

ICS 01.140.20

A 14

C A D A L 项 目 标 准

CADAL 41201—2012

数字图书馆知识组织服务标准

Digital Library Knowledge Service Standard

第一稿

2012-05-08 发布

2012-05-09 实施

CADAL 项目管理中心 发 布

目 次

前言	95
引言	96
1 范围	97
2 规范性引用文件	97
3 术语和定义	97
3.1 知识服务	97
3.2 知识组织	97
3.3 知识关联	97
3.4 参考文献	97
3.5 概念	98
3.6 引证文献	98
3.7 同被引文献	98
4 知识服务系统总论	98
5 知识导航	99
5.1 按知识源导航	99
5.2 按学科知识点导航	99
5.3 按热门专题导航	99
6 知识检索	100
6.1 知识点检索	100
6.2 字段检索	100
6.3 知识检索技术	101
7 知识网络	103
7.1 知识关联网	103
7.2 文献关联网	106
8 用户知识空间	107
8.1 用户知识交互	107
8.2 用户知识管理	108
8.3 用户社区空间	108
参考文献	109

CADAL 项目标准规范汇编(四)

CADAL 41201—2012

图 1	数字图书馆知识体系结构	98
图 2	知识导航——图书推荐示例	99
图 3	知识导航——热点推荐示例	99
图 4	一站式检索示例	101
图 5	知识脉络——发展趋势示例	105
图 6	知识脉络——发展比较示例	106
图 7	文献引证关系	106
表 1	布尔逻辑检索运算符号示例	101
表 2	位置检索运算符号示例	102
表 3	知识呈现维度分类	104

前 言

《数字图书馆知识组织服务标准》主要包括以下四个方面的内容：

- 知识导航服务；
- 知识检索服务；
- 知识发现与网络构建；
- 用户知识交流空间。

CADAL 数字图书馆知识组织服务标准是依据中美双方对 CADAL 内容的共同要求，参照国家科技部科技基础研究重大科技专项“我国数字图书馆标准与规范建设”项目之《数字图书馆集成服务描述标准发展现状》、《知识组织体系应用标准规范研究》、《数字图书馆相关领域标准规范现状与发展研究——数字科研》等研究成果，对数字图书馆的服务方式、服务对象及相关技术发展进行分析后制定的。

本标准由大学数字图书馆国际合作计划(CADAL)项目管理中心归口。

本标准起草单位：宁波大学图书馆。

本标准主要起草人：豆洪青、刘柏嵩、董其军、毛海波。

引 言

基于知识组织技术的知识服务标准化是为了把知识组织成果更好地呈现给用户，通过标准化方法，定义数字图书馆知识服务所涉及的服务方式与内容，指导数字图书馆知识服务工作的发

数字图书馆知识组织服务标准

1 范围

本标准规定了数字图书馆知识服务的总体要求和方法。

本标准适用于文献资源数据库服务平台。

2 规范性引用文件

下列文件对于本标准的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本标准。凡是不注日期的引用文件，其最新版本(包括所有的修改版)适用于本标准。

数字图书馆标准规范发展战略子项目组. 我国数字图书馆标准规范建设与应用的实施指南. [2012-04-05]. <http://cdls2.nstl.gov.cn/2003/Whole/TecReports.html>.

潘淑春, 盛玲玉, 牛离平. 数字图书馆相关领域标准规范现状与发展研究(数字科研). [2012-04-05]. <http://cdls2.nstl.gov.cn/2003/Whole/TecReports.html>.

张晓林. 数字图书馆集成服务描述标准发展现状. [2012-04-05]. <http://cdls2.nstl.gov.cn/2003/Whole/TecReports.html>.

富平, 等. 数字图书馆知识组织体系标准规范应用机制研究. [2012-04-05]. <http://cdls2.nstl.gov.cn/2003/Whole/TecReports.html>.

3 术语和定义

3.1 知识服务

知识服务是以信息知识的搜寻、组织、分析、重组的知识和能力为基础, 根据用户的问题和环境, 将其融入用户解决问题的过程中, 提出能够有效支持知识应用和知识创新的服务。

3.2 知识组织

知识组织指对知识单元进行标引, 并构建其语义关系。

3.3 知识关联

知识关联指用于指示知识之间关系的组织形态。

3.4 参考文献

参考文献是文章或著作等写作过程中参考过的文献。

3.5 概念

概念指人们对事物本质的认识,逻辑思维的最基本单元和形式。在此表示最基本的知识点。

3.6 引证文献

引证文献指引用本部分的文献。

3.7 同被引文献

同被引文献指与本部分同时被作为参考文献引用的文献,是与本部分共同作为进一步研究的基础。

4 知识服务系统总论

知识服务系统是依据知识组织技术,把知识之间的关系在网站上进行呈现,为用户提供知识级的服务。知识服务系统具有知识检索功能、知识揭示功能、知识关联分析、个性化知识服务功能;用户与用户之间、用户与系统之间实现动态交互功能;网站内的信息可以直接和其他网站相关信息进行交互,能通过第三方信息平台同时对多家网站的信息进行整合使用。

数字图书馆知识体系结构如图 1 所示。

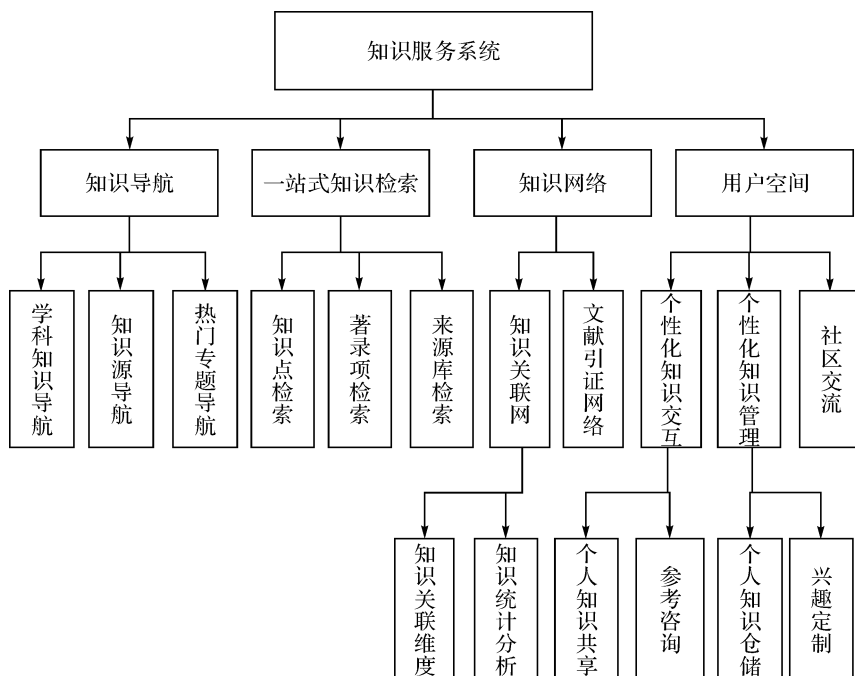


图 1 数字图书馆知识体系结构

5 知识导航

知识导航指通过门户网站，提供馆藏数字化资源及互联网相关资源的导航功能。

5.1 按知识源导航

根据知识的来源类型进行导航，如知识在期刊、图书、报纸、图谱、视频、互联网资源不同来源中进行呈现。图书推荐示例如图 2 所示。



图 2 知识导航——图书推荐示例

5.2 按学科知识点导航

按学科知识点导航指根据每个学科下的知识点进行知识导航。

5.3 按热门专题导航

按热门专题导航指根据统计数据，自动生成热门专题，或可根据用户个性化需要进行功能定制。热点推荐示例如图 3 所示。

点击TOP10	搜索TOP10
▪ 红楼梦(一)	▪ 论语
▪ 中国书法全集--商周金文卷	▪ 红楼梦
▪ 围城	▪ 教育杂志
▪ 中国书法全集--颜真卿一	▪ 申报
▪ 玉簪先生文集(一)	▪ 金瓶梅
▪ 绘图真本金瓶梅	▪ 教育
▪ 存素堂集续编(四)	▪ 史记
▪ 金瓶梅①	▪ 东方杂志
▪ 新青年	▪ 桐旧集
▪ 中国书法全集--王羲之王献之二	▪ 新青年

图 3 知识导航——热点推荐示例

6 知识检索

知识检索是使用知识本体构建的领域查询推理系统，主要包含三个部分：领域知识本体、知识本体之间的映射和内部的推理引擎。知识本体库是提供推理引擎进行推理的事实基础；知识本体之间的映射处理则是建立几个知识本体之间联系的桥梁；推理引擎是进行查询推理的基础。

提供一站式(跨库)检索，以多个分布式异构数据源为对象的检索系统，集成并兼容不同类型、不同渠道、不同格式的资源数据库，向用户提供统一的检索接口，只需在检索框中输入检索词，便可在多个电子资源中检索到相关信息，甚至可链接到全文。

6.1 知识点检索

知识点检索是根据以揭示各学科专业词汇(关键词、术语、主题词)所代表的概念之间的关系为基本内容的专业知识库系统。词和词之间含各种关系：同义、缩略语、译名、上下位等。设置检索智能扩展功能和概念导航功能。检索智能扩展基于概念关系词典相关语义场运算技术，实现不同字段间的智能扩展，能够提高查准查全率，便于发现新知识，实现知识漫游服务。概念导航基于概念关系词典相近语义场运算技术，以文本和图像两种形式表现概念之间的相近关系，实现概念导航，能够提高查全率，便于发现新知识。

6.2 字段检索

字段检索是把检索词限定在某个/些字段中，如果记录的相应字段中含有输入的检索词，则为命中记录，否则检不中。

6.2.1 主题词检索

主题词检索是指用叙词、标题词、单元词、关键词等表达信息需求而进行的检索。主题词是规范化的检索语言，它对文献中出现的同义词、近义词、多义词以及同一概念的不同书写形式等进行严格的控制和规范，使每个主题词都含义明确，以便准确检索，防止误检、漏检。主题词表是对主题词进行规范化处理的依据，也是文献处理者和检索者共同参照的依据。

6.2.2 句子检索

检索与句子相似的结果，利用句子检索，实现对事实的深度检索。

示例：上肢麻木和头痛的症状是由什么疾病引起的？运用“句子检索”来找到相关答案。

第一步：选择句子检索，输入关键词。

检索在句子中同时含有“上肢麻木”和“头痛”的句子。

第二步：浏览句子，进一步筛选。

在查到的结果中，进行浏览，若每个句子是一条事实，那么检索结果提供了许多可能的答案。

6.2.3 外部特征检索

外部特征检索包括时间、出版来源、ISSN(与CN号自动关联)、作者、基金、机构、版本等字段限定检索功能。

6.2.4 来源库检索

在一站式检索中,如图4所示,可根据来源数据库进行选择检索,并自动选择该来源数据库,提供来源可检索字段限定。



图4 一站式检索示例

6.3 知识检索技术

6.3.1 布尔逻辑检索

布尔逻辑检索运算符号示例见表1。

表1 布尔逻辑检索运算符号示例

运算符	检索功能	检索含义	举例	适用检索项
逻辑“与”:运算符为 And 或 *	AND,并且包含。逻辑与检索能增强检索的专指性,使检索范围缩小,此运算符适于连接有限定关系或交叉关系的词	包含 str1 和 str2	TI= 转基因 * 水稻	所有检索项
逻辑“或”:运算符为 Or 或 +	OR,或者包含。逻辑或检索扩大了检索范围,此运算符适于连接有同义关系或相关关系的词	包含 str1 或者 str2	TI= 转基因 + 水稻	
逻辑“非”:运算符为 Not 或 -	NOT,不包含。逻辑非和逻辑与运算的作用类似,可以缩小检索范围,增强检索的准确性。此运算适于排除那些含有某个指定检索词的记录	包含 str1 不包含 str2	TI= 转基因 - 水稻	

使用“AND”、“OR”、“NOT”等逻辑运算符,“()”符号将表达式按照检索目标组合起来。

注意事项:

- 所有符号和英文字母,都必须使用英文半角字符。

- “AND”、“OR”、“NOT”三种逻辑运算符的优先级相同；如要改变组合的顺序，请使用英文半角圆括号“()”将条件括起。
- 逻辑关系符号（与(AND)、或(OR)、非(NOT)）前后要空一个字节。

6.3.2 位置检索

对词之间的邻近关系进行检索，适合全文检索精确检索，以达到深度检索的目的。词与词之间通过运算符相连。

位置检索运算符示例见表 2。

表 2 位置检索运算符示例

%str	包含	包含词 str 或 str 切分的词	TI%转基因水稻	全文、主题、题名、关键词、摘要、中图分类号
=str	包含	包含检索词 str	TI=转基因水稻	
=‘str1 /SEN N str2’	同段，按次序出现，间隔小于 N 句		FT=转基因 /SEN 0 水稻	
=‘str1 /NEAR N str2’	同句，间隔小于 N 个词		AB=转基因 /NEAR 5 水稻	主题、题名、关键词、摘要、中图分类号
=‘str1 /PREV N str2’	同句，按词序出现，间隔小于 N 个词		AB=转基因 /PREV 5 水稻	
=‘str1 /AFT N str2’	同句，按词序出现，间隔大于 N 个词		AB=转基因 /AFT 5 水稻	
=‘str1 /PEG N str2’	全文，词间隔小于 N 段		AB=转基因 /PEG 5 水稻	
=‘str /with N’	要求检索词必须按指定顺序紧密相连，词序不可变，词之间除允许有空格、标点、连字符外，不得夹单词或字母；(Nw)表示连接的两个词之间最多可夹入 N 个词(N 为自然数)，词序不得颠倒		TI=转基因/with 水稻	

6.3.3 智能检索

智能检索以文献和检索词的相关度为基础，综合考查文献的重要性等指标，对检索结果进行排序，以提供更高的检索效率。智能检索的结果在排序的同时考虑相关性和重要性。相关性采用各字段加权混合索引，相关性分析更准确；重要性指通过对文献来源进行权威性分析和引用关系分析等实现对文献质量的评价。

适用于题名、关键词、摘要、主题等检索项用叙词、标题词、单元词、关键词等表达信息需求而进行的检索。自由词自动与《汉语主题词表》进行匹配。输入的检索词能自动纠错，并与知识标引的领域本体自动关联。包括同义扩展和属性扩展两种类型。

同义扩展：这是最基本的扩展类型，即通过本体定义获取检索提问中概念的同义词。

属性扩展：通过本体定义获取检索提问中与主语相关的属性定义。

7 知识网络

知识网络是对检索结果内容在知识层次进行呈现与增值服务。主要体现为概念地图技术。

概念地图是用层级结构的方式表示概念之间的关系。在概念地图中,概念是用层级结构的方式来呈现的。其中,最广泛、最一般的概念置于地图的上端,次一般和更具体的概念按等级排在下面。特殊知识领域的层级结构根据知识应用或思考的情景而定。概念地图的另一个特征是交叉连接。交叉连接用于表示概念地图中概念之间的关系(命题)。交叉连接表明了呈现在地图上的某些领域,知识是怎样相联系的。

7.1 知识关联网

知识关联主要是通过知识元链接和引文链接将文献间的知识关联起来。通过文献之间、知识元之间、分类导航之间的交叉链接,构建起节点丰富、交织纵横的知识网络系统。知识元链接包括作者、机构、刊名、关键词、相关作者群、相关研究机构、相关关键词等,通过这些链接可检索知识网络中心配制的数据库的相应字段,获取检索结果。也可结合知识库查找知识元的概念解释。同时,还可方便查看其他相关类别下的文献。知识网络系统强调库间关联,多个库联成有机整体,库间引文链接和知识网络中心均体现了这一功能特点。

知识关联有知识内容之间本质的内在属性的关联和知识附加属性之间的关联两种形式,根据知识的内在属性和附加属性,知识关联的多维度分为基本维度和扩展维度两类。

7.1.1 知识关联维度

(1) 学科维度:知识所属的学科范畴。

任何知识都需要用语言或其他符号表示,这些语言或符号在不同的语境中可能会表达不同的内容。只有先进行学科维度的归类,才能准确地定义该知识。

任何知识都可按照一定的标准,进行其学科的划分,如根据知识的普遍意义可分为世俗知识、科学知识、人文知识、社会科学知识、艺术知识、没有文字的知识(如视听艺术)等;还可以根据杜威十进分类法、中国图书馆分类法等专门的学科分类体系对知识从学科的维度进行呈现。

对海量知识进行学科维度描述的基础在于领域本体的构建,根据概念之间的相互关系来确定学科领域,并根据已构建的领域本体进行知识的学科呈现。

学科维度的关联,是通过检索结果进行学科聚类的方式呈现。

(2) 时空维度:知识来源。

知识的实效性存在于一定的时空里,也就是说在一定的时间里和一定的空间里才能显现出它的价值。利用知识与知识间时空属性作为中介,将知识关联起来,以获取范围更小的一个关联结果。知识的时空维度主要体现在其载体的来源上,如出版时间、出版物名称等。

通过文献出版时间及单位分别进行聚类呈现。

(3) 呈现维度：显性与隐性。

根据知识能否清晰地表述和有效地转移,可以把知识分为显性知识(explicit knowledge)和隐性知识(tacit knowledge)。隐性知识通过知识关联、挖掘可以转换为显性知识。

显性知识又称明晰知识、外显知识,是指能明确表达的知识,即人们可以通过口头传授、教科书、参考资料、期刊、专利文献、视听媒体、软件和数据库等方式获取,可以通过语言、书籍、文字、数据库等编码方式传播,也容易被人们学习的知识,包括可以写在书本和杂志上,能说出来的知识。

知识呈现维度分类见表 3。

表 3 知识呈现维度分类

	隐性知识	显性知识
隐性知识	社会化(意会知识)	外化(概念化知识)
显性知识	内化(操作化知识)	综合化(系统知识)

常见的有关键词关联,以知识内容中的关键词作为关联纽带,有相同关键词的知识进入关联体系中。知识发布时填写或自动提取关键词,知识库中与该关键词相同的知识会被提取出来,作为关键词关联结果呈现给用户参考。

通过对定量知识的分析,聚类出该知识中相关性较强的内容。举个例子,当我们在研究“南极”的知识内容中,可以聚类出“企鹅、气候变暖、冰层融化”等相关性很高的内容。

不同的知识表示对于知识的描述都不尽相同,通过概念之间的这些关系,可将存在于知识中的隐性知识挖掘出来,将隐性知识转化为显性知识。

通过知识单元的共现关系等呈现。

(4) 特征维度:知识所依附的载体类型,如学位论文、期刊论文、古籍等。

知识载体关联是指基于各类知识物质载体所形成的知识关联,如文献知识关联和非文献知识关联,文献知识关联中有文献学科关联、文献主题关联、引文关联等,这些知识关联都是基于文献外部特征和内容特征的关联,文献作为一个整体不仅包含着反映文献主题内容的内容特征如主题词、关键词、摘要等,而且包含着大量的外部特征,如分类号、作者、引文、标题、机构、期刊等。这些特征中隐藏着许多有意义的信息,通过分析文献内容的关联,可以发现该领域中的研究热点问题,而通过分析作者或机构的关联,可以发现该领域中的核心作者和核心机构。

通过文献载体类型进行聚类。

(5) 关系维度:从属关系、并列关系、交替关系、相关关系。

在关系维度中,类别、种属的关联是最为常见的一种关联方式,即属于同一个知识分类中的知识之间的关联。与该知识在同一分类中的其他知识,会被作为相关知识被呈现出来,供用户作为延伸阅读使用。

以一个问题为核心,将解决该问题的知识,层层推理出来。诊断关联是 kmpro 的独创关联模式,系统通过人工建立知识与知识之间的关联关系。这种关联以问题为核心,以解决问题的思路为延展,由一个问题关联多个解决方案,每个解决方案下面又可延展出相关知识,其相关模式类似于医生为患者选择诊断方案,这种推理诊断模式是精准、有效的“最

(3) 知识比较分析：根据词频，自动比较高频次主题词的发展脉络,并通过可视化的方法展示给用户。发展比较示例如图 6 所示。

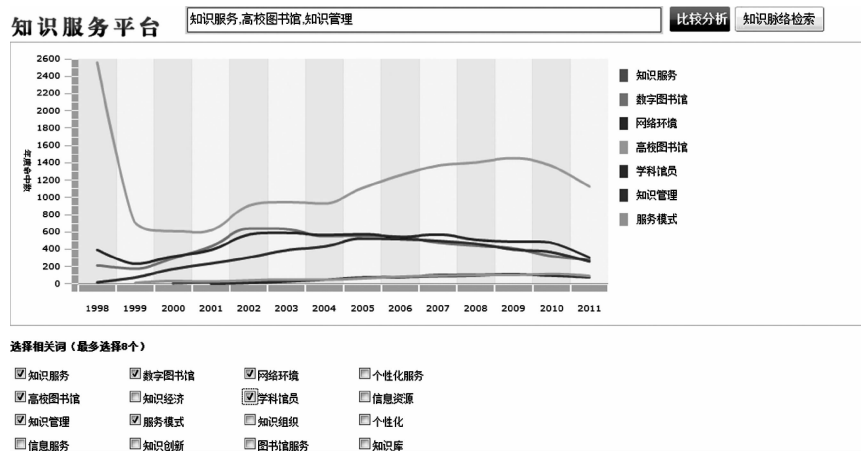


图 6 知识网络——发展比较示例

7.2 文献关联网

7.2.1 文献引证与被引证

文献引证关系如图 7 所示。

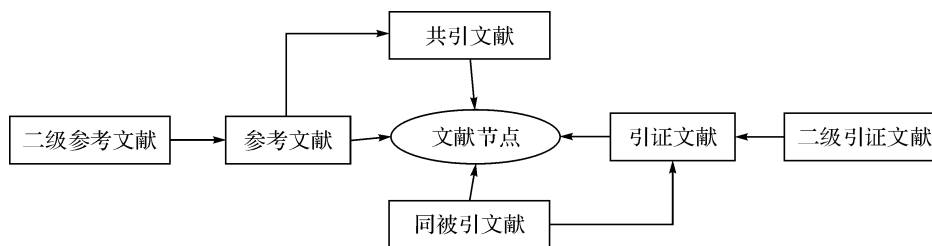


图 7 文献引证关系

7.2.2 同被关注文献

用户浏览、下载某文献时,与该文献同时被其他用户下载浏览过的其他文献。

7.2.3 合作者文献

与该作者合作过的,并可进一步呈现合作者的其他文献。

7.2.4 文献特征关联

(1) 相关基金文献:

根据基金名称与编号,自动聚类该基金的所有文献。

(2) 相关作者文献:

根据参考文献和引证文献情况,以及合作者信息,分析相关作者文献。

(3) 相关机构文献:

根据参考文献和引证文献情况,分析相关机构文献。

(4) 分类:

根据知识组织标引自动学科分类结果,构建该文献在此学科的底层分类,以使其从上层导航可以浏览更多相关领域的文献。

7.2.5 出版年

根据文献出版时间进行聚类。

8 用户知识空间

(1) 知识服务提供增值的过程中,要针对用户进行个性化服务,根据用户的行为特点,发掘新的知识点与知识关联,并为用户提供一个知识交互的社区交流空间。

(2) 用户根据自己的兴趣目标,进行相关的知识检索或浏览,并了解其他用户对相关知识的知识评价、服务评价及其讨论,对自己的知识需求进行个性化描述,获得初步的个性化知识需求表达。

(3) 根据个性化知识需求表达,获取相关知识(原子本体模块或本体模块),对其进行分解、合并,进行个性化知识组织,获得个性化知识服务。

(4) 根据个性化知识服务的结果,对个性化知识需求描述和组织作出相应的调整,如此反复,获得自己所需的个性化知识需求和知识组织方式,并建立用户兴趣模型(用户兴趣模块化本体)。用户兴趣模型随着用户的目标、兴趣的变化而发生动态变化,它的改进依赖于用户获得知识服务的历史和来自其他用户的评价,当用户的目标、兴趣发生变化,用户兴趣模型亦发生改变。

8.1 用户知识交互

用户可将自己的特色资源通过知识共享平台发布出来,同时也可以发布自己的知识需求信息,以寻找其他用户的帮助。

8.1.1 用户知识共享

用户可以发布自己的知识需求情况,其他用户可以针对其知识需求提供解答或服务。营造一个知识分享的平台,方便用户获取一些平台无法提供的资源。用户可以在知识获取中向图书馆提出资源需求,同时也可以选择是否在知识共享平台中进行发布。

读者推荐文献是利用互联网媒体的交互特性,由用户的行为自行编织的知识网络系统。该知识网络系统随着用户行为不断发生,系统自学习、自组织、动态更新,具备逐渐完善收敛的特性。这种自学习、自完善的系统,经过一段时间的收敛稳定,其内容关联客观准确,误差很小,反映了文献之间的一种客观关联性,便于读者参考。

8.1.2 参考咨询

通过参考咨询,实现了“用户—馆员”之间的即时沟通交流,用户登录后可以向在线的服务馆员发起即时会话请求,服务馆员通过网络服务来响应请求,并可以进行文件的即时传递。

8.2 用户知识管理

提供用户的个性化知识管理功能,包括知识获取、知识存储、个人笔记、论文写作、个人主页等,帮助用户管理和发布自己的资料,构建个人的知识管理空间。其中,知识获取服务包括参考咨询、知识推送、定题服务。

(1) 知识存储:对个人收藏的资料进行系统化、有序化的分类管理,支持对个人资料的标引、编辑、批注等功能。

(2) 知识检索:可对个人知识管理中存储的知识进行检索。

(3) 知识采集:随时将正在浏览的知识添加到个人知识存储默认的目录中。

(4) 个人笔记:随时记录自己知识学习的感悟和体会。

(5) 论文写作:提供在线论文写作功能,在论文写作的过程中,提供相关的文献信息服务。

8.2.1 个人知识仓储

个人知识仓储包括历史浏览、搜索记录、下载添加的文献、定制的信息保存、保存下载列表、群作者管理等。

8.2.2 个人知识定制

根据用户感兴趣的知识点,自动推送新的文献、引文跟踪等。根据用户需要,对相关功能的呈现进行定制。

(1) 定题服务:面向特定的用户,既可以是个体,也可以是团体(如课题组、科室)。从用户的特定需求出发,为用户提供针对性强的内容服务和解决方案。

(2) 知识推送:系统自动搜集用户感兴趣的信息,将采集的最新信息直接推送到用户指定的邮箱或个人空间中。

(3) 学科知识导航:根据用户使用历史记录,自动推动学科方面相关文献与知识点,并动态更新知识点的变化情况。

8.3 用户社区空间

与共同爱好者之间形成一个社交网络,了解相关学者的最新情况、领域最新研究进展,分享学术成果和心得,与其他学者交流协作,最终可在平台上与其他学者一起开展科研、教育和学习活动。

自动建立学者空间,与知识服务平台内容整合,揭示学者学术成果,可自由添加学者介绍。根据已发论文的主题词,自动生成学者关注点。

对学术成果根据知识服务平台内容进行聚类。

与学者外部社区(如豆瓣、人人网、科学网等)账户结合,搭建与其他学者交流的渠道。

参 考 文 献

[1] 李鹏. 基于知识管理的图书馆知识服务研究[D]. 太原: 山西大学硕士学位论文, 2010.

[2] 李霞. 知识服务平台构建的若干问题研究[D]. 沈阳: 东北大学博士学位论文, 2008.

[3] 李兴东. 面向知识服务的语义 Web 服务发现研究[D]. 合肥: 合肥工业大学硕士学位论文, 2010.