

索引软件用户手册

(2024-1)

索引软件由 4 个部分组成，分别是 1 个标引软件、2 个拼音排序软件、1 个格式检查软件。4 个软件的使用方法分别介绍。

1、索引标引软件

1.1 网址

索引标引软件的网址是：

<http://www.cane.ac.cn/auth>

注意：请使用 **360 极速浏览器（首先推荐）、google（谷歌）浏览器或火狐浏览器。**

打开网址，显示用户登录界面（图 1-1）。



The image shows a user login form titled "用户登录" (User Login). It features two input fields: "账号" (Account) and "密码" (Password). The "账号" field has a link "没有账号,新注册一个?" (No account, register one?). The "密码" field has a link "忘记密码?" (Forgot password?). Below the fields is a blue "登录" (Login) button.

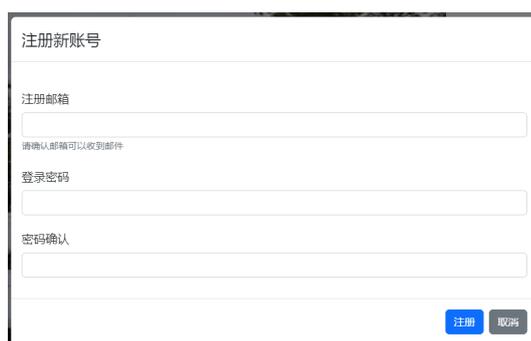
(图 1-1)

1.2 注册

新用户需要注册帐号（注意帐号是一个**邮箱号**，须牢记密码，忘记**不可**找回）。点击

账号 [没有账号,新注册一个?](#)

，出现以下信息，可以开始注册。



The image shows a "注册新账号" (Register New Account) form. It includes fields for "注册邮箱" (Register Email), "登录密码" (Login Password), and "密码确认" (Confirm Password). There is a note "请确认邮箱可以收到邮件" (Please confirm the email can receive mail). At the bottom right, there are "注册" (Register) and "取消" (Cancel) buttons.

(图 1-2)

注意：注册完成后**别忘记**点击注册界面右下角的  按钮完成注册。图 1-2。

注册完成后请及时与审核员联系（**审核员 QQ: 398536346**），以便于及时**审核**，审核通过后**重新登录**方可使用。**注意：**注册时只需要填写邮箱即可，不必填写姓名，**一定不要**填写单位等信息，这些信息待与审核员联系时另行登记。登录帐号后，出现如下界面（图 1-3）：

基本信息(使用中)

登录账号
398536346@qq.com

用户姓名

详细地址

备注

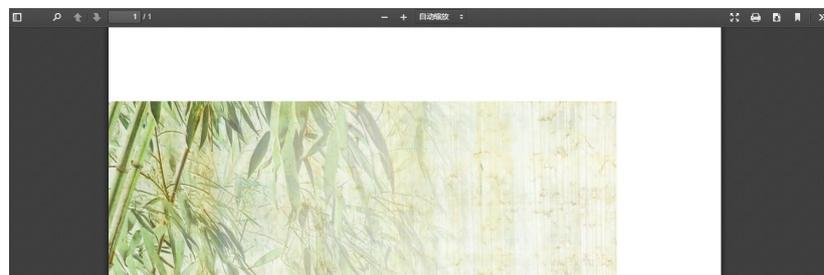
保存

(图 1-3)

登录后**不必填写**用户姓名、地址等信息，只要点 **保存** 键**保存**即可完成注册。

1.3 导入

请点击界面右上角 **工作区** 按钮，进入标引页面（注意这是一个 DEMO，图 1-4）：



(图 1-4)

再请点击页面右上角的 **切换文档** 按钮，出现（图 1-5）

选取文档

查看历史记录

上传文档

选择文件 未选择任何文件

重命名文档

确定 取消

(图 1-5)

点击（图 1-5）中的 **选择文件** 按钮，选择并导入上传用户自己准备好的需要为之编制索引的 PDF 文档（请注意需要标引索引的 **原始文档格式** 须为可复制粘贴的 PDF，如图 1-6）。并点击右下方的 **确定** 键完成上传。

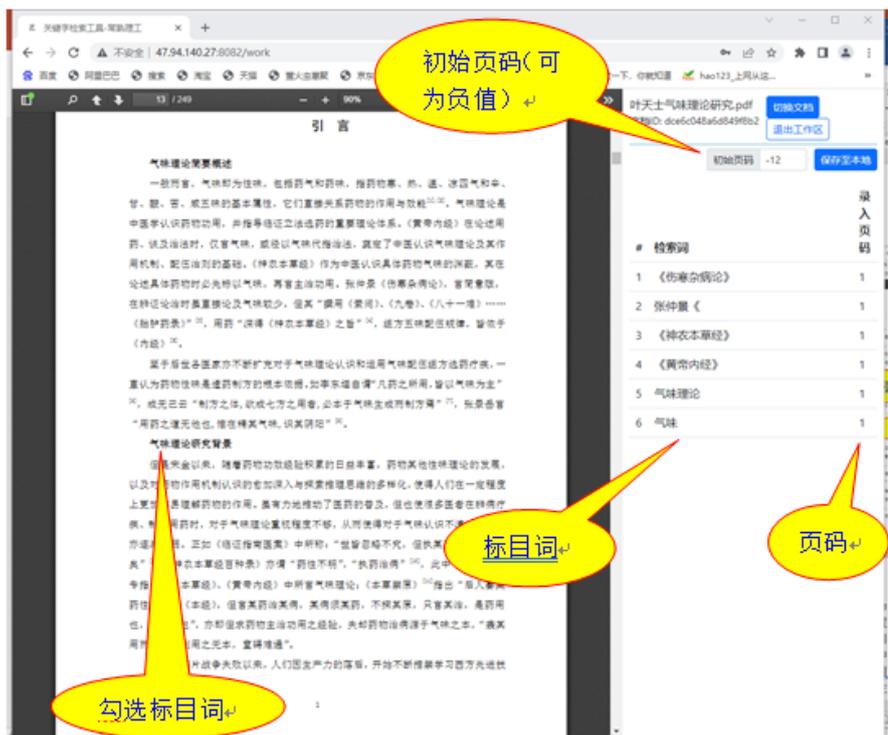
注意：导入论文时，有时时间会比较长，请多等待一会儿，如时间过长，可按 ESC 键。

1.4 标引



（图 1-6）

选择文档确定后即进入标引页面（图 1-7）：



(图 1-7)

标引分为选词和标注页码两部分。图 1-7 中的左侧为选词部分，把光标对准所要选择语词的起始位置，按下鼠标左键，拖选所要标引的语词，所选择的语词就会突出显示。图 1-7 的左侧为标引部分，当鼠标左键松开时，所选择的语词就会按选择的先后顺序显示在左侧的检索词栏目中，并且系统会自动给该语词标注出其所在原文的页码。语词和页码共同组成了一条索引款目的基本要素。如图 1-7 中左侧的**气味理论**被选中时就高亮显示，作为标目进入右侧的检索词栏目之下，并被标上页码为 1 的数字。

1.5 初始页码

标引前要先设定初始页码，初始页码默认为 0。**初始页码**的计算格式为-N（N 为正文页之前的页面数，**不含正文页**）。例如，正文前的页面有 12 页，则初始页码设为-12。例如本例中的初始页码“-12”。



如果原始文档的第一页就是正文的第一页，即原始文档正文页之前的页码为 0，则初始页码默认为 0。

由于系统的问题，有时初始页码设置为负值时，导出文档页码会出现错误。因此，**强烈推荐建议**用户的标引文档起始页从正文的第 1 页开始（如果正文页前有内容，标引之前可先**删除** PDF 文档正文之前的内容）。

1.6 删除

当某一条语词选错或其它原因需要删除时，鼠标右键点击页面左侧的语词，就会出现删除提问（图 1-8），点击确定就会删除该词。也可以在之后导出的 TXT 标引文档中进行修改，操作更简便。

（图 1-8）



1.6 导出

标引完成后，就可以导出标引文档。点击图 1-7 右上角的 **保存至本地** 按钮，选定路径，即可保存（图 1-9）。

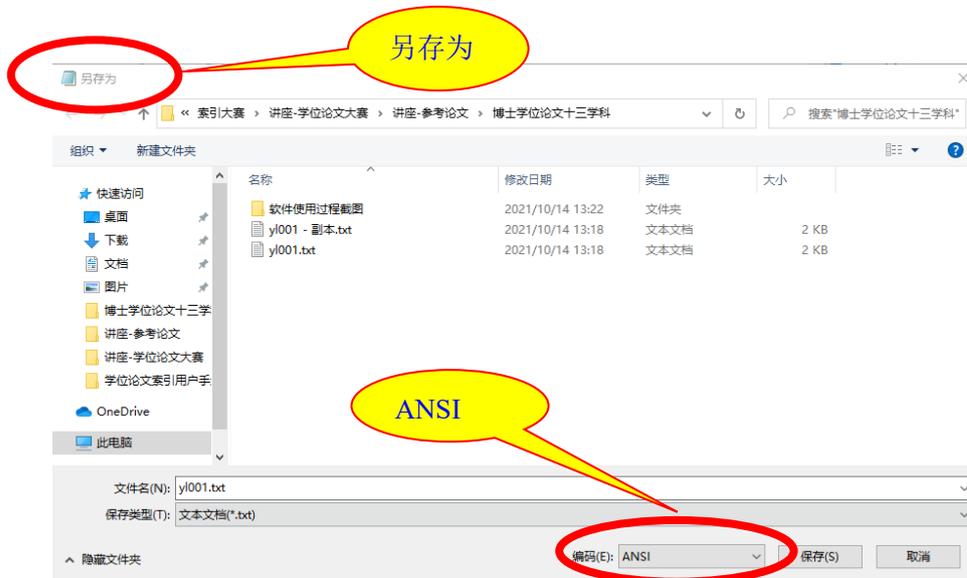
（图 1-9）



打开保存的**标引文档初稿**（部分），初稿为.txt 格式（如图 1-10）：

（图 1-10）

排序准备工作：须把**标引初稿**打开后，将 TXT 标引初稿由 UTF-8 另存为 ANSI 格式，因为后面的排序系统只认编码格式为 ANSI 格式的 TXT 文档（图 1-11）：



(图 1-11)

提示：1.点击文件按钮；2.选择另存为；3.在编码下拉中选择“ANSI”；4.点击保存。

2、拼音排序软件

拼音排序软件包括 2 个软件，用法相近。

2.1 UTF-8 格式拼音排序

网址：<https://26o711c516.zicp.fun>

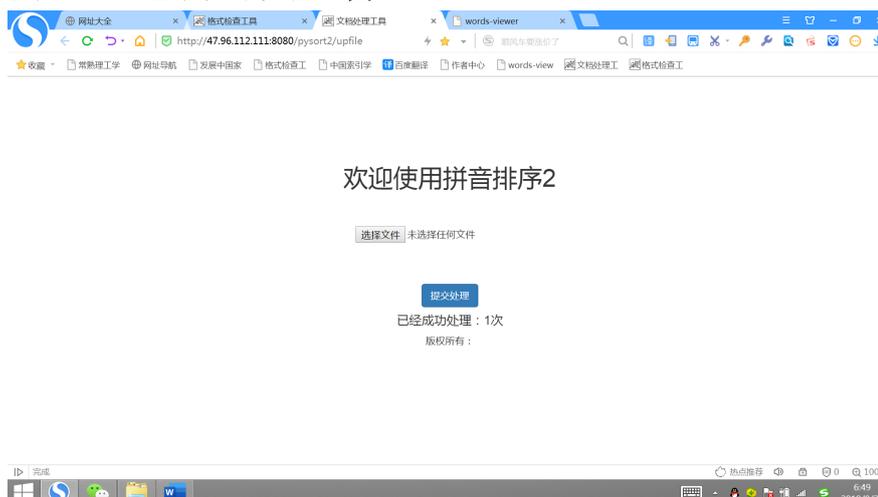
用于排序的 txt 文档须为 UTF-8 格式（从系统保存到本地的标引文档即为 UTF-8 格式，纠错后可直接排序，无需转换编码。）

2.2 ANSI 格式拼音排序

网址：

<http://168.138.152.190:8080/pysort2/upfile>

打开网址，出现如下页面（图 2-1）：



(图 2-1)

点击图 2-1 中的选择文档按钮 选择文件 未，选择导出生成标引初稿 TXT 文档见图 1-10（注意：如前所述，标引初稿 TXT 一定要另存为 ANSI 格式）。

提交处理

点击图 2-1 中的提交处理按钮，生成排序后的索引排序稿（图 2-2）。

B
北京地铁;2
C
财产性人格;9
D
德国;10
电子人格;10
E
二元论观点;8
G
高雅宁;2
《国家人工智能研发战略规划》;3,[10](#)
H
邯郸车主状告特斯拉公司案;2
J
机器人“人格”;6

（图 2-2 索引排序稿）

编辑索引排序稿，如校对、加标题、编制说明、分栏、排版等（图 2-3）。

索 引	
本索引包含主题索引和文献名索引。索引按标目的汉语拼音排序；标目和出处之间以半角分号“;”隔开；同一标目有多个不同页码时，页码按照从小到大的顺序拼接起来，页码之间以半角逗号“,”隔开；标目中的标点符号如引号“”、书名号《》、圆点·等不参与排序。索引由郝建华编制。	
主题索引	主题索引
B 北京地铁;2	人工智能创造物为知识产权的新客体;6
C 财产性人格;9	人工智能犯罪案;1
D 德国;10 电子人格;10	人工智能机器人的民事法律地位;4
E 二元论观点;8	人工智能机器人民事主体资格;6
G 高雅宁;2	人工智能机器人侵权的责任主体;4
H 邯郸车主状告特斯拉公司案;2	人工智能机器人侵权责任;7
J 机器人“人格”;6	人工智能机器人侵权责任的构成要件;4
K 科尔;11	人工智能立法;2
M 美国;10 美国 Uber 自动驾驶汽车撞人事件;2 美国白宫科技政策办公室;10 美国高速公路交通安全管理局;11	人身性人格;9
O 欧盟;10	S 三元论观点;8 手术机器人;1
R	W 为电子法人;9 无辜代理人;11
	Y 亚利桑那州坦佩;2 一元论观点;8 英国;11 约翰·弗兰克·韦弗;10 约翰·金斯顿;11
	Z 智能聊天软件;1 自动驾驶汽车的侵害风险;11 自发侵权;2

（图 2-3 索引稿）

文献名索引

文献名索引

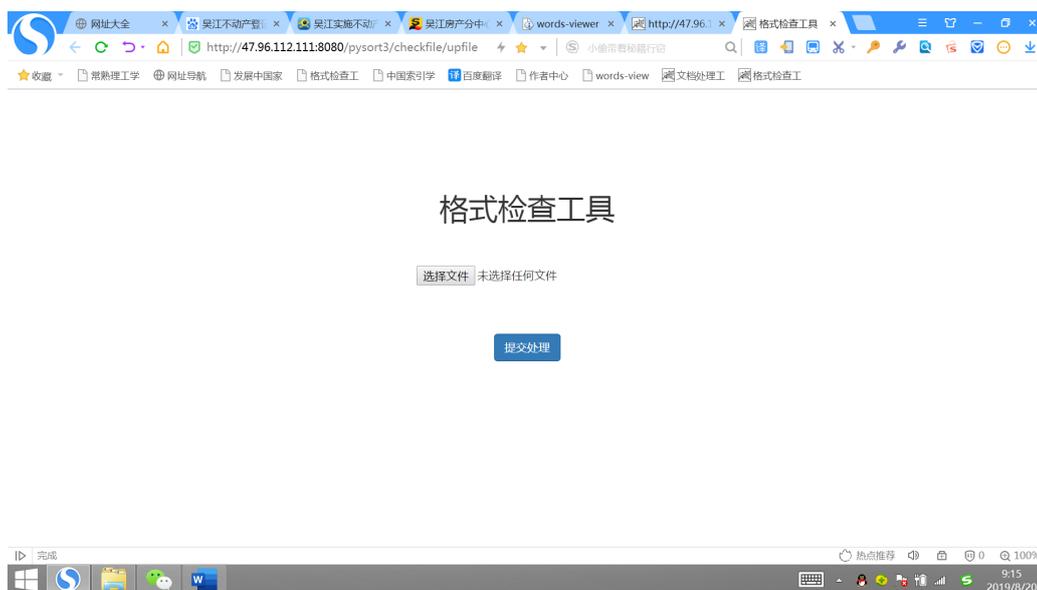
- G**
《国家人工智能研发战略规划》;3,10
- J**
《就机器人民事法律规则向欧盟委员会提出立法建议的报告草案》;2,10
- O**
《欧盟机器人民事法律规则》;2,10
- Q**
《侵权责任法》;7
- R**
《人工智能白皮书》;10
- S**
《上海市智能网联汽车道路测试管理办法》;3
- W**
《为人工智能的未来做好准备》;3,10

3、格式检查软件

格式检查软件的网址是:

<http://168.138.152.190:8080/pysort3/checkfile/upfile>

打开网址,出现如下页面(图3-1):



(图 3-1)

点击图 3-1 中的 **选择文件** 按钮,选中标引初稿。然后点击 **提交处理** 按钮。如果没有格式错误,就会出现如下界面(图 3-2):

检查报告

文件校验通过共找到0错误

(图 3-2)

如果有格式错误，则检查报告中会出现**报错信息** (图 3-3):

检查报告

文件校验通过共找到5错误

第5行存在空行错误

第7行存在空行错误

第9行 (高雅宁2)错误

第12行 (人工智能立法)错误

第23行 (机器人“人格”)错误

你可以复制错误内容在文件中快速查找到错误位置

[点击重新上传文件](#)

(图 3-3)

根据检查报告的错误提示，查看标引文档初稿，用复制、粘贴、查找的方法，可以发现原文中的格式错误，并加以改正。如图 3-4 中的第 5、第 7 行有**空行**、第 9 行的**高雅宁**标目和页码之间没有半角分号；第 12 行的**人工智能立法**之后没有页码；第 23 行的**机器人“人格”**既无分号，也没有页码 (图 3-4)。



(图 3-4)

把错误全部改正之后才可以正常排序。排序完成后还要进行人工校对，有时还有错误。如图 3-5 中的多音字拼音 **Z** 部的**曾隆兴**、**曾世雄**排在了 **C** 部，需要手工调后到相应位置，然后进行排版等编辑工作。

多音字拼音Z部的曾隆兴、曾世雄排在了C部

B
 巴里·尼古拉斯;161
 边沁;161
 卜晶玮;161

C
 蔡自兴;160

Z
 曾隆兴;160
 曾世雄;159
 陈金钊;160
 程啸;160
 茨威格特;162

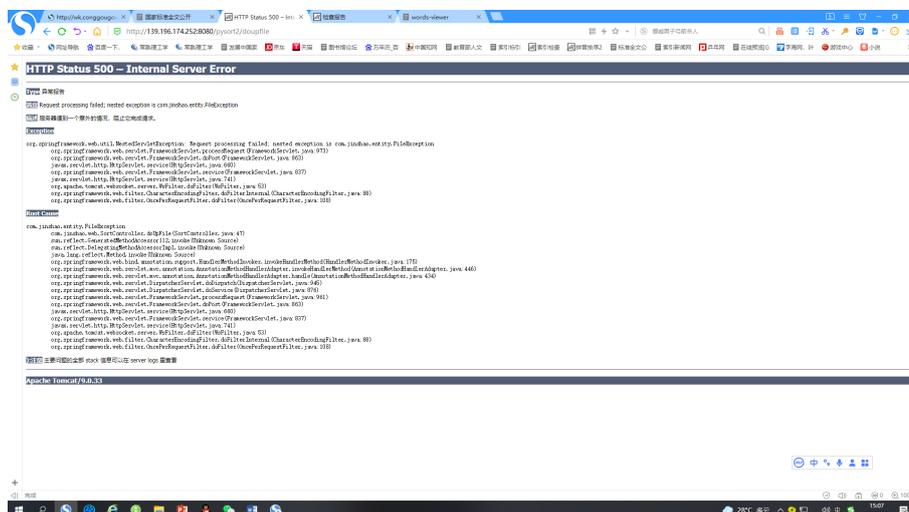
D
 刁荣华;159
 丁凤楚;160

(图 3-5)

4、常见错误汇总

4.1 格式错误

标引，无论是手工标引、软件标引，还是其它方式标引，所获得的 TXT 标引文档中如果有格式错误，排序软件就不能正常排序，而是会出现如下的错误信息。



(图 4-1)

(1) ANSI 编码格式

如前所述，利用标引软件标引导出的 TXT 标引文档默认的是 utf-8 格式，需要另存为 ANSI 格式。

(2) 其它错误

手工标引过程中会出现软件标引中出现的编码格式不是 ANSI、或一行中出现两个及以上分号 (;) 等格式方面的错误，还会出现如前所述的标目和页码之间缺少起分隔作用的半

角分号 (;)、缺少页码、或二者皆缺，以及空格、空行等格式错误，需要加以注意校对和修改。如果有上述错误中的任一项，排序软件都会不能完成对标引文档的排序。

(3) 发现错误的有利工具

上述这些格式错误，如果完全靠人工校对，有时难免不能全部校对出来加以改正。格式检查软件可以帮助用户快速发现这类格式错误，提醒用户去修改错误。需要注意的是，一行多个分号 (;) 这项错误，目前软件还检查不出，而是要靠人工经验作为纠错的第一步手工修改。

4.2 内容错误及解决

索引标引过程中，不论是手工标引还是机器标引，都有可能出现内容错误，例如多标字、少标字、原文中有但不能用于排序的符号（尤其是不能有半角分号;），这类错误需要进行人工校对。

需要注意的是，用于排序的标引文档，每行只能有一个标目、一个分号和一个页码，即已经经过排序的文档不能再次利用排序软件进行排序。

4.3 编制经验

(1) 如果用于标引的学位论文的 PDF 文档选择过程中有选不中字等现象时，可以把 PDF 另存为 WORD，校对并与原版本页码对应准确后，再转换为 PDF 格式进行标引。如果 PDF 是扫描件，文字不够清晰度，则转变成 Word 时可能变成图片格式，则需要将图片中的文字通过文字识别等办法转换为可识别文本。

(2) 标引时原文的页面在标引软件框内不要出现翻页跨页的情形，即不要在一个标引框内出现两个原文页面，那样会引起页码标注错误。

(3) 多音字排序中会出现排列位置错误的现象，需要手工校对。

(4) 用于标引的论文如果较长，可以只保留正文（即把封面、目次、摘要、参考文献、致谢等正文之前、之后的内容裁剪掉），以加快上传速度。

(5) 如果不能一次性完成标引，中途而需要休息或转到其他作业时，建议不必退出系统，返回界面可继续标引，这样不用反复上传文档，以减少上传花费的时间。

(6) 《GB/T 41210-2021 学位论文内容索引编制规则》

<http://c.gb688.cn/bzgk/gb/showGb?type=online&hcno=880E61B530D05F7F63340483FD2283CF>

中国索引学会创新实践基地

2023年5月14日