



MyBox：简易工具箱
用户手册 - 文档工具

作者：Mara

版本：6.7.1

日期：2023-3-13

内容目录

1 简介.....	3
1.1 主菜单.....	3
1.2 资源地址.....	4
1.3 文档.....	5
1.4 工具的菜单.....	7
2 树形信息.....	8
3 笔记.....	9
4 PDF 工具.....	10
4.1 查看 PDF 文件	10
4.2 批量转换.....	10
4.3 批量提取.....	11
4.4 其它功能.....	11
5 文本编辑的基础.....	12
6 Markdown 工具.....	13
6.1 Markdown 编辑器	13
6.2 批量转换.....	13
7 文本工具.....	14
7.1 文本编辑器.....	14
7.2 其它功能	14
8 网页工具.....	15
8.1 网页编辑器	15
8.2 查找网页中的字串.....	16
8.3 查找网页中的元素.....	17
8.4 网页截图.....	18
8.5 其它功能.....	18
9 微软的文档格式.....	19
9.1 处理 PPT 文档	19
9.2 处理 Word 文档	20
9.3 其它功能.....	20
10 编辑字节.....	21
11 系统粘贴板中的文本.....	22
12 MyBox 粘贴板中的文本.....	23

1 简介

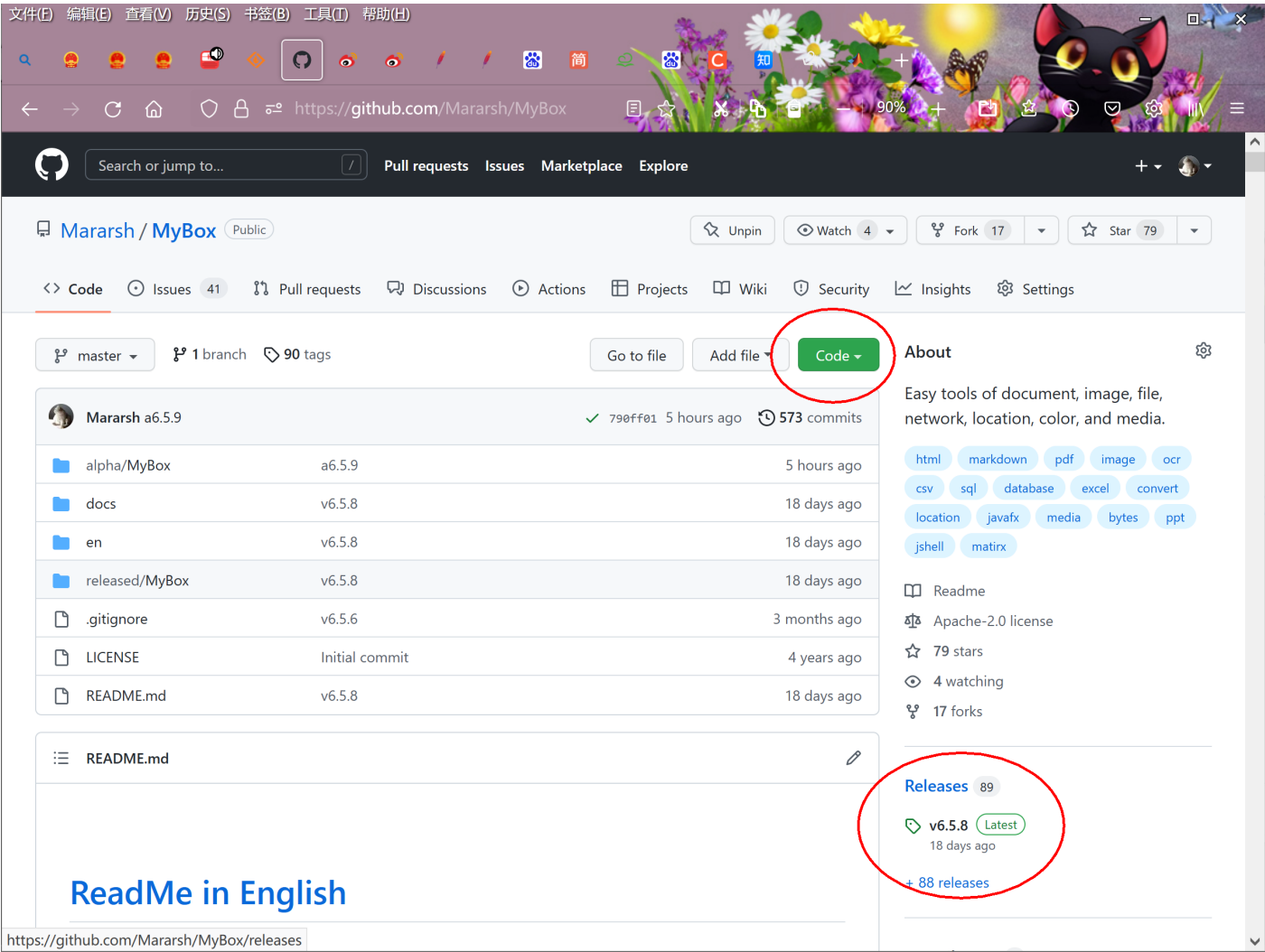
这是利用 JavaFx 开发的图形化桌面应用，目标是提供简单易用的功能。免费开源。

1.1 主菜单



1.2 资源地址

内容	链接
项目主页	https://github.com/Mararsh/MyBox/
源代码和编译好的包	https://github.com/Mararsh/MyBox/releases
提交软件需求和问题报告	https://github.com/Mararsh/MyBox/issues
数据	https://github.com/Mararsh/MyBox_data
文档	https://github.com/Mararsh/MyBoxDoc
镜像	https://sourceforge.net/projects/mara-mybox/files/
云盘	https://pan.baidu.com/s/1fWMRzym_jh075OCX0D8y8A#list/path=%2F



1.3 文档

文档名	版本	修改时间	英文	中文
开发日志	6.7.1	2023-3-13	html	html
快捷键	6.7.1	2023-3-13	html	html
功能列表	6.7.1	2023-3-13	html	html
打包步骤	6.7.1	2023-3-13	html	html
开发指南	2.1	2020-8-27	PDF odt	PDF odt
用户手册-综述	6.7.1	2023-3-13	html PDF odt	html PDF odt
用户手册-数据工具	6.7.1	2023-3-13	html PDF odt	html PDF odt
用户手册-文档工具	6.7.1	2023-3-13	html PDF odt	html PDF odt
用户手册-图像工具	6.7.1	2023-3-13	html PDF odt	html PDF odt
用户手册-文件工具	6.7.1	2023-3-13	html PDF odt	html PDF odt
用户手册-网络工具	6.7.1	2023-3-13	html PDF odt	html PDF odt
用户手册-媒体工具	6.7.1	2023-3-13	html PDF odt	html PDF odt
用户手册-开发工具	6.7.1	2023-3-13	html PDF odt	html PDF odt
示例-笔记	6.7.1	2023-3-13	html	html
示例-树形	6.6.1	2022-11-16	html	html
示例-收藏的网址	6.6.1	2022-11-16	html	html
示例-SQL	6.6.1	2022-11-16	html	html
示例-JShell	6.6.1	2022-11-16	html	html
示例-JEXL	6.6.1	2022-11-16	html	html
示例-JavaScript	6.6.1	2022-11-16	html	html
示例-数学函数	6.6.1	2022-11-16	html	html
示例-行过滤	6.6.1	2022-11-16	html	html
关于-数据分组	6.7.1	2023-3-13	html	html
关于-颜色	6.6.1	2022-11-16	html	html
关于-坐标系统	6.6.1	2022-11-16	html	html
关于-媒体	6.6.1	2022-11-16	html	html
关于-数据分析	6.6.1	2022-11-16	html	html
关于-Java 编程的一句话事项	6.7.1	2023-3-13	html	html
关于-JavaFx 编程的一句话事项	6.7.1	2023-3-13	html	html
调色盘-美术色相环(RYB)24 色	6.7.1	2023-3-13	rgba all	rgba all

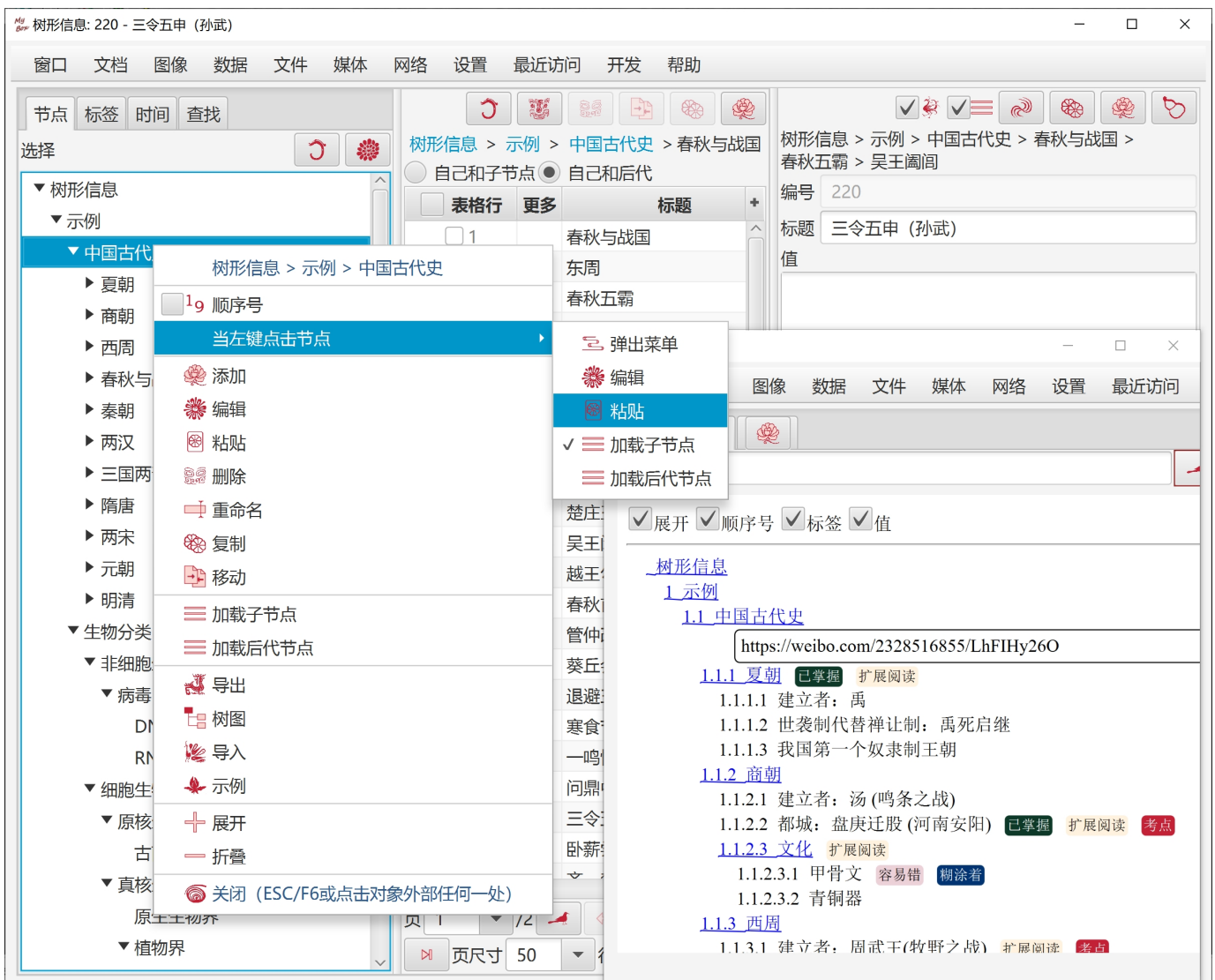
文档名	版本	修改时间	英文	中文
调色盘-常用网页颜色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-中国传统颜色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-日本传统颜色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-来自 colorhexa.com 颜色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-美术色相环(RYB)12 色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-美术色相环(RYB)360 色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-光学色相环(RGB)12 色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-光学色相环(RGB)24 色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-光学色相环(RGB)360 色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-美术颜料	6.7.1	2023-3-13	rgba all	rgba all
调色盘-MyBox 的颜色	6.7.1	2023-3-13	rgba all	rgba all
调色盘-灰阶	6.7.1	2023-3-13	rgba all	rgba all
图片的故事	6.7.1	2023-3-13	html PDF odt	html PDF odt

1.4 工具的菜单



2 树形信息

1. 信息被组织为一棵树。
2. 对任意节点可执行：增删改子节点、重命名、移动、复制、导出、显示树图、展开、折叠。
3. 编辑节点：
 - 节点名不应包含字符串“>”（“>”前后是空格）。
 - 一个节点可以设置多个标签。
 - 可选显示节点的序列号。
4. 导出：
 - 格式：文本（用于导入）、单个网页、网页框架、xml。
 - 可选择是否导出时间、标签。
 - 可选择文件字符集。
 - 可设置网页风格。
5. 导入。提供示例。
6. 查询：
 - 节点的子节点或所有后代。
 - 按标签。可选多个标签。
 - 按更改时间。 可选多个时间。
 - 按标题或内容包含的字符串。 可输入多个关键字。

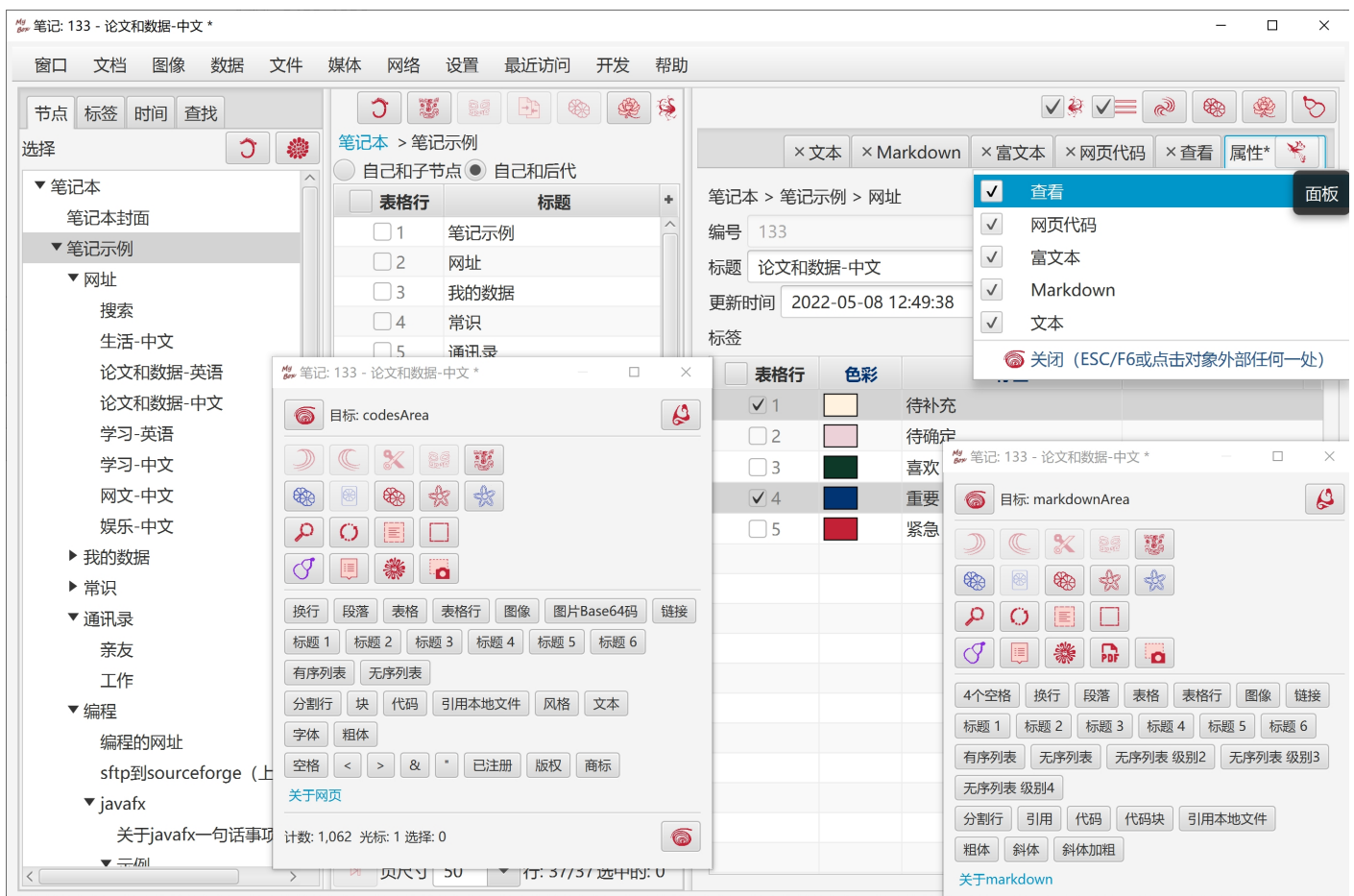


3 笔记

1. 笔记是 html 格式的信息片段：

- 笔记可以用五种模式来编辑，即：“网页代码”、“DOM”、“富文本”、“Markdown”、“文本”。
- 每个编辑模式都可以单独修改网页。
- 当前页签下的编辑模式为“当前编辑模式”。
- 点击按钮“同步”(F10)，把当前编辑模式中的修改应用到其它模式。
- 点击按钮“保存”(F2)，保存当前编辑模式的修改，并将修改自动同步到其它模式。
- 笔记的 html 代码应是"body"的内容，不要包含标签“html”、“head”、和"body"。
- 可对笔记设置风格，仅用于显示，不会保存在笔记的代码中。

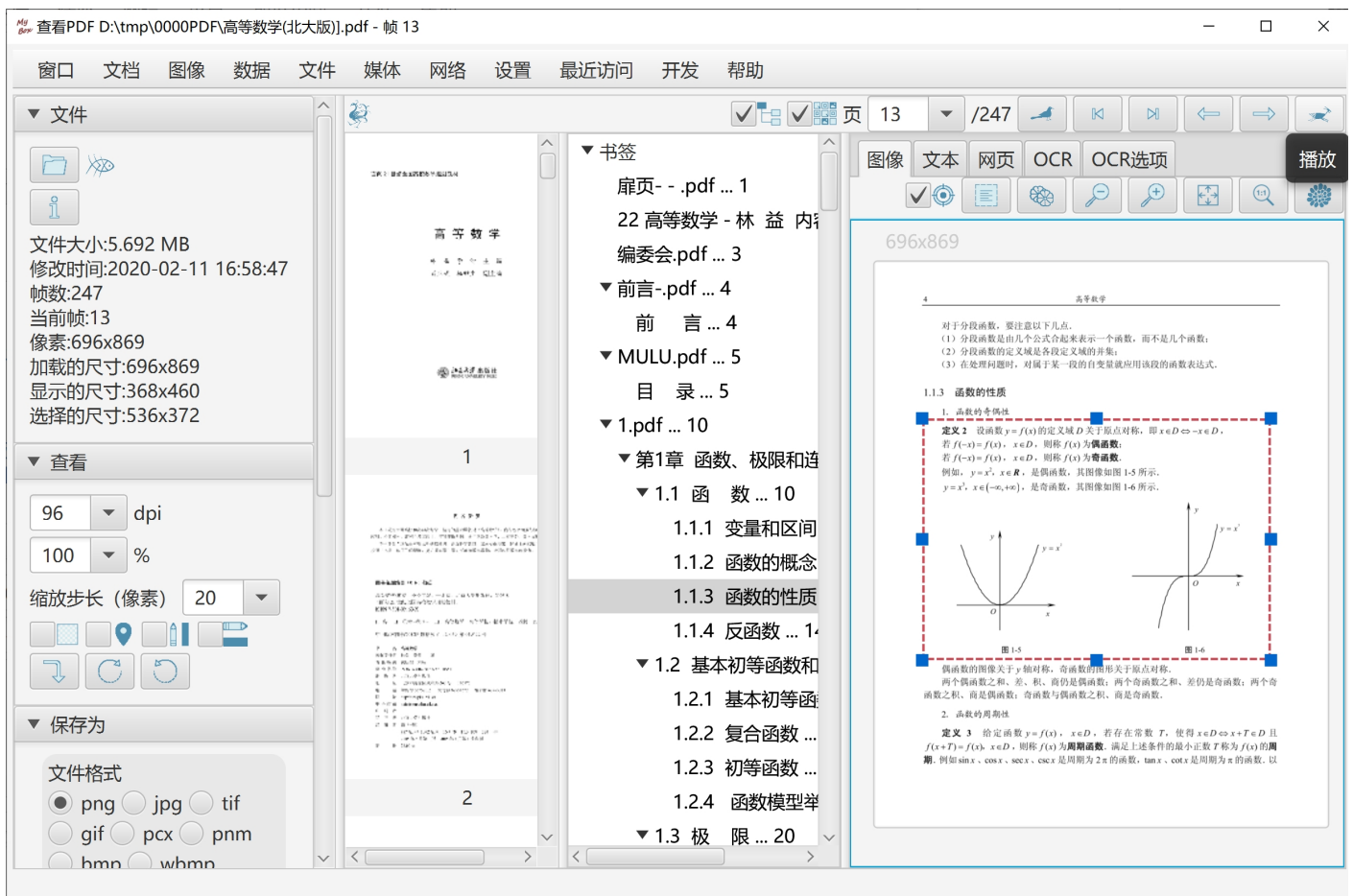
2. 笔记本是由笔记构成的树状信息。



4 PDF 工具

4.1 查看 PDF 文件

1. 标签和缩略图
2. 每页被转换为一张图片。可设置 dpi 以调整清晰度
3. 提取页面上的文字
4. 转换页面为一个网页
5. 识别页面中的文字 (OCR)



4.2 批量转换

1. 将 PDF 文件的每页转换并保存为一张图片文件，包含图像密度、色彩、格式、压缩、质量、色彩转换等选项。
2. 转换 PDF 文件中被选择页的图片，并保存为新的 PDF。
3. 压缩 PDF 文件中的图片，保存为新的 PDF。可设置 JPEG 质量或者黑白色阈值
4. 将 PDF 转换为网页，可选：每页保存为一个 html、还是整个 PDF 保存为一个 html；字体文件/图像文件是嵌入、单独保存、还是忽略。

4.3 批量提取

1. 将 PDF 中的图片提取出来。
2. 将 PDF 文件中的文字提取出来，可以定制页的分割行。
3. 识别 PDF 文件中图片的文字（OCR）。

4.4 其它功能

1. 分割 PDF 文件为多个 PDF 文件，可按页数或者文件数来均分，也可以设置起止列表。
2. 合并多个 PDF 文件。
3. 将多个图片合成 PDF 文件
4. 写 PDF 的选项：页面尺寸、图片属性、字体文件、页边、页眉、作者等。
5. 修改 PDF 的属性，如：标题、作者、版本、修改时间、用户密码、所有者密码、用户权限等

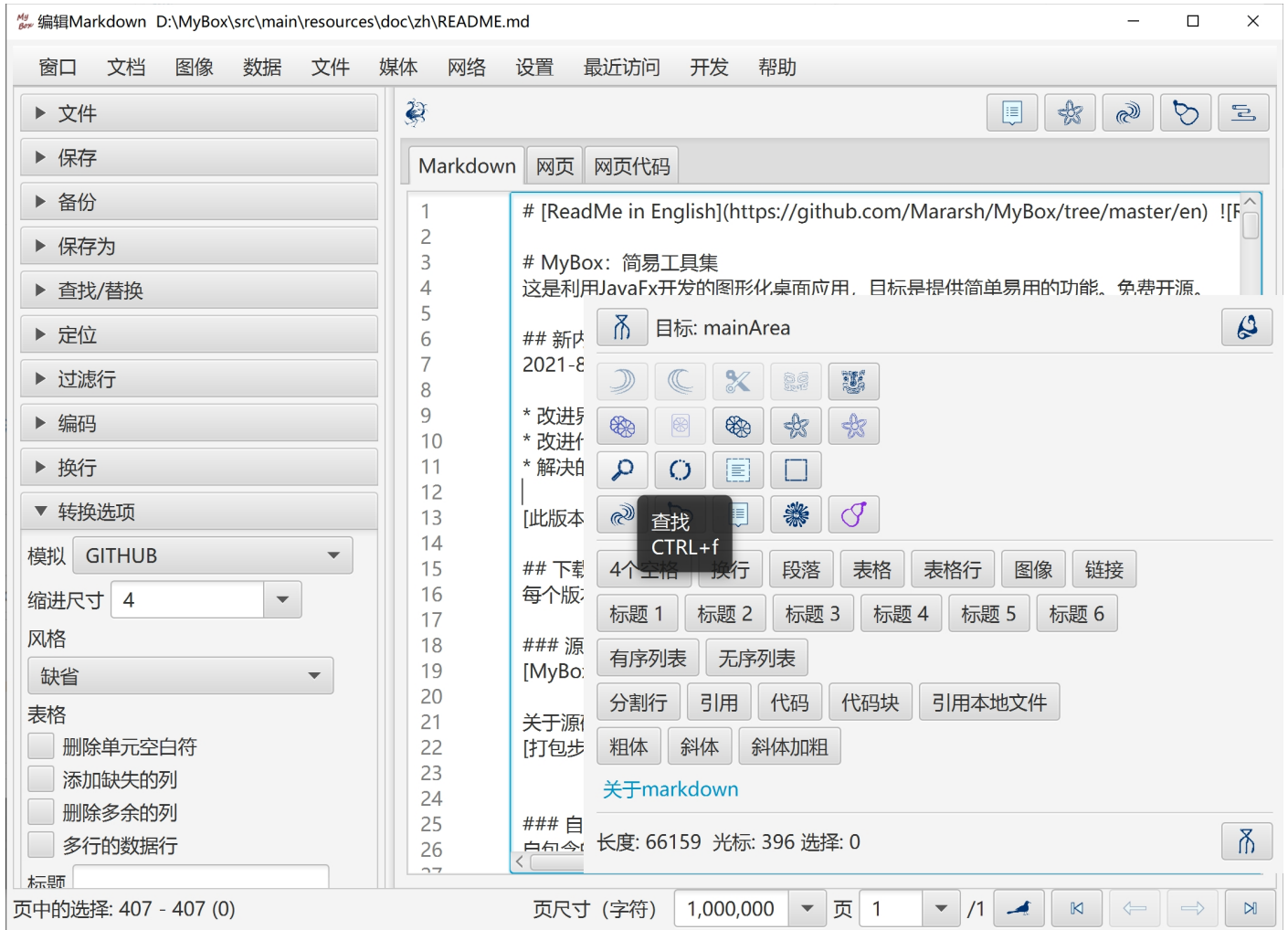
5 文本编辑的基础

1. 编辑功能（复制/粘贴/剪切/删除/全选/撤销/重做/恢复）及其快捷键。
2. 查找与替换：
 - 选项：忽略大小写、回绕。
 - 查找字串和替换字串都支持支持多行。其中的换行符按照当前文件换行符的定义来处理。
 - 支持正则表达式，提供示例。
 - 计数
 - 由于算法的限制，对于多页文档查找规则表达式时，假设：匹配的字符串的最大长度小于当前 JVM 可用内存的 1/16。
3. 定位：
 - 选择指定行号的行。
 - 选择指定位置的字符。
 - 选择指定范围的行。
 - 选择指定范围的字符。
4. 行过滤：
 - 条件：包含/不包含任一、包含/不含所有、包含/不包含规则表达式、匹配/不匹配规则表达式。
 - 可累加过滤。
 - 可保存过滤结果。
 - 可选是否包含行号。
5. 分页。可用于查看和编辑非常大的文件，如几十 G 的运行日志。
 - 设置页尺寸。
 - 页面导航。
 - 先加载显示首页，同时后端扫描文件以统计字符数和行数；统计期间部分功能不可用；统计完毕自动刷新界面。
 - 对于跨页字符串，确保查找、替换、过滤的正确性。
6. 定时自动保存。只在有修改时保存。
7. 保存时自动备份。可恢复到备份。

6 Markdown 工具

6.1 Markdown 编辑器

- 提供输入格式的按钮。
- 同步显示转换的网页和网页代码



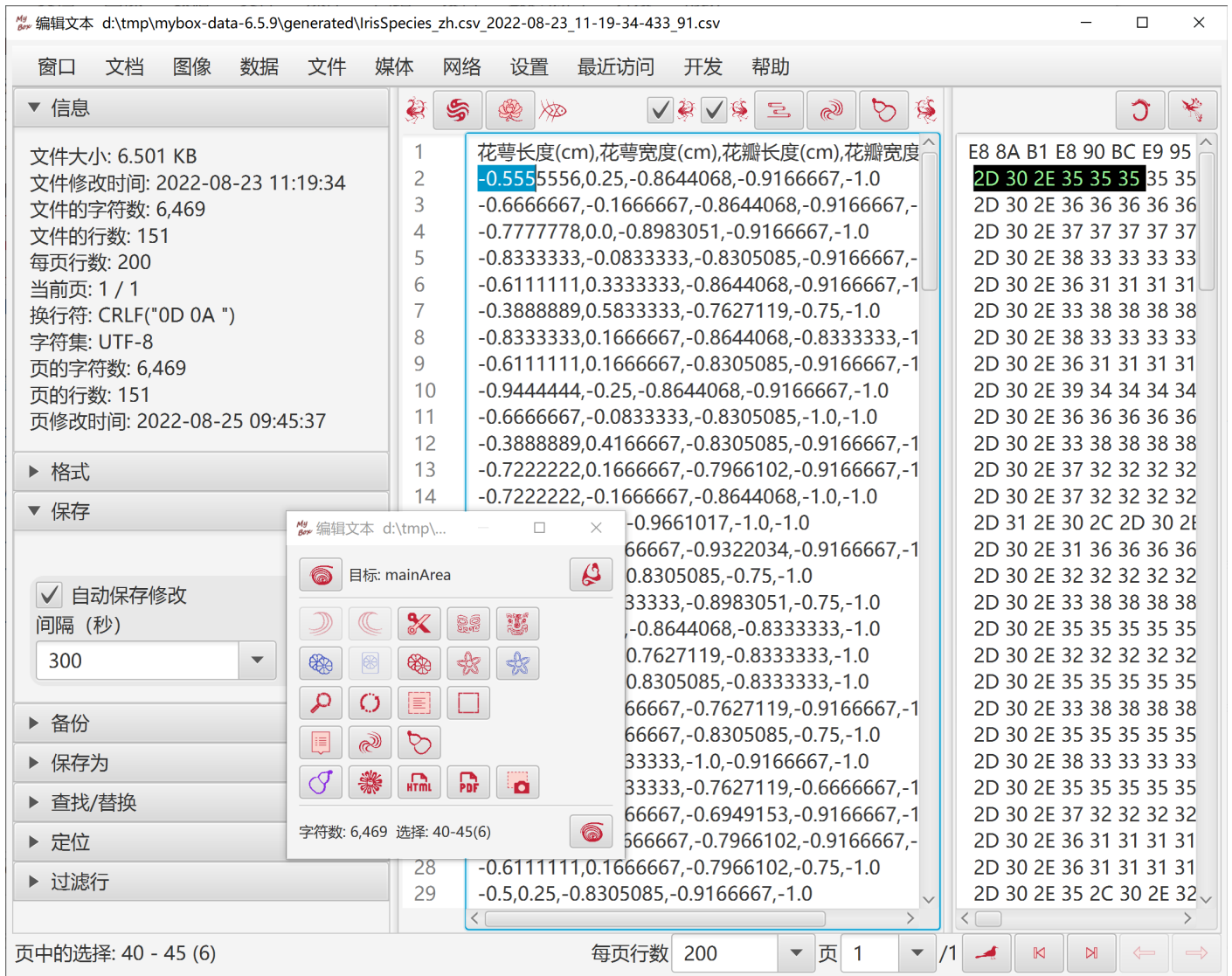
6.2 批量转换

1. 批量转换 Markdown 为网页。
2. 批量转换 Markdown 为文本。
3. 批量转换 Markdown 为 PDF。

7 文本工具

7.1 文本编辑器

1. 自动检测或手动设置文件编码；改变字符集实现转码；支持 BOM 设置。
2. 自动检测换行符；改变换行符。显示行号。
支持 LF（Unix/Linux）、CR（Apple）、CRLF（Windows）。
3. 字符集对应的编码：字节的十六进制，同步显示、同步选择，可选同步滚动、同步更新。



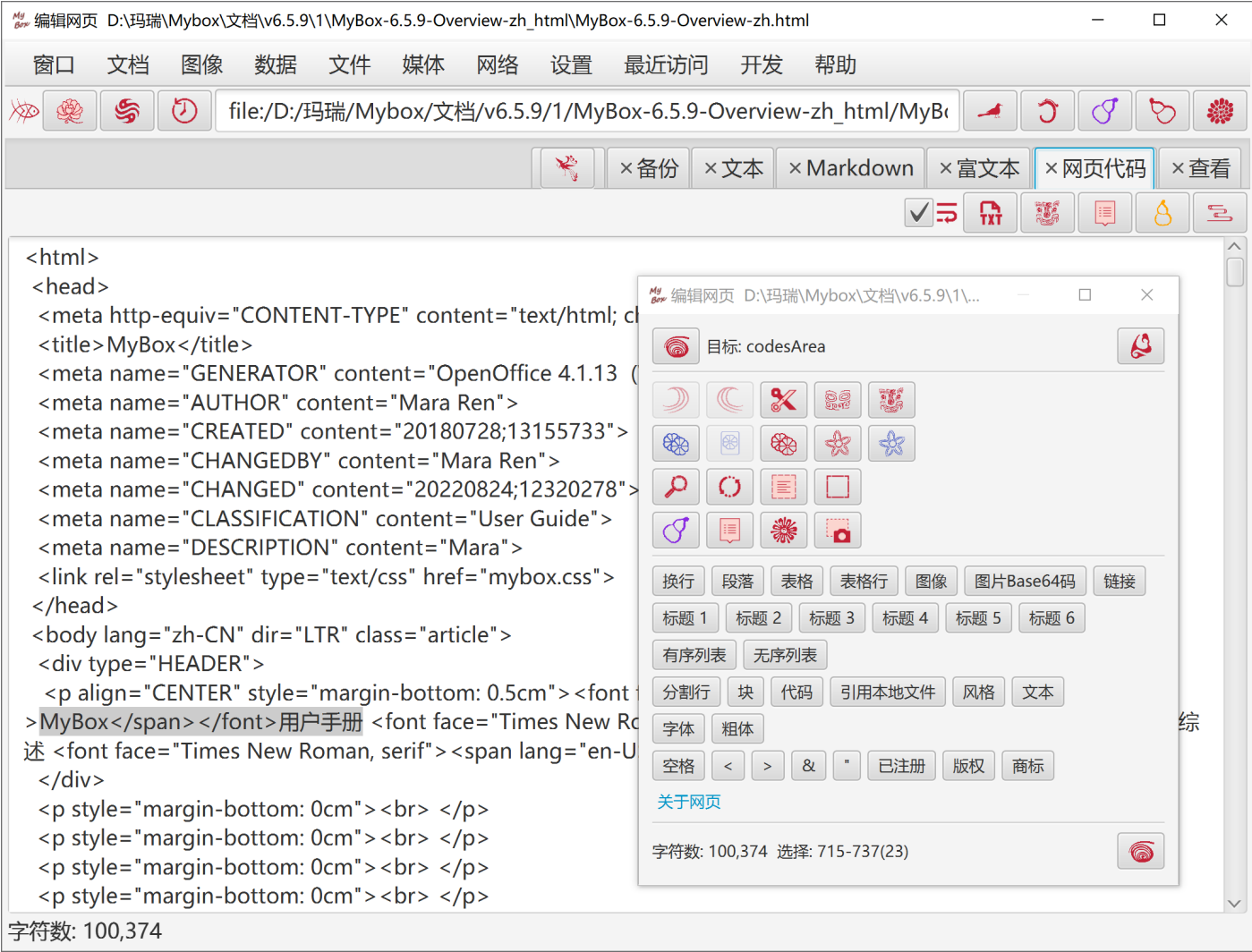
7.2 其它功能

1. 批量转换/分割文本文件。
2. 合并文本文件。
3. 批量转换文本文件为网页/PDF。
4. 批量替换文本文件中的字符串。

8 网页工具

8.1 网页编辑器

1. 加载网页的方式：打开文件、创建文件、或输入网址。
2. 网页可以用五种模式来编辑：
 - 编辑模式：“网页代码”、“DOM”、“富文本”、“Markdown”、“文本”。
 - 每个编辑模式都可以单独修改网页。
 - 当前页签下的编辑模式为“当前编辑模式”。
3. 以下按钮对当前编辑模式生效：
 - 点击按钮“同步”(F10)，把当前编辑模式中的修改应用到其它模式。
 - 点击按钮“保存”(F2)，把当前编辑模式的修改保存下来，并将修改自动同步到其它模式。
 - 点击按钮“弹出”(CTRL/ALT+p)以把当前编辑模式的内容显示在一个新窗口中。
 - 点击按钮“菜单”(F12)或右键点击编辑区域，以弹出右键菜单。
 - 点击按钮“MyBox 粘贴板”(CTRL/ALT+m)以弹出文本粘贴板来方便粘贴。
4. 对于 frameset 可以选择 frame 来编辑。



8.2 查找网页中的字串

MyBox 查找网页中的字串

窗口 文档 图像 数据 文件 媒体 网络 设置 最近访问 开发 帮助

版本迁移

1. 每个版本有自己的配置文件，新版本可以复制已安装版本的参数。

2. 每个版本处理的所有数据都在它指向的“数据目录”下。多个版本可以指向同一数据目录。

3. **MyBox**向后兼容：新版本可以处理旧版本的数据目录。而不保证向前兼容：旧版本处理新版本的数据目录时可能出错。

配置

配置文件在"用户目录"下：

平台	MyBox 配置文件的目录
win	C:\用户\用户名\mybox\MyBox_v6.5.9.ini
linux	/home/用户名/mybox/MyBox_v6.5.9.ini
mac	/Users/用户名/mybox/MyBox_v6.5.9.ini

可以临时改变配置文件：在命令行启动jar包时设置参数"config="配置文件名"”。利用“设置”功能也可以修改配置参数。

资源地址

内容	链接
项目主页	https://github.com/Mararsh/MyBox

A

条件

☐ 正则表达式

☐ 忽略大小写

字符串

MyBox

显示

找到的项目

文字颜色

背景色

当前项目

文字颜色

背景色

字体 1.2em

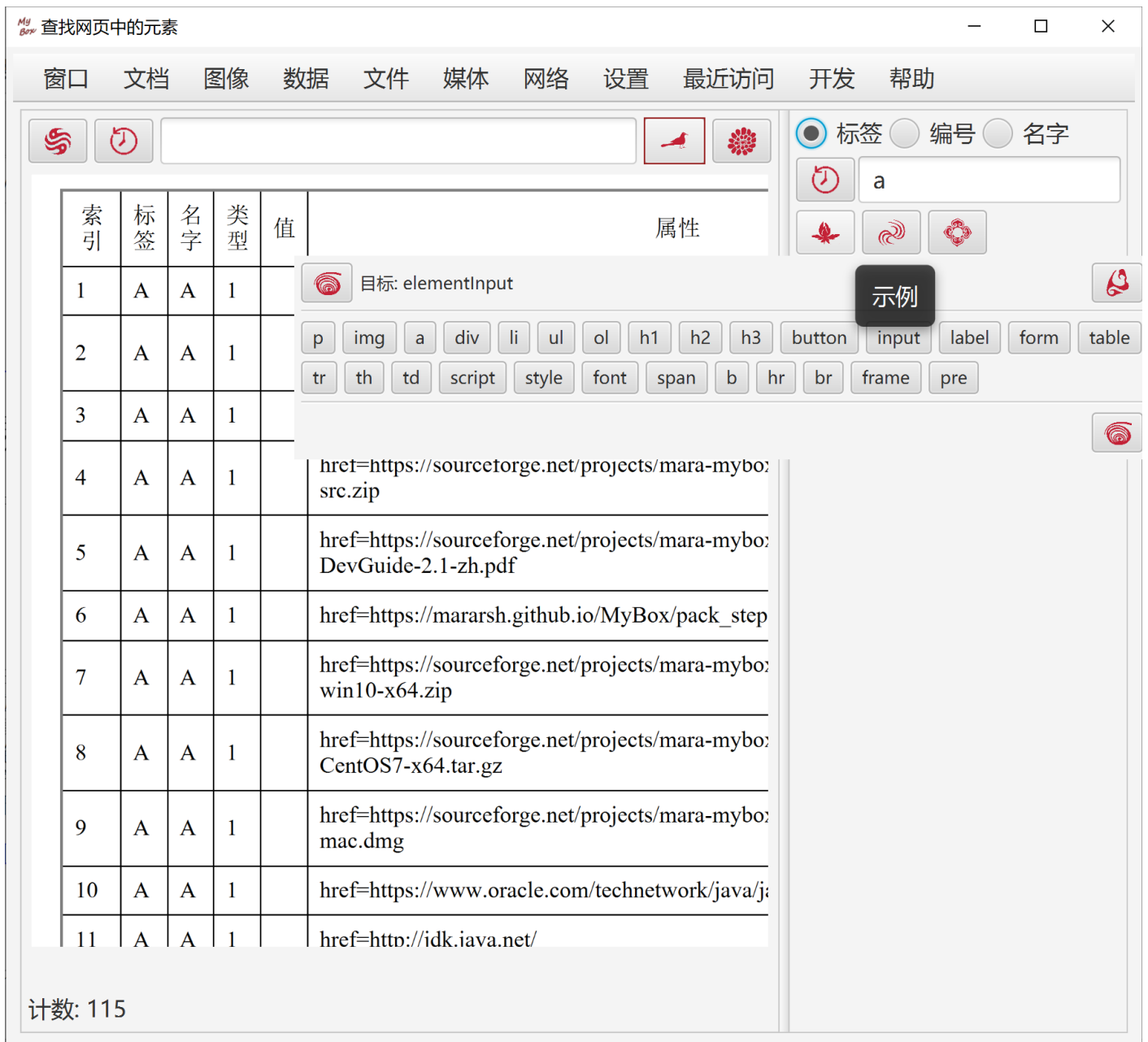
结果

找到: 33

项 18

☐ 回绕

8.3 查找网页中的元素



8.4 网页截图



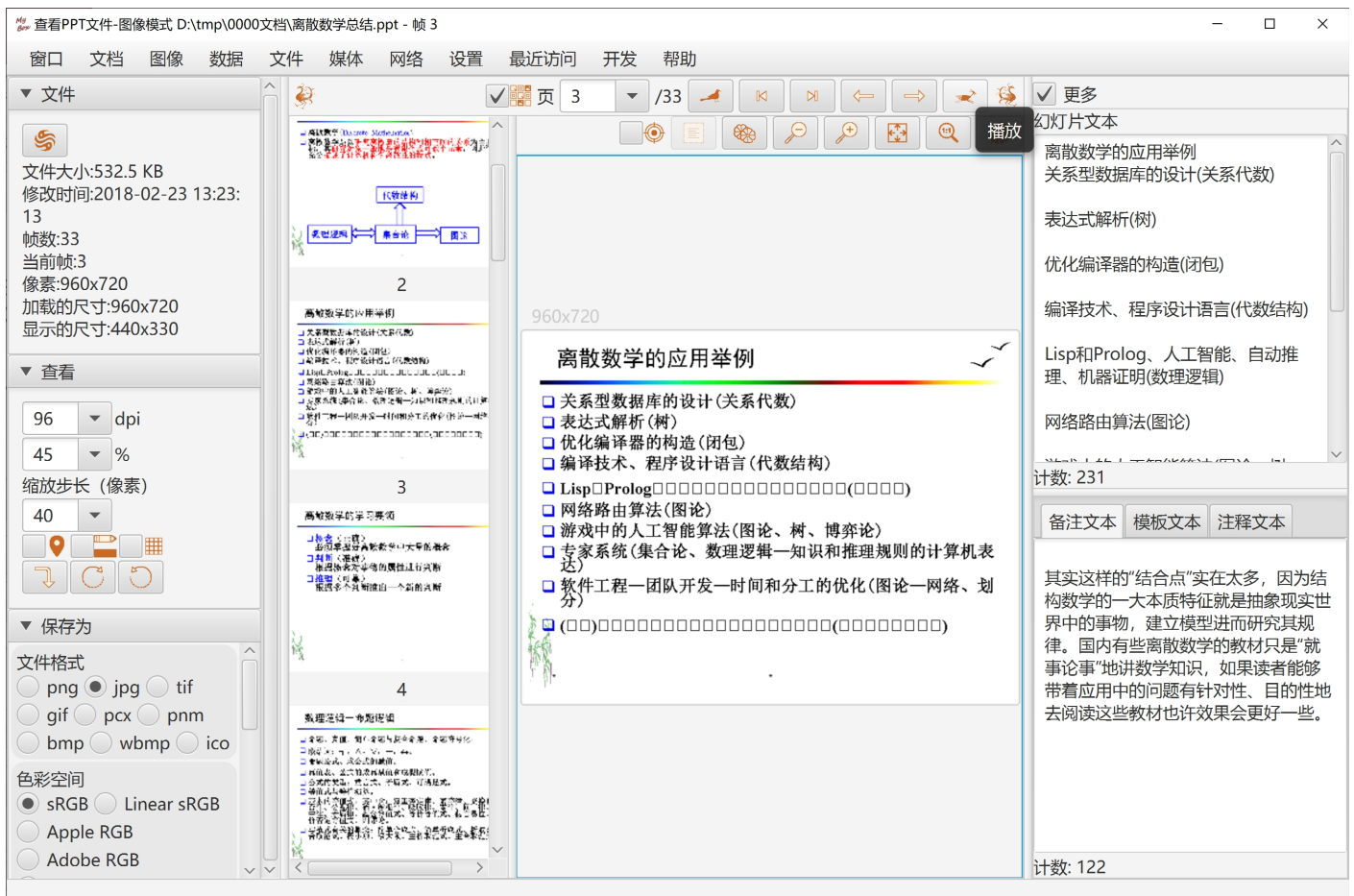
8.5 其它功能

- 1. 批量提取网页中的表格数据：
- 2. 批量转换网页为 Markdown/文本/PDF 文件：
- 3. 批量设置网页的编码/样式/http-equiv。
- 4. 合并多个网页为一个网页/Markdown/文本/PDF 文件。
- 5. 对多个文件生成框架文件。

9 微软的文档格式

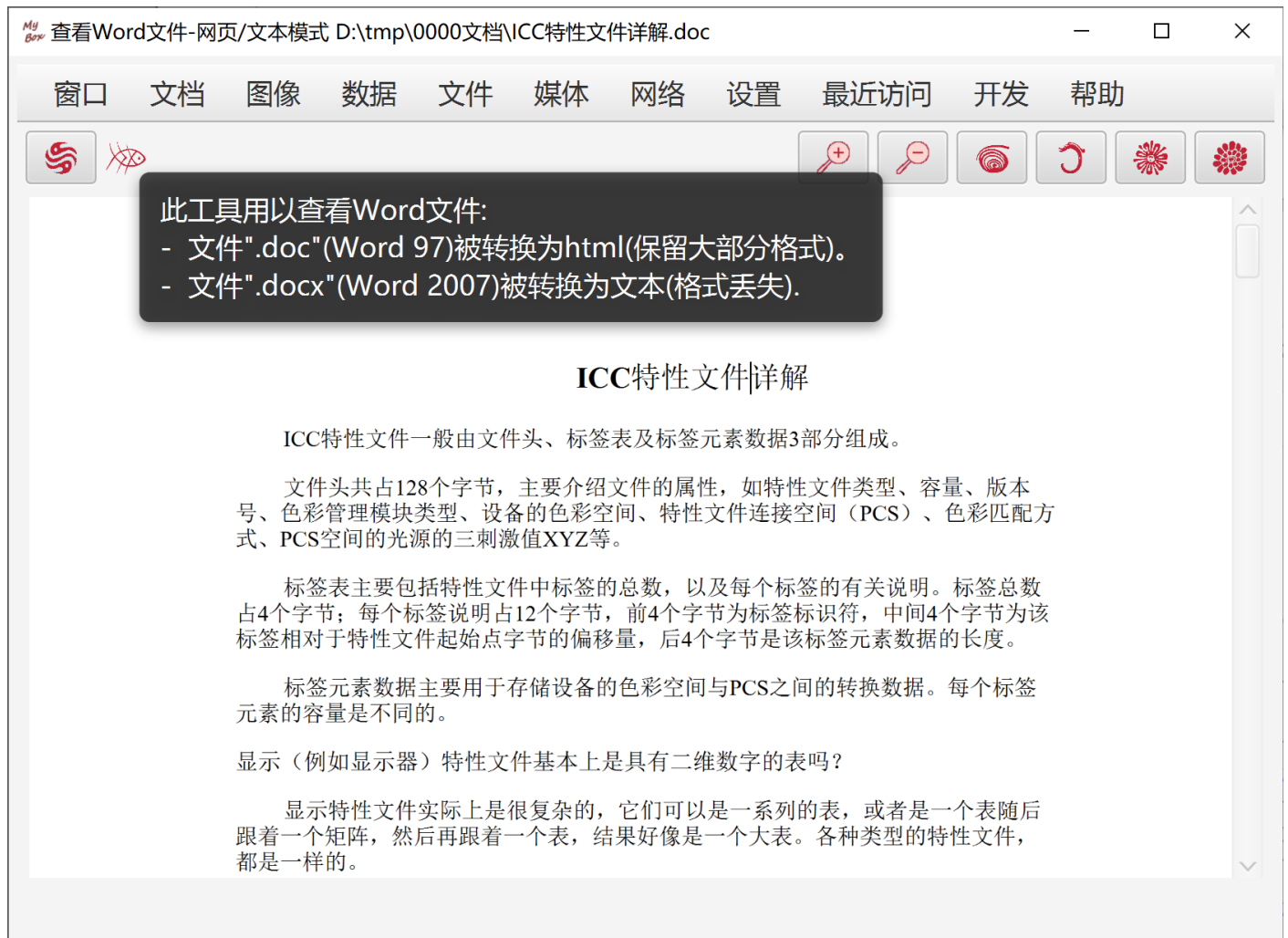
9.1 处理 PPT 文档

- 格式：
 - 文件".ppt"(PPT 97, 即 OLE 2 格式)。
 - 文件".pptx"(PPT 2007, 即 OOXML 格式)。
 不说明则同时支持这两种格式。
- 查看 PPT 文件-图像模式。逐页显示：
- 页面被转换为图片
- 幻灯片文本和备注文本
- 批量转换 PPT 文件为图片/PDF。
- 批量提取 PPT 文件中的对象：
 - 可选择提取：幻灯片文本、备注文本、模板文本、注释文本、图片、声音、OLE(Word/Excel)
 - 不支持从".pptx"中提取声音。
- 批量分割 PPT 文件。
- 合并".pptx"文件。
- 将多个图片合成为".ppt"文件。
- 播放 PPT 文件。



9.2 处理 Word 文档

1. 格式：
 - 文件".doc"(Word 97, 即 OLE 2 格式)被转换为 html, 保留大部分格式。
 - 文件".docx"(Word 2007, 即 OOXML 格式)被转换为文本, 格式丢失。
2. 查看 Word 文件。
3. 批量转换 Word 文件为 html/PDF。

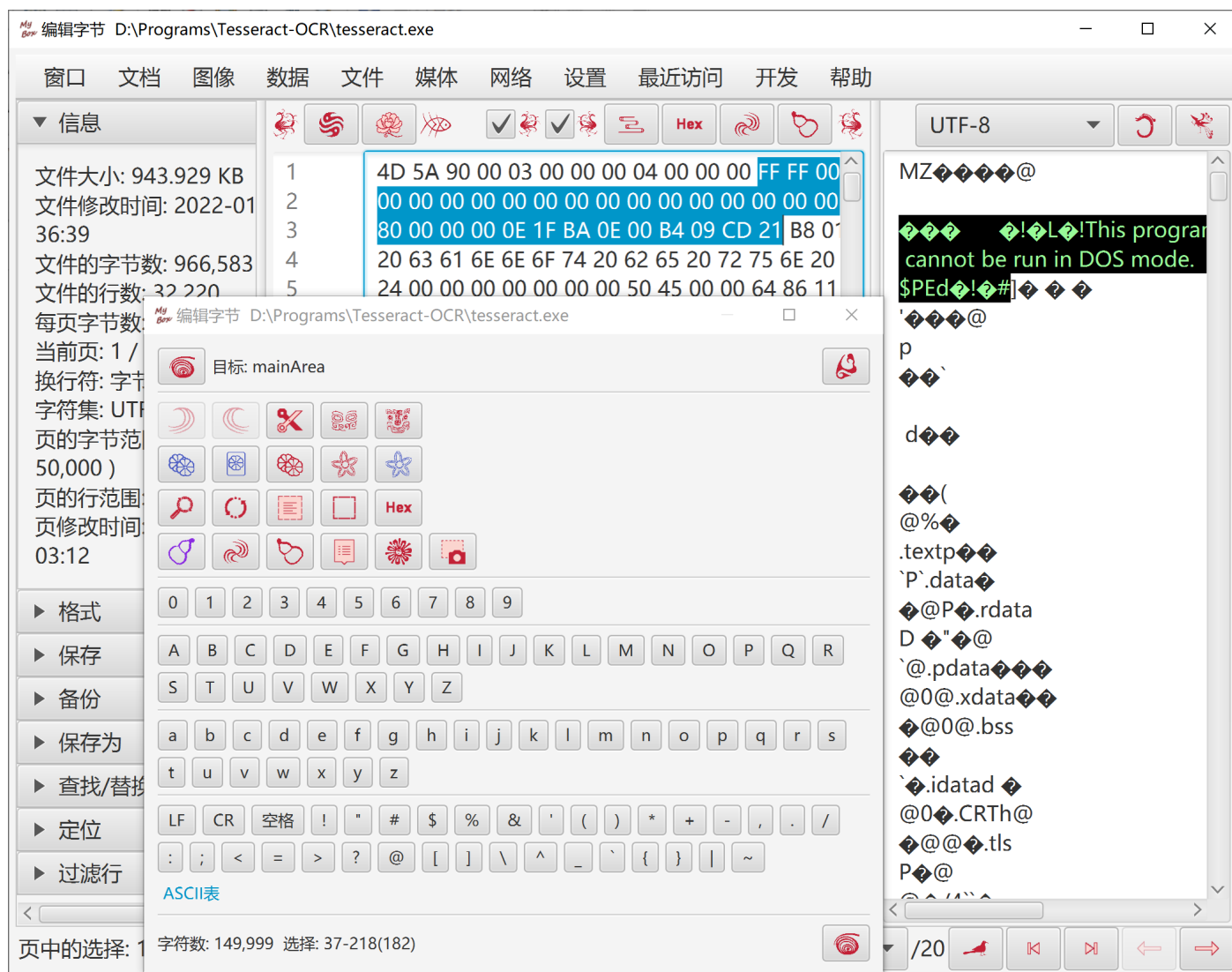


9.3 其它功能

从 Excel/Word/PowerPoint/Publisher/Visio 中提取文本。

10 编辑字节

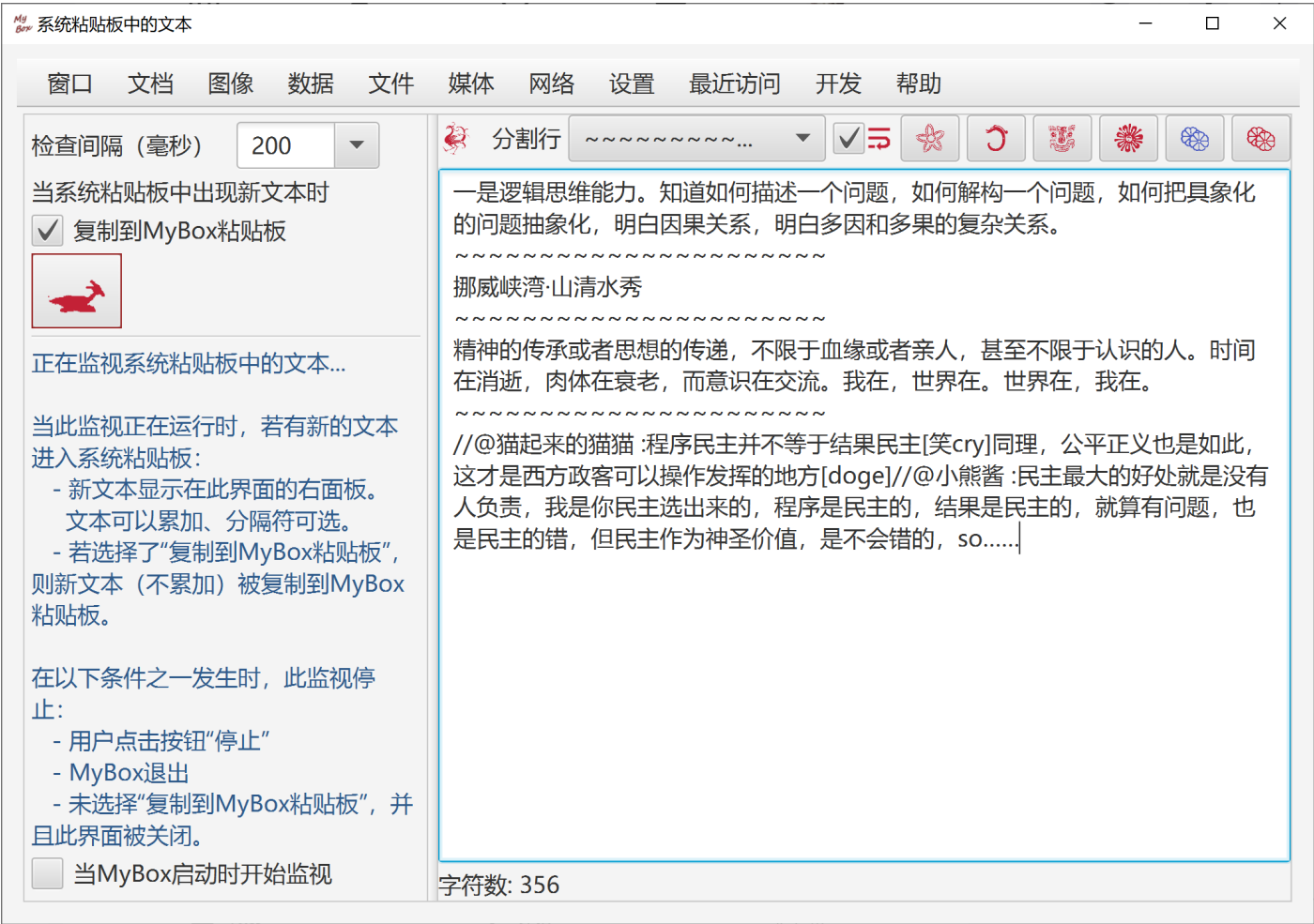
1. 字节被表示为两个十六进制字符。所有空格、换行、非法值将被忽略。
2. 常用 ASCII 字符的输入选择框。
3. 换行。仅用于显示、无实际影响。显示行号。可按字节数换行、或按一组字节值来换行。
4. 选择字符集来解码：同步显示、同步滚动、同步选择。非字符显示为问号。
5. 分页。若按字节数换行，则行过滤时不考虑跨页。



11 系统粘贴板中的文本

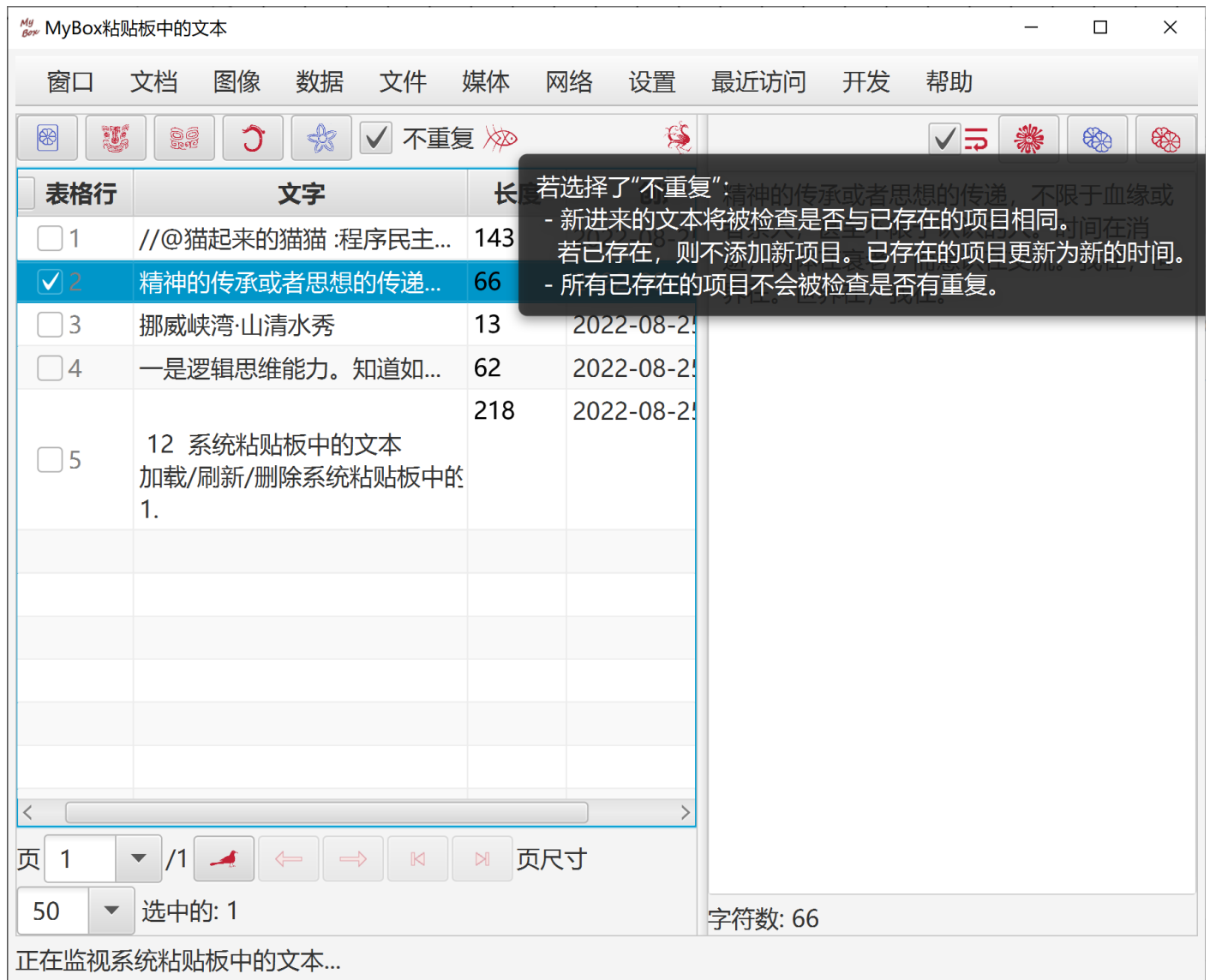
加载/刷新/删除系统粘贴板中的文本：

- 1. 点击按钮后，开始监视系统粘贴板的新文本，将它们保存到 MyBox 粘贴板中。
- 2. 可以设置监视间隔。
- 3. 在监视界面上可以累加监视到文本。分隔符可选。
- 4. 在以下条件之一发生时，此监视停止：
 - 用户点击按钮“停止”
 - MyBox 退出
 - 未选择“复制到 MyBox 粘贴板”，并且此界面被关闭。
- 5. 可选在 MyBox 启动时开始监视。



12 MyBox 粘贴板中的文本

1. 增删查 **MyBox** 粘贴板中的文本。
2. 手动添加系统粘贴板中的文本。
3. 把选择的文本复制到系统粘贴板。
4. 在所有文本输入控件的右键菜单中，可以弹出 **MyBox** 粘贴板以供选用粘贴。



<文档结束>