



2019年新信息环境下 CASHL 资源与服务拓展设计研究

双一流建设背景下基于 CASHL 资源的 人文社科领域学术影响力评价体系 设计

课题单位：四川大学图书馆

课题负责人：党跃武

课题组成员：张盛强（四川大学图书馆）、舒予（四川大学图书馆）、雷琴（四川大学图书馆）、范金晶（四川大学图书馆）、龚桢程（四川大学公共管理学院）

结项时间：二零二零年五月

摘要：从科研总量、科研质量和高水平研究三个角度设计标准化指标计算框架，同时引入传统文献计量学指标和 Altmetrics 指标，构建了一套适用于不同类型评价对象的科研评价指标体系，消除由于不同学科的差异性带来的学术影响力不可比性。选取人文社科领域的具有代表性的研究学者、学科领域以及学术期刊三种不同类型的评价对象进行实证研究，结果表明该评价体系能够消除评价对象之间由于学科差异造成的影响，并且可以推广至学者、学科以及期刊等不同类型的学术影响力评价。作为该评价体系的应用，可以为 CASHL 文献信息资源的建设和配置提供决策支持：学者学术影响力评价可以帮助发现学科领域高水平学者，进而以其论著纳入到 CASHL 文献信息资源的范围，为 CASHL 文献资源建设提供一种思路；学科学术影响力评价可以揭示各高校的优势学科，从而精准保障和优化各高校人文社科领域的文献信息资源；期刊学术影响力评价可以挖掘学科领域的高水平期刊，为 CASHL 遴选学术期刊提供数据参考。

关键词：学术影响力 学科标准化 指标体系

目 录

引言.....	4
1 研究现状、目的及意义.....	5
1.1 国内外研究现状.....	5
1.2 研究目的及意义.....	6
1.3 研究内容、方法与创新点.....	6
2 人文社科领域学术影响力评价体系的设计.....	7
2.1 科研总量指数.....	8
2.2 科研质量指数.....	9
2.3 高水平研究指数.....	10
3 实证研究与分析.....	11
3.1 指标数据获取说明.....	11
3.2 实证分析.....	11
4 结论.....	17
4.1 研究总结.....	17
4.2 研究不足.....	18
4.3 优化高校人文社会科学领域学术评价初步思考.....	19
参考文献.....	22

引言

近年来,科技评价制度改革、学术评价改革成为国内科研管理的热议话题,国家层面陆续出台若干政策文件来推动科技评价、学术评价的科学发展。2018年7月,国务院印发《关于深化项目评审、人才评价、机构评估改革的意见》《国务院关于优化科研管理提升科研绩效若干措施的通知》两个文件,为全面贯彻党的十九大精神,针对当前国内科技评价的问题和深入推进科技评价制度改革,提出了克服、清理“唯论文、唯职称、唯学历、唯奖项”等“四唯”现象的要求。同年10月,科技部等五部门联合发文开展清理“四唯”行动并提出具体举措,以清理各科研管理部门中涉及“四唯”的做法。多部门的联合行动反映出当前科技评价制度改革的迫切性和需求性,同时也为科研项目评审、人才评价、机构评价提供了改革导向,即评价不可套用“一刀切”的标准,同时要避免评价中只参看某一方面的极端现象,应尊重事实规律和坚持科学公正以形成以科技创新质量、贡献、绩效为导向的分类评价体系。今年,教育部与科技部为贯彻落实习近平总书记在全国教育大会上提出“深化教育体制改革”“扭转不科学的教育评价导向”的指示,专门印发《关于规范高等学校SCI论文相关指标使用树立正确评价导向的若干意见》文件,强调要进一步规范高校学术评价标准、评价指标的使用,建立健全分类分体系的评价体系等。该文件延续了国务院和科技部改革科技评价制度文件坚持科学评价的精神,将改革的对象聚焦于高校这个科研重地,这说明高校学术评价是当前高校科学研究及其科研管理工作的重点和难点,有不断改善和进步的空间,规范性与科学性地开展分类评价是当前高校学术评价改革的一个方向。

高校是科研生产的主要阵地。尤其是在目前高校“双一流建设”的背景下,迫切需要帮助高校总结和分析其科研实力的现状、全面了解院系的科研发展情况、识别出具有科研优势、发展潜力以及创新能力的学科或者团队。因此,需要构建一套科学合理的学术评价指标体系肩负适应不同学科学术要求、量化科研绩效、树立正确科研导向的责任。人文社会科学贴近社会发展主战场,其学术成果形式多样、特点复杂。基于这些现状,针对人文社会科学研究领域的学术影响力评价的指标及其体系的构建和完善是学术评价中具有挑战的部分。当前这一方面的研究取得了比较丰富的成果,具有代表性的指标体系国外有ESI(Essential Science Indicators)指标,国内有中文社会科学引文索引评价体系(CSSCI)。虽然前者ESI简单易行且国际认可度高的优势,但是,它对人文学科、社会科学学科的划分模糊不精确,仅以论文作为衡量学者、机构的学术水平和学术影响力的评价标准较为偏重自然科学评价,用来反映人文社学科、社会科学的真实学术水平却有失偏颇;后者CSSCI从人文社会科学的

内涵和自身特点出发,打造适用于中国人文社会科学学科的学术评价体系,具有评价对象多元、国内认可度高的优势,但也难免有学科分布不均、重引用轻价值的局限,如何立足于国内社会科学学术的需求,构建与完善适用于中国社会科学学术评价的指标体系是其需要努力的方向。

此外,在大数据时代,科研学术成果数量激增,学科的细分与交叉融合越来越多,传统的评价方法和单一的评价指标体系应用于当下的学术评价中显得较为吃力,且高校学术评价对评价体系具备“个性化”与“定制化”的要求也越来越高。人文社会科学学术评价的研究方法和评价指标体系需要继续创新和完善,这不仅是高校中对学术人才的筛选与引入、科研管理的高效运作、学术资源的更优分配和学科建设的发展导向的需要,也是促进社会科学研究理论与方法不断丰富和创新的需要。

1 研究现状、目的及意义

1.1 国内外研究现状

在不同的评价体系中,评价对象可以是机构、学科、研究团队、个人或者期刊,不同类型的评价对象均以学术论文作为其学术影响力的主要表现形式,因此从文献计量的角度,可以通过引入文献计量指标,对评价对象的论文产出、影响力等方面进行测度,揭示出评价对象学术影响力的表现,而这种基于文献计量的学术影响力评价方式具有客观性和公正性,能够克服同行评议由于主观判断带来的评价偏差,与之形成良好的互补,在各种评价工作中起到了重要的作用^[1-3],同时随着 Altmetrics 的发展,越来越多的学术评价中也引入了 Altmetrics 指标用以弥补传统计量学的缺陷^[4,5]。另一方面,不同的评价对象的研究领域是有差异的,由于不同学科之间发文数量、引用和使用规律存在着明显的区别,导致不同学科的学术影响力不能纳入到同一个尺度下进行比较,因此需要构建学科标准化的指标用以消除这种由于学科不同造成的差异,目前基于学科标准化指标的方法已有相当多的研究成果,并且也广泛应用于机构学科、研究团队、个人或者期刊的评价^[6-12],但是由于同一个评价对象中的研究成果也可能属于不同的学科范畴,而这些学科标准化指标往往只用于单一的学科评价,即是限定机构、研究团队或者个人在某一个具体学科的学术影响力进行测度和评价,这样不利于对评价对象整体的学术影响力水平进行描述,因此需要一个合理的集结模型将其所涉及不同学科的学术成果以统一的形式综合起来加以表征,这样就需要将用于单一学科的标

标准化指标适当地推广至个人、学科或者任意类型的学术影响力评价，进而构建一个能够适用于不同评价对象的统一的学术影响力评价指标体系。

1.2 研究目的及意义

对人文社会学科领域的学术影响力的评价一直都是学术评价领域中不断探索的研究方向之一，学术成果是由人经过科研工作生产的，它的质量与价值可以反映学者在科研学术中的工作质量与成效。对人文社会科学领域的学术成果进行科学客观、平等公正的评价是促进其社会价值与影响力实现的一种途径，也是对人文社会科学学者丰硕智力劳动的尊重。因此，为社会科学的学者建立科学合理的评价指标体系具有极大的现实意义。

在已有研究基础上，综合具体学科内涵和自身特点，力图构建一套能够适用于人文社科领域学术影响力的多维度评价体系，可以为高校的人文社科领域的科研管理决策、学科建设规划提供定量的数据支持和依据，对优化科研资源配置、提高科研管理水平具有重要的意义。

通过构建的学术影响力评价体系，为 CASHL 遴选文献信息资源、优化高校文献信息资源配置提供重要的数据参考。

1.3 研究内容、方法与创新点

1.3.1 主要研究内容

该课题综合文献计量指标和 Altmetrics 评价指标，引入科研总量指数、科研质量指数和高水平研究指数三个学科标准化的评价指标，从评价对象整体层面的科研产出能力和学术影响力、评价对象的科研水平、关注度和效率、评价对象的高水平研究表现三个维度对评价对象的学术影响力水平进行一个全面的描述，进而构建一套能够同时适用于对学科、个人、期刊等不同类型的学术影响力评价指标体系。

1.3.2 研究方法

(1) 文献调研法

对国内外学术影响力评价的相关文献进行梳理和分析，扩宽研究的深度和广度，充分对比和挖掘各种研究方法的特点。

(2) 多属性决策方法

设计指标标准化方法，利用多属性决策理论确定合理的评价体系计算框架，提高指标的区分度和精准性。

(3) 实证分析法

选取人文社会科学领域具有代表性的学者、期刊和学科领域为实证样本进行评价,验证该课题研究所构建的学术影响力评价指标体系在实际学术评价中的可行性。

1.3.3 课题创新点

该课题相关理论和方法涉及科技情报分析、不确定性决策和评估、多元统计等多个研究领域,属于涵盖图书情报、信息检索与分析、统计学理论等多学科交叉的综合性研究课题。

课题创新之处在于构建了针对人文社会科学领域的学术影响力评价指标体系标准化计算框架,通过该评价体系可以全面地反映高校科研实体的学术产出、学术质量和高水平研究三个维度的科研表现,并且该指标评价体系不仅可以将不同类型的研究实体涉及不同学科的学术影响力以量化的形式有效加以综合,从而实现来自不同学科背景的评价对象学术影响力的横向对比,同时还可以推广到个人、学科、期刊和机构等任何层面的研究实体的学术评价,有助于高校在“双一流”建设中针对其内部的各种不同层次的研究实体的学术绩效评价。

2 人文社科领域学术影响力评价体系的设计

针对评价指标的选取,俞立平等从科技投入和科技产出两个方面对科研的效率进行了评价^[13],花芳用产出指数、影响指数和被引指数描述学术影响力的水平^[14],Costas R 等利用因子分析法将不同的指标聚类,形成了产出、影响力和期刊质量三个维度的学术影响力评价指标体系^[15],李贞针对高校人文社科团队绩效评价的“投入—中介—产出”(IMO)模型,构建了三层指标结构的评价体系^[16],时云从科研项目与经费、研究成果、科研奖项、科研交流四个方面建立以质量为导向的高校人文社科评价指标体系,用以评价教师、科研人员的基本科研工作业绩^[17],郭凤娇在期刊影响因子、被引次数等传统的学术影响力评价指标基础上,引入了 Altmetric 评价指标,进而对学术论文的影响力进行评价^[18],张献锋从学术创新、学术影响、办刊能力、学风建设等四个方面构建多层次的中国社科类学术期刊评价指标体系^[19]。李品等结合大数据时代背景构建大数据思维下的哲学社会科学学术成果评价指标体系,以学术价值和社会效益作为一级指标,学术创新、学术影响、学术质量、社会影响、社会作用 5 个二级指标以及 34 个不同侧重点的三级指标^[20]。对相关文献调研发现,国内外对学术影响力评价指标体系设计的研究视角主要集中在基于 Altmetrics 的学术影响力评价指标设计以及多维度指标体系构建两个方面。

该课题将论文数量、被引次数等传统学术评价指标与 Altmetrics 的学术影响力评价指标充分融合,构建一套包含 3 个一级指标、7 个二级指标的学术影响力评价指标体系,见表

1. 在此指标体系的基础上，设计描述评价对象科研总量、质量和高水平研究水平的标准化指标，具体思路如下：选定评价对象，以该评价对象在某一时间段的各个学科的科研表现作为最小分析单元，针对其在每个学科单元下按指标体系中的各二级指标进行标准化，再将各二级指标加权综合，得到评价对象在各学科的总量、质量和高水平研究三个维度的标准化指标，最后将评价对象所涉及的不同学科指标进行加权综合，得到评价对象整体的总量、质量和高水平研究指数。

表 1 学术影响力评价指标体系

一级指标	二级指标
科研总量	论文总数量
	论文被引总次数
	论文篇均被引次数
科研质量	论文被引百分比
	论文篇均下载量
	被引次数排名前 20%的论文百分比
高水平研究	下载量排名前 20%的论文百分比

2.1 科研总量指数

现假设有共计 J 个评价对象（评价对象可以是个人、研究团队或者机构等）作为评价对象，第 j 个评价对象在第 i 年的研究论文涉及一共 K 个学科，分别发表了 $A_{i,1}^j, A_{i,2}^j, \dots, A_{i,K}^j$ 篇论文，被引用次数为 $C_{i,1}^j, C_{i,2}^j, \dots, C_{i,K}^j$ 次。

定义 1. 评价对象 j 在第 i 年学科 k 的相对产出度指标为

$$I_a^j(i, k) = \frac{A_{i,k}^j/A_i}{W_a(i, k)/W_a(i)}$$

其中， $A_i = \sum_{j=1}^J \sum_{k=1}^K A_{i,k}^j$ 表示所有待评价的评价对象在第 i 年发表的总论文数， $W_a(i, k)$ 表示参考论文集合（可以是全球范围的论文，也可以是国内范围的论文）第 i 年学科 k 的论文数量， $W_a(i)$ 表示第 i 年参考论文集合的总数量。

$I_a^j(i, k)$ 可以看作是评价对象 j 在学科 k 的相对所有评价对象的论文总数量按参考论文集合学科份额的归一化处理，代表了评价对象 j 在学科 k 上的相对产出能力和评价对象 j 在学科 k 上的优势。不难发现，如果所有的评价对象中都只有一位研究人员时， $I_a^j(i, k)$ 可以作为个人的科研产出评价指标，而当评价对象 j 扩大到整个机构的人员时， $A_{i,k}^j$ 即为整个机构第 i 年在学科 k 发表的论文数量， A_i 即为整个机构第 i 年发表的论文数量，于是 $I_a^j(i, k)$ 则转变为 Frame 提出的 AI 指数^[21]。

定义 2. 评价对象 j 在第 i 年学科 k 的相对活跃度指标为

$$I_c^j(i, k) = \frac{C_{i,k}^j / C_i}{W_c(i, k) / W_c(i)}$$

其中, $C_i = \sum_{j=1}^J \sum_{k=1}^K C_{i,k}^j$ 表示所有待评价的评价对象在第 i 年发表论文的总被引次数, $W_c(i, k)$ 表示参考论文集合第 i 年学科 k 的论文被引总次数, $W_c(i)$ 表示参考论文集合第 i 年的论文被引总次数。

$I_c^j(i, k)$ 可以看作是评价对象 j 在学科 k 的相对所有评价对象的论文总被引次数按参考论文集合学科份额的归一化处理, 是评价对象在学科 k 上的相对活跃程度的表征。与相对产出指标类似的, 当所有评价对象中都只有一个研究人员时, $I_c^j(i, k)$ 可以作为个人的科研活跃度评价指标, 而当评价对象 j 扩大到整个机构的人员时, $I_c^j(i, k)$ 则转变为 Braun 和 Schubert 提出的 AAI 指数^[22]。

定义 3. 评价对象 j 在第 i 年学科 k 的科研总量指数为

$$T_{i,k}^j = \alpha \cdot I_a^j(i, k) + \beta \cdot I_c^j(i, k)$$

在不同的评价体系中, 学科的产出能力和活跃度以会被认为具有不同的重要性, 因此可以在 $I_a^j(i, k)$ 和 $I_c^j(i, k)$ 前分别加入代表着不同重要性的归一化系数 α 、 β 。在该课题中取 $\alpha = \beta = 0.5$, 在实际应用中, 可以根据评价的需要对 α 、 β 赋予不同的权值。 $T_{i,k}^j$ 是评价对象 j 在学科 k 上的相对产出和活跃度的综合体现, 可以用于测度评价对象的优势学科和主要研究方向。

定义 4. 评价对象 j 在第 i 年的科研总量指数为

$$\bar{T}_i^j = \sum_{k=1}^K T_{i,k}^j \quad (1)$$

由上式可知, 评价对象在第 i 年的总量指数是其在各学科上的总量指数之和。从直观上来看, 评价对象的总量是由其在各个研究方向上的产出和活跃程度组成的, 而各个研究方向上的科研总量在进行了标准化后, 已经纳入到了一个统一的尺度下, 因此将各学科上的科研总量指数直接取和是符合现实意义的, 并且不会产生由于学科差异性造成的误差。

2.2 科研质量指数

定义 5. 评价对象 j 在第 i 年学科 k 的相对影响力指标为

$$I_{cpp}^j(i, k) = \frac{C_{pp_{i,k}}^j}{F_{CS_{i,k}}}$$

其中, $C_{pp_{i,k}}^j = \frac{C_{i,k}^j}{A_{i,k}^j}$ 是第 i 年评价对象 j 在学科 k 发表的论文篇均被引次数, $F_{CS_{i,k}}$ 是参考论文集合第 i 年学科 k 的篇均被引次数。

定义 6. 评价对象 j 在第 i 年学科 k 的相对被引率指标为

$$I_{cr}^j(i, k) = \frac{C_{r_{i,k}}^j}{F_{cr_{i,k}}}$$

其中, $C_{r_{i,k}}^j$ 是第 i 年评价对象 j 在学科 k 发表的论文被引用率, $F_{cr_{i,k}}$ 表示参考论文集合第 i 年学科 k 的被引率。

定义 7. 评价对象 j 在第 i 年学科 k 的相对关注度指标为

$$I_i^j(i, k) = D_{i,k}^j / D_{i,k}$$

其中, $D_{i,k}^j$ 表示评价对象 j 的平均下载量, $D_{i,k}$ 表示参考论文集合第 i 年学科 k 的论文下载量的基准值, 该基准值可以根据实际需求以及数据获取的可行性等因素确定。

定义 8. 评价对象 j 在第 i 年学科 k 的科研质量指数 $Q_{i,k}^j$ 为

$$Q_{i,k}^j = \alpha \cdot I_{cpp}^j(i, k) + \beta \cdot I_{cr}^j(i, k) + \gamma \cdot I_i^j(i, k)$$

与定义 3 类似, α 、 β 和 γ 是重要性系数, 代表了论文的影响力、被引率和被关注程度在评价体系中的重要程度, 在该课题中取 $\alpha = 0.5$, $\beta = 0.2$, $\gamma = 0.3$ 。

定义 9. 评价对象 j 在第 i 年的科研质量指数为

$$\bar{Q}_i^j = \frac{1}{K} \sum_{k=1}^K Q_{i,k}^j \quad (2)$$

由上式可知, 评价对象在第 i 年的科研质量指数是其在各学科的科研质量指数的算术平均值。不难发现, 当相对影响力指标 $I_{cpp}^j(i, k)$ 的系数 $\alpha = 1$ 以及相对被引率指标 $I_{cr}^j(i, k)$ 的系数和 $I_i^j(i, k)$ 的系数 $\beta = \gamma = 0$ 时, 评价对象 j 在第 i 年的科研质量指数 \bar{Q}_i^j 实际上就是 Opthof 和 Leydesdorff 提出的 MNCS 指标^[23]。由于相对影响力、相对被引率和相对关注度 3 个指标的基准值均为参考论文集合的平均值, 于是, 当 $\bar{Q}_i^j < 1$ 和 $\bar{Q}_i^j > 1$ 时, 分别表示评价对象的科研质量低于和高于参考论文集合的平均水平。

2.3 高水平研究指数

定义 10. 评价对象 j 第 i 年学科 k 的高水平研究指数 $H_{i,k}^j$ 定义如下:

$$H_{i,k}^j = \alpha \cdot Pt^j(i, k) + \beta \cdot Dt^j(i, k) + \gamma \cdot J^j(i, k)$$

其中, $Pt^j(i, k)$ 、 $Dt^j(i, k)$ 和 $J^j(i, k)$ 分别表示在第 i 年评价对象在学科 k 发表的论文中的被引次数进入前 20% 的论文比例、下载量进入前 10% 的论文比例以及高水平期刊论文比例。

高水平指数是评价对象的高水平研究水平的标准化表征, 从多个不同的角度描述了表示了论文中包含的高价值论文, 而论文比例本身即为一个相对指标。 α 、 β 和 γ 为重要性系数, 在该课题中, $\alpha = \beta = \gamma = 1/3$ 。

定义 10. 评价对象 j 第 i 年的高水平研究指数为

$$\bar{H}_i^j = \frac{1}{K} \sum_{k=1}^K H_{i,k}^j \quad (3)$$

评价对象在第 i 年的高水平研究指数是其在各学科的高水平研究指数的算数平均值。

通过上述 (1)、(2)、(3) 式, 定义了科研总量指数、科研质量指数和高水平研究指数的计算方式, 三种指数均以学科为分析单元进行了标准化处理, 不仅能够将同一个评价对象中属于不同学科的学术影响力合理地综合表征, 而且能够实现不同评价对象之间的横向比较,

其中科研总量指数是评价对象科研产出和科研活跃度的综合体现，具有“总量”的特征，科研质量指数是评价对象科研影响力和利用率的综合体现，代表了科研成果的价值，高水平研究是评价对象在前沿领域的表现。当评价对象为个人时，三种指数可以作为个人学术影响力的评价指标，而当评价对象扩大为学科或整个机构时，三种指数又可以推广为学科或者机构的评价指标，因此三种指数能够适用于不同类型评价对象的学术影响力评价，具有一定的通用性。

3 实证研究与分析

3.1 指标数据获取说明

数据来源：论文数量数据和下载次数数据来自中国知网数据库-中国学术期刊(网络版)，并且限制来源期刊类别为 CSSCI。引用次数数据来自中国引文数据库，限制被引文献类型为“CSSCI 收录刊”，来源文献范围为期刊库、学位论文库和会议论文库。

学科分类方式：按中国知网“文献分类目录”下的 10 个专辑进行分类，该课题将“社会科学 I 辑”和“社会科学 II 辑”合并为一个学科类别“社会科学”。

时间范围：2016 年-2020 年。

参考论文集合：在该课题中所需的参考论文集合均选取数据库中的所有国内论文。

参考论文集合基准线：选取参考论文集合论文的篇均被引次数、被引率作为计算相对影响力指标和相对被引率的基准值，选取参考论文集合下载量前 10%的阈值作为计算相对关注度的基准值（这里需要说明的是，选取下载量的平均值或者中位数更能够代表基准值，但是数据库只能提供前 6000 篇论文的信息，无法获取到全部论文信息，因此也就不能获取到下载量的平均值或者中位数，在该课题中只能以前 10%的阈值代表基准值）。

3.2 实证分析

3.2.1 用于学者的学术影响力评价

选取四川大学人文社科领域具有代表性的 3 位学者作为评价对象，其中学者 1 的主要研究领域集中在“哲学与人文科学”，学者 2 的主要研究领域分别为“经济与管理科学”和“社会科学”，学者 3 的主要研究领域集中在“社会科学”。3 位学者最终的结果见表 2，科研总量指数趋势、科研质量指数趋势和高水平研究趋势分别见图 1-3。从结果可知，学者 3 的整

体科研产出较高,而学者2的科研产出呈现上升趋势;学者3的科研质量也保持着较高水准,并且3位学者的科研质量在近五年均基本超过了国内的平均水平;在高水平研究方面,学者3的表现最好,除了2020年之外,3位学者在高水平研究方面也都超过了国内的平均水平。此外,学者2的学术研究横跨两个研究领域,不论是论文数量,还是被引次数、被引率以及下载次数,相关数据隶属两个不同的学科,因此学者2的各项指标值不能直接通过类似综合加权的数学处理获得。该课题设计的指标标准化计算框架首先将不同学科的学术表现做了标准化处理,使得各学科的指标值已经纳入到了一个统一的尺度下,在此基础上将学者2分属两个学科的学术表现进行综合处理,保证了其合理性。

在实际的应用中,可以通过海量的学术数据,挖掘出学科领域的高水平人才和领军学者,将其主要的论著,例如教材、专著、研究报告等,也纳入CASHL文献信息资源的选择范围中,从而丰富CASHL文献信息资源的建设方式。

表2 3位学者的科研总量指数、科研质量指数和高水平研究指数

	学者	2016年	2017年	2018年	2019年	2020年
科研总量指数	学者1	1.05	0.69	0.64	0.74	0.55
	学者2	0.30	0.49	0.61	0.50	0.92
	学者3	1.21	1.12	1.01	1.12	0.42
科研质量指数	学者1	1.33	1.83	3.52	2.42	0.22
	学者2	1.97	0.68	5.49	1.87	1.09
	学者3	2.91	5.59	7.10	4.20	0.32
高水平研究指数	学者1	5.21	5.00	6.67	5.63	1.67
	学者2	2.50	2.92	5.52	4.58	0.42
	学者3	6.11	6.88	7.14	6.50	1.67

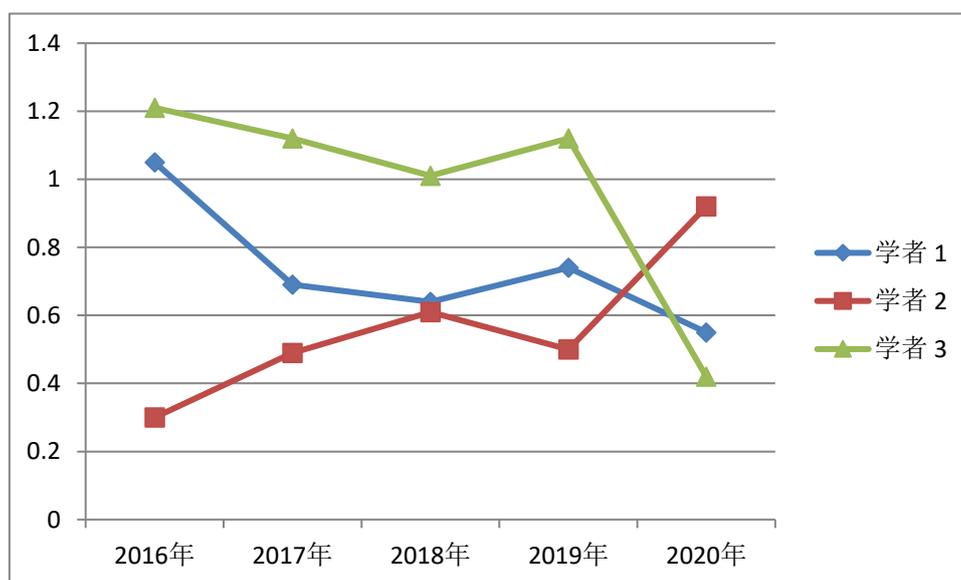


图1 3位学者科研总量指数趋势图

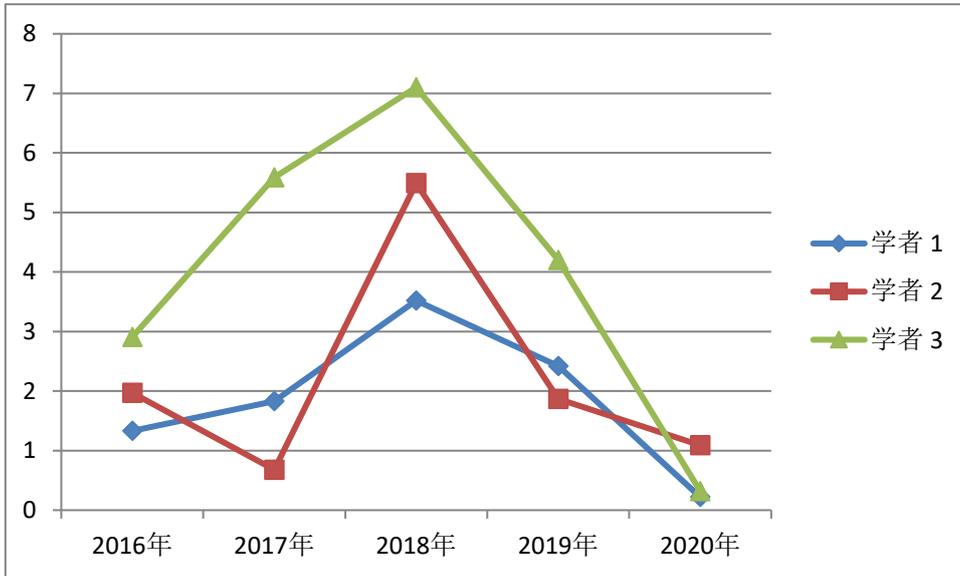


图2 3位学者科研质量指数趋势图

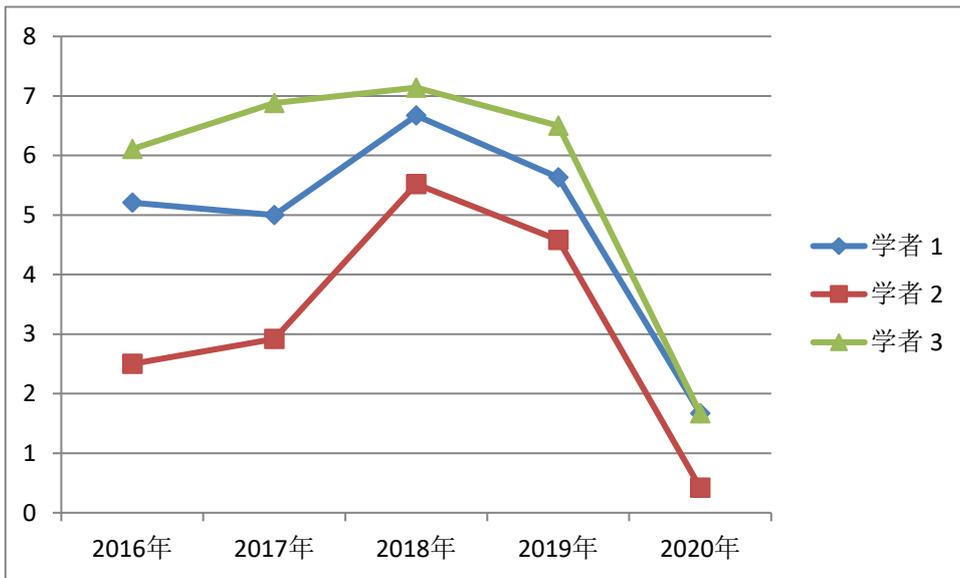


图3 3位学者高水平研究指数趋势图

3.2.2 用于学科的学术影响力评价

选取四川大学“哲学与人文科学”、“社会科学”与“经济与管理科学”3个学科进行评价与对比分析。最终结果见表3，3个学科的科研总量指数趋势、科研质量指数趋势和高水平研究趋势分别见图4-6。从结果可知，“哲学与人文科学”学科的整体产出最高，并且在各年份都超过了国内的平均水平，“经济与管理科学”学科的产出趋势较为平稳，在国内平

均水平线附近徘徊；在科研质量方面，“社会科学”学科的表现最好，但在 2020 年有所下降，而“哲学与人文科学”学科的科研质量处于稳步上升的趋势，除 2016 年之外，表现均超过了国内平均水平；在高水平研究方面，除开 2020 年的数据，三个学科均表现较好，超过国内水平较多，其中“社会科学”学科发展态势较好，一直处于上升趋势。

在实际的应用中，可以依托该评价体系揭示出各高校的优势学科和具有发展潜力的学科，从而为优化文献资源配置提供重要的数据参考，使得不同学科的文獻资源能够精准、有效地服务于各高校的相关科学研究。

表 3 3 个学科的科研总量指数、科研质量指数和高水平研究指数

学科		2016	2017	2018	2019	2020
科研总量指数	哲学与人文科学	1.36	1.43	1.40	1.00	1.59
	社会科学	0.66	0.72	0.87	0.96	0.85
	经济与管理科学	1.11	1.06	0.91	0.95	0.97
科研质量指数	哲学与人文科学	0.97	1.11	1.28	1.28	1.25
	社会科学	1.06	1.11	1.44	1.52	0.91
	经济与管理科学	1.09	0.97	1.04	1.14	0.70
高水平研究指数	哲学与人文科学	3.63	2.93	3.28	3.76	0.39
	社会科学	2.38	2.95	3.63	3.73	0.26
	经济与管理科学	2.35	2.38	2.51	1.53	0.52

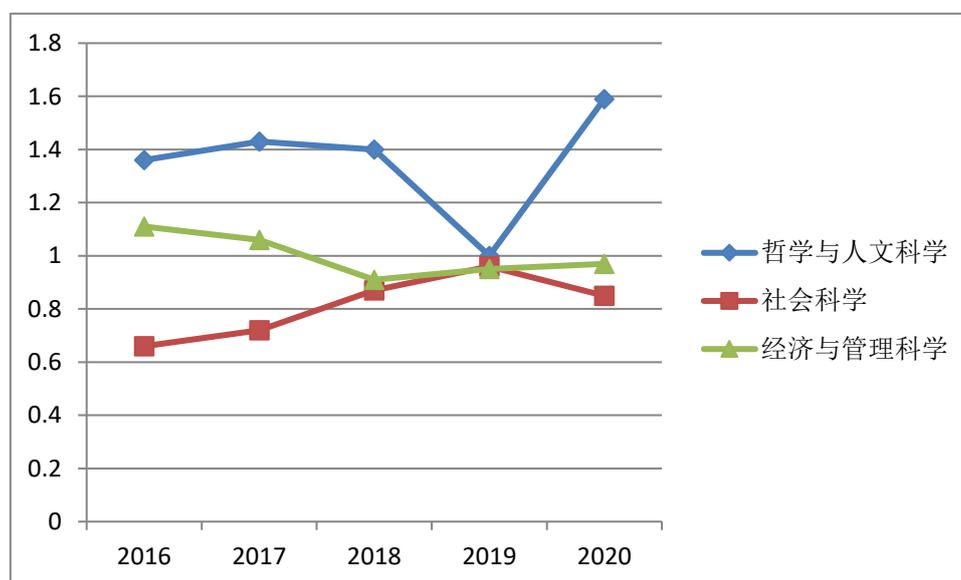


图 4 3 个学科科研总量指数趋势图

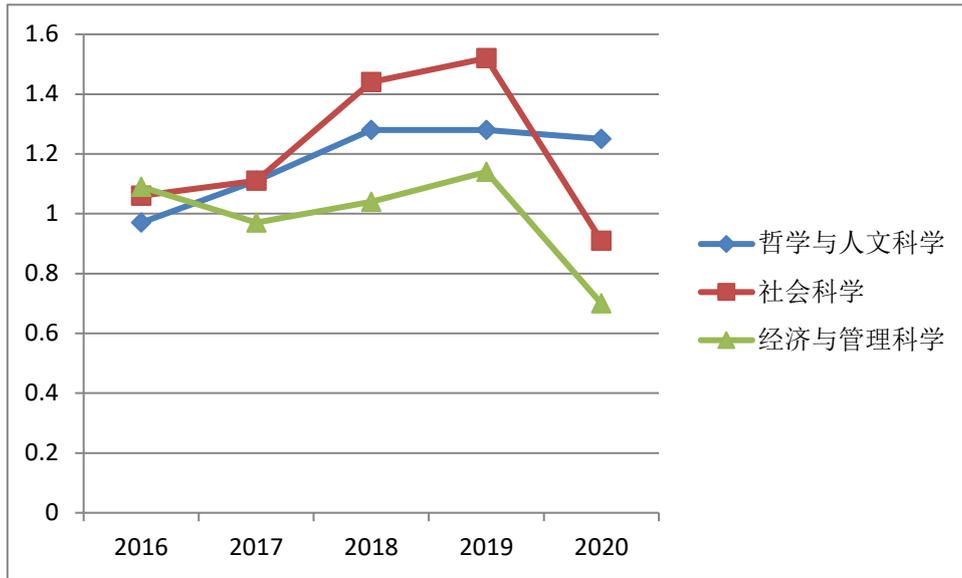


图 5 3 个学科科研质量指数趋势图

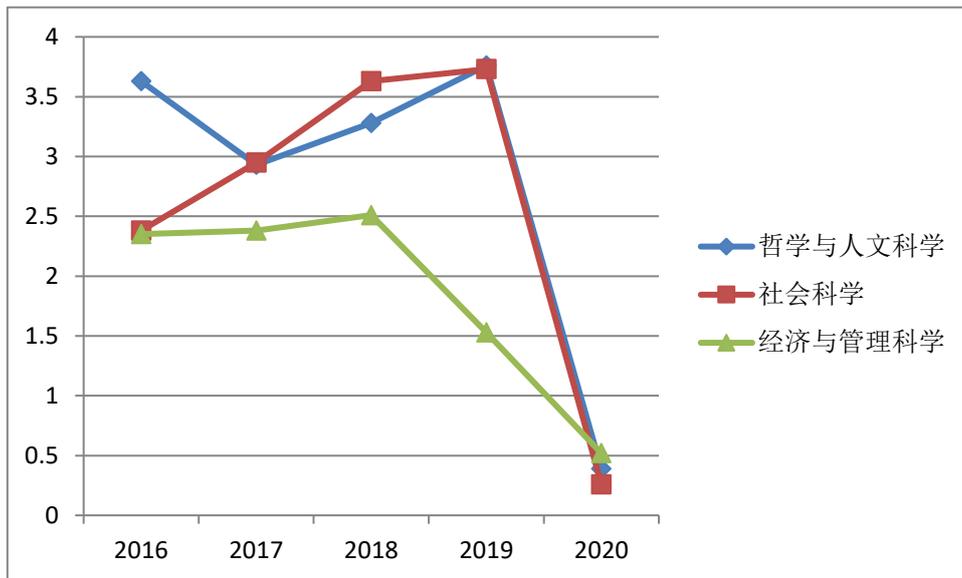


图 6 3 个学科高水平研究指数趋势图

3.2.3 用于期刊的学术影响力评价

选取《经济研究》《哲学研究》与《法学研究》3种期刊作为评价对象，3种期刊均被收录整合在CASHL期刊平台中，发文分别集中在“经济与管理科学”、“哲学与人文科学”和“社会科学”3个学科。最终结果见表4，3种期刊的科研总量指数趋势、科研质量指数趋势和高水平研究趋势分别见图7-9。从科研产出看，3种期刊呈现出不同的特点，期刊《经济研究》产出相对较高；从科研质量看，3种期刊的表现远超过国内平均水平，其中期刊《法学研究》表现最好；在高水平研究方面，3种期刊同样远远超过国内水平，其中期刊《法学研

究》表现最好，而《经济研究》和《哲学研究》2种期刊表现较为接近。3种期刊在科研质量和高水平研究2个方面的表现都远远超过了国内的平均水平，这也反映出CASHL平台整合的期刊在国内具有一定的影响力。

在实际的应用中，可以借助该评价体系挖掘出不同学科领域的高水平期刊，评价结果能够作为遴选期刊的参考，从而为CASHL用户提供高质量的期刊，并提升CASHL在国内的影响力。

表4 3种期刊的科研总量指数、科研质量指数和高水平研究指数

	学科	2016	2017	2018	2019	2020
科研总量指数	经济研究	1.54	1.73	1.57	1.66	1.44
	哲学研究	1.00	0.91	0.94	1.04	1.29
	法学研究	0.40	0.40	0.51	0.43	0.55
科研质量指数	经济研究	3.38	3.21	3.36	3.99	1.40
	哲学研究	1.63	1.97	2.22	2.71	1.12
	法学研究	3.84	3.62	4.62	3.87	1.58
高水平研究指数	经济研究	5.03	5.13	5.26	4.52	2.95
	哲学研究	4.35	5.00	5.11	4.65	1.11
	法学研究	7.05	6.70	6.74	5.98	2.21

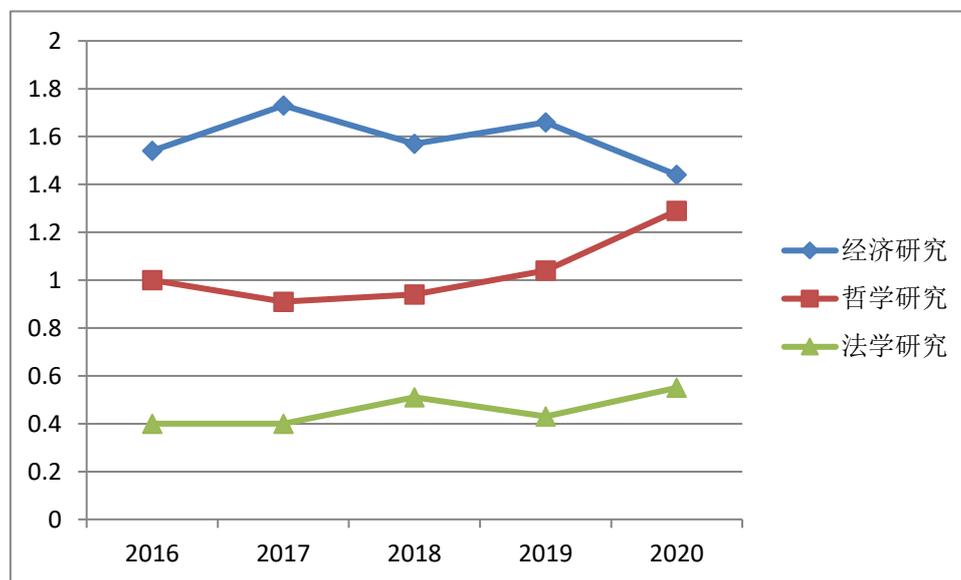


图7 3种期刊科研总量指数趋势图

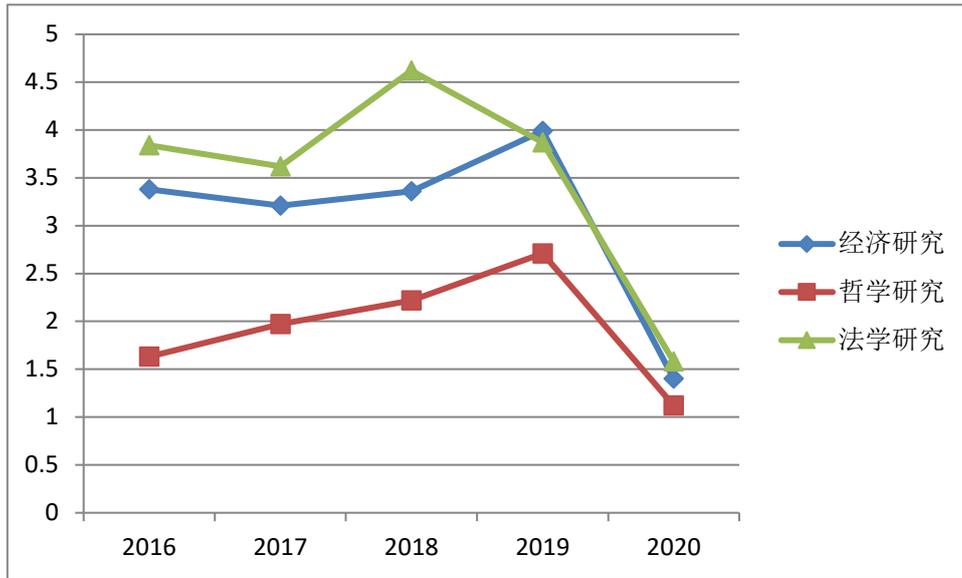


图 8 3 种期刊科研质量指数趋势图

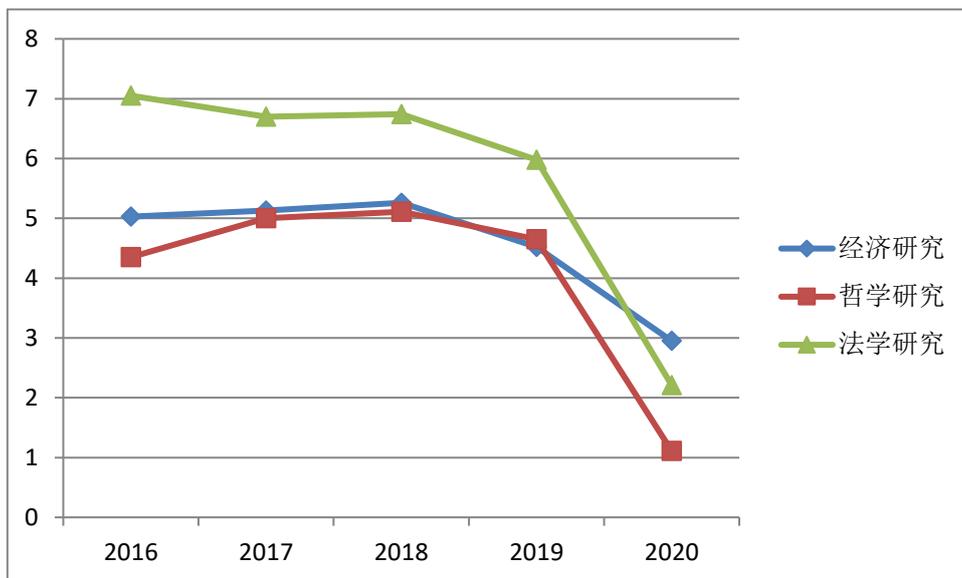


图 9 3 种期刊高水平指数趋势图

4 结论

4.1 研究总结

该课题充分综合了文献计量指标和 Altmetrics 指标，从科研总量、科研质量和高水平研究三个维度设计了学科标准化的科研评价指标，由此构建了一套适用于不同类型评价对象的科研评价指标体系，从实证分析中可以看到，该评价体系能将同一评价对象属于不同学科的学术成果有效地加以综合表征，不仅消除了评价对象之间由于学科差异造成的影响，并且

可以推广至个人、学科、期刊等不同类型的评价对象的学术影响力评价。

作为该指标评价体系的应用,可以为 CASHL 的文献信息资源建设与配置提供一个有效的工具和思路:

(1) 利用该评价体系,挖掘出人文社科学科领域的高水平人才和领军学者,将其主要的论著,例如教材、专著、研究报告等,也纳入 CASHL 文献信息资源的选择范围中,从而丰富 CASHL 文献信息资源的建设方式。

(2) 依托该评价体系,揭示出各高校的优势学科和具有发展潜力的学科,从而为 CASHL 优化文献资源配置提供重要的数据参考,使得不同学科的文献资源能够精准、有效地服务于各高校的相关科学研究。

(3) 借助该评价体系挖掘出不同学科领域的高水平期刊,评价结果能够作为遴选期刊的参考,从而为 CASHL 用户提供高质量的期刊,并提升 CASHL 在国内的影响力。

4.2 研究不足

(1) 评价体系“通用性”和“精确度”的矛盾

该课题的目的是建立一套适用于不同评价对象、不同研究实体、不同学科的评价体系,该评价体系的特点是具有“普适性”、“通用性”,因此在构建指标体系的过程中,选取的是能够体现不同评价对象学术影响力共性的指标,由于学术论文是不同类型的研究实体学术成果最基本的体现,因此课题设计的评价体系采用的均为论文指标。而一些特点鲜明、能够精确体现研究实体学术影响力内涵的指标,例如获奖次数、项目数量、影响因子、决策报告等,由于只适用于一类或者少数几类研究实体,最终只能舍弃掉。

在构建评价体系时,需要明确评价体系的用途及目的,研究探讨各类评价对象的内涵、特点以及相互之间的联系和区别,厘清评价指标的使用范围,确保评价体系的科学性、规范性与准确性。

(2) 未找到合适的多源数据的整合方案

大数据与开放科学背景下的学术数据的来源是丰富多样的,纸质、多媒体、专用数据库、网络平台等多渠道的学术数据获取途径构成了当前的学术数据的来源多源的特点,既有同类型学术成果数据的来源多样,也有多类型学术成果数据来源的多样。当前整合学术数据程度较高的商业数据库数量繁多,各有特点,但收录的数据范围重叠,为科学研究提供结构化、高质量数据集的能力有限。在未能从内部平台获取精准数据的条件下,使用单一的数据源进

行科学研究，操作相对简易和可实现，但实验结果被信服的程度却较难提高；使用多源的数据源进行科学实验，工作量大、研究难度也大，没有合适的指导方案可能无法按目标完成研究，也容易造成研究资源的浪费。缺少合适多源数据整合方案来指导这些多源异构学术数据的使用，就难以将它们深层次的研究价值和社会效益更好地体现出来。因此，在学术评价中，多源数据的整合方案是研究工作中的重要一环。该课题在研究过程中未能针对评价对象找到合适的多源数据整合方案，无法避免评价结果可能存在一定误差和偶然性。

（3）数据样本范围全面性有待提高

该课题在实证研究部分选择四川大学的学科、学者以及相关学科的期刊作为实验样本，对于该课题研究要实现综合和有效评价人文社会科学领域学术影响力的目标来说，数据样本采集的范围还不够全面：实证研究选取的评价对象所涉及的学科未对整个人文社科领域全覆盖；在样本选择的学者、期刊不够全面，例如未在不同年龄结构、不同职称的学者、不同水平层次的期刊中均匀采样；学科划分略显粗糙，该课题选取的是中国知网提供的一级学科划分的方式，学科范围较大，事实上在同一一级学科下的不同学者、期刊等的研究领域也存在着明显的差异。

（4）缺少有效的数据采集手段

数据的采集也是课题研究过程中遇到的主要困难，虽然选择的数据库提供了较为丰富的数据，但与国外一些知名的科研分析工具（例如 InCites、SciVal）相比，相关数据仍然缺少的深度组织和整合，课题研究中需要的一些数据只能通过手工采集和清洗，这使得课题研究效率较低。

（5）社会科学中交叉学科评估需要给予重视

一些具有多学科交叉性质的学科不能完全从自然科学的角度或人文社会科学的角度来单独评价。在学科发展的过程中，一些交叉学科融合了自然科学与社会科学的特点，其学术成果的学科性质也可能出现与所划分的学科类型特征相反或差异较大的情况。该课题研究对这些交叉方向、跨科学科的学术评价未做深入探讨，交叉学科也是社会科学的重要组成部分，对这类学科精准与全面的学术评价的是今后可以研究的方向。

4.3 优化高校人文社会科学领域学术评价初步思考

课题组对人文社科领域评价的优化对策进行了初步的思考，一则是对未来高校人文社会科学学术评价的优化路径的探究，以此来推进高校人文社会科学学术评价研究的完善进程；

一则可为构建人文社会科学学术影响力评价指标体系的相关研究提供一定的理论借鉴和参考。

(1) 结合实际探索有效评价方法和途径

从国家的宏观视角来看,应在学术评价中注重人文社会科学学术成果本土化的特点以提升评价成果的适用化。我国是一个人文社会科学大国,产出的人文社会科学学术成果种类丰富、数量繁多,且国内有相当一部分的人文社会科学学术成果都具有本土化特点,充分反映了国情、传统文化以及思想特色。它们的价值与贡献不宜直接套用国外的评价体系来衡量,应先思考这些本土化特点强烈的社会科学学术成果的内涵价值发挥和外部贡献彰显的途径是什么,再进行下一步的评价工作。立足于我国国情,结合具体学科背景,探索具有中国特色、符合不同学科规律和特色的高校人文社会科学领域学术评价途径和方法,而对于西方的现代知识体系应保持谨慎态度和坚持合理借鉴,在学术上重视自主性和本土价值,如避免过度使用国内外权威数据库、核心期刊来主导评价而落入“核心至上”的局限中。

从高校的视角来看,运用复合评价理念指导学术评价、提倡创新与价值导向、建立健全分类分体系的评价体系都是应加强的方向。如前今年教育部与科技部《关于规范高等学校SCI论文相关指标使用树立正确评价导向的若干意见》中,倡导高校在学术评价中要规范SCI论文相关指标的使用、完善学术同行评价、提倡分类评价体系等意见,明晰了当前和未来高校学术评价发展和改革的趋势。由于不同高校类型与资源不同、不同学科建设现状不同、不同学术梯队学者能力不同,高校在对不同主体的学术评价中,应结合自身学者、学科、学校实际对上方文件中的意见正确解读与合理吸纳,探究适合自身的评价体系、方法和途径。针对当前部分高校在学术评价中存在过度倚重核心期刊论文指标、质与量评价关系失衡、评价成果适用范围有限等问题,要因地制宜地解决。如在评价理念上选择复合评价理念来代替单一评价理念,破除“论文至上”的错误认识,从多维度有效引导学术评价的方向。在中国人民大学人文社会科学学术成果研究中心与中国人民大学书报资料中心联合发布的复印报刊资料系列人文社科评价成果中,运用了“以学术成果为指导,以同行评议为主导,以价值判断为引领,以数据分析为支撑”的“复合评价”理念,就为国内社会科学学术评价改革提供了良好的示范。在评价导向上,要坚持创新水平和科学价值的导向并合理平衡“质”与“量”评价的权重,破除“唯量化”论。目前学术评价的理论研究往往以学术论文作为研究实体学术绩效的主要表现形式,应当以目前我国学术评价的实际问题为导向,探索行之有效的评价方案。

(2) 多途径提升科学评价的实证能力

实践是检验真理的唯一途径。实证研究是学术评价中验证评价理论方法、评价框架模型和评价指标体系在具体实践中是否切实可行、是否发挥功能作用、是否体现社会效益重要途径之一，实证研究的结果和结论是指导学术评价工作进行迭代优化的重要参考依据。所以，实证研究是学术评价的生命周期中不可缺少的重要一环，加强实证研究、提升实证能力也是学术评价研究中迫切的需求。在人文社会科学学术评价中如何加强实证研究和提升实证能力？首先，在研究意识上，要对实证研究这一环节引起足够的重视，没有有效的实证研究是难以让自己的研究成果被大众承认接受和信服的，也难以让自己研究成果价值真正地体现出来。其次，在研究实践上，要善于寻找适用的实证方法和提升自身的实证研究技能，如评价指标体系的构建要抛弃庞杂冗余的指标，力求精简可行；要合理借助软件或相关工具在过程中提高数据采集、数据处理的效率。此外，社会科学的学术成果本身具有的特殊性、价值难量化，在难以被数据量化和检验的情况下是否能通过其他实例来验证其价值和贡献，这也是未来研究中需要思考的问题。

（3）构建合理的高校学术评价管理模式

对高校人文社会科学领域的学术评价，在高校层面可以强化学术评价管理的功效，一是可观测高校各学科的学术水平和发展情况，因地制宜地推进学科建设；二是通过公平的科研绩效来保证学者们的切身利益，挖掘和嘉奖优秀学者，督促学术懈怠的学者向上，引导在高校的科学研究形成健康的学术风气。高校从事社会科学学科研究的教师和科研人员对本领域学术成果的价值体现方式、对本学科当前学术评价的弊漏是最有发言权的，因此，要引导高校社会科学学者参与到高校学术评价管理的工作中是重要也是有必要的。例如在同行评价制度无法完全落实的领域可否考虑将同领域学者互评制度纳入补充范围，高校管理部门考虑是否能适当下放部分学术自治权力给学者来完善学术成果发表与出版机制、学术评价奖惩机制、职称评审机制等。

（4）大力促进学术信息交流和学术资源共享

在信息技术快速发展、日益成熟的背景下，为在线科研提供了良好的发展空间，也为学术信息交流与学术资源共享提供了便捷与帮助。不同地域不同高校不同学科的学术评价改革的进展，以及学术评价研究的成果、学术评价工作的经验，都具有进行更大范围、更为顺畅地开展交流与共享的需求。当前国内还鲜有这类能满足这些需求并供高校使用的工具或平台，如何充分利用信息技术工具来助力高校学术评价发展也是应该思考的一个方向。如建立高校学习共同体并提供无障碍研讨平台，让高校学者共同监督和管理学术评价从启动、执行到落地的整个生命周期，在评价工作中有任何质疑与问题及时反馈和被响应；在问题的探讨和方

案的选择上,倡导扁平化的交流,不仅提倡不同职级的学者在相同或相近学科下的积极交流,而且鼓励不同院校、不同学科的学者开展正常的学术争鸣。

CASHL 作为全国性质的文献信息资源平台,也可以构建和完善专门用于人文社会科学学者学术成果及科研活动数据的平台,为高校学术评价提供真实准确的数据,对多种类型的学术成果进行有效整合并构建相关数据库提供给研究使用,以及探究多源学术数据的整合方案,支撑国内人文社会科学领域学术评价理论研究的开展。

参考文献:

- [1] Bornmann L, Daniel H D. Functional use of frequently and infrequently cited articles in citing publications. A content analysis of citations to articles with low and high citation counts[J]. *European Science Editing*, 2008, 34(2): 35-38.
- [2] Van Raan A F J. Measuring science[M]//*Handbook of quantitative science and technology research*. Springer Netherlands, 2005: 19-50.
- [3] Daniel H D. Publications as a measure of scientific advancement and of scientists' productivity[J]. *Learned publishing*, 2005, 18(2): 143-148.
- [4] 杨思洛.新时期我国人文社科成果国际影响力评价: 问题与趋势[J].*情报资料工作*,2020,41(03):20-25.
- [5] Siluo Yang, Mengxue Zheng. Performance of citations and altmetrics in the social sciences and humanities[J]. *Proceedings of the Association for Information Science and Technology*,2019,56(1).
- [6] Abramo G, D'Angelo C A, Di Costa F. Assessment of sectoral aggregation distortion in research productivity measurements[J]. *Research Evaluation*, 2008, 17(2): 111-121.
- [7] Bornmann L, Mutz R. Further steps towards an ideal method of measuring citation performance: The avoidance of citation (ratio) averages in field-normalization[J]. *Journal of Informetrics*, 2011, 5(1): 228-230.
- [8] Waltman L, van Eck N J, van Leeuwen T N, et al. Towards a new crown indicator: Some theoretical considerations[J]. *Journal of Informetrics*, 2011, 5(1): 37-47.
- [9] 郑芳,周群,陈仕吉,左文革. 基于百分位数指标的个人科研绩效评价研究[J]. *科研管理*,2013,S1:189-194.
- [10] 戴星. 评估机构影响力的标准化指数——MZE 指数:以图情学科为例[J/OL].*情报杂志*:1-9[2020-06-17].<http://kns.cnki.net/kcms/detail/61.1167.G3.20190710.1239.028.html>.
- [11] 刘雪立,申蓝,王燕,孟君,魏雅慧,郑成铭,董建军,范艳芬,方红玲,李常艳,单卫华,盛丽娜,杨博,周二强. 学术期刊的跨学科评价指标 JIPR8 及其应用[J].*中国科技期刊研究*,2019,30(06):663-670.
- [12] 赵镇,陈金源,曹蓓,韩栋,程玲,陈征,白玉,郑童桐,赵醒村.基于美国《科学引文索引》分区的新系数算法对不同学科间学术论文评价方法探讨[J].*科技管理研究*,2019,39(01):61-64.
- [13] 俞立平,张晓东. 基于熵权 TOPSIS 的地区高校科技竞争力评价研究[J]. *情报杂志*,2013,11:181-186.
- [14] 花芳. 学科基准化的科研绩效定量评价方法及指标研究[J]. *图书情报工作*,2014,03:78-84.

- [15] Costas R, Van Leeuwen T N, Bordons M. A bibliometric classificatory approach for the study and assessment of research performance at the individual level: The effects of age on productivity and impact[J]. *Journal of the American Society for Information Science and Technology*, 2010, 61(8): 1564-1581.
- [16]李贞,张体勤.中国高校人文社科团队绩效评价指标体系研究——基于 IMO 模型分析[J]. *重庆大学学报(社会科学版)*,2018,24(02):77-86.
- [17]时云,赵中洲,周欣.建立以质量为导向的高校人文社科科研评价指标体系研究[J]. *黑龙江教育(理论与实践)*,2017(06):22-25.
- [18]郭凤娇,赵蓉英,孙劲敏.基于科学交流过程的学术论文影响力评价研究——以中国社会科学国际学术论文为例[J]. *情报学报*,2020,39(04):357-366.
- [19]张献锋.中国社科类学术期刊评价体系的若干思考[J]. *重庆大学学报(社会科学版)*,2015,21(05):116-120.
- [20]李品,杨建林.大数据时代哲学社会科学学术成果评价:问题、策略及指标体系[J]. *图书情报工作*, 2018, 62(16): 5-14.
- [21] Frame J D. Mainstream research in Latin America and the Caribbean[J]. *Interciencia*, 1977, 2(3): 143-148.
- [22] Schubert A, Braun T. Relative indicators and relational charts for comparative assessment of publication output and citation impact[J]. *Scientometrics*, 1986, 9(5-6): 281-291.
- [23]Waltman L, van Eck N J, van Leeuwen T N, et al. Towards a new crown indicator: Some theoretical considerations[J]. *Journal of Informetrics*, 2011, 5(1): 37-47.