

面向档案的知识聚合服务研究现状及趋势展望*

魏扣¹ 李子林² 郝琦¹

(1. 中国人民大学劳动人事学院, 北京 100872; 2. 中国人民大学信息资源管理学院, 北京 100872)

摘要: 立足档案知识聚合服务的核心概念, 从档案信息资源整合、档案知识服务、档案利用服务三大方面梳理档案知识聚合服务研究现状。在此基础上提出档案知识聚合模式研究, 档案知识聚合服务技术研究, 社交媒体环境下档案知识聚合服务研究, 图、情、档协同参与档案知识聚合服务研究四大方向, 为我国档案领域深入开展档案知识聚合服务研究提供参考和指导。

关键词: 档案; 知识聚合; 知识服务

中图分类号: G270

DOI: 10.3772/j.issn.1673-2286.2018.03.004

知识经济时代, 知识聚合作为知识处理的一种重要理念和技术被广泛应用于信息服务和知识服务领域。档案资源作为公众获取信息的一种重要资源形态, 凭借其原始记录性、凭证性和权威性特点, 逐渐成为社会公众认可的重要知识资源。在此背景下, 深化档案资源的知识性开发, 实现档案知识的科学、合理、广泛传递, 充分发挥档案的知识价值, 满足社会公众的决策知识需求和文化知识需求, 成为档案领域亟需关注和重视的问题。目前, 国内图书情报领域知识聚合的模式、方法、技术等已有深入的积累, 也产生了一些典型实践应用案例。这些研究成果为档案学界开展档案知识聚合研究提供参考和借鉴。同时, 档案学者需立足档案学视角, 结合档案知识自身特点, 选取恰当的聚合模式和聚合方法, 在相关信息技术的支撑下深入开展档案知识聚合的理论研究和实践应用。

1 核心概念界定

1.1 知识聚合

知识聚合是在社会信息量剧增、用户信息需求涌现

及新兴信息技术发展应用等因素共同作用的时代背景下产生的新概念。知识经济时代, 人们逐渐意识到知识的重要性并将其看作组织机构的重要资产, 知识聚合研究逐渐引起国内外学者的关注。贯君等^[1]提出知识聚合指在统计分析、数据挖掘、人工智能等方法基础上, 对可能存在隐性关联的知识单元进行凝聚的新型信息技术, 以提出知识单元的内在关联为手段, 目的是构建多维度、多层次、多关联的知识体系。目前, 国内多数学者认为知识聚合是通过对文献、信息、数据等资源进行知识化处理, 抽取其中的知识单元, 结合用户需求, 并将这些知识单元按照一定关联重新聚集和组合, 形成满足用户需求的高聚合度、强关联性的新知识集合。换言之, 知识聚合是从用户需求和解决用户实际问题出发的一种精细化、细粒度层次的知识单元聚合活动。

知识聚合包含知识聚合主体、知识聚合客体、知识聚合流程及知识聚合模式和技术。知识聚合主体是知识聚合过程中的主要能动因素, 是知识聚合流程的发起者和知识聚合结果的发布者, 牵引和指导整个知识聚合过程。知识聚合客体是知识聚合的“原材料”, 指需要被聚合的零散知识单元。知识聚合是面向应用和服务的一项实践工作, 其实施离不开聚合流程、聚合模式和

* 本研究得到国家社会科学基金青年项目“社交媒体环境下公共档案资源知识聚合与服务研究”(编号: 16CTQ032)资助。

技术的支持。知识聚合流程包括知识单元处理、知识单元间的关联挖掘、知识单元聚合、知识聚合结果发布、知识聚合效果评估等环节。在知识聚合流程中,针对不同聚合要求选择相应的聚合模式,同时要综合应用知识工程、本体理论、关联语义等知识聚合技术。

1.2 档案知识聚合

档案知识聚合目前在国内档案学界还未被广泛地提出和应用。通过对“档案知识”“知识聚合”的分析,本文将档案知识聚合定义为通过对数字档案资源进行相应的知识抽取和知识表示,形成档案知识单元并存入档案知识库,按照档案用户需求,充分挖掘知识单元间的关联,对其重新聚集和组合,形成用户感兴趣的、能够解决实际问题的、具有较强知识性的档案知识集合。

档案知识聚合由聚合主体、聚合客体、聚合环境构成。档案知识聚合主体主要是档案工作者(人工智能技术的发展与应用可能会实现无人工干预的智能化档案知识自聚合模式)。档案知识聚合客体是指档案知识,即立足档案内容,经过挖掘、处理、组织形成的计算机可实施聚合操作的知识单元。档案知识聚合环境主要包括宏观的社会环境和微观的实施环境。宏观的社会环境即社会政治、经济、文化、教育等方面的发展情况,以及此背景下的社会档案意识、社会公众文化素养等。微观的实施环境即理论环境、硬件环境、技术环境和需求环境,理论环境指档案知识聚合、档案知识管理、服务等相关理论的发展与应用情况;硬件环境指档案知识聚合实施所需的计算机、网络、服务器等基础实施配备情况;技术环境指档案知识聚合实施所需知识挖掘技术、聚类技术等的应用情况;需求环境指档案用户对档案知识及档案工作者的要求和期望,这是影响档案知识聚合效果的重要因素。

2 我国档案知识聚合服务研究现状

图书情报界在知识聚合的理论研究、技术应用和实践推广方面已形成一定的研究成果,但档案学界对档案知识聚合的研究较为鲜见,主要集中在以下三个方面。

2.1 档案信息资源整合研究

虽然国内档案学者对档案知识聚合的相关研究甚

少,但档案信息资源整合研究自2003年出现,一直是档案学界的研究热点。档案信息资源整合指根据一定需求,将分散的档案实体、档案信息及相关的档案管理要素进行融合、重组,形成能够反映一定主题的、高效率的档案资源新系统^[2]。档案信息资源整合的思想和档案知识聚合的思想在本质上存在共同点,两者关键的区别在于档案信息资源整合是一种较为粗粒度的整合方式,整合对象为档案信息;而档案知识聚合是基于知识语义的深层次、细粒度的聚合方式,聚合对象为档案知识。因此,档案信息资源整合研究可为档案知识聚合服务研究的开展提供理论、方法和技术上的参考。

国内档案界对档案信息资源整合的研究主要集中在3个方面。①档案信息资源整合内容研究。蒋冠^[3]认为,网络环境下档案信息资源的整合超越档案部门的界限,实现档案部门与社会信息服务机构对档案信息的整合,包括系统的整合、信息的整合、管理的整合和服务的整合。杨红仙^[4]指出,档案信息资源整合包含档案实体整合、信息整合、技术标准整合、系统整合、人才结构整合5个方面。②档案信息资源整合模式研究。陈永生^[5]、滕霞^[6]认为虚拟档案信息资源整合模式是借助信息技术和系统平台实现群体层次的档案信息整合。③档案信息资源整合技术及标准研究。刘明^[7]认为,网络档案信息资源整合的技术保障涉及数据的组织、转换技术,数据存储与压缩技术,分类、索引与检索技术,用户界面与系统集成技术等。梁作华^[8]提出,国家应出台包括档案数据库结构、信息存贮和著录格式、软硬件配置、网络体系结构、信息处理界面等完整的档案信息资源整合标准体系。

2.2 档案知识服务研究

知识经济时代,知识管理、知识服务理念的发展与应用为档案服务理论与实践研究带来新的思路和方向。档案知识服务指在信息技术和人工智能等技术支撑下,将知识管理理念引入档案部门,对档案资源进行一系列加工处理后形成知识资源,根据用户需求,针对不同问题提供相应解决方案^[9]。档案知识聚合服务是档案知识服务发展的新阶段,是更高质量、高价值、高难度的档案知识服务。档案知识服务研究为档案知识聚合服务夯实前期研究基础。

国内档案界对档案知识服务的研究主要集中在3个方面。①档案知识服务模型研究。张斌等^[10]提出,从档

案知识提供与档案知识检索、档案知识定制与交流两方面构建基于档案知识库的档案知识服务模型。毕建新等^[11]通过对档案信息集成、知识采集、知识组织、知识存储、用户建模、知识通道等要素进行规划、组合来构建用户目标驱动的档案知识服务模型。②档案知识服务方式研究。孙艳丽^[12]提出,以平台的方式将档案知识资源系统、应用服务系统、网络服务系统和信息用户连接,实现档案知识服务的提供与开展。③档案知识服务技术研究。裘丽^[13]认为,语义网技术、智能Agent技术是实现档案知识服务的核心技术,档案部门应以多元视角,结合档案知识服务的实际情况选择科学、合理、实用的知识服务技术。

2.3 档案利用服务研究

档案利用是档案管理的最终目的,也是档案工作价值实现的关键路径。档案知识聚合服务是档案利用服务的重要组成部分,其在服务观念、服务原则、服务方式和服务内容上始终保持与档案利用服务的一致性。

立足档案知识聚合服务主题,总结归纳出与其相关的档案利用服务研究。①档案利用服务观念研究。冯惠玲^[14]提出,档案馆要从“官本位”思想转变为“民本位”思想,要将“国家档案观”转变为“社会档案观”,并采取具有亲和力的服务方式以便公民对档案资源的利用。张斌等^[15]认为我国“社会档案观”已经逐步形成,要在馆藏资源多元化的基础上发展档案服务社会化,通过创新服务机制,完善服务制度,丰富利用服务方式来拉近档案与民众的距离。②档案利用服务方式研究。黄霄羽等^[16]提出运用社交媒体工具创新档案利用服务方式,借助社交媒体的优势拓展档案利用服务的范围。韩云惠^[17]提出运用APP实现个性化、社会化、多元化的档案利用服务,细化档案利用的具体形式,包括智能档案地理信息服务、3D档案展览、即时反馈档案资源等。③档案利用服务技术创新研究。郑光辉^[18]提出借助云计算技术构建档案云利用系统,通过提供虚拟计算、虚拟存储及虚拟网络等功能支撑档案利用系统的运行,以门户网站形式向全国档案利用者提供档案信息资源共享和利用服务。向泽红^[19]提出将云计算技术、3D虚拟技术应用于智能化档案信息服务中,实现档案信息可视化检索,降低档案服务成本,提高档案信息利用的便捷性。

3 面向档案的知识聚合服务研究方向

通过对我国档案知识聚合服务研究情况进行梳理,发现当前我国档案界关于档案知识聚合服务的研究深度不足,缺少基于档案知识特点的知识聚合模式研究,对档案知识聚合服务方法、技术的研究大多停留在可行性分析和探索性引入阶段,无法实现新技术和信息环境下档案知识聚合服务的创新发展。鉴于此,本文提出档案知识聚合服务研究的下一阶段重点可以从以下四个方面展开。

3.1 档案知识聚合模式研究

档案知识聚合模式是档案知识聚合服务实现的关键。当前,国内档案学者对档案知识聚合模式的研究还处于空白阶段,相关研究成果主要集中于档案信息资源整合模式、档案知识服务模式、档案利用服务方式方面。一方面,档案信息资源整合模式研究的对象为文献级别的档案信息资源,这种整合模式还停留在档案资源聚合的初级阶段,即对传统粗粒度档案资源的聚合。然而,针对档案资源知识单元的聚合,即对深度聚合模式的研究有待推进和深入。另一方面,档案知识服务模式、档案利用服务方式的研究虽创新性地提出用户需求导向、知识库建设、运用社交媒体工具等发展理念和实施路径,但仍与档案知识聚合模式研究存在较大差距。如何挖掘档案资源中的知识,立足档案用户需求对档案知识单元进行语义关联和聚合,形成用户所需的聚合型、关联型档案知识集合是未来研究的方向。

当前,知识聚合模式主要包括基于计量的知识聚合、基于语义的知识聚合、基于情景的知识聚合3类。档案知识聚合模式研究可在借鉴现有图书情报等领域知识聚合模式研究成果的基础上,结合档案资源整合、档案知识服务、档案利用服务的理念与要求,从优势、劣势、应用场景的维度分析,选择与档案知识聚合对象、流程、用户需求相适应的档案知识聚合模式。

3.2 档案知识聚合服务技术研究

目前国内档案学者对档案知识服务涉及的知识抽取、知识组织、知识存储和检索技术的研究仍停留在介绍性引入和理论性探讨阶段。同时,在档案利用服务技术研究方面,虽有学者提出引入云计算技术、3D虚拟技

术以实现档案信息服务的场景化、便捷化,但具体的、可操作的技术应用和实施路径研究缺乏。鉴于此,可供档案知识聚合服务技术研究借鉴的前期成果不多,档案知识聚合服务技术研究的开展成为档案领域面临的艰巨却必须克服的疑难问题。

档案知识聚合服务主要由档案知识聚合、知识聚合服务两部分组成,其实现需要一系列相关技术的支撑与运用。档案知识聚合是从微观视角分析如何挖掘、处理、聚合档案知识,因此,在一定程度上是对档案知识服务中知识处理技术的扩展,即从语义关联的角度,根据用户的需求,对档案资源进行基于知识单元的细粒度、深层次聚合,形成统一的档案知识集合。该过程的实施涉及对知识抽取技术、知识表示技术、知识组织技术和知识库建构技术的应用。以知识抽取为例,采用词典标引法、切分标记标引法、单汉字标引法、统计标引法、语义分析标引法或人工智能标引法,对结构化(主题词表、数据表)、半结构化(Web信息)、非结构化文件(新闻报道、科技文献、多媒体文件)等进行内容和形式的多重分析处理,把知识点(知识元)逐条挖掘出来。知识聚合服务是通过构建知识服务平台来实现的。用户通过访问服务平台的用户界面,向档案部门提出知识服务需求,档案部门根据目前用户需求为用户提供相关主题的知识集合或根据用户以往需求特征通过服务界面向用户主动推荐相关知识。该过程的实施涉及知识检索技术、语义识别技术、智能推理技术、用户界面设计技术、数据可视化技术和平台安全维护技术的应用。

3.3 社交媒体环境下的档案知识聚合服务研究

在社交媒体时代,档案部门应积极借助社交媒体平台,开展档案利用服务。档案知识服务是档案利用服务与知识服务的结合,是发挥档案知识价值的重要手段。档案知识聚合服务是对档案知识服务的发展和完善,是档案知识服务未来的重点研究方向。社交媒体平台的广泛使用及其在信息传播方面的独特优势,使在社交媒体环境下开展档案知识聚合服务成为档案部门提升档案服务水平、实现自身价值的重要举措。从当前国内档案学者的研究内容来看,主要集中在如何利用社交媒体进行档案工作宣传和档案开发利用等方面,但是如何应用社交媒体进行档案开发利用的研究深度不够,多数文章的研究停留在现象描述、理论分析、建

议性措施方面,缺少将档案利用服务与社交媒体优势深入结合,通过研究档案知识处理、构建用户需求模型、分析具体实施机制等来为档案用户提供档案知识服务的微观细节研究。

社交媒体环境下档案知识聚合服务概念的构成要素主要包含服务资源的聚合、服务平台的聚合、服务方式的聚合及服务对象的聚合。基于此,档案知识服务资源聚合对象的范围与类别、聚合实施的流程、聚合模式与核心技术的选择,多种类型知识服务的组配与聚合方式,档案知识聚合服务的社交媒体平台选择、构建,多元档案用户群的知识需求分析及知识需求模型构建等成为社交媒体环境下档案知识聚合服务研究亟需关注的具体问题。

3.4 图、情、档协同参与档案知识聚合服务研究

相较于档案领域,图书情报领域长期关注知识组织、管理、服务等问题,形成并积累了多样的、成熟的知识组织方法、技术与系统。当前,图书情报领域知识聚合的方法主要分为:基于情报检索语言的聚合,涉及元数据、主题词和分众分类等常用检索语言;基于知识网络的聚合,涉及基于网络的统计分析、计量分析与社会网络分析方法等;基于语义网的聚合,涉及本体、微格式和关联数据等一系列语义关联方法;基于主题的聚合,是通过数据挖掘将主题相似或相同的知识聚合起来而建立的一种主题模型^[20]。

档案知识资源呈现出离散、多源、异构等特征,因此档案知识聚合难度较大,开展档案知识聚合服务对知识组织的方法和技术提出更高要求。鉴于此,图、情、档协同参与档案知识聚合服务研究成为新时期的发展趋势。一方面,图书情报领域基于元数据、社会标签、关联数据和计量等方法的聚合方法体系为档案知识聚合提供参考,帮助档案工作者实现档案知识元的有效组织与聚合;另一方面,立足档案资源自身特征,从档案利用服务视角探索知识聚合服务,有利于满足社会公众的档案知识需求,具有重要的现实意义。

4 结语

档案知识聚合服务作为知识经济时代档案开发利用形式的创新体现,为档案知识价值增值与档案服务

社会化的实现提供路径。相较于图书情报领域长期关注知识组织、管理、服务等问题,形成并积累了多样的、成熟的知识组织方法、技术与系统,国内档案学界对档案知识聚合的研究刚起步,有待进一步深入研究和体系化发展。档案领域应积极寻求与图书情报领域的交流合作,吸纳已有知识聚合服务研究成果,立足档案自身特点,建立档案学科自身的知识聚合服务研究体系,深入推进档案知识聚合服务理论与应用研究。与此同时,档案部门应借势社交媒体新工具,改变传统思维,不断提升档案服务水平,实现档案馆工作转型升级、提质增效,最大程度地发挥档案知识资源价值。

参考文献

- [1] 贯君,毕强,赵夷平. 基于关联数据的知识聚合与发现研究进展[J]. 情报资料工作, 2015(3): 15-21.
- [2] 余厚洪. 2003—2011年我国档案信息资源整合研究综述[J]. 档案管理, 2012(1): 63-65.
- [3] 蒋冠. 网络环境下档案信息资源整合研究[D]. 湘潭:湘潭大学, 2005.
- [4] 杨红仙. 信息化背景下档案信息资源的整合与共享[D]. 昆明:云南大学, 2010.
- [5] 陈永生. 政府信息资源整合共享研究——从国家档案馆的角度[J]. 档案学研究, 2010(1): 46-51.
- [6] 滕霞. 档案信息资源整合对策探析[D]. 哈尔滨:黑龙江大学, 2008.
- [7] 刘明. 网络档案信息资源整合研究[J]. 档案与建设, 2007(12): 15-17.
- [8] 梁作华. 浅谈网络环境下档案信息资源整合策略[J]. 山东省青年管理干部学院学报, 2008(6): 156-157.
- [9] 杨力,姚乐野. 基于知识管理的数字档案馆服务体系构建[J]. 档案学通讯, 2010(1): 58-60.
- [10] 张斌,郝琦,魏扣. 基于档案知识库的档案知识服务研究[J]. 档案学通讯, 2016(3): 51-58.
- [11] 毕建新,郑建明. 用户目标驱动的档案知识服务模型研究[J]. 浙江档案, 2014(8): 16-19.
- [12] 孙艳丽. 基于用户需求的档案知识服务平台及保障体系建设研究[J]. 情报科学, 2013, 31(10): 88-92.
- [13] 裘丽. 后保管时代下构建档案知识服务模式探索[J]. 云南档案, 2015(9): 52-55.
- [14] 冯惠玲. 论档案馆的“亲民”战略[J]. 档案学研究, 2005(1): 10-13.
- [15] 张斌,徐拥军. 档案事业:从“国家模式”到“社会模式”[J]. 中国档案, 2008(9): 8-10.
- [16] 黄霄羽,郭煜晗,王丹,等. 国外典型档案馆应用社交媒体创新档案服务的实践特点[J]. 档案学通讯, 2016(3): 87-93.
- [17] 韩云惠. APP在档案利用服务中的应用图景展望——基于拉斯韦尔的5W模式分析[J]. 档案学研究, 2017(4): 78-82.
- [18] 郑光辉. 使用云计算技术创造开放数字档案利用新模式[J]. 档案学研究, 2013(5): 46-50.
- [19] 向泽红. 基于云计算和3D虚拟技术的智能化档案信息服务研究[J]. 山西档案, 2014(3): 59-61.
- [20] 李亚婷. 知识聚合研究述评[J]. 图书情报工作, 2016(21): 128-136.

作者简介

魏扣,男,1986年生,博士,研究方向:档案学基础理论、知识管理, E-mail: weikoukyc@ruc.edu.cn。
李子林,女,1995年生,博士研究生,研究方向:档案学基础理论、档案信息服务。
郝琦,男,1988年生,博士,讲师,研究方向:知识管理。

The Current Status and Development Trend of Archives-Oriented Knowledge Aggregation Service

WEI Kou¹ LI ZiLin² HAO Qi¹

(1. School of Labor and Human Resources, Renmin University of China, Beijing 100872, China;
2. School of Information Resource Management, Renmin University of China, Beijing 100872, China)

Abstract: Based on the definition of the core concept of archive knowledge aggregation service, the article combs the current status of archive knowledge aggregation service from three aspects including the integration of archive information resources, the research of archive knowledge service and the research of archive utilization service. On the basis of this, the paper puts forward four major research directions including the research on archive knowledge aggregation mode, the research of archive knowledge aggregation service technology, the research on archive knowledge aggregation service in social media environment and library, information and archives collaborative participation of archive knowledge aggregation service, which will provide reference and guidance for deepening the research of archive knowledge aggregation service in Chinese archive field.

Keywords: Archives; Knowledge Aggregation; Knowledge Services

(收稿日期: 2018-01-25)