

标准文献内容指标服务系统^①

计雄飞¹, 张宝林¹, 王玉琢²

¹(中国标准化研究院, 北京 100191)

²(北京理工大学, 北京 100081)

摘要: 标准在一定范围内具有法律法规效应, 它是国民生产、贸易往来的基础规范. 标准文献的收集整理与挖掘利用是标准咨询与服务的必要条件. 本文结合标准文献资源的现状与实际用户需求, 设计了从标准文献影印件到标准内容指标的加工流程, 实现了标准文献内容指标服务系统.

关键词: 标准; 标准文献; 标准内容; 技术指标; 系统设计

Standard Document Content Indicator Service System

Ji Xiong-Fei¹, ZHANG Bao-Lin¹, WANG Yu-Zhuo²

¹(China National Institute of Standardization, Beijing 100191, China)

²(Beijing Institute of Technology, Beijing 100081, China)

Abstract: Standards have legal effect within a certain range, and it is basic norm of national products and trade contacts. The collection, mining and using of standard documents are necessary conditions of standard information consulting and service. Combining situation of standard document resources and actual user needs, this paper designed a processing flow from standard document photocopy to standard content indicators, and implemented a standard document content indicator service system.

Key words: standard; standard document; standard content; technical indicator; system design

1 引言

标准文献简称标准, 是指由公认机构批准发布, 能够共同使用和重复使用的规范性文件^[1], 是国民生产的各个环节的基础规范. 按照标准化对象的基本属性, 标准分为技术标准和管理标准两大类^[2]. 随着我国生产力的急速提高, 各行各业的快速发展以及进出口贸易的蓬勃发展, 国内外标准的相关技术规范和政策法规越来越受到重视. 在过去的三十年中, 我国标准行业迅速发展, 标准化队伍逐步壮大. 标准服务的模式也正经历着不断创新的过程: 从标准产品来讲, 由标准影印版(纸件)到数字化标准信息在网络推送; 从标准品种来讲, 由主营国内标准到主营 60 多个国家和地区的相关标准^[3]. 与此同时, 各行业的迅猛发展也加快了标准的更新速度, 这对标准化信息咨询与服

务提出了更高的要求.

标准文献信息产品主要服务于政府决策机构、相关情报机构以及企业, 前两者的应用服务主要由高校、科研院所纵向基金项目定向承担, 而后者(企业)是标准服务机构最大的需求用户. 2005 年在国家科技基础条件平台中心的支持下推出标准文献共享服务平台是当前国内最权威的标准文献对外服务窗口^[4], 在很大程度上满足了企业用户对标准文献的基本需求. 但专业化、个性化的标准信息服务一直以来都是企业的宿求, 如近年来由国家标准馆推出的标准推送、领域专题集成和情报分析等服务模式^[5]. 本文结合企业用户需求与国家标准馆馆藏资源, 整合了标准文献内容指标加工流程及其管理规范, 建立了较为完善的领域性标准文献内容技术指标库, 实现了标准文献题录、

^① 收稿时间:2014-06-04;收到修改稿时间:2014-06-24

内容和指标的综合检索与内容对比服务系统.

2 分析与设计

标准文献与其他文献资源相比具有一定的特殊性,就目前而言国家标准馆的馆藏标准原件主要以标准影印件为主,这就造成加工标准文献数字化产品具有一定的难度.因此,标准文献内容的加工是标准信息资源深度挖掘与利用的前提条件,也是提供专业化标准文献信息服务的基础工作.如图1所示,标准文献内容加工从数字化标准影印件开始,首先,加工信息标准文献题录数据库,进行一对一的资源关联说明;其次,依据领域性或专题性定向服务需求,从标准资源数据库中抽取相关待加工的标准,建立目标标准资源库(主要指标准文献影印件数据集);第三,对目标标准资源库中的全部标准影印件进行OCR全文识别,组建成标准全文数据库,从这部分内容开始支持了真正意义上计算机全文检索;第四,依据标准文献内容结构拆分标准全文数据库中的全部文档,组建新的标准结构化内容数据库,该部分支持标准文献结构化对象的检索;第五,对标准结构数据库进一步加工,提取关键技术指标,组建标准内容指标数据库;最后将标准题录、标准影印件、标准全文数据库,标准结构化内容数据库和标准内容指标数据库集成在一起,建立相互的关联关系,提供题录、内容、指标三个角度的检索与对比服务.

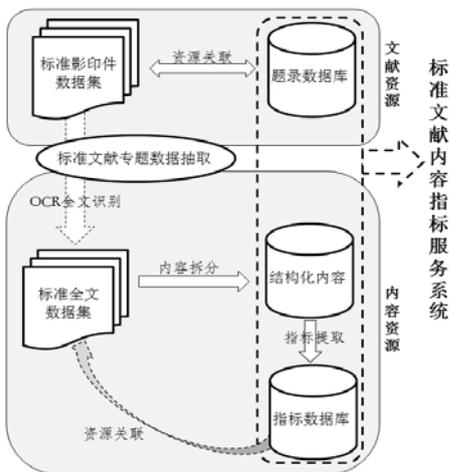


图1 标准文献加工流程

标准文献内容指标服务系统以标准题录数据库、

标准结构化内容数据库和标准内容指标数据库为直接信息资源,同时以标准影印件(原版)和标准全文数据库为基础资源,如图1所示.五种资源以题录资源为中心,系统以题录检索为入口可以之间关联到其他四种资源.同时其他四种资源也可由题录资源进行间接关联,即五种资源以题录资源为媒介相关管理.如图2所示,标准题录资源主要包含标准号、标准名称、国别、发布机构和实施日期等属性;全文影印件记录着文本类型和存储资源,其资源以文件形式存储;数字化全文与全文影印件相似,以文件形式存储;结构化内容包括标准化对象、结构化ID、文本内容和引用等信息;指标数据库包括指标名称、测试方法、技术指标和适用范围等信息.

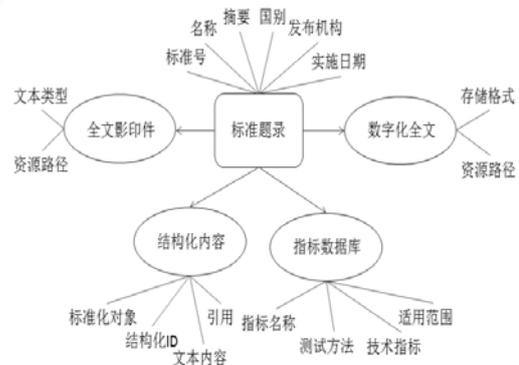


图2 资源关联结构

基于上述分析本系统设计了检索、比对和专题三个核心模块以及一个统计分析辅助模块.如图3所示,检索模块主要包括全文检索、内容检索和指标检索.比对模块包括内容比对和指标比对.在资源检索与对比的基础上设计了统计分析和专题模块,统计分析模块旨在为用户提供当前系统资源的实际情况,如资源种类、数量等信息;其包括题录关联统计(题录资源之间的相互联系)、内容关联统计、指标关联统计和专题统计分析等.专题模块提供了开放式接口,可以直接挂

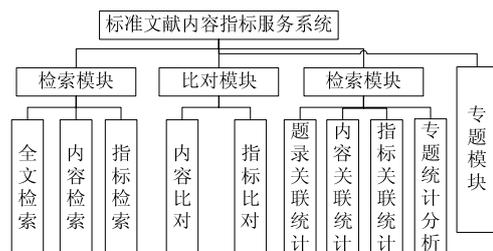


图3 系统功能结构

接专题性标准内容指标, 提供更权威的信息集成服务.

3 系统实现方案

系统技术实现主要从系统资源, 应用环境, 内容管理、数据安全和研发成本等方面来考虑以何种技术方式实现. 标准文献内容指标服务系统其资源包括五种, 主要关系数据库(题录数据、结构化内容和指标数据库)和数据文件存储(标准影印件和标准全文数据集). 就目前业务开展而言, 从领域性专题标准文献开展内容指标的加工与服务, 其数据量属中小型, 因此采用 Microsoft SQL Server 2005. 但在数据库方法实现过程中内置三种数据访问方法: SQL Server, MySQL 和 Oracle, 以配置文件的形式配置数据库连接方式, 支持后续系统升级以及数据迁移. 从资源更新、内容管理和数据安全角度考虑, 系统为每一个资源项设置了唯一资源代码, 如图 4 所示. 资源代码包括资源类型、预留空位、入库时间、数据版本和资源顺序号五个部分. 其中资源类型明确了资源的种类, 预留空位可以为后续系统资源扩建提供了空间. 入库时间受限保证了资源顺序号在一定程度上的有效性, 另外它是系统资源故障跟踪的有效方法. 数据版本表示同一批次入库的所有资源的版本号, 方便数据维护与更新.

顺序号由系统依据当前数据资源情况和当天资源数量自动生成唯一编码, 一般以十进制阿拉伯数字为编码基础, 并允许扩展.



图 4 资源代码

就当前而言, 建立标准文献内容指标服务是一种基于国家标准馆藏数据资源的一种探索性创新服务模式. 在应用环境和研发成本上适当降低要求, 以经济实用的方式快速搭建系统平台, 进一步验证完善研发方案. 因此本系统确定使用 .NET 编程技术, 采用 B/S 结构. 一方面微软技术平台的框架下增强用户体验, 缩短系统研发周期, 另一方面以 B/S 结构建立 WEB 系统, 方便数据更新以及内外网应用. 从数据安全角度考虑, 三台服务器(操作系统 Windows 2003)用于构建系统应用平台, 第一台用于系统存储数据资源, 包括 Microsoft SQL Server 2005 和 FTP 文件资源系统. 第二台用于系统 WEB 应用, 安装 .NET Framework 4.0 和 IIS 6.0. 第三台用系统资源和 WEB 系统应用定期备

份.

在完成系统研发与应用的整体规划之后, 本部分围绕标准文献内容指标检索、对比和专题集成等主要系统功能展开设计与实现方案的介绍. 从系统架构实现角度而言, 系统体系结构自下而上分为数据库层、数据访问层、业务逻辑层和展现层. 数据库层是指数据资源的存储、管理、访问和权限控制, 由系统管理员配置 SQL Server 2005 和 FTP 服务器完成. 数据访问层采用 C#语言开发, 建立了通用的数据方法规则, 并支持三种类型的数据(SQL Server、MySQL 和 Oracle), 增强了系统的可移植性. 业务逻辑层根据实际业务逻辑, 实现具体的业务流程, 为展现层提供算法调用. 展现层是系统应用的窗口, 在 .NET Framework 4.0 框架下加强了客户端用户体验, 主要由 html 和 aspx 文件以及客户端应用脚本 js 构成. 四个系统控制层次之间, 单方向访问和服务, 即上层访问下层, 下层服务上层. 依据这样的系统实现策略, 逐一实现各系统模块.

(1) 检索模块

系统检索模块主要分三个层次检索全文、内容、指标, 但大部分用户最直接面对是题录检索, 这一点与科技文献检索相似. 题录检索就是检索标准的名称、标准号(标准的唯一标识符), 国别、摘要、关键词、发布机构、实施时间和替代关系等属性, 在本系统设计中采用数据库全字段模糊匹配检索, 即同一标准输入任意属性值均可检索到结果. 这一设计简单快捷, 但在数据容量较大时, 会产生检索延时情况. 目前本系统对数据已经做了领域性抽取, 缩小了检索范围, 因此这一影响相对较小. 以题录检索为例, 其检索结果以部分摘要和标准号信息为基本结果, 同时在系统实现中支持用户对检索结果显示方式的个性化设置. 如图 5 所示:



图 5 题录检索

检索“变压器”部分结果被展示. 点击蓝色摘要内容, 直接浏览该标准详细内容信息. 单击“标准号”查看该标准的指标信息. 单击发布机构名称, 如“沈阳变压器研究所”查看该单位拟定的相关标准.

(2) 对比模块

系统对比模块旨在为标准分析人员提供有力的技术支持. 本模块建立两种对比模式, 一种为标准结构化内容对比, 另一种为标准内容指标对比. 如图6所示:

两项标准 BBT 0035-2006 和 GB/T 10228-1997 结构化内容对比展示. 这一系统功能可以根据用户需求任意选定比较范围. 标准之间可以进行比较分析主要源于标准具有一定的法律法规效应, 其结构相对固定. 可归纳为前言, 适用范围, 规范性引用文件, 术语和定义, 符号和缩略符, 分类和命名, 要求, 试验和测试, 标志、包装、运输和贮存, 附录以及修改单和技术勘误. 在此基础上进行对比分析可以直观地分析到行业标准、国家标准、地方标准和其他标准之间的差异. 为

当前执行标准的分析对比和拟定新标准提供必要的支撑. 由于系统基础资源数据是由 OCR 全文识别而得到的数据, 虽然经过人工校正, 但仍与原文有出入, 因此系统支持用户即时浏览原版信息. 如图6中, 单击下划线“内容”可直接浏览原版信息, 系统自动定位到当前内容.

指标对比模块是系统的核心功能之一, 如图7所示, 本模块提供任意多个相关标准的同类指标对比. 直接展示标准指标的测试指标和测试方法. 这在过去的标准信息指标对比中, 需要翻阅大量的标准技术文件, 耗时巨大. 本系统提供的指标对比功能, 建立在领域性权威标准信息基础之上, 为企业、政府、研究院所标准化从业人员提供了权威的技术方案和信息支撑. 单击“标准号”可查阅该标准的全文技术指标, 单击“指标名称”可查阅当前数据库中相同指标名称和同类型全部技术指标, 单机“测试指标”或“测试方法”可浏览原版标准.

标准文献内容指标服务系统			
标准信息 专题指标 统计分析 站点介绍 咨询服务 关于我们			
标准内容对比			
序号	标准化结构	标准一: BBT 0035-2006	标准二: GB/T 10228-1997
1	前言		本标准是修改采用 GB 10228-88《三相空气自冷干式电力变压器技术条件》和 ZB K41 003-88《三相树脂绝缘干式电力变压器技术条件》两个干式电力变压器技术条件标准合并编写而成的. 本标准保留了原来两个标准中仍然适用的内容, 如技术要求、试验方法、检验规则、标志、包装、运输和贮存几章的部分条文. 同时, 为适应变压器技术的发展, 对 10 kV 级无励磁调压配电变压器性能参数进行了修订; 将非包封线圈系列的容量范围扩展到 2500 kVA; 增加了六个系列产品参数表 (表 2、表 3、表 5、表 6、表 7、表 8); 在参数表中列出了绕组绝缘材料耐热等级为 B、F、H 级时的负载损耗数值. 本标准从实施之日起, 同时代替 GB 10228-88 和 ZB K41 003-88. 本标准由中华人民共和国机械工业部提出. 本标准由全国变压器标准化技术委员会归口. 本标准负责起草单位: 沈阳变压器研究所、沈阳第二变压器厂、上海变压器厂. 本标准参加起草单位: 北京变压器厂、顺德特种变压器厂、金乡变压器厂、福州变压器厂、三门变压器厂、江西电气设备厂. 本标准主要起草人: 刘永珍、李桂林、刘大为、朱建荣、周炎士. 本标准 1988 年首次发布.
2	适用范围	本标准规定了家用电冰箱包装技术要求、试验方法、检验规则、标志、运输和贮存. 本标准适用于质量 150 kg 以下的家用电冰箱包装.	本标准规定了非包封线圈的和包封线圈的三相干式电力变压器技术参数和要求、试验方法及检验规则. 本标准适用于电压等级为 6、10、35 kV 级, 额定频率为 50 Hz, 额定容量为 10000 kVA 及以下, 线圈绝缘材料耐热等级为 B、F、H 级, 户内使用的, 无励磁调压或有载调压的, 非包封线圈的或包封线圈的三相干式电力变压器. 本标准不适用于 SF6 气体绝缘的干式变压器.
3	规范性引用文件	GB 1019 家用电器包装通则 GB 2828 逐批检查计数抽样程序及抽样表 (适用于连续批的检查) GB 2829 周期检查计数抽样程序及抽样表 (适用于生产过程稳定性的检查) GB 3095 大气环境质量标准 GB 4456 包装用聚乙烯吹塑薄膜 GB 4706.1 家用和类似用途电器的安全通用要求 GB 4706.13 家用和类似用途电器的安全家用电冰箱和食品冷冻箱的特殊要求 GB/T 4857.1 包装 运输包装件 试验时各部位的标示方法 GB/T 4857.3 包装 运输包装件 静载荷堆码试验方法 GB/T 4857.5 包装 运输包装件跌落试验方法 GB/T 4857.7 包装 运输包装件 定频正弦振动试验方法 GB/T 4857.11 包装 运输包装件 水平冲击试验方法 GB 6543 瓦楞纸箱 GB 6544 瓦楞纸箱 GB/T 8059.1 家用制冷器具 冷藏箱 GB/T 8059.2 家用制冷器具 冷藏冷冻箱 GB/T 8059.3 家用制冷器具 冷冻箱 GB/T 8059.4 家用制冷器具 无霜冷藏箱、无霜冷冻箱、无霜冷冻食品储藏箱和无霜食品冷冻箱 GB 8166 缓冲包装设计方法 GB/T 12023 塑料打包带 QB/T 1649 聚乙烯低发泡塑料包装材料	下列标准所包含的条文, 通过在本标准中引用而构成本标准的条文. 本标准出版时, 所示版本均为有效. 所有标准都会被修订, 使用本标准的各方应探讨使用下列标准最新版本的可能性. GB 191-90 包装、储运指示标志 (neq ISO 780: 1985) GB 1094.1-1996 电力变压器 第一部分 总则 (eqv IEC 76-1: 1993) GB 4208-93 外壳防护等级 (IP 代码) (eqv IEC 529: 1989) GB 5273-85 变压器、高压电器和套管的接线端子 GB/T 5465.2-1996 电气设备用圆形符号 GB 6450-86 干式电力变压器 (eqv IEC 726: 1982) JB/T 501-91 电力变压器试验导则 ZB K41 005-89-6~220 kV 级变压器声级

图6 内容对比

(3) 专题模块

专题模块旨在为用户提供专业权威的领域性标准内容指标汇总. 如图8所示燃气灶具标准内容指标一览, 左侧为系统构建的树状指标导航, 包含了80多种

指标类型; 右侧为当前指标数据库中全部的技术指标, 可以通过左侧导航, 定位到用户关注的标准技术指标. 通过勾选多个当前浏览的指标的小方框, 单击“指标对比”用户可浏览当前选中指标对比详细页面, 如图7

所示. 单机“标准号”可浏览当前标准的全部指标信息; 单击“指标名称”可浏览当前指标的具体信息. 另外单

击右侧表格列名“标准号”, “指标名称”, “指标类型”或“国家”可动态进行显示排序.

标准文献内容指标服务系统

标准信息 专题指标 统计分析 站点介绍 咨询服务 关于我们

锂电池指标对比详情

序号	标准号	指标名称	测试指标	测试方法
1	SJ 20941-2005	强制放电/过放电	电池应不起火、不爆炸。	用10V的直流电源和合适的电子负载或条件相当的放电设备, 以0.2ItA的电流进行放电12.5h, 放电过程不间断。
2	SN/T 1828.14-2006	短路	没有发生解体或起火, 没有发生自测可观察到流出物为毒性、易燃或腐蚀性物质质量损失要求: 电池或电池组质量/g ≤1.0, ≤5.0g 最大质量损失/% 0.5; > 1.0g, ≤5.0g 最大质量损失/% 0.2; >5.0g 最大质量损失/% 0.1	在试验温度55℃下, 电池和电池组用同一根导线连接正极和负极使其短路。直到表壳温度回到55℃之后至少1h。
3	SN T 1414.3-2004	短路	不爆炸、不起火, 电池外部温度不应超过150℃。	1.电池应在外电路总电阻小于50 mΩ的条件下, 分别在20℃±5℃和55℃±2℃的环境温度下进行短路试验, 当电池电压低于0.1 V或电路(或限流设备)不再放电, 同时电池壳温度恢复至不高于初始环境温度以上10℃时, 结束试验。2.测10只电池
4	UL 1642-2012	高倍率充电/非正常充电	无起火或爆炸现象	选用电池为经过5周充电的满充电状态的电池。测试温度: 20 ±5℃ (68 ±9°F)按照出厂说明设置的放电截至电压, 恒定电流放电速率: 0.2 C/1 h, 随后按照出厂说明的最大输出电压和三倍于最大出厂电流进行持续7小时充电。
5	UN38.3	短路	电池壳体温度不超过170℃, 6h观察组装无破裂, 无断裂, 无着火	55±2℃温度下, 电阻小于100 mΩ导线连接电池正负极, 当壳体温度恢复到55±2℃后至少1h结束短路状态, 观察6h后测试结果。

总共1页 合计5条 第1页 [最首页](#) [前一页](#) [后一页](#) [最后一页](#) 跳转到 页

版权所有: 中国标准化研究院 技术支持: 国家标准馆系统运行部

图 7 指标对比

标准文献内容指标服务系统

标准信息 专题指标 统计分析 站点介绍 咨询服务 关于我们

电性能

- 持续充电
- 充电
- 大电流(I C5A) 放电
- 低温放电
- 电池内阻
- 电压
- 短路
- 额定容量/20℃放电容量
- 高倍率充电/非正常充电
- 高温放电
- 过充电保护性能
- 强制放电/过放电

<input type="checkbox"/>	序号	标准号	指标名称	指标类型	国家
<input type="checkbox"/>	1	UN38.3	短路	滥用试验	国际
<input type="checkbox"/>	2	UN38.3	过充电保护性能	滥用试验	国际
<input type="checkbox"/>	3	UL 1642-2012	高倍率充电/非正常充电	滥用试验	美国
<input type="checkbox"/>	4	UL 1642-2012	短路	滥用试验	美国
<input type="checkbox"/>	5	UL 1642-2012	强制放电/过放电	滥用试验	美国
<input type="checkbox"/>	6	UL 1642-2012	温度循环测试	滥用试验	美国
<input type="checkbox"/>	7	SN T 1414.3-2004	短路	滥用试验	中国
<input type="checkbox"/>	8	SN T 1414.3-2004	强制放电/过放电	滥用试验	中国
<input type="checkbox"/>	9	SN T 1414.3-2004	高倍率充电/非正常充电	滥用试验	中国
<input type="checkbox"/>	10	SN T 1414.3-2004	过充电保护性能	滥用试验	中国
<input type="checkbox"/>	11	SN T 1414.3-2004	持续充电	适用性试验	中国
<input type="checkbox"/>	12	SN T 1414.1-2004	高倍率充电/非正常充电	滥用试验	中国
<input type="checkbox"/>	13	SN T 1414.1-2004	过充电保护性能	滥用试验	中国
<input type="checkbox"/>	14	SN T 1414.1-2004	强制放电/过放电	滥用试验	中国

图 8 专题指标浏览

4 结语

在各行各业高速发展的今天, 一些生产、管理、经济等问题日益突出, 如环境污染、食品安全和网络安全等. 而标准在此过程中扮演着重要角色, 也越来越受到各界的重视. 同时在信息技术的推动下标准信息产品应运而生, 一方面为政府决策提供必要的支持, 一方面为行业提供基础技术规范. 当前, 标准信息产品尚不完善, 仍处于发展阶段, 如何更好地迎合用户需求是标准化咨询与服务人员一直探索的问题. 本文基于国家标准馆馆藏资源整合了标准文献内容加

工流程, 建立从标准影印件到标准内容指标的数据加工方法, 形成了基于五种数据资源(影印件、OCR 全文、题录数据、内容数据和指标数据)的标准内容指标服务系统. 该系统采用.NET Framework 4.0 平台技术, 完成了检索、对比、专题和统计分析辅助模块, 实现了标准内容指标在线可查可比可统计的技术方法与应用系统.

参考文献

1 国家标准.GB/T 20000.1-2002 标准化工作指南.第 1 部分:

- 标准化和相关活动的通用词汇.
- 2 李春田.标准化概论(第5版).北京:中国人民大学出版社, 2010:24.
 - 3 卢丽丽,等.永不衰竭的宝库-走进国家标准馆.北京:中国质检出版社,2012.
 - 4 计雄飞,张宝林,王霞,魏利伟.专题服务方式探讨-以标准文献服务为例.标准科学,2014,(2):29-32.
 - 5 计雄飞,张宝林,李抵非,甘克勤.标准文献内容挖掘与比对.标准科学,2012,(8):16-19.
 - 6 蔡华利,施慧斌.国内外标准法规比对分析系统设计与实现.微型机与应用,2011,30(20):77-80.

www.c-s-a.org.cn

www.c-s-a.org.cn