

唐库听读网介绍



唐库听读网介绍。

目 录

[唐库听读平台介绍](#)

[导入导出文本](#)

[鼠标悬浮查词功能](#)

[内容隐藏显示功能](#)

[查找替换及信息提取统计](#)

[如何为标题或关键词添加连续序号？](#)

[Markdown格式简要说明](#)

[正则表达式](#)

[版权保护期限](#)

[关于我们](#)

[征稿启事](#)

你好，欢迎来到唐库听读网！

唐库听读网是以文字阅读和语音听读相结合的文库网站，你可以阅读文字，眼睛看累了你可以听（鼠标双击文本即可）。上传文本即可听读自己的内容，支持导入txt, pdf, docx, epub文件(支持目录导入)。

文库专辑分为公开版与私有版，公开版所有用户均能看到，私有版仅能自己看到，但可以通过阅读码分享给其他人。

唐库听读网语音朗读功能调用了浏览器的语音朗读引擎，朗读质量取决于浏览器：

- 1，实测微软的Edge浏览器朗读质量最好，音色也多，断句也更智能；
- 2，手机安卓浏览器仅有Chrome，Edge支持朗读，国内的均不支持，但即使是安卓Edge浏览器，其语言朗读功能较PC削弱很多，只有一种音色，朗读时也不支持游标跟随；
- 3，苹果Iphone, Ipad, Mac的Safari浏览器均支持全功能朗读（比安卓就是优秀），但也比不过Edge，建议苹果系统上也安装Edge。

[Edge浏览器安卓版下载](#)

[Edge浏览器Mac版下载](#)

[Edge浏览器Windows版下载](#)

[微软官网下载Edge](#)

苹果Iphone, Ipad, Mac可在AppStore上下载Edge。

Edge的多种朗读角色中，不同角色智能不同，很多角色标了Online，实测并不产生网络流量。

唐库听读网是一个自助式的阅读平台，可以将txt文件轻松导入唐库，会自动拆分章节，请参考导入导出文本。

唐库听读网朗读时文字同步滚动，章节可以自动切换，尤其适合大屏幕，电视机/宽屏显示器+千元级迷你PC主机（无风扇静音那种），即可达到天花板级的阅读体验，再也不用担心眼睛疼了。

可以将一个文本文件的内容拆分为多个文章，甚至形成树状结构，都是支持的，下面来举例说明。

在编辑界面顶部工具栏从右端数第5个为“导入导出文本”，点击弹出对话框，点击“打开txt文件”，再点击“智能拆分章节”，即可看到拆分好的章节，继续点“导入拆分章节”，导入完成，刷新网页即可看到，再点存盘按钮右侧的发布按钮，发布完成，再点右上角“阅读界面”，即可听读制作好的内容。

点击“智能拆分章节”时，在第2个输入框会给出当前标题的正则表达式，例如：

```
(\n|^) * *.{1,26} * *\n
```

下面进行解释（你也可以让AI解释正则表达式）：

“`(\n|^)`” 这部分里\n表示换行符，|表示或者，它需要括起来，^表示文件开头，整体含义就是标题的最前面是换行符或者文件开头；

“`* *`” 这部分星号表示0到多个，两个星号前面的是英文和中文空格，意思是说标题以换行或文件头位置开始，接下来可以有0到多个中英文空格；

“`.{1,26}`” 这部分里句点表示除了换行符号以外的其它任意字符，{1,26}表示最少1个最多26个，意思是标题最少1个字，最多26字；

“`* *\n`” 表示标题的后面也有0到多个中英文空格；

\n 标题最后以换行符结尾。

你需要检查拆分好的章节里面有无错误，如有错，可调整上面的正则表达式，然后点“用标题关键词拆分章节”，当然如果原始文本不符合规律你需要改原始文本。

如想了解如何拆分标题，可看下面例子：





最上面的内容包含两篇文章：

文章1：标题1（内容：aaa bbb ccc）

文章1：标题2（内容：aaa bbb ccc）

如何将这两篇文章拆分呢？

关键词是标题2字，输入标题，点击“预览拆分文本”，即可看到拆分结果：

- 1, 标题:标题1, 长度:11字符
- 2, 标题:标题2, 长度:15字符

如何形成树状目录结构呢？如下图所示：



ddd
ccc

第1章 | 第2节 yyyy
ddd
eee
fff
ggg

第2章 | 第1节 zzzz
hhh
iii
jjj

节

替换词，默认支持用正则表达式。

打开文本文件 查找下一个 查找上一个 替换 删单个换行符

1, 标题:第1章 | 第1节 xxxx, 长度:11字符
2, 标题:第1章 | 第2节 yyyy, 长度:15字符
3, 标题:第2章 | 第1节 zzzz, 长度:11字符

导入文本文件到服务器 预览拆分文本 导入拆分文本到服务器

不同层级的标题用'|'连接，拆分关键词输入“节”（注意这要求文章内容不包含这个字，如果包含，拆分关键词可输入“第\d+节”，这是正则表达式，\d+表示1到多个数字）
这样导入后就能形成树状目录结构，例如像《资治通鉴》就是双层目录，300多万文字，都是一键导入的。

唐库导出的文本是可以原样导入的，参考一下导出的文本，你就知道如何导入目录了。

唐库目前可以导入txt, pdf, docx, epub 四种文件，可以导入txt（通过标题拆分），epub中的目录；pdf, docx暂不支持，可以通过LibreOffice(免费软件)先转为epub再导入。

还可以导入目录下的所有txt, docx文件，磁盘目录结构可转变成文章的目录结构。

在唐库听读平台的阅读界面及写作界面，将鼠标移到文字上面，即可弹出这个词条的解释。
如果是英文单词，即弹出英文单词解释：

👁 浏览 23 📄 扫码 🔄 分享

World War I : It killed innocence in literature too 第一

News / Insight 新闻/洞

innocence

['ɪnəsəns]

n. 无罪, 无知, 天真无邪

[法] 无罪, 无罪的人, 清白无辜者

The dreadful war's soldier - writers "shed the convent understated irony." 可怕的战争使得士兵作家们“抛弃了

A scene from director Lewis Milestone's All Quiet on t Maria Remarque. 路易斯·迈尔斯通导演的电影《西线：

By : Jim Coyle News, Published on Fri Jun 20 2014 ↑

如果是汉字，即弹出百科词条：

《了不起的盖茨比》 语录

了不起的盖茨比

- 1 弗·司各特·菲茨杰拉德著短篇小说
- 2 美国1974年杰克·克莱顿执导电影
- 3 2000年劳拉·邓恩主演电影
- 4 2010年发行解谜游戏
- 5 2011新翻拍电影《了不起的盖茨比》

1 弗·司各特·菲茨杰拉德著短篇小说

《了不起的盖茨比》是美国作家弗·司各特·菲茨杰拉德所著的一部以20世纪20年代的纽约市及长岛为背景，被设定在现代化的美国社会中上阶层的白人圈子中展开。《了不起的盖茨比》问世，奠定了菲茨杰拉德在美国文学史上的地位，成了20年代“爵士时代”的代表作家之一。20世纪末，美国学坛长河中选出一百部最优秀的小时，它傲然跻身当代经典行列。另有根据小说改编的

简介

内容简介

作者简介

写作界面可关闭此功能。

如果在手机上，轻触文字即可弹出提示。

唐库的文章支持部分内容的隐藏及显示。

例如双语阅读读者希望能在隐藏中文和显示中文之间来回切换。

唐库通过内容切换标签实现此功能：

将文章的内容用<tang1> </tang1>包围，例如：

```
<span><tang1> aaa </tang1></span>
```

```
<span><tang2> bbb </tang2></span>
```

```
<span><tang3> ccc </tang3></span>
```

内容“aaa”，“bbb”，“ccc”就可以隐藏及显示了。

唐库最多三种内容切换标签（tang1, tang2, tang3），每种标签名字均可自定义。

以下是测试：

这是内容1

这是内容2

这是内容3

点击顶部工具栏上的“内容1”，“内容2”，“内容3”试试看。“内容1”，“内容2”，“内容3”的名称是可以随意设置的，在专辑设置界面的内容切换标签处设为为“内容1,内容2,内容3”即可。

唐库的查找替换具有以下特殊功能：

- 1) 用户自定义词表替换功能，把多个词替换成为另外的词；
- 2) AI替换功能，查找一个词条，用AI处理该词条，替换为AI处理的结果；
- 3) 信息提取统计功能，支持关键词提取，统计；
- 4) 预览替换功能，在替换前先预览一下，避免出错；
- 5) 可以无限扩充的词条转换功能（例如转大写，转小写）。

1 用户自定义词表替换功能：

用户有一个自定义替换词表：

aaa,bbb

ccc,ddd

意思是 aaa 要替换为 bbb, ccc 要替换为 ddd, 如何操作呢？

The screenshot shows a 'Global Content Replacement' dialog box with the following fields and options:

- 源字符串 (Source String):** (?<userlist>)
- 选项 (Options):** A, H, * (selected), AI替换, 替换词表, 历史, 帮助, 新开网页
- 替换内容 (Replacement Content):** \$1 = \$@1
- 替换词表 (Replacement Table):**

```

aaa,bbb
ccc,ddd
eee,fff

```
- 指定行 (Specify Line):** 所在行关键词, 全部行 (selected), 前, 后
- 指定文章 (Specify Article):** test4, 指定文章 (selected), 包含目录
- 预览替换 (Preview Replacement):**

```

预览替换, 找到 2 处:
test4 | 1行 || aaa 这是原文
test4 | 1行 || aaa = bbb 这是原文

test4 | 3行 || ccc 测试
test4 | 3行 || ccc = ddd 测试

```
- 底部按钮 (Bottom Buttons):** 查找, 查找全部, 预览替换 (highlighted), 替换, 统计

查找目标 (?<userlist>) 的含义是：

这是一个正则表达式，括号组表示分组，“?<userlist>”定义了userlist是分组名，userlist是预定义分组名，有特殊含义，就是查找用户自定义词表中的词。

替换目标 \$1 表示第一个分组（小括号组）匹配的内容（\$0表示整个查找串），如果该分组是userlist, 则是用户自定义词表中的逗号前的词，\$@1表示第一个分组经过运算后的结果，如果该分组是userlist则是用户自定义词表中的逗号后的词。

从“预览替换”可看到，\$1的内容是“aaa”，\$@1的内容是“bbb”。

更加常用的用法是查找 `\b(?<userlist>)\b`

`\b` 的含义匹配单词边界，例如空格或标点符号，不加`\b`像“1aaa1”会替换为“1bbb1”，加了`\b`后只有单独的“aaa”替换为“bbb”了。

2 AI替换功能：

AI替换功能可以用于英文文章的人名地名的翻译，找到这些名称，然后用AI翻译，将翻译结果替换原名。这功能必须借助[正则表达式](#)的分组，下面举例说明一下：

全局内容替换 ✕

源字符串

选项 A H * [AI替换](#) [替换词表](#) [历史](#) [帮助](#) [新开网页](#)

替换内容

Prompt

选择AI 唐库初级AI

指定行 全部行 前 后

指定文章 指定文章 包含目录

预览替换，找到 1 处：
test4 | 5行 || **John** is my friend.
test4 | 5行 || **John = 约翰** is my friend.

John,约翰

查找
查找全部
预览替换
替换
统计

解释：

查找目标 (?!<ai>[A-Z][a-z]+) 的含义是：

- 1) 括号表示正则表达式分组；
- 2) (?!<ai>...)表示该分组名称为“ai”，该名称为预定义名称有特殊含义，会调用AI查询该分组匹配的词条；
- 3) [A-Z] 表示A-Z之间任意大写字符；
- 4) [a-z]+ 表示a-z之间任意小写字符，数量为1到多个；

替换目标 $\$1 = \$@1$ 的含义是：

$\$1$ 代表第一个小括号中的字符串($\$0$ 代表整个字符串)

$\$@1$ 代表第一个小括号中匹配的字符串经过运算后的结果

从“预览替换”可看到， $\$1$ 的内容是“John”， $\$@1$ 的内容是“约翰”，这是AI翻译的结果。

例子2：



全局内容替换

源字符串: 式子 (?<ai>[0-9\\(|\\)|+|-|*|V|+)

选项: A H * AI替换 替换词表 历史 帮助 新开网页

替换内容: $\$1 = \$@1$

Prompt: \$source
计算上面的数学式子，请返回最简短答案，不要加任何备注或解释。

选择AI: 唐库初级AI 78691881

指定行: 所在行关键词 全部行 前 后

指定文章: test4 指定文章 包含目录

预览替换, 找到 1 处:
test4 | 7行 || 计算式子 $1+(2+3)*2$
test4 | 7行 || 计算 $1+(2+3)*2 = 11$

$1+(2+3)*2, 11$

查找 查找全部 预览替换 替换 统计

解释：

查找目标 (?<ai>[0-9()+-.*V|+]) 的含义是：

- 1) 括号表示正则表达式分组；
 - 2) (?<ai>...)表示该分组名称为“ai”，该名称为预定义名称有特殊含义，会调用AI查询该分组匹配的词条；
 - 3) [0-9()+-*/]+表示由0-9字符和小括号以及 + - * / 构成的式子；
- 替换目标 \$1 = @\$1 的含义是：
- \$1 代表第一个小括号中的字符串(\$0 代表整个字符串)
- @\$1 代表第一个小括号中匹配的字符串经过运算后的结果
- 从“预览替换”可看到，\$1的内容是“1+(2+3)*2”，@\$1的内容是“11”，这是AI计算的结果。

类似上面userlist, ai这样的预定义名称还有：

- 1) tolower 转小写功能
- 2) toupper 转大写功能
- 3) index 返回连续序号12345
- 4) indexcn 返回连续序号一二三四五
- 5) len 返回捕获组匹配的字符的个数

3 信息提取统计功能:

全局内容替换

源字符串: [A-Z]{3,20}

替换内容: 请输入要替换成的内容

指定行: 所在行关键词 全部行 前 后

指定文章: 文章所在目录关键词 全部文章 统计关键词

查找全部, 找到 1316 处:
 共有关键词种类: 46
 关键词1: ROMEO, 数量: 199
 关键词2: JULIET, 数量: 147
 关键词3: CAPULET, 数量: 128
 关键词4: NURSE, 数量: 115
 关键词5: BENVOLIO, 数量: 78
 关键词6: MERCUTIO, 数量: 74
 关键词7: FRIAR, 数量: 73
 关键词8: LADY, 数量: 69
 关键词9: LAWRENCE, 数量: 67
 关键词10: PARIS, 数量: 38
 关键词11: PETER, 数量: 38
 关键词12: TYBALT, 数量: 31

A H * 退出 查找 查找全部 预览替换 替换

信息提取统计功能需要[正则表达式](#)，上图是通过正则表达式提取《罗密欧与朱丽叶》专辑中的大写人名，“[A-Z]{3,20}”表示要查找的词由A-Z等大写字母组成，字符个数3-20个，这样取出来的基本都是人名。

4 查找替换换行符的问题

\n代表换行符，可在替换内容中使用\n即可。

当查找串中不含\n时，唐库按行内容查找，正则表达式中的 ^, \$ 匹配每一行的开头及末尾。

当查找串中含\n时，唐库按文章内容查找，正则表达式中的 ^, \$ 匹配整篇文章的开头及末尾。