

# 民法总则数据保护路径： 概括式保护及与知识产权协调

曹建峰

腾讯研究院版权研究中心, 广东 深圳 518057

## 摘要

从设立数据保护条款必要性等方面论证提出,《中华人民共和国民法总则(草案二次审议稿)》第124条相对其他规定,更为合理。一方面,从体系科学的角度来看,符合知识产权特征的数据已被纳入知识产权体系,并无必要将其创设为新的知识产权客体;另一方面,随着大数据产业发展,保护数据极具产业基础与时代意义,但在制度设计上,应着重考察大数据产业基础、《中华人民共和国民法总则》条文设计与既有法律体系协调以及法律实施效果等多重因素。

## 关键词

数据保护;知识产权;协调

中图分类号:D92

文献标识码:A

doi: 10.11959/j.issn.2096-0271.2017011

## *Data protection approach in General Principles of Civil Law: generalized protection and coordination with intellectual property*

CAO Jianfeng

Copyright Research Center of Tencent Research Institute, Shenzhen 518057, China

## *Abstract*

From the necessity of establishing the data protection clause, a conclusion that the Article 124 of General Principles of Civil Law (Draft for Second Reading) is reasonable was obtained. On one hand, from systems perspective, since the data that accords with the characteristics of intellectual property has been covered by intellectual property law, therefore there is no need to create a new subject matter of intellectual property for data; on the other hand, with the development of big data industry, legal protection for data both is of significance and based on industry development. When designing legal framework for data protection, the legislator should take big data industry development, the coordination with existing legal systems, effects of law implementation and other relevant factors into account.

## *Key words*

data protection, intellectual property, coordination

## 1 引言

随着大数据、云计算、物联网、人工智能、虚拟现实等新兴技术和商业模式发展,数据保护被提上日程。对数据进行分类是配置不同数据保护制度的前提。一般意义上可以将数据分为3类:个人数据、开放数据及商业数据。个人数据需要个人数据保护制度,比如《中华人民共和国民事诉讼法总则(二次审议稿)》(以下简称二审稿)第109条规定的个人信息保护规定;开放数据需要政府数据开放制度,比如美国、英国等国家的数据开放法律制度。对于商业数据保护,《中华人民共和国民事诉讼法总则》(以下简称民法总则)在修法过程中展现了较大分歧。2016年6月审议的《中华人民共和国民事诉讼法总则(草案)》第108条将“数据信息”纳入知识产权进行保护;但2016年11月18日公布的二审稿第124条规定:“法律对数据、网络虚拟财产的保护有规定的,依照其规定。”在此背景下,本文从设立数据保护条款的必要性、与知识产权协调、未来具体制度安排等方面进行进一步论证,以期为我国数据保护制度完善提供有益的理论支持。

## 2 民法总则设立数据保护条款的现实意义

一方面,随着大数据、云计算、物联网等技术和产业发展,对数据的使用和分析成为一项核心能力,数据的重要性日益凸显,与之相关的窃取、盗用等违法行为也日益增多;另一方面,现有的知识产权法律制度基于独创性、创造性等理论,在大数据产业持续快速发展的形势下,对商业数据

已不能提供充分的法律保护。因此,有必要在民法总则中针对数据保护设立专门条款,这既具有现实的产业基础,又具有激励发展的实践意义。

### 2.1 保护劳动投入以激励大数据行业创新发展

(1) 大数据行业正在高速发展,对数据的使用、分析成为一项核心能力

人们正处于社会、经济以及技术的变革中。互联网已经发展到了大数据时代。以算法、机器学习、人工智能等为核心的大数据正在变革人们的世界。大数据几乎无法使用大多数的数据库管理系统处理,而必须使用“在数十、数百甚至数千台服务器上同时平行运行的软件”。由此产生了大数据行业,从硬件、软件到技术、算法,再到数据服务商,形成了一个完整的产业链。

《2015年中国大数据交易白皮书》<sup>①</sup>显示,2014年全球大数据市场规模约为285亿美元,预计2015年将达到421亿美元;在全球大数据市场中,行业解决方案、计算分析服务、存储服务、数据库服务和大数据应用是市场份额排名最靠前的细分市场,分别占据35.40%、17.30%、14.70%、12.50%和7.90%的市场份额。大数据行业的发展促进了大数据交易。作为国内首家大数据交易所,贵阳大数据交易所为包括社交大数据、电商大数据、交易大数据、消费大数据、金融大数据、医疗大数据、交通大数据等在内的30多种大数据提供交易。但是,《贵阳大数据交易所702公约》<sup>②</sup>显示,贵阳大数据交易所交易的并非底层数据,而是基于底层数据,通过数据清洗、分析、建模、可视化等获得的大数据成果。

搜索引擎、社交网络、新闻网站、音乐网站、视频网站、网络交易平台等领域的主流互联网公司在提供服务的过程中收

① <https://files.acrobat.com/a/preview/71d613c8-3d25-45ec-9c35-d63fa2882ebd>

② <https://files.acrobat.com/a/preview/737b39d6-a7db-4381-b9a6-fa4da5ac3776>

集了海量的数据,日益注重其中的商业价值,开始在数据商业利用方面进行探索,比如利用大数据挖掘、分析等技术对用户进行画像,以便更精准地向用户推荐内容、产品、服务或者投放广告。Google、Facebook、亚马逊等拥有海量用户的互联网公司正向大数据公司转变,对数据进行收集、融合、清洗、加工、挖掘、分析、可视化,并形成可以帮助商业决策的大数据成果,正成为一项核心能力。此外,在国内外,作为数据中介商(data broker)的大数据服务商近年来迅速涌现,从数据收集、存储、融合、加工再到挖掘、分析,为企业提供一整套的商务智能(business intelligence, BI)解决方案,帮助企业精准营销、提高产品或者服务的质量等。

(2)就商业数据而言,企业在大数据开发方面的劳动投入需要获得法律的保护

按照洛克的财产权劳动理论,当人们将他的劳动与处于共有状态的某个东西混合在一起的时候,他就取得了该东西的所有权<sup>①</sup>。因此,当企业投入大量劳动(无论其为机械劳动还是智力劳动)进行大数据开发,并形成数据集、数据库等大数据成果时,即使这些成果因为缺乏独创性或创造性而不能获得版权等知识产权保护,也可以因为这些成果中体现的劳动投入而获得某种程度的法律保护。比如,在国际新闻服务公司诉美联社案<sup>②</sup>中,美国最高法院就认为,信息、设计等无形物可以因劳动、金钱等投入而产生一种“准财产权(quasi-property right)”,从而可以基于反不正当竞争法禁止他人不当盗用。因此,从保护劳动和投资、法律激励等理论出发,给予大数据行业中形成的相关大数据成果以充分、合理的法律保护,对于促进大数据行业创新发展是很必要的,正如欧盟1996年出台《数据库指令》<sup>③</sup>是为了保护数据库制作者在数据库(无论其是否

具有独创性或创造性)制作上的时间、金钱、劳动等投入,从而促进数据库行业发展一样。

## 2.2 既有知识产权法律体系对数据的保护不充分

在大数据时代,数字数据(digital data,一般以二进制“0”和“1”形式存在)以史无前例的速度增长。谷歌搜索每天收到并处理35亿条搜索请求;Facebook每日上传图片3.5亿张。全球数据总量的80%是过去两年产生的。数据的形式和种类不断丰富,从文本数据到图片数据、音视频数据,新类型的数据不断出现。数据的体量和种类无疑增加了数据分析的难度和复杂性,也给数据保护提出了挑战。我国既有知识产权法律制度在数据保护上的局限性开始彰显。我国既有知识产权法律制度对数据的保护主要体现在3个层面:著作权法、商业秘密法和反不正当竞争法。

(1)著作权法不保护不具有独创性的事实数据

在著作权法上,我国对作品独创性或原创性的要求比世界很多国家都更严格,体育赛事类和游戏竞技类的视频节目是否是作品,在我国依然存在争议,互联网企业从用户或者通过多种物联网设备收集、聚合的事实数据(比如购物偏好、信用记录等数据)就更难被认定为具有独创性。此外,虽然数据汇编作品可能因其内容的选择或者编排而体现出独创性,从而可以作为版权作品(比如具有独创性的数据集或者数据库)予以保护,但是这一保护却不能延伸到其中的数据。正如菲斯特出版公司诉农村电话服务公司案<sup>④</sup>一样,即使原告在地址簿的收集、整理上投入了巨大的劳动成本,但只要地址簿内容的选择和编

<sup>③</sup> International News Service v. Associated Press, 248 U.S. 215 (1918)

<sup>④</sup> [http://europa.eu/rapid/press-release\\_IP-96-171\\_en.htm?locale=zh](http://europa.eu/rapid/press-release_IP-96-171_en.htm?locale=zh)

<sup>⑤</sup> Feist Publications, Inc., v. Rural Telephone Service Co., 499 U.S. 340 (1991)

排不具有独创性,就不能获得版权保护;版权保护的唯一基础在于独创性,而非劳动。

(2) 商业秘密保护具有局限性,与大数据产业诉求格格不入

在我国,企业的商业秘密受到反不正当竞争法保护。一般而言,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。互联网企业收集、加工、融合、挖掘的数据落入这个定义,并不存在问题。但是,能够作为商业秘密保护的数据要求具有机密性并被采取保密措施,这在某种程度上与大数据时代所要求的数据自由流动、公开交易、分享和利用等是不相容的,难以适应大数据行业的发展需求。不仅如此,针对海量的数据采取保密措施,对企业而言是很大的负担。然而,即使企业就相关数据享有商业秘密并采取了严格的保密措施,但任何人都可以通过反向工程合法破解其商业秘密;这种做法虽然客观上鼓励了竞争,但从经济和社会的角度看,无异于重复开发和浪费。此外,商业秘密权无法保护具有商业价值但不符合商业秘密构成要件的那些数据,即使投入了很大的劳动收集、整理、获取那些数据。

(3) 反不正当竞争法在保护的范围和标准方面具有很大的不确定性

虽然在理论上,从他人数据集或者数据库中提取不受版权保护的数据,制作竞争性的数据库或者用于其他商业目的,可以构成不正当竞争<sup>[2]</sup>。然而,反不正当竞争法虽然对竞争者的不正当竞争行为具有规制作用,但是相对于著作权法而言,它在保护的范围和保护的的标准方面又具有很大的不确定性,因此其缺陷也表现得非常明显。一方面,虽然被世界知识产权组织纳入知识产权的范围,但是反不正当竞争权

的具体内容、权利的期限等都不明确,权利人难以有效利用。比如,当发生数据许可或者转让时,许可使用或者转让的到底是什么,是不明确的。另一方面,反不正当竞争法不能给数据集或者数据库创设一种稳定的财产权,从而导致反不正当竞争法理论缺乏一个稳定的分析基础,因此难以理解和适用<sup>[3]</sup>。美国在通过反不正当竞争法保护不受版权法保护的数据方面,具有较为成熟的司法实践,即便如此,通过“盗用理论”保护他人通过投入劳动获取的数据时,也需要满足较为苛刻的条件。此外,反不正当竞争法无法规制竞争者以外的一般行为人的恶意破坏、损害等行为,也缺乏平衡私人利益和公共利益的机制。

### 3 “数据信息”不宜概括纳入知识产权体系

《中华人民共和国民法总则(草案)》把“数据信息”纳入了知识产权客体的范围,在产业界和学术界引发了广泛争议。对此,笔者认为,“数据信息”不足以成为新的知识产权客体。

(1) “数据信息”缺乏法律概念应具有明确性,是一个不恰当的概念

一方面,“数据信息”是一个高度概括、高度抽象的一般概念,包容万象,难以确定具体所指,作品、商标、商业秘密、专利等几乎所有的知识产权客体都可以落入“数据信息”这一范畴。有学者甚至认为知识产权的客体就是信息。因此,一旦将“数据信息”纳入知识产权客体,势必造成知识产权法律体系的混乱,可能带来更多争议。即使在对数据进行保护或者曾经寻求对数据进行保护的美国,都从来没有使用一个如此抽象、高度概括的概念,而是采取“数据

库(database)”“数据集(collection of data)”等较为明确的概念。

另一方面,如果“数据信息”意指“数字化信息”,这是不是表示数字化作品、数字化商标等都可以纳入“数据信息”进行保护?如果是这样,既有的知识产权法律体系势必受到重大冲击,这更加说明“数据信息”难以与作品、商标等并列为知识产权客体。此外,“数据信息”包括个人数据、开放数据以及商业数据,将“数据信息”纳入知识产权客体势必导致对个人数据、开放数据等主张知识产权。

(2)将不具有独创性的“数据信息”纳入知识产权客体会拉低知识产权保护门槛

事实、观点或者知识、思想、交流的其他基础材料不当被垄断,是知识产权法的一个基本原理。于是,在著作权法上,有了思想—表达二分法以及延伸出来的融合理论(即思想与表达融合为一);在专利法上,发现、科学理论等不能被授权专利;在商标法上,商品或者服务的通用名称等不能被授权商标权,诸如此类。贸然将“数据信息”纳入知识产权体系进行保护,既可能导致公有领域的信息被私人占有,又会拉低知识产权保护标准,引发制度性冲击。

事实上,所谓的“数据信息”不具有独创性,否则通过著作权法予以保护是不成问题的。具有独创性的可能是数据加工、分析的过程(算法和分析工具的独创性),并且考虑到这个过程付出的劳动和金钱投入,所以有必要在一定程度上保护其投资,而这正是欧盟数据库立法的初衷。然而,欧盟《数据库指令》出台后遭到多方批评,欧盟数据库资料工作小组甚至建议终止《数据库指令》而采用美国不正当竞争法上的盗用理论来保护数据库。其主要的理由是认为《数据库指令》采取的授予宽泛的专有财产权结合少量例外的方式,在很

多方面大大超出了必要和适当的程度,过度保护了数据库制作者的利益,损害了社会公共利益,不应成为国际数据库保护的模板。

国际社会也曾力推通过特别法保护数据,但均未成功。比如,1996年欧盟、美国向世界知识产权组织(World Intellectual Property Organization, WIPO)提议了一个数据库保护条约,但最终遭到国际社会的激烈反对, WIPO最终表态说这一问题需要进一步研究和论证。在美国,《信息集反盗版法》保护“通过大量金钱或者其他资源投入而收集、整理或者维护的信息集”,这一保护几乎不设门槛,不要求数据库的创新性,少有例外规定,没有存续期限的限制,相当于赋予数据库制作者数据垄断权,可以阻止任何人提取、使用或者重新使用数据库中的实质性部分。但这样的立法最终被认为是为企业创设了一个合法的垄断权,而缺乏利益平衡机制,最终在强烈的反对声中被否决了。目前,除了俄罗斯和欧盟,世界上几乎没有其他国家和地区通过特殊权利保护数据或者数据库。

因此,如果贸然将不构成作品的数据库、数据集或者数据库纳入知识产权的保护范围,不但会导致私人对原本不该受到保护的“数据信息”主张知识产权,而且将极大拉低知识产权的保护门槛和标准,可能导致后续其他同样不具有创造性的新事物也要求获得知识产权保护,给知识产权制度带来根本性冲击。

(3)“数据信息”知识产权保护的价值追求难以明确

实际上,大数据时代的核心不是让私人把数据控制起来,专为其所用,而在于数据的自由流动、利用和共享,这无论是在互联网行业、科学研究方面抑或医疗行业都是成立的。恰恰是知识产权制度难以在保护私人在数据上的投资和公众的数据利

益之间实现平衡,传统知识产权制度下的合理使用、强制许可制度等是不充分的。因此,通过知识产权保护“数据信息”所欲达成的目的有悖于大数据时代数据共享的理念。

## 4 数据法律保护的制度设计

通过知识产权制度保护不具有独创性的数据是不可取的,因为这有悖于知识产权制度的基本原理。未来是否需要对数据库、数据集等大数据成果设定特殊权利,从而对不具有独创性的数据以特殊权利进行保护,需要经过严格论证,这一方面美国和欧盟都是前车之鉴。

### 4.1 民法总则宜对数据保护作概括式、宣示性规定

二审稿在第120条不再将“数据信息”纳入知识产权客体,并在第124条规定“法律对数据、网络虚拟财产的保护有规定的,依照其规定”,这一概括式、宣示性规定是较为合理的。未来在经过严格论证后,可以考虑对数据提供一定形式的法律保护。这样的制度安排至少有以下好处。

第一,在涉及数据的不正当竞争案件中,可以为反不正当竞争法的适用提供一个统一的法律基础。在美国,通过“盗用理论”保护企业搜集的数据实际上存在很大困难,所以一般只限于“热点新闻”类的盗用案件,就是因为法律对不具有独创性的数据的保护是不明确的。一旦法律规定投入时间、金钱、劳动或者其他资源获取的数据受法律保护,就可以在此类不正当竞争案件中,为反不正当竞争法的适用提供一个统一的法律基础。

第二,回应大数据等产业发展需求,

可为产业创新发展提供法律上的激励。在大数据成果无法律保障的情况下,企业势必不会投入更多的金钱和其他资源进行创新,因此有必要通过原则性规定给大数据行业的参与者吃一颗“定心丸”。

第三,为后续立法提供基本法依据,保证后续立法的弹性空间。当前大数据相关产业才开始起步,业务和产业实践不断变化,针对数据的进一步立法还需要结合产业实践进行充分论证,不宜贸然进行过度立法,否则可能造成适得其反的法律效果。

### 4.2 未来进一步立法的考量因素

第一,在具体规则设计方面,应当具有适当的门槛,避免以极小的投入对数据主张权利。比如,即使在对数据库进行特殊法保护的欧盟,也不是所有的数据集都可以成为数据库,而需要满足特定的条件。

第二,引入利益平衡机制,一方面需要平衡个人信息的保护,可以通过匿名化、假名化、加密等技术手段以及个人信息保护法等法律手段实现,比如,二审稿第109条规定“自然人的个人信息受法律保护。任何组织和个人不得非法收集、利用、加工、传输个人信息,不得非法提供、公开或者出售个人信息”,这就是较为合理的;另一方面,需要平衡数据产权与数据自由流动、共享和使用。

第三,强化责任规则和反不正当竞争规则。在强调信息自由流动的大数据时代,应加强从责任规则、反不正当竞争规则等方面保护数据,毕竟企业的目的是从数据的利用中获得经济利益,而非一律阻止他人使用其数据。因此,通过强化适用责任规则(比如对他人数据的何种使用需要承担法律责任)和反不正当竞争规则(比如对他人数据的何种使用构成不正当竞争行为),既可以保护企业在大数据上的开发

投入及其成果,又可以较为合理地平衡各方利益,避免造成对数据的绝对垄断。

## 5 结束语

民法总则在修法过程中,将数据保护纳入其中,就是在回应大数据产业发展所带来的数据保护诉求,这值得肯定。但是在具体保护路径的选择上,笔者认为,知识产权体系有其自身的理论基础和法理依据,生硬地将数据保护塞入既有知识产权法律体系,有违体系科学之原则。当前而言,数据无论是对大数据行业,还是对云计算、物联网、人工智能等新型业态,都是极为重要的“资产”,是商业竞争的核心,因此在民法总则中给予概括式、原则性保护,是很有必要的。未来随着对产业实践了解的加深以及学界与业界交流的加深,在严格

论证的基础上,可以考虑进一步立法。

## 参考文献:

- [1] 李杨. 再评洛克财产权劳动理论——兼与易继明博士商榷[J]. 现代法学, 2004(1): 171-172.  
LI Y. Re-comment of Locke's theory of property-labor: discussing with doctor Yi Jiming [J]. Modern Law Science, 2004(1): 171-172.
- [2] 李杨. 数据库的反不正当竞争法保护及其评析[J]. 法律适用, 2005(2): 56-60.  
LI Y. Anti-unfair competition law protection for database and its comments [J]. Journal of Law Application, 2005(2): 56-60.
- [3] 郑胜利, 崔国斌. 数据库保护的立法现状与基础理论[M]. 北京: 法律出版社, 2002: 106.  
ZHENG S L, CUI G B. Legislative status quo and fundamental theory concerning database protection [M]. Beijing: Law Press · China, 2002: 106.

### 作者简介



**曹建峰** (1989-), 男, 腾讯研究院版权研究中心研究员, 主要从事网络安全、数据保护、网络版权与内容产业等互联网法律政策研究工作。曾参与编写腾讯研究院法律研究中心《互联网+时代的立法与公共政策》和《网络空间法治化的全球视野与中国实践》两本著作。代表论文包括《被遗忘权在国内外的的发展及对其本土化可能性的反思》《从欧盟“被遗忘权”看网络治理规则的选择》《欧盟网络中立规则评述——一个更开放的欧盟互联网》《人工智能: 机器歧视及应对》等。

收稿日期: 2016-11-25