

衍生数据知识产权保护的论证成与权利构造

王婉仪

南京理工大学知识产权学院, 江苏 南京

收稿日期: 2024年3月14日; 录用日期: 2024年5月6日; 发布日期: 2024年5月15日

摘要

衍生数据作为数据要素中的重要组成部分, 蕴藏着巨大的经济价值。衍生数据符合劳动赋权理论与创新激励理论, 赋予其知识产权保护具备法理基础; 从知识产权客体扩张性逻辑以及原型范畴理论来看, 衍生数据可因“家族相似性”而纳入知识产权客体体系内, 但考虑到衍生数据自身的独特属性, 既有的知识产权制度难以对其进行合理的规制, 因而需要为衍生数据在知识产权体系内创设新型知识产权即衍生数据专有权。具体而言, 该权利的构造需以满足多方利益平衡和促进数据要素市场流通原则, 以非独创性、公开性以及来源合法的衍生数据为客体, 基本权能包括发布权、使用权、传播权、禁止权。为避免衍生数据权人的权利过分扩张, 权利的行使受到保护期限、合理使用制度以及强制许可制度的限制。

关键词

衍生数据, 知识产权客体, 数据产权

Theoretical Proof and Rights Construction of Intellectual Property Protection for Derived Data

Wanyi Wang

School of Intellectual Property, Nanjing University of Science and Technology, Nanjing Jiangsu

Received: Mar. 14th, 2024; accepted: May 6th, 2024; published: May 15th, 2024

Abstract

Derived data, as an important component of data elements, contains enormous economic value. Derived data conforms to the theories of labor empowerment and innovation incentives, giving it a legal basis for intellectual property protection; from the perspective of the logic of expansion of

intellectual property objects and the theory of prototype categories, derived data can be included in the system of intellectual property objects due to “family similarity”. However, considering the unique attributes of derived data itself, existing intellectual property systems are difficult to regulate it reasonably. Therefore, it is necessary to create new types of intellectual property rights, namely exclusive rights to derived data, within the intellectual property system. Specifically, the construction of this right needs to meet the principles of balancing multiple interests and promoting the market circulation of data elements, with unoriginal, public, and legally sourced derivative data as the object. The basic rights include the right to publish, use, disseminate, and prohibit. To avoid excessive expansion of the rights of derivative data rights holders, the exercise of rights is limited by protection periods, fair use systems, and mandatory licensing systems.

Keywords

Derived Data, Intellectual Property Objects, Data Property Rights

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

数字经济时代，数据要素被比作石油逐渐成为国家基础性资源和生产要素，其所蕴含的价值对社会的经济发展有巨大驱动作用。中共中央国务院关于构建数据基础制度更好发挥数据要素作用的意见》(以下称为数据二十条)提出“要建立数据产权制度，建立公共数据、企业数据、个人数据的分类分级授权制度”，要求对数据进行保护，并对数据的权益客体进行类型化的划分[1]。近年来，学界对于数据权益问题展开了一系列讨论，但对数据的划分、数据权利的客体、权利构造等方面产生了分歧。区分数据的不同类型是认识数据是否可以纳入知识产权保护的重要前提，不能将数据视为一个整体进行保护。因此需要通过合理的数据分类标准，将衍生数据从数据整体中抽离出来，以衍生数据为保护对象，明晰衍生数据的界定范围，使其与其他的数据区分开来。

对于衍生数据的保护首先需要回答问题：一、衍生数据的法律属性。二、衍生数据的保护模式。三、衍生数据权利的内容与限制。因此，基于上述问题，本文从数据资源类型化分析入手对衍生数据界定，厘清其与其他数据的关系。接着对衍生数据进行知识产权保护的正当性进行证成，从衍生数据作为知识产权客体的正当性和法理基础两个角度着手。虽然衍生数据具备纳入知识产权体系内的条件，但因其所拥有的独特属性无法被既有的知识产权法所完美保护，因此笔者提出衍生数据属于新型知识产权客体，应创设“衍生数据专有权”。最后，本文通过对衍生数据的权能进行了规范化构造，完善权力的保护与行使边界。

2. 数据类型化中衍生数据的界分

2.1. 数据资源类型化分析

传统的将信息内容作为数据分类标准具有局限性，过分关注内容而忽略数据的形式必然会使不同类型的信息内容重叠。若将数据权利赋予任何一个主体都是不当的，被授予数据权利的主体会享有一定的垄断权力，其余主体对于数据资源利用的意愿会降低，容易形成“数据孤岛”进而阻碍数据的流通利用率。简而言之，以信息内容进行分类的传统标准不能满足对不同阶段的数据进行全面的保护，并且容易

引起同一数据处于上下游的数据主体对其权属进行无意义的争夺[2]。因此，只有充分认识数据的利用过程才能对数据进行合理的分类。

数据资源的利用通常会经历三个阶段，分别是数据收集、数据整合以及数据分析，以此为基础厘清不同数据之间的转化、利用关系，构建合理的数据类型化标准。数据资源转化首先是原始数据的生产，接着通过对数据进行汇集形成集合数据，最后对数据进行分析行的加工形成衍生数据。笔者认为应当根据数据在利用过程中的处理程度对数据进行分类。因此可以将数据资源分为原始数据、集合数据以及衍生数据。原始数据是指通过合法的方式直接获得的未经过处理的数据，不具有直接应用的价值功效。集合数据是指将原始数据进行收集、分类、整理等加工方式处理而形成的数据。目前，若将数据整体纳入知识产权客体的范畴之中缺乏正当性。首先对于原始数据而言，原始数据的来源广泛，对其界权难以保持法律上的确定性。对于原始数据应当强调对其进行利用而不是强调归属[3]。其次对于集合数据可以按照著作权法中的汇编作品进行保护。因此，下文以衍生数据为保护对象对其进行分析论证。

2.2. 衍生数据的界定

第一，经过实质性加工。衍生数据相较于原始数据和集合数据的加工方式而言，具有“精加工”的特点，但这种加工具体而言需要达到什么程度才能使数据成为衍生数据。有学者认为衍生数据是经过智力劳动的深度加工所形成的数据[4]。但也有观点认为衍生数据并不以实质性的加工为要件，只要通过合法获得具有市场价值即可[5]。笔者认为，衍生数据应当具备实质性加工要件，由此区别于集合数据。首先若只要通过简单的加工即可形成衍生数据，那么成为衍生数据的门槛就会过低，导致劣币驱逐良币，背离促进数据要素流通的目的。

第二，经过脱敏处理。去标识化、匿名化都属于数据脱敏技术。衍生数据需要在市场上流通，为了保护个人信息和敏感信息使各方利益达到平衡的状态，需要对数据进行脱敏处理，在对数据处理的过程中要对原始数据中的敏感信息进行过滤和清洗。虽然无法保证是否能够达到“无法识别且无法复原”的程度，但是数据脱敏在数据处理过程中仍是非常重要的过程。

第三，具备财产属性。稀缺性、可支配性以及价值性等特征是衍生数据具备财产属性的本质特征。首先衍生数据具有稀缺性，衍生数据是以原始数据和集合数据为基础，这两种数据通常由数据收集者享有，具备实际的稀缺性，而衍生数据是对其投入了实质性的劳动因此而产生了现实的价值。其次衍生数据具有可支配性。财产的可支配性是指客体在人力的适使用下能够被支配。对于衍生数据而言，数据处理者可以对其进行访问、修改、删除等操作，因此具备可支配性。最后衍生数据具有较强的经济价值，在立法层面《民法典》将“数据”与“网络虚拟财产”并列规定，后者的经济价值已经得到了广泛认同，立法上的规定在一定程度上保留了认可衍生数据经济价值的空间[6]。衍生数据所具备的财产属性不言而喻。综上所述，对本文研究的保护对象进行进一步的限定，应当是“原始数据和集合数据经过实质性的加工和脱敏化处理所形成的具备财产属性的衍生数据”。

3. 衍生数据赋予知识产权保护的正当性论证

3.1. 衍生数据赋权的法理基础

3.1.1. 劳动赋权理论

从洛克的劳动赋权论看来劳动是财产获得正当性的基础。虽然该学说最初描述的是有形财产例如麦子、果实之类[7]，但对于无形的智力成果也同样适用，人们通过脑力活动将现有的资源转化为智力成果，该智力成果也应当成为财产权的客体。衍生数据在处理的过程中通常要付出大量的时间和劳动成本，需要对海量的原始数据或者集合数据进行采集、清洗以及分析，这每一个过程都需要付出智力劳动。虽然

衍生数据最后是由计算机自动处理和生成的，但是算法的选择却是通过数据处理者进行智力投入的，并且计算机仅仅只是辅助性的工具，想要生成有价值的衍生数据仍需要人员与资本来维护其数据的运行。因此数据主体通过智力活动形成的衍生数据应当要获得财产权的保护。

3.1.1. 激励创新理论

对衍生数据赋予知识产权保护可以以激励创新理论来论证。知识产权制度的核心在于通过激励机制促进人们对技术创新的不断投入，促进社会进步和技术发展，保护知识产权就是保护创新，这也是知识产权制度的经济价值所在。形成衍生数据的任何一个环节都需要数据处理者不断投入资金和技术，若否定数据处理者对衍生数据产品的排他权利，使他人可以利用爬虫等低成本的手段获取他人的数据成果，将会使数据处理者失去处理信息的动力这将直接导致低质量的数据在市场上流通[8]，不利于技术的进步。同时也可能会引发数据垄断的现象，因为数据处理者会采取更加保密的措施来对其数据进行保护，造成事实上的垄断，使数据处于封锁状态。亚当斯密就认为，最直接的激励因素来源于利益，是利益的激励引发了所有经济活动参与者勤勤恳恳、兢兢业业的经济行为。激励机制希望人们通过趋利避害的本性，通过利益诱导促使人们进行某种活动。因此，要想实现数据的社会化利用离不开给予数据处理者一定的激励措施[9]，对衍生数据赋予知识产权制度的保护，目的在激励更多的数据生产和更新[10]，促进数据产业的繁荣。

3.2. 衍生数据作为新型知识产权客体的正当性

3.2.1. 符合知识产权客体的扩张性逻辑

知识产权的客体是人们在知识形态领域中创造的劳动成果，本质是信息，但并不意味着所有的信息都能成为知识产权的特定客体[11]。对于知识产权客体的形态存在“智力成果说”、“形式说”、“符号说”等，但无论哪种术语在外延的意义上都只是一种类型化的概念，无需纠结其与实体法上权利的设定是否完全匹配。相反，法律的制定都有一定的滞后性，未百分百对应反而给予了知识产权客体一定的开放性和弹性。因此，笔者不欲对上述学说展开具体的讨论，而是聚焦于知识产权客体的扩张性逻辑分析。

现代知识产权作为区别于有形财产权的新兴财产权利正处于一种“攻城略地”的无序扩张状态。知识产权法的客体最初仅以著作权法、专利法、商标法为主干，但现代知识产权客体却已呈现出外部扩张和内部扩容的趋势。一方面，新的客体不断被纳入知识产权的法律体系之中，如植物新品种、商业秘密等。另一方面，知识产权法内部也进行了扩张，传统知识产权法中的专利法、著作权法的客体类型也进行了扩容，例如计算机软件、基因等。在现代的知识产权制度中，客体的扩张是其制度变革最显著的外在体现。随着新的客体形态的出现，知识产权客体的范畴化逐渐失去了包容性，但对其不能盲目的扩张，应当结合实践或者利益来推动扩张而不能仅仅依靠理论去推动。回顾历史，可以发现知识产权的客体是随着市场经济的发展和科学技术的进步在谨慎的扩张[12]。知识产权制度的实践已经证明了知识产权客体的开放性，这也为衍生数据成为知识产权的新型客体之一提供了可能性。综上所述，从知识产权客体的扩张性的角度出发，笔者认为知识产权的客体范畴并不是封闭式的，它是一个开放、包容的体系，现代知识产权客体的扩张是具有必然性的。

3.2.2. 衍生数据与知识产权客体的相似性

回顾知识产权客体当前的形态可以发现虽然很难通过合适的知识产权语境将它们串联起来，但它们的特征之间却存在着交叉关联性。例如商业秘密和专利共同为具有经济价值的技术信息提供保护；发明和作品都是创造性的智力成果；商标和作品都属于“符号表达”等。因此知识产权客体之间虽然不能以一个共同的特征进行概括，但它们可以因为相似的特征而成为一个体系。从知识产权客体演变的脉络历

史可以看出,新的客体与某个原有的客体因具备一定的相似性而逐渐成为了知识产权客体中的新成员,即家族相似性。原型范畴理论认为随着其他新成员的不断加入,这个新成员会逐渐具备所有成员中的共同特征,从而变成了中心化的原型。典型的商标的加入,另外集成电路布图设计、植物新品种以及著作权中客体的扩张均是适应现实的需求,通过与既有的知识产权客体之间的相似性而逐步纳入知识产权体系之中。

根据上述分析,原型范畴理论为衍生数据纳入知识产权体系中提供了借鉴意义。衍生数据作为数字经济时代的产物,其与知识产权客体之间具有高度的相似性。第一,兼具知识产权法意义上的稀缺性。当前知识产权制度对于财产的保护是通过授予权利人一种有限的排他性权利制造稀缺来保护客体利益[13]。而如上文所提到的,衍生数据也恰好具备这种稀缺性,数据在流通过程中与其他知识产品相似,均会因被复制而打破其稀缺性。第二,非物质性。衍生数据与知识产权客体都是无体无形物,都不具有物理的外观,不能被人们现实占有。第三,非冲突性。衍生数据和其他知识产品一样可以在一定程度上满足不同地点不同主体的使用,并且不会因此而发生占有上的冲突。衍生数据的重复使用也不会因此而减损其价值。因此笔者认为根据原型范畴理论,衍生数据可以因其与知识产品的相似性而逐渐被纳入知识产权体系之中。

3.2.3. 既有知识产权制度保护适用困难

既有的知识产权制度保护衍生数据主要存在两种途径,一是知识产权专门法的保护路径,二是反不正当竞争法的“一般条款”,但这两种保护途径均有其局限性。

对于著作权法保护路径而言,著作权法只能通过汇编作品的方式对衍生数据进行保护,对其内容的选择需要具备“独创性”,作品中的独创性是作者进行选择、取舍、编排的结果。因此著作权法只能对具有独创性的衍生数据进行保护,而不适用非独创性的衍生数据;对于专利法保护路径而言,我国专利法明确规定,智力活动的规则和方法不授予专利权。而衍生数据是由事先选定的算法自动加工而成的,仅仅涉及特定的算法或者计算机软件,其属性与智力活动的规则和方法类似,往往不具备可专利性。此外由于衍生数据的复杂的形态以及共享的特征,专利法保护也明显不当。其次,在目前的司法实践中,对于衍生数据的纠纷通常采用反不正当竞争法的一般条款来解决。但是其一般条款具有不确定性,对于行为的构成要件、客体保护范围等方面缺少明确的规定,存在滥用的风险[14],因此不适合将此作为长期保护衍生数据的方法。

结合前文可以发现,衍生数据可以作为知识产权的客体之一,但却不能被现有的知识产权客体所容纳,因此可以将衍生数据视为数字经济时代的新型的知识产品来保护,使其成为知识产权的新型客体。

4. 衍生数据的权能构建

4.1. 确立衍生数据知识产权保护的价值选择

4.1.1. 追求多方利益原则

法律的目的在于个人原则与社会原则之间形成平衡。所以在衍生数据权利的构建中遵循利益平衡原则尤为重要,在权利保护的过程中要平衡多方主体的利益诉求与利益关系。例如,数据处理者在采集到的原始数据进行数据分析时,要遵循合理目的、最小必要和合理预期原则,对数据中所包含的个人信息进行清洗、匿名化处理。此外,在社会公共利益方面,利益平衡原则要求在授予衍生数据一定程度的排他权进行保护的同时,也要保障其他数据企业能够合法的访问和利用数据,促进数据的流动和开发,实现知识产权法律的效益目标即是利益平衡并达到最大化。知识产权法所强调的利益平衡精神,主要是通过权力配置与权利限制的规则来实现的,因此对于衍生数据权利的构建应当要实现权利保护与限制的平衡。

4.1.2. 促进数据要素市场流通原则

衍生数据权利的构造必须要考虑促进数据要素在市场上的自由流动和开放共享。在现实层面,限制数据流通不仅会影响数据价值增值同时也需要投入更高的管理成本,一些数据主体由于无法获取有价值的的数据而不能进行创新或者是不得不支付高昂的使用费,这对于数据要素的流通是非常不利的。数据的价值在流通过程才能得以实现,数据的共享可以促进资源的合理配置以及各数据主体之间的交流合作。在政策层面,国内外对于数据的制度规范基本都是以促进数据流通共享为导向。例如,欧盟《通用数据保护条例》中指出“限制或者限制个人数据在欧盟内部流通”,该规范明确表明了一数据流通为基本的制度目标。在我国促进数据流通也是各地数据规范的主要内容,《上海市数据条例》明确提出“促进数据依法有序自由流动”。而江苏、浙江等地也纷纷出台数字经济促进条例。所以在对衍生数据权利构造的过程中同样也需要以促进数据要素市场为原则。

4.2. 衍生数据权利的规范化构造

在知识产权体系内,衍生数据权利的规范化构造主要从两方面出发:一是衍生数据专有权的创设,从权利的主体、客体以及运行机制等方面配置与衍生数据自身特有的属性相匹配的保护模式;二是根据上文确定的衍生数据知识产权保护的价值选择来对衍生数据权利进行一定的限制,促进衍生数据的市场流通性和各方主体之间的利益平衡。

4.2.1. 衍生数据专有权创设

为了实现衍生数据专有权保护制度的有序运行,在权利创设的过程中需要对衍生数据专有权的主体、客体以及内容等方面进行进一步的明确,以期符合知识产权价值观念,并形成与著作权、商标权和专利权等既有知识产权制度平行的权利构造。

首先,在权利主体的范畴上,由于衍生数据生产过程中会涉及多方主体,如果设置共同所有权或者多个主体对衍生数据都享有独立的权利,那么在进行制度构建时会大大增加立法的难度。因此笔者认为,衍生数据专有权为单一主体更为合理。衍生数据的生产过程中需要以海量的原始数据为基础并辅之以大量的人力和技术,因此权利应该归属于数据处理者。在职务过程中所形成的衍生数据可以参照《专利法》中关于职务发明的认定,在没有特殊约定的情况下,衍生数据专有权应该归属于衍生数据的生产企业。

其次,在权利客体范畴,衍生数据作为新型知识产权客体,需满足非独创性、公开性、合法性。第一,非独创性。具有独创性的衍生数据可以通过现行的著作权法进行保护,为了避免与现行衍生数据知识产权保护机制之间发生冲突,应将衍生数据专有权的客体限定为不具备独创性的衍生数据。第二,具备可公开性。所谓公开性,是指将权利保护的客体公之于众,另社会公众知悉的一种状态。数据的价值在于它的动态性和可利用性,与商业秘密所保护的为公开的数据不同,公开的衍生数据并不会影响它的价值,反而会因为不断的流通利用而产生增值。可公开的衍生数据具有“公用性”的特征,但是并不能其所具备的公共属性而让他人随意“搭便车”。因此,为了进行利益平衡,公开的衍生数据需要受到保护,以促进数据资源开发利用的立法宗旨。第三,合法性。数据的收集、处理过程必须要合法,如果其来源一开始就存在不法行为,那么该数据也不存在保护的必要性。

最后,在权利内容的配置上,立足于利益平衡和促进数据要素市场流通原则,笔者认为衍生数据专有权应当是不同于有体物所有权的“有限排他性”权利。因此,可将数据权能分为发布权、使用权、传播权、禁止权。发布权类似于著作权法中的发行权,指企业首次向社会公开其衍生数据的权利,一经使用该权利即穷竭。传播权可以参照著作权法中对于信息网络传播权的规定,禁止未经许可向公众传播其数据。禁止权是为了防止第三方采用不法手段获取其数据的行为。

4.2.2. 衍生数据权利限制

根据衍生数据知识产权保护的价值选择, 为了避免权利的过分扩张, 因对其进行一定的限制, 笔者认为可以通过以下方法进行限定: 一是设定保护期限。由于衍生数据具有很强的时效性, 数据是动态变化的。因此需要设定合理的期限, 使数据在失去时效性后能进入公共领域, 变成一项公共资源, 促进数据的流通; 二是建立数据合理使用制度, 这是目前著作权法普遍采用的制度, 借鉴知识产权中的合理使用制度具有合理性。笔者认为在遵守公开原则、适当原则以及比例原则的前提下, 合理使用的具体内容可以包括个人基于非商业目的使用衍生数据、出于教学或科学研究的目的获取数据、为了社会利益和国家利益获得数据。三是建立特殊领域的数据强制许可制度。以必要设施原则为基础, 如果某个数据主体在该领域内处于绝对的支配地位, 容易造成垄断和限制竞争的结果。必须要建立公平、合理、非歧视性的数据访问规则, 这可以参照必要标准专利的 FRAND 原则。在保护衍生数据专有权的同时, 也应到要限制权力的滥用。

5. 结语

信息技术的发展对数据财产的保护提出了新的要求, 产生了一种全新的法律关系即衍生数据专有权。本文以知识产权制度为切入点, 法哲学的劳动赋权理论是衍生数据赋权的底层逻辑, 创新激励理论是衍生数据与知识产权制度契合的法理基础。通过回顾知识产权客体的扩张性历史, 可以看出知识产权客体范畴是一个开放、包容的体系, 这为衍生数据纳入知识产权客体体系中提供了可能。而原型范畴理论又以“家族相似性”为原理分析了衍生数据与知识产权客体的相似性, 应当将衍生数据作为知识产权的客体之一。但考虑到衍生数据的独特属性, 发现既有的知识产权制度无法与其完美适配, 笔者提出应当将其作为新型知识产权客体创设衍生数据专有权。紧接着后文对衍生数据专有权进行了详细的构建, 以追求多方利益平衡原则和促进数据要素市场流通原则为价值选择, 对衍生数据专有权进行保护与限制。具体而言数据处理者为权利主体, 以非独创性、可公开以及合法获取的衍生数据为客体, 以发布权、使用权、传播权、禁止权为权利内容对衍生数据专有权进行保护。为防止形成数据垄断现象, 从设立保护期限、合理使用制度以及强制许可制度对衍生数据专有权进行限制, 以期对衍生数据保护提供合理的构想, 促进数据资源的流通与共享。

参考文献

- [1] 孙莹. 企业数据确权与授权机制研究[J]. 比较法研究, 2023(3): 56-73.
- [2] 黄武双, 邱思宇. 数据资源的类型化赋权[J]. 吉首大学学报(社会科学版), 2023, 44(5): 111-121.
- [3] 吴汉东. 数据财产赋权的立法选择[J]. 法律科学(西北政法大学学报), 2023, 41(4): 44-57.
- [4] 高阳. 衍生数据作为新型知识产权客体的学理证成[J]. 社会科学, 2022(2): 106-115.
- [5] 许可. 从权利束迈向权利块: 数据三权分置的反思与重构[J]. 中国法律评论, 2023(2): 22-37.
- [6] 刘双阳, 李川. 衍生数据的财产属性及其刑法保护路径[J]. 学术论坛, 2020, 43(3): 39-47.
- [7] 崔国斌. 著作权法原理与案例[M]. 北京: 北京大学出版社, 2014: 8-9.
- [8] 蔡伟钊. 数字经济时代企业数据的法律保护与模式构建[J]. 社会科学家, 2023(12): 86-92.
- [9] 高宁宁. 企业数据财产权的法律属性界定[J]. 甘肃政法大学学报, 2022(5): 97-110.
- [10] 陶乾, 李衍泽. 论衍生数据的知识产权保护模式[J]. 大连理工大学学报(社会科学版), 2023, 44(4): 94-101.
- [11] 易继明, 李春晖. 知识产权的边界: 以客体可控性为线索[J]. 中国社会科学, 2022(4): 120-139+207.
- [12] 和育东. 从权利到功利: 知识产权扩张的逻辑转换[J]. 知识产权, 2014(5): 9-14.
- [13] 肖冬梅. “数据”可否纳入知识产权客体范畴? [J]. 政法论丛, 2024(1): 137-148.
- [14] 冯晓青. 知识产权视野下商业数据保护研究[J]. 比较法研究, 2022(5): 31-45.