

基于规范文档的学术会议导航系统 构建研究*

赵捷, 杨代庆, 王星

(中国科学技术信息研究所, 北京 100038)

摘要: 为实现NSTL学术会议资源深层揭示和高效利用, 在综合归纳相关研究的基础上, 对学术会议导航系统构建进行分析, 提出基于规范文档的学术会议资源系统设计和构建方案, 开发了轻量级的导航验证系统对相关方案的可行性进行验证, 并对验证结果进行分析。

关键词: 学术会议; 规范文档; 导航系统; 系统设计; 系统建设

中图分类号: G311

DOI: 10.3772/j.issn.1673-2286.2015.07.004

学术会议一般由相关学科领域的学术机构主办, 一些新的科学观点、科学发现、研究进展等往往率先在学术会议上发表并通过学术会议文献公诸于众, 从而使学术会议文献资源具备良好的时效性、新颖性和较高的学术价值。根据国际大会及会议协会(International Congress & Convention Association, 简称ICCA)统计, 全世界每年召开的学术会议达到上万个, 产生的学术会议文献资源数量极为庞大^[1]。会议录是学术会议文献资源的主要载体, 有印刷版、光盘版、网络版、E-Only等多种载体类型。因会议录收集出版机构众多, 交流传播方式灵活, 使得会议录的出版形式呈现出一定的灵活性和复杂性。

也正因为学术会议资源所具有的诸多特性, 导致相关会议名称、主承办机构名称等内容呈现较大程度的多样性和不规范性, 从而给学术会议资源的揭示利用带来不便^[2], 特别是给学术会议导航系统的构建带来不便。所以, 需要通过学术会议资源清理规范和导航系统构建, 实现学术会议资源的深层揭示与高效利用。

1 相关研究

在学术会议导航系统构建中, 相关研究成果值得参考与借鉴。朱江等^[3]认为学术会议资源具有学科范围广、数量多、学术水平差异大、出版形式多样等特点, 与图书、期刊相比, 形式特征、内容特征、著录格式、检索习惯等存在明显区别。梅梅等^[4]对医学学术会议资源特点进行总结归纳, 认为医学学术会议资源内容新颖, 前瞻性、动态性、延续性、时效性、隐蔽性等特性。潘晓蓉等^[5]对1999年至2010年国际出版的重要学术会议录信息进行分析, 揭示了会议举办国、会议文献数量、出版机构、出版形式和出版时滞性等5个方面的出版规律及特征。在学术会议资源应用研究方面, 王韬等^[6]针对外文学术会议资源复杂多变的特点, 讨论了造成检索相关资源命中率低的原因, 并结合用户使用习惯介绍了通过会议名称、会议简称、专题题名、会议主题等途径检索外文学术会议资源的技巧。任平^[7]结合网络环境下学术会议资源呈现出的发布形式、获取

* 本研究得到国家自然科学基金项目“机构规范文档结构及构建方式研究”(编号: 15BTQ015)和国家科技图书文献中心项目“外文会议编目数据清理”(编号: 2013XM47)资助。

方式多样化等新特点,从用户需求角度系统介绍了网络环境下如何利用国内外学术会议馆藏目录(OPAC)数据库、会议录数据库、会议论文数据库、搜索引擎等检索工具查找学术会议题录、会议录、会议论文、会议日程等信息的相关途径与获取技巧。谢静^[8]等以“中国学术会议在线”网站为研究对象,从交互平台、资源情况、学术交流氛围等方面对其运营模式进行了分析,指出了服务优势,并针对资源与服务方面存在的问题提出了改进意见。

目前许多主流数据库系统均提供学术会议资源导航服务功能。国际上,SPIE、IEEE、EI、AIAA等数据库系统均具备学术会议资源导航功能,可以按不同维度、不同层次对相应机构召开的会议、文集、论文进行关联聚类,还可以将不同年度连续召开的会议及其文集组织在一起进行关联展示,使读者能够从会议、文集、主题等维度浏览查找所需的学术会议资源,也能够检索结果中从会议、出版者、主题等维度快速定位到所需内容,极大地方便了使用。国内方面,万方、CNKI等数据库系统具备中文学术会议资源导航功能,主要从学科、主办单位等维度对会议、文集及论文进行聚类,但相关系统在名称统一表达(如会议名称、主承办单位名称等)、会议多层揭示、会议关联聚类等功能实现方面存在缺陷,会议与文集之间的关联性不够,特别是没有将连续召开的会议及其文集组织在一起关联展示。例如,不能采用统一规范的名称表达同一会议、同一机构,不能逐层揭示具有一定层级关系的会议,不能关联揭示具有一定相关关系的会议,这给读者查找所需会议文献带来不便。究其根源,主要是相关会议数据存在多样性问题。为解决相关问题,需要基于规范文档来构建学术会议导航系统。

2 导航系统设计

2.1 导航服务需求分析

构建学术会议导航系统,对相关资源的数据规范性和一致性要求较高。NSTL采用基于CNMARC的编目规则描述并保存学术会议资源,按照遵循实物的著录原则,无法以统一规范的方式对学术会议名称、主办者名称、承办者名称、会议届次、会议时间、会议地点等属性进行表达,更无法对会议间存在的连续关系(如连续召开的会议)、并列关系(如联合召开的会议)、

从属关系(如大会与子会)、用代关系(如会议规范名称与会议不规范名称的对照关系)等关系进行表达,因而很难对相关内容进行统一表达和关联聚类。为解决相关问题,需要采用会议元数据描述方案对学术会议资源各种属性进行规范表达,采用会议名称规范文档对会议名称进行统一表达并对会议间的各种关系进行描述,采用机构名称规范文档对机构名称进行统一表达。利用学术会议元数据以及相关规范文档构建学术会议导航系统,实现学术会议资源的多维表达、深层揭示与关联聚类,进而引导读者有效利用相关资源。

2.2 导航系统功能设计

导航系统功能框架如图1所示,划分为导航管理功能和导航服务功能。导航管理功能可按一定策略规划导航树,对导航页面内容布局进行管理。导航服务功能可根据导航管理功能所规划的导航树和所定义的导航页面内容布局,对所有学术会议资源或检索命中的部分学术会议资源进行信息组织与聚类,生成导航树和其他导航内容,从而实现多维导航和多层次导航。学术会议导航服务形成机制如图2所示。

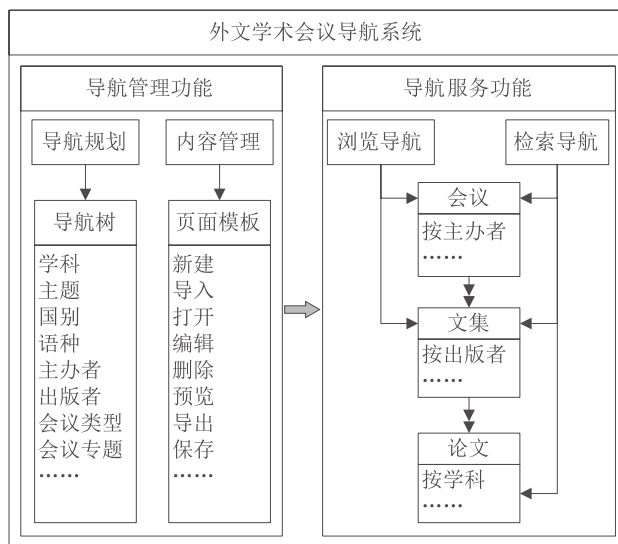


图1 导航系统功能框架

导航管理主要用来定义、命名、设置、调整导航树。导航树是导航系统定义的数据对象,可被导航系统识别。导航系统根据相关属性生成导航树,导航管理采用所见即所得的编辑方式对相关属性进行管理。内容管理功能主要用于预先定制导航页面模板,采用可被导航系统识别与理解的脚本语言描述导航页面,设定

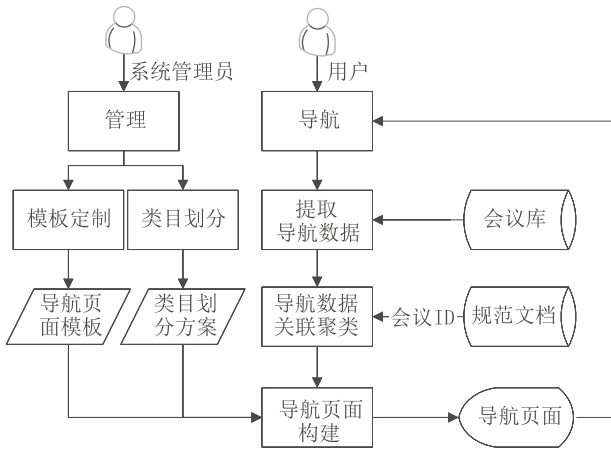


图2 导航服务形成机制

标题、分类树、表格等内容的位置与布局，通过预览功能，可以查看导航页面的呈现效果。

导航服务包括浏览导航和检索导航功能。浏览导航服务功能按照预先制订的信息组织策略对全部学术会议资源进行多维组织，通过逐层递进的方式引导读者发现利用所需的相关会议资源。检索导航服务功能根据检索结果集，对命中的相关会议资源进行动态组织，通过分面方式缩小检索范围，并通过逐层递进或递归的方式帮助读者快速定位所需的会议资源。

2.3 学术会议资源描述

学术会议资源蕴含着丰富的科研实体，如涉及各个科学主题的创作作品，创作、生产、传播、保存相关作品的个人与组织等，相关实体之间存在广泛的、相互联系的关系网络。学术会议导航系统正是借助这种关系网络，引导读者从网络中的某一节点航行到网络中其它任意节点。对学术会议资源的描述需要反映其中各实体间的相互关系，以便导航系统能够识别并按相应关系进行导航。

为更清晰地描述学术会议资源各类实体及其相关关系，采用书目记录功能需求模型（Functional Requirements of Bibliographic Records, 简称FRBR）对其进行描述。FRBR是一种描述书目记录结构及关系的概念模型，它为探讨书目记录的结构与关系并为实现书目记录与用户基本需求的关联提供了一个定义清晰的结构化框架。模型分3组定义了10个实体及其属性，并定义了实体间的关系、实体及其属性与用户任务的映射关系^[9]。其中，实体及基本关系如图3所示。

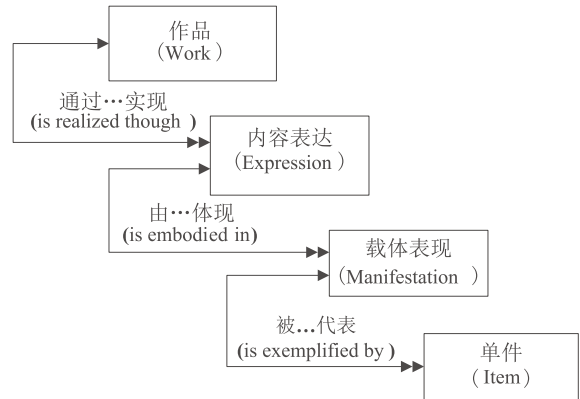


图3 实体及基本关系

学术会议资源主要涉及会议、会议录、会议论文、馆藏信息及规范文档5个实体元素集，描述框架如图4所示。它客观地描述了各实体元素集间的相互关系，学术会议导航系统多层次导航功能正是基于这些关系。

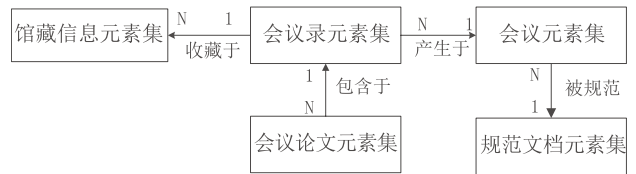


图4 学术会议资源元数据描述框架

实体元素集元素定义需要考虑导航系统各项功能需求，提出各元素集中与导航功能实现密切相关的元素简表。

2.4 导航系统构建

针对NSTL学术会议资源现状和学术会议导航系统构建需求，本研究采用了图5所示的构建路线。



图5 学术会议导航系统构建路线图

首先，需要深入开展用户需求调研和数据质量分析，明确NSTL用户的实际需求，掌握NSTL相关数据质量现状，发现其中存在的主要问题；第二，根据用户实际需求优化调整系统设计；第三，根据系统最终设计，建立健全相关元数据标准规范；第四，结合数据质量现状，有针对性地开展相关数据的清理规范，构建规范文档，解决相关数据质量问题；第五，依据系统设计及相

关元数据标准规范,进行系统集成开发,最终实现导航系统各项功能。

通过分析调研,构建学术会议导航系统重点要将数据清理规范,以解决数据质量问题。为保证清理工作有章可循,本研究首先对《NSTL联合目录系统会议文献元数据规范》、《国家科技图书文献中心会议录编目手册》等标准规范提出了修订完善建议并得到采纳,在此基础上制订了《会议规范文档元数据方案》、《会议信息元数据著录规则》、《会议数据清理规则》等规范文件,形成了清理规范所需的标准规范体系,从而为清理规范工作提供必要的工作准则。

会议数据清理规范由NSTL各成员单位共同完成。为保证清理规范协同有序,本研究设计开发了网络化的会议清理规范协同工作平台。平台具备三类基本功能:一是可以实现CNMARC数据的检索、浏览与编辑,二是可以实现元数据的抽取、转换、校验与保存,三是可以同步进行会议规范文档建设。利用该平台可以有效管理清洗规范各环节的工作,明确分工、划清责任;可以统计工作数量,追溯工作质量;可以简化操作,提高工作效率;还可以作为元数据上传的预处理工具,对新增数据进行质量控制,防止不合格数据再度产生。图6为基于会议清理规范协同工作平台的工作流程。

按照会议规范文档元数据方案,本研究提出了会议规范库构建方案并依据该方案建成了会议规范库,其中包括:会议名称规范文档、会议主承办机构名称规范文档、规范名称/原有名称对照文档、会议关系文档等内容。会议名称、主承办机构名称规范情况如表1所示。

表1 名称规范情况简表

	原有名称	规范名称	归一率
会议	116580	36946	3.1
主承办机构	135166	15773	8.5

会议名称规范效果如图7所示,主承办机构名称规范效果如图8所示。

3 导航设计方案验证

为验证设计方案的可行性和会议规范文档的适用性,评估清理规范后的数据质量并为学术会议导航系统的最终开发提供参考,本研究根据学术会议导航系统设计方案,利用会议规范文档开发构建了导航

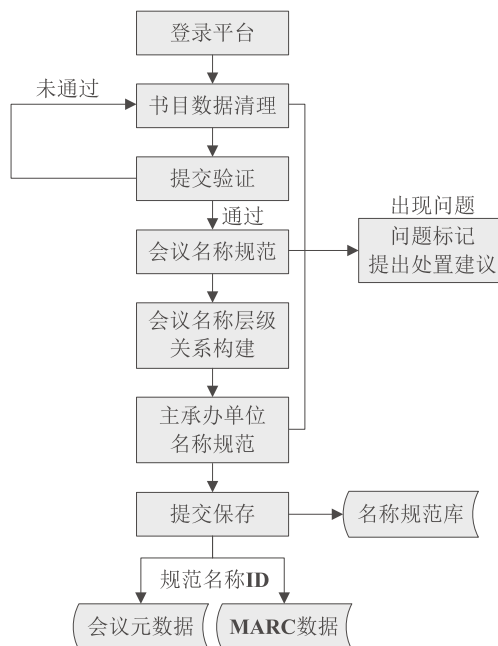


图6 基于会议清理规范协同工作平台的工作流程

规范会议名称	原会议名称
Electrochemical Society meeting	Electrochemical Society(ECS) Meeting
	Electrochemical Society
	ECS Meeting
	Electrochemical Society meeting
	ECS Meeting/PRIME
	Meeting of the Electrochemical Society
	Organic and biological electrochemistry posters (General)-ECS meeting
	Meeting of Electrochemical Society
American Chemical Society meeting	American Chemical Society Meeting
	American Chemical Society National Meeting
	American Chemical Society
	ACS National Meeting
	Spring national ACS meeting
	American Chemical Society meetings

图7 会议名称规范效果

规范机构名称	原机构名称
American Society of Mechanical Engineers	American Society of Mechanical Engineers
	ASME Material handling Engineering Division
	American Society for Mechanical Engineers
	ASME
	American Society of Mechanics Engineers
	ASME International
	American Society of Mechanical Engineer
	American Society of Mechanical Engineering
	02American Society of Mechanical Engineer
	02American Society of Mechanical Engineers
	American Society of Mechanical Engineering(US)
	American Society of Mechanical Engineers
	American Society of Mechanical Engineers

图8 主承办机构名称规范效果

验证系统。系统使用的数据为清理后的86383条会议数据,采用C/S系统架构,数据库管理系统为MS SQL SERVER,运行环境要求不高,可以联网使用,也可以单机安装使用。

导航验证系统主要验证了浏览导航服务功能设计,如图9所示。由于条件所限,未能对检索导航和导航管理功能设计进行验证。采用清理规范后的会议数据及会议规范文档可以实现会议名称、主承办机构名称等会议信息的统一、规范、有序,将会议关联聚类,实现学术



会议名称	年份	届次	会议类别
IEEE Conference Report	2000	第11届	IEEE Conference Report
IEEE Conference Report	2001	第12届	IEEE Conference Report
IEEE Conference Report	2002	第13届	IEEE Conference Report
IEEE Conference Report	2003	第14届	IEEE Conference Report
IEEE Conference Report	2004	第15届	IEEE Conference Report
IEEE Conference Report	2005	第16届	IEEE Conference Report
IEEE Conference Report	2006	第17届	IEEE Conference Report
IEEE Conference Report	2007	第18届	IEEE Conference Report
IEEE Conference Report	2008	第19届	IEEE Conference Report
IEEE Conference Report	2009	第20届	IEEE Conference Report
IEEE Conference Report	2010	第21届	IEEE Conference Report
IEEE Conference Report	2011	第22届	IEEE Conference Report
IEEE Conference Report	2012	第23届	IEEE Conference Report
IEEE Conference Report	2013	第24届	IEEE Conference Report
IEEE Conference Report	2014	第25届	IEEE Conference Report
IEEE Conference Report	2015	第26届	IEEE Conference Report
IEEE Conference Report	2016	第27届	IEEE Conference Report
IEEE Conference Report	2017	第28届	IEEE Conference Report
IEEE Conference Report	2018	第29届	IEEE Conference Report
IEEE Conference Report	2019	第30届	IEEE Conference Report
IEEE Conference Report	2020	第31届	IEEE Conference Report
IEEE Conference Report	2021	第32届	IEEE Conference Report
IEEE Conference Report	2022	第33届	IEEE Conference Report
IEEE Conference Report	2023	第34届	IEEE Conference Report
IEEE Conference Report	2024	第35届	IEEE Conference Report
IEEE Conference Report	2025	第36届	IEEE Conference Report
IEEE Conference Report	2026	第37届	IEEE Conference Report
IEEE Conference Report	2027	第38届	IEEE Conference Report
IEEE Conference Report	2028	第39届	IEEE Conference Report
IEEE Conference Report	2029	第40届	IEEE Conference Report
IEEE Conference Report	2030	第41届	IEEE Conference Report
IEEE Conference Report	2031	第42届	IEEE Conference Report
IEEE Conference Report	2032	第43届	IEEE Conference Report
IEEE Conference Report	2033	第44届	IEEE Conference Report
IEEE Conference Report	2034	第45届	IEEE Conference Report
IEEE Conference Report	2035	第46届	IEEE Conference Report
IEEE Conference Report	2036	第47届	IEEE Conference Report
IEEE Conference Report	2037	第48届	IEEE Conference Report
IEEE Conference Report	2038	第49届	IEEE Conference Report
IEEE Conference Report	2039	第50届	IEEE Conference Report
IEEE Conference Report	2040	第51届	IEEE Conference Report
IEEE Conference Report	2041	第52届	IEEE Conference Report
IEEE Conference Report	2042	第53届	IEEE Conference Report
IEEE Conference Report	2043	第54届	IEEE Conference Report
IEEE Conference Report	2044	第55届	IEEE Conference Report
IEEE Conference Report	2045	第56届	IEEE Conference Report
IEEE Conference Report	2046	第57届	IEEE Conference Report
IEEE Conference Report	2047	第58届	IEEE Conference Report
IEEE Conference Report	2048	第59届	IEEE Conference Report
IEEE Conference Report	2049	第60届	IEEE Conference Report
IEEE Conference Report	2050	第61届	IEEE Conference Report
IEEE Conference Report	2051	第62届	IEEE Conference Report
IEEE Conference Report	2052	第63届	IEEE Conference Report
IEEE Conference Report	2053	第64届	IEEE Conference Report
IEEE Conference Report	2054	第65届	IEEE Conference Report
IEEE Conference Report	2055	第66届	IEEE Conference Report
IEEE Conference Report	2056	第67届	IEEE Conference Report
IEEE Conference Report	2057	第68届	IEEE Conference Report
IEEE Conference Report	2058	第69届	IEEE Conference Report
IEEE Conference Report	2059	第70届	IEEE Conference Report
IEEE Conference Report	2060	第71届	IEEE Conference Report
IEEE Conference Report	2061	第72届	IEEE Conference Report
IEEE Conference Report	2062	第73届	IEEE Conference Report
IEEE Conference Report	2063	第74届	IEEE Conference Report
IEEE Conference Report	2064	第75届	IEEE Conference Report
IEEE Conference Report	2065	第76届	IEEE Conference Report
IEEE Conference Report	2066	第77届	IEEE Conference Report
IEEE Conference Report	2067	第78届	IEEE Conference Report
IEEE Conference Report	2068	第79届	IEEE Conference Report
IEEE Conference Report	2069	第80届	IEEE Conference Report
IEEE Conference Report	2070	第81届	IEEE Conference Report
IEEE Conference Report	2071	第82届	IEEE Conference Report
IEEE Conference Report	2072	第83届	IEEE Conference Report
IEEE Conference Report	2073	第84届	IEEE Conference Report
IEEE Conference Report	2074	第85届	IEEE Conference Report
IEEE Conference Report	2075	第86届	IEEE Conference Report
IEEE Conference Report	2076	第87届	IEEE Conference Report
IEEE Conference Report	2077	第88届	IEEE Conference Report
IEEE Conference Report	2078	第89届	IEEE Conference Report
IEEE Conference Report	2079	第90届	IEEE Conference Report
IEEE Conference Report	2080	第91届	IEEE Conference Report
IEEE Conference Report	2081	第92届	IEEE Conference Report
IEEE Conference Report	2082	第93届	IEEE Conference Report
IEEE Conference Report	2083	第94届	IEEE Conference Report
IEEE Conference Report	2084	第95届	IEEE Conference Report
IEEE Conference Report	2085	第96届	IEEE Conference Report
IEEE Conference Report	2086	第97届	IEEE Conference Report
IEEE Conference Report	2087	第98届	IEEE Conference Report
IEEE Conference Report	2088	第99届	IEEE Conference Report
IEEE Conference Report	2089	第100届	IEEE Conference Report

图9 学术会议导航验证系统实现情况

会议资源的多维导航、多层揭示。导航验证系统的开发为今后NSTL构建学术会议导航系统提供了参考。

4 结语

学术会议导航系统是用户发现和利用学术会议资源的重要途径。鉴于学术会议资源的复杂性，构建功能强大的学术会议导航系统必须保证较高的数据质量，会议名称、主承办机构名称等需要具有高度的一致性和规范性。规范文档既是保证数据一致性和规范性的重要手段，又是强化导航系统功能的重要保障。因此，实践中要不断加强规范文档相关研究，逐步改进规范文档构建方法，在学术会议导航系统中充分利用规范文档，以强化和丰富系统功能。

参考文献

[1] ICCA.The International Association Meetings Market 2002-2011[EB/

OL].[2014-06-29]. <http://www.iccaworld.com/dcps/doc.cfm?docid=1520>.
[2] 李树国. 西文会议录检索中常见问题及检索对策[J].中国图书馆学报,2002,28(3):75-78.
[3] 朱江,陈漪红. 科技会议文献管理与服务现状及改进对策[J].四川图书馆学报,2005(4):40-43.
[4] 梅梅,唐小利. 医学会议文献的传播模式和获取策略研究[J].医学信息学杂志,2012,33(4):69-72.
[5] 潘晓蓉,丁道劲. 国外学术会议录出版特征研究[J].科技与出版,2013(5):49-52.
[6] 王韬,何小明. 西文会议录的著录与检索[J].图书馆学研究,2004(6):65,77-78.
[7] 任平. 网络环境下会议文献的新特点与获取之道[J].图书情报工作,2007,51(1):103,108-110.
[8] 谢静,李璐,赵康. “中国学术会议在线”运营模式研究[J].科技传播,2014(1).
[9] 国际图联书目记录的功能需求研究组.书目记录的功能需求最终报告[R/OL].王绍平,王静,林明等,译.德国绍尔出版社,1998.[2014-7-29]. <http://www.nlc.gov.cn/newtsgj/ijflaygt/jtjtlzwyzyx/tlzlby/201011/P020101130497421213155.pdf>.

作者简介

赵捷,男,1959年生,中国科学技术信息研究所信息资源中心副主任,研究馆员,研究方向:数据挖掘与知识服务, E-mail: zhaojie@istic.ac.cn。
杨代庆,男,1975年生,中国科学技术信息研究所信息资源中心副主任,高级工程师,研究方向:信息系统构建。
王星,男,1977年生,中国科学技术信息研究所工程师,研究方向:信息管理与信息系统。

Research of Conference Navigation System Construction Based on Standardized Documentations

Zhao Jie, Yang DaiQing, Wang Xing
(Institute of Scientific and Technical Information of China, Beijing, 100038)

Abstract: In order to achieve the conference resources of NSTL reveal deeply and use efficiently. On the basis of summarized relevant research , the paper analysis the construction of academic conference navigation system and put forward a scheme for designning and constructing a conference resources system which based on standardized documentation, and then developed a lightweight navigation validation system to validate the feasibility of the relevant solutions. At last, the paper analysed the verification results.

Keywords: Conference; Standardized Documentation; Navigation System; System Design; System Construction

(收稿日期: 2015-07-01; 编辑: 雷雪)