

论数据确权的一般化路径

阮神裕

摘要 《关于建构数据基础制度更好发挥数据要素作用的意见》(以下简称《数据二十条》)颁布之后,不少观点主张将权利客体区分为原始数据、数据集合和数据产品,分别赋予不同类型的数据产权,形成数据确权的复杂化路径。但实际上,区分确权缺乏必要性和可行性,不是最优的确权路径。绝对权意义上的财产权通常具有一般化的特征,表现为客体的抽象性、消极权能的显著性和积极权能的总括性。财产权之所以是一般化的,是由其调整权利人和不特定第三人法律关系的功能决定的。数据产权同样应该是绝对权意义上的财产权,因而数据确权应当采取一般化路径。数据产权的客体应该只是数据,不应该区分原始数据、数据集合和数据产品;数据的物理边界由数据在虚拟层面上的可访问性决定;以数据的可访问性为标准,可以区分建构公开数据产权和非公开数据产权。在数据确权一般化路径的基础上,还应当引入治理策略,一是对法定在先权益进行分离保护,二是明确规定法定访问权和使用权,从而更加精确地调整数据之上承载的多元利益。

关键词 数据确权;《数据二十条》;数据产品;可访问性;公开数据;数据权益

中图分类号 D913 **文献标识码** A **文章编号** 1672-7320(2024)05-0148-12

基金项目 最高人民法院 2023 年度司法研究重大课题(ZGFYZDKT202314-03)

2022 年 12 月 2 日,中共中央、国务院印发了《数据二十条》,强调建立保障权益、合规使用的数据产权制度。数据确权的研究重心逐渐转向了数据产权的法律构造。在建构数据产权的众多方案中,存在一种区分原始数据、数据集合和数据产品分别确权的路径,由此引发了数据产权的复杂化。问题在于,财产权在实践推理中的功能是指导海量的、不特定陌生人的行动,这一功能决定了财产权的结构不能太复杂,而应当是标准化或者一般化的^[1](P145)。过于复杂化的数据产权体系是否符合财产权的功能和运行原理?数据确权能否以及如何采取一般化的路径?在一般化数据产权的法律构造基础之上,又该如何更细致地保护数据生产各参与方的合法权益?本文拟对以上问题作反思性的讨论,为数据确权提供参考。

一、数据确权的复杂化路径及其问题

在对数据确权的讨论中,理论界和实务界均有不少观点主张依据数据的全生命周期,区分原始数据、数据集合和数据产品的客体类型,分别确立不同类型的数据产权。此即本文所称的数据确权的复杂化路径。然而,该路径缺乏必要性和可行性。

(一) 数据确权的复杂化路径

《数据二十条》提出要建立数据资源持有权、数据加工使用权和数据产品经营权等分置的产权运行机制,该政策文件试图在权利客体的维度上区分原始数据、数据集合和数据产品,并且分别进行确权。原始数据是指由数据持有者收集的关于客观世界的数字化原始记录;数据集合(或称数据资源)是指经

过汇聚、清洗、整理以及分类等初步加工形成的集合；数据产品则是指经过深度加工之后形成的数据产品或服务。这是《数据二十条》的独创之处。欧盟《数据法》(Data Act)^①“鉴于条款”第15条明确指出，从数据中推断或衍生出的信息，即从数据中赋予价值或洞察的额外投资结果，特别是使用专有的复杂算法得出的结果，不适用该法。

在《数据二十条》颁布后，我国学界提出的数据确权方案也采取了复杂化的路径。概言之，该路径主张：随着数据持有者对数据的劳动与投入的增加，数据的独创性程度或者数据价值生成逐渐深化，区分原始数据、数据集合和数据产品三种不同阶段的客体，数据持有者应当对各个阶段的客体享有一种逐渐变强的数据产权。兹举几例对数据确权的复杂化路径加以说明。

第一种观点主张，数据来源者享有数据所有权，数据处理者享有派生的数据用益权。依据价值生成阶段的不同，数据用益权进一步区分为数据资源持有权、数据集合的加工使用权和数据产品经营权。其中，数据资源持有权是一项弱权利，只能对抗他人的不当爬取，须经授权才能积极支配和使用数据。数据集合的加工使用权来自数据来源者的授权，须经登记方可对抗第三人。数据产品是从数据中提炼出来的具有价值的信息，成为独立的权利客体，其上成立数据产品所有权，包括占有、使用、收益和处分四项权能^[2](P39-47)。

第二种观点主张，随着劳动程度的逐步深化，在原始数据、数据资源和数据产品上分别成立三种不同法律构造的持有权。具体而言：首先，原始数据持有权的法律构造取决于其是否包含个人数据或公共数据。若不包含个人数据和公共数据，原始数据持有权就具有完全支配性和绝对排他性；反之，则仅具备有限支配性和有限排他性。其次，数据资源持有权的法律构造也取决于数据资源中是否残存着个人数据和公共数据，分别成立完全支配绝对排他和有限支配有限排他的两种权利。最后，数据产品蕴含着企业的实质性加工，因此数据产品持有权的法律构造为完全支配和绝对排他。至于加工使用权和经营权则是持有权的许可使用的产物，它们与持有权可以并行不悖、同时存在^[3](P62-67)。与前一观点不同的是，该观点主张原始数据、数据资源和数据产品分别成立三种不同的持有权，加工使用权和经营权则是从持有权中分离出来的子权利。可以想象，这一观点确立的权利类型的数量相对较多。

第三种观点主张，依据独创性程度的不同来区分原始数据、数据集合和数据产品，其中只有数据集合和数据产品有确权的必要性。数据集合虽然蕴含着企业的实质性劳动，但是还没有达到独创性的程度，应当成立数据集合邻接权。数据产品达到了独创性程度，应当成立类似于著作权的企业数据产品持有权与企业数据产品使用权。企业数据产品持有权表征的是数据产品的产权归属，企业数据产品使用权包括经营权、收益权或者销毁权等^[4](P193-199)。

以上三种观点的共同之处在于，在原始数据、数据集合和数据产品的生命周期中，随着价值生成阶段、劳动程度或者独创性程度的逐步深化，数据持有者享有一种逐渐增强的数据产权。一方面，在积极权能的层面上，随着数据生产程度的逐步深化，数据持有者的积极权能也逐渐增强，越来越不受个人信息权益或公共利益的限制。另一方面，在消极权能的层面上，随着数据生产程度的逐步深化，数据持有者的消极权能也逐渐增强，数据产权具有越来越强的排他性特征。有的学者形象地使用游尺来比喻：在游尺的两端，一端完全无排他性，另一端具有完全支配性和绝对排他性，数据产权落在该谱系的哪一位置，取决于数据的种类为何^[5](P168)。

(二) 数据确权的复杂化路径面临的问题

依据数据的生命周期区分原始数据、数据集合和数据产品分别进行确权，其合理性表现在，其尝试根据数据类型的劳动程度不同，赋予不同强度的数据产权，在较为精细的颗粒度上实现劳动程度与权利

^① 《数据法》的全称为 Regulation of The European Parliament and of The Council on Harmonised Rules on Fair Access to and Use of Data and Amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act)。

强度的精确匹配,从而最大限度地契合财产权的劳动理论。但是笔者认为,过分复杂化的确权方案既不必要,也不可行。详言之:

1. 欠缺必要性

数据确权的复杂化路径面临的首要问题在于,是否有必要对原始数据、数据集合或数据产品进行区分确权?该路径主张不同数据类型上分别成立不同类型的数据产权。然而,这些所谓的不同类型的数据产权之间的区别只是表面性的。不管最初对数据类型的分类多么复杂,数据产权的法律构造总是大同小异。

一种观点主张数据集合成立有限财产权,数据产品成立新型财产权。前者的权利内容是对数据集合进行控制、使用、许可他人使用或转让的权能,以及禁止他人对数据集合的实质内容进行复制使用的权能。后者的权利内容是对数据产品的控制、使用、传输和处分的权利,以及请求恢复数据产品完整性和删除数据产品的请求权^[6](P116-120)。然而,数据集合有限财产权与数据产品新型财产权之间没有实质性差异。就积极权能而言,权利主体对数据集合享有的是控制、使用、许可他人使用或转让的权能,对数据产品享有的是控制、使用、传输和处分的权利,二者没有本质区别。就消极权能而言,数据集合有限财产权的相关义务,乃是未经同意不得复制数据集合的实质性内容,数据产品新型财产权的相关义务,亦是未经同意不得随意复制数据产品,二者同样没有本质区别。可以说,该二者之间的差别仅仅是客体,权利内容几乎没有差异。

还有一种观点认为,数据集合因其不具有独创性而成立邻接权,数据产品因其独创性成立类似于著作权的企业数据持有权和使用权^[4](P195)。的确,邻接权与著作权存在一定的差异——邻接权人享有的专有权利在数量上少于著作权人享有的专有权利——但是,仅仅宣称数据集合成立邻接权、数据产品成立著作权,只是提出了数据集合与数据产品的权利结构应当存在差异,还没有说明二者的实际差异是什么。《中华人民共和国著作权法》(以下简称《著作权法》)没有关于邻接权的一般规定,而是列举了表演者权、录音录像制作者权和广播电台、电视台的权利,因而难以从中归纳出邻接权的法律构造。故此,仅仅声称数据集合成立邻接权,还不足以揭示数据集合邻接权的具体内容。有的学者认为,不管是数据集合邻接权,还是数据产品持有权,均不具有完全排他性^[7](P81)。如此一来,数据集合邻接权和数据产品持有权的排他性几乎没有差异了。

数据确权的复杂化路径的初衷是,依据数据的劳动程度之差异进行类型化,赋予各种数据类型不同强度的数据产权。然而以上论证表明,尽管该路径对数据类型的区分十分精致,但数据产权的权利内容没有实质性差异。如此一来,区分原始数据、数据集合与数据产品分别确权的必要性将会大打折扣。

2. 缺乏可行性

数据确权的复杂化路径还面临着缺乏可行性的问题。复杂确权论者主张根据劳动程度或独创性程度,将数据区分为原始数据、数据集合和数据产品分别确权。然而,即使是不对数据产业实践进行考察,也可以想象,这几个概念之间的界分将十分困难。数据的价值需要得到多大程度的实现,数据持有者需要投入多大的劳动,数据的独创性需要达到何种程度,原始数据才能蜕变为数据集合,数据集合才能进化为数据产品?这个问题恐怕没有办法得到具有确定性的答案。由于原始数据、数据集合和数据产品之间的界分存在不确定性,因此当一个陌生人在面对数据时,可能因为无法作出准确判断而实施不法行为,这将严重限制陌生人的行动自由。

为了弥补这种不确定性,有的论者提出通过“举证责任的分配机制”进行平衡,由数据产权的权利主体对自己享有的数据产权承担更重的证明责任^[8](P76)。还有一种观点认为,由于数据资源可能同时承载着个人利益和公共利益,因而数据资源持有权的排他性应当受到限制,至于该排他性在多大程度上受到限制,不宜预先作出绝对化规定,而是应当结合具体个案进行判断,而相应的事实举证责任应当由企业承担^[3](P66)。倘若数据产权主要是在司法裁判中发挥作用,也即数据确权所要确立的是裁判规

范,那么可以通过举证责任的分配机制,将依据标准作出裁判的不确定性及其风险转嫁给实力雄厚的一方,从而最大限度地修正“价值生成阶段”“劳动程度”或者“独创性程度”这些标准带来的不确定性。但是,包括数据产权在内的任何一项财产权发挥规范性的场景,主要不是在法院,而是在日常生活中。因此,若是将数据产权置于日常生活中的实践推理中,那么界定原始数据、数据集合和数据产品的可行性问题将会被放大,成为一个难以轻易克服的理论难题。

二、数据确权一般化路径的理论证成

数据确权的目标之一是确立一个具有绝对权意义上的财产权。在现行法中,诸如所有权、著作权、商标权等财产权具有绝对性,因而通常具有一般化的特征。这是因为财产权承载着调整权利人与海量的、不特定的陌生人之间法律关系的法律功能。这就决定了,数据产权要想被确立为一项具有绝对权意义上的法定财产权利,那么数据确权也应当采取一般化路径。

(一) 法定财产权利的一般化特征

《中华人民共和国民法典》(以下简称《民法典》)的体系建构以民事权利为中心展开,总则编专设“民事权利”一章对我国民事权利体系进行总括性规定,其中包含了物权和知识产权等绝对权意义上的法定财产权利。由于这些法定财产权利具有对世性,因而其法律构造通常呈现出一般化的特征。

第一,法定财产权利的客体具有抽象性。现实生活中存在纷繁复杂的外在客体,但是法定财产权利的类型是有限的,其不总是随着客体类型的变化而有所差异。以所有权为例,在现实世界中,既有土地、房屋或者其他建筑物等不动产,也有原材料、半成品或者产品等动产;但是在法律世界中,这些外在的表象被抽象为不动产和动产,并且只成立一种类型的财产权即《民法典》第240条意义上的所有权。之所以作此抽象,是因为不管客体是不动产抑或动产,也不管客体是动产中的原材料、半成品或者产品,它们在所有权所要调整的法律关系中,均是无关紧要的。相反,倘若现行法针对动产的全生命周期,分别赋予原材料、半成品或者产品以不同类型的所有权,那么《民法典》物权编将会呈现出极其复杂的局面。故此,客体的抽象性决定了法定财产权利的一般性。

第二,法定财产权利的积极权能具有总括性。财产权所要实现的目的通常是权利人从财产中获得物质利益。正如边沁所说的:“财产不是别的东西,只是期待的基础,即从我们所拥有的某一事物中,根据我们与这一事物的关系产生得到某种利益的期待。”^[9](P138)《民法典》第205条规定“本编调整因物的归属和利用产生的民事关系”,其中的“利用”通常是指权利人从动产或者不动产中获取经济利益的行为。尽管《民法典》第240条将“利用”进一步具现化为占有、使用、收益和处分四项典型权能,但是理论界通常认为,所有权的积极权能不是四项典型权能的算术和,而是总括性地涵盖任何可以从不动产或动产中获得物质利益的利用权能^[10](P195)。这是因为,权利主体如何利用其财产,无需法律加以规定,而是其与生俱来的自由。法律可能为了公共利益或者他人私益而对财产权的积极权能稍加限制,但是不可能事先对权利主体如何利用其财产作出事无巨细的安排。因此,法定财产权利的积极权能通常具有总括性,这也使得其具有一般性的特征。

第三,法定财产权利的消极权能具有显著性。法定财产权利的消极权能,是指法定财产权利给不特定第三人施加的指向性义务(或称相关义务)。所有权的消极权能是他人未经同意不得以任何方式干涉所有权人之物。专利权的消极权能是他人未经同意不得以任何方式实施专利。消极权能的显著性表现为,其通常由权利客体的物理边界决定。权利客体的物理边界就是权利的边界。例如,传统民法在定义所有权概念时,将动产或不动产分配给所有权人,与此同时,动产或者不动产的物理边界决定了所有权的消极权能。所有权给他人施加的义务,原则上是通过动产或不动产的物理边界规定的。他人在判断是否存在义务,以及存在何种义务时,只需要明确物理边界即可。此时,物理边界作为简单信号向不特定第三人传达了“远离”的义务^[11](P1694)。又如,在知识产权中,专利权也是通过类似的简单信号,即是

通过“发明的物理边界”界定权利边界的。以化学发明为例,申请者通过声明化学发明的结构来对某一物质主张权利。对该物质的任何使用,无论申请者在申请时是否预见,都受专利权的禁止^[12](P179)。事实上,一项法益若想作为绝对权受到法律的保护,那么该法益应当具有可识别性。这种可识别性源于两个基础,一是社会与文化的共通认识,二是存在可感知的客体。故此,德国民法理论中的绝对权通常具有社会典型公开性^[13](P107-111)。总之,法定财产权利的消极权能通常是由权利客体的物理边界决定的,进而赋予了法定财产权利一般性的特征。

(二) 数据确权应当一般化的理由

绝对权意义上的财产权之所以应当具备一般性的特征,是由财产权在实践推理中发挥的指导行动的功能决定的。

首先,财产权在日常生活中具有指导不特定第三人的行动的功能。任何一个社会均面临对客体进行分配的问题。财产权制度承载着对外在客体进行分配的功能。一方面,法律体系将客体分配给某一主体,该主体有权按照自己的意愿处置这些客体,从中获得物质利益的满足;另一方面,法律体系还要求他人未经同意不得(或者不得以特定方式)干涉分配给权利人的客体。正如康德所述:“我与它(东西)如此结合在一起,以至于一个他人未经我的许可而使用它就会伤害我。”^[14](P40)故此,财产权的存在决定了不特定第三人行动自由的边界。

其次,由于不特定第三人通常是海量的,因此财产权的法律构造应该是一般性的。财产权的对世性决定了,每个人都必须认识到其有义务不干涉他人财产权。如此一来,必然产生巨大的信息成本和集体行动的协调问题。财产权的存在及其边界应该通过简单信号进行传达,以便被大量具有不同经验和智慧技能的人理解^[15](P1853)。在这个意义上,财产权应当具有一般性。相反,若是财产权的存在及其边界取决于“标准”,这有可能诱使每个人为了自己的利益解读这个标准,难以产生确定性。

最后,财产权对行动的指导功能,应当放在日常生活的场景中,而非诉讼场景中进行观察。财产权首先是行为规范,其次才是裁判规范。在数据确权的讨论中,有的论者意识到了数据确权的复杂化路径将会引发不确定性,并且提出用举证责任的分配机制化解这种不确定性,将因不确定性引发的风险分配给数据持有者^[8](P76)。这一思路在诉讼场景中或许是正确的,然而在日常生活的场景中却有问题。人们在日常生活中想要利用某一财产时,需要知道该财产之上是否存在法律权利以及相关义务,以便据以作出决策。在这样的场景中,很难想象人们可以要求某一主体对财产之上的法律权利及其相关义务进行举证。故此,通过举证责任的分配机制化解财产权的边界模糊带来的不确定性问题,忽视了财产权的行为规范的功能。

基于以上分析可知,数据产权若要发挥法定财产权利指导他人行动的功能,那么数据产权的法律构造同样应当具有一般化的特征。数据确权的复杂化路径的主要问题在于,这些方案没有提供关于数据产权施加给他人的相关义务的具体内容,因而无法在日常生活中为人们的行为提供明确的指引。

(三) 数据确权一般化的步骤

当然,数据确权的复杂化路径并非毫无价值,这些观点在探索数据产权与其所要实现的终极目标(如对劳动价值的保护)之间的关系上作出了贡献。但是,过于复杂的数据确权方案将会导致不特定第三人在判断自己是否可以实施某一行行为时陷入困境。因此,数据确权的工作不应当一开始就设计一个异常复杂的数据产权体系,而是应当分为以下两个步骤。

其一,建构一般化的数据产权的法律构造,明确数据产权的排他性。数据在何种情况下应当专属于一方主体,并且排斥其他主体的访问或者使用。为了实现这一目标,数据确权的第一步采取的手段应该是排他策略,即通过诸如客体的物理边界这样的简单信号向不特定第三人传递权利的存在与边界。不管是有形财产还是无形财产,法定财产权利的客体及其物理边界,在决定法定财产权利的权利边界以及不特定第三人行为自由的界限时,均发挥着重要作用。同样,在数据确权工作中,必须明确数据产权的

客体及其物理边界,并且通过该物理边界初步锚定数据产权的轮廓。

其二,在一般化的数据产权法律构造的基础上,通过更加精细的利益衡量和权利配置,对数据确权的一般化路径进行规则补充。为了实现这一目标,数据确权的第二步应当采取的手段应该是治理策略。所谓治理策略是指在界定财产权时,通过更直接地界定谁可以或者不得实施何种行为的方式,来界定财产权的内容与边界。治理策略界定的是一个人有权利行使某种行为^[12](P151)。以所有权为例,所有权界权中的排他策略体现为,以动产或者不动产的物理边界为界,他人未经同意不得以任何方式干涉所有权人之物;所有权界权中的治理策略则表现为,《民法典》第288条对相邻关系的特别规定,如不动产权利人因建造、修缮建筑物以及铺设电线、电缆、水管、暖气和燃气管线等必须利用相邻土地、建筑物的,该土地、建筑物的权利人应当提供必要的便利(《民法典》第292条)。同样,数据确权在排他策略的基础上,也应当采取治理策略对数据生产各参与方享有的合法权益进行更加细致的界定。

三、数据确权一般化路径的具体展开

鉴于数据产权在实践推理中发挥着指导他人行动的功能,因此,数据确权所要建构的数据产权的法律构造应当采取一般化的路径。数据确权的一般化路径包括两个层次:一是数据产权的客体应当是数据,没有必要区分原始数据、数据集合或数据产品;二是应当通过数据的物理边界进行确权,而数据的物理边界是由可访问性决定的。依此路径,可以初步形成公开数据产权和非公开数据产权的二元构造。

(一) 数据作为数据产权的客体

数据确权的一般化路径首先意味着数据产权的客体应当只有一种,即数据,而不需要区分数据集合、数据产品。复杂确权论者倾向于根据数据的类型不同而分别确权,根据劳动投入的程度、创造性程度或者其他标准,将数据分为原始数据、数据集合和数据产品,赋予从弱到强的财产权利。这种确权方案缺乏必要性和可行性。为了克服这种复杂性与不确定性,应当将数据产权的客体界定为数据。

第一,原始数据与数据集合均属数据,没有必要区分二者分别进行确权。二者的差异在于,数据集合蕴含着更多的劳动投入或者创造性较高,但是它们本质上均属于数据。程啸认为,从民事权利客体的角度而言,没有必要区分数据、数据集合和数据产品,它们只是从经济学的角度出发对数据进行划分^[16](P141)。这一观点值得赞同。在工业制造业中存在原料、材料和成品的区分,该区分只是对制造工艺与会计账目存在意义;但在法律世界中,原料、材料和成品上的民事权利、权利变动规则以及权利救济规则没有实质性差异,均可涵摄至《民法典》第240条的所有权。同样,原始数据与数据集合均属数据,没有必要作区分确权。

第二,数据产品不应该成为数据确权的客体。首先,数据产品概念的外延十分宽泛。在最广义上,任何涉及数据处理的产品均可以被称为数据产品。一些业界人士将收集数据的埋点技术、汇聚数据的数据中台等称为数据产品^[17](P52)。这些用于处理数据的产品要么是代码,要么是经营方案,无法成为数据确权的客体。在讨论数据确权时,数据产品只能是基于数据产生的产品,而非用于处理数据的产品。其次,基于数据产生的产品可能是衍生数据,即通过大量智力劳动投入,运用算法及分析模型将海量原始数据深度加工、分析与提炼生成的具有市场价值的商业数据集^[18](P106)。例如,数据处理者将他人收集的真实路况的数据进行标注形成衍生数据,并将该衍生数据作为产品出卖给自动驾驶的开发者。但在笔者看来,衍生数据仍是数据,其应当与数据成立统一的数据产权。最后,基于数据产生的衍生产品可能是智慧决策、计算机软件程序、商业方案等,但其并非数据确权的适当客体。一方面,诸如ChatGPT等生成式人工智能是基于海量数据训练而产生的,虽然属于数据的衍生产品,但是它本身不是数据,其或许因具有独创性受到著作权的保护,但是不再是数据确权的对象,否则将会造成“破坏性的权利重叠”^[19](P71)。另一方面,诸如淘宝商城提供的“生意参谋”等提供实时数据的应用程序也被定性为数据产品。此类数据产品的功能实现依赖于两个要素,一是数据,二是程序(或算法)。第三人侵害行为的目

标可能是程序,例如复制代码或者破坏程序的稳定运行。但是这类侵害行为与数据产权的问题没有直接关系。当第三人侵害行为的目标是数据时,该数据产品的保护问题实质上是数据保护的问题。例如在一个案例中,某甲公司创立了一个互助平台,该平台允许“生意参谋”的用户在该平台上共享账户和密码,第三人可以通过购买或者租赁账号密码的方式访问“生意参谋”提供的数据。尽管审理法院使用了“数据产品”的表达方式(浙江省杭州市中级人民法院(2018)浙01民终7312号民事判决书),但是某甲公司的行为实质上针对的是该应用程序中的数据。

第三,《数据二十条》作为政策文件不应该成为法律上确立数据产权的直接依据。《数据二十条》提到了原始数据、数据资源和数据产品等概念,例如“推动数据处理者依法依规对原始数据进行开发利用”,又如“建立数据资源持有权、数据加工使用权、数据产品经营权等分置的产权运行机制”等。但是,参与该文件制定的学者指出,之所以在持有、使用和经营这些行为动词之前加上资源和产品等前缀,主要是为了符合商业用语习惯,而与数据是否经过匿名化处理或者是否系标准化产品等问题关系较弱,在法学说和实定法表达层面,这些限定用语的意义并不明显^[20](P1156)。也就是说,数据资源持有权、数据加工使用权和数据产品经营权的构词规则只是修饰性的,确立数据资源持有权的同时,也应该确立数据产品持有权;或者反过来说,确立数据产品经营权并不意味着数据资源不能成立经营权,即便是原始数据也可以发生流转交易,从而产生原始数据经营权,只不过原始数据的经营行为需要得到审慎对待。除此之外,《数据二十条》作为政策性文件,不能简单等同于将来可能的数据产权立法。政策性文件提出的经济目标若要转化成法律规范,仍有相当长的一段道路要走。“经济学的一般理论主要是在讨论抽象的概念,如财产、交易、企业、资本和劳动,律师和法学教授处理的则是这些概念的实际状况。”^[21](P14)故此,不能以教义学的态度对待《数据二十条》的具体内容,将其作为数据确权立法的直接依据。

(二)数据可访问性作为确权标准

财产权之所以具有一般性,是因为财产权通常使用财产的物理边界作为简单信号传递该财产权施加给他人的相关义务。同样,在数据确权的一般化方案中,数据产权的消极权能也应当通过数据的物理边界这一简单信号进行描述。这就产生了一个问题,数据存在物理边界吗?如果存在的话,应当如何描述?对此问题,笔者认为应当在虚拟层面上通过可访问性界定数据的物理边界。

首先,数据的物理边界应当在操作系统和应用程序呈现的虚拟层面上界定。在纯粹物理意义上,数据不具有显著的物理边界。电子数据的本体是数据载体中的电磁结构。数据载体或许具有显著的物理边界,但是以微米级别存在的电磁结构很难产生显著性。因此,数据的物理边界难以通过纯粹物理意义上的电磁结构进行界定。然而,数据通常不是以其本体形式参与人类社会的日常生活与市场活动的,而是通过计算机的操作系统和应用程序以虚拟形式呈现在人类社会中的。不管是普通用户对数据的访问、复制或者简单处理,还是专业人士对数据的清洗、分析或者挖掘,均以计算机的操作系统和应用程序为媒介^[22](P249)。换句话说,人们对数据的实际支配是在虚拟层面上进行的,数据的物理边界也应当在虚拟层面上进行界定。

其次,在操作系统和应用程序呈现的虚拟层面上,数据的物理边界是由数据的可访问性决定的。在普通用户日常浏览网页的场景中,用户浏览行为的边界,是由数据的可访问性决定的。在数据处理者处理个人信息时,个人信息若是已经公开的(即可访问的),依据《中华人民共和国个人信息保护法》(以下简称《个人信息保护法》)第27条的规定,数据处理者可以在合理的范围内自由处理,除非信息主体予以拒绝。在数据处理者爬取数据时,亦应根据数据的可访问性进行区分评价,数据若是公开可访问的,爬取行为原则上不具有违法性,除非该爬取行为构成不正当竞争;相反,数据若是不可访问的,或者受限访问的,那么攻击技术保护措施爬取数据的行为将会落入刑法的评价范畴。在现实世界中,动产或者不动产的物理边界既决定了权利主体的支配范围,也决定了其他主体行为自由的边界。而在计算机的操作系统和应用程序提供的虚拟世界中,数据的可访问性决定了数据持有者对数据的使用范围,同时也决定

了其他主体的行为边界。

最后,数据的访问行为与其他利用行为难以截然区分,因此在允许访问的同时,通常也应当允许其他利用行为。一方面,数据的访问行为是一切后续利用行为的基础。复制、使用、分析、挖掘或者公开数据,甚至是破坏数据的行为,均建立在访问数据的基础上。消费者对文本、图片或者视频等非结构性数据的利用,以访问数据为前提条件;对于消费者没有购买访问权限的数据,消费者没有访问可能性,因此也就无法利用数据。经营者对其他经营者的业务数据的利用,也以访问数据为前提条件。提供数据的一方可以将数据直接复制给需求方,也可以开放API端口给需求方,这些机制均旨在实现需求方对供给方的数据的访问。因而有的学者认为,一个数据驱动的业务模式和经济模型,主要不是取决于一个企业通过占有载体实际控制的数据的大小,而是取决于该企业可以访问的数据的多少^[23](P1224)。另一方面,访问数据通常难以和其他利用行为进行有效的区分。倘若一个网站允许不特定人访问数据,通常意味着这些人也可以复制数据。一些网络服务提供者尝试在提供数据访问的同时,禁止用户复制数据。但是,用户往往在访问之后采取各种手段对数据进行复制。由此引发了网络服务提供者进行加密、用户通过各种手段进行解锁的拉锯战,极大地阻碍了信息流通。由此可见,区分数据的访问和利用将会造成较高的管理成本。故此,不妨以数据的可访问性为标准,可访问的数据通常也是可利用的,不可访问的数据通常就是不可利用的。

综上,尽管数据不像动产或者不动产一样具有明确的四至边界,甚至不像作品、专利或者商标一样具有可视化的边界,但是数据的可访问性界定了一条非常明确的物理界线,为数据确权奠定了基础。

(三) 公开数据与非公开数据的产权构造

依据数据的可访问性,可以区分公开数据产权和非公开数据产权。公开数据产权与非公开数据产权的差异主要体现在消极权能上,而非积极权能上。

一方面,在积极权能的层面上,不管是公开数据产权,还是非公开数据产权,其积极权能均具有总括性,即只要不与其他主体的合法权益或者法律规范发生冲突,数据产权的权利主体可以按照自己的意愿处分数据,包括事实上的处分行为,如数据的清洗、分析、挖掘、删除或者销毁行为,也包括法律上的处分行为,即转让或者许可他人使用数据。故此,公开数据产权与非公开数据产权的法律构造没有实质性差异。

事实上,一些学者提出的数据确权方案虽然看起来存在较大差异,但是落实到其所确立的数据产权的积极权能时,表面上看似各异的方案又呈现出高度的相似性。例如,有的学者主张数据产权应当采取权利束理论进行界定,在权利束理论下,数据产权的积极权能包括持有、使用、收益和处分的权能^[24](P71)。有的学者主张数据产权的性质应当是新型财产权,数据财产权的积极权能包括利用权能、收益权能、占有(持有)权能和处分权能^[25](P161-163)。以上观点共同认为数据不应当成立所有权,但其提出的数据产权的积极权能却与《民法典》第240条的规定如出一辙。

数据产权的积极权能之所以存在相似性,是因为法律没有必要一一描述和界定数据产权的积极权能,不管现行法是否规定数据产权,也不管规定何种类型的数据产权,数据持有者均有权按照自己的意愿事实性处分数据。这是因为,在市场经济体制下,数据持有者对数据进行事实性处分的自由,不是由法律赋予的,而是原本就享有的。目前我国还没有颁布确立数据产权的法律或者行政法规,然而数据持有者始终可以按照自己的意愿对数据进行事实性处分,从而实现自己的经济目标。因此,明确列举数据产权的积极权能意义不大。数据持有者以何种方式利用数据,取决于其技术能力和市场需求,而不取决于法律如何规定。当然,不管是非公开数据还是公开数据,数据持有者的积极权能可能受到其他主体的合法权益,或者法律法规的限制。

另一方面,在消极权能的层面上,公开数据产权和非公开数据产权之间存在显著的差异性,具体而言:数据若是可访问的或公开的,不特定第三人不受这些义务的拘束,原则上可以自由访问、复制或者使

用数据,除非数据利用行为构成不正当竞争。数据若是不可访问的或非公开的,那么不特定第三人负有未经同意不得访问、复制或者使用数据的义务。当然,不管数据是公开还是非公开的,不特定第三人均应承担不得破坏数据完整性的义务。故此,非公开数据产权和公开数据产权的差异体现在排他性上,前者具有较强的排他性,而后者仅享有有限的排他性^[26](P10)。

有的学者主张通过创设高度标准化的“数据一般财产权”进行数据确权,这一权利的标准化程度如此之高,以至于公开数据和非公开数据的差异无关紧要,即应当赋予二者相同的数据一般财产权^[20](P1155)。依此观点,公开数据的持有者也可以禁止他人访问、复制或使用数据,但这不会影响社会公众对公开数据的访问和流通。这是因为,公开数据持有者虽然有权禁止他人访问、复制或者使用数据,但是不会真的执行这一权利^[27](P60)。笔者认为,这一观点值得商榷。该确权方案先是规定公开数据持有者具有禁止他人访问、复制和使用的权能,再将是否行使权能的选择交给公开数据持有者。该持有者或许在大多数情况下均不行使权能,然而在其认为有利可图时,可能通过提起诉讼行使权能。这将给数据需求方造成极大的不确定性。

综上,公开数据产权和非公开数据产权的积极权能十分相似,数据持有者有权按照自己的意愿随意处置其数据;二者的差异主要表现在消极权能上,非公开数据产权的不可访问性决定了其排他性强于公开数据产权。

四、数据确权一般化路径的规则补充

依据数据的可访问性初步界定了数据产权的轮廓,明确了权利人与不特定第三人之间的法律关系。然而,数据之上可能承载着错综复杂的利益关系,一是数据中的部分内容可能成立法定在先权益,二是数据生产涉及的多方主体均有利益主张。为此,以数据产权的一般化法律构造为基础,仍需通过治理策略更加精细地调整数据生产的各参与方的合法利益。此即数据确权一般化路径的规则补充。

(一) 法定在先权益的分离保护

数据之上承载的合法权益,除了数据持有者的数据产权之外,还可能包括个人信息权益、著作权或者商业秘密等法定在先权益。数据是任何以电子或者其他方式对信息的记录,信息既可能是外在客体的信息,也可能是自然人、法人或者非法人组织各类主体的信息。这些信息上可能承载着一定的人格利益或者经济利益。例如,电商平台收集的数据涉及消费者的身份、联系方式、浏览记录或者购买记录等个人信息。又如,绘画型生成式人工智能的训练数据涉及海量作品,这些作品受到《著作权法》的保护。《数据二十条》提出建立健全个人信息数据确权授权机制,并且充分保护数据来源者合法权益,这就表明数据确权工作始终涉及数据产权与法定在先权益的关系问题。

问题在于,该法定在先权益应当以何种方式参与数据产权的法律构造?一种观点认为,数据来源者既享有法定在先权益,也享有数据财产权,而数据持有者享有的,乃是从数据财产权分离出来的次级权利^[28](P117-127)。另一种观点认为,根据是否存在法定在先权益的限制,数据产权应当进行差异化配置,不受法定在先权益限制的,应当成立完全支配绝对排他的数据产权,受到法定在先权益限制的,则依情况分别成立完全支配无排他性或有限支配有限排他的数据产权^[3](P63-67)。这些确权方案的共同之处在于:法定在先权益的存在影响了数据产权的法律构造。

笔者认为,法定在先权益固然应当受到法律保护,但其不应当影响数据产权的法律构造,而是应当从数据产权中分离出来予以单独保护。具体而言:

第一,法定在先权益的存在不应当影响数据产权的法律构造。一方面,就消极权能而言,非公开数据产权施加给不特定第三人的义务,即未经同意不得访问、复制或者使用,不受到法定在先权益的影响。公开数据原则上允许他人的访问、复制或者使用,但是其中某些数据片段成立的法定在先权益可能限制他人的访问自由。例如,公开数据若是包含作品的,作品的著作权将会施加未经同意不得复制、发行或

者公开传播的义务；公开数据若是包含专利的，专利权将会施加未经同意不得实施专利的义务。需要注意的是，法定在先权益没有改变公开数据产权的法律构造，只是在公开数据产权这一层次之上又施加了另一个层次的限制。法定在先权益与公开数据产权是相互分离的。另一方面，就积极权能而言，数据产权的积极权能的确可能受到法定在先权益的限制。例如，数据包含个人信息时，对该数据进行处理的行为，将受到个人信息权益的限制。但是，此种限制只是影响了数据产权的权利行使，而不改变数据产权的法律构造。可供参照的是《民法典》对所有权的规定。《民法典》只规定了一种所有权，但所有权在不同场景下受到的限制存在差异。《民法典》第279条第1句规定，业主不得违反法律、法规以及管理规约将住宅改变为经营性用房。这决定了住宅所有权的权利行使不同于商业店铺所有权，但这不会导致住宅所有权的法律构造不同于商业店铺所有权，二者均属《民法典》第240条意义上的所有权。同样，法定在先权益的存在可能使数据产权的权利行使受到限制，但是没有必要根据权利行使的差异而建构不同类型的数据产权。

第二，法定在先权益与数据产权的一般构造之间具有可分离性。所谓可分离性是指，法定在先权益与数据产权是两个相对独立的模块，只要明确了法定在先权益是如何限制数据产权的，那么法定在先权益的内部构造就可以进行相对独立的调整，而不会引起巨大的连锁效应；同样，数据产权也可以相对独立地进行设立或者修订，也不会对法定在先权益造成体系性波动。这就是“财产权的模块化理论”——模块内部的互动是紧密的，模块之间的互动则是稀少的，并且通常是标准化的，因而其中任何一个模块发生变动，都不会给其他模块造成严重影响^[12](P164)。事实上，现行法中不乏模块化的现象。肖像作品上既有肖像权，又有著作权，但二者具有可分离性。肖像权的法律构造自《中华人民共和国民法通则》（以下简称《民法通则》）至《民法典》发生了一定变化，《民法通则》第100条规定他人未经同意不得以营利为目的使用肖像，而《民法典》第1019条第1款规定他人未经同意不得使用肖像，不以营利为限。同样，著作权的法律构造自《著作权法》1990年颁布以来也发生了一定的变迁。但是，只要明确了肖像权与著作权的交互规则（《民法典》第1019条第2款），那么无论肖像权内部发生何种调整，均不影响著作权的法律构造；反之亦然。在这个意义上，肖像权与著作权具有可分离性。同样，只要明确了法定在先权益与数据产权的交互规则，那么法定在先权益就可以从数据产权中分离出来。法定在先权益的变更或者废除，均不影响数据产权的法律构造。

（二）法定访问权和使用权的明确规定

在实践中，数据的生产过程涉及多方主体，各方主体往往对该数据均有利益主张。由此产生了数据产权在数个协作生产者之间难以配置的问题。例如，诸如淘宝、京东或者拼多多等电商平台的经营模式决定了，电商经营的数据是由平台、店铺以及消费者共同生产的。平台提供基础设施投入了大量成本，其希望通过数据分析进一步优化平台运营。店铺的经营活动实际生产了数据，其希望通过数据分析优化营销方案，甚至希望将数据迁移到其他平台。数据产权如何在协作生产者之间进行配置的问题，涉及的是权利人的内部关系，而不涉及权利人与不特定第三人之间的外部关系，第三人未经同意擅自访问、复制或者使用非公开数据的，不管是平台还是店铺，均可主张防御性请求权。问题在于协作生产者的内部关系应当如何规范。

对此问题应当区分两种情况。第一种情况是每个协作生产者“各自持有”其生产的不同数据^[29]（P79-81），或者“平行持有”共同生产的相同数据^[20]（P1157-1159），即每个生产者均可以在自己的服务器上访问和使用数据，此时每个生产者当然享有数据的访问权和使用权。这种情况相对简单。第二种情况则较为复杂，即协作生产的数据是由处于强势地位的一方实际控制，处于弱势的一方对数据的访问和使用取决于另一方提供的渠道。以电商平台为例，店铺通常没有能力使用自己的服务器存储数据，其对数据的访问，依赖于平台提供的应用程序。此时，平台可能通过格式条款限制或者剥夺另外一方的数据访问权或者使用权。

为了解决这一问题,欧盟的《数据法》的立法模式可供借鉴。该法尝试通过建立明确和公平的数据访问和使用规则,实现数据价值的公平分配。特别是随着物联网的普及,由于联网产品(如智能汽车、智能家居等)的数据是由是联网产品的设计商、制造商或者相关服务的提供者持有的,实际使用联网产品及其相关服务的用户(可以是自然人、法人或者非法人组织或消费者)往往很难访问和使用相关数据。而这些数据对于售后市场、辅助服务或其他服务而言是重要的生产要素。为了实现数据共享、促进数据驱动的创新经济,《数据法》致力于为用户单方配置数据访问权和使用权:一是赋予用户直接访问数据的权利,即联网产品或者相关服务的设计、制造或提供,应当具备允许用户直接访问数据的功能;二是授予用户请求提供数据的权利,用户若是无法直接在联网产品或相关服务中访问数据的,其有权请求数据持有者提供数据;三是规定用户有权按照自己的意愿使用联网产品或者相关服务生成的数据,包括向第三方共享数据。总之,该法通过单方面赋予用户数据访问权和使用权的方式,协调了用户与联网产品、相关服务的提供者之间的数据控制能力的不对等关系,弥补了用户访问、使用和分享数据的不足。可以说,欧盟《数据法》规定的数据访问权、使用权和共享权,实质上是通过单方赋权调整用户与联网产品、相关服务的提供者之内部关系的“倾斜保护型法”^[30](P450-453),而非赋予任何一方主体对数据进行独占控制的排他权。

同样,在数个协作生产者之间存在地位不对等,以至于其中一方实际控制数据时,数据确权工作的重点应该是单方赋权,即赋予处于弱势地位的一方访问和使用数据的权利。当然,单方赋予一方主体法定的访问权和使用权,仍然是建立在数据产权的一般化方案的基础之上的:一方面,对于不特定第三人而言,其行为自由的边界仍然是由数据的可访问性决定的。在协作生产者内部进一步细致规定法定访问权与使用权,并不影响不特定第三人的行动。另一方面,对于潜在交易者而言,只要该访问权与使用权是通过法律法规的方式赋予的法定权利,而非约定权利,那么其就应当合理预期通过交易获得的数据可能负担着访问权与使用权的义务。因而法定的访问权与使用权的存在,不会给数据交易造成不确定性。

在数据确权的讨论中,不少观点寄希望于设计一个精细的数据产权体系,以实现劳动投入程度与数据产权强度之间的精确匹配,形成了数据确权的复杂化路径。然而,本文论证表明数据产权作为绝对权意义上的财产权,在实践中发挥着指导行动的功能。这决定了数据产权应当采取一般化的路径:第一,将数据作为数据产权的唯一客体,依据数据的可访问性标准,分别建构公开数据产权和非公开数据产权;第二,在数据确权一般化路径的基础上,针对数据之上承载的多元利益格局,制定数据产权的治理规范,调整权利人内部之间的法律关系。

参考文献

- [1] 熊丙万. 实用主义能走多远? ——美国财产法学引领的私法新思维. 清华法学, 2018, (1).
- [2] 申卫星. 论数据产权制度的层级性:“三三制”数据确权法. 中国法学, 2023, (4).
- [3] 孙莹. 企业数据确权与授权机制研究. 比较法研究, 2023, (3).
- [4] 吴桂德. 商业数据的私法保护与路径选择. 比较法研究, 2023, (4).
- [5] 周樾平. 大数据时代企业数据权益保护论. 法学, 2022, (5).
- [6] 姬蕾蕾. 企业数据保护的司法困境与破局之维:类型化确权之路. 法学论坛, 2022, (3).
- [7] 张素华. 数据产权结构性分置的法律实现. 东方法学, 2023, (2).
- [8] 徐玖玖. 利益均衡视角下数据产权的分类分层实现. 法律科学(西北政法学报), 2023, (2).
- [9] 吉米·边沁. 立法理论. 李贵方等译. 北京:中国人民公安大学出版社, 2004.
- [10] 崔建远. 物权法. 北京:中国人民大学出版社, 2021.
- [11] Henry E. Smith. Property as the Law of Things. *Harvard Law Review*, 2012, 125(7).

- [12] 亨利·E. 史密斯. 作为财产的知识产权——界定信息权利//易继明. 私法. 武汉: 华中科技大学出版社, 2016.
- [13] 于飞. 侵权法中权利与利益的区分方法. 法学研究, 2011, (4).
- [14] 康德. 道德形而上学. 张荣、李秋零译. 北京: 中国人民大学出版社, 2013.
- [15] Thomas W. Merrill, Henry E. Smith. *The Morality of Property*. William and Mary Law Review, 2007, 48(5).
- [16] 程啸. 论数据产权登记. 法学评论, 2023, (4).
- [17] 杨楠楠, 李凯东, 姚向雁等. 数据产品经理: 解决方案与案例分析. 北京: 机械工业出版社, 2022.
- [18] 高阳. 衍生数据作为新型知识产权客体的学理证成. 社会科学, 2022, (2).
- [19] 塞巴斯蒂安·洛塞等. 数据交易: 法律·政策·工具. 曹博译. 上海: 上海人民出版社, 2021.
- [20] 熊丙万. 论数据权利的标准化. 中外法学, 2023, (5).
- [21] 大卫·D. 弗里德曼. 经济学与法律的对话. 徐源丰译. 桂林: 广西师范大学出版社, 2019.
- [22] Johannes Kevekordes. *Daten Als Gegenstand Absoluter Zuordnung*. Berlin: Duncker & Humblot, 2022.
- [23] Udo Kornmeier, Anne Baranowski. Das Eigentum an Daten — Zugang statt Zuordnung. *Betriebs Berater*, 2019, (22).
- [24] 王利明. 数据何以确权. 法学研究, 2023, (4).
- [25] 张新宝. 论作为新型财产权的数据财产权. 中国社会科学, 2023, (4).
- [26] 崔国斌. 大数据有限排他权的基础理论. 法学研究, 2019, (5).
- [27] 熊丙万, 何娟. 数据确权: 理路、方法与经济意义. 法学研究, 2023, (3).
- [28] 申卫星. 论数据用益权. 中国社会科学, 2020, (11).
- [29] 宁园. 从数据生产到数据流通: 数据财产权益的双层配置方案. 法学研究, 2023, (3).
- [30] 丁晓东. 法律如何调整不平等关系? 论倾斜保护型法的法理基础与制度框架. 中外法学, 2022, (2).

On the Generalization of Data Property Right Delineation

Ruan Shenyu (Renmin University of China)

Abstract After the release of the "20 Key Measures", many advocate a complex path for data property right delineation by distinguishing the objects of rights into original data, data sets and data products, thereby assigning different types of property rights to different types of data. This complex path is not an optimal one for want of necessity and feasibility. Property rights, in the sense of absolute rights, normally have generalized characteristics, which are reflected in the abstractness of the object, the prominence of negative powers and the comprehensiveness of positive powers. The fact that property rights are generalized is determined by their function in adjusting legal relationships between right-holders and unspecified third parties. Data property rights should also be absolute rights; therefore, the delineation of data property rights should adopt a generalized path. The object of data property rights should be limited to data itself, without distinguishing between original data, data sets or data products; the physical boundaries of data are determined by the accessibility of data on a virtual level; based on the accessibility of data, it is possible to distinguish between public data property rights and non-public data property rights. On the basis of a generalized path, governance strategies should also be introduced, first by separating the protection of statutorily prior interests and second by clearly specifying statutory access and usage rights, thereby more precisely adjusting the multifaceted interest structure borne by data.

Key words data property right delineation; "20 Key Measures"; data products; accessibility; public data; data rights and interests

■ 作者简介 阮神裕, 中国人民大学法学院助理教授, 中国人民大学未来法治研究院助理研究员, 北京 100872。

■ 责任编辑 李 媛