

基于Latex的标准文档在线编辑协作平台的设计与实现

沈宇帆¹, 冯桂林², 沈 炜¹

¹浙江理工大学信息学院, 浙江 杭州

²浙江省自然资源厅信息中心, 浙江 杭州

Email: mail@syficy.com

收稿日期: 2020年9月22日; 录用日期: 2020年10月5日; 发布日期: 2020年10月12日

摘 要

在如今的工业生产上,“标准先行”理念成为共识,国内外标准体系管理手段却仍然非常原始并且影响标准化的发展。标准的制定需要企业、标准化技术委员会、标准管理部门共同协作,且随着生产的发展,标准的迭代也在加速。现有的标准起草软件工具无法满足低门槛、社会性协作、快速迭代、泛编码等需求。本文设计的基于latex排版系统、Vue前端框架、Spring Boot后端框架的标准文档在线编辑平台,结合了latex排版系统和云计算技术,将标准体系文档解构成一个树形结构体,对正文部分进行文本块划分,可进行不同字体编码、表格、公式的编辑。用户只需要通过浏览器对标准文档内容进行在线编辑,系统在云端生成标准格式的pdf文档,实现了标准文档的在线编辑和云协作。

关键词

标准文档, Latex, Vue, Spring Boot

Design and Implementation of a Standard Document Online Editing and Collaboration Platform Based on Latex

Yufan Shen¹, Guilin Feng², Wei Shen¹

¹School of Information, Zhejiang Sci-Tech University, Hangzhou Zhejiang

²The Information Center, Department of Natural Resources of Zhejiang Province, Hangzhou Zhejiang

Email: mail@syficy.com

Received: Sep. 22nd, 2020; accepted: Oct. 5th, 2020; published: Oct. 12th, 2020

Abstract

In today's industrial production, the concept of "standards first" has become a consensus, but domestic and foreign standard system management methods are still very primitive and affect the development of standardization. The formulation of standards requires the collaboration of enterprises, standardization technical committees, and standards management departments. With the development of production, the iteration of standards is also accelerating. Existing standard drafting software tools cannot meet the requirements of low threshold, social collaboration, rapid iteration, and general coding. The standard document online editing platform based on latex typesetting system, Vue front-end framework, and Spring Boot back-end framework is designed in this paper combines latex typesetting system and cloud computing technology to decompose the standard system documents into a tree structure and perform the text part. The text block is divided, and different font codes, tables, and formulas can be edited. Users only need to edit the content of standard documents online through the browser, and the system generates pdf documents in standard formats in the cloud, realizing online editing and cloud collaboration of standard documents.

Keywords

Standard Documents, Latex, Vue, Spring Boot

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近些年来,随着产业多样化的发展,无论是提供产品还是服务的企业,“标准先行”已经成为共识和市场门槛,企业在构建和管理标准体系方面有着迫切需求。一方面,随着各类新兴技术的产生与发展,标准种类的增多,标准管理工作量大,部分文档需要更多复杂的公式、字体等;另一方面,标准的制定需要企业、标准化技术委员会、标准管理部门共同协作,且更新换代的速度很快。

而现有的标准文本编辑、协作与维护模式仍然较为原始。在国内,当前的通用标准编写软件是 TCS 2010,是 MS Word 的插件,但与高版本 Word 软件和 64 位操作系统存在兼容性问题,且无法满足多用户云端协作等需求。在国外,以美国为例,世界上最大的标准化组织 ASTM [1] (美国材料与试验协会)仍在使用较为原始的手工管理的模式,通过邮件来传递各个委员会起草、修订的标准文本,并手工加以整合,以 pdf 文件的形式,再发送给各个委员会进行审议。这样的单机、手工操作方式,已经不能适应现代工作的快捷化、网络化的要求。

在这样的背景下,基于 latex 的标准文档在线编辑平台可以实现低门槛、社会性协作、快速迭代、泛编码等需求,有助于帮助企业之间交流标准文件,并将标准制定工作流程化,规范化,从而加快标准化实践的进程。

本文通过设计与实现基于 latex 的标准文档在线编辑平台,且使用者无需特意学习 latex 的语法,降低了用户的学习门槛。软件系统前端将文档结构解构为树形数据结构,将正文内容分解为单元格,提高了编辑的便捷性。对于公式、表格、图片,系统都设计了独立的编辑器,实现了复杂内容的编辑的简化。

软件后端使用 Spring Boot 框架，完成编译生成基于 latex 的标准文档的工作。

系统实现了对标准文档在线的所见即所得编辑，从而在 Web 端实现了对标准文档的协作，极大地提高了标准文档编辑和协作的效率。

2. 需求分析

本系统在需求分析阶段，主要考虑的是标准文档编辑、协作流程的简单化、编辑功能的多样化。系统总体需求图如图 1 所示，具体需求可以分为：

- 1) 用户管理，包括登入注册，以及个人所属项目管理等。
- 2) 项目管理，包括项目的创建者、创建日期，项目成员所具有的编辑权限等。
- 3) 标准文档编辑，由于标准文档可以被解构为一个树形的数据结构，根据每一部分不同格式和编辑要求，将编辑功能分为封面编辑、正文编辑、附录编辑、参考文献编辑四大板块，实现不同的编辑优化。
- 4) 公式表格图片管理，由于产业的多样化，公式、表格、图片等在标准文档中的应用也在逐渐增多，复杂度也在增大，用户在本系统平台上完成这些内容的编辑与管理。
- 5) 标准文档展示，用户在前端完成输入后，后端即时编译并且生成 pdf 的文档返回前端浏览。

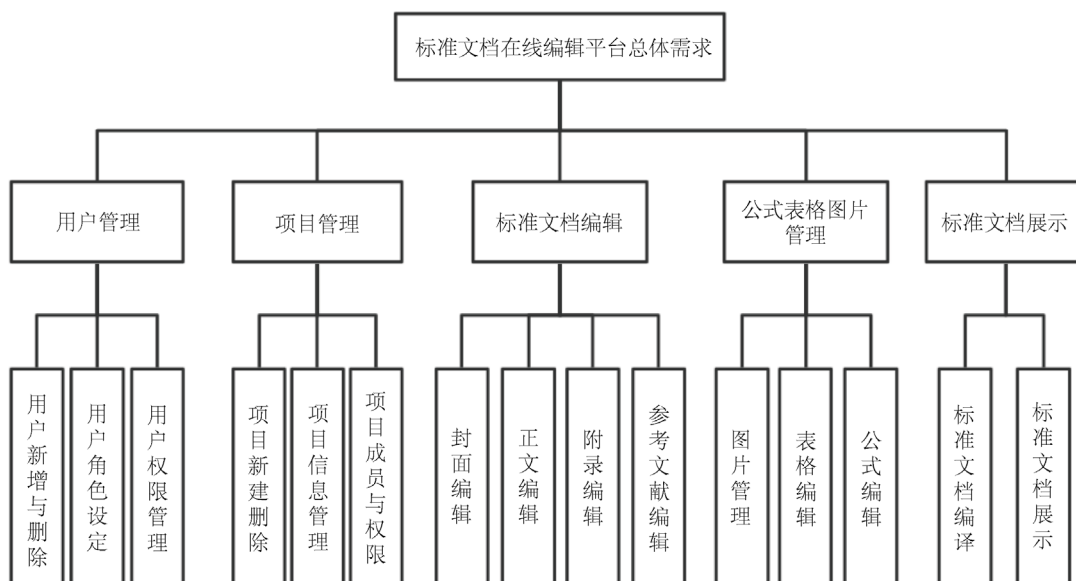


Figure 1. Analysis of overall system requirements

图 1. 系统总体需求分析

3. 系统设计

本系统前端部分采用 Vue 框架，Vue 框架采用的是 MVVM (Model-View-ViewModel) 的设计模式，MVVM 采用了双向的数据绑定(data-binding)，使得 View 层与 Model 层的同步工作是完全自动化的。因此开发者只需关注业务的逻辑，而不需要手动去操作 DOM，同时 Vue 还拥有组件化开发等特性，提高开发效率。

本系统的前端可以分为两个部分：用户、项目管理部分和文档编辑、展示部分。用户、项目管理部分需要对项目成员进行权限划分，以实现多用户协作功能。而文档编辑、展示部分要对标准文档的在线编辑提供特殊优化。传统的 web 端在线编辑往往采用富文本编辑的方法，由于本系统存在云协作、多人编辑等功能模块，传统的富文本在线编辑的设计模式既无法满足精准排版、功能多样化的需求，也难以

降低用户的操作门槛。

Latex 是由美国计算机科学家莱斯利·兰波特(Leslie Lamport)开发的格式化排版系统[2]，应用 latex 排版系统能够排出复杂的图标和精美的数学公式，到目前为止，国内外公认的文本排版系统仍是 Latex 排版系统，其能够在较高层次上实现文章排版的美感[3]。本系统选择 latex 作为底层排版工具，即有利于标准文档的精准排版，也对公式表格等元素的编辑提供底层技术支持。

前端部分设计上，本系统创新地使用将整体标准文档划分为树型结构，根据各部分不同的编辑特点，划分为：封面编辑、正文编辑、附录编辑、参考文献编辑四个部分，分别进行优化编辑，提高效率。

对于正文部分，本系统采用单元格划分的方式，由使用者将正文进行划分，便于用户编辑和社会性协作。每一个正文单元格从内容种类上可以划分为普通文段、图片、表格、公式四大类型。同时，本系统对正文部分在 web 端的编辑提供多种快捷键、可视化优化，提高编辑和协作的效率。

本系统后端采用 Spring Boot 框架，该框架可以快速构建企业级的应用程序，简化开发的繁琐配置，从而极大的提高了开发的效率，简化了部署流程。考虑到后期应用量大，本系统在服务器端基于 Docker 容器技术部署，Docker 容器技术是一种虚拟化技术[4]，具有很好的性能和安全性，该技术既简化服务器更换流程，也为分布式部署提供支持。

系统整体架构设计如图 2 所示：

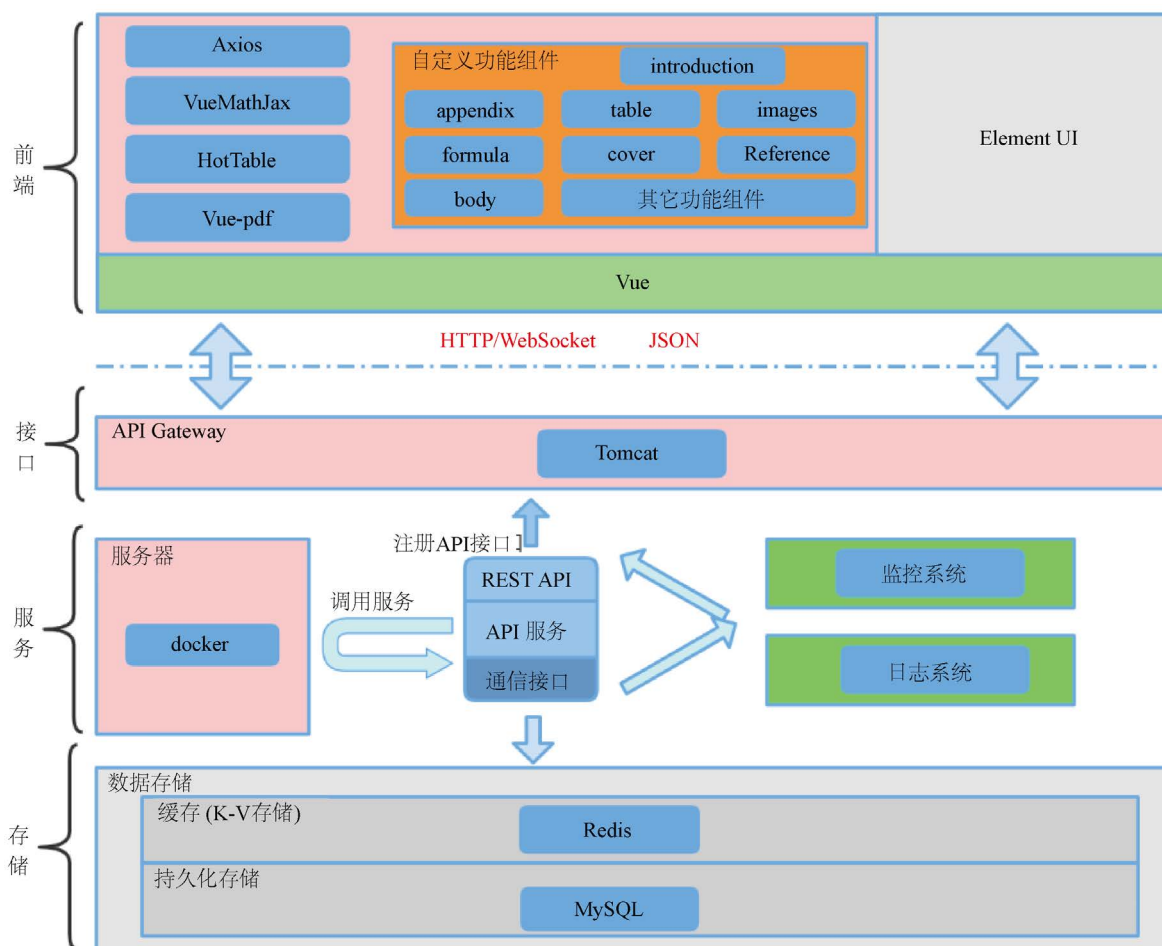


Figure 2. System architecture diagram

图 2. 系统架构图

4. 系统实现

4.1. 前端实现

1) 项目管理与协作部分, 系统通过分配项目成员不同权限的形式实现云端协作。项目创建者为项目管理员, 享有最高权限, 可邀请其它用户加入项目。项目编辑时, 其它用户(除管理员外)登入系统后处于锁定状态, 无法进入编辑。管理员可以停止正在编辑的用户的编辑行为和撤销用户的编辑。通过 diff 算法, 项目成员可以清楚的看到每个成员的编辑变动, 从而提高协作效率。

2) 封面部分, 系统抽取了关键的封面元素: ICS (International Classification for Standards, 国际标准分类法)号, CCS (Chinese Classification for Standards, 中国标准文献分类法)号, 标准名, 标准类型(国家标准、地方标准、行业标准、企业标准), 标准号, 代替标准号, 标准标题, 标准标题英文译名, 发布时间, 实施时间, 发布单位。对这些关键元素进行优化编辑, 通过联级选择器等方式, 提高使用者的编辑效率, 减少编辑错误的概率。同时, 为了减少系统的开销, 提高体验感受。

3) 公式编辑部分, 系统使用 VueMathJax 插件, 通过前端 latex 渲染成 HTML (超文本标记语言)的方式实现, 使用针对不同的数学符号的不同按钮实现进行公式的编辑, 由于存在大量的不同数学符号, 通过编译渲染的方式会造成页面卡顿, 本系统直接显示不同数学符号的图标, 减轻前端显示的负担。最后通过 HTTP 请求直接将公式的 latex 文本存入后端数据库。

4) 表格编辑部分, 系统使用 HotTable 组件。对于表格的编辑功能, 本系统实现了上方、下方插入行, 左方、右方插入列, 移除行、列, 多种对齐方式, 合并单元格, 复制, 剪贴等功能。由于 HotTable 的响应式数据是基于 JavaScript 对象存储在浏览器中的, 因此本系统在前端使用 JavaScript 将 HotTable 中对象格式的数据转化为 latex 语法文本, 以便后端文档编译。

5) 图片管理部分, 系统使用项目图片单独管理的方式, 将上传的每个图片设置一个无重复的编号, 在正文部分中通过引用图片编号的方式插入图片。

6) 正文编辑部分, 由于传统的富文本编辑的模式不能满足需求, 系统采用了划分单元格的方式进行正文部分的编辑, 每个单元格具有“层级”和“样式”两个属性, 如图 3 所示。标题内容为独立的单元格, 通过“层级”属性区别不同的标题层级; 非标题内容可以由用户自主决定划分为多个单元格, 这些单元格的“层级”标签属性固定为“p”, 通过“样式”标签可以引入图片、公式、表格等非文本内容, 当“样式”标签修改为非文本内容时, 单元格内部的输入框将转变为选择器, 通过选择非文本内容的 Id 选择对应的内容。另外, 为了提高正文部分的编辑效率, 本系统设计了一套快捷键系统, 通过浏览器的 api 监听不同的键盘输入值, 可以让使用者在脱离鼠标的情况下只使用键盘完成大部分的文本编辑操作。

7) 与后端交互部分, 本系统采用基于 promise 的 axios 库, promise 是 JavaScript 在 ES6 (ECMAScript 6)中新增的功能, 通过回调以及其它技术, 避免了前端网络请求中回调地狱(Callback Hell)等问题, 可以通过链式调用连续执行多个异步操作, 大大简化了异步编程。由于本系统编辑部分内容多, 数据分散, 因此使用异步编程是十分有必要的。Axios 库支持 node 端和浏览器端, 使用 promise 管理异步, 并且支持拦截器等高级配置。

4.2. 后端实现

本系统后端采用 Spring Boot 框架, Spring Boot 将常用的依赖分组整合到一个依赖中, 可以自动配置, 简化 Spring 应用的配置。通过 jdbc (Java Database Connectivity)访问数据库, 通过系统命令的方式调用 latex 的编译模块, 对数据库中的有关项目 latex 文件进行编译, 输出 pdf 文件返回给前端。

数据库采用 mysql, 系统的数据库设计主要分为用户和项目两部分, 具体系统数据库设计 ER 图如图 4 所示。



Figure 3. Realization diagram of text editing part

图 3. 正文编辑部分实现图

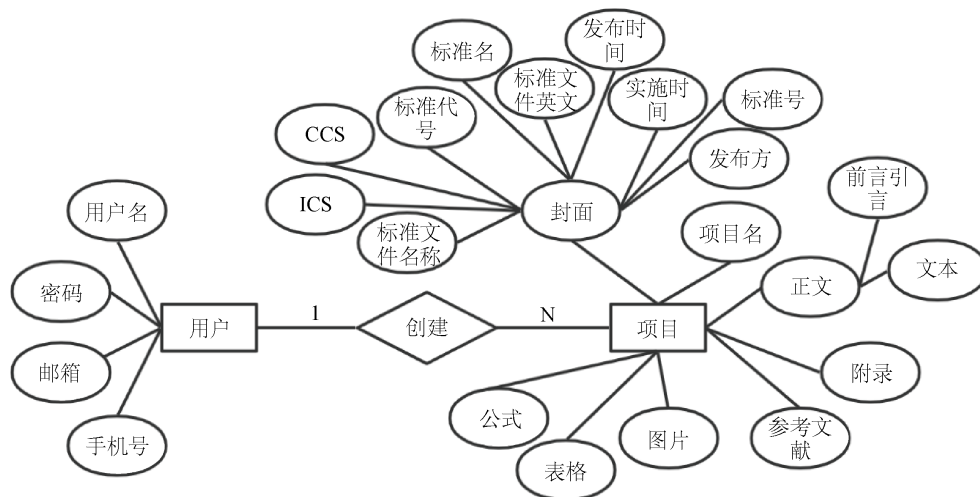


Figure 4. System database design ER diagram

图 4. 系统数据库设计 ER 图

编译模块使用 XeTeX，是一种使用 Unicode 的 Tex 排版引擎，支持现代化字体技术，由于默认其输入文件为 UTF-8 编码，XeTeX 可以直接使用自身操作系统中的字体，而不需要额外的配置。

每年国家、各个省和行业协会都会颁布、修订和更新标准，用户基于这些标准而建立的标准体系，必须进行相应的修改，以体现标准的更新。这对企业是一个既有专业性，又耗时费力的工作。在时间的维度上，本系统可以实时更新每个标准文件的引用状态，获取最新标准以确保本平台上的每个标准文件引用的都是最新的国家标准或者地方标准，保证标准体系与标准的更新相适应。通过字符串比对算法、LSTM 神经网络的实践，让各方对标准文本的意见，都可以通过平台进行比对和展示，使标准协作真正实现现代化。

考虑到标准文档后期数据量的增加可能降低编译速度，本系统根据每一篇标准文档分为封面、前言

引言、正文、附录、参考文献、图片管理、公式编辑、表格编辑八个部分的特点，后端系统进行多线程独立编译，最后整合为一的方式，提高了编译速度。当用户进行数据更改时，后端系统会识别用户更改的模块，只对该模块进行独立的重新编译，从而优化编译效率。

5. 性能评估

本系统前端部分，通过采用浏览器缓存机制储存、HTTP 异步请求等优化工作，提升了前端性能。本系统后端部分，采用了多线程编译、模块独立编译等优化工作，编译速度得到了很大的提升。但前端在调用后端编译接口时，调用时间仍然较长，影响用户体验，因此本系统尝试寻找其它提高编译速度的方式。

对于一篇 latex 文档，编译程序需要多次遍历才能完成编译工作，因此影响 latex 编译的主要是文件读写速度。EXT4 是第四代扩展文件系统，是 Linux 系统下的日志文件系统[5]。本系统实验通过在同一款电脑上使用相同的编译程序 XeTeX，编译相同内容的 latex 文档(140 行，共 3603 字符)，从而对比不同文件系统下 latex 文档的编译速度。实现记录如表 1 所示。实验结果表明，在 Linux 操作系统/EXT4 文件系统下的编译速度达到了 windows 操作系统/NTFS 文件系统下的速度的四倍，提高了本系统的编译性能和用户体验。

Table 1. Compilation speed comparison table

表 1. 编译速度对比表

文件系统	NTFS 文件系统	EXT4 文件系统
耗时	0.997 s	0.228 s

6. 结束语

本文基于企业、标准化技术委员会、标准管理部门对标准化文档编辑的便利性、协作性的需求，运用云计算技术、前后端分离技术，打造了一款基于 latex 的标准文档在线编辑平台。该平台能够在多浏览器端进行标准文档的可视化编辑与协作，文档解构为树形结构与正文部分块化编辑等创新点为标准文档的编辑提供了极大的便利性。并且本系统基于 latex 排版，不仅可以满足公式、表格、图片等多种复杂的文档输入需求，而且能够使得最终编译后的文档精美且精准。通过实践，本平台能够满足企业在标准化实践上的迫切需求，解决目前在标准化协作方面的痛点，加快企业标准化工作的进程，同时，该技术又能够改变国内外在此方向上的软件较为原始和落后的局面，具有非常好的实用价值，对相关的其它类型文档编辑与协作提供了参考意义。

致 谢

感谢浙江理工大学信息学院计算技术研究所对本项目所需要的软件、硬件提供支持，感谢浙江理工大学教务处对大学生创新创业训练计划客观公正评审及经费支持。

基金项目

国家级大学生创新创业训练计划项目(201910338015)。

参考文献

- [1] 郑鹰. 美国材料与试验协会(ASTM)的标准制定工作[J]. 世界标准化与质量管理, 2001(9): 33-34.
- [2] Lamport, L. (1994) LATEX: A Document Preparation System: User's Guide and Reference Manual. Addison-Wesley,

Boston.

- [3] 王春燕. 应用 LaTeX 系统排版自然科学类期刊的优势分析[J]. 出版科学, 2007, 15(3): 77-79.
- [4] Merkel, D. (2014) Docker: Lightweight Linux Containers for Consistent Development and Deployment. *Linux Journal*, No. 239, 76-90.
- [5] Mathur, A., Cao, M., Bhattacharya, S., *et al.* (2007) The New Ext4 Filesystem: Current Status and Future Plans. *Proceedings of the Linux symposium*, Vol. 2, 21-33.