

基于发现系统的资源调度知识库研究

曾建勋

【摘要】[目的/意义]立足于发现系统中情景敏感的文献资源调度需求,探讨资源调度知识库框架及构建策略。[研究设计/方法]在分析发现系统的功能与实质、资源调度的缘起与内涵的基础上,进行资源调度知识库框架设计与构建策略研究。[结论/发现]基于发现系统的资源调度知识库应包含用户权限及偏好知识、文献资源特征及质量知识、文献资源服务方式及约束知识、情景信息处理知识四类知识;知识库构建中,需要在明晰资源调度知识库元数据框架基础上,强化与发现系统的协同,形成服务于用户的资源调度规则。[创新/价值]提出的资源调度知识库框架及构建策略,能够为发现系统的资源调度知识库建设提供参考。

【关键词】文献资源发现;发现系统;资源调度;知识库;情景敏感

【作者简介】曾建勋(ORCID:0000-0002-0432-9618),中国科学技术信息研究所,博士,研究馆员,研究方向:数字图书馆与数字出版、知识组织与知识链接,E-mail:zeng@istic.ac.cn(北京 100038)。

【原文出处】《图书情报知识》(武汉),2019.6.12~18

【基金项目】本文系国家社会科学基金重点项目“基于知识组织的图书馆资源发现服务体系研究”(17ATQ002)的研究成果之一。

随着搜索引擎的深入人心,将元搜索引擎与联合目录以及数字资源电子清单相结合而形成的文献资源发现系统,加速了文献资源的开放融合和协作共享。一方面突破了图书馆联盟间的馆际互借和文献资源发现系统的交换范围,使更多的学术出版机构通过协议或者开放获取方式开放数字资源,实现图书馆与出版商、数据商、发行商、馆配商等的广泛合作,逐渐形成超越馆际的跨界数据交换模式,另一方面资源来源途径的多元化和资源载体的多样化,使得用户有更大的资源选择权利和获取空间,不仅为用户提供资源多来源渠道列表,而且可以结合用户场景和需求,根据IP权限、地理位置、资源价格和质量等,基于情景敏感为用户优选相对贴切的资源。而这些都依赖于资源调度知识库的构建,才能实现在合适时间点、以合适形式、将合适文献请求传递给合适的文献服务方,提高文献资源交换共享效率。有鉴于此,本文试图在分析资源发现系统之资

源调度缘起和内涵基础上,分析情景敏感因素和调度知识,设计资源调度知识库框架结构,提出资源调度知识库构建策略。

1 发现系统功能与实质

资源发现系统作为一种基于文献元数据资源大规模集成整合而形成的新一代图书馆系统,自2009年由Serials Solutions推出Summon以来,迅速在国内外主要图书馆得到应用。它不仅继承了以往图书馆的书目数据,而且整合了出版商、数据库商、信息机构及网络上的各类异构学术资源,形成大规模的元数据仓储,并通过数据映射、数据转换、数据规范、数据归一,建立预索引集合,实现多源异构资源的一站式搜索;此外,它不仅依据用户场景对检索结果进行相关性排序,而且针对不同获取渠道实现情景敏感式优选;其云服务模式不仅与图书馆的用户认证系统进行挂接,而且与图书馆编目系统相沟通,与图书馆馆际互借或原文传递系统相衔接,实现与图书馆

相关系统在云端的集成融合^[1]。

1.1 多元异构元数据集成,彰显开放融合特征

发现系统的元数据可能来自上千家不同的单位,而很多单位采用的元数据标准各异,描述粒度层级不一,厚薄不均,因此需要按照统一的元数据标准进行字段映射、格式转换和逻辑验证,形成统一标准格式描述的资源信息,并从母体、卷期、篇级元数据和来源渠道四个方面对元数据进行查重、规范、归并和归一。通过不同类型资源的查重规则,对重复资源进行母体合并、编辑,规范不同卷期表述形式,对冲突母体的卷期信息以及篇级信息进行冲突合并,使薄数据合并成为厚数据,并对同一对象元数据的不同来源进行合并,针对一个元数据对象实体形成多来源渠道。这样,融化各自的数据孤岛,形成统一的、同构的元数据体系。

1.2 实现与编目/发布系统整合,突出跨界开放链接功能

图书馆都拥有自身的书目借阅系统、用户认证系统、文献传递系统等完整的图书馆业务流程系统,建有自身的图书馆门户网站及书目检索系统、电子资源导航系统等多种资源服务系统,但是无论是业务系统抑或是服务系统均未完全实现互联互通,不乏信息孤岛的存在,这在一定程度上造成业务流程重复和数据冗余^[2]。通过发现系统,基于云架构对印本、自建数据库、电子资源、开放资源等进行数据、流程和系统的统一整合与协同管理,将发现系统与图书馆自身的内部业务管理系统与服务系统,以及出版社发布系统等进行系统互联、流程整合,形成统一的工程化管理模式和一体化服务入口,从而提升管理效率,优化服务效率。

1.3 多途径/渠道信息服务聚合,实现情景敏感的资源调度

任何一个单位都无法完全收集收藏所有文献资源,各图书馆文献信息资源具有互补性,在数字环境下很多出版资源更是存储在出版社或数据商的云平台中,机构知识库则存储于科研机构服务器上。发现系统通过定期更新图书馆或科研单位的馆藏资源信息(例如印本MARC数据、数据库订购信息、特色资源列表),定期获取出版社资源信息,一方面通过感

知用户订购权益提供全文获取途径,与出版社电子资源、开放资源相链接,另一方面对馆藏资源实现资源调度系统与图书馆原文传递服务系统对接,对用户检索请求进行文献传递、代查代借服务回应,来提高全文获取的保障率。以此将资源与服务共享层级延伸至相关机构,整合相关图书文献机构的服务能力。

1.4 语义知识关联融合,实现个性化信息推荐

资源发现系统不仅可以强化文献获取服务的针对性,而且可以通过科研实体抽取、引文网络耦合、规范文档建立、概念主题识别、资源解析关联等操作,利用元数据的智能化标引以及实体概念的语义关系组织,形成“科研人员—科研机构—学科主题”的各类知识关联网络,推进知识服务和数据服务。同时,可以通过对用户不同时期的检索或获取行为的感知,或对不同用户对同一资源的利用行为感知等,实现资源对不同用户或同一用户对同类资源的感知推荐,形成多层次服务功能的整合。

2 资源调度的缘起和内涵

2.1 资源调度的缘起

联合目录是揭示、报道多个文献收藏单位所藏文献的目录,通常由若干文献收藏单位遵循统一的著录项目和标准合作编制而成。一方面用户可以借此了解文献来源渠道,获得馆藏指引服务,另一方面可以揭示系统、地区,甚至全国书刊品种的分布情况,为馆际互借、交换和调整创造条件,提升图书馆资源使用和馆藏效益。

联合目录自13世纪的《英格兰图书馆登记册》创建以来不断发展。但是随着数字出版的快速发展,伴随学术搜索引擎的加速普及,各文献收藏单位OPAC或出版社电子资源目录正在融入时下的资源发现系统,形成馆藏信息在统一元数据索引上的全面集成,依据馆藏机构资源变化动态地为用户呈现恰当的资源地址和所能获取的服务。这样,资源共建形式从单一的印刷型文献资源发展为数字化电子资源;共知手段由书本式联合目录发展为联机联合目录直至资源发现系统;共享方式也从传统的馆际互借、文献传递发展为网络远程传递。联合目录从图书馆单方面编制,已经发展成与出版商、数据商等达成更广泛的统一战线,构建新的集成化联合目录。

2007年OCLC将书目数据与FirstSearch平台上的四个文摘数据库的文章级元数据在WorldCat.org进行融合,从2014年推出基于云的应用程序WorldCat发现服务,资源整合粒度从书目数据深入到篇章级元数据,在此基础上基于用户地理位置将来源图书馆进行排序,并为用户推荐在同一城市的图书馆,实现地理位置敏感^[3]。中国科学院联合目录整合了院内400余家成员馆所订购的印本及电子资源,针对不同用户订购资源的多样性实现了基于用户IP的权限敏感。清华大学引进Primo构成的“水木搜索”整合清华大学的馆藏印本资源与购买的数据库资源,通过检测用户IP,实现IP位置敏感,让用户能够直接下载订购电子资源,并通过SFX开放链接服务为用户提供了书籍资源的网上书店购买链接和豆瓣读书讨论地址^[4]。百度学术搜索系统整合了各数据库和网络上的免费电子资源,按照开放程度区分、标注了获取该资源所需要的权限,并提供了文献互助平台,将文献传递服务的提供方从图书馆转变至所有用户,以众包的方式提供文献传递服务。电子资源的产生与飞速增长拓展了用户获取文献资源的途径,开放链接技术使得资源的提供方和文献元数据可以相互分离,资源提供者可以是数据库商,也可以是图书馆等机构。

因此,文献资源调度思想源于联合目录系统,不仅在书目层面,而且在篇章层面,不仅在图书馆间,而且跨界连通出版商、数据商和网络开放平台,实现文献信息的开放链接与共享;将图书馆丰富的书目数据资源和出版社即时的出版元数据资源整合集成起来,在弘扬原有联合目录体系的基础上,构建集印本资源与电子资源、订购资源与开放资源于一体的统一书目集成平台。在创新原有目录索引方法的基础上,形成包含各类资源目录及其用户情景信息的资源调度知识库,并依据用户场景进行情景敏感式调度服务指令,展示某一资源对象的多渠道提供或多途径获取来源^[5],继而向用户提供更适合的资源及其获取途径,提高文献资源从查询、订购到获取的一站式服务效率,推进文献资源的聚合和共享。

2.2 基于情景敏感的资源调度内涵

发现系统集成融合了多方面的元数据资源,既

有物理馆藏资源,也有电子资源;既有文本资源,也有声像资源、可视化资源,以及数值型资源或科学数据资源,集不同类型、版本、媒介和颗粒度资源于一体;往往一个元数据资源对象涉及多种媒介,有多种来源或多个获取途径,同时这个元数据资源对象可以从文本型资源关联到其他如科学数据、声像类型的资源,所以,从资源来源形式上既有图书馆馆藏、也有图书馆订购出版社电子资源,还有网上OA资源,以及出版社单篇或会员订购资源等;从资源提供方式上,既可以进行馆藏指引提供代借代查、原文传递、网上书店链接,又可以提供文献链接直接下载或付费下载等方式。在资源调度过程中,要综合考虑多种因素,例如印本资源的地理位置、电子资源订购权限、用户身份权限、用户对资源质量的要求、用户能够承受的资源价格、资源被用户利用情况、资源提供的便捷性、资源提供方的信誉等^[6]。资源调度就是通过主动感知用户实时需求,根据用户情景信息及其变化,与相关资源和服务进行匹配,建立一种与资源和服务相适应的调度机制,并将情景信息匹配引入推荐系统中,动态地计算用户所需文献与其提供途径的优先调度指令,提高资源获取的准确性和服务提供的可靠性,为用户提供一站式泛在服务。

因此,资源调度是一种主动、自适应、智能性的实时服务,其内涵是在馆藏信息统一元数据索引集成基础上,依据馆藏或出版机构资源情况,自动感知实时情景信息,智能判断用户信息行为及需求,识别用户的位置、类型、属性、特征、偏好及其对资源与服务的使用权限,进行需求与资源、服务的情景化匹配,为用户推荐或调度恰当的资源地址和所能获取的资源服务,实现在不同情境下的个性化服务,继而优化用户一站式体验^[7],提升用户获取的针对性,实现模型如图1所示。

3 资源调度知识库框架

资源发现系统集成多类型、多渠道、多载体、多层级和多途径的文献资源,不仅包括图书馆馆藏资源,还包括出版商、数据库商、发行商和互联网等资源,实现印本与电子资源的同步揭示。针对这些分布式文献资源,为实现情景敏感的资源调度,除了需要获取用户、资源、服务的实时情景信息,设计调度

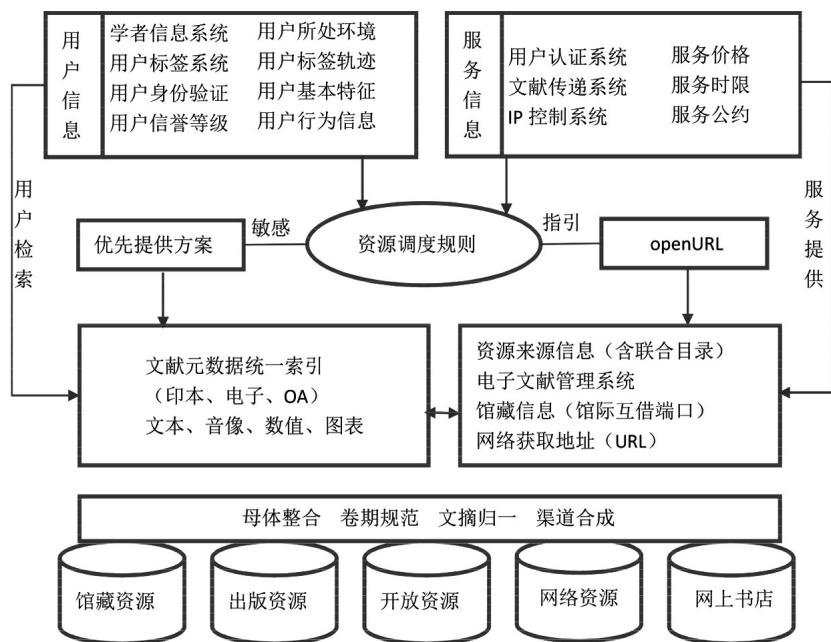


图1 基于情景敏感的资源调度

算法外,还需要用户权限及偏好知识、资源对象的特征及质量知识、资源服务及约束知识、情景信息处理知识作为支撑^[8]。因此,资源调度知识库是用户、资源和服务等不同要素及其相互关系的信息集合。其中,用户权限及偏好知识可以通过用户管理系统提取,资源特征及质量知识可以通过元数据基础系统提取,服务方式及约束知识可以通过各类服务公约/协议提取,情景处理知识来源于情景感知系统,这些知识最终通过基于情景敏感的资源调度规则发生作用,故而资源调度知识库基本框架如图2所示。

3.1 用户权限及偏好知识

用户权限及偏好知识存储的是与文献资源调度相关的用户知识。用户权限知识用于明确界定用户可以在什么条件下对哪些对象进行哪些操作,具体而言主要包括:①作用对象,可以是服务方、数据库、资源母体甚至是单个资源;②操作权限,例如可在线浏览、获取全文、批量获取等;③权限约束,例如权限的时间约束、单日可获取文献数量等。用户偏好知识则是在符合权限约束的前提下,用户希望优先对哪些对象行使什么权利^[9]。偏好涉及的要素主要包

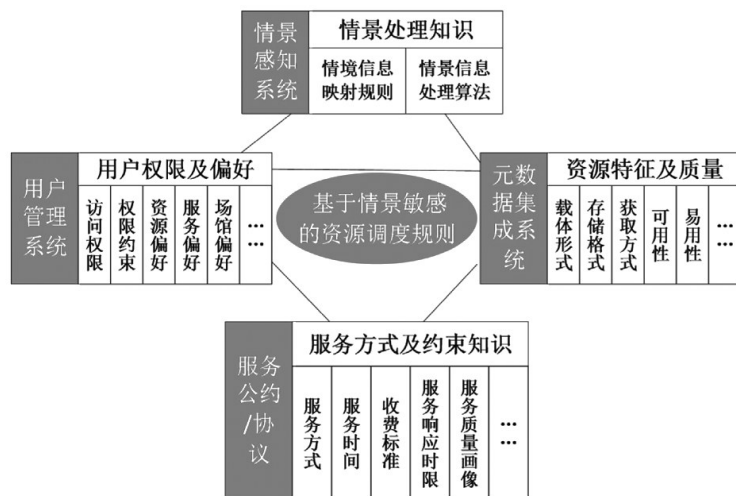


图2 资源调度知识库框架

括时间、位置、距离、语言、资源特征、资源质量、服务方、服务质量等方面。此外,用户的偏好知识既可能只涉及一个要素,也可能涉及多要素组合。

值得指出的是,此处的用户既包括个体用户,也包括机构用户;既包括可以通过IP地址等标识的匿名用户,也包括可以通过ID等信息精确定位的用户;而且不同类型用户间还存在交叉关系,例如一个通过ID访问的用户,其IP刚好处于某机构用户的IP段内。基于此,在用户权限及偏好知识相关的元数据体系构建中,需要区分不同类型用户并进行差异化处理;同时,还需要针对几类不同用户的权限、偏好出现冲突的情况,建立冲突处理规则。

3.2 文献资源特征及质量知识

发现系统的资源调度对象是一个一个的文献资源,既可以是存储在某服务机构服务器上的数字副本,也可以是纸质副本、缩微胶卷、光盘磁盘等。基于此,资源调度知识库中存储的也是与用户资源获取偏好相关的、关于某个具体物理资源的特征及质量的知识信息。文献资源特征方面需要明确载体类型,在此基础上,对于数字化资源,需要进一步记录其存储格式、资源大小、获取方式、资源价格等特征;对于非数字化资源,则需要进一步记录其存储位置、副本数量、资源状态等信息^[10]。文献资源质量方面则重点关注资源的完整性、可用性、易用性等信息。实践中,为衡量资源的质量信息,还需要对不同载体形式的文献有针对性地设计指标。

3.3 文献资源服务及约束知识

不同的文献资源服务机构在业务组织上具有不同的约束条件及服务方式,这些也将对用户的偏好产生影响,进而影响文献资源的调度,因此也应将其纳入资源调度知识库中。在服务约束方面,主要包括服务时间(例如周末及节假日是否服务、每天的具體服务时间等)、资源获取速度和总量控制条件、网络端口、服务认证方式、服务响应时限、收费标准、知识产权保护策略等。服务质量方面,需要获取的知识信息包括服务的稳定性、便捷性、易用性、经济性、安全性、可信性等。服务效果评价中,需要区分数字化与非数字化资源,分别建立服务评价指标体系与打分标准,使得不同细分类型的数字化服务之间、非

数字化资源服务之间具有可比性。

3.4 情景信息处理知识

情景敏感的资源调度中,发现系统采集到的实时情景信息不能直接作为调度算法的输入信息,需要先进行预处理,转换成算法可用的知识信息。因此,虽然用户情景信息具有动态性,无法直接存储于资源调度知识库中,但为便于情景信息的处理与利用,需要在资源调度知识库中存储情景信息处理知识。这些知识主要包括两类,一类是情景信息映射规则,经这些规则处理后,原始情景信息可以转变成可枚举的少量取值,例如根据用户的IP信息,可以按网络端口映射为中国联通、中国电信和中国移动等,如果其处于未登录状态,则从权限方面可以将其映射为匿名机构用户(可对应到具体机构)、普通匿名用户;另一类是情景信息处理算法,经这些算法处理后,原始情景信息可转换成连续数值,例如用户的位置坐标信息,经过处理后可以转换成与各个图书情报服务场馆的距离信息^[11]。

除上述信息外,资源调度知识库还存储一些用于链接不同数据表或外部系统的关联信息或接口信息,包括与一些外部系统合作完成的从用户登录到用户获取资源的完整流程信息,以及链接用户与用户行为、用户所选资源的字段信息等。

4 基于发现系统的资源调度知识库构建策略

资源调度知识库是发现系统进行情景敏感的资源调度,保障用户体验的基础,同时其建设也依赖于发现系统,一方面资源调度知识库中的部分知识直接来源于发现系统,例如用户权限、文献特征、服务约束知识等;另一方面资源调度知识库中的部分知识需要建立在发现系统提供的数据库基础上,例如基于用户行为数据进行用户偏好知识的提取等。因此,在资源调度知识库构建中,需要在明晰资源调度知识库元数据框架基础上,强化与发现系统的协同,形成服务于用户的资源调度规则^[12]。

4.1 探索资源调度的形式化表达,明晰知识库元数据框架

实现资源调度知识库建设需要将相关情景信息以机器可处理的形式定义、表示和存储。对于用户情景信息,主要表示用户的基本信息和用户在具体

的时间、地点、使用具体网络与应用系统时的行为,其行为可以是浏览、标注、检索、下载等,而这些行为的表示则需要转换成相应的行为参数,即用标签、检索词、下载内容等具体概念和数字表示。用户偏好是用户在特定时期一种相对稳定的兴趣,也需要转换成具体的本体概念及其权重来表示。这样,通过概念属性与关系的描述可以更好地表达用户的情景化需求。资源情景模型主要表示该信息资源的基本情况,例如图书借阅记录、专利文献下载记录、期刊论文引用记录、读书评论记录等,而每条情景记录下又分为更细的粒度,例如某篇论文被谁、在何时、通过何种方式下载以及下载次数等,也需要进行形式化表示以更详细地描述信息资源。在资源调度时,可以选择最适合用户情景的信息资源,实现用户情景与信息资源的情景化匹配,实现最合适的资源调度。

以往的联合目录都有其具体的元数据框架及其各类元素的标准规范,同样对于资源调度知识库也需要构建其元数据框架,既有用户、资源和服务情景信息的业务元素,也有管理元数据、结构元数据、关联元数据等不同功能性元素,并且要形成业界能够达成共识的元数据标准,以便于相关元数据的互操作和交换共享。

4.2 构筑元数据跨界合作机制,形成嵌入服务的知识库构建流程

资源发现系统的元数据有三种来源:一是内容提供商签约直接获得授权的元数据,二是采用元数据收割的方式,收集网络开放资源以及未授权的第三方元数据,三是本地馆藏资源的元数据和电子资源清单上载,需在系统部署实施时对本地资源的数据进行映射和上载,并定期更新。发现系统中的元数据及其资源获取途径的收集、整理和关联,是形成资源调度知识库的重要来源之一,同时在资源被用户检索发现之后,在联合服务之中的用户行为和需求等情景信息的动态获取,也是知识库相关用户和服务情景信息的主要来源渠道。所以,发现系统中元数据获取采集方式、用户管理方式和跨界协同服务机制等,直接关系到资源调度知识库的构建,需要从用户注册、访问、检索、点击、获取等行为过程中动态采集情景数据,将知识库的构建嵌入发现系统元

数据获取、用户注册与日志管理、用户网络行为和服务履行等过程之中,动态地收集和获取资源调度知识库构建所需情景信息^[13]。

4.3 多途径获取情景信息,强化信息清洗和场景分析

用于资源调度知识库建设的情景信息种类繁多,不仅包括物理情景,还包括计算情景、网络情景、用户社会关系情景等,因此情景组成复杂多样、异质化明显。同时,情景信息分布广泛,范畴宽泛,获取方式多种多样,既可通过服务平台或交互日志捕获,又可以从已有信息中抽取、由用户直接注册或选择。一些如位置、时间、网络链接等原始情景数据需要通过物理方式获取,而习惯、偏好等需要通过逻辑方式获取。由于用户行为受所处环境影响、资源也在不断更新,因此,资源调度知识库需要针对用户行为状态及资源情景内容的变化多途径动态获取各种情景信息。

由于情景信息来源各异、动态生成、表征对象不同,所以通常是模糊、不精确、不稳定甚至有冲突、不一致的,所以,构建资源调度知识库需要对初始信息进行清洗、过滤、转换、关联和融合。既要将其中的模糊、不准确、不稳定、矛盾和错误的信息清洗过滤掉,又要将不同来源、不同格式的情景数据用统一的格式进行表示,还要根据用户和服务场景,对获取的情景和已有的情景库进行统计关联分析推理出复合情景,才能得到资源调度所需的情景知识。

4.4 凝练情景敏感算法,形成切合用户需求的资源调度规则

对于电子资源,既可能有免费OA资源或用户已经开通的数据库资源,还有可能是需要用户实时购买的电子资源,这样用户既可以实时下载OA资源或开通订购资源,也可以前往出版社网站即时购买;对于物理资源,既可能是某图书馆馆藏资源,也可能是网上书店销售图书,这样用户既可以自行前往借阅或请求原文传递或代借代查,又可以在线购买。所以,在资源调度过程中,首先要识别登录用户的IP或ID,来判断用户所在机构和身份类型,明确其能够利用的资源及访问权限,进而结合资源类型、资源开通权限或价格、获取距离和提供方式等实现资源优先

排序。这里,要根据用户身份、偏好与所处情景,凝练出情景敏感算法,在感知用户对资源的访问权限、质量要求、价格承受度、距离远近、服务信任度等的同时,根据用户登录IP与机构地址计算用户电子资源访问权限或与收藏机构间距离,结合资源本身特征属性、服务规则和合约,来制订切合用户的资源调度规则,继而对资源的不同来源和途径进行优选和排序,形成适合于用户的各种资源调度优先排序指令,将用户引导至相应的资源来源或服务。例如形成权限敏感规则用于判断用户对于该资源访问的具体权限,资费敏感规则表示资源价格以及用户对价格的在意程度,距离敏感规则表示资源与用户的距离以及用户对距离的在意程度,质量敏感规则表示资源质量以及用户对质量的在意程度,以及服务敏感规则表示用户喜爱的服务类型。在知识库积累一定量的各类型用户于各种不同情境下的决策行为后,采用机器学习方法丰富和凝练调度规则。

参考文献:

- [1]王卫军.高校图书馆情景敏感服务及其情景模型构建[J].情报理论与实践,2015,38(2):88-91.
- [2]袁静.图书馆情景感知自适应个性化服务平台构建[J].图书馆学研究,2012(17):37-40.
- [3]苏建华.图书馆选择资源发现系统的策略分析——以资源发现系统与学术搜索引擎的较为视角[J].情报科学,2015(6):91-94,105.
- [4]袁静.情景感知自适应:图书馆个性化服务新方向[J].图书情报工作,2012,56(7):79-82,97.
- [5]包凌,蒋颖.图书馆统一资源发现系统的比较研究[J].情报资料工作,2012,33(5):68-73.
- [6]叶亚娜,肖婵,耿骞.数字资源的集成检索与资源调度[J].中国教育网络,2007(11):66-67.
- [7]曾建勋,丁道劲.基于语义的国家科技信息发现服务体系研究[J].中国图书馆学报,2017(4):51-62.
- [8]曾建勋.资源发现系统的颠覆性[J].数字图书馆论坛,2016(2):1.
- [9]丁道劲,曾建勋.文献元数据集成管理研究[J].情报学报,2019(6):568-577.
- [10]曾建勋.从联合目录到资源调度知识库[J].数字图书馆论坛,2017(5):1-1.
- [11]Shiri A, Oliphant T. Temporal Patterns of Searching in a Public Library Discovery System[J]. Canadian Journal of Information and Library Science, 2017, 41(1): 1-17.
- [12]Gallaway T O, Hines M F. Competitive Usability and the Catalogue: A Process for Justification and Selection of a Next-Generation Catalogue or Web-Scale Discovery System[J]. Library Trends, 2012, 61(1): 173-185.
- [13]李勃慧.面向发现服务的图书馆资源调度知识库结构研究[D].北京:中国科学技术信息研究所,2018.

Research on Resource Scheduling Knowledge Base Built upon Discovery System

Zeng Jiangxun

Abstract: [Purpose/Significance]Based on the requirements of context-sensitive literature resource scheduling in discovery system, this paper discussed the framework and construction strategy of resource scheduling knowledge base.[Design/Methodology]After analyzing the functions and essence of discovery system, and the origin and connotation of resource scheduling, this paper studied the framework and construction strategy of resource scheduling knowledge base.[Findings/Conclusion]The resource scheduling knowledge base built upon discovery system should contain four categories of knowledge, which are users' rights and preference, characteristics and quality of literature resource, modes and constraints of literature resource services, and contextualized information processing. In the process of knowledge base construction, it is necessary to strengthen the collaboration with discovery system and form resource scheduling rules to serve users based on clear metadata framework of resource scheduling knowledge base.[Originality/Value]The framework and strategy proposed in this study can provide a reference for the construction of resource scheduling knowledge base of discovery system.

Key words: Literature resource discovery; Discovery system; Resource scheduling; Knowledge base; Context sensitive