# 基于数据支持的学科评估实践*

## ——东华大学2015年SCI-E收录期刊分区数据监测分析

陈惠兰，董政娥，任瑞荣

（东华大学图书馆，上海　201620）

**摘要：**针对大数据、云服务和“互联网 +”特点，绘制了由数据源层、数据仓库、数据处理层（数据分析和数据挖掘）和服务层四部分构成的基于数据支持的学科评估路径。并以东华大学 2015 年 SCI-E 收录期刊分区数据监测为案例，基于数据探讨了学科评估问题，以期为大数据时代决策服务提供理论基础和数据支持，最终突破学科服务发展瓶颈。

**关键词：**学科评估；期刊分区；数据支持；决策服务

**中图分类号：**G647，G256　　**文献标识码：**A　　**文章编号：**2095-3380（2017）01-0013-06

# Subject Evaluation Practice Based on Data Support

## —A Case Study of Prediction on Quartile in Category Indexed by Science Citation Index Expanded in 2015 at Donghua University

Chen Huilan, Dong Zheng'e, Ren Ruirong

（Library, Donghua University, Shanghai 201620）

**Abstract:** Based on features of big data, cloud services and the Internet+, the path of the subject evaluation based on data analysis which consists of four parts such as data source layer, data house, data processing layer（data analysis and data mining）and decision service layer was drawn. Prediction on quartile in category indexed by Science Citation Index Expanded in 2015 at Donghua University was taken as an example, in which the subject evaluation was practiced based on the data to provide a theoretical basis and data support for decision service under the big data era, and to breakthrough bottleneck of subject service finally.

**Keywords:** Subject evaluation; Quartile in category; Data support; Decision service

## 一、前言

学科服务作为图书馆创新服务的最大亮点被寄予了改变图书馆地位和形象的使命。[1]尽管我国最早于 1998 年首创学科馆员制度开展学科服务，但是至今各图书馆的学科服务均处于探索期，没有现成的模

* 基金项目：教育部人文社会科学规划项目“纺织史研究资源保障及其服务体系构建研究”（13YJA870001）；上海市高等教育学会项目“大数据运用于高校发展研究”（ZCGJ07-15）；中央高校基本科研业务费专项资金资助项目“大数据时代图书馆基于 ESI、InCites 和 DII 数据库支持学科发展的服务路径实践研究”（14D123702）；东华大学终身教育研究所 2015 年度课题研究项目（2015LE0092）；ISTIC-ELSEVIER 期刊评价研究中心开放基金项目“基于 SCOPUS 和 SCIVAL 的决策支持服务模式研究”。

**收稿日期：**2016-11-08，**修回日期：**2016-12-08

式可以遵循,也没有既定的经验可以借鉴。当前,大数据、云服务和“互联网 +”作为数字化时代的重要概念、技术和方法,正在社会经济领域快速渗透,日益强烈地影响着人们的观念、生活和决策选择。在“大众创业、万众创新”被视作中国新常态下经济发展“双引擎”指引下,各行各业均在积极寻求升级变革。同样,大数据、云服务和“互联网 +”将为图书馆突破学科服务瓶颈提供新思路、新技术和新方法。[2]

大数据时代,在商业、经济及其他领域中,决策将日益基于数据和分析而做出,而并非基于经验和直觉。[3]因此,数据驾驭能力向每一位学科服务人员(或学科馆员)提出了新的挑战,如何通过对海量数据(结构化和非结构化)的挖掘和运用,为学科发展和高校发展决策提供数据支持和理论依据成了图书馆未来发展的方向之一。而“互联网 +”时代是一个靠数据说话的时代,面对大数据和互联网的挑战,学科服务必须进行角色转变:以资源建设和推送为中心转变为以数据服务为中心支持决策,即通过大数据、云服务加强对用户的研究,开展知识关联服务、数据支持和精准服务。为此,本文在探讨基于数据支持的学科服务路径基础上,以东华大学 2015 年 SCI-E 收录期刊分区数据监测为例,探讨和实践了决策支持服务。

## 二、基于数据支持的学科服务路径

如何驾驭大数据?如何让海量数据转化为提高高校领导层决策科学性的工具?基于此,图 1 绘制了基于数据支持的学科服务路径,具体可分为数据源层、数据仓库、数据处理层(数据分析和数据挖掘)和服务四部分,具体路径如图 1 所示。

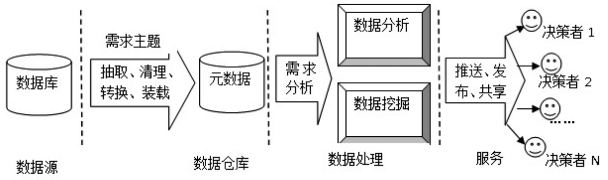


图 1 基于数据支持的学科服务路径图

Ⅰ层(数据源):从各种信息源获取数据信息,图书馆作为信息资源集散地,各种馆藏资源包括数字资源和纸质资源均为数据支持服务提供了数据源泉。

Ⅱ层(数据仓库,Data House):多种来源异构数据的存在使得预处理成为必要环节,它使数据格式规范化进而转化成统一的数据模式,这是数据支持服务的首要任务。数据仓库是一个面向主题、集成、相对稳定、随时间不断变化的数据集合。[4]其主要用于支持管理中的决策制定过程,并通过对数据进行深层挖掘、多维数据分析、动态查询报表等帮助有关人员制定改进工作流程的决策。

Ⅲ层(数据处理):这一层是建立在数据库基础上的。针对决策者需求,主要以数据分析和数据挖掘(Data Mining)为主,即利用各种分析挖掘工具进行多角度、多层次分析,从中发现变化趋势,挖掘隐藏在大量数据背后的潜在规则和知识,最终基于事实数据辅助决策者做出正确的决策。

Ⅳ层(支持服务):针对不同决策者,可以将数据分析和数据挖掘结果通过不同媒介进行共享、发布和推送等,实现决策支持服务。

## 三、SCI-E 收录期刊分区视角下东华大学 2015 年的学科发展

1. 期刊分区总体情况

以“donghua”为检索策略在地址字段检索了 2015 年东华大学被 SCI-E 收录论文情况,检索表明,截至 2015 年 11 月 20 日,东华大学共有 893 篇论文被 SCI-E 收录(表 2)。由表 1 可见,东华大学 2015 年被 SCI-E 收录期刊小类分区中,1 区、2 区、3 区和 4 区的期刊百分比为 13∶23∶27∶37;大类分区中,对应百分比为 11∶25∶26∶39,也就是说无论按照大区分类

表 1 东华大学 2015 年 SCI-E 收录期刊分区分布情况

| | | 期刊数目 | 百分比 /% |
|---|---|---|---|
| 小类分区 | 1 区 | 47 | 13 |
| | 2 区 | 79 | 23 |
| | 3 区 | 96 | 27 |
| | 4 区 | 129 | 37 |
| | 合计 | 351 | |
| 大类分区 | 1 区 | 42 | 11 |
| | 2 区 | 96 | 25 |
| | 3 区 | 101 | 26 |
| | 4 区 | 150 | 39 |
| | 合计 | 389 | |
| Top 期刊 | 是 | 63 | 17 |
| | 否 | 307 | 83 |
| | 合计 | 370 | |

注:在小类分区统计中,如果一个期刊在不同学科领域有不同分区,则以最高分区计算

表 2 东华大学 2015 年 SCI-E 收录期刊各小类分区分布情况

| 所属小类 | 论文数 | 期刊种数 | 大类分区 | | | | Top 期刊 | |
|---|---|---|---|---|---|---|---|---|
| | | | 1 区 | 2 区 | 3 区 | 4 区 | Y | N |
| 1 区 | 105 | 47 | 22 | 22 | 3 | | 25 | 20 |
| 2 区 | 263 | 79 | 15 | 44 | 18 | | 26 | 49 |
| 3 区 | 254 | 96 | | 24 | 42 | 24 | 7 | 78 |
| 4 区 | 223 | 129 | | | 20 | 109 | | 122 |
| 合计 | 845 | 351 | 36 | 90 | 83 | 133 | 58 | 269 |
| 总计 | 893 | 353 | 42 | 96 | 101 | 150 | 63 | 307 |

还是小区分类，分布在 1 区的期刊均只有总期刊的 13%左右，分布在 4 区的期刊则占总期刊的百分比最高,小类分区和大类分区分别为 37%和 39%;分布在 2 区和 3 区的期刊则基本相当,而 2 区和 3 区期刊总数则是总期刊数的一半以上，小类分区和大类分区分别为 50%和 51%。在这些期刊中 17%的期刊为 TOP 期刊。

2. 小类 1 区期刊及论文分布

2015 年，东华大学共有 105 篇论文被 47 个小类 1 区期刊收录，而在这 47 种期刊中,22 种期刊同时分布在大类 1 区中。如图 2 所示,对于这 105 篇被收录论文,26 种期刊均只收录论文 1 篇,10 种期刊收录论文 2 篇,4 种期刊收录 3 篇,2 种期刊分别收录 4 篇和 6 篇论文,3 种期刊分别收录论文 5 篇、8 篇和 14 篇。

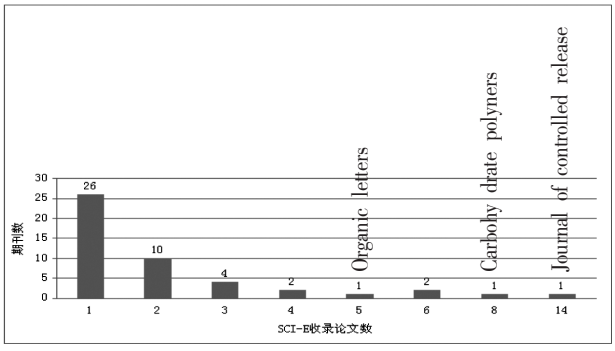


图 2　小类 1 区论文及其期刊分布情况

在学科分布方面，被收录论文覆盖 34 个小类学科,学科聚类(如图 3)表明,论文分布最多的学科依次为化学、工程、材料和药学,分别为 31 篇、23 篇、16 篇和 14 篇。

在学科聚类过程中,如图 4 所示,子学科最多的为材料、工程和化学,具体如下:

(1)在材料科学领域,综合学科、纸与木材和生物材料论文分布最多，分别占该学科论文总数的 44%、25%和 13%,其次为硅酸盐、复合材料和纺织材料;

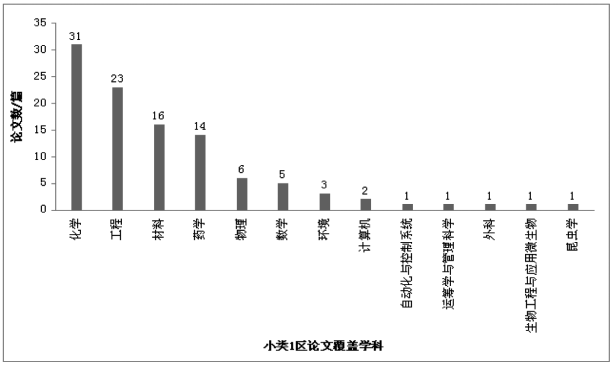


图 3　小类 1 区论文涉及学科分布情况

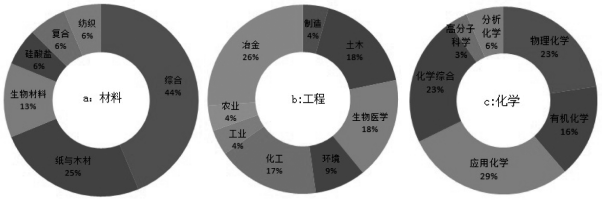


图 4　小类 1 区论文涉及较集中学科分布情况

(2)在工程学科领域,论文分布在最前面的分别为冶金(26%)、土木(18%)、生物医学(18%)、化工(17%)，其次还涉及 9%的环境工程和均为 4%的制造、工业、农业;

(3)在化学学科领域,论文分布最多的分别为应用化学、化学综合、物理化学和有机化学,分别占该学科论文总数的 29%、23%、23%和 16%;其次涉及 6%的分析化学和 3%的高分子科学;

(4)其他:在计算机学科,主要涉及硬件和控制论;在环境学科,主要涉及环境科学和水资源;物理学科,主要涉及热力学、声学和力学;而数学主要以应用数学为主。

3. 小类 2 区期刊及论文分布

2015 年东华大学被 SCI-E 收录的 893 篇论文中,262 篇论文分布在小类 2 区，而这 263 篇论文分布在 79 种期刊上,其分别分布在大类 1 区、2 区和 3 区,涉及的期刊数分别为 15 篇、44 篇和 18 篇。而这些期刊有 26 种期刊为 TOP 期刊，另外 49 种则不是 TOP 期刊。如图 5 所示,分布在小类 2 区的 79 种期刊中,37 种期刊均只收录论文 1 篇,16 种期刊收录 2 篇,8 种期刊收录 3 篇,3 种期刊分别各收录 4 篇、5 篇和 6 篇论文,2 种期刊收录了论文 7 篇,收录论文 8 篇、9 篇、10 篇、13 篇、23 篇的各有 1 种;*Electrochimica Acta* 和 *Textile Research Journal* 均收录论文 24 篇。

在学科分布方面,共涉及 15 个学科(如图 6),其中,收录论文最多的学科分别为材料科学(88 篇)、化

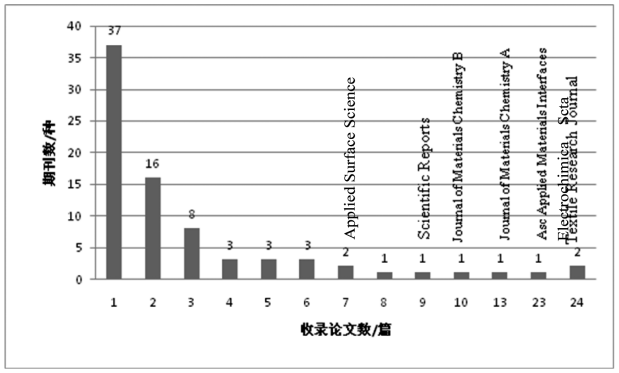


图 5 小类 2 区论文及其期刊分布情况

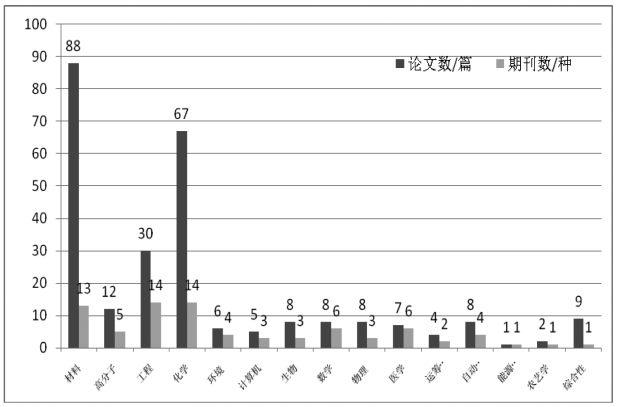


图 6 小类 2 区论文涉及学科分布情况

学(67 篇)和工程(30 篇)。

对于这些学科，每一学科均有自身特点，如图 7 所示,具体如下：

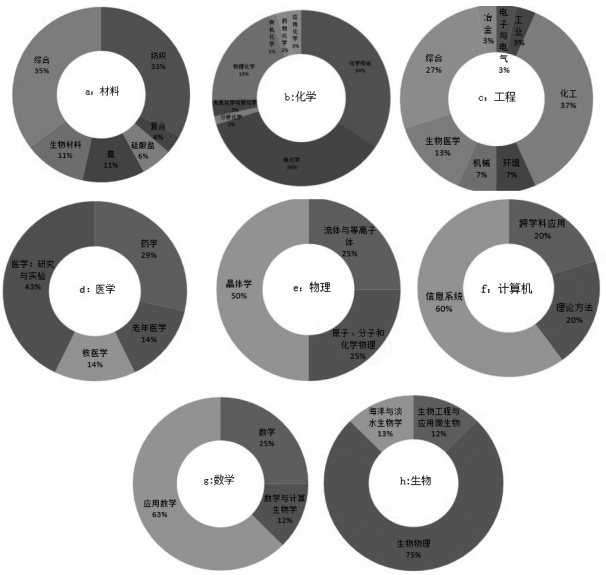


图 7 小类 2 区论文涉及较集中学科分布情况

(1)如图 7(a)所示,材料学科除综合学科外,主要以纺织材料、生物材料和膜材料为主,分别占该学科的 33%、11%和 11%,此外,还涉及 6%的硅酸盐和 4%的复合材料；

(2)对于如图 7(b)所示的化学科学来讲,其最多的分别为电化学(36%)化学综合(34%)和物理化学(19%)，其次为无机化学与核化学（3%)、应用化学(3%)、分析化学(2%)、药物化学(2%)和有机化学(1%)；

(3)如图 7(c)所示,工程学科主要以化工(37%)、综合(27%)、生物医学(13%)为主,其次是环境(7%)和机械(7%),此外,还涉及工业(3%)、电子与电气(3)和冶金(3%)；

(4)如图 7(d)所示,医学主要以 43%的医学研究与实验和 29%的医学为主,另外,还涉及 14%的核医学和老年医学；

(5)对于物理学科(如图 7e)来讲,除 50%的晶体学外，还分别涉及 25%的流体与等离子体和原子、分子和化学物理；

(6)对于计算机学科(如图 7f),除 60%的信息系统外,还分别涉及 20%的理论方法和跨学科应用；

(7)数学学科涉及 63%的应用数学、25%的数学和 12%的数学与计算生物学；

(8)生物学科涉及 75%的生物物理、13%的海洋与淡水生物学和 12%的生物工程与应用微生物。

4. 小类 3 区期刊及论文分布

2015 年,东华大学共有 254 篇论文被 96 个小类 3 区期刊收录，而这 96 种期刊分别分布在大类 2 区、3 区和 4 区,所占比例为 24:42:24;此外,这些期刊有 7 种期刊为 TOP 期刊，其余 78 种期刊为非 TOP 期刊。如图 8 所示，对于这 254 篇被收录论文,55 种期刊均只收录论文 1 篇,13 种期刊收录论文 2 篇,10 种期刊收录 3 篇,5 种期刊收录 4 篇,2 种期刊收录 5 篇论文,6 种期刊分别收录论文 6 篇、7 篇、8 篇、10 篇、20 篇和 62 篇。

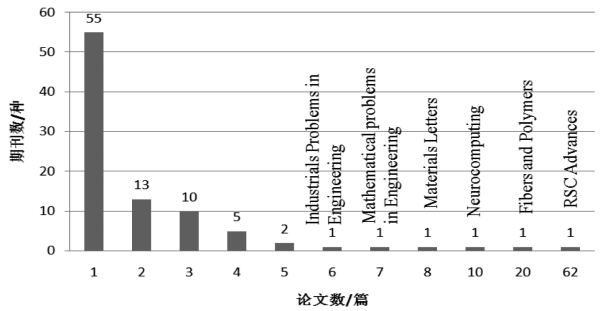


图 8 小类 3 区论文及其期刊分布情况

在学科分布方面,经聚类得到如图 9 所示学科及

其收录论文分布,分析可见,化学、材料、工程和物理收录论文居前;其次是计算机、生物和数学;此外还涉及环境、综合和管理科学。

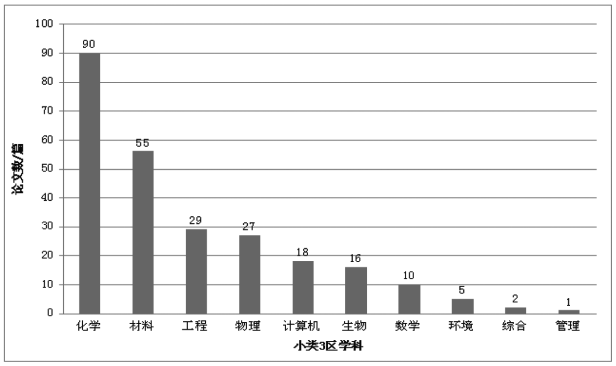


图 9　小类 3 区论文涉及学科分布情况

为了进一步剖析各细分学科分布情况，图 10 绘制了各学科的分布图。具体如下：

(1)由图 10(a)可见,化学学科中化学综合论文分布最多,为 75%;其次为高分子科学和有机化学,论文分布分别为 12%和 8%;其次是 2%的应用化学、2%无机化学与核化学以及 1%的电化学;

(2)由图 10(b)可见,材料学科种纺织和综合论文分布最多,分别为 41%和 36%;其次为生物材料、复合材料和硅酸盐,分别为 9%、7%和 5%;此外,还涉及 2%的纸与木材;

(3)如图 10(c)所示,工程学科主要以综合、生物医学、化工、电子与电气和土木工程为主，分别为 28%、28%、27%、14%和 3%;

(4) 如图 10(d)所示,物理学科主要以物理化学、力学和应用为主,分别分布论文 37%、26%和 22%;其次为 8%的流体与等离子体和 7%的综合学科;

(5)如图 10(e)所示,计算机学科主要覆盖 56%的人工智能、22%的自动化与控制系统、17%的理论方法和 5%的硬件;

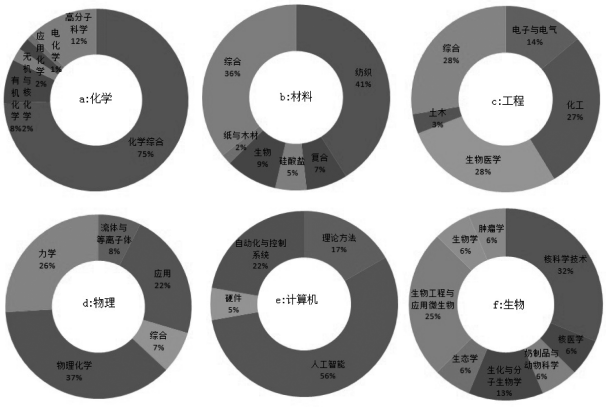


图 10　小类 3 区论文涉及较集中学科分布情况

(6)如图 10(f)所示,生物学科主要涉及核科学技术(32%)、核医学(6%)、奶制品与动物科学(6%)、生化与分子生物学(13%)、生态学(6%)、生物工程与应用微生物(25%)、生物学(6%)和肿瘤学(6%)。

5. 小类 4 区期刊及论文分布

2015 年,东华大学共有 223 篇论文被 129 个小类 4 区期刊收录，而这 129 种期刊分别分布在大类 3 区和 4 区,所占比例为 20∶109。如图 11 所示,对于这 223 篇被收录论文,93 种期刊均只收录论文 1 篇,18 种期刊收录论文 2 篇,7 种期刊收录 3 篇,2 种期刊收录 4 篇,3 种期刊收录 5 篇论文，收录论文 6 篇、8 篇、15 篇和 21 篇的各有 1 种。

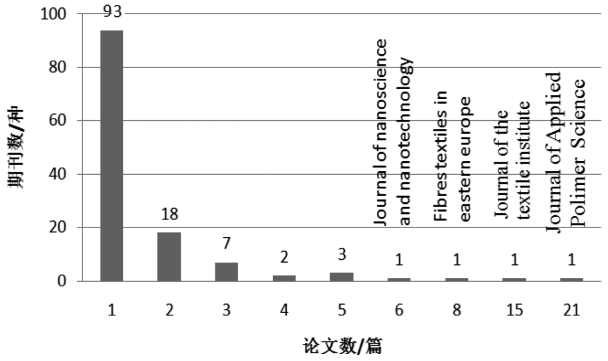


图 11　小类 4 区论文及其期刊分布情况

在学科分布方面，经聚类得到如图 12 所示学科及其收录论文分布,分析可见,材料、化学、工程和物理收录论文居前;其次是计算机、数学、环境、和生物;此外还涉及医学。

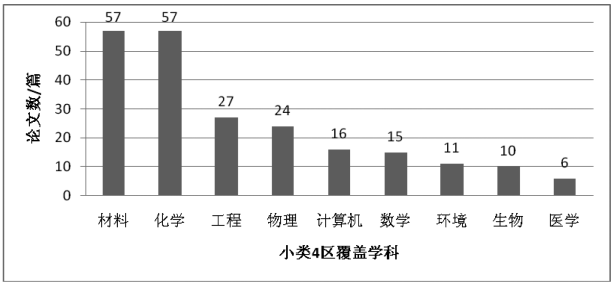


图 12　小类 4 区论文涉及学科分布情况

为了进一步剖析各细分学科分布情况，图 13 绘制了各学科的分布图。具体如下：

(1)由图 13(a)可见,材料学科中纺织和综合论文分布最多,分别为 56%和 26%;其次为硅酸盐,占该学科的 8%;此外,分别涉及 2%的生物材料、纳米科技、表征与测试和 2%的纸与木材;

(2)由图 13(b)可见,化学学科中高分子科学论文

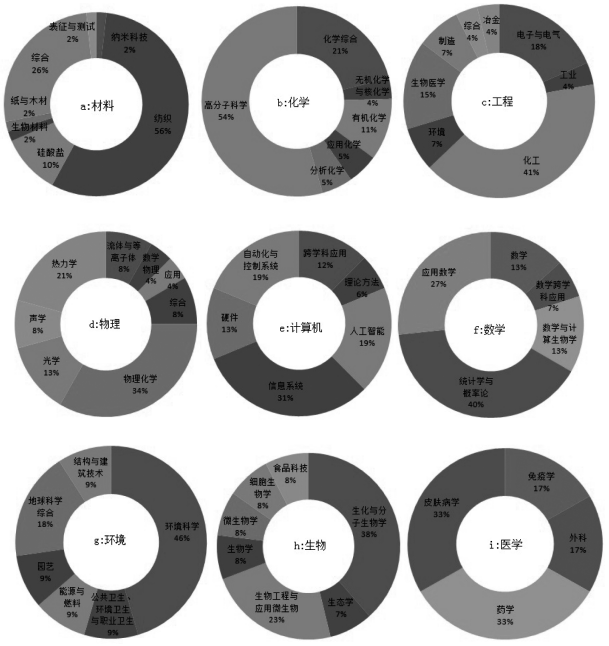


图 13　小类 4 区论文涉及较集中学科分布情况

分布最多，为 54%；其次为化学综合和有机化学，论文分布分别为 21%和 11%；此外是 5%的应用化学、5%分析化学以及 4%的有机化学与核化学；

（3）如图 13（c）所示，工程学科主要以化工、电子与电气和生物医学为主，分别为 41%、18%h 和 15%；其次涉及环境和制造，均为 7%；此外，还涉及 4%的综合、工业和冶金；

（4）如图 13（d）所示，物理学科主要以物理化学、热力学和光学为主，分别分布论文 34%、21%和 13%；其次为 8%的声学、流体与等离子体和综合学科；此外还涉及 4%的数学物理和应用；

（5）如图 13（e）所示，计算机学科主要覆盖 31%的信息系统、19%的自动化与控制系统、19%的人工智能、13%的硬件、12%的跨学科应用和 6%的理论方法；

（6）如图 13（f）所示，数学学科以统计学与概率论和应用数学为主，分别占该学科总数的 40%和 27%；其次为 13%的数学与计算生物学、13%的数学和 7%的数学跨学科应用；

（7）如图 13（g）所示，环境学科主要以环境科学和地球科学综合为主，分别占该学科总数的 46%和 18%；其次分别涉及 9%的能源与燃料、结构与建筑技术、园艺、公共卫生和环境卫生与职业卫生；

（8）如图 13（h）所示，生物学科主要涉及生化与分子生物学（38%）、生物工程与应用微生物（23%）、生物学（8%）、微生物学（8%）、食品科技（6%）、细胞生物学（8%）和生态学（7%）；

（9）如图 13（i）所示，医学学科除主要涉及 33%的医学和皮肤病学外，还分别涉及 17%的外科和免疫学。

## 四、结果与建议

学科服务作为目前引领图书馆服务发展的一个重要航标，其专业化、个性化、学科化的服务模式扭转了传统图书馆笼统化、一般化的服务格局，试图从纵深方向提升图书馆服务的品质并赋予其知识化、学科化特质。[5]面对大数据、云服务、“互联网 +”，甚至是智慧图书馆等的挑战，学科服务的“学科”化特色在推动图书馆发展转型的同时，图书馆的服务能力受到了空前的挑战：在用户需求分析基础上进行数据价值提纯、数据分析和数据挖掘，最终为用户提供满意、精准的服务。以东华大学 2015 年 SCI-E 收录期刊分区数据监测为案例的数据支持学科服务实践和曾晓牧等人[6]的学科服务实践表明，学科有效信息行为对于开展学科服务至关重要。SCI-E 收录期刊分区数据监测表明，通过 SCI-E 收录期刊不同分区的分布及其特征分析，可以获取学科有效信息行为，不仅有助于为馆藏结构调整和采购方案确定提供数据支持，而且有助于为学科规划决策提供数据支持。因此，有必要基于学科数据，构建完善的决策支持系统，不断实现和完善智慧学科服务。

**参考文献**


[1] 初景利.我国图书馆学科服务的难点与突破[J].中华医学图书情报杂志，2012(4)：1-4.

[2] 曾晓牧，林佳. 基于学科有效信息行为的学科服务实践——以清华大学数学学科为例 [J]. 大学图书馆学报，2014(3)：85-90.

[3] 大数据时代来临决策不能只凭经验[EB/OL].(2012-04-11). http://finance.jrj.com.cn/tech/2012/04/11022912710354.shtml

[4] 陈超祥，丁健龙.基于数据挖掘的图书馆决策支持系统建模与实现 [J]. 浙江树人大学学报（自然科学版），2007(2)：1-4.

[5] 燕辉，魏小盈，杨华.学科服务背景下“智慧与服务”新解读[J].图书与情报，2014(1)：122-127.

[6] 曾晓牧，林佳. 基于学科有效信息行为的学科服务实践——以清华大学数学学科为例[J].大学图书馆学报，2014(3)：85-90.