

## 大数据之于学术评价:机遇抑或陷阱? ——兼论学术评价的“分裂”

■ 朱 剑

(南京大学学报编辑部 江苏 南京 210093)

**【摘要】**作为行政权力部门分配学术资源和进行学术管理重要依据的学术评价如今已发生了分裂,定性评价和定量评价由一般的评价方法升格成为两种不同的评价种类。两类评价的主体分别为学术共同体和专业评价机构。两个评价主体展开了对评价权力的角逐,致使学术评价失去了完整性而变得残缺不全。正是在如此情境下,学术评价迎来了大数据时代。大数据之于学术评价,既提供了评价回归统一的机遇,也布下了分裂走向极致的陷阱。重建科学合理的学术评价机制的关键,是处理好大数据时代与前大数据时代学术评价的关系、科研体制与学术评价的关系以及学术评价内部诸要素间的复杂关系。唯此,分裂的评价方能在科学的意义上重新归于统一。

**【关键词】**大数据 学术评价 评价主体 评价方法

DOI:10.16034/j.cnki.10-1318/c.2015.04.013

“大数据”概念的提出不过短短四五年的时间,之所以能在如此短的时间里风行于世,不仅因为互联网时代有着“数量巨大、结构复杂、类型众多数据构成的数据集”这一事实基础,更因为技术进步所带来的看得见的效用:人们可以从几何级数增长的天量数据中挖掘可资利用的信息,并通过“基于云计算的数据处理与应用模式,通过数据的集成共享,交叉复用形成的智力资源和知识服务能力”<sup>[1]</sup>,为经济建设和社会发展提供其无可比拟的服务。

作为一场新的技术革命,大数据不同凡响的作用正在某些行业中显现,这就使学界中人联想到,在学术研究和学术评价领域,大数据是否也会带来某些变化,甚至是革命性的变化?答案似乎是肯定的。在学术研究中,采集和运用数据进行研究已是一种得到公认的方法,大数据无疑会使可资学术研究的数据在量与质的方面都有所提高,会使研究方法更加丰富多彩。与学术研究相比,在学术评价中,数据的作用更是无处不在:一方面,根据数据制作的各种排行榜和排名表大行其道;另一方面,围绕数据的意义和作用的争论也从未平息,对学术评价的质疑和诟病几乎都与数据有关。两方面的对立呈胶着状态,迫切需要新元素的介入来改变现状。在这样的情境下,大数据自然被寄予了更多的期望。

的确,由于评价数据与数据处理之于评价的不可或缺性,大数据完全有可能对评价产生影

收稿日期:2015-04-10

作者简介:朱剑,南京大学学报编辑部编审,主要研究学术期刊与学术评价。

基金项目:本文系教育部人文社会科学研究规划基金项目“高校学术期刊发展战略研究”(课题编号:14YJAZH116)的研究成果。

响。但是,影响的程度有多大,影响是正面的还是负面的,则很可能取决于我们对大数据与评价关系的看法。要让大数据发挥正面的积极作用,我们首先应该弄清楚学术评价的现状及其问题的症结到底何在;其次应该准确地描述和分析这些问题与数据和数据处理方法之间有着何种关系;最后再来看大数据是否有助于这些问题的解决以及解决的条件是否具备,从而判断大数据能在多大程度上改变学术评价现状。本文将循着这个思路对大数据与学术评价的关系作一尝试性的探讨。

## 一、学术评价的现状:分裂的评价

### (一) 学术评价的分裂

学术评价的对象是多元的,学术成果(作品)、作者、期刊、出版机构、科研机构等都可归入评价对象,但只有对学术成果(作品)的评价是基础性的,对其余各种对象的评价都是建立在对成果(作品)评价的基础上的,所以,核心的评价是对成果(作品)的评价。学术成果(作品)一旦公开发表或发布,对它的评价就是所有读者和应用者的基本权利,这就决定了评价的主体也是多元的。一般说来,读者和应用者通过学术批评来行使他们的评价权利。学术批评的方式当然也是可以多种多样的,最常见的批评方式则是撰写和发表学术评论。多主体决定了学术批评永远也不可能完全一致,而且,被批评者拥有反批评的权利,评价不会一次性地结束。换言之,在学术研究领域,一锤定音的评价几乎不会存在。学术评价在一定时段内的不确定性对于学术研究的进展是十分必要的,学术研究正是在这样反复的批评与反批评中实现超越的。

在现实中,学术评价的作用并不局限于研究领域和应用领域,它还是行政权力部门行使学术资源分配和科研管理权力的依据,从而与权力相关,而资源分配和科研管理需要的是结论明确的评价,不然就会无所适从。显然,多主体和多时段反复进行的学术批评无法满足行政权力部门的需要。为了使评价具有明确的结论,最好的办法就是对评价主体的身份作出限定。如果有一个主体能够博采众长,短时间内作出让多数人信服的权威评价,那么,这个主体作出的评价就是行政权力部门最好的也是最自然的选择。这是一种特殊的评价,它不是学术评价的全部,却需要建立在对全部评价进行判断和扬弃的基础之上。这样的评价在科研体制中的重要性不言而喻,如果这样的评价出了问题,必将危及整个科研事业,也必将引发被评价者群起而攻之。不幸的是,群起而攻之的局面至少在十多年前就出现了,且久久不能平息。可以说,今天对学术评价的所有批评和抨击都集中于为行政权力部门服务的评价上。只要集中分配和管理学术资源,不管分配和管理的权力集中于行政权力部门,还是转移到专门机构,为其服务的学术评价都是不可或缺的。在可以预见的将来,这样的集中分配和管理资源的方式都不可能根本的改变,这是我们今天讨论学术评价问题无法改变的前提。因此,本文所讨论的学术评价也主要限于这样的评价,而不是针对多元主体的学术批评。必须指出的是,为分配和管理资源服务的评价如果能做到科学和合理,对学术研究同样是有意义的。从根本上说,分配和管理学术资源也是为学术研究服务,它与学术研究的利益应该是一致的。

如果说,为行政权力部门服务的学术评价应该由权威的评价主体作出,那么,迄今为止,得到最多承认的、最可能具备这种能力的评价者则是学术共同体,尽管它总是遭遇各种挑战。学术共同体的评价又称为“同行评议”。直到20世纪90年代末,来自学术共同体的专家也还是这种评价的唯一主体,同行评议也还是这种评价的唯一方式。但是,随着国家对学术研究投入的

不断加大,学术资源日渐增多,评价与利益的关系越来越紧密,同行评议开始“异化”,“这些起源于欧美、并且长期行之有效的东西,却在中国发生了变异,甚至是‘异化’。其中‘异化’最明显、危害最严重的,莫过于‘同行评议’”,“权力支配,人情主导,标准缺失……三者只要居其一,都会使学术评价的意义受到严重损害”,“在我们当前的学术评价中,往往是三者一起发生作用,多路夹击,来自欧美的同行评议,怎么可能不水土不服以致彻底变质呢?”<sup>[2]</sup>于是,以“定量评价”的数据统计和运算见长的专业评价机构应运而生,并开始与学术共同体抢夺评价主体的资格。

随着专业评价机构的问世并积极地参与到为行政权力部门服务的行列中,通行的学术评价已分裂为定性评价和定量评价两类。需要特别指出的是,将学术评价区分为独立的定性评价和定量评价,是存在着对学术评价认识的误区的。定性也好,定量也罢,原都只不过是学术评价的一种具体方法,互相之间并不对立和排斥,在专业评价机构出现以前,即当学术共同体还是唯一的评价主体时,评价者可以根据不同的评价目的,采取或定性或定量或两者皆用的方法。专事

%

经指出,评价机构与评价对象之间存在着“信息不对称”:“对于各学科前沿问题、艰深问题……评价机构掌握的实质性信息远低于被评价者。”既然如此,评价机构又是依凭什么获取学术评价主体身份的?这是因为,评价机构不仅擅长于对一些论文和期刊外在形式数据的统计,并且进行了复杂的数据处理,形成了大量的形式数据,“从而造成了占据信息优势的假象”<sup>[3]</sup>。正是凭借对这些关于评价对象的量化数据的拥有以及运算的貌似具有排他性的能力,更因其制作的各种排行榜和排名表使复杂的学术评价变得十分简单和明确,得到了行政权力部门的青睐,

值判断的结论内涵的丰富性。其评价结论可以是排名表,但更可以是鉴定性的文字,甚至长篇大论,如果评价的公正性和专业性得到保证,就既能让公众信服,也能让评价对象从中获益。但是,在放弃了定量分析等非直接判断的方式、同行评议在成为单纯的定性评价之后,其缺陷也是明显的,本来具有的优势甚至变成了劣势:(1)单纯的定性评价有可能使评价成为一种纯粹的主观判断,必然受到评价者自身学识、德行等多方面的制约,以及外界各种因素的干扰,公正性难以得到保证。(2)定性评价的过程难以展示,很难做到公开;如果将评价视为一种科学研究或实验,其过程难以重复验证,科学和公平也难以体现。(3)由于前两点缺陷,评价的随意性难以避免,公信力难以保证。

需要指出的是,学术评价走向分裂的标志并不在于评价方法的不同和对立,而在于两个独立的评价主体的出现。主体的分裂甚至对立,才导致了方法的割裂和抵牾。作为一般的评价方法,无论单纯的定性评价还是单纯的定量评价都不可能独立应对所有的评价需要,本应统摄于同一评价主体之下,才能给评价提供科学和合理的基本保证。这里所说的“统摄”与所谓“定量评价与定性评价相结合”的含义是有本质区别的,关于这一点,下文将详加分析,这里只是强调,定性评价与定量评价各自的独立乃至互相对立,只能使学术评价因分裂而走入歧途。

### (三) 评价分裂的后果

因评价方法上升为评价类别而出现的评价分裂,对学术评价造成的最严重后果就是使学术评价变得残缺不全。

在学术评价中,评价主体是否具备学科专业优势是起决定作用的,所以,学术共同体这一评价主体的身份毋庸置疑。在评价发生分裂之前,同行评议并不意味着对主观判断的必然滥用和对定量评价的必然排斥,同行学者不能进行定量评价是因专业评价机构的产生而制造出的一种假象。今天的同行评议,受到两方面的制约,一是在事关利益的评价中,同行评议对评价环境和评价主体自身的道德水准有苛刻的要求。若没有一个干净的学术环境,单纯的定性评价可能比单纯的定量评价更加荒唐。二是在学术成果(作品)数量巨大的时代,离开了数据统计和定量分析,同行评议的效率将十分低下,不确定性也必然增加,要提高效率和保持一定的稳定性,同行评议除了经验判断外,也离不开数据分析。当然,同行所做的数据分析与评价机构的定量评价并不是一回事。评价分裂后的同行评议,既缺乏清洁的外部环境,又割舍了定量分析,当然无法令人信服。

在评价分裂后,独立的定量评价主要由专业评价机构作出,其评价主体的身份是令人怀疑的。如果说,只有学术成果(作品)的读者和使用者才具备评价主体的资格,那么,评价机构既不是读者,也不是使用者,其对评价对象各种形式数据的统计大多与阅读和使用无关,应该是连起码的评价主体的资格也不具备的。由于不懂专业的根本性缺陷几乎无法靠自身的努力得到足量的弥补,所造成的外行评价内行,很可能失之毫厘而差之千里;对评价信息的数据化处理虽然是其优势,但不具备学科专业知识,其数据分析难以深入内容层面,内容信息的缺失也使得评价机构的定量评价只有数据,只有运算,只有排序,却没有学术,也没有真正的价值判断。仅此两点,都足以致命。因此,在传统的评价环境内,评价机构的定量评价绝不可能是一种独立存在的评价。

在现实评价中,数据运算与经验判断已成为两种对立的评价,在这背后,其实隐含了这样一种理论预设:定性评价与定量评价应该由不同的人或机构分别来做,因为从事定性评价的学术共同体做不了定量评价,而专业评价机构也做不了定性评价,所以将学术评价分为各自独立的定性评价和定量评价是有道理的。这样的理论预设是值得怀疑的。第一,专业评价机构的确没有能力做定性评价,但学术共同体并非不能做定量评价,与定性评价的专业门槛高企不同,定量

评价并没有高不可攀的技术门槛,学术共同体在经过必要的训练后,完全有能力做定量评价。第二,专业评价机构不仅没有能力做定性评价,其实也没有能力独立完成定量评价,定量评价包含有数据运算与价值判断这两层意思,评价机构的最终产品,即其制作的所有排行榜和排名表都只是运算的结果,而不是专业分析和价值判断的结论,充其量也只是未完成的评价。第三,成功的定量评价离不开学术共同体的参与和主导,从数据挖掘的方向、数据评价意义的设计、数据运算的公式,到运算结果的解析和最终具有价值判断意义的评价结论的得出,哪一项也离不开学术共同体。

可见,评价的分裂和分裂后定性和定量评价的互相对立,不仅使同行评议沦为狭隘的主观判断,而且使定量评价变成了一场数字游戏。评价再也无法统摄于同一评价主体,残缺不全也就是必然的了。这样的结果不仅导致了现实中学术评价的不合理,而且也直接导致了学术评价自身的危机与困境。那么,大数据时代的到来是否有助于学术评价走出分裂的困境?

## 二、大数据的介入:机遇还是陷阱?

当我们将改进学术评价的希望寄托于大数据时,面对的却是分裂的评价,分裂必定导致评价的残缺,但统合后的评价未见得就一定科学合理,因此,对大数据之于评价作用的判断,需要回答两个问题:一是大数据能否将分裂的评价统合起来;二是统合后的评价能否借助于大数据变得科学和合理。

### (一) 信息、数据与评价的关系

在回答大数据对评价的作用这个问题之前,有必要重新讨论评价信息、评价数据与学术评价的关系。信息与数据的作用主要体现在评价过程中,所以,让我们从评价的过程说起。无论在前大数据时代还是大数据时代,学术评价的过程简单说来实际上就是这样几个步骤:(1) 确定评价的目的(必须是具体的而不是抽象的目的);(2) 确定评价对象(评价客体);(3) 确定谁来评价(评价主体);(4) 确定评价方法、程序和标准;(5) 评价主体收集与评价客体相关的各种可能有用的信息;(6) 根据确定的方法、程序和标准,对所收集的信息进行比较、分析和处理;(7) 得出评价结论。

在一般理解中,似乎只有评价方法和作为评价标准的指标体系才与数据相关,其实不然。如果说,评价数据来源于评价信息,那么,以上所言评价过程的七个方面,每一个方面都与评价信息有关。比如,确定评价目的时,除了考虑此项评价的必要性外,还要考虑可行性,即实现此项目的是否有必要的信息作为评价的基础;再如,确定评价主体时,必须有一个评价专家数据库,可以提供专家的各种相关信息;至于信息收集和处理,更是直接作用于信息;而评价结论其实也就是评价者通过信息处理而对评价对象作出的价值判断,理应包含丰富的信息。所以说,信息是评价的基础,没有相关信息,也就无从评价。

信息、数据与评价三者间的关系主要体现在转换和分析两个方面。首先,是对与评价有关的信息进行数据转换。在前大数据时代,这种转换基本是指向性的和局部性的。所谓“指向性”,指的是具有明确指向的信息(高密度高价值信息)到数据的转换,比如文摘、引用等指向明确的信息,转换为可以运算的数据,而离开了这些有明确指向的信息,数据生成几无可能,这与对数据的认识以及挖掘和处理数据的能力有关。所谓“局部性”,指的是只是部分信息的部分方面被转换为数据,比如文摘和引用信息都属于相对比较容易转换为数据的,

转换的指向性和局限性依然存在,但是,由于对数据认识的深度和数据挖掘的广度大大增加,处理的技术更加先进,我们有理由期待,除了指向明确的信息转换为数据以外,还有可能出现低密度高价值的新数据源;我们也有理由相信,可以有更多的信息能够比较全面地转换为数据。其次,是对由信息转换而来的数据进行运算和分析并得出结论。在前大数据时代,由于信息到数据转换的指向性和局部性,数据不可能完全代替信息,因此,数据对评价的作用是有限的,单纯的数据运算结果也是代替不了评价结论的。在大数据时代,不管数据量有多大,也不管运算方法有多先进,单纯的数据运算仍然无法代替价值判断,但是,数据的运算结果可以更好地用于学术评价,特别是对数据的利用不再是专业评价机构的专利,数据对同行评议同样有意义,甚至有更大的意义。因此,引入大数据,的确会对学术评价发生重大影响。

## (二) 大数据对评价的影响

“大数据对社会经济生活产生的影响绝不限于技术层面,它为我们看待世界提供了一种全新的方法,即决策行为将日益基于数据分析做出,而不是像过去更多凭借经验和直觉做出。”“这是一场革命,庞大的数据资源使得各个领域开始了量化进程,无论学术界、商界还是政府,所有领域都将开始这种进程。”<sup>[4]</sup>显然,大数据时代的到来,凭借巨量数据的获取和运算的模式正在颠覆“凭借经验和直觉”的模式,而被广泛地运用于判断、预测和决策之中。如果数据的量足够庞大,覆盖的面足够宽广,挖掘得足够彻底,处理得足够精细,那么,这种凭借“云计算”得出的评价性结论或预测性判断就可以相对准确,在某些领域或某些问题的判断上,它有可能胜过经验和直觉。

正是基于大数据的特点,在学术评价领域,当“凭借经验和直觉”的定性评价与凭借客观数据的定量评价孰优孰劣争得不亦乐乎之时,人们似乎找到了一把打开这个几乎成了“死结”的问题之锁的钥匙。比如,有评论者认为“在‘前大数据时代’,期刊的评价方式多采用专家定性评价、内容评价……评价的模糊性与主观性是这一阶段期刊评价存在的主要问题。主观性指评价缺少客观依据和数据支撑,具有想当然的嫌疑;模糊性则因为评价依据不能使人信服而导致评价不够精确”,“这些‘前大数据时代’期刊评价存在的主要问题,随着大数据产生而得以‘缓解’。比方说,由于大数据的支撑,定量评价、形式评价等客观的评价方式可以得到更好的实现……促使学术评价更趋于精确性与客观性。从某种意义上来说,这是对‘前大数据时代’期刊评价的超越。”<sup>[5]</sup>这里说的虽然是学术期刊评价,但因为现行学术评价机制的核心就是学术期刊评价,通过对学术期刊的评价来间接地评价学术成果,即发表在学术期刊上的论文,在没有找到使人信服而又操作性良好的直接评价学术论文的方法之前,这样的“以刊评文”实则就是学术评价的主要表现。显然,大数据技术被寄予厚望,希望通过它可以大大提升定量评价的精准度,以实现评价的“超越”。

不难看出,以上对大数据的期待,仍然是建立在定性评价与定量评价相分裂和相对立基础上的,由于定性评价被认定为“缺少客观依据和数据支撑,具有想当然的嫌疑”,自然就会推论出大数据对定性评价的影响是极为有限的,亦即大数据挽救不了定性评价,转而对大数据时代有“客观依据和数据支撑”的定量评价寄予了厚望,期望它能超越“前大数据时代”。也就是说,要通过加强定量评价,使之压倒性地战胜定性评价,从而使分裂的评价一方“吃掉”另一方而实现评价的统一。

然而,以上看法并没有弄清信息、数据与评价之间的关系,也就没有理解数据在评价中的真实作用和意义。大数据时代云计算处理的数据不管量有多大,都离不开与其所代表的信息的关系,丢失了信息的数据不管设计了多么复杂的算式,其运算结果都没有任何意义,只会将评价扭曲。由于大数据时代数据的价值虽大,但密度极低,较之前大数据时代,数据的挖掘和运算都更

离不开对数据的解读。这种解读,核心是对数据与信息关系的解读,需要有计算技术,更需要有各学科专业知识。专业评价机构虽然具备计算技术,却永远也无法具备各学科的专业知识,所以,他们可以对基于大数据的评价提供技术支持,但却无力得出科学的评价结论。学术评价不是简单的或复杂的数字游戏,而是一项严肃的科研活动,它必然要受到科研活动一般规律的支配和约束。大数据于科研的意义并非仅及于学术评价一项,如今,许多学科的研究都引入了大数据技术,但这些科研活动的主体,无一例外,都是本专业的科研人员,而不是只懂技术而不懂专业的外行。如果非本专业的评价人员能够对该专业的学术成果作出科学评价,那么,他们也就可以在所有的学科的研究中取代专业人员,而这不啻天方夜谭。学术共同体之所以被国际学术界公认为最合适的学术评价主体,更多的是因为只有这样的评价才具有专业水准,才能对学术研究作出真正的贡献。

到目前为止,所有有关大数据的理论方法和实践经验并不能直接应用于学术评价。大数据理论和方法在其他领域的成功运用并不意味着它就能成功地移植到学术评价中来。学术评价特别是人文社会科学学术评价较之一般运用于生产领域、经济领域的基于大数据的评价要复杂得多。其不同之一是,不管学术评价的目的和对象如何多元,学术质量都是评价的中心问题,而学术产品与现代工业产品最大的区别在于,前者是十足的个性化产品,很难进行深入到内容层面的关于学术质量的量化比较;后者则是规整的标准化产品,很容易进行包括产品质量、用户体验、市场需求在内的各方面的量化比较。其不同之二是,学术评价主要是对既有成果的价值判断,而生产和市场领域的评价更多的是对市场前景的预测和营销策略的设计,两者对数据的挖掘方向和运算方法因目的不同而存在巨大差异。因此,在生产领域、经济领域中风生水起的大数据理论和方法,在学术评价领域要想一显身手,需要评价理论和方法的及时跟进。其最大的瓶颈就是,必须找到一条能够让个性化产品可以不受干扰地利用大数据技术进行量化数据类比的路,这条路,在前大数据时代根本就没找到。在评价环境没有根本改变、评价理论和方法没有根本突破之前,要找到这条路基本不可能。因此,大数据技术要在学术评价方面真实地投入运用并发挥积极作用,恐怕还有很长的路要走,在现实与愿景之间,还横亘着一条巨大的鸿沟。

### (三) 机遇与陷阱并存

其实,不管是否在大数据时代,包括学术评价在内的所有评价都会受到以下几个方面因素的影响和制约:其一,是否具备成熟的、并且为公众(特别是评价对象)普遍认可的评价理论。其二,是否具备在评价理论指导下成熟的、并且为公众(特别是评价对象)普遍认可的科学评价方法。其三,是否具备学术评价所必需的干净整洁的学术环境。其四,是否具有高学科专业水准和公信力的评价主体,能够担当起进行公正合理的学术评价的责任。其五,能否收集到足以支持评价方法的准确的评价信息。其六,是否具备对评价信息进行有效处理的能力。其七,是否具有对处理结果进行专业解释的能力。只有同时满足了上述条件,才可能得出科学合理的评价结论。

笔者并不怀疑大数据技术将会给学术评价带来某种机遇,大数据以及相应技术显然比前大数据时代更容易部分地而不是全部地满足上述第五和第六个条件,从而有可能在评价中更好地发挥数据的作用。但是,科学评价所必须具备的七个条件,特别是前四个和最后一个条件并不会因为大数据时代的到来自然而然地就具备了,即使是第五和第六这两个条件,也不仅仅是数据的量级和运算的能力问题,所以也不会因为大数据时代的到来就一定能取得突破。

在分裂的评价中,专事定量评价的各评价机构几乎垄断了所有的数据挖掘与运算排序,而这些机构显然不具备上述基本条件,撇开前四个条件不论,仅就后三个条件而言,即使其最擅长的数据处理也不可能完成,因为他们没有能力对数据及运算结果予以专业解析,只能将经运算



排序的结果作为评价的结论发布出来,评价的过程在此已告中断,今天的各种排行榜和排名表大多是这样的未完成产品。没有学术共同体的主导<sup>①</sup>,这样的所谓定量评价,不仅无法发挥大数据的威力,而且已成为大数据在学术评价中应用的障碍。

正因为如此,过分地迷信评价机构的大数据技术,就有可能掉入工具理性的陷阱。“当前我国社会经济、文化、信息越来越向全球化发展,各行各业都在进入科学管理时代,可以说工具理性已经逐渐渗透到社会生活的方方面面……从西方科学技术体系移植的各种科学计量工具以一种学术研究的姿态凛然应用于中国学术管理的方方面面,信息技术的不断进步使得海量的信息都可以被储存、被分离、被分析,于是核心期刊排行榜、来源期刊排行榜、中国大学排行榜以及研制各种名号排行榜的技术门槛已经不是问题,越来越多的研究机构和传媒企业甚至个人都以无比的热情投入到对他人的排序中。”<sup>[6]</sup>然而,这些基于大量数据及运算而制作的各种排行榜和排名表的科学性却始终受到学界的质疑。比如在定量评价中普遍使用的引文分析法的“合法性问题”在学术界就一直争议不断,“向来不乏学者对以引文分析为基础的量化学术评价提出异议或批评。争议源自于学术界对引用行为和引文本质没有形成普遍性的认识论共识”<sup>[7]</sup>。可见,在评价理论不能让学术界信服、评价方法不能得到科学验证之时,大数据技术纵然能增加数据的量以及运算方法,其在学术评价中也难有用武之地。不仅如此,在定量评价一家独大的情形下,数据的作用往往被片面夸大和扭曲,大数据如果盲目地运用,只会围绕评价对象制造出更多的形式数据,而深入内容层面的数据仍将付诸阙如,其片面性和扭曲度将更为加大,亦即前大数据时代定量评价的缺陷将被成倍地放大,这样的评价,恐怕只会添乱。从这个意义上来说,大数据不仅不能让定量评价更精细和准确,相反,会让已扭曲的定量评价走向极致。

与任何新技术一样,作为一项新技术的大数据的运用也具有一定的条件限制,不具备必要的条件,机遇就可能成为陷阱,而大数据之有可能成为学术评价的陷阱是缘于评价的分裂,因此,要抓住大数据提供的机遇以改进学术评价,要务就是让分裂的评价走向统一。首先,要破除只有评价机构才具备挖掘数据和运算能力的神话,各专业学者既然有能力将大数据运用于学术研究,就有能力将大数据运用于学术评价,定量评价不是评价机构的专利。其次,要看到所有评价机构发布的各种排行榜和排名表都不是严格意义上的定量评价,更不是完整的学术评价。再次,要看到迄今为止,所有评价机构所做的与数据和运算有关的工作都仍停留在前大数据时代,其发布的排行榜和排名表所依凭的基本都是指向性明确的单一数据,其算法也基本是简单的加权统计。对于评价机构来说,如何真正运用大数据技术,其难度一点不比专业学者小,甚至更大,因为不懂学科专业,就搞不清什么数据是有价值的。在这方面,他们恰恰需要专业学者的指导。最后,大数据运用于学术评价,只能由专业学者主导,评价机构的所谓定量评价不管它的数据量有多大,运算有多精细,也只能是学术评价的一个工具。纵然大数据有可能使这个工具更为精细,但是,工具永远不可能完全代替人。因此,我们在谈论大数据对评价的积极作用时,评价的统一是必要前提。

### 三、大数据之于学术评价:从分裂走向统一?

以上我们分析了学术评价的现状及其问题的症结所在以及数据与评价的关系,显然,大数据可以给学术评价带来某些机遇,但同时也存在着某种陷阱,那么,应该如何抓住大数据的机遇

<sup>①</sup> 目前专业评价机构大多号称其排行榜和排名表是“定量评价与定性评价相结合”的产物,因为他们引入了同行专家参与他们的制作。而事实上,专家在其中的作用是很可疑的。专家充其量也只是起到了“打分机器”的作用,根本谈不上“专家主导”,与学术共同体为主体的评价更是相差甚远。

并有效避免其陷阱?在笔者看来,关键是要让分裂的评价走向统一,为此,必须处理好以下几个关系。

### (一) 大数据时代学术评价与“前大数据时代”学术评价的关系

目前,对现行评价的诟病可谓不绝于耳,批评的焦点多集中于现行的评价方法,改造或重建学术评价的努力也主要体现在评价方法层面,似乎只要找到合理的方法,评价问题就能迎刃而解。大数据正是因为有可能带来评价方法的改进才被寄予厚望。然而,评价绝不是单纯的方法问题,大数据并没有也不可能改变学术评价的基本属性,除了为学术研究服务以外,学术评价在今后相当长的时间内仍将是学术资源分配和学术管理的依据。因此,大数据时代的学术评价,并不意味着能成为割断历史的全新评价,而只能是过去评价的一种延续。这就决定了这样一个基本事实:过去评价中存在的基本问题,大数据时代同样存在,这些问题不解决,是改变不了评价现状的。

同时,大数据时代又是一个以技术革命为突破口的观念革命时代。大数据并不是一项孤立的技术,它的产生与信息时代的到来是紧密相关的,以互联网思维为标志的新思维不仅孕育了新的产业,而且正在使许多传统行业发生变革,所以,我们有理由期待基于大数据技术的新的评价理念的产生,来推动评价理论的创新,破解评价中的既有问题。但是,迄今为止,作为互联网思维最好体现的新技术、新模式和新平台在学术评价中尚未得以呈现,学术评价与新的时代仍然格格不入,因此,如何通过引入大数据实现评价理论的革新才是突破旧问题的关键。

### (二) 科研体制与学术评价的关系

科研体制与学术评价之间存在着一定的因果关系,从某种意义上来说,有什么样的科研体制,就有什么样的学术评价,学术评价的问题恰恰是科研体制问题的表现。在科研体制不变的前提下,要想让大数据时代的学术评价走出定性评价与定量评价分裂的状态,构建健全的学术评价体系,难度极大。当然,评价体系对科研体制也会有反作用,评价的革新,可以倒逼科研体制的改革,但这有待于学术界的广泛参与。最本质的学术评价乃是内容评价,而最好的内容评价来自于学术共同体内的学术批评。缘此,评价体系的重建只能寄希望于学术民主,而不仅仅是某种方法和技术的突破。学术是否民主,民主的程度如何,有一个标志,那就是看是否具有正常的学术批评。然而,在目前的科研体制和评价机制之下,我们还看不到正常的学术批评的回归。

如果说,大数据能对学术评价产生一定的正面影响,那么,大数据首先要在学术批评中发挥作用。大数据无疑是可以用于学术批评的,可以为学术批评提供更多的资源和便利。现在的问题首先是,谁来做真正的学术批评?其次才是大数据在评价中的运用。道理很简单,没有学术批评的先行并作为评价的主要信息源和数据源,所有的信息和数据都是无法深入到内容层面的,不能深入到内容层面的评价,纵然再怎么“客观”、“公正”,又能有多大意义?因此,正常的学术评价何以溃退?如何回归?这的确值得一直在批评学术评价现状的学术界深思和反省。

### (三) 学术评价内部各要素间的复杂关系

#### 1. 学术共同体与专业评价机构的关系

如果我们认同学术评价的信息应源于学术批评,那么,谁是评价主体的问题就不言而喻,只有学术共同体才是合适的学术评价主体。今天来自学术共同体的评价即“同行评议”之所以因极度“异化”而陷入困境,一个十分重要的原因就是正常的学术批评的缺失,致使同行评议失去了最基本的依据和最有效的监督。如果学术民主和学术批评能够走向正常,“同行评议”真正成为学术共同体意志的体现,那么,以学术共同体为主体的学术评价机制就有望重建。与学术共同体相比,专业评价机构因不是共同体成员而缺乏的恰恰是专业性,所以不可能是合适的学

术评价主体。这是一个基本的判断,但作这样的判断,并非将专业评价机构排除在学术评价之外,专业评价机构自有其独特的作用,那就是评价数据的生产者。其实,每个评价机构都制作有大量的评价数据,大多建有自己的数据库,只是这些数据库到目前为止都没有很好地发挥应有的作用。一方面,评价机构为自己的数据库的开放设置了种种障碍,数据制作也远非透明,其数据来源、统计口径等也是各行其是,缺乏公认的科学标准。另一方面,学术界对数据库的价值还不了解,数据库还很少为学术研究和学术批评所利用。这就造成了专业评价机构依凭其对数据库的垄断来争夺评价主体的地位、却遮蔽了数据库的真正价值这样的买椟还珠的结果。因此,正确的选择应该是专业评价机构从评价主体回归其数据生产者的定位。在大数据时代,专业评价机构在数据的提供、算法的设计等方面都是大有可为的,但其作用也仅限于此,真理多行一步就会变成谬误,所有的数据及算法都只是为学术研究和学术评价提供工具,而不是评价本身。利用这些数据来进行研究和批评或评价的应该是各专业学者,亦即各学术共同体成员。如何认识和利用数据这一工具,是大数据时代各专业学者面临的新考验,这也是他们回归学术评价主体地位必须跨越的一道门槛。

## 2. 定性评价与定量评价的关系

在今天的评价体制中,定性评价和定量评价已被视为两种不同的评价,由此产生了一系列问题,于是,从管理部门到学术界以及专业评价机构都提出了定量评价与定性评价相结合的原则。之所以要“两结合”,一般的解释是让这两种各有所长、也各有缺陷的评价互补,以使评价变得相对全面和公正。这种已被普遍接受的观点貌似公允,其实似是而非。如前所述,其要害在于分裂了的评价纵然能够“两结合”,也只能貌合神离。比如,在一些所谓的“两结合”评价中,专家实际只是根据量化指标进行打分的机器,其学术专长无法发挥,徒有其名。在我们明确了学术评价的主体只能是学术共同体后,就不难发现科学的评价并不在于定性评价与定量评价这两种评价相结合,而在于让定性评价和定量评价都回归到一般的评价方法,由评价主体根据不同的评价目的来选择和确定最合适的方法。这个方法,可以是定性的,也可以是定量的,更多的时候则可能既有定性的又有定量的。如同学术研究一样,研究者可以根据不同的研究目的和所掌握的不同资料,采用不同的方法。方法可以多种多样,但不同的方法只能由同一个主体来统摄,而不应分别成为不同主体的专利。由此观之,大数据时代意味着用于评价的资料更为丰富,相应地,评价方法也会更趋多元,定性评价与定量评价都只是方法之一种,而且,随着数据的精细化、多元化,定性与定量之间的分野也会日渐模糊,而不会如现在这般泾渭分明。能够灵活而准确地选用和统摄评价方法的主体显然只能是学术共同体。

## 3. 评价目的与指标体系的关系

学术评价的目的是多种多样的。无论何种评价,其能否成功进行,评价目的的清晰明确都是具有决定意义的。“不同的评价目的决定着不同的评价标准和指标、不同的评价方法和评价专家的选择以及评价程序的确定,它是评价的龙头,分类评价的动因,规定、制约和导引着整个评价的方向和具体做法。”<sup>[8]</sup>但自从蜕变为专事为行政权力部门分配学术资源和进行科研管理的工具之后,学术评价的发展一直有一种追求或趋向,即忽略评价目的的不同,而刻意寻求通适性的评价指标体系,让一个评价指标体系能够适应所有目的的评价,以使资源分配和科研管理更为直观和简单。制定“科学合理的评价指标体系”成为行政权力部门的明确要求,并写入了一些行业指导类的文件,比如原新闻出版总署制定的《全国报纸期刊出版质量综合评估指标体系(试行)》就是这类文件的典型,“总署赋予了这个指标体系两个鲜明的特征:其一,纯量化;其二,通适性”<sup>[9]</sup>。这个指标体系,应对的是包括学术期刊在内的全国各种各样的报纸期刊,其评价结果如何,也就可想而知了。这样的指标体系,通适性与纯量化实乃互为表里:因其通适,必

然要求纯量化;因其纯量化,故而能够通适。指标体系本质上只是一种评价标准,指望一个标准普适于不同目的、不同对象、不同内容的一切评价,是不切实际的。这样的指标体系完全抹杀了不同的评价目的对评价方法和标准的选择性差异。在这样“万能”的指标体系约束下,不管评价目的如何不同,定性评价实际已被排除在评价方法之外,定量评价已成为唯一的方法。这已成为当今行政权力部门主导的学术评价中最常见的现象,也已成为前大数据时代评价的一个特征。如果这样的趋势不予改变,随着大数据时代的到来,在通适性的量化评价面前,再丰富的数据特别是所有个性化的数据都将失去意义,评价必定走入死胡同。因此,在大数据时代,必须破除通适性、纯量化评价的神话,根据不同的评价目的,选择不同的评价方法,唯此,大数据才可真正有用武之地。

#### 4. 评价信息与评价数据的关系

大数据作为一种技术,在各种可资评价的信息获取以及信息到数据的转换方面,其作用是前大数据时代无法比拟的。但大数据要在评价中发挥作用,还有以下五大瓶颈需要突破:一是对信息的挖掘,可资评价的信息散布于方方面面,如何将其挖掘出来?二是从信息到数据的转换,如何将具有丰富内涵的信息在尽可能不丢失其意义的前提下转变为数据,对丢失的信息如何弥补?三是对数据意义的界定,每项数据都有其特定的含义,缩小和夸大其意义都会导致评价结果的偏颇,如何作出准确的界定?四是对数据的处理和运算,不同的处理方法和运算公式会得出不同的处理和运算结果,什么样的数据用什么样的处理方法和运算公式最合理?五是对运算结果的解读,对数据的处理和运算都是脱离了原始信息的处理手段,对数据处理和运算结果的解读实际就是让数据还原为带有价值判断意义的信息的过程,如何保证这样的还原不会走样?以上五大瓶颈在前大数据时代即已存在,随着大数据技术的运用,被挖掘数据量必然出现井喷式的激增,突破瓶颈的难度将更大。所有这些,都不仅仅是单纯的技术问题,更是各学科的专业问题。

#### 5. 评价结论与评价效果的关系

不管出于何种目的、采取何种方法的评价,都是要让被评价者从评价中获得教益,否则,评价就失去意义。因此,评价效果的好坏往往通过评价结论对被评价者有无助益及助益多少来体现。即使为行政权力部门服务的评价,也理应如此。今天的学术评价之所以被诟病连连,一个重要的原因就是评价结论过于简单,简单到就剩下一张排行榜或排名表,被评价者几无学术或质量进益方面的收获。但这样的评价在现行的科研体制中却与学术利益甚至学术研究和发表的资格紧紧地绑在一起而变得如此重要,迫使被评价者——如学者、期刊主编等,不得不将评价机构作为研究的对象,以从事各专业学术研究者的智商,“破译”如此简单的榜(表)的奥秘当然不在话下,于是,或者选择激烈的批评,或者选择无奈的迎合,或者两者兼而有之,从而造成了学术界围绕评价而出现的一些怪象,“某些机构……竞相公开打出了期刊评价甚至学术评价的旗号……迫不及待地要充当评价主体”;“学术期刊不再关心学者的需求,而一味迎合评价机构的偏好,甚至不惜造假”<sup>[10]</sup>。评价结论的极简化和行政权力部门简单化管理需求以及由此造成的对评价目的差异性的忽视、对定性评价的排斥和对定量评价的偏好以及由此造成的评价方法的单一、对评价指标通适性和纯量化的追求以及由此造成的评价信息和数据的片面性之间,有着明确的因果关系。换言之,学术体制不改,大数据也难以改变评价结果仍然简单的现状;而简单的结果不仅对评价对象的作用有限,而且难免偏颇,甚至背离真实,评价仍然难逃尖锐的批评乃至猛烈的抨击。

## 四、结 语

现在让我们回到本文开头提出的问题:大数据可能给学术评价带来怎样的影响?

正是在评价分裂的情境下,学术评价迎来了大数据时代。大数据之应用于学术评价,对数据挖掘和运算能力、对数据和运算结果的解析能力都提出了新的要求,而分裂的评价显然满足不了这样的要求,从而为学术评价从分裂走向统一提供了某种契机,而结束分裂状态,也是大数据成功地运用于学术评价的前提条件。但是,被权力和利益撕裂了的学术评价要走向统一,并不取决于技术,而是取决于造成评价分裂的科研体制是否有所改变,如果体制不变,那么,评价分裂依旧,大数据之于定性评价则几无助益,而对定量评价的作用更不容乐观,当面对剧增的数据和复杂的运算不能作出具有专业水准的解析时,定量评价只能与学术研究渐行渐远,越来越背离评价的初衷。

因此,要让大数据在学术评价中发挥积极作用,并借助大数据的推动,重建科学合理的学术评价,前提就是:让评价不再分裂,让方法不再被垄断;让评价机构做其力所能及的事——专心地做数据,而把数据的最终解释权交还给学术共同体;让正常的学术批评重回学术研究之中;让同行评议真正代表学术共同体的意志。而这一切,都有待于科研体制的改革和评价理论的重建,但愿大数据能成为促成这一改革和重建的催化剂。

### [ 参 考 文 献 ]

- [1][4]《带您了解大数据》<http://www.thebigdata.cn/YeJieDongTai/8470.html>
- [2]李剑鸣《自律的学术共同体与合理的学术评价》,载《清华大学学报(哲学社会科学版)》2014年第4期。
- [3]朱 剑《重建学术评价机制的逻辑起点——从“核心期刊”、“来源期刊”排行榜谈起》,载《清华大学学报(哲学社会科学版)》2012年第1期。
- [5]郝日虹 张清俐《寻找大数据与学术期刊评价体系创新的“契合点”》,载《中国社会科学报》2015年2月11日。
- [6]陈 燕 崔金贵《学术评价中工具理性与价值理性的主导成因及平衡机制》,载《清华大学学报(哲学社会科学版)》2012年第6期。
- [7]刘 宇 李 武《引文评价合法性研究——基于引文功能和引用动机研究的综合考察》,载《南京大学学报(哲学·人文科学·社会科学)》2013年第6期。
- [8]叶继元《人文社会科学评价体系探讨》,载《南京大学学报(哲学·人文科学·社会科学)》2010年第1期。
- [9]朱 剑《量化指标:学术期刊不能承受之轻——评〈全国报纸期刊出版质量综合评估指标体系(试行)〉》,载《清华大学学报(哲学社会科学版)》2013年第1期。
- [10]仲伟民《缘于体制:社科期刊十个被颠倒的关系》,载《南京大学学报(哲学·人文科学·社会科学)》2013年第2期。

(责任编辑:任天成)