

知识自组织与他组织方法类比与融合研究^{*}

孙中秋, 陈晓美, 毕强

(吉林大学管理学院, 长春 130022)

摘要: 大数据时代, 知识资源呈现数量大、种类多、关系复杂等特征, 增加了知识组织的难度, 仅依靠知识自组织方法或知识他组织方法很难缓解这一困境。针对这一问题, 文章尝试融合知识自组织和他组织方法。首先, 对知识自组织和他组织方法进行简要介绍, 其次, 从知识组织的主客体、理念、维度、工具、方法与技术等方面比较知识自组织与他组织方法, 挖掘两者的融合点。最后, 探讨知识自组织和他组织方法融合机理, 并构建知识自组织和他组织方法融合模型。本研究不仅深化了知识组织理论, 还为知识组织实践工作提供新的方法, 具有重要意义。

关键词: 知识自组织; 知识他组织; 本体; 分众分类法; 社会网络分析

中图分类号: G254

DOI: 10.3772/j.issn.1673—2286.2014.09.04

1 引言

自1929年美国图书馆学家布利斯提出“知识组织”一词后, 在业界引起了广泛关注。按照组织力量来源, 知识组织可划分为自组织和他组织两种方式。一般将组织力来自系统内部的知识组织方式称为知识自组织。相反, 组织力来自系统外部的, 则称知识他组织^[1]。随着语义Web和网格技术等相关概念的提出, 知识组织研究在广度和深度上不断扩展, 与此同时, 知识组织方法的比较融合研究也逐渐进入学者的研究视野, 目前, 业界研究较多的知识组织方法主要有主题词表、本体 (Ontology)、社会网络分析 (SNA)、关联数据 (Linked Data) 以及分众分类等。

对于知识组织方法的比较融合研究开展较早, 相关研究主要涉及分众分类法与本体^[2]、分众分类与文献计量^[3]、关联数据与本体^[4]、主题词表与本体^[5]等。本文则从知识组织视角对知识自组织与他组织方法进行比较和融合研究, 从知识组织主客体、理念、方法技术、维度、工具五个方面对知识自组织和他组织方法进行比较分析, 在此基础上, 探讨知识自组织和知识他组织方法的融合机理和融合模型, 以期对相关研究提供有益帮助。

2 知识组织方法概述

分众分类法, 最早由信息学家Thomas Vander提出, 是适应互联网发展的新型知识自组织方法。Folksonomy由“Folk”和“Taxonomy”两个单词组合, 意为“以用户为中心的基于大众一致意见产生的分类体系”^[6], 分众分类方便灵活, 与用户认知紧密结合, 强调群体智慧和协同合作的重要性, 便于用户信息获取与交流, 但分众分类等级结构划分不明确, 准确性和时效性差, 应用范围较为狭窄。

主题词表是较为传统的知识他组织方法, 是指专家利用领域叙词表对词语以及词语之间的语义关系进行定义和组织^[7]的方法。词表通常按照既定的层级结构填充, 结构严密、等级清晰, 但扩展性差, 此外, 词表编制较为复杂, 需要大量人力和时间资源, 影响词表的时效性。社会网络分析法, 是对社会网络中行动者之间的关系进行量化研究的方法^[8]。社会网络分析法最大特征为分析重点不再是行动者 (用户、资源) 而是行动者之间的关联关系, 研究者利用专门的社会网络推断模型、程序和技术对数据进行统计研究, 揭示网络的整体结构, 挖掘关联关系。本体是指对共享概念模型的规范化、形式化和明确化

^{*}本研究得到国家自然科学基金项目“语义网络环境下数字图书馆资源多维度聚合与可视化展示研究” (编号: 71273111) 资助。

的说明, 其强调领域中的本质概念以及概念之间的关联^[9], 本体强调高度形式化和语义丰富, 构建过程必须有领域专家的参与。

3 知识自组织与他组织方法比较分析

知识自组织和他组织方法有很多, 如分众分类法、本体、关联数据、语义链、主题词表等, 其中业界研究较多的为以下四种, 见表1。表1主要从知识组织的主客体、理念、方法与技术、维度和工具等角度对知识自组织和他组织方法进行比较分析, 为下文二者融合机理研究奠定理论基础。

3.1 主客体

分众分类法的执行主体是普通网络用户, 用户可独立完成知识组织任务, 不需借助专业人员或工具的帮助。分众分类法主要用于网络信息资源组织。分众分类时效性强、形式化要求低等特点符合网络信息资源更新速度快、种类多样的特征, 但由于知识自组织方法执行主体的大众化使得知识自组织应用客体带有局限性, 即大众不具备组织专业知识资源的能力。专业人员追求高程度形式化和语义化的知识组织效果, 分众分类非形式化、语义程度低等弊端导致其不适合应用于专业知识资源组织, 知识自组织方法的适用范围较为狭窄。

主题词表、社会网络分析和本体等知识他组织方法的执行主体则是具备专业知识和专业技能领域的专家。如在数据获取和分析阶段, 要求研究人员具备计算机编程知识, 并掌握相关的软件使用技能, 技术要求高, 对于非计算机专业人士而言较为困难。主题法、社会网络分析法等知识他组织方法的应用客体包括网络信息资源、学术资源、人际关系等。知识他组织方法的高度形式化、结构严密、等级明确、语义丰富等特征迎合了众多研究人员的需求, 因此被广泛利用, 使用范围较为宽泛。

3.2 组织理念和维度

分众分类法坚持“以用户为中心”的知识组织理念, 用户可以按照自己的信息需求和情感体验对知识单元进行组织, 添加资源标识, 强调用户认知对于知识组织的重要性, 组织结果的全面性高。其从“用户、资源、标签”三个维度组织知识资源, 形成平面化的知识组织体系, 方便用户沟通与交流, 有利于知识的传递扩散, 但不能从纵深角度挖掘知识资源之间的关联关系。知识他组织方法强调“以资源或关系为中心”的知识组织理念, 根据知识本身特点或知识之间的关联关系组织知识资源, 用户不能根据自己的喜好改变知识资源间的关联关系, 组织结果的准确性高。他组织方法的知识组织维度具有层级性的特征, 如社会网络分析法从中

表1 主要知识自组织和他组织方法

类型	名称	主客体	理念	方法与技术	维度	工具
自组织方法	分众分类法	普通用户	形成以“用户为中心”的检索与标注系统	标签过滤技术、标签分类技术、标签推荐技术	用户、资源、标签	R语言、Wordle、tagCloud发生器、ImageChef等
他组织方法	主题词表	领域专家	形成结构严密、等级清晰的知识组织体系	叙词表编制技术	用、代、属、分、参	《汉语主题词表》、《医学主题词表》、《停用词表》
	社会网络分析法		数字资源网络结构化剖析, 语义关系挖掘	矩阵构造、可视化、编程	网络中心性、网络群聚性和网络关联性	Ucinet、NetDraw、Pajek
	本体		揭示资源的概念层次与语义结构	骨架法、企业建模法、七步法、数据挖掘	类、实例、属性	Protégé、Ontolingua、WebOnto

心性、群聚性、关联性三个一级维度组织知识资源，而中心性又可下设点度中心性、中介中心性等二级维度，维度指标的层级性有利于知识组织的层级化和形式化。

3.3 技术方法和工具

分众分类的主要技术为标签云，利用统计上浮原理绘制知识云图，通过知识单元大小和颜色来显示知识单元的重要程度，可视化效果好，用户能从图上简单快速且准确获取需要的信息。随着信息技术的发展，互联网上存在许多标签云生成工具，如Tagxedo、Tagcloud等，这也大大降低标签云制作的难度，知识自组织方法具有平民化特征，技术门槛较低，深受普通用户的喜爱。

知识他组织技术主要有可视化技术、词表编制技术、本体构建技术等，通过知识图谱形象揭示知识节点之间的关联关系。相较于自组织方法，他组织方法（尤其是社会网络分析和本体方法）形成的知识图谱所含内容较为复杂，不仅能揭示知识单元的重要程度（如社会网络分析法从点度中心度、中介中心度），还可挖掘知识单元之间的关联（如社会网络分析法通过聚类实现子群凝聚挖掘关联），用户需要具备专业知识并有充足的时间才能完全掌握图谱传递的信息。他组织方法借助的软件和工具主要有叙词表、可视化软件、本体构建软件等。知识他组织方法技术门槛高，普通用户很难快速掌握。

4 知识自组织与他组织方法融合研究

4.1 融合机理研究

通过上述分析可以看出，知识自组织和他组织方法实现知识组织各有利弊。自组织方法简单易用，检全率高，时效性强，适用于普通大众，但语义关联度低，结构性和等级性差。他组织方法语义丰富、高度形式化、检准率高，但时效性差、复杂难懂，不适于普通用户。可见，知识自组织与他组织方法具有互补性，存在融合可能性。本文认为知识自组织与他组织方法的融合应是根据不同需求进行的不同形式融合，不仅要集成两者的优势，也要尽量弥补两者的劣势，因此，将自组织与他组织的优劣特性放置于两条数轴上，分析自组织与他组织方法融合机理，如图1所示，分为四个象限。

知识自组织和他组织方法的融合旨在寻求“1+1>2”的效果，由图1可知，知识自组织和他组织方法融合存在以下四种情形，即图1所示的四个区域。

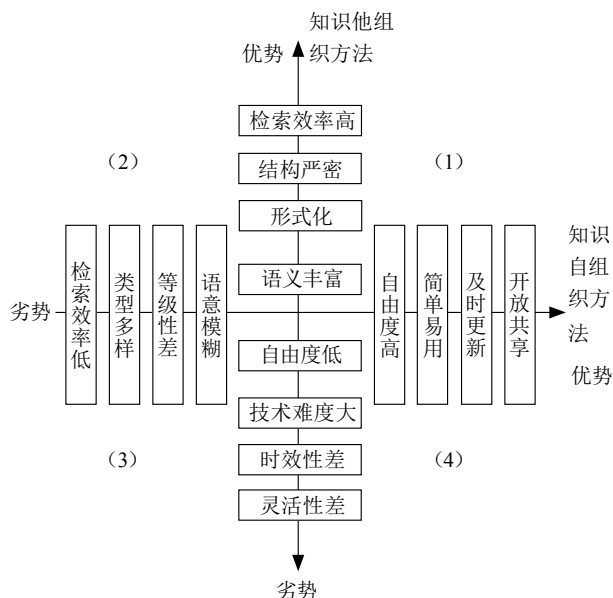


图1 知识自组织与他组织方法融合机理图

(1) 优势集中区，此种融合是双向融合，充分发挥知识自组织和他组织方法的优势。知识组织结果既富有丰富语义又便于普通用户使用，是较为理想的知识组织状态。基本融合思路为：将两者放在平等的层次上，通过计量和社会网络分析等方法聚类相关标签，映射成高度形式化和语义化的本体，可用于知识发现、知识存储和知识检索等^[10]。

(2) 利用知识他组织方法改造知识自组织方法，弥补知识自组织方法的不足，旨在追求规范化、形式化和语义丰富的知识组织效果。基本融合思路为：以分众分类法为受众，研究人员利用社会网络分析法、本体等知识他组织方法深度挖掘知识单元之间的关联关系和语义关系，对分众分类标签进行规范化和结构化处理。形式化和语义化的知识组织形式固然好，但高度形式化的组织效果很难达到，因此，此种融合最好选择折中的知识组织效果。

(3) 采用其他的方法、技术、工具和理论优化知识自组织和他组织方法。现有的知识组织方法并不完善尽美，研究人员可从深度和广度等方面继续研究，弥补现有知识组织方法的缺憾。

(4) 利用知识自组织方法改造知识他组织方法，优化知识组织效果。Folksonomy为知识他组织研究者提供新的视角，不仅有助于缓解知识他组织过程中用

户参与过少导致理论脱离实践的情况,也能充分利用群体智慧,达到知识组织的预期目的和效果。基本融合思路为:从分众分类标签中获取大量且能反映用户需求的原始标签数据集,利用知识他组织方法对标签数据进行规范化和语义化处理,使得标签资源的普适性和实用性大大加强。

从上述分析不难发现,知识自组织和他组织方法具有多方向融合的可能性,基于双方优势的融合提高了两者的应用层面,可用于解决更高层次的问题;基于一方优势的融合则有利于弥补另一方的劣势,达到预期的知识组织效果;基于双方劣势的融合则推进研究者完善相关理论,发掘新的技术方法和工具,最大限度减少两者融合带来的负面影响,使得知识组织效果接近预期的知识组织目标。

4.2 融合模型研究

本节在知识自组织和他组织方法融合机理基础上,试图探讨知识自组织和他组织方法融合模型,旨在形象展示知识自组织和他组织方法融合过程。该模型以分众分类法为基础,融合其余三种知识他组织方法,借助数据挖掘、可视化、词表编制等技术和数据库、停用词表等工具,以期最大限度地发挥知识自组织和他组织方法融合的优势。图2为知识自组织和他组织方法融合模型。

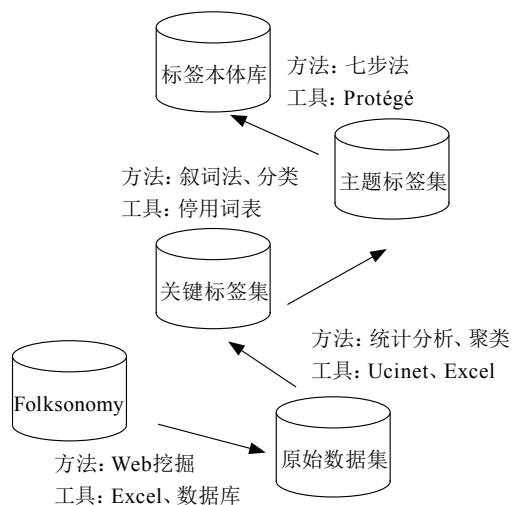


图2 自组织和他组织方法融合模型图

由图2可知,该模型以知识自组织方法的Folksonomy为基础,借用知识他组织方法弥补Folksonomy的不足。具体步骤为:利用Web页挖掘方法从社会化标注网站中

抓取标签资源,形成原始数据集,经过Excel处理,导入数据库中;利用统计分析方法,依据标签出现频次对标签进行排序,利用社会网络分析工具Ucinet对高频标签聚类,识别关键标签和标签群;利用叙词法和停用词表,确定标签主题词,建立主题标签集,对标签进行规范化处理;利用本体工具建立主题标签本体库,使得标签资源形式化、结构化和语义化。由于此种融合与图2箭头顺序相同,故本文定义此种融合是正向融合,用以区分下文的其他融合情形。正向融合最大限度地发挥知识他组织方法的优势,提高用户检索的准确性,旨在将非形式化、语义模糊的知识资源转化为规范化、语义丰富和结构化的知识资源。

上述模型并不是单一方向的融合,也可逆向利用,即借助Folksonomy的优势和“用户标注”思想弥补知识他组织方法的不足。具体步骤为:提取本体库中的知识资源的主题标签→制作网络主题词表,将规范化的主题标签和网络词汇关联起来→利用社会网络分析法挖掘关键标签→生成分众分类标签系统,向用户推荐。此种融合为逆向融合,逆向融合能最大限度发挥知识自组织方法的优点,提高用户检索的全面性,旨在将形式化和规范化的标签资源,转化为通用的、时效性强的标签资源。此外,该模型还可用于指导知识自组织方法和知识他组织方法的折中融合,即研究人员既不片面追求知识自组织的高度自由化和平民化,也不最大限度地要求知识组织效果的高度结构化、语义化和形式化,而是选择合适的界点,折中知识组织效果。

需要指出的是,本节所探讨的融合模型尝试尽可能地包含知识自组织方法和他组织方法,即涵盖上述五种融合方法,实际研究过程中,研究人员可根据自己的研究需求确定知识组织方法个数,以减轻工作量。

5 结语

知识自组织方法和他组织方法融合是知识组织发展的必然趋势。二者的相互借鉴、相互改造有利于弥补劣势、发挥优势。本文在对知识组织方法简单介绍的基础上,从知识组织主客体、理念与工具、方法和技术等视角对知识自组织和他组织方法进行比较分析,发现两者具有极大互补性,在此基础上,分析研究二者的融合机理和融合模型,总结发现,无论是正向融合、负向融合还是折中融合,知识自组织与他组织方法融合过程均可大致分为三个阶段:第一阶段为关键标签提取,

第二阶段为词汇转换,第三阶段为标签本体库构建(分众分类标签构建)。

知识自组织和他组织方法的融合将知识组织方法推向全新的应用层面,为知识自组织基础上的他组织研究和知识他组织基础上的自组织研究提供研究方法。知识自组织和他组织方法融合不仅是方法的结合,更是技术的集成、理念的渗透和研究者的互动。目前,知识组织方法融合仍处于初步研究阶段,需要克服许多研究瓶颈,如网络主题词表编制、标签规范化处理等,研究人员应从深化融合理论和方法、挖掘新的技术和工具等方面继续开展研究,推动知识自组织和他组织方法的深度融合。

参考文献

- [1] 苗东升.自组织与他组织[J].中国人民大学学报,1988(4):67-70.
- [2] GRUBER T. Ontology of folksonomy: A mash-up of apples and oranges [J]. International Journal on Semantic Web and Information Systems (IJSWIS), 2007, 3(1): 1-11.
- [3] MARKINES B, CATTUTO C, MENCZER F, et al. Evaluating similarity measures for emergent semantics of social tagging [C]// Proceedings of the 18th international conference on World Wide Web. ACM, 2009: 641-650.
- [4] 陈毅波.基于关联数据和用户本体的个性化知识服务关键技术研究[D].武汉大学,2012.
- [5] 赵东霞,赵新力.基于政务主题词表的本体构建研究[J].现代图书情报技术,2008(3):73-77.
- [6] MATHES. Folksonomies-cooperative classification and communication through shared metadata [J]. Computer Mediated Communication, 2004, 47(10): 1-13.
- [7] 贾君枝,杨洁,卫荣娟.《汉语主题词表》简单知识组织系统表示的自动转换设计[J].情报理论与实践,2011(5):54-57.
- [8] HANNEMAN R A, Riddle M. Social network analysis [M]. Riverside: University of California, 2001: 1-154.
- [9] 杨建林.基于本体的文本信息检索研究[J].情报理论与实践,2006,29(5):598-601.
- [10] 张云中.本体与自由分类法的融合机理研究[J].情报理论与实践,2012(2):35-40.

作者简介

孙中秋,女,1990生,吉林大学管理学院情报学硕士研究生,E-mail: 466386531@qq.com。
陈晓美,女,1973年生,吉林大学管理学院副教授。
毕强,男,1954年生,吉林大学管理学院教授,研究方向:知识管理。

Knowledge Self-organization and Knowledge Other-organization: Analogy and Integration

SUN ZhongQiu, CHEN XiaoMei, BI Qiang
(School of Management, Jilin University, Changchun 130022, China)

Abstract: In the era of big data, knowledge resources present features of large quantity, varieties and complex relationships, which increase the difficulty of knowledge organization, and relying knowledge self-organizing methods or knowledge other-organization methods alone is very difficult to alleviate this dilemma. Aiming at this problem, this paper attempts to integrate knowledge self-organization and knowledge other-organization methods. Firstly, it introduces knowledge self-organization and knowledge other-organization methods briefly. Secondly, it compares knowledge self-organization and knowledge other-organization methods in terms of subject and object, concepts, dimensions, tools, methods and techniques of knowledge organization, mining integration points. Finally, it discusses the fusion mechanism and building fusion model of knowledge self-organization and knowledge other-organization methods. The research not only deepens the theory of knowledge organization, but also provides a new method for practical work of knowledge organization, which has great significance.

Keywords: Knowledge self-organization; Knowledge other-organization; Ontology; Folksonomy; Social network analysis

(收稿日期: 2014-09-05)