

董建文. 高校毕业论文格式设置软件的设计与实现[J]. 智能计算机与应用, 2025, 15(2): 70-76. DOI:10.20169/j.issn.2095-2163.24070603

高校毕业论文格式设置软件的设计与实现

董建文

(南京审计大学金审学院, 南京 210023)

摘要: 本研究利用 Python 语言开发了一款高校毕业论文格式自动设置软件,旨在提升论文格式设置的效率,减轻学生在遵循格式规范时的工作负担。经测试检验,软件运行稳定、格式设置准确率高。通过调整函数参数,软件可适应不同高校的毕业论文格式要求,从而具备广泛的适用性和推广潜力。运用本软件与手工调整方法相比,显著缩短了论文格式设置的时间,使师生能将更多的时间和精力投入到论文内容的研究和表述上,从而有效提高高校毕业论文整体质量。因此,本软件具有一定的实践意义和应用价值。

关键词: 格式设置; Word 对象模型; Python 语言

中图分类号: TP391.4

文献标志码: A

文章编号: 2095-2163(2025)02-0070-07

Design and development of software for formatting college graduation theses

DONG Jianwen

(Nanjing Audit University Jinshen College, Nanjing 210023, China)

Abstract: The present study has developed a software for automatically setting the format of university graduation thesis using Python, aimed at enhancing the efficiency of thesis formatting and reducing the workload for students in adhering to formatting standards. Upon testing, the software has proven to be stable and highly accurate in formatting. By adjusting function parameters, the software can adapt to different formatting requirements of graduation theses from various universities, possessing broad applicability and potential for promotion. Utilizing this software in comparison with manual adjustment methods significantly reduces the time required for thesis formatting. This enables both educators and students to allocate more time and effort to the research and articulation of thesis content, thereby effectively enhancing the overall quality of university graduation theses. As a result, the software holds certain practical significance and application value.

Key words: Format Setting; Word Object Model; Python Language

0 引言

随着中国高等教育的蓬勃发展,高校招生规模持续扩大,毕业生人数也呈现稳步增长的趋势。为了进一步提升高校教学质量,2007 年教育部和财政部联合启动了“高等学校本科教学质量与教学改革工程”。作为高校教学中不可或缺的一环,毕业论文(含设计,以下简称毕业论文)既体现了学生专业知识掌握与应用能力,也展示了高校的教学质量水平。

为深化高校教育教学改革,全面提升人才培养质量,教育部于 2019 年发布了《关于深化本科教育教学改革全面提高人才培养质量的意见》^[1],其中特别强调对本科毕业论文的质量监控。高校积极响

应,结合各自的专业特征和教学实际,制定了相应的毕业论文质量控制措施。依据教育部 2021 年起实施的《本科毕业论文(设计)抽检办法(试行)》^[2],本科毕业论文需接受一年一度的抽检评议。抽检结果不仅关乎学生的学业成绩,还将作为评估高校本科教育教学水平、一流本科专业建设、专业预警以及教育资源配置的重要依据,因此高校对毕业论文质量的把控愈发严格。毕业论文的质量涉及内容和格式两个关键维度。在格式方面,尽管高校已经制定了详细的规范标准,但许多学生在遵循这些标准时仍然面临挑战。具体而言,格式要求的复杂性往往导致学生在操作过程中出现疏漏^[3],进而需要进行多次修订,消耗了大量的时间和精力。

基金项目: 2023 年江苏省高校教育信息化研究重点课题(2023JSETKT025)。

作者简介: 董建文(1979—),女,硕士,副教授,主要研究方向:信息管理与信息系统,信息化教学。Email:103023826@qq.com。

收稿日期: 2024-07-06

哈尔滨工业大学主办 ◆ 系统开发与应用

近年来,设计开发毕业论文格式检测系统已成为研究领域的一个热点。例如,研究者提出了流式文档排版效果自动化测试方法^[4]、毕业论文格式检查算法^[5],基于 OOXML 标准创建了论文格式检查平台^[6],设计开发了电子设计文件审查系统^[7]、学位论文格式规范性自动检测系统^[8]以及毕业论文格式检测与校正系统^[9]等。这些成果充分反映了学术界对于提高毕业论文格式准确性和规范性的关注,但格式检测准确率仍在 90.34%至 96%之间,还有待提高。维普和知网两大毕业论文格式检测系统经过不断发展与创新,数据库覆盖范围大幅扩展,检测算法深度优化,检测精度显著提升,通过便捷的在线功能,为众多高校提供了准确全面的毕业论文格式检测服务。然而,在利用这两个系统进行论文格式检测时,必须提交原始 Word 文件,而且文档必须明确设置大纲级别,否则无法对相关部分进行格式检查。虽然格式检测报告能够精确地标注论文的格式问题,但学生仍需下载、参照报告逐一手工修正这些问题。因此,尽管有自动化工具辅助检测,最终的格式调整仍依赖于人工核查与操作。

在当前技术发展的背景下,人工智能获得了广泛的应用和关注,但在 Word 文档的格式编辑方面仍存在明显的局限。尽管人工智能系统可以提供关于文档格式设置的建议及指导信息,但并不具备按照特定格式要求自动编辑和调整 Word 文档格式的能力,不能直接操作文档以实现格式的自动化调整。

因此,实现完全自动化的文档格式编辑和调整仍需进一步的技术突破和创新。相较于按照上述检测报告或指导建议手工调整论文格式的方法,利用自动

化工具对内容完整但未设置格式的论文直接进行精确的格式设置,可以提高操作效率和准确率。基于此目的,本研究开发了高校毕业论文格式设置软件。

1 功能与设计

目前,学生在撰写毕业论文时普遍采用的文档编辑工具包括 Microsoft Word 或 WPS Word。这些工具支持多种文件格式,其中.docx 或.doc 格式因其广泛的兼容性和丰富的功能而被广泛采用。而处理 Word 文档格式的主流技术有 Open XML SDK^[10]和基于 Word 对象模型^[11]两种,本研究选用基于 Word 对象模型构建框架,利用 Open XML SDK 设置个别格式细节。

Python 编程语言以其庞大的第三方库生态系统而闻名^[12],与基于 Visual Basic 的宏语言 VBA^[13]相比,开发效率更高。且 Python 与 Word 软件之间具有良好的对接机制^[14],Python 的 docx 库、win32com 库、os 库、re 库和 tkinter 库能够协同工作,直接对 Word 文档进行格式设置和编辑,极大地提高工作效率和准确性,因此,选择 Python 作为开发语言,运用 docx 库将 Word 文档及其段落、文本、字体等元素视作对象,通过操作这些对象设置毕业论文格式。

本软件基于输入-处理-输出(IPO)模型设计,是一款单机版应用程序。软件接受内容完整但尚未格式化的毕业论文 Word 文档作为输入,通过执行一系列格式设置和转换处理,最终输出一个以学生信息命名的、格式经过修改的文档。软件包含选择毕业论文、设置论文格式、生成格式修改文档 3 个核心功能模块,这些模块及其相互关系如图 1 所示。

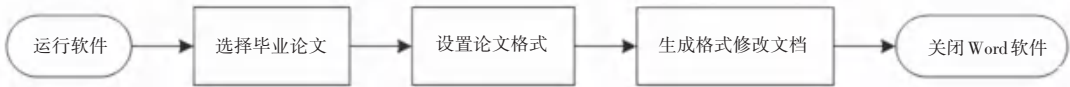


图 1 毕业论文格式设置整体功能流程

Fig. 1 Overall functional flowchart of graduation thesis format setting

本软件设计采用从整体到局部、逐步细化的策略:首先配置论文的总页面布局;然后构建清晰的三节文档框架,并对各部分进行精细的格式处理,涉及中英文摘要、关键词段落的格式设置,各级标题的样式设定,图表的标准化处理,以及页眉页脚和参考文献的格式调整等;最后,将 Python 代码封装并转换为 exe 可执行文件,以确保软件的易用性及其在不同操作系统中的稳定性和兼容性。

1.1 选择毕业论文

软件启动时,使用 `filedialog.askopenfilename()`

函数打开文件选择对话框,为用户提供一个直观界面,以便于浏览和选择文件(参见图 2)。一旦用户选择并打开文件,程序便运用 `split()` 函数对文件路径进行分割,获取文件所在的目录路径及文件名。值得注意的是,docx 库仅支持 Word2007 及以后版本的.docx 文档^[15],然而用户可能会尝试打开旧版 Word 软件所保存的.doc 文件。为了确保文档的兼容性,作者通过条件语句检查文件后缀是否为'.doc',若检查为真,则使用 `SaveAs("{}x".format(path), 16)` 函数将.doc 文件转换为同名的.docx 文

件,并通过 `messagebox.showinfo()` 函数弹出信息框(参见图 3),提示用户文档类型已成功转换,需要重新选择新文档进行格式设置。这一设计不仅简化了文档的打开过程,还有效解决了文件格式兼容问题。

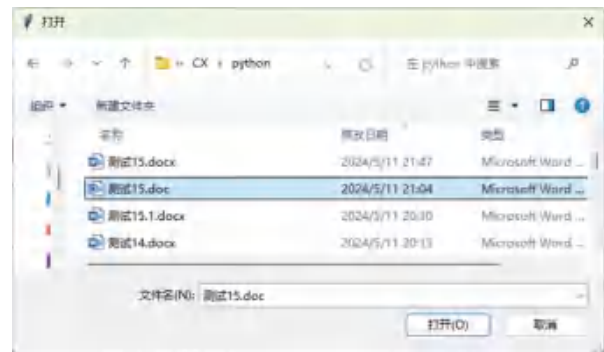


图 2 打开毕业论文
Fig. 2 Open graduation thesis



图 3 提示论文后缀并自动生成同名 docx 文件
Fig. 3 Prompts the paper suffix and automatically generates a docx file with the same name

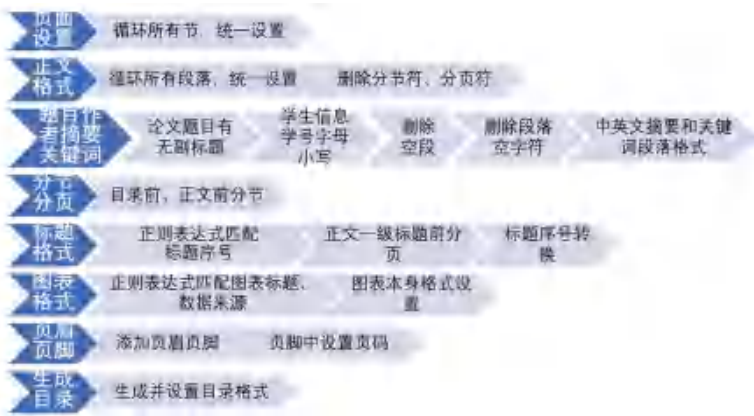


图 4 毕业论文格式设置功能模块结构图
Fig. 4 Structure diagram of graduation thesis format setting function module

1.3 保存文档

在本模块中,论文的作者信息(包括班级、学号和姓名)被用作文件名,通过调用 `save()` 函数,将完成格式设置的 Word 文档保存到指定的文件路径。保存操作结束后,利用 `tkinter` 库中的 `messagebox.showinfo()` 函数弹出提示框(见图 5),提示用户文件

1.2 设置论文格式

设置论文格式是软件的核心功能,依据金审学院毕业论文的格式要求,利用 `docx` 库及 `oxml` 组件逐步设置论文的文档格式。鉴于毕业论文页眉和页脚内容在不同部分有所变化,文档被分为 3 个主要部分:第一节“前文”,涵盖论文标题、作者信息、中英文摘要与关键词;第二节“目录”,展示论文的三级目录结构;从引言/绪论开始至论文末尾,包括致谢、参考文献等,均属于第三节“正文”。

格式设置软件的自动操作包括多个方面:首先,配置页面和段落的格式,确保文档的整体风格一致。接着,从原文档中提取“前文”部分的内容,包括论文标题、作者信息、中英文摘要以及关键词等,并对其应用特定的格式设置。然后,进行分页和分节操作,在首节“前文”之后插入“目录”节,自动生成目录并设定目录格式。最后,将“正文”部分复制到新文档的末尾,构成一个完整的三节结构文档。此外,按照规定的格式标准,对“正文”节中的一至三级标题、致谢、参考文献、图表等分别进行精准的格式设置,以保证文档格式的规范性和统一性。最后,分别为“前文”和“正文”两节设置页眉和页脚的内容及格式,满足论文格式的多样化要求。软件的功能模块结构设计如图 4 所示,该设计及其逻辑旨在确保操作的高效性,并显著提升用户体验。

格式已经设置完成并成功保存。用户单击提示框中的“确定”按钮后,将调用 `Quit()` 函数和 `os.system()` 函数,自动关闭所有已开启的 Word 文档,并终止软件的运行。实时保存功能和用户反馈机制相结合,不仅确保了操作流程的连续性,还实现了错误的最小化,从而提升了整体工作效率。



图 5 论文格式设置完成提示信息

Fig. 5 Prompt message for completing paper format setting

2 难点解决方案

2.1 文档分节

根据格式标准,毕业论文需要被划分为 3 节,即在文档的适当位置插入新节,同时确保内容的连贯性和完整性不受影响。考虑到 docx 库仅限于在文档末尾位置添加新节,如何在文档中间位置插入新节成为技术难点,本研究提出了一种创新解决方法:首先,将论文内容分成 A、B 两个文档,A 文档包含论文的标题、学生信息、中英文摘要和关键词,B文

档则包含从引言/绪论开始直至文档结束的所有内容;随后,在 A 文档末尾添加一节,并在该新节中创建目录项,然后再添加另一新节;最终,将 B 文档的内容复制并整合到 A 文档的第三节中,以此构建出符合格式要求的 3 节结构。具体的操作流程参考图 6,核心代码见表 1。

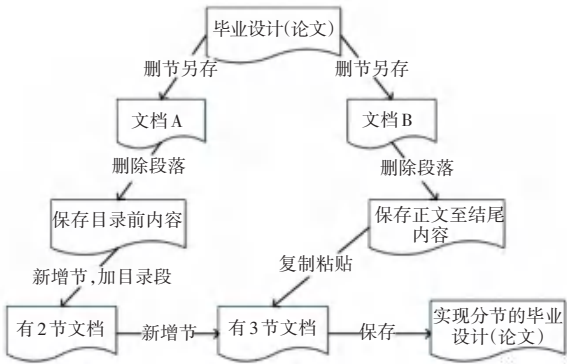


图 6 文档分节流程

Fig. 6 Document segmentation process diagram

表 1 文档分节核心代码

Table 1 Core code for document segmentation

功能	核心代码
打开正文 word 文档,并复制内容	app = win32com. client. Dispatch('Word. Application') doc = app. Documents. Open(dirname+'/' +nfn) doc. Content. Copy()
打开前文 word 文档,光标移至前文 word 文档结尾处	doc1 = Word. Documents. Open(dirname+'/' +filename+'. docx') s = Word. Selection s. MoveRight(1, doc1. Content. End)
粘贴正文内容	s. Paste()

2.2 中英文摘要、关键词段落格式

在毕业论文的中英文摘要与关键词部分,起始标题(如中英文摘要、关键词)与后续文本的格式要求存在差异。由于同一段落内无法直接应用多种字体和样式,传统的格式设置方法无法满足此要求。本软件的解决方案是将中英文摘要、关键词段落分为标题(如‘摘要:’、‘关键词:’、‘Abstract:’、‘Key Words:’)和具体内容两部分。首先,清空原有段落文本,然后在重新加载标题与内容时,分别设置适当的格式,最后,将正确格式化的标题与内容合并,构成完整的段落。图 7 展示了针对中英文摘要及关键词段落格式设置的具体步骤,而表 2 则提供了相应核心代码。

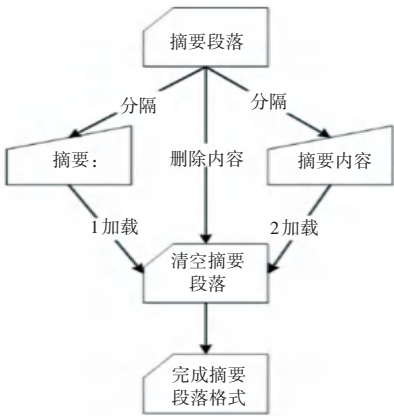


图 7 中英文摘要、关键词格式设置流程

Fig. 7 Process for formatting Chinese and English abstracts and keywords

表 2 中英文摘要、关键词格式设置核心代码
Table 2 Core code for formatting Chinese and English abstracts and keywords

功能	核心代码
创建标题列表	<code>zy_list=['摘要:', '关键词:', 'Abstract:', 'Key Words:']</code>
判断段落中是否包含标题列表内容	<code>if zy_list[n] in pa. text:</code> <code>zy=zy_list[n]</code>
依据 zy 变量将段落分割为标题和内容两部分	<code>rest=r. text. split(zy)</code> <code>r. text=""</code>
加载标题到空段落中,并设置格式	<code>for text in rest[:-1]:</code> <code>run=pa. add_run(text=text)</code> <code>run=pa. add_run(zy)</code> <code>run. font. size=Pt(12)</code>
加载内容到标题之后,并设置格式	<code>run=pa. add_run(rest[-1])</code> <code>run. font. bold=False</code>

2.3 正文标题编号转换

高校毕业论文需遵循国家标准,其中正文部分的一级、二级和三级标题编号应按照“1”、“1. 1”、“1. 1. 1”的形式编排。然而,部分学生习惯使用“一”、“(一)”、“1.”等形式进行标题编号。为减轻学生调整论文标题编号的工作量,软件设计了标题编号转换功能,该功能需确保转换精确,即准确定位二级、三级标题所属的上级标题。软件运用正则表达式识别非标准的“一、(一)1.”标题编号,通过遍历其前后文的标题段落,比较其段落次序与上一级标题的段落次序,将其精确转换为符合标准的格式。

在转换二级标题编号时,软件会查找此标题之前最近的一级标题编号,通过计算两个标题段落在段落列表中的索引号之差来确定距离,并将二级标题编号更新为相应的形式。例如,将“(二)”转换为 $n. 2$,其中 n 代表其所属的一级标题编号。同理,三级标题编号的转换逻辑相似,以三级标题编号“1.”转换为标准形式“ $m. n. 1$ ”为例,其中 $m. n$ 代表其所属的二级标题编号,正文标题编号形式转换的主要步骤见图 8,核心代码见表 3。此功能的实现,不仅提高了论文格式调整的效率,还确保了论文格式的规范性和一致性。

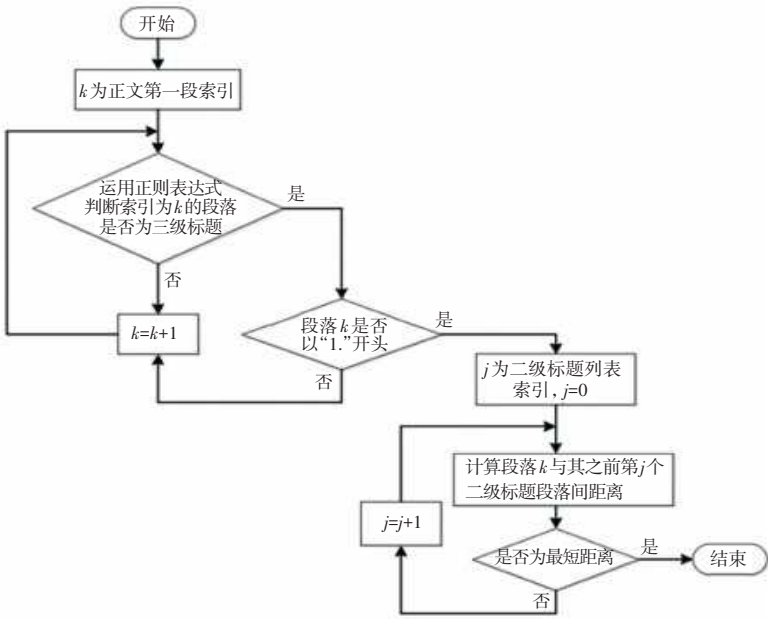


图 8 正文标题编号转换流程
Fig. 8 Main text heading numbering conversion process

表 3 正文标题编号转换核心代码

Table 3 Main text heading numbering conversion core code

功能	核心代码
运用正则表达式构建标题编号模板	<code>pattern = r'\d+(\.\d+){2}'</code>
创建二级标题的段落索引列表	<code>yejbt_list=[ejbtp11,ejbtp12,]</code>
for 循环定位需被替换的三级标题编号所在段落	<code>for k in range(w, 论文最大段落数)):</code> <code>if re.match(pattern,段落文本内容) and not :</code>
计算并赋值最短距离	<code>diff=k-int(ele)</code> <code>nearest=yejbt_list[0]</code> <code>min_diff=k-int(nearest)</code>
定位距离最近的二级标题段落	<code>if diff>0 and diff<min_diff:</code> <code>nearest=ele</code>
三级标题'1.'格式转换	<code>document1.paragraphs[k].text=pc.replace("1.",pcnearest[0:3]+'1')</code>

3 测试与效果

3.1 软件测试

为验证并提升软件性能和适应性,作者在一台配备 2.3 GHz 处理器和 16 GB 内存的计算机上,对 253 篇毕业论文进行了全面测试。测试过程中,记录了软件完成论文格式设置所需的时间,并仔细检查了设置后的论文格式效果,以验证其是否严格遵循了既定的格式标准。

在深入分析测试结果的基础上,对程序进行了相应的修改和完善,从而显著提升了软件的适应性。尤其在文本模板识别与处理方面,软件通过扩展识别范围和优化逻辑代码,能够精准地识别并格式化与标准模板不完全匹配的内容。例如,在为图表下方的“数据来源”标注设置格式时,作者通过将学生论文中多种相近表述纳入识别范围,增强了处理能力,能够准确识别并统一处理“图片来源”“图像来源”“资料来源”“材料来源”等多种表述方式,有效应对了内容多样性带来的挑战。这不仅提高了软件的健壮性,还确保了软件运行的稳定性和格式设置的精确性。

综合测试结果表明,本软件能够满足毕业论文格式设置的要求,确保每篇论文的格式都符合规定的标准。

3.2 应用效果

经过全面测试,本软件展现了良好的性能。所有未设置格式且内容完整的毕业论文均能通过本软件成功转换为既定格式,充分体现了软件的高稳定性;格式设置的完成时间均不超过 10 s,证明了软件的高效性;虽然 docx 库在处理页码时存在固有的局

限性,导致部分目录页码显示异常,但除此之外,软件能够生成符合毕业论文格式要求的文档,确保了较高的格式一致性。

软件界面设计简洁直观,无需用户额外安装程序或插件,有效降低了操作难度,使得师生能够快速掌握和便捷使用。同时,软件对计算机硬件的要求不高,展现出良好的适应性和稳定性。这些特点保证了软件能够高效、准确地辅助师生完成毕业论文的格式设置,提高了毕业论文完成的效率和质量。目前,本软件已获得国家版权局的著作权认证,并被师生应用于学年论文和毕业论文的格式设置中。

4 结束语

本研究设计并实现了一款基于 Python 语言开发的高校毕业论文格式设置软件。软件不仅运行稳定、使用便捷,还具备高度的适应性,通过调整函数参数,可满足不同高校的毕业论文格式要求,为广大师生提供高效、精准的毕业论文格式设置服务,使其能够更专注于论文内容的深度研究和创新表达,进而有效提升毕业论文的整体质量。因此,软件具有一定推广潜力和应用价值。

展望未来,本研究计划将软件的测试和优化扩展至包括 Mac OS 在内的多个平台,以确保其能够满足更广泛的用户需求。目前,软件在参考文献格式设置方面仅提供了基本功能,如字体、字号和段落悬挂缩进等调整,尚未完全遵循国家标准规定的参考文献著录规则。因此,后续工作将集成 Python 的 PyZotero 库和 CSL 引文格式文件,以完善参考文献的格式设置,进一步增强软件的实用性和规范性。

参考文献

[1] 中华人民共和国教育部. 教育部关于深化本科教育教学改革全面提高人才培养质量的意见[EB/OL]. (2019-10-08)[2024-04-27]. http://www.moe.gov.cn/srcsite/A08/s7056/201910/t20191011_402759.html

[2] 中华人民共和国教育部. 教育部关于印发《本科毕业论文(设计)抽检办法(试行)》的通知[EB/OL]. (2021-01-04)[2024-04-27]. http://www.moe.gov.cn/srcsite/A11/s7057/202101/t20210107_509019.html

[3] 刘树林. 高校本科毕业生学位论文格式存在的问题与解决路径[J]. 中国教育技术装备,2023,558(12):121-124.

[4] 左阔,李宁,田英爱,等. 流式文档排版效果自动化测试方法[J]. 计算机工程与应用,2021,57(2):273-278.

[5] 袁敏. 学术论文格式检查和内容校对的研究[D]. 北京: 北京交通大学,2019.

[6] 蒋金敏. 基于 OOXML 标准的论文格式检查平台[D]. 济南: 济南大学,2017.

[7] 郭晓倩. 电子设计文件审查系统的设计与实现[D]. 西安: 西安电子科技大学,2019.

[8] 张巍巍. 学位论文格式规范性自动检测系统设计与实现[D]. 大连: 大连理工大学,2019.

[9] 徐俊. 毕业论文格式检测与校正系统研究与实现[D]. 重庆: 重庆邮电大学,2022.

[10] 纳利军. 基于 Open XML 的毕业论文格式修订系统的设计与实现[J]. 电脑知识与技术,2022,18(9):41-43.

[11] 罗帅,罗斌. Python 辅助 Word+Excel: 让办公更高效[M]. 北京:清华大学出版社,2022.

[12] 嵩天,礼欣,黄天羽. Python 语言程序设计基础[M]. 北京: 高等教育出版社, 2017.

[13] 董建文. 运用 Excel VBA 实现高校新生分班[J]. 智能计算机与应用,2018,8(3):208-211.

[14] 王国平. Python+Office[M]. 北京: 电子工业出版社,2021: 303.

[15] 侯伟波. 软件项目文档格式审查系统的设计与实现[D]. 西安: 西安电子科技大学,2016.