

Doi:10.11835/j.issn.1008-5831.fx.2020.04.003

欢迎按以下格式引用:田小楚,高山行.论大数据在著作权法保护中的冲突与协调[J].重庆大学学报(社会科学版),2021(1):141-151. Doi:10.11835/j.issn.1008-5831.fx.2020.04.003.



Citation Format: TIAN Xiaochu, GAO Shanxing. On the conflict and coordination of big data in the protection of copyright law[J]. Journal of Chongqing University(Social Science Edition), 2021(1):141-151. Doi:10.11835/j.issn.1008-5831.fx.2020.04.003.

论大数据在著作权法保护中的冲突与协调

田小楚^a,高山行^b

(西安交通大学 a.法学院;b.管理学院,陕西 西安 710049)

摘要:现代网络技术与科技手段使大数据的利用与传播打破了传统的交易模式,使数字时代的盗版行为成为不受时空限制的低成本行为,甚至可能以牺牲“创作”为代价换取信息社会的快速发展。大数据资源法律救济保护的滞后性难题,引发了人们对大数据在著作权法保护中的冲突与协调的思考。为实现对大数据权利人的保护和社会利益最大化目的,文章以大数据及其权属论争为基点,通过法经济学与法哲学的理论范式推演探知大数据的权属划分与价值收益,得出大数据适用于著作权法保护的可行性结论。并借鉴大数据在英美法系与大陆法系法律框架下的经验与共识,解构大数据适用著作权法的助益与阻滞因素,从而尝试以设立“数据权”或与其他法律制度兼容的形式检验法学科学处理现代科学的技术理性,为大数据的法律保护提供更多可能性。

关键词:大数据;著作权法;数据产权;数据权属;复制权;合理使用

中图分类号:D923.41 **文献标志码:**A **文章编号:**1008-5831(2021)01-0141-11

一、问题的提出

在数字时代,大数据权利人维权之术可谓举步维艰。“数字技术使复制和传播的盗版行为几乎无需任何成本”^①。人机交互的进步推动人类分享意识的觉醒,人们以低成本的使用、复制、传播等方式黏着于大数据不限时空的迭代传递中,甚至迷失在以公益批判和阶级意识的生成作为声讨大数据权利人解除“道德奴役”的乌托邦构想中,以至于忽略了“按需经济”是影响实践理性的导发诱

修回日期:2020-03-17

基金项目:教育部规划基金项目“大数据资源治理的产权界定及管理机制研究”(19YJAZH019)

作者简介:田小楚,西安交通大学法学院知识产权研究中心法律治理学专业博士研究生,Email263485437@qq.com;高山行,西安交通大学管理学院教授,博士研究生导师。

①参见 House Report105-551,Part 2,105th Congress,2nd Session,p.25.

因。诚然,各国法律通过不同形式赋予大数据权利人诸如“复制权”“发行权”和“信息网络传播权”等专有权利,并规定了各种法律救济手段以保护其合法权益,但此类方法在数字时代已经很难有效地起到预设的保护作用,原因在于法律救济只是一种事后救济,大数据资源一旦被侵权,在极短时间内便可以形成数以万计的非法复制件被“无偿”利用,法律的滞后性对此局面已无力回天。在此背景下,许多大数据权利人开始通过技术手段开启利己的法律保护模式,然而,相应的破解、规避方式也一脉而生。例如我国发生的首例涉及数据侵权的技术措施诉讼——“精雕诉奈凯案”^②,原告所称的技术措施所保护的客体——数据(文件)——成为本案的争议焦点,该案最终以精雕公司所采用数据文件并非计算机程序或《著作权法》保护的客体为由宣告原告败诉。针对上述问题,本文力图解构引发现实问题中更深层次的原因,推演大数据权属及其特点的规律有可能被挪移的法律领域,从而理性地对待大数据保护过程中的立法调试问题,从而助益著作权法与其他法律制度的兼容性,化解复杂的框架效应。

二、大数据及权属论争

(一)什么是“大数据”?

“大数据”这一概念首次亮相于《自然》(Nature)杂志2008年刊登的题为“Big Data”的专题,并广泛应用于经济、计算机等不同领域。Kusnetzky Dan在What is “Big Data”一文中将大数据定义为:人工无法在短时间内管理、处理、截取并整理成为人类可读的大规模信息。此外,大数据研究机构Gartner亦给出更详尽的定义,“大数据”实则是指具有更强的流程优化能力、洞察发现力和决策力的海量、高产且多元化的信息资产。就大数据的运行轨迹、程序与呈现结果来看,主要包括以下特点:(1)规模性(Volume,耗费大量存储空间);(2)多样性(Variety,来源广泛、形式多样);(3)复杂性(Complexity,不易分割,难以标准量化);(4)多栖性(Polytropic,携带、复制方便);(5)价值稀疏性(Value,知识密度低、价值总量高)。

数据确权作为大数据法律保护的逻辑起点,是建立数据秩序和数据规则的基石,也是大数据法律保护制度的前提要件。由于数据权属界定不清而导致的大数据开发利用无法受法律保护的事实已经成为时下备受关注的焦点。《中国信息资源开发利用指数报告》显示,2017年我国各省(市、自治区)的数据开发利用指标均值为29.83,相较于2009年的均值(32.09)下降了2.26个百分点^[1]。表明我国数据开发利用的整体水平偏低。数据权属立法的缺失,公开主体对自身权利和义务的惶惑与不确定,已严重影响了数据开发利用之深度与广度。为此,数据权利与数据产权适值应定下奥义为实践舆情断然解惑。数据权利作为私权利的表现形式,其人格权与财产权的界定最为复杂。其中数据人格权的核心价值是维护数据主体之为人的尊严,享有隐私不被窥探、自由不受剥夺、信息不被盗用、名誉不受侮辱的权利^[2]。而数据财产权是权利人直接支配特定的数据财产,从而排除他人干涉所享有的占有、处分、收益和使用权利。数据产权的界定则相对复杂。首先,我们需清楚认知什么是“产权”?所谓产权(Property right),即指“一种通过社会执行而实现的对某种财产的多种用途进行选择的权利”^[3]。而这种财产,无论公有、私有,唯有被合理使用,才赋予经济价值。而想要了解产权,必然离不开法经济学和法哲学理论奠基,从而使大数据与产权之间的关系蔚然明晰,得知大数据产权究竟缘起何处,其权属究

^②参见上海市第一中级人民法院(2006)沪一中民五(知)初字第134号民事判决书。

竟花落谁家。

(二) 数据产权的经济学探源

1. 科斯定理与麦特卡夫定律共识

互联网时代下,“摩尔定律”归纳了信息技术快速发展的同时,亦引发了人们对于技术进步伴随海量数据资源滥用所面临的数据权属不清、利益失衡等尴尬困境,冷静思考这一现象所反映的技术失控与法律意旨,不禁担忧大数据权属问题将成为互联网平台之间大数据流转与交易埋藏的“定时雷弹”。经济学家哈罗德·德姆塞茨(Harold Demsetz)认为,只有当大数据的产权收益大于成本时,确权行为才会产生经济价值。要确定大数据的产权收益大于成本,需通过大数据的稀缺性为之增值,固定其财产属性。我们通过科斯定理(Coase theorem)分析大数据产权在不同主体之间资源配置的效率,试图探寻产权与大数据的影响因素为何。

科斯定理 I:在交易费用为零时,当事人之间的市场交易无论怎么进行初始配比,资源配置都会达到帕雷托最优。

科斯定理 II:在交易费用不为零时,不同的产权初始配比会带来不同的资源配置效率^③、^[4]。

科斯定理 I 析出的假设前提旨在提供预设的理论指向,而现实中交易费用为零的可能性极低,科斯定理 II 所升格的产权配比假设更符合大数据的价值指引。倘若个人数据不流入市场、不公开使用,则数据价值与使用效率几乎为零;若一旦将大量个人数据汇集到指定主体,当数量级达到千亿级以上,则数据价值不可估量。当然,科斯定理也并非经济产权领域亘古不变的真理,其片面性与局限性仍无法避免,例如对盲目追求资源配置效率所达到帕雷托最优状态的真实存疑。

倘若科斯定理只是片面追求资源配置效率,我们不妨换一种解释方法——麦特卡夫定律,即:数据价值以用户数量平方的倍速增长,当 V 代表数据总价值, n 为用户数,那么以定律公示 $V=n^2$ 可以推算出数据对于每个人的价值与数据权利人的数量成正比。假设一个数据对数据权利人的价值为 1 元,那么规模为 100 倍的数据总价值为 10 000 元,规模为 1 000 倍的数据总价值为 100 万元,数据总价值恒定为数据规模的平方倍速增长。由此可以看出,无论是科斯定理抑或麦特卡夫定律均表明,数据的初始配比将影响数据资源的最终配置效率,且数据资源越集中效率越高、价值越大。

2. 大数据适用著作权法的模型推演

基于引证大数据产权的法经济学探源,我们不妨关注这样一个问题:大数据是否适用著作权法保护?参考法经济学对大数据交易价值的假设,选择利用“理性人”的谱系理论作出回应的结果是:假设著作权人为“理性人”^④,当理性人趋利避害的本性充分实施于交易过程中,其经济效益最大化、资源集中以及供求定律等基本原则方能发挥作用^⑤、^[5]。但遗憾的是,现实生活中不存在完全理性人。为此,著作权法中的合理使用制度将成为大数据权利人迎合不完全理性人最好的佐证。

模型 I:当大数据的使用行为带来的潜在收益高于交易成本时,不利于纳入著作权法合理使用的范畴。由于数据使用者通过对数据的使用可以带来收益,收益大于交易成本的动机促使数据使用者寻求

^③科斯定理是由诺贝尔经济学奖得主罗纳德·哈里·科斯(Ronald H. Coase)命名。他于 1937 年和 1960 年分别发表了《厂商的性质》和《社会成本问题》两篇论文,这两篇文章中的论点后来被人们命名为著名的“科斯定理”,是产权经济学研究的基础,其核心内容是关于交易费用的论断。

^④所谓“理性人”是人的行为是理性选择的结果,在既有的约束条件下,理性人会追求自身利益的最大满足。

^⑤平衡理论认为,著作权人从自私的角度出发,只会考虑其享有著作权的私人成本,而刻意回避权利的社会成本,甚至故意将私人成本转化为社会成本,在收获私人利益的同时损害社会整体效益。

数据权利人的许可,此模型情境不利于纳入著作权保护的合理使用的范畴。原因在于,将个人数据有偿公开,数据权利人对数据使用者存在潜在的价值收益,若收取高额费用,该交易成本不易被数据使用者所接受。

模型Ⅱ:当大数据的使用行为带来的潜在收益低于交易成本时,有利于纳入著作权法合理使用的范畴。鉴于大数据使用的商业性质并非合理使用原则的决定性因素,对著作权制度的规约作用影响不大。因此,即使是商业性使用大数据的行为也可能纳入著作权保护的合理使用范畴。例如新闻报道所用数据具有较强时效性,难以在短时间获得数据权利人许可,即便获得许可,若许可成本高昂导致新闻价值下降,那么此数据所带来的新闻的社会公益效应则微乎其微。在此情境下,大数据的交易成本不易被公众接受,而公益新闻报道却应当为之,著作权法中的合理使用制度便需要被引入并施以法律保护作用。

模型Ⅲ:当大数据的使用行为未带来任何收益时,可以纳入著作权法合理使用的范畴。该情境最常见的便是纯粹为了个人欣赏、研究、学习等目的使用大数据,这些数据使用行为不会增加明显的收益,如果有也是增加些许社会公共利益。比如因教学目的使用大数据是为了“授人以渔”、传播知识和促进社会知识文明的进步与发展。此类大数据的使用行为不应当产生交易成本,数据使用者也不会对数据权利人给予利益补偿,否则将会损害社会公共利益、阻碍人类文明的进步。

综上所述,大数据适用著作权法保护是可行的。在公众对大数据的使用过程中如若得不到著作权法的保障,那么便会产生不符合市场规律的交易成本(数据使用者与数据权利人进行谈判所产生的费用),导致交易成功率锐减。而著作权法中的合理使用制度有效解决了大数据交易市场中的许可交易失败隐忧,克服了市场失灵的弊端与阻滞困境。鉴于此,通过法经济学的分析,大数据适宜于纳入著作权法的保护范畴。

(三) 大数据著作权合理性之法哲学范式

对于财产权的法哲学谱系理论可以参考约翰·洛克^⑥(John Locke)的财产权劳动学说。洛克认为,财产权是“天赋之权”,上帝将一切给予人类并为之共有,人类进步所演化的劳动分工又使人们所共有的一切拨归私用,社会为保障公平与正义,以法律的形式创造了私人所有权。这一理论为解释大数据这一无形财产权的合理性及适用著作权制度的正当性奠定了理论基础,并揭示了法律这一系统容许著作权与大数据之间的因果联系。对于著作权制度而言,以劳动分工为基础和前提确立的财产权制度,本质上激发了著作权创造者的积极性,为人类社会之发展增添动力。财产不单纯是物质的派生物,也是人类选择的反映,有形财产抑或无形财产均为人类原初状态的劳动产物,大数据也属其中一种。虽然财产权的重要性正逐渐被使用权所粉饰,分享意识的启蒙对于大数据成果的保密、设限等做法极力反对,声称一切消减共享经济的规范都是违背人类进步的符咒。但著作权作为唯一包含人格权的知识产权类型,不会迁就某些以“人类”“国家”“集体”为受益主体的“投机”立场,毕竟许多问题及观点都缺少旁征博引的求证与深入骨髓的辨析以证立真伪。即便社会倡导知识共享,即便法律不保护它无力保护之物,但大数据权利人所付出的合法的、积极的劳动收益相较于数据使用人无偿获取的利益,谁更值得保护不言而喻,当然,法律应给予人类进步齿轮中无限轮回、创造的劳动者(权利人)应有的尊

^⑥约翰·洛克是17世纪英国资产阶级哲学家和政治法律思想家,自由主义的奠基人,古典自然法学派的杰出代表之一,创立了对后世影响深远的财产权劳动学说。

重与保护。

黑格尔的劳动理论与人格正当性理论极为相似,被用于著作权时带着直觉般的吸引力和理想主义,他认为“人有权将自己的自由意志体现在任何物中,因而使该物成为我的东西;这就是人对一切物据为己有的绝对权力”^[6]。这一论断并非单纯地主张创造者自我价值实现是至高无上的,运用于著作权制度中,可以理解为著作权无需通过与有体财产权的类比加以论证,否则可能会歪曲与意志有关的人格和心理特征的尊崇地位。黑格尔的劳动理论为著作权保护客体的扩充提供了合理性论证,即大数据作为人类意志的体现,当然可以成为著作权应保护之物。

建立公共福利基础之上或工具主义的机理理论是以“激励”智力创造成果为切入点,为大数据著作权提供正当性的另一种重要理论。在激励理论中,大数据著作权保护所需要论证的合理性可归结为经济需要、精神需要和社会需要三个面向。以上需求合理性的满足条件,均以激励大数据权利人创作为基础和前提。法产生于利益的需求源自自然法学派“无救济,无权利”的社会契约论——人出于保护生命、财产等基本权利的需要,自然让渡部分权利以达成人人应当遵守的公共契约^[7]。而著作权法的产生根源于保护著作权人的所有权和促进人类文明的传播,然而这两者的利益平衡主要参照的是对著作权人利益的考量。通常,著作权法在赋予个人复制作品权利的同时,也会限制其获取著作权作品的行为,但大数据发表简易、使用方便、复制便捷、获取手段多样等特点,使得大数据的法律保护专属性被弱化。若想实现公众利益的有的放矢,需要著作权法在网络安全与数字技术的运用发展中增强保护。

三、大数据适用著作权法之域外引论

在大数据时代,权利人如何运用著作权法保护正当权益的破题关键在于适用著作权法的边界定位,基于大数据信息处理流程的四个阶段,即:采集与预处理—存储和管理—处理和分析—应用。其中,第二、四两个阶段可以调用著作权制度的层级结构寻求保护。首先,第三阶段“处理和分析”所形成的数据成果在进行第四阶段的数据“应用”时,出于具象化的数据成果形式主要以软件方式存在,符合著作权法对软件作品保护的范畴。其次,第二阶段“存储和管理”过程中,数据经过集合、汇总,形成了数据库、数据仓库、云数据库等形式,对这一部分的数据利用著作权法保护值得探究,关注焦点在于如何定性未固定成果形式的大数据与著作权法保护客体范畴的边界弹性。此时,“限制和例外”制度便成为大数据之于著作权法保护栖身立命的正当性工具,而唯一与大数据息息相关的便是著作权法中“汇编作品”(compilation)这一种类。

严格讲,汇编作品在著作权法中并非单纯的作品种类,几乎任何一类作品(除立体美术、艺术作品外)都能以汇编的形式出现^[8],而汇编作品本身也居于作品分类中的其中一类。那么,大数据作为数据的集合,其构成形式可能是作品、作品的片段或其他形式,甚至可以理解为汇编作品是大数据组成的一部分。从而,人们试图通过汇编作品这一“媒介”,将大数据以谦抑性的方式纳入著作权法保护的客体范畴。

法学上的论证虽是普通性实践论证的一种,但自有其特色,就满足著作权作品的先决条件——“独创性”标准而言,其论证的形式和规则受到不同法系的限制。在英美法系国家,著作权法中的作品“独创性”标准是司法认知作品来源、区别其归属问题的判断要素,作品所蕴含的经济价值和社会效益才是左右立法者考量的决断因素。可以看出,英美法系国家的著作权法对作品的“独创性”标准要求

较低。例如美国 H. R. 3531. Sec. 3(a)^⑦ 法案规定:“数据特殊权利保护的标准是在数据库内容的组合、组织、收集、证明或表现上有质量的或数量的实质性的技术、人力、财力或其他资源的投资。”这一规定与 WIPO 数据库条约草案和欧盟数据库指令相类似,成为大数据在有效力的法规上进行法律治理下的概念诠释。值得注意的是,这一条文所采取的措施是以特殊权利保护的“投资”取代了著作权法的“独创性”标准,意味着无法纳入著作权法“独创性”标准的大数据将会落入特殊权利保护的范畴,这与前文法经济学就大数据产权价值的初始配比分析相契合,使大数据的实质性“投资”成为法律适用合理性的推论和证明。

大陆法系国家与英美法系国家在著作权“独创性”标准制定上反差鲜明,前者的保护标准更高,标准制定更为严苛。大陆法系国家主张,作品的“独创性”标准显著区别于实用艺术、外观设计等标准,且需要同时兼顾著作权人的财产权利、经济利益与精神权利。这无疑成为大陆法系国家著作权保护高标准的审查门槛。以大数据而论,其个性创作的程度需要达到或高于“艺术性”创作的高度才有可能赢得客观新颖性审查的要求,成为作品类型被纳入著作权法的保护范畴。

早在 1991 年,世界知识产权组织(WIPO)便肯定了《伯尔尼公约》第 2 条第 5 款有关数据及数据库受著作权保护的规定。这一规定明确将“汇编作品”类型纳入“选择”或“编排”的以计算机储存形式所固定的数据及数据库,实现了立法者初设法律体系时有关大数据问题的漏洞补位。此外,《TRIPS 协议》第 10 条第 2 款^⑧所特指的经过“编排”或“选取”数据只要能体现智力创造的过程,无论以何种方式呈现都应当受到著作权法的保护,该条款中“数据或其他内容的汇编”不仅是《伯尔尼公约》第 2 条第 5 款“文学或艺术作品的汇集本”的上位概念,而且就“文学或艺术作品”以外的其他材料的汇编成果,已然也成为了完全意义上的大数据。与此同时,《世界知识产权组织著作权条约》(简称 WCT)第 5 条“数据汇编”的标题后注明“数据库”字样,再一次将“大数据”的国际著作权法地位赋予了清晰明确的确认与解释。

当然,《欧盟数据库指令》对数据的结构问题也有所涉及,“鉴于”部分(15)规定^⑨认为,著作权法对数据的保护标准设定不仅要考量数据库内容的“选择”与“编排”,更应当注重数据库结构是否达到智力创造的条件。而“鉴于”部分第 35 规定则补充了数据及数据库结构的例外选择与限制行为的例外,对此,也可以将“限制行为的例外”理解为著作权法中“合理使用”的情形,类推适用目的性限缩的法理解释。换言之,欧盟成员国对于大数据这一规定目的在于,为大数据提供著作权法保护的同时,又利用“合理使用”制度将其排除出“侵权行为”的范围。

四、大数据适用著作权法之制度殊异

著作权法与专利法和商业秘密法不同,在大数据权属划分的问题上,著作权不对方法和过程进行保护,只对其最终表现的智力成果形式采取排他性保护。当然,局限性也不可避免,为了进一步促进大数据运用著作权制度保护的优越性,需先清晰认识与定位制度适用过程中的隔阂与矛盾。

^⑦美国 1996 年 5 月 23 日提出了《1996 年数据库投资及凡知识产权侵权法》(Database Investment and Intellectual Property Antipiracy Act of 1996),即 H. R. 3531 法案。

^⑧《TRIPS 协议》第 10 条第 2 款:“数据或其他内容的汇编,无论是采用机器可读方式或者其他方式,只要其内容的选取或编排构成了智力创造,就应对其提供保护。这样的保护不应扩展到数据或内容本身,不应影响对数据或内容本身所获得的任何的著作权。”

^⑨《欧盟数据库指令》“鉴于”部分(15)规定:“鉴于用以确定数据是否可以获得著作权保护的标准应规定为数据库内容的选择于编排的实际是作者自己的智力创造;鉴于这种包含应包括数据的结构。”

(一) 法律规定的原则化

前文已对不同法系国家及国际组织有关大数据著作权法保护相关立法进行了梳理与归纳。不难发现,无论采用何种立法模式,这些国家均对大数据的定义、保护范围、例外情形,以及与“合理使用”冲突的机制均列有明文规定,构建了一个相对完备的法律试运行体系。相比之下,我国的大数据立法则成为法律治理环境下的附庸之物,其法律规定显得过于原则化和简单化。具体而言,与大数据相关的法律规定也不过是《著作权法》《计算机软件保护条例》《反不正当竞争法》的个别条款在特殊情况下的适用。例如《著作权法》对汇编作品保护的主张是将故意避开或者破坏权利人为其作品或其他受保护客体采取的保护著作权的技术措施定为侵权行为。就此来看,尽管我国《著作权法》借鉴发达国家的立法,对“汇编作品”的定义、保护范围和例外情形有较详尽的规定,但这并不能完全弥补《著作权法》对新兴技术适用过于原则化的缺陷。

此外,《著作权法》(2020)第15条^⑩虽然对“汇编作品”定义充分诠释,但其中的“独创性”也仅仅包括“对数据材料内容的选择或编排”,而不包括“不构成作品的数据库”,因为经过编排的数据材料是用来“运行”的,不是用于“浏览、欣赏”的。如此一来,对于大数据的合理使用究竟如何界定,《著作权法》同样未予以解释。

(二) 例外情形规定虚化

我国《著作权法》以汇编作品形式纳入大数据这一客体的例外情形可谓佶屈聱牙,主要在于未列明对大数据禁止规范的具体例外情形,且《著作权法》(2020)第53条^⑪所列举的可适用大数据侵权行为采取“技术措施”的限制都无法解决大数据权属与“合理使用”之间的冲突^[9],立法本意在解决现实问题上如蜻蜓点水,仅形成微澜,缺少精微深奥的保护屏障。

此外,就“汇编作品”而言,《著作权法》《世界知识产权组织版权条约》(WCT)第5条规定和《与贸易有关的知识产权协议》(TRIPS协议)第10条第2款的规定立场相同,即著作权保护延及大数据。然而,却并未明确将经过“编排”“选择”的数据纳入“汇编作品”这一类别,极具价值的数据库亦未于法有据地被全面覆盖,致使大数据适用著作权法保护条件过高。

(三) 侵权行为定性混同化

鉴于著作权法主要针对大数据形成的“汇编作品”进行保护,大数据筛选作为汇编作品中“挑选”的形式之一,符合原创性要求,但这种保护并不能有效减少非必要性的复制行为。原因在于数据的易获取性使得抄袭者或非法用户轻而易举地盗取数据资源,且不需要复制数据库内数据的具体选择或者编排。此外,根据《著作权法》(2020)第53条规定,直接规避技术措施的行为并不构成侵害专有权利的行为,即使禁止直接规避行为,也不能将该行为认定为是对大数据汇编作品的直接侵权或间接侵权。而《著作权法》将直接规避技术措施的行为定为“侵权”,极易引起误解,使人们误以为未经许可的大数据汇编作品直接规避行为与未经许可的复制和传播一样,都是侵害专有权利的行为。

而类似于大数据的软件作品抄袭、屏蔽、恶意篡改等著作权侵权行为也常有发生,侵权方有可能会

^⑩《著作权法》(2020)第15条规定:汇编若干作品、作品的片段或者不构成作品的数据库或者其他材料,对其内容的选择或者编排体现独创性的作品,为汇编作品,其著作权由汇编人享有,但行使著作权时,不得侵犯原作品的著作权。

^⑪《著作权法》(2020)第53条第6款规定:有下列侵权行为的,应当根据情况,承担……(六)未经著作权人或者与著作权有关的权利人许可,故意避开或者破坏技术措施的,故意制造、进口或者向他人提供主要用于避开、破坏技术措施的装置或部件的,或者故意为他人避开或者破坏技术措施提供技术服务的,法律、行政法规另有规定的除外……

提出技术中立之抗辩。但值得关注的是,侵权方式在内容上中立还是在技术上中立存在本质差别,目前涉及此类纠纷的争论较大,暂未给出定论。另外,在大数据软件雷同的司法判定过程中,如何对实质性相似问题给出公平、合理的决断,也是大数据的司法实践中事案有别的另一难题,亟需著作权法调试、修缮后的科学保护。

(四) 制度选择的差异化

著作权与邻接权虽然都属于著作权保护制度的权利类型,但侧重点略有不同。前者倾向于保护作品产生过程中的智力创造,后者倾向于保护作品传播过程中所投入的劳动^[10]。因此,对于“有价值无独创性”的大数据而言,人们想到了邻接权保护模式正契合保护大数据价值“投入”的命门,并在1985年将计算机软件列为法国修订后著作权法中关于邻接权保护的客体,成为大数据软件作品开启邻接权保护模式的先河。

那么,究竟应不应当采用邻接权的模式保护“有价值无独创性”的大数据呢?就其内容而言,大数据内容以作品、作品片段、数值、事实等其他形式信息材料所构成,若是大数据作品或作品片段的传播行为从“作品传播权”扩大解释为“数据传播者权”,从而将后者纳入邻接权保护范围,但这一做法并未从根本上解决大数据权属保护的问题。退一步讲,倘若赋予那些产生重大投资又具“独创性”的数据双重保护,使它们在著作权与邻接权的“呵护”下拥有更多的选择权利,那么问题也接踵而至,双重保护下的大数据传播者此时会获得一项新的邻接权,这一权利与该数据本身所享有的著作权和邻接权的关系应当如何辨别?完全按照邻接权保护模式不难分析,其优点在于能够在原有的著作权法权利体制内为有限的大数据作品提供部分保护,而不必另行创设新的权利体系。但缺点也不容忽视,这种模式与新产生的邻接权会打破原有法权体系结构的稳定性,造成著作权与邻接权运用权利异象。因此,邻接权保护模式的选择并非明智之举,与其因为扩大解释后的“数据传播者权”牵强引论到邻接权保护范畴,不如微调著作权法施以更科学、合理的法律保护模式。

五、新兴权利的扩张与法律制度的衔接

本文尝试切入合同法、物权法、反不正当竞争法等其他法律保护模式,促使大数据在著作权保护下更好地融合其他法律制度,待时机成熟时不妨增设新兴权利予以配合,以求大数据及其产权配比最优化发展。

(一) 增设数据权

鉴于大数据通常不以公开形式示众,侵权行为人较少选择通过反向工程的手段窃取数据内容的特点,不妨增设采取特殊权利保护形式——新增数据权——为大数据的存储、管理及应用提供新兴权利。

通常,数据使用者或下游用户无法跳过著作权许可使用的权限对汇编作品形式呈现的大数据进行浏览、查阅或学习,而数据权作为根植于著作权法基础之上的新兴权利,其保护的客体范围可以更大,无论是单位数据、数据库、数据结构,抑或数据应用软件、数据库系统和数据资料库,均能成为数据权的客体保护对象。并且为了权衡数据权利人与使用者的利益关系与公众获取资源的公益性,数据权利人有权通过禁令救济的手段禁止非授权用户的侵权行为,前提是在数据权法定有效期限内行使该权力,且具有有限的排他性特征,即数据使用者合理使用、复制和传播该数据内容(不包括数据权利人通过合同等方式的特殊限制)^[11]。当然,立法者在设定数据权时应充分考虑数据权利人滥用权力的情形,规避损害数据使用者学习、欣赏、创作的自由,以及妨碍社会进步的特权发生。

在此还需说明,前文法经济学与法哲学表明,大数据权利人重视数据权属排他性权利的程度超过数据本身的收集、获取和使用,因而商业秘密的保护需求远不及著作权法保护。故而,数据权的增设目的并非鼓励新的或有价值的数据之公开,而是保护有公开数据内容的权利人所披露的数据能得到经济效益上的收益和激励。原因在于大数据获取方式的独特性决定了数据权所保护的数据内容和披露方法并非一致,正如数据权的公开要求类似于专利法中要求申请人在申请专利中披露其发明,这一要求明显低于著作权法中要求作者在有效期限内以有形客体形式呈现的要求。

(二) 网络授权合同与著作权限制之调和

利益导向式的立法在发挥重要的调和作用时,也存在预想与现实反差过大的失意情况。正如大数据著作权限制^⑫与网络授权合同所产生的矛盾是法之发现过程中脱离法律解释范畴的漏洞,尚需透过制度与实践的互动探知大数据运行下的权属嬗递与规则创制。以《欧盟计算机软件指令》为例,该指令第9(1)款针对计算机软件的复制、改编、备份等设置了限制权利,“任何违反第6条^⑬以及5(2)、(3)款中的例外的合同条款皆无效”。这些限制权利进一步明确了对反向工程侵权的认知偏差,为大数据的著作权限制提供了理据^[12]。同时,《欧盟数据库指令》有关使用者强制性权利的规则设置也可以为著作权法提供参考,即赋予数据使用者接触数据库内容的行为和对非实体部分数据提取、再利用的输出,均成为限制合同自由但又无法剔除的两项权利^⑭。诚然,著作权法与该类网络授权合同中的强制性权利可续叠适用,大数据权利人能够通过著作权授权合同将一系列限制和例外排除在外,所赋予数据使用者的自由亦会是最公平、最正义的权利。

然而,大数据著作权授权合同往往还需司法机关介入进行裁断,司法程序的调和避免新兴权利立法实践中的缺漏与空白,最具争议的当属在解决著作权授权合同方面有关著作权滥用禁止原则的弹性需求^[13]。禁止著作权滥用原则是类推禁止专利权滥用原则的派生产物,美国1948年的M. Witmark & Sons v. Jensen案建立该原则以来并未被广泛认可^[14]。但在我国,类似于大数据的著作权滥用案件与竞争法之间互补已不鲜见。就大数据权利人构成著作权滥用的行为不以违背反垄断法为前提的这一论断,只会成为大数据侵权诉讼中的抗辩事由,无法起到两法衔接融洽的实质性作用。准以此言,若想实现大数据在司法实践中的能动保护,大数据权利人著作权的行使需建立在网络授权合同“限制”与“例外”范畴之内,当超“度”用权发生时,必然引起禁止著作权滥用原则的审查,这便是网络授权合同与著作权滥用之间的耦合交涉,大数据无法脱离两者间权利效能的较量存在。

(三) 对添附制度之借鉴

添附制度^⑮始于罗马法,主要包括混合、附和和加工,这是大陆法系国家物权法中确认财产的一项重要规则,也是物权变动所运用的重要方法和制度^[15]。对“大数据”而言,数据的汇集、储存、分析、处理的过程更接近于添附制度中的加工^⑯。而关于加工物所有权的归属,《德国民法典》第950条规定的

^⑫著作权限制分为四类:第一类是以言论自由、隐私保护、通讯自由为基础的基本自由限制;第二类是基于公共利益及信息流通为前提的复制和翻译限制;第三类是著作权的工业竞争管制限制,如软件的反向工程等;第四类则是调整“市场失灵”的使用限制,如非商业性的私人复制等。

^⑬《欧盟计算机软件指令》第6条需要满足以下要件:第一,这些行为的实施人首先得有权使用该软件备份;第二,行为人为获交互性操作的必要信息不能从已有处获得;第三,这些为获交互操作的行为限于原软件的内容。

^⑭参见EU Software Directive Art. 9(1), Art. 5(2)(3), Art. 6, Art. 15.

^⑮添附制度是指不同所有人的物结合在一起而形成不可分离的物或具有新物性质的物,若想恢复原状在事实上不可能或者在经济上不合理,却仍需确认该新财产归属的制度或方法。

^⑯加工,是指一方使用他人财产加工改造为具有更高价值的新的财产。

立法例、《日本民法典》第246条以及《法国民法典》第570条至第572条^①均认为,动产经过加工形成新的物质属性后,其产权归属于使原物价值增益较多的一方。这也印证“大数据”虽源自用户所产生的基础信息,但是其聚合价值远远大于基础数据信息简单加合所带来的获益。因此,物权法的添附理论在大数据的著作权制度构建中亦具有参考价值。

将经过“加工”的数据分析结果所析出的产权赋予平台运营者(数据的“加工者”),将原始信息的数据产权赋予数据的产生者(互联网平台用户),这一做法不仅可以鼓励大数据的“深加工”行为使数据产权“物尽其用”,而且有利于保护涉及原始数据的用户数据财产、人身权益(包括隐私权及个人信息安全)等相关法益。与此同时,鉴于大数据的非穷竭(Inexhaustibility)的特征^②,平台运营者在进行分析、处理、应用类似于即时通讯系统、App、网站等网络平台用户发表的文学作品、评论等数据信息时,必须把添附的范围限制在著作权法允许的适当引用的范围内。

(四)反不正当竞争法之填补

《著作权法》对大数据的保护还可以配合《反不正当竞争法》双管齐下,对原始数据的使用、汇总和管理不仅可以适用《著作权法》(2020)第15条规定,还可通过《反不正当竞争法》第2条^③有关“诚实信用”基本原则和第9条第3款^④“商业秘密”相关规定施以保护,对于无法通过《著作权法》保护却以《反不正当竞争法》施以救助的部分“加工前”的大数据,其在聚合前的海量原始数据流转是否具有法律依据,还需要根据许可使用协议来规范大数据信息的自由流通与市场交易,其典型代表即为百度文库、期刊数据库的运营和使用,无论是利用反不正当竞争法保护抑或是商业秘密保护,其原始数据的法益皆可以通过许可使用协议得以实现。因此,大数据的法律保护体系是各部门法之间相互配合与制度衔接的多维度规范与防范的科学系统,绝非《著作权法》独大的一方尊崇,需结合大数据信息处理的流程和特点适用不同的法律规定为其保驾护航。

总之,在探求现实世界的真相与起源的历史长河中,法经济学与法哲学既可以作为理论指南,也可以作为裁量标杆,既可以作为科学的具象化依据,也可以作为治国理政的标尺,每个面向都为我们提供看待新兴事物的独特视角。当下在对大数据的法律保护未臻确意之际,著作权保护制度或许是值得参酌的。虽然,这一视角未必能够以最全面、最科学的路径关照大数据实践中的种种异象,但并不妨碍检验真知过程中不断修缮与完备的初衷,这也正是法学科学性在探求现实世界司法缺失补足过程赖以存续的基础。大数据权属划分及权利归属允许我们沉降到认知活动底层去了解著作权法与其他法律制度的相通之处,更加理性地处理现代科学技术带来的法律难题。

参考文献:

- [1] 中国人民大学数据工程与知识工程重点实验室. 中国信息资源开发利用指数报告(2018)[R]. 北京:中国人民大学电子版,2019:1487.

^①《法国民法典》第570条至第572条、《日本民法典》第246条均以加工物属于材料所有人为原则,而在加工所增加的价值远远超过材料的价值时才属于加工人为例外。而依《德国民法典》第950条规定,加工于他人动产者,以由加工人取得加工物所有权为原则,在加工的价值显然少于材料的价值时,由材料所有人取得加工物所有权为例外。

^②非穷竭(Inexhaustibility)是指:共享产品被生产出来之后,对它的消费无论扩大到多少用户,其品质和功用也不会有任何下降,用户之间也不会有任何相互干扰;或者换成经济学的术语来描述,共享产品提供的服务扩展到其他消费者的社会成本近乎于零。

^③《反不正当竞争法》第2条规定:“经营者在市场交易中,应当遵循自愿、平等、公平、诚实信用的原则,遵守公认的商业道德。”

^④《反不正当竞争法》第9条第3款规定:“本条所称的商业秘密,是指不为公众所知悉、能为权利人带来经济利益、具有实用性并经权利人采取保密措施的技术信息和经营信息。”

- [2] 连玉明, 大数据战略重点实验室. 重新定义大数据[M]. 北京: 机械工业出版社, 2017: 5.
- [3] 姚剑波, 杨朝琼. 大数据安全与隐私[M]. 成都: 电子科技大学出版社, 2017: 3.
- [4] 高建伟, 牛小凡. 科斯《社会成本问题》句读[M]. 北京: 经济科学出版社, 2019: 56.
- [5] 熊琦. 著作权激励机制的法律构造[M]. 北京: 中国人民大学出版社, 2011: 115.
- [6] 黑尔格. 法哲学原理[M]. 范扬, 张企泰, 译. 上海: 商务印书馆, 1961: 57.
- [7] 杨小兰. 网络著作权概论[M]. 北京: 知识产权出版社, 2012: 7.
- [8] 郑成思. 版权法[M]. 北京: 中国人民大学出版社, 1997: 189.
- [9] 王迁. 版权法对技术措施的保护与规制研究[M]. 北京: 中国人民大学出版社, 2018: 398.
- [10] 韦之. 著作权法原理[M]. 北京: 北京大学出版社, 1998: 12.
- [11] 朱真真. 大数据时代数据公开与知识产权保护的冲突与协调[J]. 北京: 中国科技论坛, 2019(3): 117-123.
- [12] 杨斌, 刘智鹏. 论网络授权合同与著作权限制的冲突与协调[J]. 湖北社会科学, 2012(5): 152-157.
- [13] De WERRA J. Moving beyond the conflict between freedom of contract and copyright policies: In search of a new global policy for information licensing transactions[J]. Columbia Journal of Law and the Arts, 2001(4): 187-193.
- [14] FRISCHMANN B, MOYLAN D. The evolving common law doctrine of copyright misuse: A unified theory and its application to software[J]. Berkeley Technology Law Journal, 2000, 15(3): 865-931.
- [15] 杨亮. “大数据”产权的法律分析[J]. 理论与当代, 2019(4): 29-31.

On the conflict and coordination of big data in the protection of copyright law

TIAN Xiaochu^a, GAO Shanxing^b

(*a. Law School; b. School of Management, Xi'an Jiaotong University, Xi'an 710049, P. R. China*)

Abstract: The use of modern network technology and scientific and technological methods has broken the traditional transaction model in the use and dissemination of big data, making piracy in the digital age a low-cost dependency that is not limited by time and space, or even sacrificing “creation” for the rapid development of the information society. Facing the lagging problem of the protection of big data resources through legal relief, people are thinking about conflicts and coordination of big data in the protection of copyright law. In order to achieve the purpose of protecting big data rights holders and maximizing social benefits, this paper takes big data and its tenure debate as the starting point, and uses the theoretical paradigm of law economics and legal philosophy to derive the ownership division and value gains of big data. The feasibility conclusion that big data is applicable to the protection of copyright law is drawn. Drawing on the experience and consensus of big data in the legal framework of the common law system and civil law system, deconstructing the benefits and hindrances of the application of copyright law to big data, the paper tries to establish a “data right” or a form compatible with other legal systems to examine the scientific rationality of jurisprudence when facing with modern science, and provide more possibilities for the legal protection of big data.

Key words: big data; copyright law; data property rights; data ownership; reproduction rights; fair use

(责任编辑 胡志平)