

## ·专稿·

编者按：本期专稿领衔学者邓仲华博士，男，湖南常德人，武汉大学信息管理学院教授、博士生导师。邓仲华教授的研究集中在信息系统理论与开发、知识组织与服务。曾主持开发多项日本外包的应用系统，主持教育部人文社会科学规划项目“多语种叙词本体构建理论与自动维护模型研究”(07JA870013)、“企业集成信息平台研究与实现”、“移动化信息管理系统结构研究”，参与教育部哲学社会科学研究重大课题攻关项目“创新国家的知识信息服务体系研究”子课题(项目编号:06JZD0032)，负责其中的“知识信息服务系统功能集成研究”。出版有《电子商务系统分析与设计》、《手机应用程序开发》、《信息系统分析与设计》、《WEB 程序设计》等，发表学术论文 60 余篇。

云服务是云计算得以实现的具体形式。各种云计算的研究与开发，最终都要归结到提供更加灵活、方便、适应个性需求的服务，即所谓的云服务。因此云服务也是云计算实现的重要标志。信息资源管理与云计算的相关研究中，云服务是最需要关注而且也是最重要的领域。

本期专稿发表的一组文章，是邓仲华教授主持的人文社会科学重点研究基地重大项目“信息资源云体系与服务模型研究”(项目批准号:11JJD630001)和国家自然科学基金资助项目“云计算环境下图书馆的信息服务等级协议研究”(批准号:71173163)的部分研究成果。内容包括:《面向信息资源云服务的知识分类研究》探讨适合于云服务的知识分类组织方法;《虚拟化与信息资源云服务》分析信息资源虚拟化以实现云服务的技术问题;《图书馆云服务的参与方关系模型研究》结合图书馆云服务的特点，研究云服务中的各个参与方及其特征，提出了图书馆云服务参与方的关系模型;《美国图书馆的云服务》考察美国图书馆的云服务并分析对我国的启示。

## 面向信息资源云服务的知识分类研究 \*

邓仲华 苏 娟 李志芳 陆颖隽

(武汉大学信息资源研究中心 湖北武汉 430072)

**摘 要:**知识分类是信息资源云中重要的基础部分。文章考察了现有的典型的知识分类方法，分析总结了他们各自的特点，针对信息资源云构建过程中的特征，从用户需求和知识的使用频率稳定性两个维度出发，探讨了适合于信息资源云服务的知识分类方法。

**关键词:**信息资源 云服务 知识分类

中图分类号: G203

文献标识码: A

文章编号: 1003-6938(2012)03-0001-05

### Research on Knowledge Classification for Cloud Service of Information Resources

**Abstract** Knowledge classification is the important basis of cloud service of information resources and ways of classification are varied. This paper examines the existing typical methods of knowledge classification, analyzes and concludes each characteristic. According to the feature of cloud service of information resources and from the dimension of user's needs and the stability of the use frequency, this paper designs a knowledge classification which is suitable to cloud environment of information resources.

**Key words** information resources ; cloud service ; knowledge classification

\* 本文系教育部人文社会科学重点研究基地重大项目“信息资源云体系及服务模型研究”(批准号:11JJD630001)国家自然科学基金资助项目“云计算环境下图书馆的信息服务等级协议研究”(批准号:71173163)研究成果之一。

收稿日期: 2012-04-15, 责任编辑: 魏志鹏

知识是一个庞杂系统,必须通过适当的分类方法,才能将知识资源进行有效的组织。知识分类是知识检索及知识服务的必要条件,特别是在构造知识性云服务时,服务内容的提取既涉及到知识体系本身,又与传播过程中知识受体的背景知识相互作用。这就要求知识分类既要适应传统的知识组织,又要满足云服务所提出的新要求。

在当前云服务的发展中,通常将云服务分为三个层次:应用层服务(AaaS)、平台层服务(PaaS)和基础设施层服务(IaaS)。这种分层的概念,较好的揭示了层次上的特点,因而成为目前普遍采用的云服务层次模型<sup>[1]</sup>。但这种三层方案忽略了信息资源这一重要因素,不利于云服务体系中的资源组织与管理,影响了信息资源在云服务中的应有作用。事实上,信息资源可作为一个辅助层(独立于三层模型之外)进行单独的组织与构建,它可以对上面三层服务进行“预服务”,即进行资源的合理组织,以使上面三层在提供服务时更高效<sup>[2]</sup>。

信息资源层要对各层提供支持,所以要根据各层的特点采用合适的知识分类方法。各种知识分类方法并没有好坏(或优劣)之分,而只有适合的程序有高有低,在云服务中这种表现尤为明显。因此分析它们的特点和适用的场合就是信息资源云的构建研究的重要任务。

## 1 现有的典型的知识分类

知识分类是根据知识管理的需要和标准,通过比较,把人类的全部知识按照相同、相异、相关等属性划分成为不同类别的知识体系,以此显示其在知识整体中的应有位置和相互关系。知识分类可作为寻求知识的出发点,它犹如学海的航标,既可指明求学的门径,也可为从事知识管理的人提供理论指导,以便使纷繁复杂的知识得以依类分理,各有归属,从而建立起次序化、规范化、系统化的知识世界<sup>[3]</sup>。

各种知识分类方式都有其明确的目的,如隐性和显性知识的分类方案就是为了研究知识的转移与传播而提出的。从应用目标的视角分析,知识分类方式可以归为三种,它们分别是:揭示知识本质的分类、用于管理/检索的分类以及用于知识运动规律研究的分类。

### 1.1 关于知识本质的分类

很多大家,包括孔子、柏拉图、亚里士多德、黑格

尔、文德尔班、培根、笛卡尔、恩格斯、钱学森等等,都对知识的本质和分类作过探讨。

知识起源于人们的日常生活需要。中外许多大家在研究知识本质的分类时,虽显现出不同的分类方式,但是都体现出了满足人们日常生活需要的这一特性。儒家的杰出代表孔子在讲学时就已经采用了分科授徒的方式,体现其经世致用的思想特征。西方哲学家亚里士多德对知识的分类(理论知识、实用知识和创造知识)也是源于人类的实践活动,其目的同样是为了解决人类在客观世界中所出现的问题。

### 1.2 用于管理/检索的分类

知识的积累,必须有相应的管理手段。在图书馆学和教育研究学界,这种用于管理而产生的分类方式体现的尤为明显。

在图书馆学中,主要是进行图书的分类。图书分类法以其载体类型分为传统文献分类法(受控)和网络信息分类法(不受控),传统文献分类法以纸质图书为媒介,而网络信息分类法以网络为媒介。传统的文献分类法,大多是以学科类目为基准,以文献搜索查找为目的。而网络信息分类法是以网络信息资源为对象,以主题为聚类标准的分类方法。通过对部分网络信息的标引,建立网络信息分类导航系统,提供浏览式的分类查询手段<sup>[4]</sup>。

在教育研究学界中,无论是以教育为目的的分类法(事实性知识、概念性知识、程序性知识和原认识知识),还是以研究为目的的分类法(知识二分法、学科三分法等),都是为了对知识进行更系统的管理,从而进一步实现教育研究的目标。

### 1.3 关于知识运动的分类

波兰尼按知识的形态将知识分为言传知识和意会知识(即显性知识和隐性知识),这种分类方式依照了知识的运动规律,把人自身的活动看作知识动态生成和表达的源泉,揭示了人的内在知识底蕴与人类活动的关系。简而言之,关于知识运动形态的种种分类,目的是为了研究知识的转移与传播,加强人类对知识的交流和共享。现代,人类的实践生产活动中,特别是涉及经济领域中,关于知识的分类大多是以知识的运动形态为基础,进行进一步的实际性划分。

## 2 相关文献管理的分类法

### 2.1 体系分类法

体系分类法是一种直接体现知识分类的等级概

念标识系统,是通过对概括文献信息内容及某些外表特征的概念进行逻辑分类(划分与概括)和系统排列而构成的<sup>[4]</sup>。体系分类法的主要特点是按学科、专业集中文献,并从知识分类角度揭示各类文献在内容上的区别和联系,提供从学科分类检索文献信息的途径。

《七略》:汉代的刘向、刘歆根据当时的国家藏书完成的图书分类目录。它基本上是按照图书的内容性质来划分的,综合性的系统反映藏书的分类目录,共有六大类三十八小类。我国图书馆学界普遍认为《七略》是我国第一部图书分类法<sup>[5]</sup>。

《杜威十进分类法》(Dewey Decimal Classification, DDC):是图书馆广为使用的分类法,这个分类系统最早在1837年时Melvil Dewey有此分类构想,于1876年正式出版。DDC是用传统的学科来分类,总共以10个主要的学科(main classes)来涵括所有的知识体系,每个大类下细分10类(divisions),接着又再分成10小类(sections)。它的十个大类(main classes)分别是:总论、哲学、宗教、社会科学、语言、自然科学和数学、技术(应用科学)、艺术、美术和装饰艺术、文学、地理、历史及辅助学科<sup>[6]</sup>。

《中国图书馆图书分类法》:简称《中图法》,是目前中国大陆地区图书情报界使用最广的综合性分类法。它把全部图书分为马克思主义/列宁主义/毛泽东思想、哲学知识、社会科学知识、自然科学知识、综合性图书知识5大部类,总共22个大类<sup>[7]</sup>。

《中图科学院图书馆图书分类法》:简称《科图法》,是对图书的一种分类方法,将知识门类分为5大部分,在这5个基本部类序列的基础上,组成了25个基本大类。

## 2.2 分面分类法

分面分类法,也称组配分类法或分析-综合分类法等。其构成原理为:主要依据概念的分析与综合原理,将概括文献、信息、事务的主题概念组成“知识大纲-分面-亚面-类目”的结构,按照一定的组配规则,通过各个分面类目之间的组配来表达文献主题的一种分类法。阮冈纳赞的《冒号分类法》、布朗的《主题分类法》、布利斯的《布利斯书目分类法(2版)》等是分面分类法的代表。

## 2.3 半分面分类法

半分面分类法,亦称体系-组配分类法,是介于体系分类法与分面分类法之间的一种分类法。它基

本上是体系分类法大量采用分面分类方法的结果,即大量使用各种通用复分表、专用复分表、类目仿分以及组配符号等,并且使分类号尽量保持分段的组配形式,在详细列举类目结构的基础上,广泛喜欢用各种组配方式编制的分类法。这种分类法具有直观性较强的特点,也有较强的文献主题概念的组配功能。其主要代表为《国际十进分类法》<sup>[8]</sup>。

## 2.4 网络信息分类法

网络信息分类法必须要满足网络信息组织的需要、满足网络用户信息查询的需要、适应网络技术环境的特点。其分类的对象类型多种多样,包括文本、图像、多媒体、网站等等。网络信息分类法的基本功能是通过部分网络信息的标引,建立网络信息分类导航系统,提供浏览式的分类查询手段。典型代表为分众分类法。

分众分类的英文原名“FOLKSONOMY”一词公认是由著名的信息构建专家Thomas Vander Wal所创,以“Folk”(民间的)和“Taxonomy”(分类体系)所组成。因此,通俗的讲,“FOLKSONOMY”指的就是平常百姓的分类管理系统,国内通常又翻译成“大众分类法”、“自由分类法”、“通俗分类法”、“群落分类法”等<sup>[9]</sup>。其特点为平面性、主观性、社会性、时效性。目前在社会化书签、博客、网络社区以及图片视频分享网站得到了广泛的应用<sup>[10]</sup>。

## 3 分类方法对比研究总结以及在信息资源云环境下的适用性

没有哪一种方法是适用于所有情况的,分类方法都是根据当下所要达到的目标为导向来进行的。

(1)关于知识本质的分类方法。关于知识本质的分类大多源于哲学分类。哲学分类是很多分类方法的基础,特别是在一些应用场景(领域)分类中。正是由于其基础性的地位,所以它的分类方法较为宏观,不够细化。

(2)关于管理/检索的分类方法。传统的文献分类法的设置大都是以学科类目为基准,在文献搜索查找中是具有优势的。但是如果放在网络环境或信息资源云环境下,就不一定适用。哥伦比亚大学图书馆参考馆员Ellyssa Kroski在2005年就指出,图书馆自上而下的阶层式分类法的优点是毫无疑问的,但是对网络资源进行分类显然不适宜<sup>[11]</sup>。例如:用户现在需要一个资源,涉及到数学、物理、经济和宗教

四个方面的知识,这时用户需要的是一种平面性的查找。如果按照文献分类的方法去进行查找,自上而下,层层递进,这样就给用户的使用造成了负担,而且不能最大限度的满足用户的需求。这样有悖于云服务的目的。

网络信息分类法,是以网络资源为分类对象,这一点和信息资源云环境下的分类是有共同点的,都是对资源的分类,而且都与网络相关。近年来,影响最大的是分众分类法。

分众分类以标签(TAG)的方式进行分类,通过标签的字号大小以及颜色的不同来表示其“受欢迎”的程度,能够实现快速而自动的信息聚类,生成系列加权列表和标签云(Tag Cloud),具有较强的大众趋同性、直观性和易用性<sup>[12]</sup>。分众分类所适用的环境以及它“标签”的概念,与信息资源云环境下的知识分类的应用场景相似,但存在一些问题:①分众分类法在使用过程中是“多对一”的映射关系,多个用户需求一项资源。而信息资源云环境下的分类所反映的是“一对多”的映射,因为是一个用户需求多项资源。②分众分类法在使用中的优点,是组织用户之间互相推荐和共享信息。而目前的云计算基本上还是私有云为主导,各个用户之间的资源有各自的私密性。

(3)关于知识运动形态的分类方法。关于知识运动形态的分类主要体现在具体的应用场景中。大多是以实际场景中的具体需求为导向,在一个领域内分类。这种分类方法不具有普适性,而且基本上是在传统环境中进行的,与信息资源云环境各个方面相去甚远,所以不适用于信息资源云环境。但是也有其可取性,如,云计算是一种商业服务模式,目前的应用主要是企业用户为对象,那么经济场景中的一些分类的方式可以为它提供一些参考。

每一种分类都有其优缺点,没有绝对完美的划分,也不存在好或坏。我们一个新的环境重新进行资源知识分类时,要根据所在场景的具体要求去处理、分划。

#### 4 面向用户需求的信息资源云环境下的知识分类

信息资源云环境下的知识分类有两点需要把握。一是要符合其应用场景,以寻求一种适合的知识类型的划分,这体现了用户对知识类型的需求。二是要找寻一种适用于网络环境的快速便捷的分类方

法,以提高其服务的有效性,这体现了用户对服务质量的需求。

曾伟生曾从投入和产出两个方面对知识进行分类,按知识关键度和知识可用度这两个维度把知识划分成为了四类:发展类知识;生存类知识;一般类知识;边缘类知识<sup>[13]</sup>。另一种基于知识的使用频率和稳定性差异,把知识分为常规性知识和非常规性知识<sup>[14]</sup>。这两种分类方案可作为我们在知识类型划分上的重要参考。

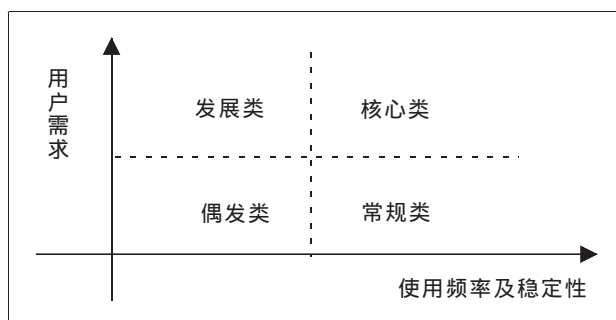


图1 面向信息资源云服务的知识分类方法

分众分类法的标签TAG也是使用频率的一个体现,且它是在网络环境下的分类方法。用分众分类法的部分可取的手段,从“知识使用频率及稳定性”这个维度进行类别划分,提供给用户快捷方便的服务。

在此,从用户需求和知识使用频率及稳定性这两个维度出发,参考曾伟生分类以及分众分类法的标签TAG,我们可以把知识分为四类(见图1)。

##### 4.1 内容划分

偶发类知识:使用频率低并且因时、因地、因认知客体而异的知识,用户需求度不高。这类知识不是人们有意识获取的知识,通常是由于偶然,或无意识地保留下来的知识。

常规类知识:此类知识已经形成了一定的规范化和程序化,稳定性很高,在用户群中也形成了约定俗成的概念,因此需求度并不高。

发展类知识:此类知识可划分到隐性知识一类,稳定性不高,但有导向作用,价值性较高。

核心类知识:此类知识是用户目前最需求的一部分。

##### 4.2 类与类之间的区别

用户高频使用,但并未进行深度挖掘的,属于成熟的,并且约定俗成的归为常规类。同样高频使用,但是在不同维度进行了深度挖掘,则归为核心类。

用户使用频率低,并且显示出只是一种表面层次的使用的归为偶发类。同样使用频率低,但是通过TAG的关联关系,显示出用户对资源有挖掘的,则称之为发展类。

#### 4.3 特点优势

类别的划分并不是一成不变的,它具有扩展性,流动性,更替性。扩展性适应了信息资源云的主要特征。流动性较好地反映了知识的游离性质,例如核心类的知识,由于知识的更新和扩展,原属于这一部分里的知识需求度降低或者说是使用频率降低,它就会流动到常规类或发展类之中去。而发展中的知识,在用户的使用过程中,随着它使用频率的升高以及进一步的知识挖掘,就流动到核心类。更替性体现在,当一类知识它的使用率越来越低,直到一个值的时候,它会被替代,而不存在于这些分类当中。

#### 4.4 缺陷及解决办法

这种理论模式在实际的使用中,需要引入一些参数。例如影响因子、权重等,来比较清晰客观的进行处理。另外还需要克服TAG辨识的模糊性。例如一词多义的时候,TAG就会有判断上的错误。这需要利用语义概念,引入同义词控制技术来进行优化<sup>[15]</sup>,以达到用户使用的最优效果。同义词控制系统采用基于语义的人工智能技术,系统基于语义分析方法自动对相同或者是相近语义的标签进行分类统计,将语义相近或相同的关键字合作,再进行使用频率统计。更高阶的同义词控制的目标就是智能化地识别TAG的低相关度的问题,采用语义近似或相同的其他关键字来避免标签的模糊性,保证分类的清晰有用。目前基于XML的同义词控制还不能完全解决标签的模糊问题,在未来的语义平台上,有望实现更加智能化的同义控制。

### 5 结语

随着技术的进步,云计算的影响越来越深,信息资源管理的发展需要传统的分类方法作出改变来适应云计算的深入。以用户需求和知识的使用频率稳定性两个维度出发,提出的知识分类方法是适应这种改变的需要。文中分析为这种分类方法的优点、缺陷及应对方案,但还有很多不完善之处。随着技术的不断加强,如何更快的解决这些不完善的地方,这是值得我们研究探讨的问题。

#### 参考文献:

- [1]刘鹏.云计算[M].北京:电子工业出版社,2011:1-3.
- [2]ZhongHua Deng,YongBo Liu, YouLin Zhao. Analysis on Integration Technology for Information Resources Cloud [C]. 2011 International Conference on Fuzzy Systems and Neural Computing (FSNC 2011).
- [3]陈洪澜.论知识分类的十大方式[J].科学学研究,2007,(1):26-31.
- [4]张丽英.中美图书分类法比较研究—从历史沿革、编订机制、应用现状与发展趋势角度出发[D].福州:福建师范大学社会历史学院,2008.
- [5]周若立,陈树年.文献分类学[M].武汉:武汉大学出版社,2001:32.
- [6]张燕飞.中外现代主要图书分类法之比较[J].图书情报知识,1990,(4):9.
- [7]穆安民.科技文献检索[M].重庆:重庆大学出版社,1996:22.
- [8]司莉.信息组织方法与原理[M].武汉:武汉大学出版社,2011:15.
- [9]中国图书馆学会.中国图书馆学会2007年年会论文汇编[C].2007.
- [10]刘洋.大众分类法的应用现状及前景分析[J].现代经济信息,2010,(5):205-206.
- [11]EllyssaKroski.The-hive-mind:-folksonomies-and-user-based-tagging [EB/OL]. [2010-02-10].<http://infotangl-e.blogspot.com/2005/10/07/the-hive-mind-folksonomies-and-user-based-tagging/>.
- [12]朱婷.情报中的序性结构初探[J].图书情报知识,2008,(3):9-11.
- [13]曾伟生.企业知识分类与管理策略[D].北京:北京交通大学经管学院,2007.
- [14]陈洪澜.知识分类与知识资源认识论[M].北京:人民出版社,2008:140-186.
- [15]周荣庭,郑彬.网络时代的新型信息分类法[J].现代图书情报技术,2006,(3):73-74.

作者简介:邓仲华(1957-),男,博士,武汉大学信息管理学院教授,博士生导师;苏娟(1985-),男,武汉大学信息管理学院研究生;李志芳(1987-),女,武汉大学信息管理学院研究生;陆颖隽(1963-),男,武汉大学信息管理学院副教授,硕士生导师。