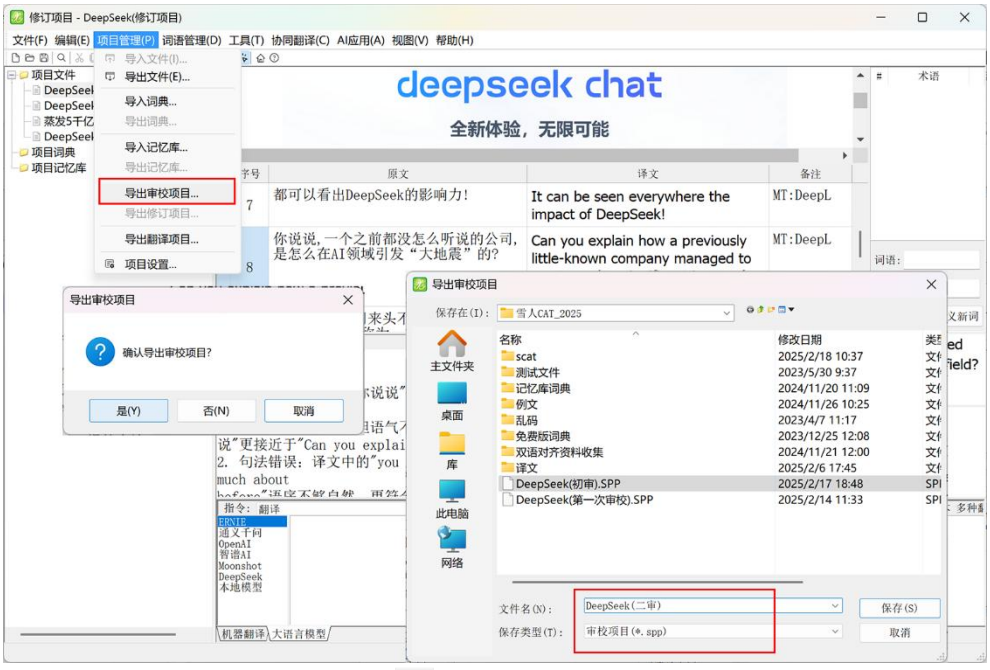
由于修订项目（*. spa）有些功能受限制，修订结束后，可以将它导出成新的翻译项目（*. spr），点击**【项目管理】➔【导出翻译项目】**菜单，导出为翻译项目后，通过文件拖放复制功能，可以进行归类、整理或并入自己以前的翻译项目中。



对于不接受的译文，可以在“讨论区”中输入不接受的原因，保留自己的原始译文，然后将这个修订项目再发送给项目组长，如下图：

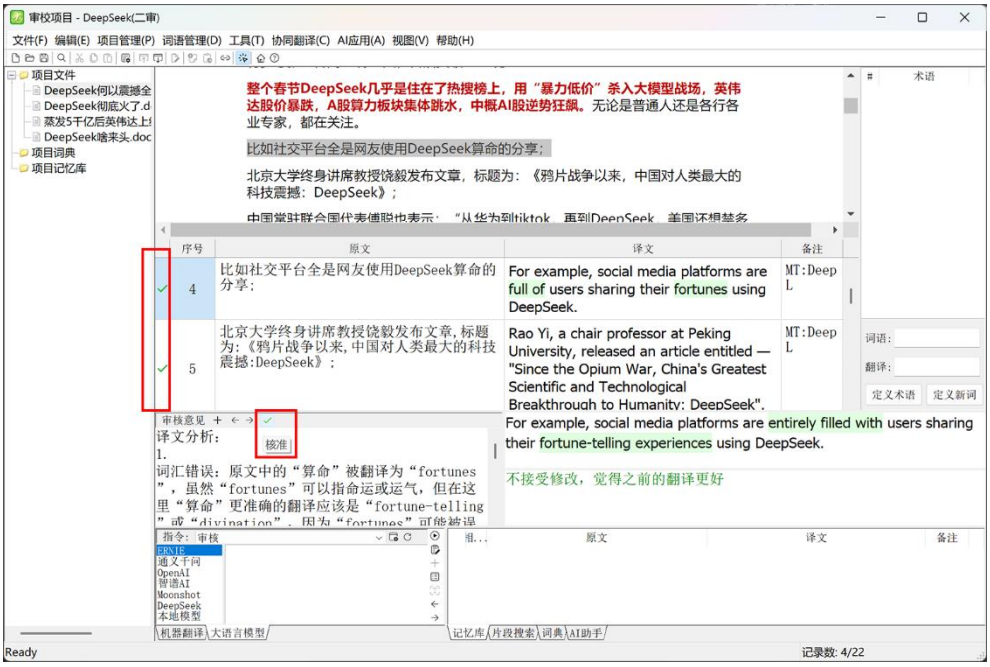


组长收到后打开它，点击【**项目管理**】➔【**导出审校项目**】菜单，生成新的审校项目文件（*. spp），发送给审校进行二次审校。如下图：



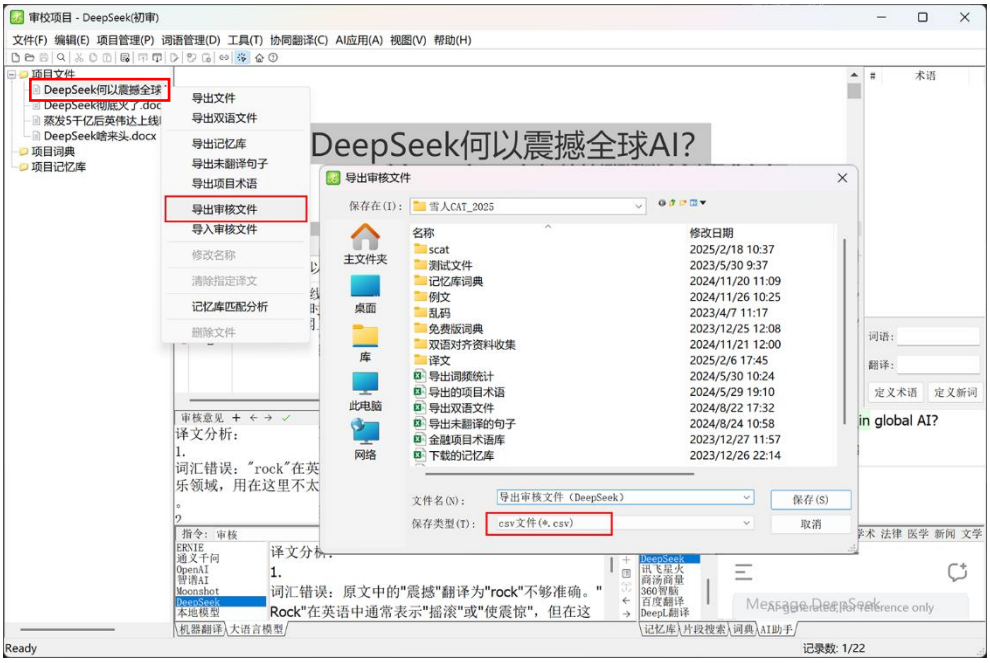
对审核无误的句子，可以点击 ✓核准按钮将其打✓，表示审核通过，一旦打了

✓的句子，译员不可以再次修改其译文。如下图：



第五节 外部审校

点击左侧某具体文档后，点击【右键菜单】➔【导出审核文件】，如下图：

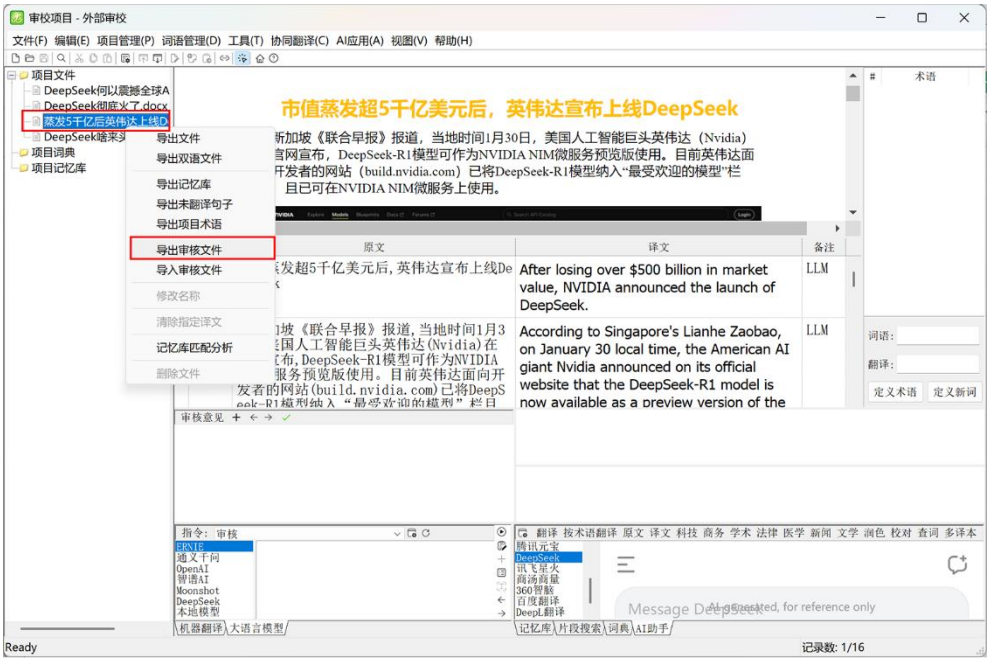


可以将雪人 CAT 中的审校信息等内容导出成“.csv”格式的文件，这个“.csv”文件可以在 EXCEL 中打开它。如下图：

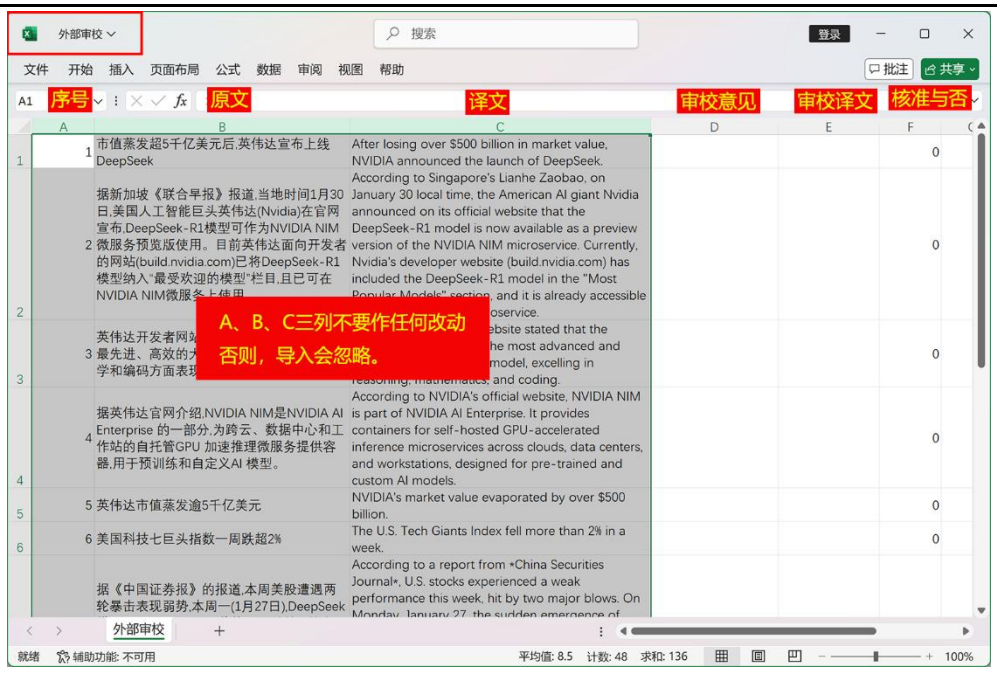


A 列为序号，B 列为原文，C 列为译文，D 列为审校意见，E 列审校译文，F 列为核准与否的状态（0 为未核准，1 为已核准，已核准的句子在雪人 CAT 中的序号前面会打上绿色的√）。

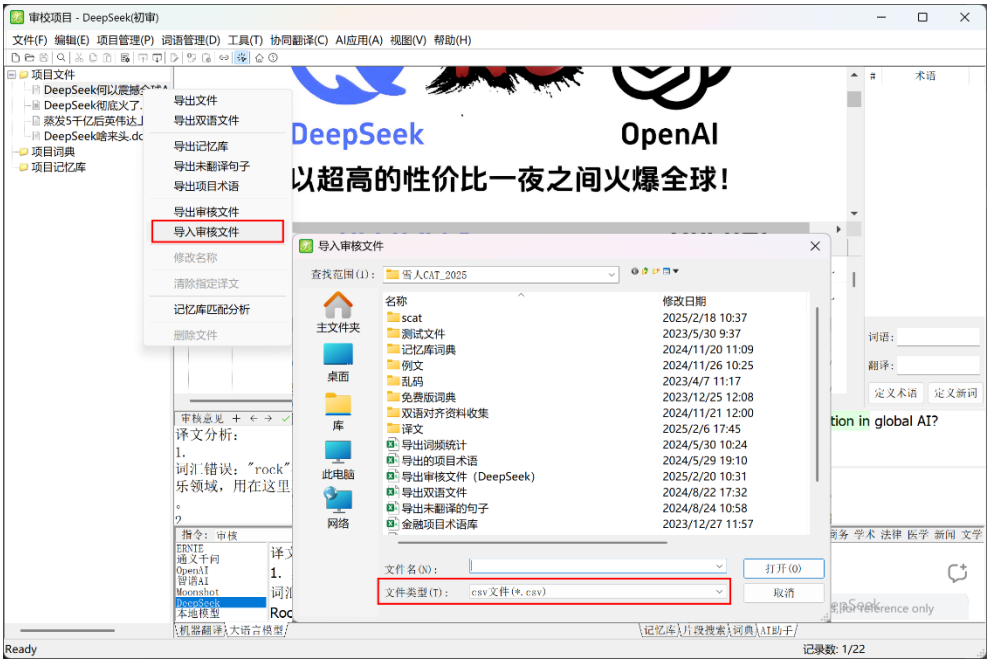
也可以在雪人 CAT 中将一个项目翻译好后，将它导出“.csv”格式的外部审核文件，交给他人在 EXCEL 中进行审核。如下图：在雪人 CAT 中，利用 AI 大预言模型，将“市值蒸发超 5 千亿美元后,英伟达宣布上线 DeepSeek”这篇文档翻译完成后，点击**【右键菜单】➔【导出审核文件】**菜单，将它导出“.csv”格式的外部审核文件。



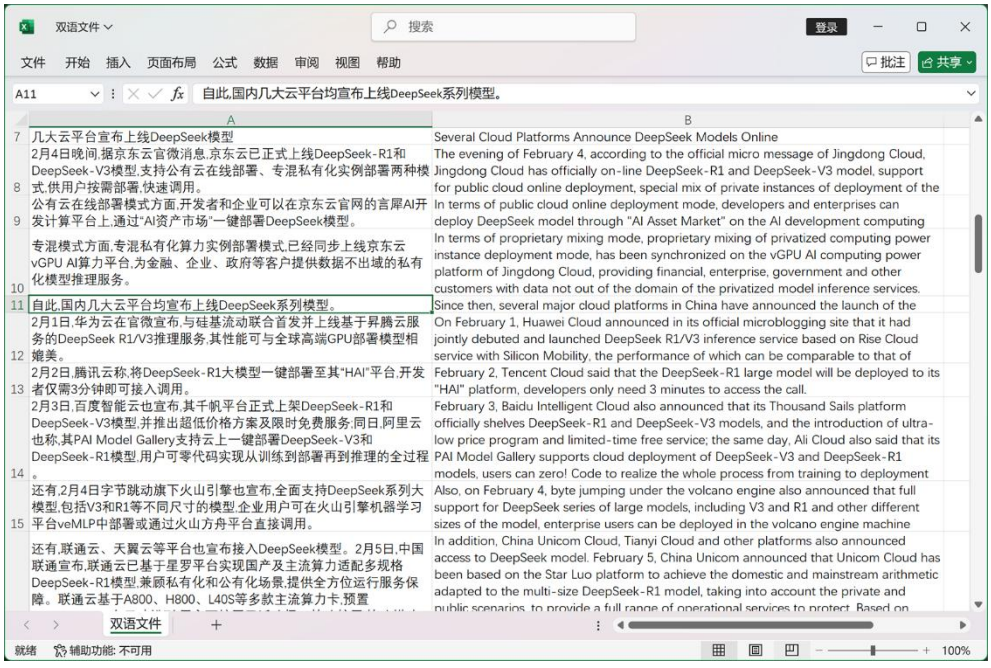
将这个“.csv”格式的文件发给他人审校，审校人员收到这个文件后，利用 EXCEL 打开它，格式如下图：



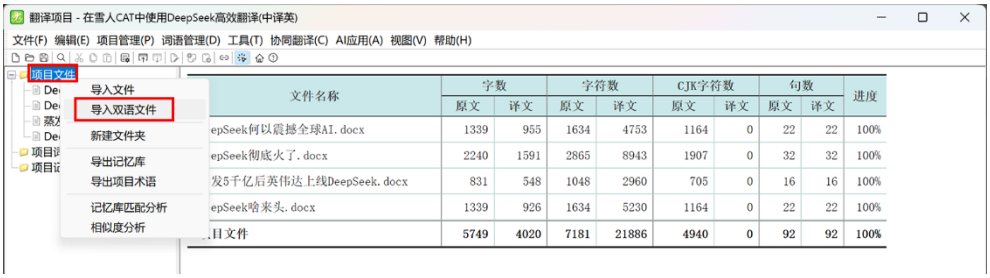
A 列为序号，B 列为原文，C 列为译文。A、B、C 三列请不要做任何修改，如果修改了，会造成导入时被忽略该句的内容。审核人员请在 D、E、F 三列中输入相应的内容即可。D 列为审校意见，E 列审校译文，F 列为核准与否的状态（0 为未核准，1 为已核准，已核准的句子在雪人 CAT 中的序号前面会打上绿色的√）。这样审核完成后，保存为“.csv”格式的文件，就可以将这些信息导入回雪人 CAT 软件中。点击**【右键菜单】➔【导入审核文件】**菜单，即可将其导入。如下图:



如果是用其他 CAT 软件做的翻译文档，希望在雪人中利用 AI 大语言模型进行审校，可以在其他 CAT 软件中将它导出成“.csv”格式的双语文件，如下图:



然后在雪人 CAT 中，点击屏幕左侧的“项目文件”后，点击**【右键菜单】➔【导入双语文件】**菜单，将它导入雪人中，如下图:



导入后，就可以按上述讲解的方法转换成“审核项目”，利用 AI 进行审校，审校后转换成“修订项目”，决定是否采取 AI 的审校译文等操作。

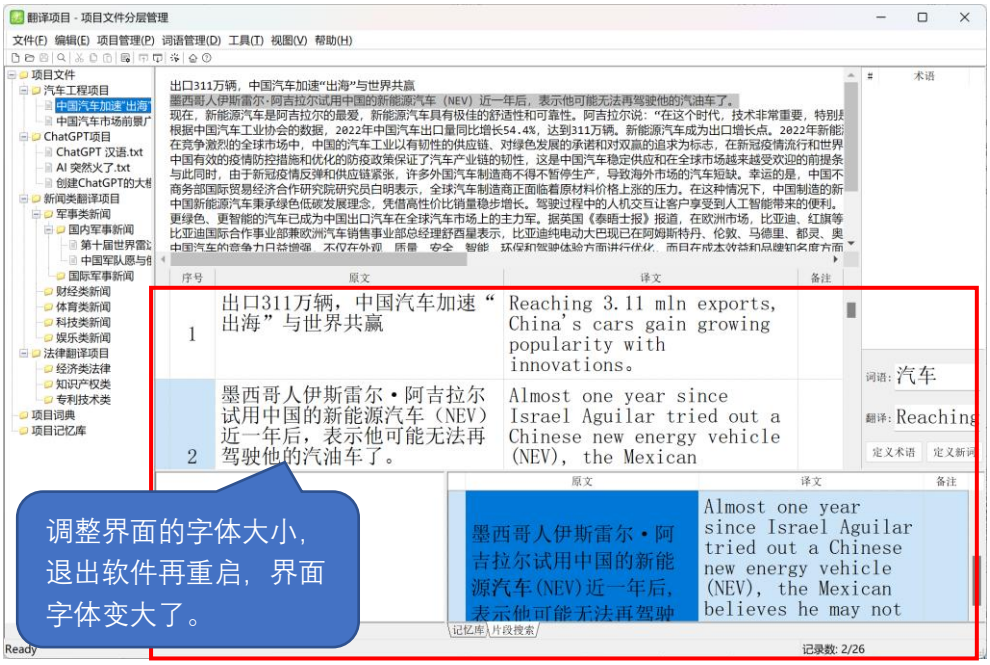
第十六章 软件设置与界面调整

第一节 调整界面字体和大小

一、调整字体的大小

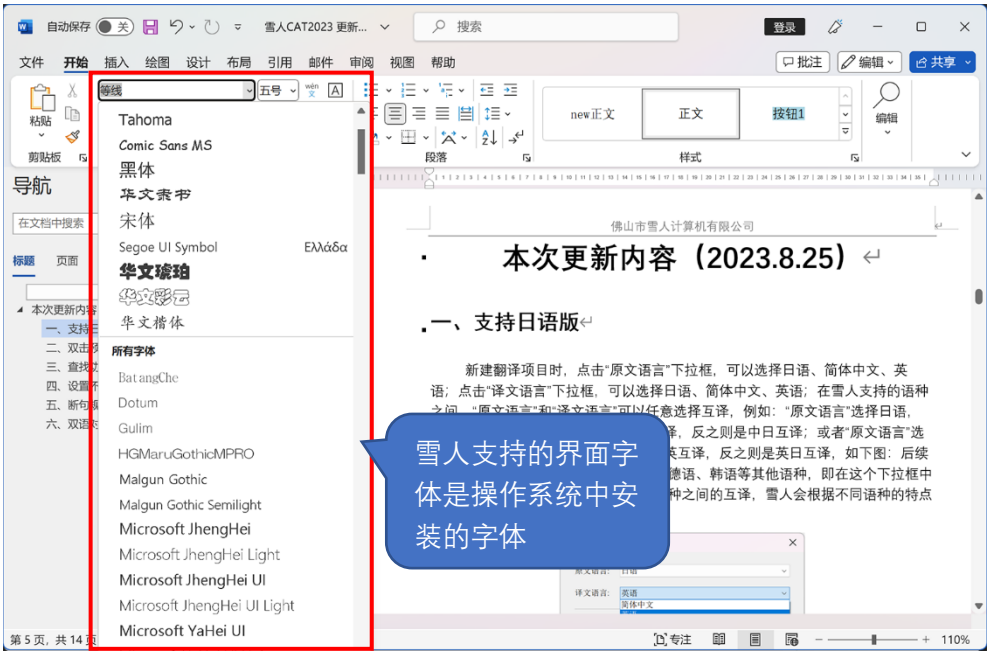
点击【工具菜单】➔【软件设置】➔【界面设置】可以调整软件界面的字体大小，如下图：进行字体调整后，需要退出软件再重新启动软件才生效。



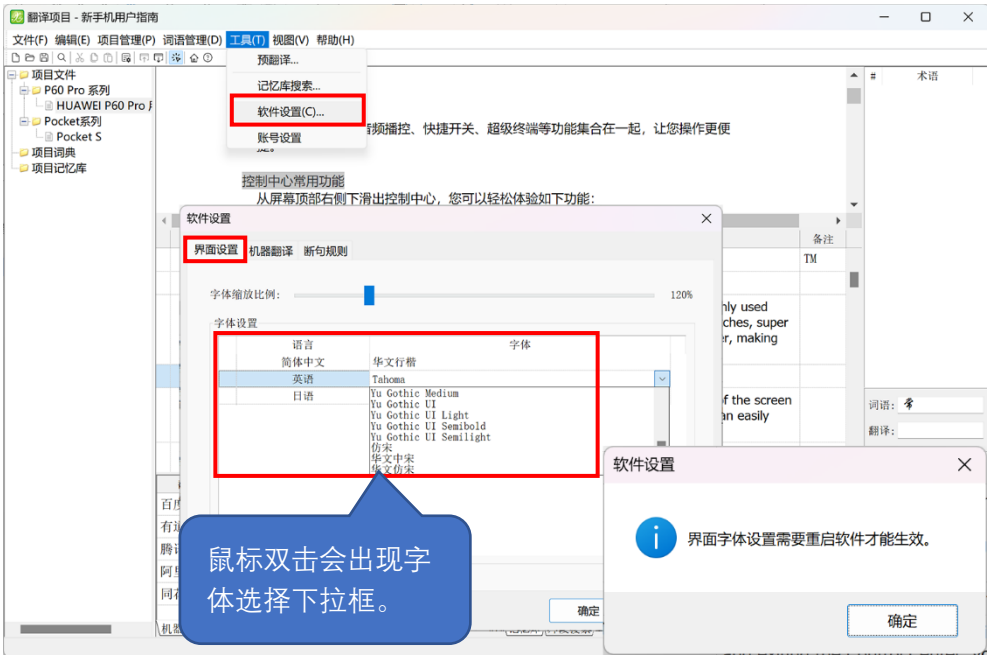


二、设置不同语种的字体

新版软件支持设置界面显示的字体，支持的字体种类和字体名称与 WORD 中支持的一样，用户可以先在 WORD 中预览测试好希望调整的字体样式，并记住该字体的名称，然后在雪人软件中根据字体名称选用这个字体。



具体设置方法：点击【工具】➔【软件设置】菜单，在“界面设置”窗口中，选择某个语种后双击鼠标，此时会出现字体下拉选择框，选择需要设置的字体，如下图：雪人此处显示的字体是用户操作系统所安装的字体，可以通过增加操作系统安装的字体来增加新的字体。



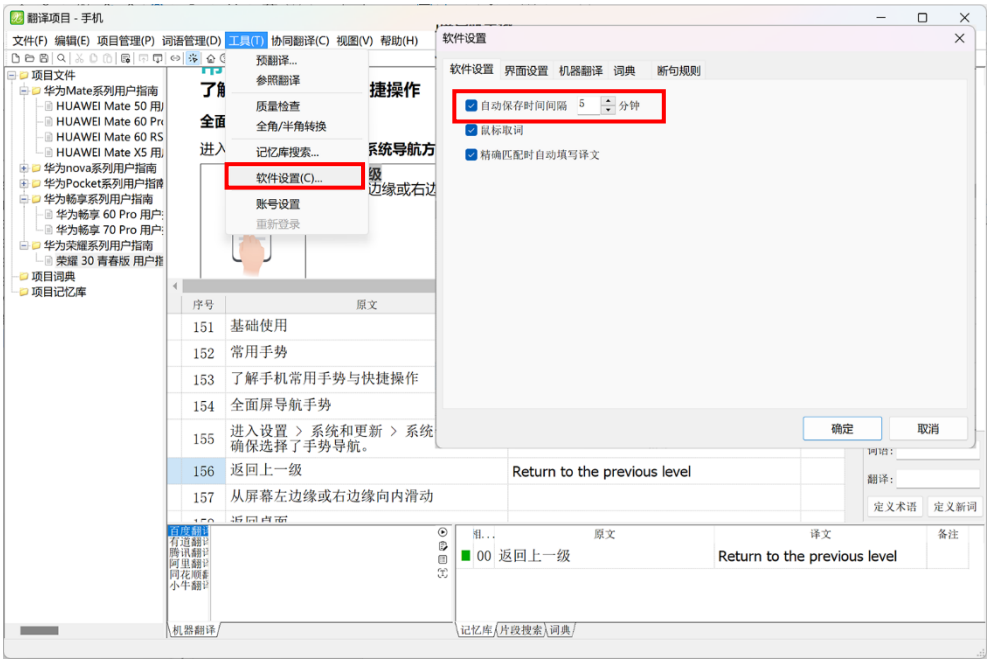
设置字体后，不是立即生效，需要退出雪人软件再重新启动软件才生效，修改字体后，软件界面字体显示如下图：翻译编辑区、机器翻译区、记忆库显示区、片段搜索区、项目记忆库、项目词典等的字体都会发生变化。



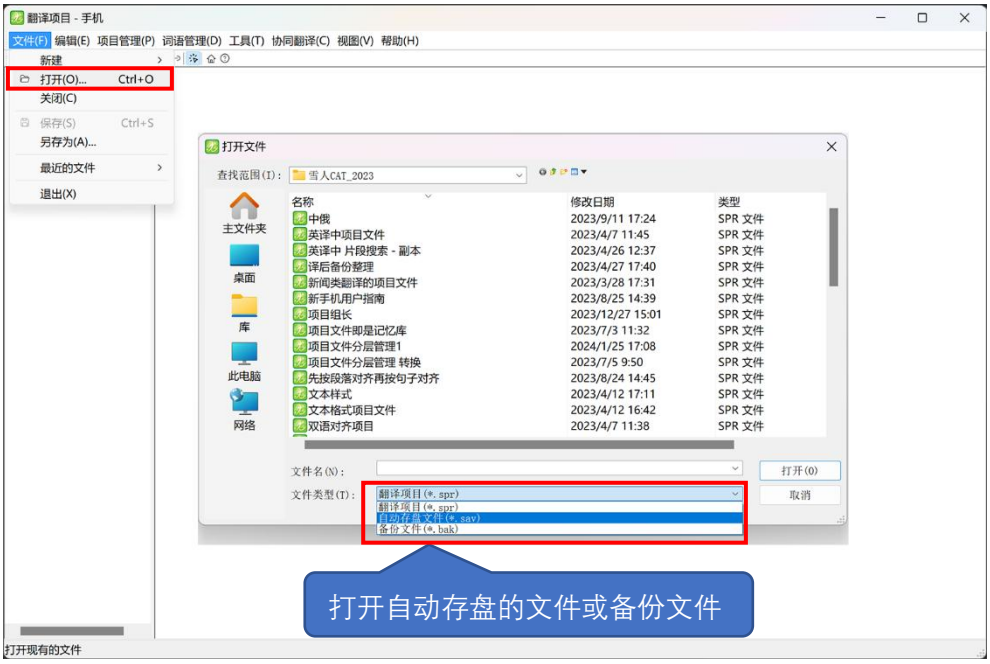
第二节 自动存盘

为了保障用户数据的安全，在保存翻译的项目文件“.spr”同时，还自动生成了“.spr.bak”和“.spr.sav”备份文件，内容与“.spr”的一样，“.bak”是上一次手动存盘的版本，“.sav”是软件自动存盘的版本，一般不用理会它们。

自动存盘的时间默认为 5 分钟，用户可以自己调整时间。点击**【工具】➔【软件设置】**菜单，在“软件设置”窗口中可以设置自动存盘的时间。如下图：



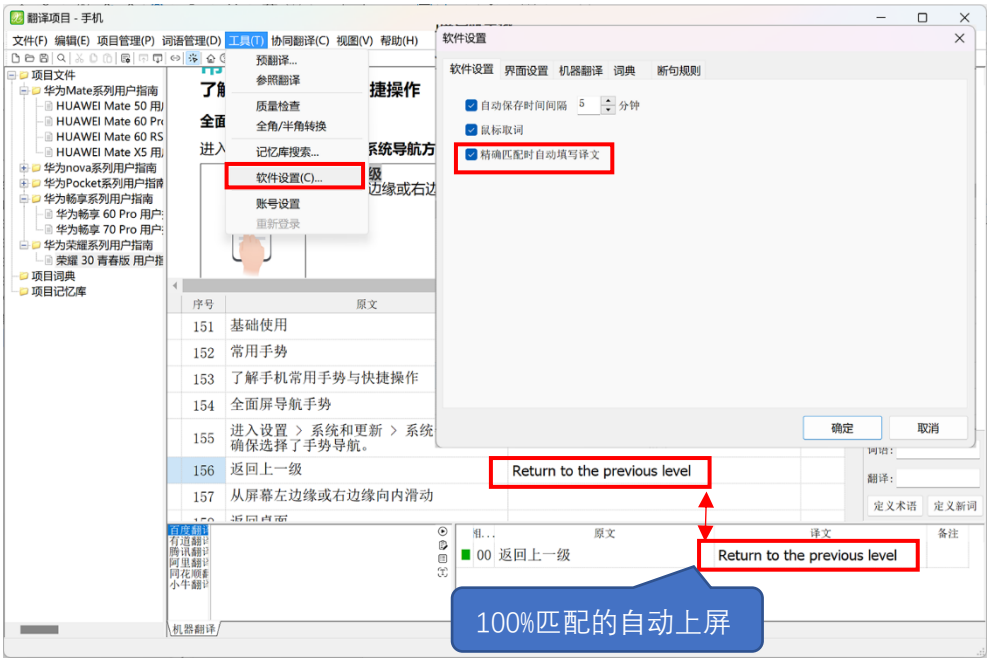
若“.spr”意外损坏，可以启动备份文件或自动存盘文件。点击**【文件】➔【打开】**菜单或工具栏的“打开”按钮，在弹出窗口的“文件类型”中点击下拉按钮，打开备份文件或自动存盘的文件。如下图：



打开后建议另存为新文件名后再继续工作，避免文件名产生混乱。

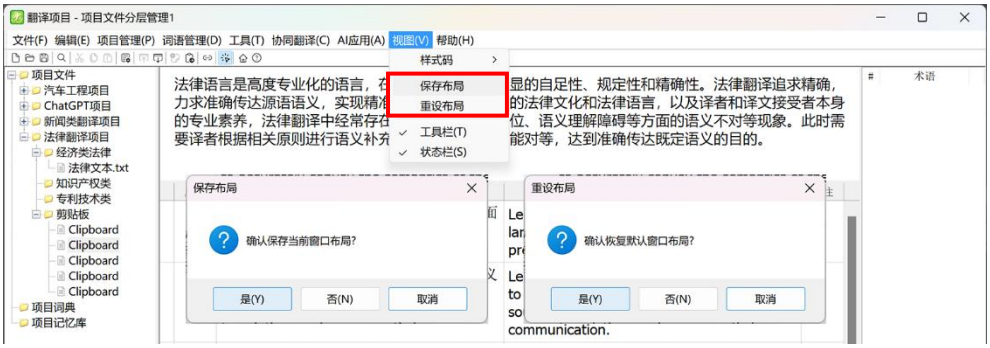
第三节 精确匹配时自动填写译文

点击【工具】➔【软件设置】菜单，在“软件设置”窗口中，勾选“精确匹配时自动填写译文”选项后，当翻译时遇到记忆库中有100%匹配的句子，光标去到那句时会自动上屏。如下图：



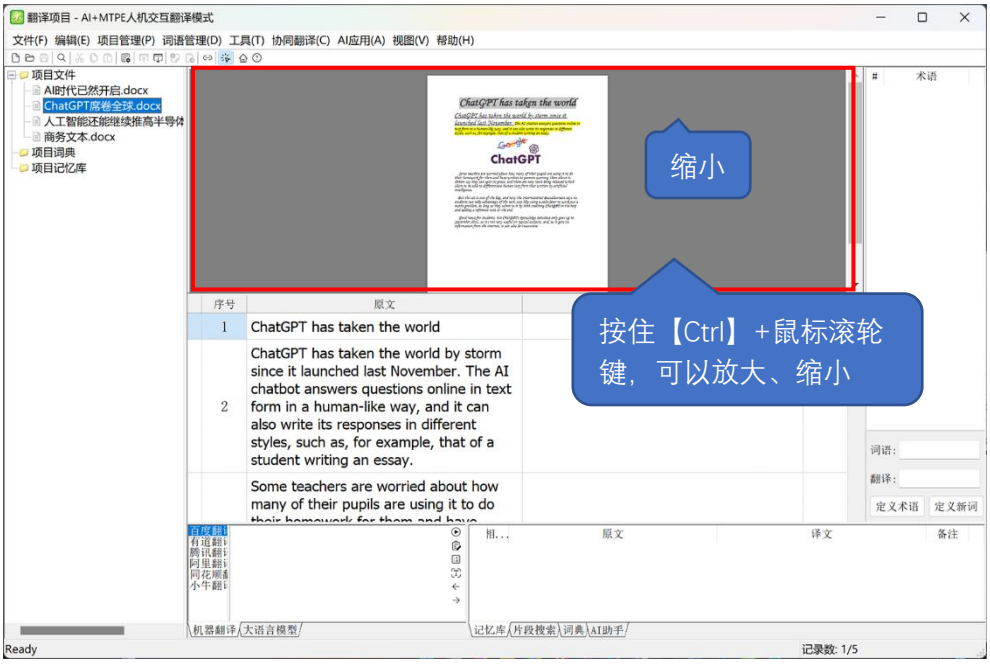
第四节 保存界面设置

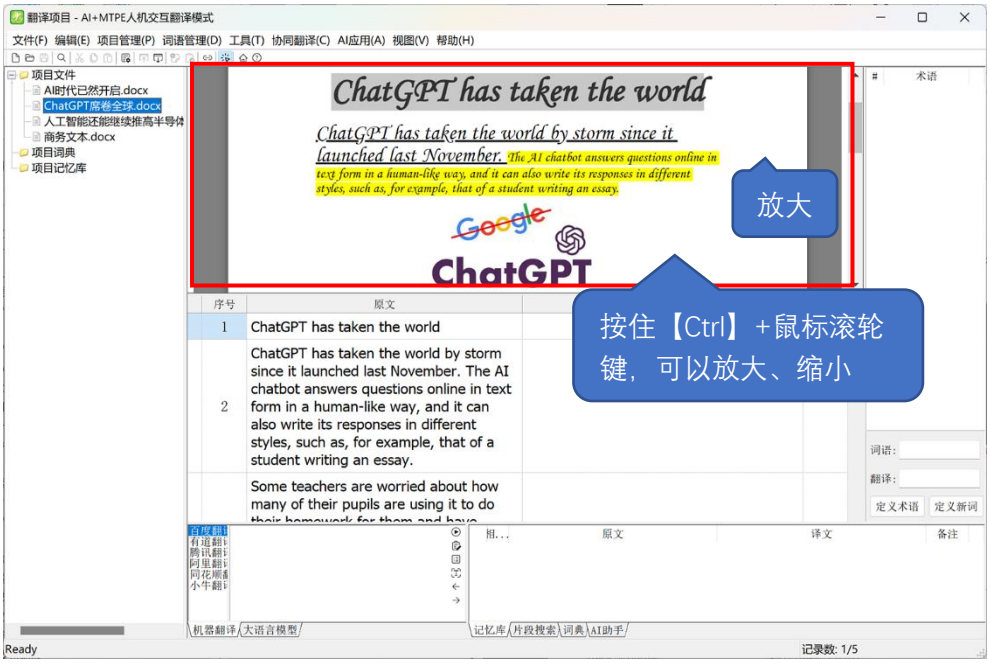
可以根据需要，灵活调整软件界面窗口的大小，进行调整后如果希望下次启动软件时，能保留自己所做的调整，可以点击主菜单**【视图】➔【保存布局】**功能将它保存起来。如果希望回复到软件的默认界面设置，可以点击主菜单**【视图】➔【重设布局】**功能来恢复。如下图:



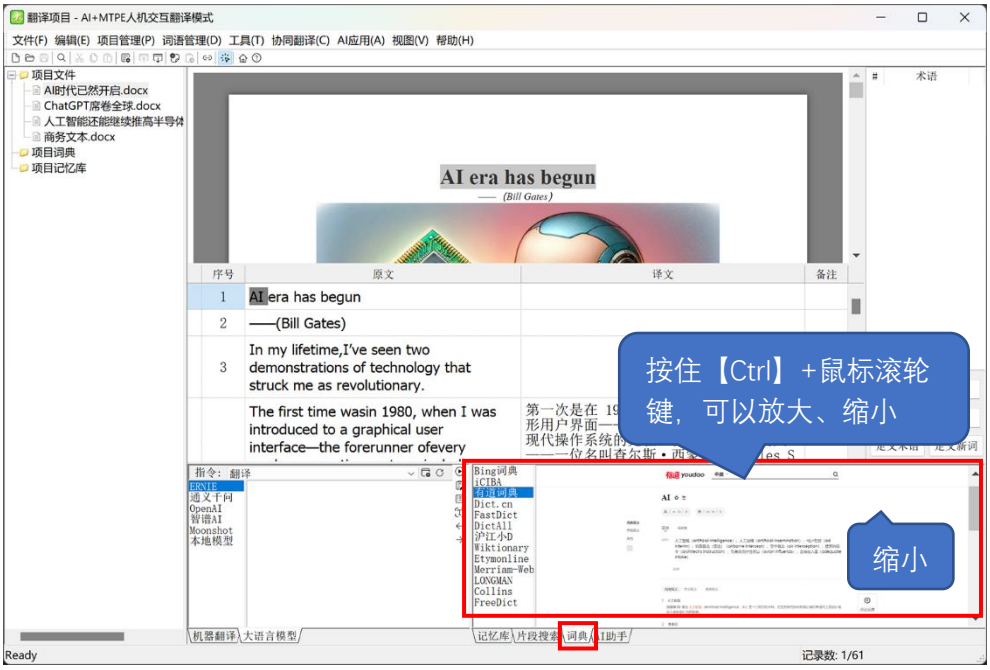
第五节 窗口的放大、缩小

在原文预览窗口，按住【Ctrl】键同时滚动鼠标滚轮键，可以实现预览内容的放大、缩小。如下图:

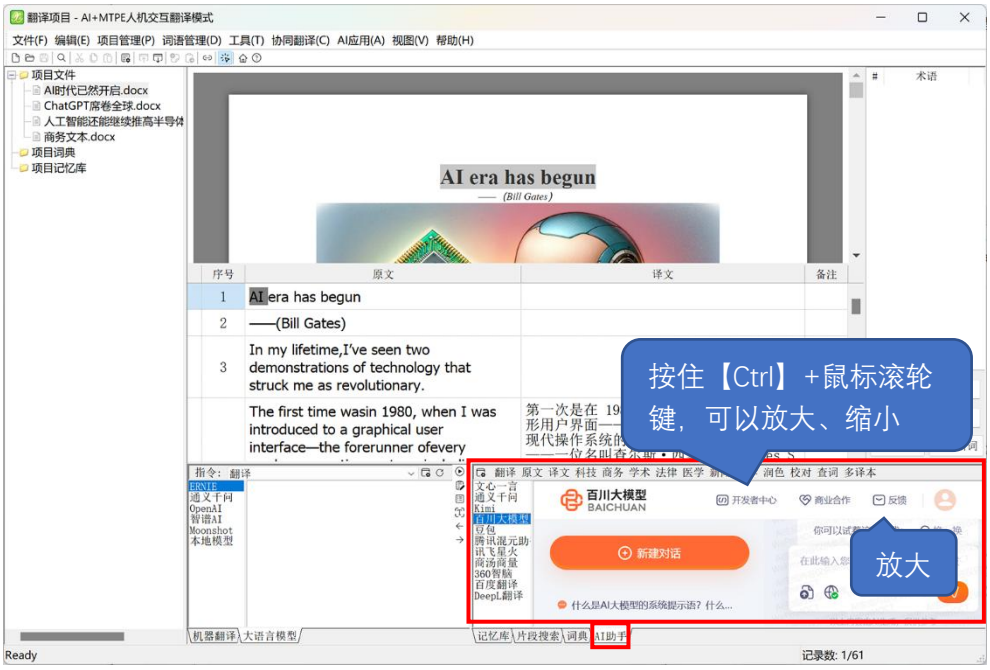


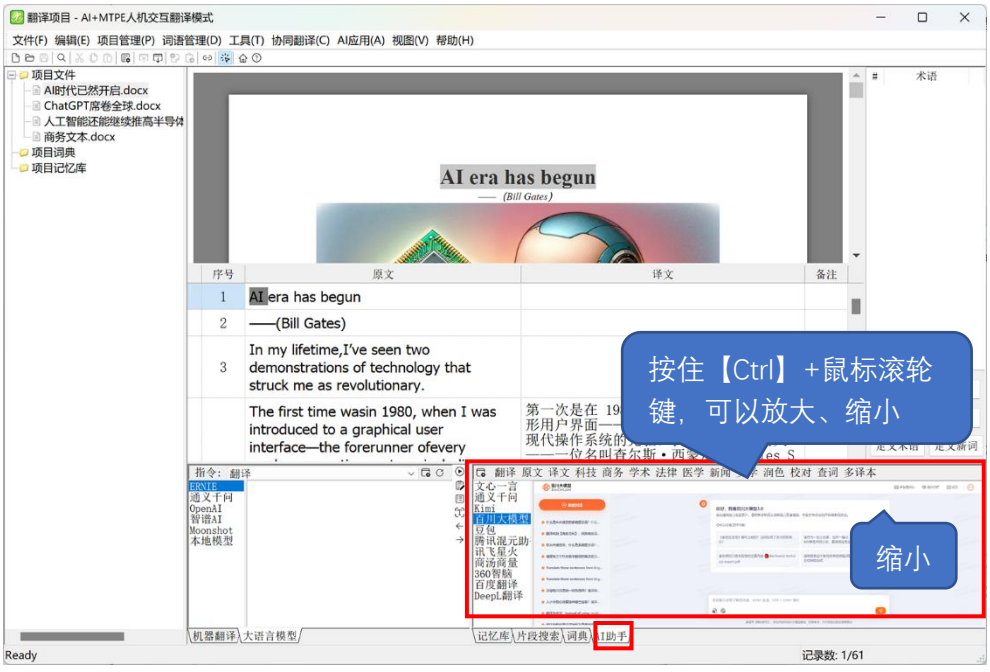


点击屏幕下方的“词典”页签，在“词典”窗口，按住【Ctrl】键同时滚动鼠标滚轮键，可以实现“在线词典”窗口内容的放大、缩小。如下图:



点击屏幕下方的“AI 助手”页签，在“AI 助手”窗口，按住【Ctrl】键同时滚动鼠标滚轮键，可以实现“AI 助手”窗口内容的放大、缩小。如下图：



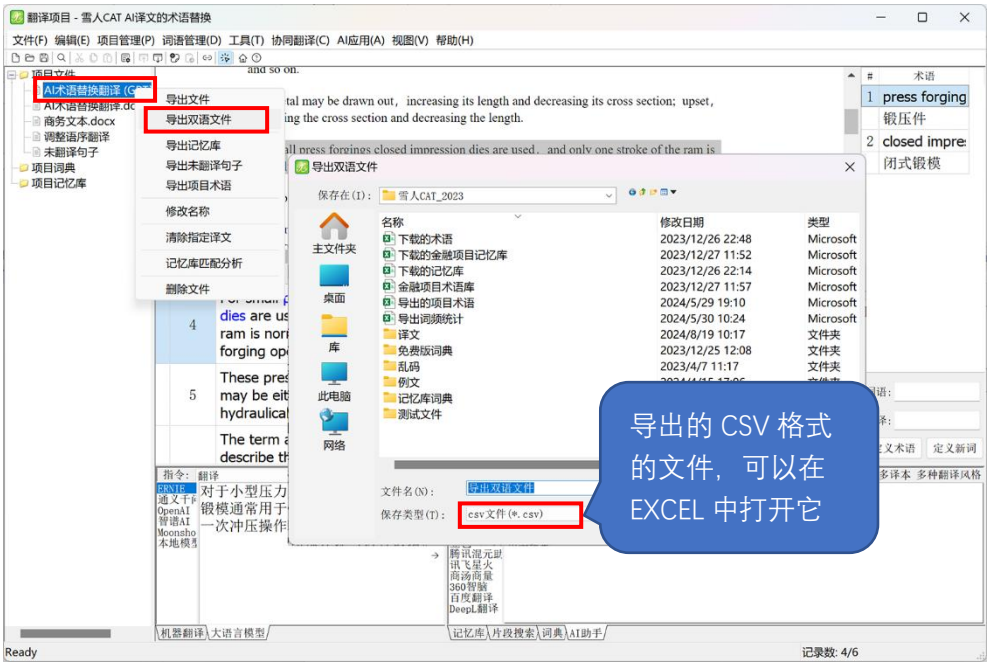


注：从此版本开始，不支持 WINDOWS7 操作系统了。

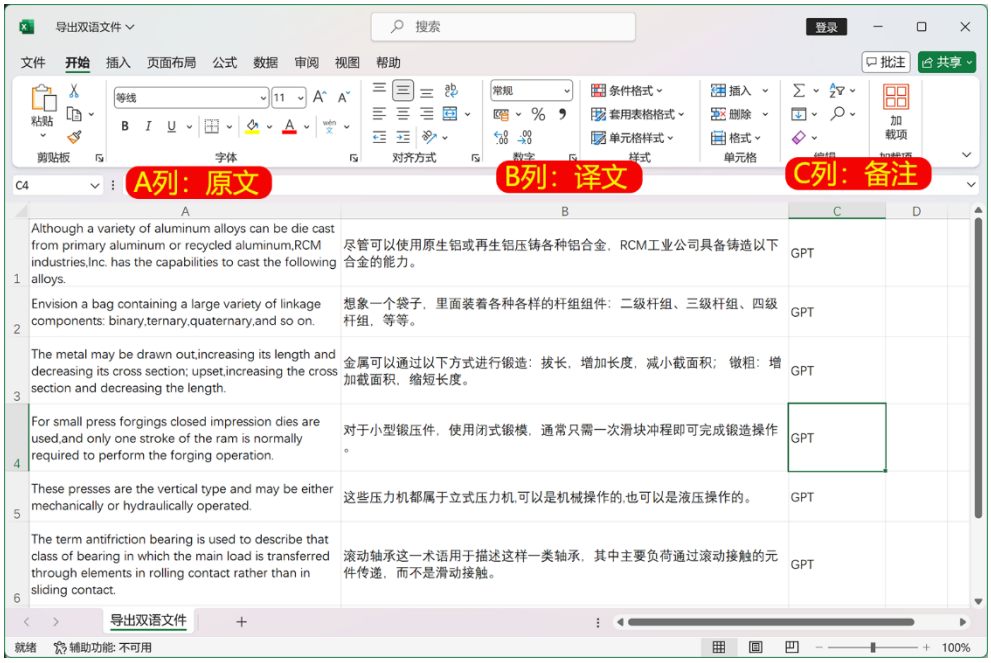
第十七章 其他

第一节 导入、导出双语文件

点击屏幕左侧的某个具体翻译文件，然后点击【**右键菜单**】➔【**导出双语文件**】功能，可以将当前文件导出“.csv”格式的表格文件，如下图：

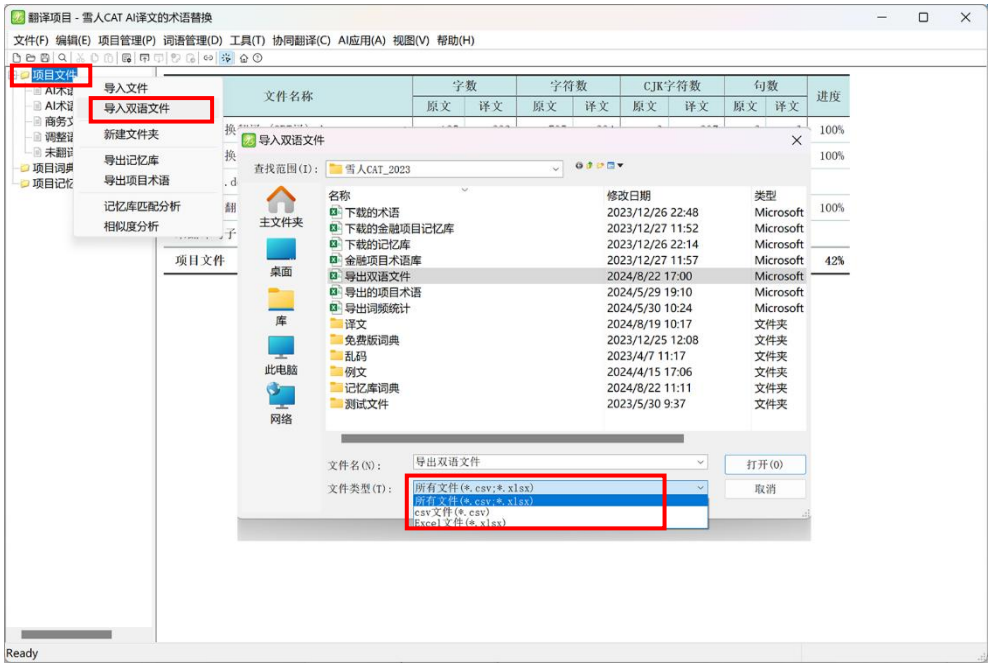


这个“.csv”文件可以在 EXCEL 打开它，A 列为原文，B 列为译文，C 列为备注。



	A	B	C	D
1	Although a variety of aluminum alloys can be die cast from primary aluminum or recycled aluminum,RCM industries,Inc. has the capabilities to cast the following alloys.	尽管可以使用原生铝或再生铝压铸各种铝合金，RCM工业公司具备铸造以下合金的能力。	GPT	
2	Envision a bag containing a large variety of linkage components: binary,ternary,quaternary,and so on.	想象一个袋子，里面装着各种各样的杆组组件：二级杆组、三级杆组、四级杆组，等等。	GPT	
3	The metal may be drawn out,increasing its length and decreasing its cross section; upset,increasing the cross section and decreasing the length.	金属可以通过以下方式进行锻造：拔长，增加长度，减小截面积；镦粗：增加截面积，缩短长度。	GPT	
4	For small press forgings closed impression dies are used,and only one stroke of the ram is normally required to perform the forging operation.	对于小型锻压件，使用闭式锻模，通常只需一次滑块冲程即可完成锻造操作。	GPT	
5	These presses are the vertical type and may be either mechanically or hydraulically operated.	这些压力机都属于立式压力机,可以是机械操作的,也可以是液压操作的。	GPT	
6	The term antifriction bearing is used to describe that class of bearing in which the main load is transferred through elements in rolling contact rather than in sliding contact.	滚动轴承这一术语用于描述这样一类轴承，其中主要负荷通过滚动接触的元件传递，而不是滑动接触。	GPT	

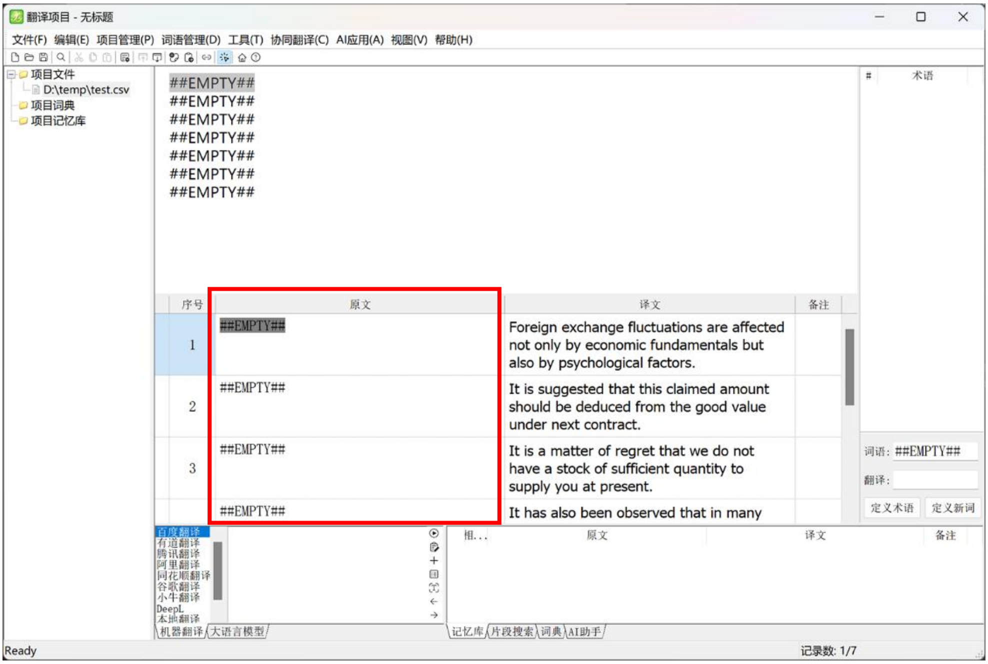
可以将这个文件发给他人校稿，校对完成后，再导入雪人中。点击屏幕左侧的“项目文件”，再点击**【右键菜单】➔【导出双语文件】**功能将它导入。如下图：



导入的双语文件，可以是 EXCEL 格式、也可以是.csv 格式。导入后如下图：

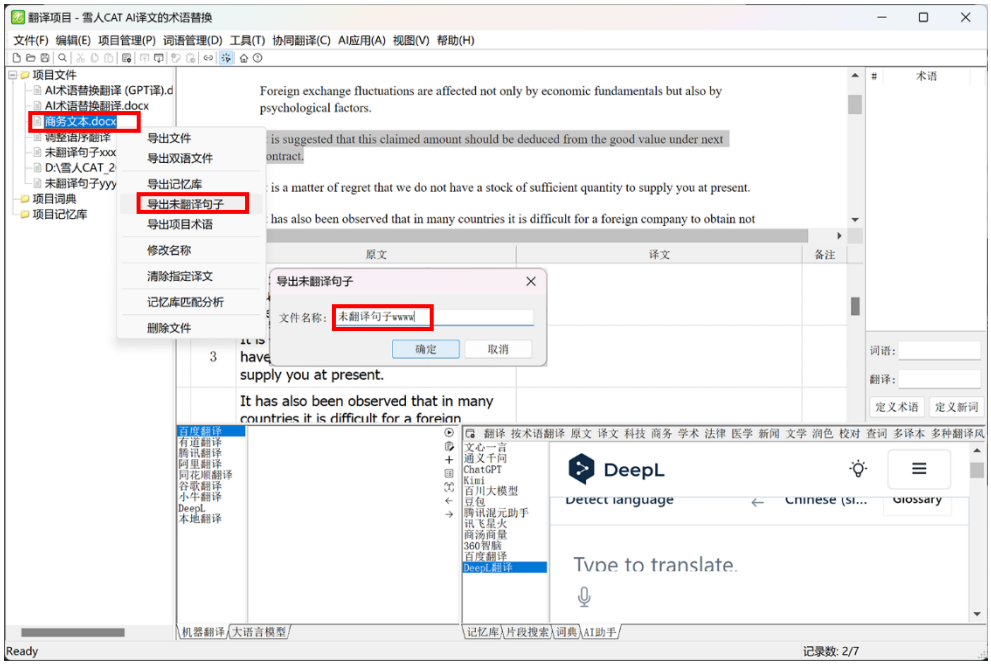


如果在原始导入的文件中 A 列原文为空白，导入雪人后，在原文列中会出现:###EMPTY### ，如下图：

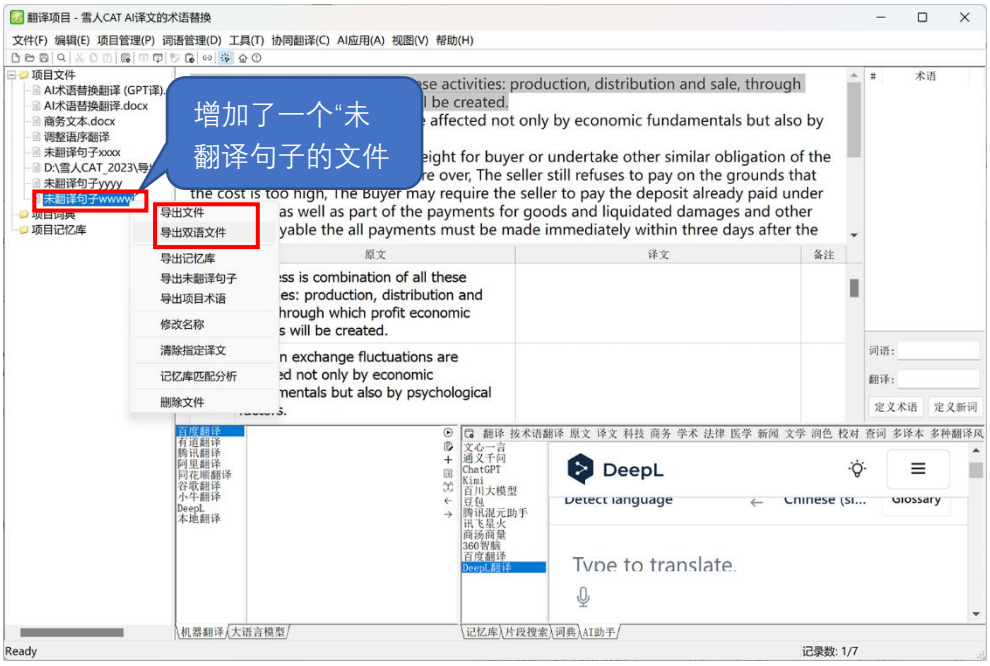


第二节 导出未翻译的句子

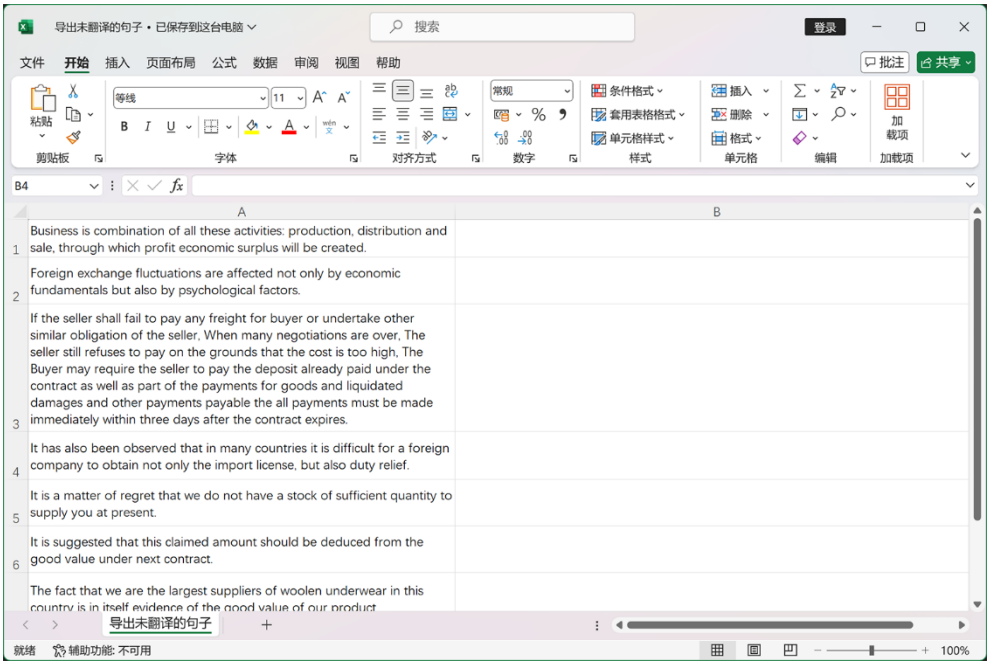
利用记忆库进行快速预翻译后，记忆库有的句子会被自动填上，未翻译的句子可以导出发给其他人翻译。点击屏幕左侧的某个具体翻译文档后，点击**【右键菜单】➔【导出未翻译的句子】**，如下图：



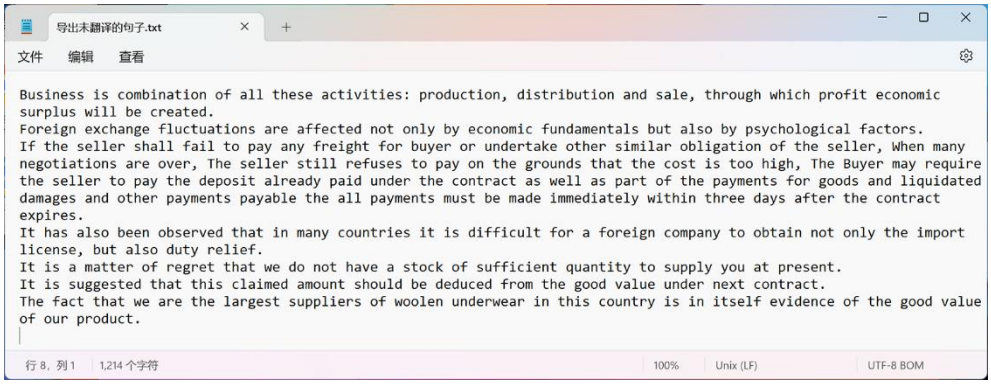
输入文件名后，在屏幕左侧的项目文件最后面会增加一个文件（例如：未翻译句子 wwww），为翻译的句子发给他人翻译时，如果对方有雪人 CAT 软件的，发送雪人的“.spr”的翻译项目文件最省事，如果对方没有雪人 CAT 软件，则可以**【右键菜单】➔【导出双语文件】**，将未翻译的句子导出成“.csv”文件。如下图：



用 EXCEL 打开这个“.csv”文件，另存为 EXCEL 格式，如下图：发送这个 EXCEL 给其他人翻译，翻译好后再利用**键菜【右单】➔【导入双语文件】**将它导入。



或者利用**【右键菜单】➔【导出文件】**，将未翻译的句子导出成“.txt”的文本文件发送给其他人翻译。如下图：

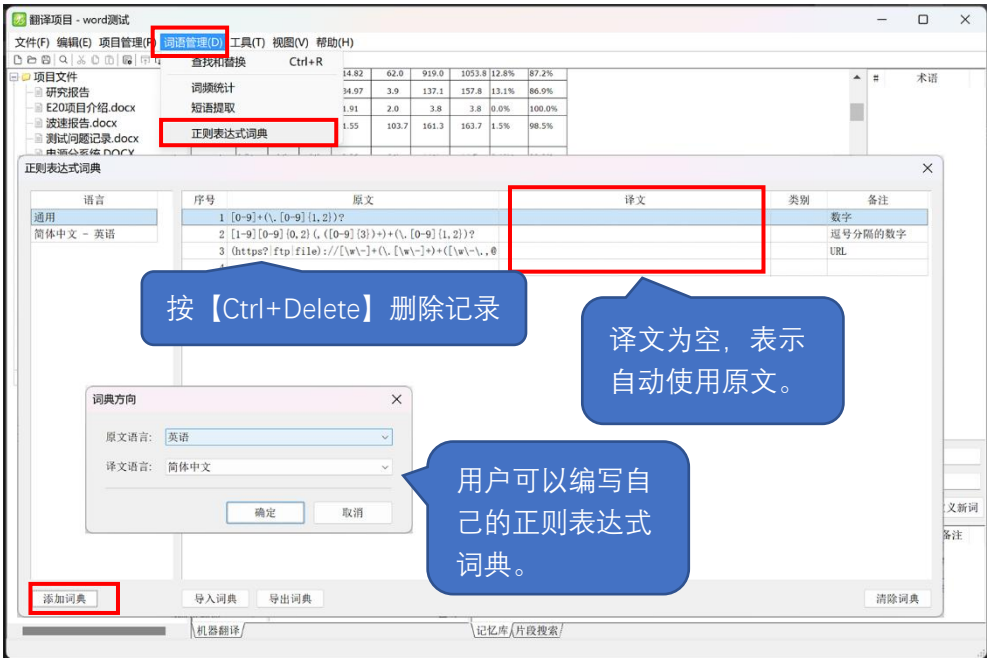


利用“导出未翻译的句子”这个功能，也可以在翻译完成后，点击它检查有没有句子被漏译？

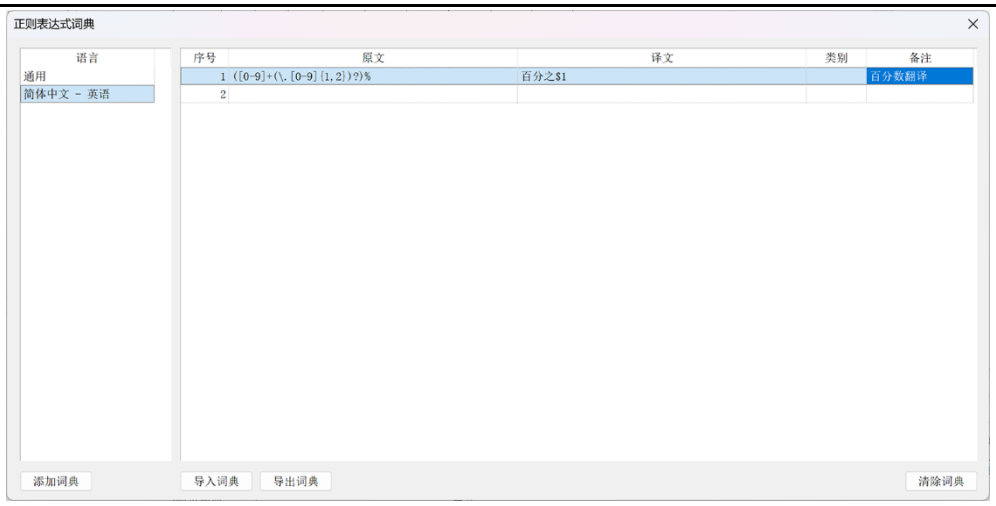
第三节 正则表达式词典

利用正则表达式词典，可以灵活处理非译成份、特殊翻译、语法规则翻译等等，正则表达式的使用有标准的语法，关于如何编写正则表达式？请用户自行查阅相关的正则表达式的书籍。熟悉正则表达式使用的用户会感受到这项功能带来帮助。

点击【**词语管理**】➔【**正则表达式词典**】菜单，会弹出“正则表达式词典”窗口，如下图：软件默认编写了几条针对数字、URL 等翻译处理的通用规则，对于数字的翻译就可以通过正则表达式词典来处理，不用将数字也发送给机器翻译引擎计算流量，节省费用。译文为空表示自动使用原文。



用户可以点击屏幕下方的“添加词典”按钮，编写自己的正则表达式词典，按【Ctrl+Delete】删除记录，译文为空时表示自动使用原文。点击下方的“导入词典”按钮或“导出词典”按钮，可以将正则表达式词典进行导入导出。例如下图中，编写了一个语言方向为“简体中文——英文”的正则表达式词典，将 xx%翻译成：百分之 xx，如：65%翻译成：百分之 65。

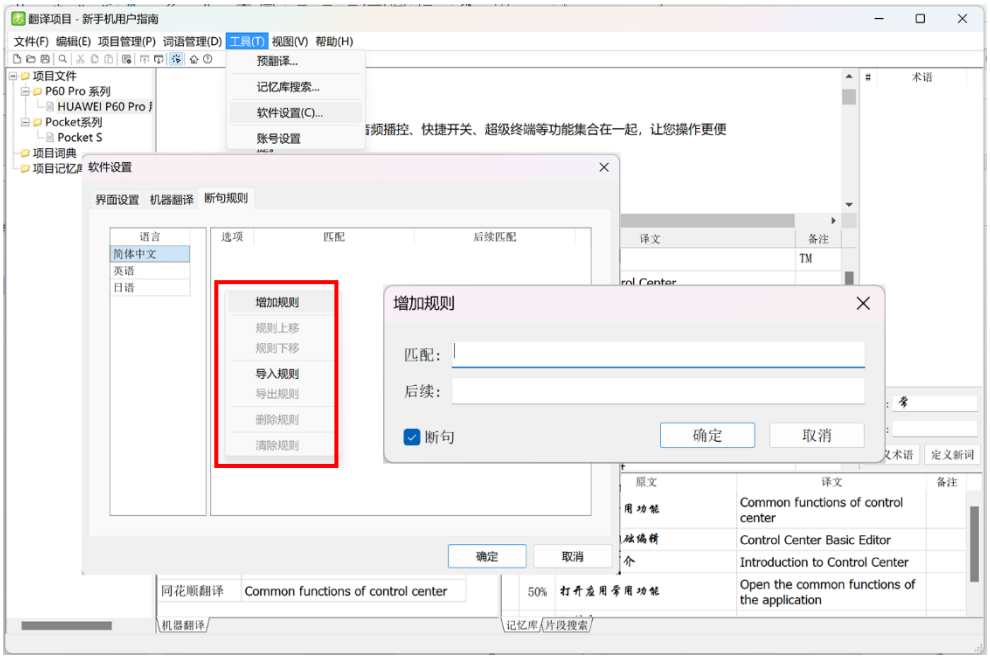


对于非译成份，会使用原文自动填充，在最左边会出现 标注，表示原文和译文完全一致，包括样式码也一致，一旦作了修改，这个图标 会消失。



第四节 自定义断句规则

雪人软件有内置的系统断句规则，但同时支持用户根据实际需要补充自己的断句规则。点击**【工具】➔【软件设置】**菜单，在“断句规则”窗口中，点击窗口左侧的语种，然后在窗口右侧点击右键，会弹出右键菜单，点击**【右键菜单】➔【增加规则】**功能，可以自己编写补充的断句规则，断句规则支持用“正则表达式”语法编写。

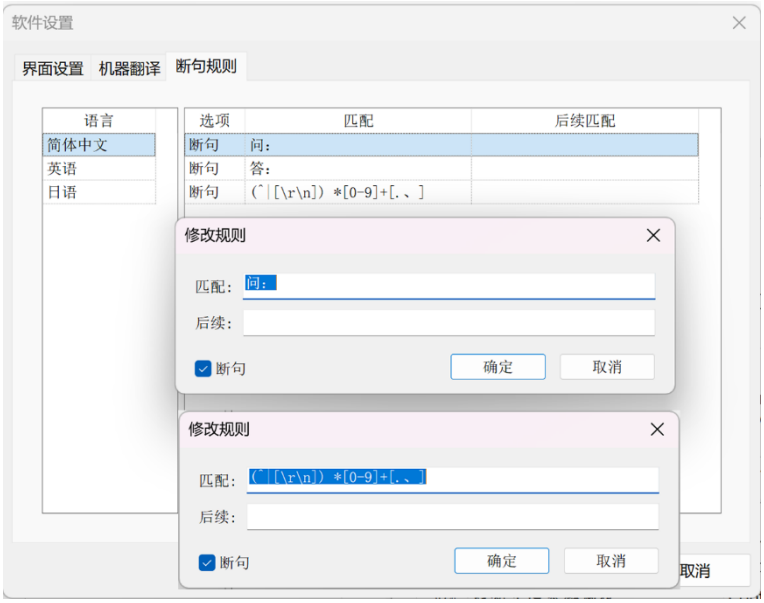


例如：有些用户为了提高记忆库的匹配率，希望按“逗号”来断句；有些特殊的格式，希望将它们单独断句等等，这些都可以通过自定义断句规则来解决。

例 1：下图中：待译文章是问答方式出现的：“序号 1、2、3…… 问：…… 答：……”的格式，若按默认的以句号断句方式，则同一行的译文也需要出现：“1、Q…”，其实我们并不希望形成的记忆库例句中含有“1、问： 答：”等等无关的词语，这些会影响该句日后的匹配率。

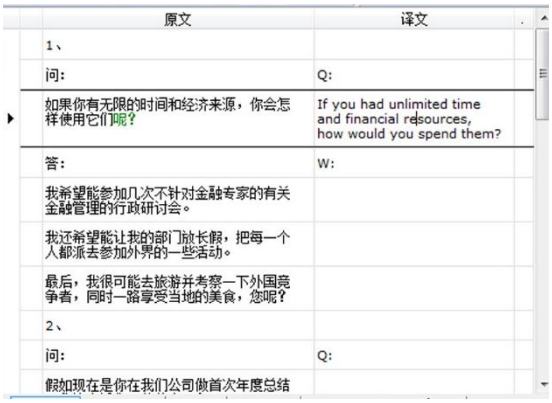


解决办法：在“简体中文”这个语种中，编写 3 个断句规则，第一个规则“问：”后面需要断句；第二个规则“答：”后面断句，第三个规则序号后面断句，具体如下图：



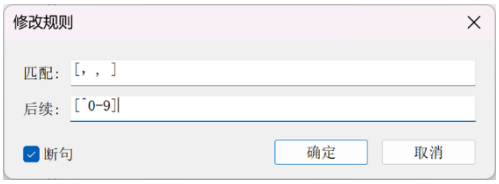
(^|[\r\n]) *[0-9]+[.、] 断句规则可以用“正则表达式”来书写，其中：(^|[\r\n]) 表示“段首”，[0-9] 表示匹配 0-9 的任意数字，[.、]表示数字后面是“.或、”，例如：1、 2、 或 1. 2. ，并勾选“断句”选项，这个规则就表示：遇到句首出现的 1、2、3、或 1. 2. 3.情况下，需要断句。

定义好这 3 个断句规则后，需要**重新导入原文，**才能按新的规则进行断句。新规则断句效果如下图：



	原文	译文
	1、	
	问：	Q：
▶	如果你有无限的时间和经济来源，你会怎样使用它们呢？	If you had unlimited time and financial resources, how would you spend them?
	答：	W：
	我希望能参加几次不针对金融专家的有关金融管理的行政研讨会。	
	我还希望能让我的部门放长假，把每一个人都派去参加外界的一些活动。	
	最后，我很可能去旅游并考察一下外国竞争者，同时一路享受当地的美食，您呢？	
	2、	
	问：	Q：
	假如现在是你在我们公司做首次年度总结	

例 2：希望按“逗号”断句，在“匹配”输入框中输入[，，]，在“后续”输入框中输入[^0-9]（表示非 0-9），同时勾选“断句”选项，如下图：这个规则表示：当遇到中文状态下的逗号或英文状态下的逗号，且后续紧接的不是 0-9 的任意数字时就断句。



注意：因为在正则表达式中，“.*()[]”这些字符属于特殊字符，如果我们在断句规则中需要匹配这些符号，一定要在这些符号前加上一个反斜杠“\”，例如“\.”表示匹配符号“.”当出现有多条断句规则时，最下面的规则最优先生效。

自定义断句规则同样可以运用到做记忆库的“双语对齐项目”，不过，如果是双语对齐项目，两个语言方向都要作调整，例如：中文按"，"断句，英文也有按“,”断句。否则对齐得不够好。

技巧：利用自定义断句规则，用户可以自己创建短语级的词典。即可以将一些长句拆成短语片段，导出 EXCEL，然后在项目词典中导入，就可以创建自己的片段短语词典了。

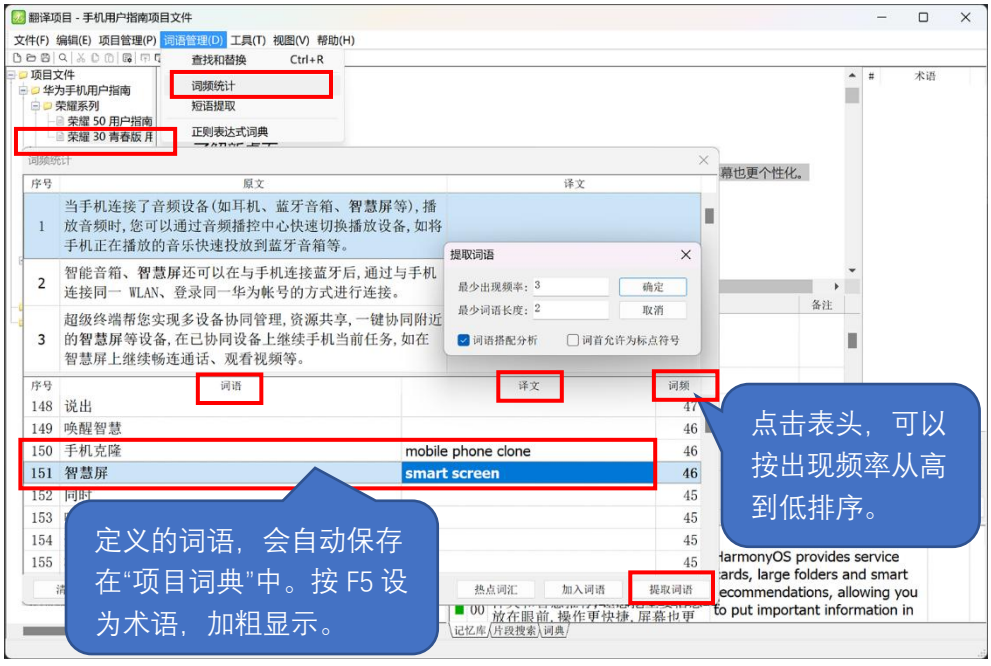
第五节 词频统计

“词频统计”就是统计出文章中的高频词语，在“翻译的项目文件”、“项目记忆库”、“双语对齐”中都支持“词频统计”功能。

（一）、在“翻译的项目文件”中进行词频统计

点击屏幕左侧“项目文件”下面的某个具体翻译文档后，再点击**【词语管理➔词频统计】**菜单，则统计的是这篇文档的词频。如果点击左侧的某个文件夹后，再点击**【词语管理➔词频统计】**菜单，则统计的是该文件夹下面所有文档的词频。如果点击左侧的根目录——“项目文件”，再点击**【词语管理➔词频统计】**菜单，则统计的是这个项目的所有文档的词频。（在双语对齐项目中也类似）

对需要翻译的文档，可以利用“词频统计”功能先统计出文章中的高频词语，在译前对它们进行准确的定义。如下图：点击“词频统计”窗口下方的“词语提取”按钮，可以提取出高频词语。点击“词频统计”窗口中部的表头，可以排序，例如点击“词频”表头，可以按出现的频率从高到低排序。按【F5】键可设为术语并加粗显示，例如下图中，定义了“手机克隆”和“智慧屏”这两个高频词，其中“智慧屏”按了【F5】键，该词加粗显示，表示这个词定义为术语。



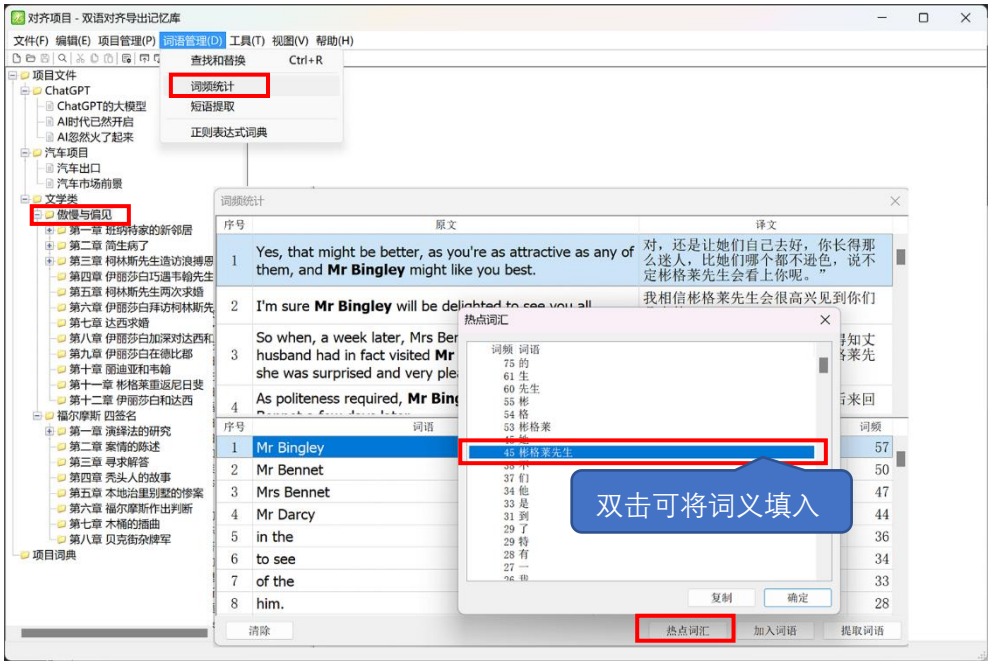
词频统计中定义的词语，会自动保存在项目词典 中。如下图：



对已经翻译好的文档也可以利用“词频统计”来提取一些高频词语加入到 项目词典 中。对记忆库也可以进行词频统计，先将语料导入在“ 项目记忆库 ”中，然后点击【**词语管理➔词频统计**】菜单，统计“ 项目记忆库 ”中的词频，从记忆库中提取词语。

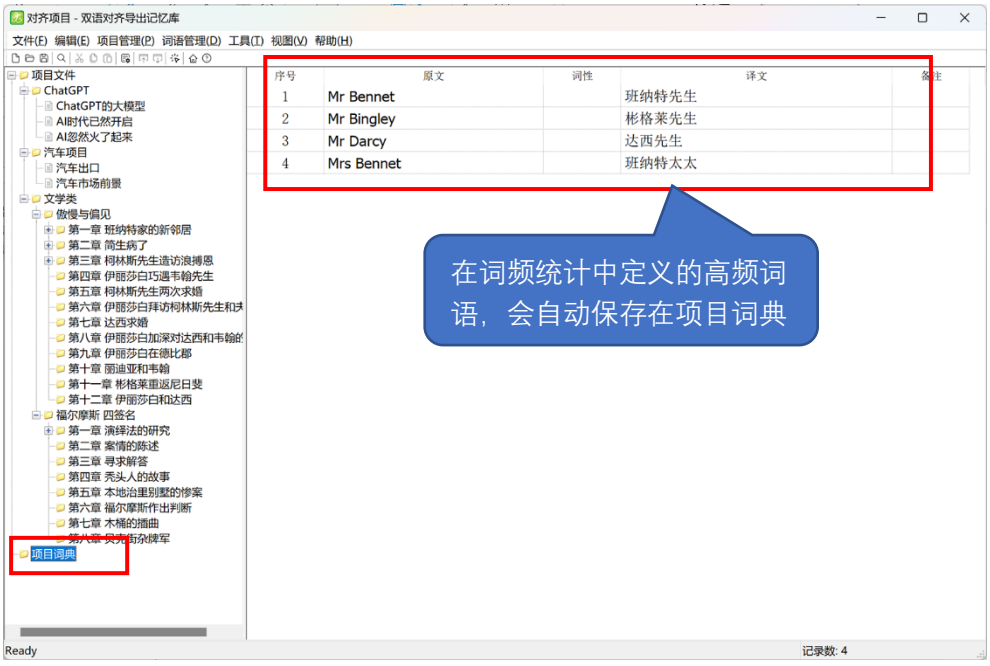
（二）、在“双语对齐”中进行词频统计

在双语对齐时，先自动对齐，然后对这个双语对齐项目进行“词频统计”，统计出高频词语并将它们加入到“ 项目词典 ”中，特别是一些地名、人名、机构名、专业术语等，先将这些词进行定义，例如：下图中，统计出文学作品中的高频出现的人名，点击“词频统计”窗口下方的“热点词汇”按钮，可以从译文中找到合适的翻译，双击即可将词义填入。





在“词频统计”中定义的词语，会自动保存在“项目词典”中，如下图：



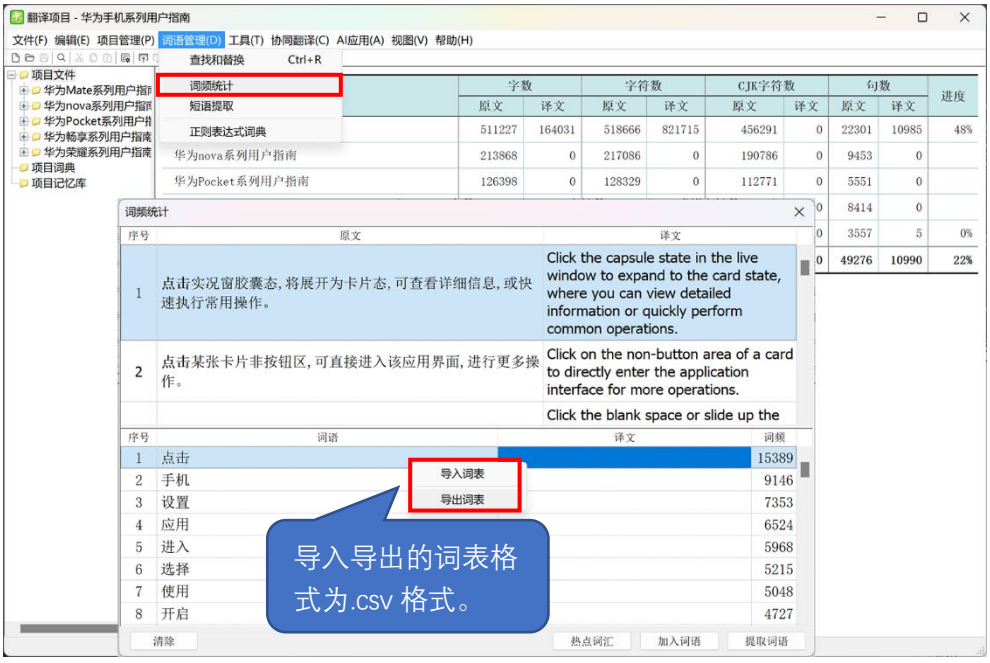
然后再次导入文件进行第二次对齐，对齐时勾选“使用词典辅助对齐”，可以提升自动对齐的准确度。如下图：



导入导出词表

为了方便对词频统计的结果进行进一步的处理，可以将统计结果导出，或导入已有的词表来分析它在当前翻译项目中重复出现情况。

首先利用词频统计功能，统计出项目中出现频率较高的词语，然后点击**【右键菜单】➔【导出词表】**功能将它导出，导出的是一个“.csv”格式的文件，可以使用 EXCEL 打开它进行词语预先定义。定义好后，可以在屏幕左侧的 项目词典 中将它导入来使用。如下图:



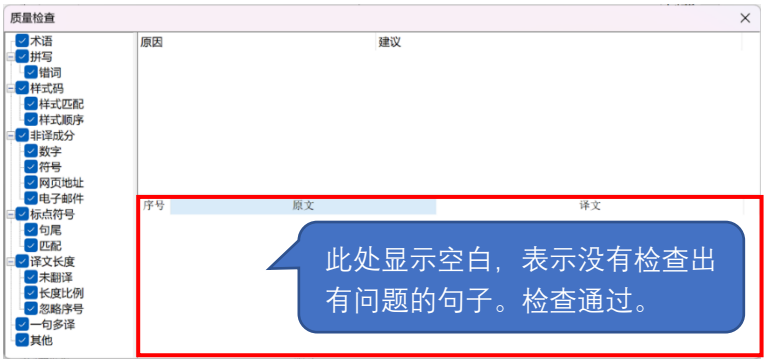
上图中，点击**【右键菜单】➔【导入词表】**功能，可导入“.csv”格式词表，A 列为原文、B 列为译文。导入后，通过每个词显示的词频数值，便可知其重复出现情况。对当前项目没有使用上的词语，其词频数值就显示为 0，点击“词频”列的表头，可按出现的次数进行排序。如果导入的词表比较大，导入会比较慢。

第六节 质量检查

翻译完成后，点击【工具】➔【质量检查】菜单，对译文进行质量检查，软件会检查出译文中可能存在的问题。质量检查的项目有：1. 术语检查；2.拼写检查；3.样式码检查；4. 数字、符号、网页地址、电子邮件等非译成分的检查；5.标点符号的检查；6. 未翻译的句子、句子长度的译文长度比对检查；7. 一句多译的检查等等。如下图：

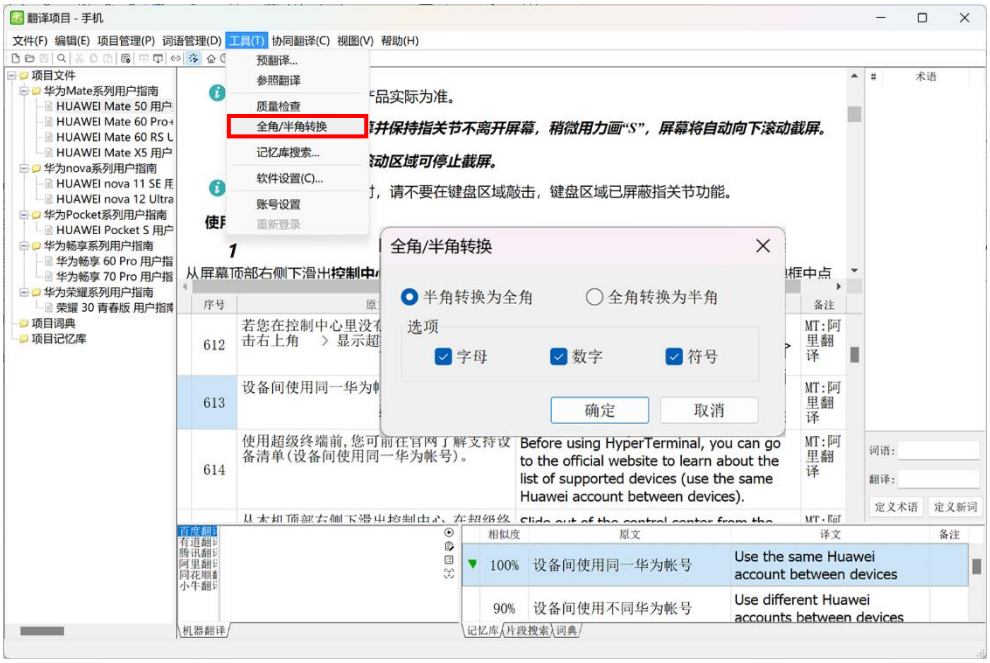


双击“质量检查窗口”中列出的可能有问题的句子，可以跳转到翻译界面的该句进行修改，逐条处理“质量检查窗口”中的句子后，对已经处理的句子，“质量检查窗口”中的记录不会自动刷新，再次点击【工具】➔【质量检查】菜单，看看还有没有检查记录？如果还有检查记录，则继续双击问题句子定位并修改，如果“质量检查窗口”下方显示空白，表示没有检查出有问题的句子，检查完成。如下图：



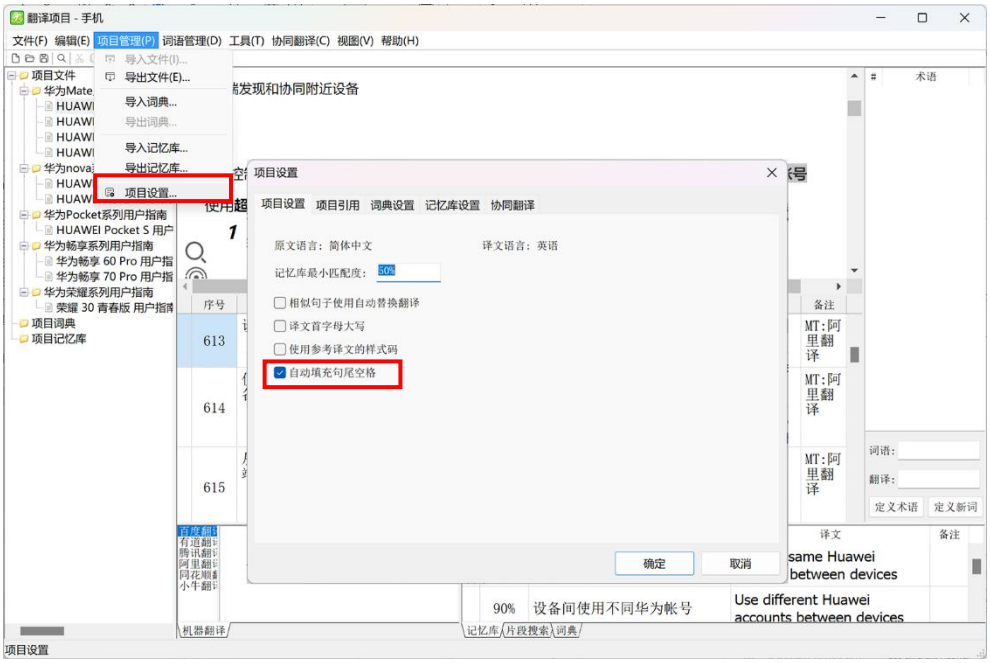
第七节 全角/半角转换

点击【工具】➔【全角/半角转换】菜单，可以对字母、数字、符号等进行全角/半角的转换。如下图：



第八节 自动填充句尾空格

点击【项目管理】➔【项目设置】菜单，在“项目设置”窗口，默认勾选了“自动填充句尾空格”的选项，例如：对于译文为英语的项目，会根据需要在句尾补充必要的空格，解决导出的译文粘连在一起的问题。如下图：



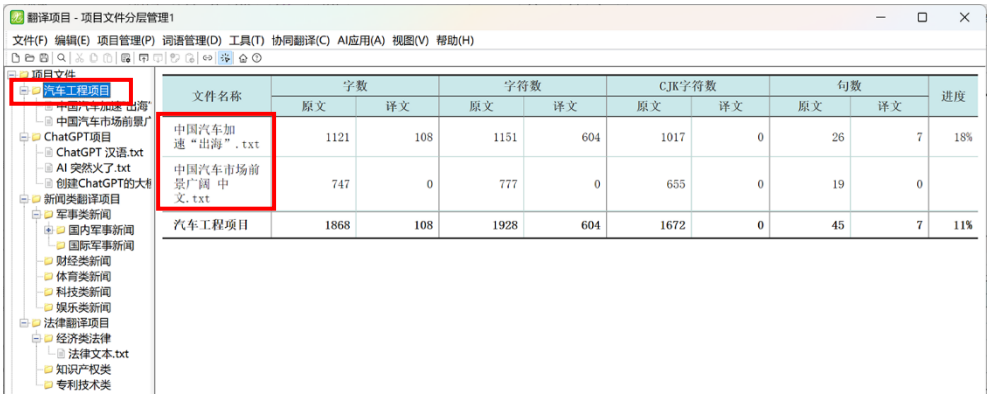
第九节 字数统计及显示翻译进度

字数统计及显示翻译进度是一项很实用的功能，软件将原文和译文的字数、字符数、句子数全部统计出来，并以表格的方式清晰显示。点击左侧最上层的“项目文件”，则以“子文件夹”为单位显示统计的结果。如下图：



文件名称	字数		字符数		CJK字符数		句数		进度
	原文	译文	原文	译文	原文	译文	原文	译文	
汽车工程项目	1868	108	1928	604	1672	0	45	7	11%
ChatGPT项目	4029	0	4331	0	3576	0	100	0	
新闻类翻译项目	2242	310	2302	1700	2034	0	52	16	24%
法律翻译项目	989	252	1575	1556	776	0	24	7	25%
项目文件	**9128**	**670**	**10136**	**3860**	**8058**	**0**	**221**	**30**	**11%**

点击某个“子文件夹”，则以该文件夹下面的各个具体文件为单位显示其统计结果，如下图：



文件名称	字数		字符数		CJK字符数		句数		进度
	原文	译文	原文	译文	原文	译文	原文	译文	
中国汽车加速“出海”.txt	1121	108	1151	604	1017	0	26	7	18%
中国汽车市场前景广阔 中文.txt	747	0	777	0	655	0	19	0	
汽车工程项目	**1868**	**108**	**1928**	**604**	**1672**	**0**	**45**	**7**	**11%**

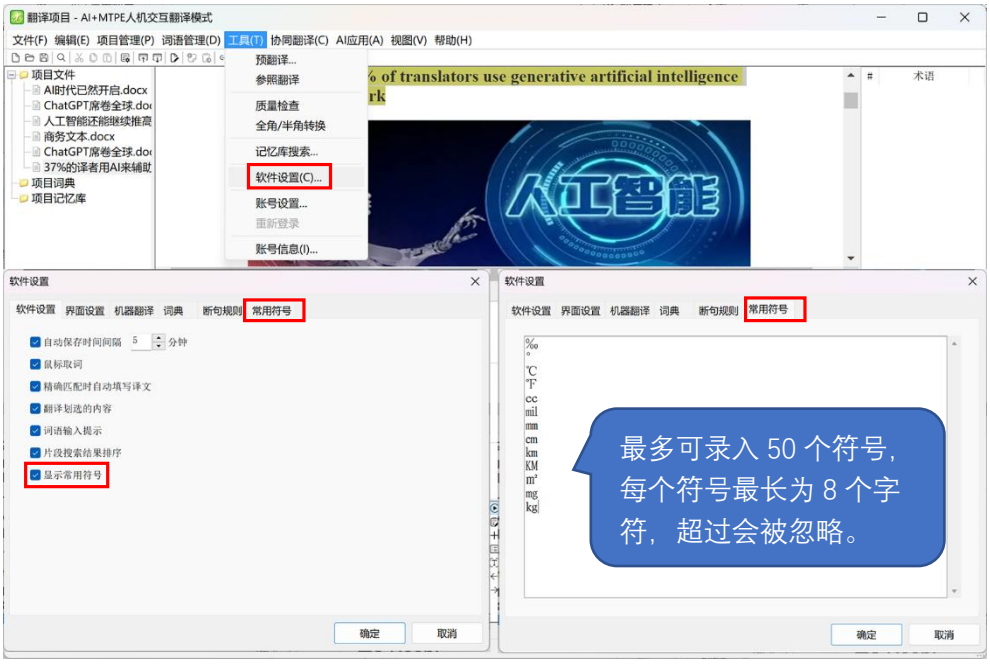
“CJK”是中文（Chinese）、日文（Japanese）、韩文（Korean）三国文字的缩写，统计时仅统计文字部分，标点符号、英文字母、数字等均不统计在内。

字数统计功能同样适用于“双语对齐”项目。

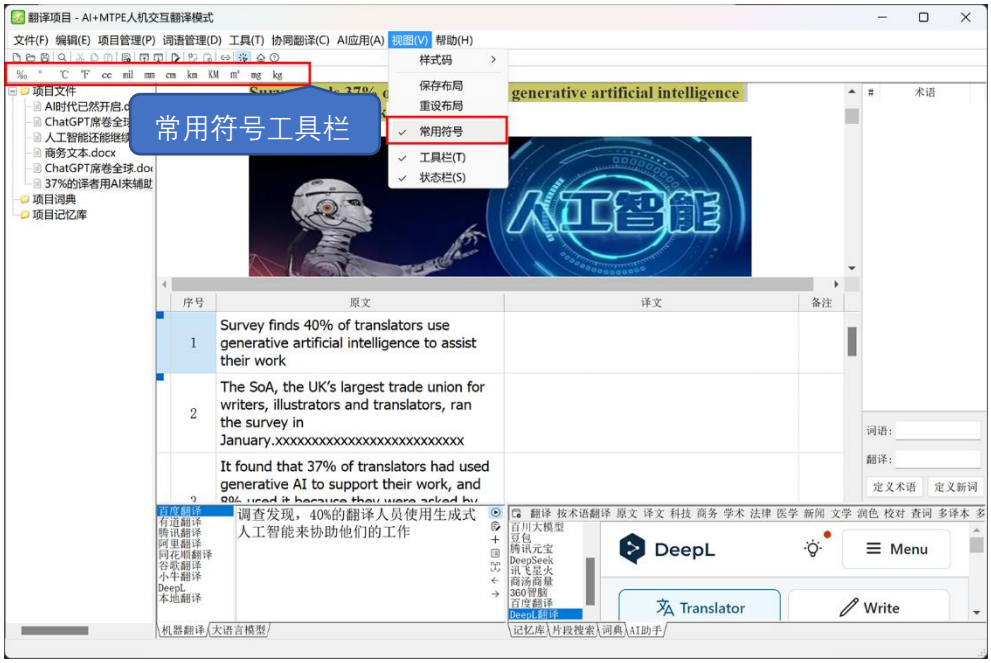
第十节 常用符号工具条

翻译时经常会遇到一些特殊符号的输入:例如‰、℃、℉、㎡等，为了方便输入，可以将这些符号录入在“常用符号”工具栏中，使用时点击工具栏中的符号即可输入。

点击【**工具**】➔【**软件设置**】菜单，在“软件设置”窗口中，有“显示常用符号”选项。如果希望默认显示出常用符号工具条，可以勾选这个选项。如果当前未显示这个工具栏，可以点击【**视图**】➔【**常用符号**】菜单设置显示状态。选择软件设置中的“常用符号”页签，可以输入自己需要的常用符号，在“常用符号”窗口中，最多可录入 50 个符号，每个符号的长度最长为 8 个字符，超过会被忽略。如下图:

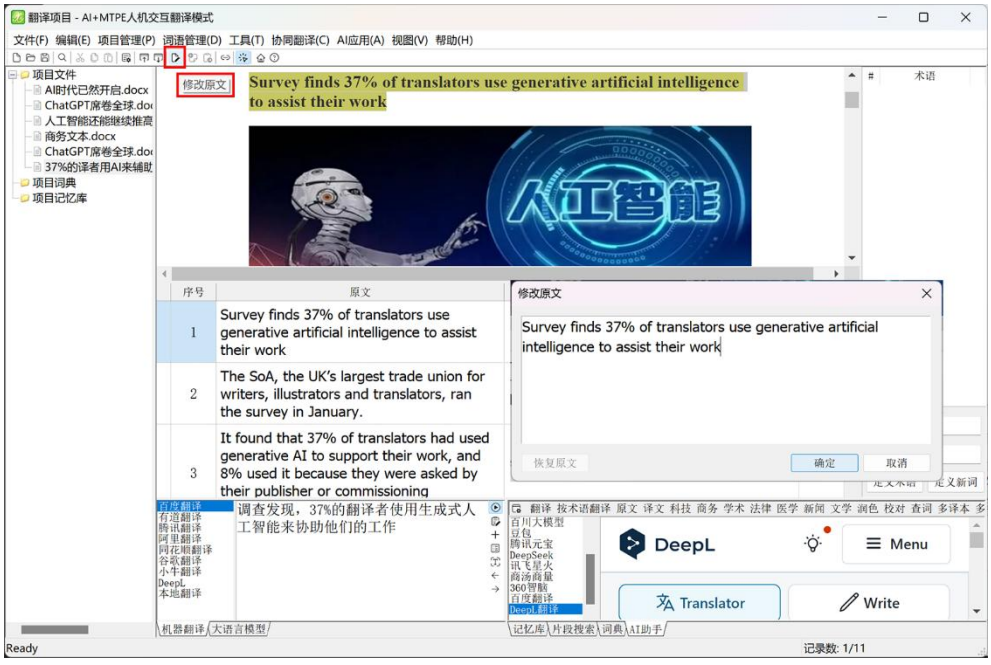


如果需要关闭“常用符号”工具栏，可以选择【**视图**】➔【**常用符号**】菜单设置隐藏。如下图:



第十一节 修改原文

点击工具栏的“ | 修改原文”按钮，在可修改当前句的原文，如下图：



对于修改了原文的句子，会在该句的左上角标注一个蓝色的小方块图标以示区分。如下图：修改原文后的句子无法拆分和合并，在恢复原文后可以重新进行调整。



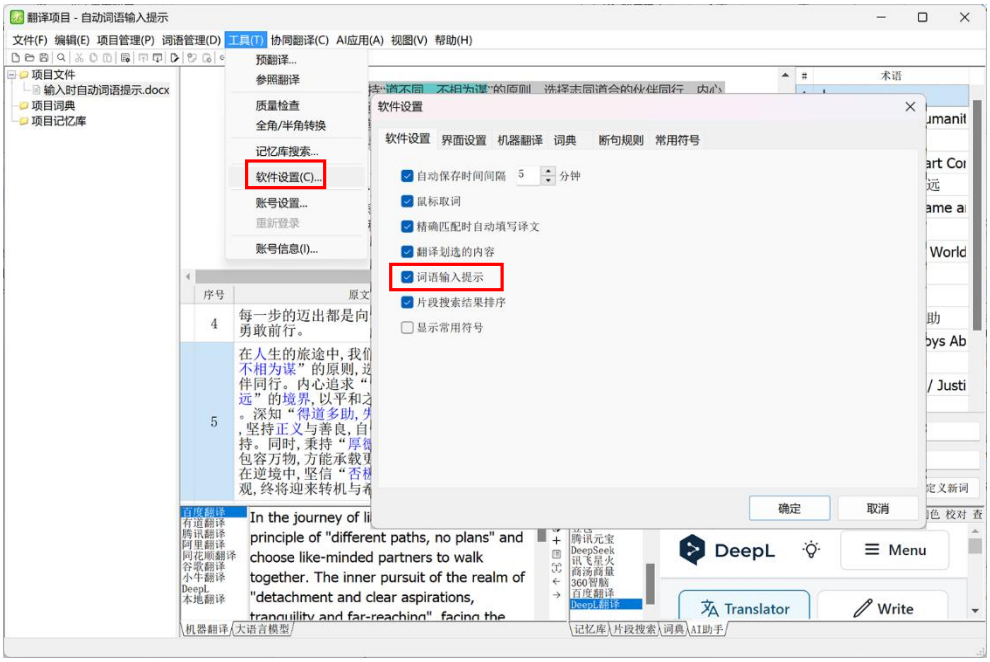
点击窗口左下角的“恢复原文”按钮，可以将原文恢复。

注意：修改原文仅用于翻译及保存的记忆库，不会对原始的原文进行修改（导入时的文档）；若译文为空的句子，导出的译文是原始的原文而不是更改后的原文。

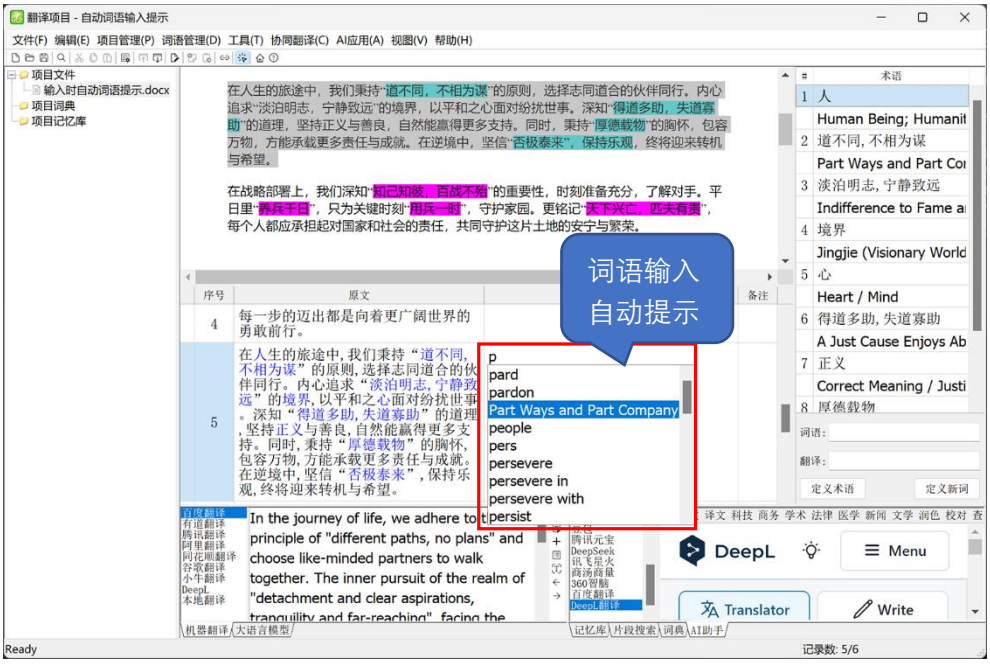
第十二节 词语输入自动提示

词语输入自动提示功能，可提升输入速度与准确性。该功能可以实时分析用户的输入，根据用户词典、或用户自己定义的词语、术语（包括正则表达式词典中定义的词语），推荐可能的词语，有效缩短输入时间。对于拼写不准确或不熟悉的单词，自动提示可以帮助用户正确拼写，减少拼写错误。

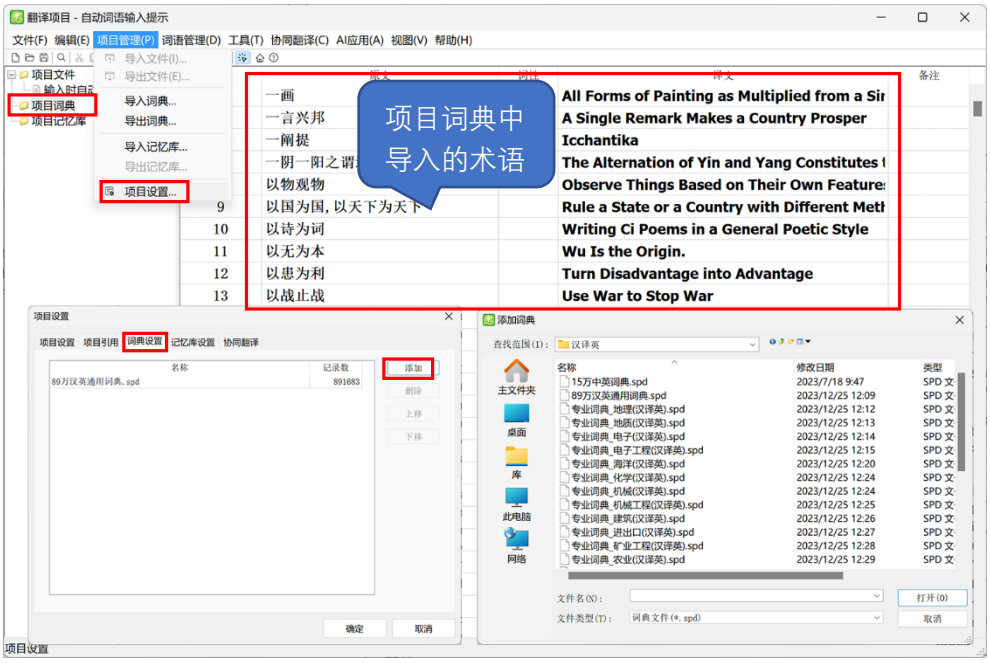
点击【**工具**】➔【**软件设置**】菜单，在“软件设置”窗口有“词语输入提示”选项，默认勾选了这个选项。如下图:



例如：将“道不同,不相为谋　Part Ways and Part Company”作为新词定义或术语添加到“项目词典”中，当翻译含有该短语的句子时，在键入第一个字母“P”时，“Part Ways and Part”会自动弹出来了，按【↑】、【↓】光标键及回车键可以直接选用，按【Tab】键亦可。

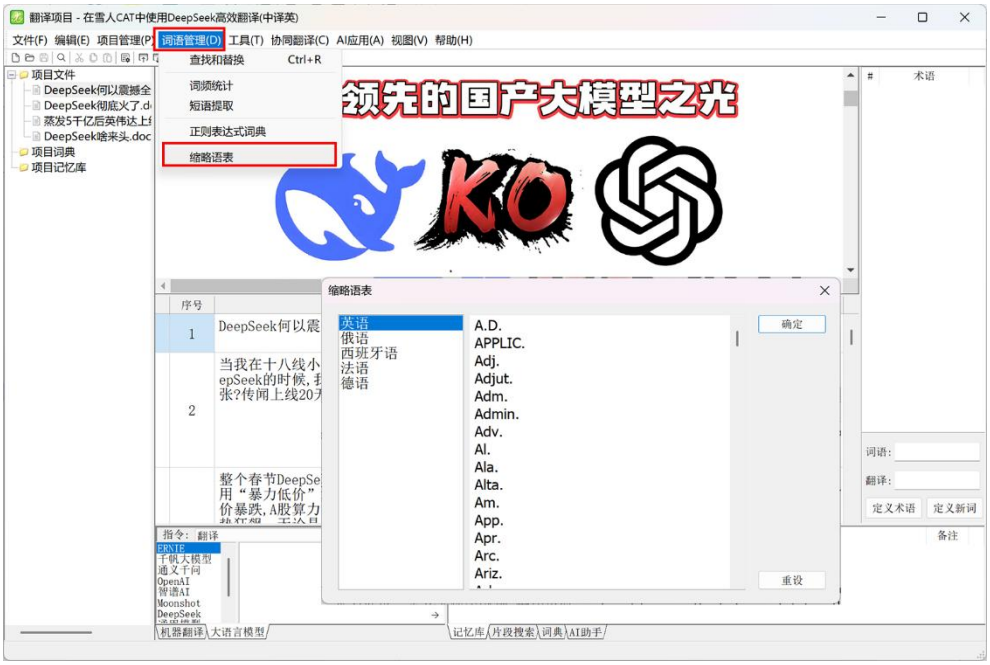


词语输入自动提示功能可以正常使用的前提是：当前翻译项目要有词典或术语。可在屏幕左侧的“项目词典”中导入自己的词典，也可点击**【项目管理】➔【项目设置】**菜单，在“词典设置”窗口中，点击“添加”按钮，添加词典文件。如下图：



第十三节 缩略语表

缩略语表主要用于解决不该断句的问题，对于断句不正确的地方，将有些词语加入到缩略语表中，再重新导入文档，就可以解决一些断句不正确的问题。点击**【词语管理】➔【缩略语表】**菜单，用户可以在“缩略语表”窗口中，添加自己的缩略语，点击下方的“重设”按钮，则恢复软件的默认“缩略语表”。



第十四节 雪人 CAT 账号级别

雪人将软件区分为：会员版与免费版。免费版功能作了适度精简。对于追求卓越翻译效率与品质的用户，敬请申请成为我们的 VIP 或会员，解锁更多高级功能！具体的功能权限介绍，请登录雪人官网：http://www.gcys.cn/vip.html 查看。

点击主菜单**【工具】➔【账号信息】**功能，可以查看当前用户的账号级别，升级为“VIP 用户”或“会员用户”后，在“账号信息”窗口中，可以看到“VIP 用户”或“会员用户”的使用时长，如下图：

