doi:10.3772/j.issn.2095-915x.2018.05.001

工程科技领域知识服务的思考与实践

潘刚 [1]　张运良 [2,3]　钟庆虹 [2,3]

1. 中国工程院　北京　100088；
2. 中国科学技术信息研究所　北京　100038；
3. 富媒体数字出版内容组织与知识服务重点实验室　北京　100038

摘 要 在大数据时代，用户对综合性知识服务的需求不断增加，亟需多来源数据的知识融合和知识服务。随着知识融合理论不断深入，多来源数据的知识融合和知识服务在政府、企业、出版、图情机构等领域已有一些探索。本文介绍了工程科技领域知识服务的架构，以及中国工程科技知识中心在建设和服务实践过程中的探索，以期为知识融合和知识服务提供参考和借鉴。

关键词：知识服务；知识融合；知识组织；中国工程科技知识中心

中图分类号：G35

开放科学（资源服务）标识码（OSID）

Thinking and Practice of Knowledge Services in Engineering Field

PAN Gang[1]　ZHANG Yunliang[2, 3]　ZHONG Qinghong[2, 3]

1. Chinese Academy of Engineering, Beijing 100088, China;
2. Institute of Scientific and Technical Information of China, Beijing 100038, China;
3. Key Laboratory of Rich-media Knowledge Organization and Service of Digital Publishing Content, SAPPRFT, Beijing 100038, China

Abstract In the era of Big Data, the demand of users for comprehensive knowledge are increasing, so the knowledge services of multi-source data fusion is urgently needed. With the deepening of the knowledge fusion theory, there already exist few research in the practice of knowledge for governments,

基金项目：中国工程科技知识中心建设项目“知识组织体系建设”(CKCEST-2018-1-26)；中国科学技术信息研究所重点工作项目“富媒体数字出版内容组织与知识服务关键技术研发与应用示范”（ZD2018-07）。

作者简介：潘刚（1976-），中国工程科技知识中心办公室副主任，E-mail: pg@cae.cn；张运良（1979-），博士，研究员，研究方向：知识组织、知识工程；钟庆虹（1995-），硕士研究生，研究方向：知识组织、媒体融合。

enterprises, publishing houses, and information service institutions. An engineering technology knowledge service architecture is introduced with the exploration of the construction and service practice of China Knowledge Centre for Engineering Sciences and Technology (CKCEST) in this study. The Knowledge Centre will continue to promote the sharing of knowledge resources and more integrated knowledge services in the future, to provide better information support and knowledge services for improving the independent innovation capability of the engineering field.

Keywords：Knowledge services; knowledge fusion; knowledge organization; China Knowledge Centre for Engineering Sciences and Technology (CKCEST)

1 引言

为国家工程科技领域重大决策、重大工程科技活动、企业创新与人才培养提供支撑和知识服务，提高国家自主创新能力，是中国工程科技知识中心(简称“知识中心”)的重要使命。中国工程院陈左宁副院长[1]指出“知识中心以跨领域专业数据融合与深度知识挖掘为目标”，可见专业数据融合在知识中心建设过程中的重要性，本文通过对知识中心相关工作做了一些梳理和思考，以期对工程科技领域专业数据融合和知识服务有所借鉴。

2 国内外现状

2.1 知识融合与知识服务理论发展

知识融合最具有代表性的定义是 Preece 等[2]在 KRAFT 项目中提出的“知识融合是指从多个异构的网络资源中定位和抽取知识，并对其进行转换以使整合的知识可以应用于解决某一问题的过程”。Alexander Smirnov 等[3]阐述了决策支持系统中基于上下文的知识融合模型，并提出知识融合有 7 种知识融合模式：简单融合、扩展融合、实例融合、参数融合、适应性融合、分面融合和历史融合。刘晓娟等[4]通过不同时期、不同学科国内外的研究者对知识融合的概念表述从处理对象、融合过程、融合结果这三个要素进行分析总结得出知识融合的结果是产生新的知识，并提出知识融合面向的是知识服务和决策问题，基础是多源异构数据。张心源等[5]提出知识融合在图书情报领域具有以丰富的知识源作为研究背景，与大数据研究联系紧密，知识融合以应用研究为主，以提供知识服务为主要融合导向等特点。

在大数据环境下，网络信息环境、信息技术和社会信息需求发生重大变革，对知识融合及服务的影响也是多方面的。张兴旺等[6]关注大数据与信息移动推荐服务的有机融合，建立了面向大数据的信息移动推荐服务模型，提出了相应的体系结构，建立了技术体系。赵蓉英等[7]采用了文献计量法对知识融合的国内外研究状况进行分析，结合大数据环境的特点，分析对知识融合在理论、需求、条件、过程、系统以及应用六个方面产生的影响。张心源等[8]、唐晓波等[9]分别独立构建了大数据环境下的知识融合框架模型，目的在于将知识融合作为大

doi:10.3772/j.issn.2095-915x.2018.05.001

数据环境下有效利用知识资源和获取新知识的手段，从而为用户提供个性化、精准化的知识服务。前者遵循“输入 - 处理 - 输出”将其分为三大模块：融合预处理、融合过程和融合服务；后者也遵循这一框架，但形成的是基于机器学习的动态反馈知识融合框架模型。Fisch 等 [10] 提出当以分布式方式从样本数据中提取诸如分类规则的知识，需要对这些规则或知识进行融合，并给出了一种基于使用概率生成分类器在分类规则参数层次上融合分类器的新方法。知识融合研究以知识融合的框架、模式等方面的理论研究为主，服务应用研究相对较少，但是也有一些研究性探索。

2.2 知识融合与知识服务研究进展

在知识经济快速发展的时代，知识融合能帮助政府、企业、出版机构、图情机构解决大量分布异构式知识资源有效关联和充分发挥作用的难题。

政府知识融合方面，陈为东等 [11] 从多维语义视野研究分布式异构多源政府网站信息资源的知识融合，构造了本体知识库，结合知识融合预处理、过程和结果 3 个关键要素以及语义环境、语义推理规则和融合算法等构建了政府网站信息资源多维语义知识融合的概念模型、结构体系及解知识空间服务模型，为知识融合领域研究提供参考借鉴。胡海波等 [12] 针对大数据环境下的电子政务知识服务的现实需求，提出了基于知识融合的电子政务知识服务模式，构建了基于知识融合的电子政务知识服务系统模型，包含本体库、政务知识库、元知识集、融合算法、解知识空间的生成、动态的反馈系统等要素。

在企业知识融合方面，古志文等 [13] 利用企业知识管理理论中的 SECI 框架界定企业知识服务和信息服务，结合知识服务驱动力、覆盖面、增值度这三个要素，构建了一个关于企业技术创新知识服务模式的三维立体模型，该模型包括缄默知识服务、信息服务两个系列各 5 种基本模式，从而体现出知识经济、知识服务产业与事业发展的一些规律。刘妍等 [14] 指出我国大多数企业均采取知识融合的方式进行创新，分析了集成创新和消化吸收再创新两种方式在知识创造生态过程的三个阶段：知识准备、知识生发、知识扩散，均表现出对应的特点，分别构建了知识生态模型，对发现提升企业融合创新效果的路径提供了思路。E.U.Kriegel 等 [15] 研究了一个独立于应用程序的知识融合工具包，用于支持智能辅助，早期预警和安全系统的决策。该工具包基于工业级 Java 企业技术，为模块化知识组件提供运行环境，这些组件包含所有与应用程序相关的功能。

在出版知识融合方面，2015 年以来基于知识服务的出版融合研究开始引起关注，逐步成为专业出版单位的数字化转型方向之一。李弘 [16] 结合具体的出版单位融合发展案例，分析专业出版单位知识服务转型的发展路径。陈建华 [17] 探讨了在媒体融合环境下，科技期刊知识服务在理论创新、载体创新、方法创新、模式创新这四个方面结合了国内外的理论实践资料，提出了科技期刊知识服务新的发展方向和应用场景。陈晓峰等 [18] 在对媒体融合背景下的学术期刊知识服务进行了充分调研的基础上，结合期刊运营实践，充分利用新媒体、大数据等技术，探索形成了知识网红 IP、在线视频课程、线上

知识问答、资讯订阅定向推送、热门主题圈、在线研讨会直播 6 种精准知识服务模式，并从“用户连接、需求分析、运营管理”三个方面提出具体实践路径。

在图情机构服务转型方面，更多的是分析面临问题并提出解决策略。郑杰等[19]从新媒体对图书馆知识服务的影响出发，列举了图书馆知识服务深度融合的三个方面，提出了在媒体融合背景下的具体发展与转变策略。叶翠等[20]探讨了馆社知识服务融合需面对的利益分配博弈问题、知识产权保护问题以及跨界融合交流问题，并给出了相应对策。

3 知识中心知识融合和服务实践

3.1 知识中心整体架构

知识中心是国家工程科技领域公益性、开放性的知识资源集成和服务平台，知识中心整体架构如图 1 所示，主要包括基础设施层、平台支撑层、资源建设层和应用服务层。

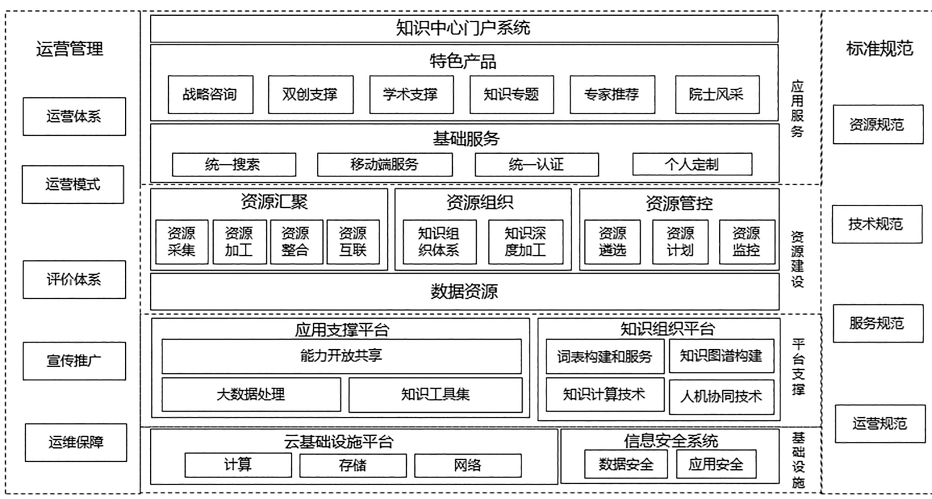

图 1 知识中心总体架构

基础设施层包括云基础设施平台和信息安全系统。云基础设施平台是知识中心平台所依托的基础云环境，包括计算资源、存储资源、网络资源及包括操作系统、数据库、虚拟化软件在内的系统软件资源。

平台支撑层包括应用支撑平台和知识组织平台。应用支撑平台将各类应用正常运行所必需的运行环境从应用系统当中剥离出来，以一体化平台的方式加以建设。

资源建设层包括资源汇聚、资源组织、资源管控。资源汇聚是将来自不同协作单位、多元异构环境的数据进行融合存储，形成互联互通的整合体，支持对资源的融合利用；资源组织是通过建设分类体系、词表、知识图谱等知识组织体系来支持知识服务的开发；资源管控是依据评价标准对各类资源进行评估、遴选和质量控制，对数据采集、加工处理、存储、应用展现等过程内容实行监控。

应用服务层是基于知识中心广泛的资源形成的用户可感知的知识融合类知识服务。知识

doi:10.3772/j.issn.2095-915x.2018.05.001

中心为用户提供的 WEB 接入界面，通过门户系统，用户可以通过统一认证，对知识中心的资源和应用服务进行一站式访问。

3.2 资源融合实践

在知识中心整体框架和知识组织体系支撑下，知识中心按照“持续清洗、多维标引、特色关联、深度融合”的原则，对资源进行持续的开发、融合和服务利用。

由于数据资源来源多样，不同来源资源体例和加工程度也不尽相同，知识中心以元数据规范为基础，结合知识组织体系，并探索了作者、空间等维度的资源融合。知识中心的数据来源主要是集中建设的资源数据以及各个分中心的特色数据，通过资源融合，主要将原来分散的工程科技数据关联打通，通过服务实现数据增值。

3.2.1 元数据规范化

利用元数据对数字化信息资源进行描述、管理、交换和共享，已成为当前数字信息资源建设的重要手段。为指导和规范知识中心各类数字资源元数据汇交、元搜索服务、资源集成整合和系统间互操作，从根本上实现知识中心资源的汇聚打通，知识中心建立了一系列与其建设目标相适应的元数据标准规范[21]。

目前，知识中心元数据规范包含了管理、主题、责任者等 13 类通用容器，并对 1 个数据集元素集和期刊、图书、论文、学者、成果、专利、标准等 22 类资源元素集，以及 27 个规范编码体系做了详尽的元数据规范描述和约定。知识中心充分借鉴了国际上相关元数据研究应用成果，制定的元数据规范在知识中心各协作单位之间资源描述、交换、复用、转换、整合和打通等环节发挥了重要作用。

3.2.2 知识组织

知识组织（Knowledge Organization）是知识中心对各类信息资源进行整理、加工、索引、分类、标引等的一系列活动，知识组织是知识中心开展知识融合和知识服务的纽带和桥梁。知识中心知识组织体系建设的框架如图 2 所示。

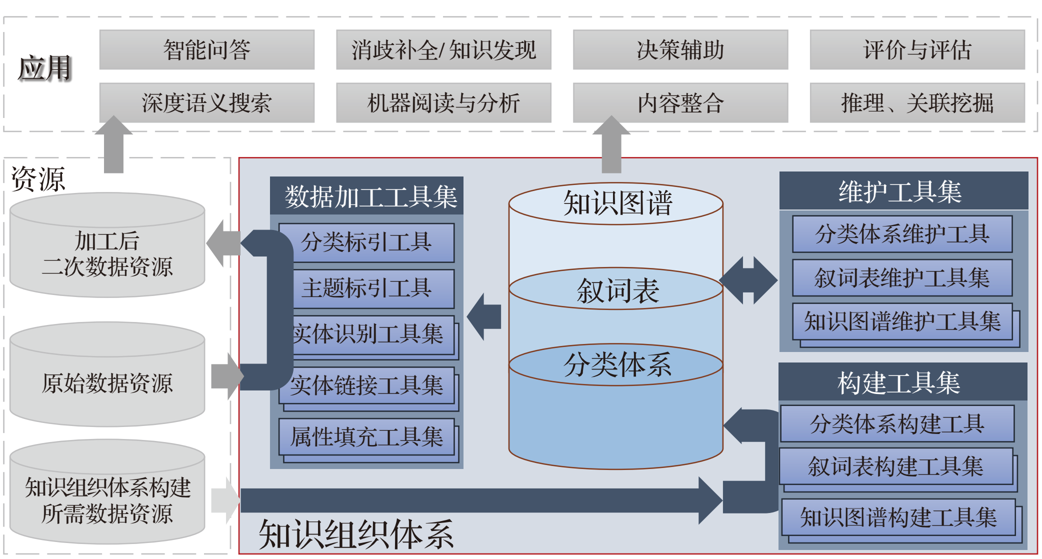

图 2 知识组织体系建设框架

知识中心结合实际需要，基于知识中心数据资源特点，建设了知识中心知识组织体系构建、维护和服务平台。一方面，知识中心各协作单位通过知识组织体系构建工具集，持续建设、更新工程科技领域知识组织体系内容（分类体系、词表、知识图谱），形成统一关联的知识组织体系内容。另一方面知识中心通过知识组织体系的数据加工工具集对不同来源的原始数据进行组织与加工，形成标引、关联后的更具价值的二次数据资源。其中，数据加工工具集中的分类标引工具、主题标引工具、实体识别工具集、实体链接工具集、属性填充工具集等将原始资源数据和建成的知识组织体系充分关联起来，为数据的打通和融合提供了工具支撑，进而能够支持包含数据融合基础的各类知识服务应用。目前，知识中心已构建了工程科技领域分类体系、约 15 万核心词的工程科技领域词表总表、近 5000 万对象的专业领域知识图谱。

通过知识组织体系建设显著提高了原始资源的加工质量和相互关联程度，还通过知识的关联和重组构建二次加工主题数据库，这部分主题数据可以提供不同需求层次的融合类知识服务。目前在数据加工工具集的建设上充分利用知识中心已经积累的经过加工的数据作为训练集合，并采用了深度学习相关算法来改进加工的效果。

3.2.3 多维度资源融合

中国工程院潘云鹤院士多次在知识中心会议中指出，数据汇集后只有通过融合打通，才可能实现数据量翻番、知识倍增和创新加速的效果。知识中心探索了多种维度的资源融合与打通。

（1）作者维度的资源融合

知识的创造主体和服务主体是都是人，大量的资源都可以通过人与人的共著关系、互引关系和共被引关系进行融合，知识中心正在通过专家库实现资源在作者维度的融合打通。知识中心在建设专家库的过程中，建立了中国名称标识符 KID(Knowledge ID) 系统。同时，为了中国名称标识符 KID 与国际名称标识符系统的互联互通，并辅助进行中外文资源的融合打通，知识中心与国际标准名称识别码（International Standard Name Identifier， 简称 ISNI）组织合作，将中国名称标识符 KID 与 ISNI 互为映射，构建了具备国际化同时兼有中国自主可控的名称标识符系统。

（2）空间维度的资源融合

空间和作者一样，也是知识资源融合的重要维度之一，大量的文献、机构、人员、产品等都可以通过空间信息进行关联融合。知识中心以文献的研究区域和文献作者的机构为位置为基础，开发和共享专题数据空间化工具，标识出相关资源作者及研究区域的空间位置信息。

知识中心通过专题数据空间化工具给地名匹配空间位置信息，并以专题可视化的方式在地图上将知识进行融合关联。图 3 是某篇论文关联推荐应用数据的案例。

3.3 知识服务实践

3.3.1 K-GO 智能搜索

用户希望以最简洁的方式获取所需要的知识，为了从来源各异的原始数据中提炼出有效的知识，知识中心初步构建了 K-GO 智能搜索

doi:10.3772/j.issn.2095-915x.2018.05.001

引擎。K-GO 智能搜索引擎内置了多元异构海量数据，提供完全自主开发的索引、性能优化、排序算法等，检索响应速度显著提升，实现毫秒级结果响应。通过集成知识组织体系内容，提升搜索的精确程度，实现了搜索提示、检索结果分类、与院士信息关联、相似内容推荐、与其他主流数据库跨库检索、轻应用检索等功能，提升了知识发现能力。

图 3 空间维度的资源融合案例

目前 K-GO 智能搜索引擎形成了以自有搜索引擎为核心，并通过对不同来源的垂直搜索引擎数据进行元数据集成、接口集成，初步实现了不同来源、不同垂直搜索引擎的数据统一揭示。

3.3.2 战略咨询智能支持

战略咨询是复杂的脑力劳动，既需要大量高水平专家团队的支持，也需要各渠道知识资源的支撑。知识中心围绕中国工程院开展战略咨询的实际需求，以文献与专利分析为核心，建成了战略咨询智能支持平台。平台通过集成技术预见、德尔菲调查、路线图等特色功能，建成了文献综述、专利分析、技术体系划分、智能制造、通用问卷等 5 个通用工具集，支撑机器人、光电子技术、航天运载器技术等 16 个课题组技术清单制定，面向 1752 名院士、专家发放网络问卷 33368 份，进行路线图绘制试点工作，平台的支撑能力初见成效。

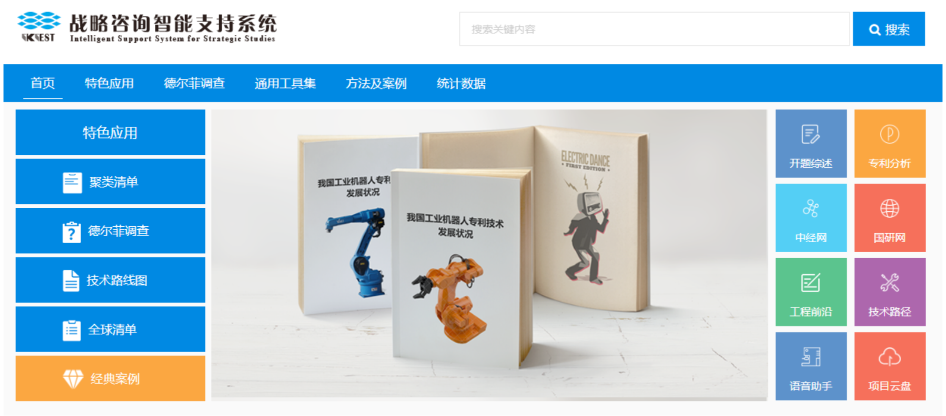

图 4 战略咨询智能支撑系统

4 总结与展望

目前大数据研究普遍存在着“只有数据、没有利用”的问题，导致搜集数据、存储数据付出的成本被浪费。以知识服务需求来驱动数据资源建设，以用户需求来驱动跨媒体、跨数据库、跨领域数据之间不同关系的发现与组织，是知识中心从建设一开始就追求的目标。未来，知识中心仍将研发智能化方法推动知识加工的深度和广度同步扩展，不断探索更多维度的资源融合，通过资源的融合打通，助力国家科技创新加速，为提高国家工程科技领域自主创新能力提供更好支撑。

参考文献

[1] 陈左宁. 中国工程科技知识中心建设发展报告 2012-2016(序) [EB/OL]. (2007-01) [2018-08-09]. http://www.ckcest.cn/aboutUs/Subcenter1.

[2] Alun P, Kit H, Alex G, et al. KRAFT: AN AGENT ARCHITECTURE FOR KNOWLEDGE FUSION[J]. International Journal of Cooperative Information Systems, 2001, 10(01n02):171-195.

[3] Smirnov A, Levashova T, Shilov N. Patterns for context-based knowledge fusion in decision support systems[J]. Information Fusion, 2015, 21(1):114-129.

[4] 刘晓娟, 李广建, 化柏林. 知识融合: 概念辨析与界说[J]. 图书情报工作, 2016(13):13-19.

[5] 张心源, 邱均平. 国内图书情报领域知识融合研究的发展与分析[J]. 数字图书馆论坛, 2016(3): 17-23.

[6] 张兴旺, 李晨晖, 麦范金. 变革中的大数据知识服务:面向大数据的信息移动推荐服务新模式[J]. 图书与情报, 2013(4):74-79.

[7] 赵蓉英, 张心源. 大数据环境对知识融合的影响研究[J]. 情报学报, 2017, 36(9):878-885.

[8] 张心源, 邱均平. 大数据环境下的知识融合框架研究[J]. 图书馆学研究, 2016(8):66-70.

[9] 唐晓波, 朱娟, 杨丰华. 大数据环境下的知识融合框架模型研究[J]. 图书馆学研究, 2016(1):32-35.

[10] Fisch D, Kalkowski E, Sick B. Knowledge Fusion for Probabilistic Generative Classifiers with Data Mining Applications[J]. IEEE Transactions on Knowledge & Data Engineering, 2014, 26(3):652-666.

[11] 陈为东, 王萍, 王益成,等. 政府网站信息资源的多维语义知识融合结构体系及策略研究[J]. 情报理论与实践, 2017, 40(6):111-116.

[12] 胡海波, 高鹏. 基于知识融合的电子政务知识服务模式及系统模型[J]. 情报工程, 2018(1):31-41.

[13] 古志文, 陈春, 吴新年. 支撑企业技术创新的知识服务模式研究--知识服务与信息服务融合发展的

doi:10.3772/j.issn.2095-915x.2018.05.001

视角[J]. 科技进步与对策, 2014(7):131-135.

[14] 刘妍, 李耘涛, 和金生. 基于知识融合的知识生态模型研究——企业创新差异视角[J]. 现代管理科学, 2016(7):100-102.

[15] Kriegel E U, Meissen U, Pfennigschmidt S, et al. A Knowledge Fusion Toolkit for Decision Making[C]. Third International Conference on Emerging Security Technologies. IEEE, 2012:90-93.

[16] 李弘. 面向知识服务的出版融合发展浅析[J]. 科技与出版, 2016(12):12-16.

[17] 陈建华. 媒体融合环境下科技期刊知识服务创新的探索[J]. 中国科技期刊研究, 2017(12):1099-1103.

[18] 陈晓峰, 云昭洁, 万贤贤. 媒体融合精准知识服务助推学术期刊供给侧改革[J]. 中国科技期刊研究, 2017(9):805-809.

[19] 郑杰, 李东鑫, 韦春伊. 媒体融合背景下图书馆知识服务[J]. 河南图书馆学刊, 2016(9):85-87.

[20] 叶翠, 刘灿姣. 数字环境下馆社知识服务融合的动因、问题及对策探究[J]. 图书馆, 2016(11):62-65.

[21] 赵瑞雪, 鲜国建, 罗婷婷,等. 中国工程科技知识中心元数据规范-(I) [M]. 北京: 中国农业科学技术出版社, 2017.