



面向知识服务的 知识库结构理论与方法

蒋 勋/著



科学出版社

国家自然科学基金项目（71774078、71303109）资助

江苏省数据工程与知识服务重点实验室项目（DEKS2014KT001）资助

江苏省“333 高层次人才培养工程”人才项目（BRA2016097）资助

江苏省“六大人才高峰”高层次人才项目（2015-DZXX-034）资助

江苏省高校“青蓝工程”人才项目（中青年学术带头人）资助

面向知识服务的 知识库结构理论与方法

蒋 勋 著

科学出版社

北 京

内 容 简 介

本书把握当今泛在数据、泛在信息和泛在知识的信息环境特征,着眼于当前知识服务的多维知识需求及知识库在大数据引领知识组织的结构体系深刻变化。主要从知识服务的角度探讨知识库结构的理论、技术和方法,并结合应急决策领域实际设计知识库结构体系。阐述经典科学理论方法对知识库结构研究的理论依据,厘清不同知识服务需求的知识库的研究边界并构建相应的结构体系,阐明知识组织、知识库与知识服务关系,架构知识库的逻辑架构与物理架构,基于粒度理论研究知识库的粒度表示及应用,分析知识服务问题与知识库粒度刻画的匹配,构建协同知识服务的模型,结合应急决策特征实现了应急响应知识库平台搭建与功能实现。

本书是知识服务实践者、知识库系统开发者的良师益友,是大数据环境知识库结构研究实践性的成果,可作为计算机科学、信息管理、情报研究的研究读本和教学用书。

图书在版编目(CIP)数据

面向知识服务的知识库结构理论与方法 / 蒋勋著. —北京: 科学出版社, 2017.12

ISBN 978-7-03-056050-6

I. ①面… II. ①蒋… III. ①知识库系统—研究 IV. ①TP311.13

中国版本图书馆 CIP 数据核字 (2017) 第 314967 号

责任编辑: 魏如萍 / 责任校对: 王晓茜
责任印制: 吴兆东 / 封面设计: 无极书装

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2017 年 12 月第 一 版 开本: 720×1000 1/16

2017 年 12 月第一次印刷 印张: 13

字数: 255 000

定价: 92.00 元

(如有印装质量问题, 我社负责调换)

作者简介

蒋勋，男，1980年12月生，江苏无锡人，副教授，管理学博士，情报学博士后。主持国家自然科学基金项目2项（青年、面上），主持江苏省社会科学基金项目2项（青年、一般），主持江苏省“333 高层次人才培养工程”科研项目1项，参与国家社会科学基金重大项目2项。先后入选江苏省高校“青蓝工程”优秀青年骨干教师、江苏省“六大人才高峰”高层次人才、江苏省“333 高层次人才培养工程”等人才项目。主要获得南京大学“优秀博士后”、江苏省教学名师、江苏省优秀科技情报工作者等荣誉。第一作者发表《情报学报》《管理学报》等 CSSCI 论文 17 篇、SSCI 论文 2 篇、专著 1 部。获江苏省哲学社会科学优秀成果奖一等奖。国家自然科学基金委项目评审专家，《图书情报工作》《情报学报》等审稿专家。



序

悄然来临的大数据，正在默默地改变着我们的一切，数据元素已经充斥于我们的生活、工作和学习之中。大数据带来的知识服务深刻改变着人们的生活方式以及企业的运作和创新模式。近些年来，知识库在经历了潮起潮落的洗礼之后显现出强大的生命力，并在国民经济活动的各类数据分析中扮演着积极角色，支撑着关系国家发展的各类数据储备、信息关联、知识传播，成为国家战略和信息进程的重要关注焦点。

与传统文献服务的知识库相比，从知识服务角度架构知识库呈现出移动性（如泛在互联、智能终端）、虚拟性（如虚拟体验、人工智能）、个性化（如精准阅读、推荐服务）、社会性（如社交媒体、社会商务）、极端数据（如富媒体、大数据）等鲜明的新特征。这些新特征是技术进步和应用创新两者交融、螺旋式演化的结果，形成了当今情报学“创造智慧”学科的鲜明主旋律。目前，学术界对知识组织领域的研究已将实现知识增值与智能重组、语义揭示与关联聚合、可视化展示与利用等作为前沿方向，也有很多成果探讨信息的本体组织和知识地图等知识的组织技术。特别是业界在重要战略技术方面的走向、新型计算模式以及社会化网络应用等对知识库提出了更高的要求。有鉴于此，解决面向知识服务的知识库最具基础性、开拓性的科学与技术难题——知识库结构，发现和诠释大数据环境下面向知识服务知识库结构的一般规律，构建知识库结构的理论体系，对“创造智慧”的情报学学科及相关学科领域的进步，以及企业竞争情报力的提升都具有十分深远的理论与实践意义。

在此背景下，我们汇聚了一批优秀学者组成具有良好研究基础、跨学科的科研团队，从知识服务视角对知识组织理论与方法开展了系列的研究工作。蒋勋博士是我们科研团队的一员，是南京大学“优秀博士后”获得者。他在承担国家自然科学基金项目时，聚焦在知识库的内在逻辑结构与整体框架结构、面向知识服务的知识组织体系、细粒度知识节点的组织模型、知识节点的知识表示模式、知识库实体对齐与扩充、协同视角的知识服务过程，并将凝练的成果拓展知识服务的研究边界，结合应急决策领域的知识服务实际在应急响应的知识表示与协同知识服务、应急响应知识库结构、应急响应知识库构建等方面的若干科学问题和技术方法进行了学术探索与科技攻关。《面向知识服务的知识库结构理论与方法》一书是上述研究的结晶，该书从知识服务的现实需求出发，以“知识的有序组织与

科学表示、知识间关联关系的深度揭示与科学链接”为依据，构建了完整的面向知识服务的知识库结构体系。其研究对于构建完善的知识组织学科体系，促进信息服务向知识服务延伸具有重要的理论与实践意义。

面向未来，我们将继续看到知识服务在我国信息化和经济建设中日益重要的作用。而知识服务伴随着知识组织技术的进步包括知识库结构的创新脉动，在带来理论和实践挑战的同时，也孕育了大量的学界与业界的机遇和广袤的探索空间。

A handwritten signature in black ink, appearing to be 'J. H. H.' or similar, with a stylized, cursive script.

2017年初冬于南京

前 言

我们正处于泛在数据、泛在信息和泛在知识的时代，由互联网、数字化、多媒体、云计算及大数据等组成的集合式多维信息环境构成了我们这个时代的主要信息特征，其影响也具有泛在特点。在这样的影响下，知识库是知识服务实现获取知识和应用知识两大职能的支撑，促进了图书情报工作从传统的文献信息服务阶段，转变和提升到用于解决实际问题的知识信息服务阶段。对数据、信息、知识的科学、有效的组织，是实现知识服务最基础、最关键的工作，也是通过知识组织来提升知识服务质量时常遇到的瓶颈问题。当前大数据的环境凸显了多（数据量大）、冗（冗余信息多）、杂（真伪信息混杂）、乱（异构信息同存）、新（实时数据快速产生）的数据特征，这对数据处理与知识组织提出了更高的要求，对更为有序地收敛于高效的知识服务提出了巨大的挑战。由此导致的科学与技术难题是相关知识的有序组织与科学表示、知识间关联关系的深度揭示与科学链接。

现有的技术手段依赖于知识库的结构来计算类的依存关系以及知识库中挂载实体的相似度。知识库结构是实现高效知识服务的基础，是破解大数据环境知识组织难题的关键。面向知识服务架构知识库结构体系需系统解决的关键科学问题：由知识库内在逻辑结构与整体框架结构的深度揭示，带来的满足用户不同服务需求与信息资源有效整合，消除知识组织与知识利用之间的鸿沟的问题；由知识库支持的知识服务改变知识获取的固有流程，形成新的知识获取过程的问题；用户高效利用、共享知识后推动知识库从单纯的知识获取向专业化、领域化的知识服务系统转变的问题。这些新问题及其交叉和渗透，为知识库结构的研究带来了前所未有的挑战和机遇。创造性地解决这些瓶颈难题，不仅是发展和完善知识组织技术的关键所在，也是提高我国知识服务科学技术含量和国际化水平的迫切需求。

围绕面向知识服务的知识库结构研究，本书结合实际应用背景探讨知识库的类型及其对知识库结构的要求和对知识服务的支撑，阐述科学理论与方法、知识组织方法、系统科学理论对知识库研究的理论依据。厘清常识性知识库、事实经验知识库、推理知识库的研究边界并构建相应的体系结构；研究知识库的组织架构，介绍知识库的架构目标，从知识序化、知识更新、知识发现、知识组织等方面阐述知识库结构的作用；阐明知识组织、知识库与知识服务的关系，在此基础上架构知识库的逻辑架构与物理架构。在体系架构基础上，进一步讨论知识库挂载的知识节点，基于粒度理论研究知识库的粒度表示；探究知识库的粒度应用，

分析知识服务问题与知识库粒度刻画的匹配,相应探讨主题词表的粒度应用与分类表的粒度应用。构建协同知识服务的模型,并以此实现应急响应的知识表示与协同知识服务,结合应急决策的特征构建应急响应知识库结构模型与平台搭建、功能实现。

本书的撰写得到了江苏省数据工程与知识服务重点实验室研究团队的大力支持。为了本书高质量地完成,与重点实验室的专家团队进行多次讨论和交流,在交流中碰撞出火花,对问题产生出更深刻的理解。2013年实验团队出版的《面向知识服务的知识组织理论与方法》对知识组织理论根源性、知识组织的模式和具体结构及发展过程中的认知路径问题进行了系统化、体系化的深入研究,这系统性的研究成果奠定了本书从知识服务的角度研究知识库结构的理论基础。本书是集体智慧的结晶,书中汇聚了团队多年来的研究成果,分别是第2章(陈祖琴)、第5章(徐绪堪、沈思)、第8章(张艳琼、唐明伟)。参加本书讨论的学者还有杨建林、章成志、孙立媛、刘喜文、郑昌兴、朱惠等。

本书系国家自然科学基金项目(71774078、71303109)、江苏省数据工程与知识服务重点实验室项目(DEKS2014KT001)成果之一,研究过程得到了国家自然科学基金的资助。本书的出版得到了江苏省“333 高层次人才培养工程”人才项目(BRA2016097)、江苏省“六大人才高峰”高层次人才项目(2015-DZXX-034)、江苏省高校“青蓝工程”人才项目(中青年学术带头人)的资助。在此对国家自然科学基金委员会、江苏省规划办、江苏省委组织部、江苏省人力资源和社会保障厅、江苏省数据工程与知识服务重点实验室表示衷心感谢。感谢南京大学信息管理学院苏新宁教授带领的研究团队,他们对本书的完成作出了很大贡献。

本书的内容是探索性的研究成果,作为新兴领域的一部专著,难免会有许多不足与值得完善之处,恳请广大读者批评指正。

蒋 勋

2017年10月于无锡

目 录

第 1 章 引论	1
1.1 研究背景	1
1.2 研究意义与内容概述	2
1.3 知识库相关研究动态分析	5
1.4 科学理论与方法对知识库研究的支撑	11
第 2 章 面向知识服务的知识库类型及构建	17
2.1 基于领域知识特征的知识库框架设计	17
2.2 常识性知识库构建	19
2.3 事实经验知识库构建	28
2.4 推理知识库构建	33
第 3 章 知识库的组织架构	40
3.1 知识库的架构目标	40
3.2 知识库结构的作用	48
3.3 知识库结构的组织	69
第 4 章 知识库的粒度表示	92
4.1 知识库的数据准备	92
4.2 知识库的粒度语义	99
4.3 粒度的知识语义片段切分	104
4.4 基于语义的知识库检索	111
4.5 语义相似度与相关度计算	115
第 5 章 知识库的粒度应用	121
5.1 粒度匹配的约束	121
5.2 主题词表的粒度应用	126
5.3 分类表的粒度应用	137
第 6 章 协同视角的知识服务过程	145
6.1 用户个体能力模型	145
6.2 社群成员选择模型	148
6.3 算例分析	153

第 7 章 应急响应的知识表示与协同知识服务	155
7.1 应急响应的知识表示	155
7.2 跨领域协同知识服务	161
第 8 章 应急响应知识库结构	169
8.1 应急响应与知识库结构的关联	169
8.2 知识节点的应急响应特征识别	171
8.3 应急响应粒度的多维组织	175
8.4 基于应急响应知识库结构的实例分析	183
参考文献	190

第1章 引 论

1.1 研究背景

大数据时代背景下,人工智能、专家系统等技术有了充分的数据支撑,使得决策分析有了准确的方向保障,保证在不失时机的前提下把握好趋势,降低决策不确定性。但大数据自身的4V[数据量大(volume)、类型繁多(variety)、价值密度低(value)、速度快时效高(velocity)]特征决定了基于大数据的决策分析更偏于概率趋势,针对具体的专指的领域性强的决策,特别是国家亟需的应急响应决策,不仅需要趋势掌控更需要精确的知识支持。知识库汇聚了各类知识实体,这些知识凝练了各领域专家学者的智慧,是减少应急处置失误,实现决策的科学化、民主化的基础,尤其对国家应急决策来说知识库是非常重要的战略知识储备。

因此,本书研究的知识库是在大数据环境下基于知识组织理论方法有效改进知识库的语义组织,提高知识检索效率,提升具有领域特征知识库的知识服务效果。

目前对知识库构建的关键技术研究,从知识组织的角度已有诸多理论与成熟实践,如利用主题图技术描述资源的知识结构和知识关系,利用资源描述框架(resource description framework, RDF)进行知识呈现和语义标注,通过关联数据实现知识库语义关联扩展,利用产生式规则、框架表示法等进行知识表示,利用主题词表进行知识组织、知识检索,利用本体实现有效的知识表示、知识分享、知识推理和知识整合等。然而也必须看到,知识库的构建是面向实践应用的,如最终用户非专业的检索表达、不同粒度的知识需求等将极大地影响用户体验,仅从知识组织角度“序化”知识库,还不足以满足各类知识服务,更需从知识服务角度“驱动”知识库的结构组织,使之与各类知识服务匹配。

知识服务对信息组织有很高的要求,强调数据间的有机关联,尤其要求通过数据间的联系从而形成相关知识。因此,如何有效存储、管理、组织和更新大规模的知识,如何有效地利用存储的知识进行推理和问题求解,并更有效地支持知识服务,是目前图书情报领域的一个重要研究课题。实现知识服务的途径有多种,其中基于专家系统提供的专家级知识,能实现一种高端的知识服务。而专家系统中知识获取的核心组成部分是知识库,尽管计算机与人工智能技术的迅猛发展支撑着知识获取手段的提升,但“知识匮乏”与“信息海洋”的对峙局面仍然存在。

为解决这一对峙局面,本书从知识服务的角度架构知识库结构模型,研究知识库的内在逻辑结构与整体框架结构,使得健全、合理的知识库结构支撑起用户不同的服务需求,并使资源和服务有效整合,消除知识组织与知识利用之间的鸿沟。由知识库支持的知识服务将改变知识获取的固有流程,形成新的知识获取过程模型;而由用户高效利用、共享知识后更将推动知识库从单纯的知识获取向专业化、领域化的知识服务系统转变,形成全新的动态知识库系统,从而丰富了知识库的结构。

1.2 研究意义与内容概述

1.2.1 研究意义

1. 对知识服务的推动及服务领域的变革

图书情报工作已从传统的文献信息服务,转变和提升到用于解决实际问题的知识服务。这样的变革更坚定了本书以知识服务为目标,对知识库结构模型中所涉及各环节要素进行理论及实践上的探讨,这种探讨有助于解决知识库在知识组织中所面临的困难,进而解决从知识源到知识获取过程中的瓶颈问题,并且能够揭示出数据库与知识库在知识获取过程中各自结构上的对应关系,构建双库协同的内在联系“通道”,最大限度地克服海量数据增加引起的算法失效及固有的知识库实时维护困难的问题。这一系统化的研究紧紧围绕知识库结构凝练了知识组织的相关理论,呈现出知识库是实现知识服务的基础和核心,最终还以知识服务为落脚点,为拓展知识服务的领域奠定了基础,顺应了知识经济的发展需要,为企业科技创新、政府管理创新提供深度且可靠的知识服务,为加快我国信息产业服务能力的提升,支持信息处理领域的知识创新,推动信息服务模式的转变,最大限度地提升我国知识服务的整体水平将作出一定贡献。

2. 勾勒知识库结构理论框架

知识存在知识库中,数据存在数据库中,如何使两者“对接”。由数据子类结构形式构成的发掘数据库的可达范畴与基于属性间关系的挖掘知识库的推理范畴,两者之间建立等价关系,这一等价关系为数据库与知识库间定向的知识挖掘与知识搜索奠定了理论基础;从而依据子类结构构建的数据库和根据知识节点网络构建的知识库逻辑等价;这逻辑等价关系打破了传统数据库知识获取的封闭系统,而与知识库结合起来,从知识库的高度,获取更深层次的知识。这在结构和功能上形成了相对于基础知识库而言一个开放、优化的可扩展体,这样的知识库结构将有助于实现知识推理和知识挖掘,并能拓展知识服务的空间,提高知识服务

质量的稳定性,为领域内新知识的快速产生提供创新思路。这项研究不仅对知识库自身结构有积极的指导作用,而且可以帮助完善和深化知识组织理论与框架,提升知识服务水平,其相关成果在未来推动科技创新以及科研环境的变革的知识服务平台上具有广阔的应用前景。

3. 形成知识库“深加工”的知识组织新模式

对知识库的应用不是简单的“拿来”,而是由“向下解构、向上建构”的探索思路对知识库整体进行基础组织模式、知识发现、关联网络、多维知识组织等深加工,实现知识发现和本体、面向决策的知识地图、关联分析与方案重组的新应用。基于此,勾勒知识库与数据库在结构和功能上形成“双库协同”知识获取模式,打破知识库中解决方案知识发现的封闭系统,构建双库协同的内在联系“通道”。为知识库内解决方案的快速产生提供创新思路,最大限度地克服大数据增加引起的算法失效及固有的智库实时知识获取困难的问题,同时也克服了知识库巨量性的问题。

4. 对未来知识库科学发展的作用

知识库的构造过程不是一次性的,而是需要很多次才能构造完成的。在构造知识库的过程中,可能会出现这样或那样的问题,一个重要的问题就是如何保证知识库的协调性,尤其避免产生知识库推导出矛盾的结论。这其中蕴含了几个迫切需要解决的关键科技问题:知识库中规则表示方法以及针对该规则的评价方法;如何将数据库的数据作为知识库的知识源;知识源是否干净以及针对非清洁数据的清洗问题;数据库中数据的更新是否会同步知识库中知识的更新。这些都影响着知识在使用过程中的有效性、一致性、完备性及共享性。本书的研究工作一方面梳理了知识库与传统数据库、文档库的不同,清晰勾勒出知识库的概念内涵;另一方面充分揭示了未来知识库的科学发​​展必然与数据库的运用有联系,牵一发而动全身。本书的研究正是解决这一系列基础性的问题,从而完善知识库自身结构,使之成为一个协同的有机整体,并以知识服务为研究知识库结构的最终落脚点,开拓未来基于知识库的知识组织创新模式。

1.2.2 内容概述

1. 数据库、知识库双库协同机制

双库协同机制主要由启发型协调器与中断型协调器来实现。启发型协调器是通过搜索知识库中知识节点的不关联态,以挖掘新的知识链,弥补知识短缺,从

而启发并激活真实数据库中相应的数据类,以产生定向发掘进程;而中断型协调器从真实数据库的大量数据中经聚焦而产生规则知识之后,使数据库的知识获取进程“中断”,而去搜索知识库中对应位置有无此规则知识的重复、冗余、循环等,若有,则取消该生成规则,若无,则继续知识获取进程。双库协同机制的研究揭示知识库与数据库之间的对应关系,为实现减小搜索空间、提高挖掘效率提供有效的技术方法。

2. 知识库中规则的简约表示

知识库与数据库最本质的区别在于知识库中的规则集合。对知识库中知识的更深层次的挖掘与提炼的驱动,必然要求建立一个能反映各知识节点之间关系的映射表示。可形象地表示成规则 $r = (s \rightarrow d)$, 其中知识节点 s 为规则 r 下的起始知识节点, d 为规则 r 下的目标知识节点, 符号“ $\rightarrow$ ”表示知识节点之间存在的映射关系。

虽然这样规则的表示显得清晰且能说清问题,但描述方式显得比较朴素,且随着知识库中知识节点数目及对应知识节点之间关系的映射数目的增加,直接使用规则 $r = (s \rightarrow d)$ 表示知识库的逻辑结构将会较为复杂。因此知识库中规则的简约表示是两难问题,在本书中将细化知识节点的粒度,进一步研究简约起始知识节点及简约目标知识节点的表示,从而研究规则的简约表示,以提高挖掘效率。

3. 知识库中的数据清洗

目前针对知识库中数据清洗的研究不多,就现有的研究工作中可以看出面临的问题:①知识库中很多非清洁的数据是很难被彻底清洗干净的;②对非清洁数据的清洗可能造成部分知识的丢失或失真;③知识库中信息更新频繁,要同步执行非清洁数据的辨识与清洗将极大影响知识库系统的效率。那么在本书研究中数据清洗的研究核心问题是如何从包含非清洁数据的知识库中获取满足非清洁度的查询结果,也就是要给出描述数据非清洁度的机制。数据非清洁度的机制将从两个方面研究:一方面是数据本身的非清洁度,表现特征是数据的错误、不准确、不完整;另一方面是数据间的非清洁度,表现特征是数据的重复、冗余、不一致。针对这两个方面,本书进一步将研究从元组上添加非清洁度来描述数据本身的非清洁度,在元组集合上定义元组相似性来描述数据间的非清洁度。

4. 构建应急决策知识库

围绕应急响应,从公众、救援人员、决策者三个角度对知识库的使用需求进行梳理,并结合应急响应各阶段过程,从常识知识、事实与经验知识以及运算推理规则三个维度探索应急响应的知识库体系框架的构建。基于知识库集结描述清

晰、反映全面、关联紧密的知识,是应急决策的智力支持,也是应急响应情报体系形成的前提,在应急响应决策中发挥着重要作用。

将理论探索中凝练的知识库结构理论提升到用于应急决策的高度,使得知识库适应信息量大、精确性高、跨领域多的情景演化要求,形成知识库协同架构有序、高效、及时地提供应急解决方案。所构建的知识库体系框架注重应急响应知识普及功能,应急过程中的快速响应功能。体系框架中以常识性知识库 *c*-KB、事实经验知识库 *e*-KB、推理知识库 *r*-KB 3 个知识库组建,为应急决策提供了精确、及时的知识支撑。

1.3 知识库相关研究动态分析

1.3.1 知识组织角度分析知识库的发展动态

国内关于知识组织的研究,还处于起步阶段,主要集中在高等院校和科研院所。在我国,“知识组织”一词最早出现在袁翰青 1964 年的文章中,认为,文献工作就是组织知识的工作,现在看来这是对知识组织的朴素理解。受国外图书馆界“知识组织”影响,刘迅于 1985 年首次将“知识的组织”作为图书馆学情报学研究的一个内容,自此“知识组织”开始进入图书情报领域研究的视野。对知识进行有效组织使其更好地被发现、利用,这是图书情报学研究的重要目的,因此知识组织也成为我国图书情报学的核心研究领域之一。

从知识组织的研究视角关注承载知识的知识库研究,王知津等(2009)的《知识组织理论与方法》在已有研究成果的基础上,全面、系统、深入地分析了知识组织的理论、方法及应用,首次对知识组织的环境进行了考究;并对知识组织的十个方面加以论述,其中就指出实现知识的有序化是知识组织的重要目标之一,而知识有序化包括了承载知识环境的有序化,并给出了支持知识环境有序化的技术手段。蒋永福和李景正(2001)清晰地勾勒知识组织的内涵,即对知识客体进行整理、加工、表示、控制等一系列过程,并给出了知识表示、知识重组、知识聚类等方法,这些经典的方法对今后从事知识库的研究依然适用,有助于我们更好地实施知识库建设。赖茂生等(2009)深入研究了知识组织领域的几个前沿问题:本体与传统知识组织工具改造的研究、受控语言与自然语言融合的研究、网络环境下知识组织研究及 Web2.0 与自由分类法,这些前沿问题的研究为进一步展开知识库的探索奠定了坚实的基础。贺德方(2010)对知识组织的汉语主题词表的研究进行了整体的回顾和展望;并从知识组织体系(knowledge organization systems, KOS)的应用角度,梳理了知识组织的主要成果,归纳了知识组织体系之间的转化、映射、标准化等方面,这些研究厘清了从知识组织角度研究知识库结构的思路。围绕 KOS,国内其他

学者也开展了建设性研究,司莉和周李梅(2010)从全球视野并通过大量的文献调研,分析了国外英文知识组织体系的研究现状,指出了 KOS 作为知识组织的载体涵盖了所有组织信息和促进知识管理的工具。王军和张丽(2008)的研究则从网络环境下对知识结构进行系统化描述与说明,以此为基础推动信息组织、信息表示和基于内容的信息检索等应用发展,其研究中继承了国际网络知识组织系统(networked knowledge organization systems/ services, NKOS)已有研究成果并针对国内知识组织工具的特点和应用环境进行了适应性与独创性的设计。朱礼军的研究围绕着由贺德方主持的“十一五”国家科技支撑计划重点项目子课题“知识组织系统的集成及服务研究与实现”,产出了重要研究成果,系统总结了基于词型、结构、语料的术语映射方法,并提出了 KOS 作为知识服务与知识组织的重要支撑工具,其结构应突破线性,采用网状结构,纵横交错、多维揭示。关于 KOS 的一系列新颖的分析方法纷纷涌现,为后续深入知识库结构的探讨,起到了指导性的、一般性的和模型化的强有力的理论支撑。屈丽娟和彭洁(2012)通用文献计量学分析了图书情报领域 20 年内围绕“知识组织”的各研究热点的动态变化情况,给出了基金论文中国家自然科学基金占基金总数的 21%,说明知识组织领域对技术的发展关注度还是很高,其研究成果揭示了知识组织沿着“理论—技术—服务—人”的路线发展。由侯汉清带领的南京农业大学科研团队承担国家社会科学基金“基于知识库的中文信息自动标引及自动分类”项目,奠定了知识库中知识组织的理论、方法,并为学者后续知识库的研究指引了研究方向。2011 年,由中国科学院文献情报中心等单位联合启动了国家“十二五”科技支撑计划“面向外文科技文献信息的知识组织体系建设与应用示范”项目,这是我国“十二五”一项重大工程项目,对我国未来向科技领域有效地提供外文科技文献服务打下了坚实基础。

国外关于知识组织的研究更重于实践方面,知识组织研究主要是以 NKOS 作为主要的研究领域开展相关研究。随着数字图书馆的研究与发展,NKOS 已经成为图书馆和信息科学领域最重要的研究课题。特别针对语义化描述的分类表、叙词表、主题词表、地名辞典等已经在国际图书情报界有较多的研究和实践,如美国国会图书馆发布的《国会图书馆主题词表》的简单知识组织系统(simple knowledge organization system, SKOS)语义描述版本。进一步基于 NKOS 向机器和各类智能应用提供服务,通过构建 NKOS 的目标为搜索引擎、学科门户、内容导航、自动分类等应用程序提供知识服务,其途径是提供开放的、统一的查询接口和访问协议,如典型的 SKOS API。

1.3.2 知识服务角度分析知识库的发展动态

中国人民解放军医学图书馆与重庆九济科技发展有限公司合作研发了中国第

一个疾病知识库《中国疾病知识总库》，目前包含 7 个库、30 多个专科系统和 7000 多种疾病信息，在医学知识服务中发挥着重要作用。毕强等（2007）比较了国内外在知识服务领域研究的差异，并揭示了信息资源只有定位描述而缺少资源内容的确切描述的不足，提出了知识库是实现知识服务的基础和核心，通过知识库将提高知识利用和共享效果。王知津等（2009）提出了信息服务未来的十大走向，认为服务目标从“信息资源”走向“问题求解”、服务内容从“信息服务”走向“知识服务”，并提出了知识服务的本质是知识创新，核心是以人为本，重点是帮助用户发现问题和解决问题。从信息服务向知识服务过渡中，温有奎和焦玉英（2010）认为知识元是实现知识服务的实体单元，构建了知识元模型使之独立于文献而直接进行知识服务。夏立新等（2008）关注到知识提供者与知识使用者的信息不对称问题，开创性地提出了知识供应链将知识提供者与最终知识用户联系起来，提出的知识流提升了知识服务的结构模式，为开展知识服务提供了新的视角。董慧和唐敏（2011）深入研究了本体知识表示的知识库，避免了不同数据源中知识异构带来的问题，在这基础上首次提出了历史领域的专家系统模型；更在其后续的研究视野中聚焦 Web2.0 服务，将语义检索应用于 Web2.0 服务中，能很好地调动了用户互动参与，增强了信息搜索的准确度，使服务更具个性化。付雅慧（2009）也注意到知识组织不仅仅是技术方法的问题，而关键是与用户需求无障碍对接，消除知识组织与知识服务之间的鸿沟，这是未来方向。王曰芬等（2008）研究了面向个性化服务的知识组织机制，形成了个性化服务的知识组织；在其后续成果中将知识服务瞄准了科技咨询，研究了将社会网络分析与社会网络可视化的先进技术引入专家库的建设中，提高了科技咨询决策效果。漆贤军和张李义（2009）分析了我国知识创新网络的结构体系，研究了知识信息服务的业务拓展与转型，并提出了知识信息服务多元化与网络化的双轨体制，有效地推动了我国自主创新能力的提高。曾建勋和邓胜利（2011）在国家科技图书文献中心资源建设和服务方面，通过知识库为各机构提供联合信息发现和揭示能力，带动各机构自身资源的发现和服务，拓展和提高了国家科技图书文献中心的知识发现与知识组织功能。李广建和李亚子（2008）指出在分布式环境下为适应 OpenURL 需要，知识库的有效利用是关键问题，知识库不仅是整个系统的核心，也是整个过程的中枢并且为构造扩展服务链接提供规则，认为只有健全、合理的知识库才能提供用户尽可能合适的知识服务。张玉峰和何超（2011）则针对非结构化的自然语言文本，深入研究了领域本体的语义文本挖掘，其研究促使更有价值的知识模式的产生，使挖掘结果在实际应用中更具指导意义。洪韵佳和许鑫（2012）围绕知识库架构与建设方法、知识组织与分类方法、知识库的检索服务、知识挖掘与发现，并从参考咨询服务的用户需求出发，对联合参考咨询知识库的发展趋势进行了探讨，更好地利用知识库的功能以提升图书馆的参考咨询服务能力。王伟军和孙晶