

数据确权困境之检讨

——基于“二次确权”的视角

李至辰, 陈琛

浙江农林大学文法学院, 浙江 杭州

收稿日期: 2023年8月12日; 录用日期: 2023年9月26日; 发布日期: 2023年10月10日

摘要

数据确权, 是数据交易的基础, 是构建数据要素市场的应有之义。基于对数据生命周期之观察, 结合大家言论, 发现数据确权中存在着双重困境: 数据与信息之间的关系, 是横亘于数据确权的第一重困境; 数据权利的内部构造, 是决定数据确权研究方向的第二重困境。厘清数据与信息的迷思, 需要明晰信息主体与数据处理者之间的关系; 突破民法“物债二分”的桎梏, 需要在数据处理者内部构建起新型的权利结构。鉴此, 数据确权问题事实上被分解为“二次确权”问题: 第一次确权, 需要明定信息主体与数据处理者对于数据享有的权利; 第二次确权则需要对数据处理者内部之间的权利进行界分。

关键词

数据确权, 二次确权, 原始数据, 次生数据

Review of the Dilemma of Data Right Confirmation

—Based on the Perspective of “Secondary Right Confirmation”

Zhichen Li, Chen Chen

School of Humanities and Law, Zhejiang A&F University, Hangzhou Zhejiang

Received: Aug. 12th, 2023; accepted: Sep. 26th, 2023; published: Oct. 10th, 2023

Abstract

Data right confirmation is the basis of data transaction and the proper meaning of building a data element market. Based on the observation of data life cycle and our comments, it is found that there is a double dilemma in data confirmation: the relationship between data and information is

the first dilemma in data confirmation; the internal structure of data right is the second dilemma that determines the research direction of data confirmation. To clarify the myth of data and information, it is necessary to clarify the relationship between information subject and data processor, break the shackles of "property debt", and build a new right structure within the data processor. In view of this, the problem of data right confirmation is actually decomposed into the problem of "secondary right confirmation": the first right confirmation needs to define the rights enjoyed by the information subject and the data processor; the second right confirmation needs to divide the boundary between the rights within the data processor.

Keywords

Data Right Confirmation, Secondary Right Confirmation, Raw Data, Secondary Data

Copyright © 2023 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

2022年6月22日,中央全面深化改革委员会第二十六次会议审议通过了《关于构建数据基础制度更好发挥数据要素作用的意见》,该意见明确指出:“要统筹推进数据产权、流通交易、收益分配、安全治理,加快构建数据基础制度体系”。数据基础制度体系涵盖了数据产权制度、数据交易制度、数据收益分配制度和数据安全治理制度四大基本制度。数据产权制度作为数据基础制度中的首要规则,是数据要素市场规范运行的底层逻辑,也是数据交易、收益分配和安全治理的制度基础。从法学的视角研究数据产权,需要在厘定数据利益的基础之上,通过制度设计以明确数据利益的归属。概言之,需要构建数据权利规则体系确认数据主体的权利,即数据确权。《民法典》和《个人信息保护法》的出台,基本上明确了数据处理过程中涉及个人信息和隐私的保护规范。遗憾的是,这两部法律尚未对数据权利归属进行明确和宣示。近年来,法学界对于数据确权的研究如火如荼,学者们分别通过不同的研究进路与范式对数据确权问题进行研究,各种言说难以定论。鉴此,本文通过梳理学界数据确权的研究范式,尝试构建起“二次确权”的研究框架,为厘定数据确权之迷思,破除数据确权之桎梏,提出一隅之见。

2. 数据基本范畴之明晰

2.1. 数据的基本内涵

数据的基本内涵,作为数据法学研究的基础,对于数据权利的确定无疑具有奠基性。学界对于数据的基本内涵的争议,主要发端于不同的视角。第一,侧重于数据的表现形式。该观点认为,“数据限于在计算机及网络上流通的在二进制的基础上以0和1的组合而表现出来的比特形式”[1]。第二,倾向于数据的本质内容。该观点认为,“数据是信息的一种形式化方式的体现,该种体现背后的含义可被再展示出来,且该种体现适于沟通、展示含义或处理”[2]。我国《数据安全法》,是将国际主流的一种定义方式进行统合对数据进行定义,“数据是指任何以电子或者其他方式对信息的记录。”我国法律对于数据如此界定,一方面将数据的表现形式从电子方式拓展到其他方式,另一方面也彰显了数据是对信息的记录这一根本性质。然而,这般定义方式仍然存在一些理解上的模糊。一方面,关于数据的表现形式中的“其他方式”指的是与“电子方式”相似的方式还是任何方式?另一方面,对于数据的实质内容“对信息的记录”应当作何解释?鉴此,有必要从法律解释的角度明确上述两个问题。

首先,关于“其他方式”的解释,应当进行语义的限缩。纵观《数据安全法》亦或是当前全国各地出台的数据条例可以发现,其中对于“数据”的语义全都限定在计算机学科背景之下。因此,这里的“其他方式”并非是不加限制的任何方式,而是为了涵盖未来数据技术的发展出现新型的类似于电子方式的其他方式。反之,如果不对“其他方式”做缩小解释,就意味着一切能够记录信息的方式都属于数据,如此宽泛的对数据定义,会导致所有能够承载信息的事物都纳入到数据的外延统摄范围内。比如,壁画、书籍、图画、磁带、录像带、人的记忆等等能够记载信息的事物。很显然,如此宽泛的定义,并不符合一般观念对于数据这一概念的认知。既然在构词上采用了“数”和“据”,这里的“数”应当是特指数字形式,对其定义的“其他方式”也应当是限于类似数字的形式,如此才能明确数据外延的边界。

其次,关于“对信息的记录”的解释。这里的“信息”应做广义理解,其含义为对事物属性及相互关系的描述。那么数据的定义就可以转化为,对事物属性及相互关系描述的记录。数据作为一种记录,其本质上也是对事物属性及相互关系的呈现,这也意味着数据实质上就是信息,只是这种信息是采用了电子或其他方式进行记录的,这种记录方式才是信息的真正载体。由此可知,数据实质上可以等同于信息,但是由于其采用了电子或其他方式对信息进行记录而与一般意义上的信息存在着区别。对于信息与数据的区分,后文将会详细阐述。

综上所述,数据的概念可以总结为,采用电子方式或类似于电子方式对事物属性以及相互关系的描述。

2.2. 数据的相关主体

2.2.1. 信息主体

明确数据基本内涵,有助于进一步厘定数据之上的相关主体。如前文所述,数据是对信息的记录,这意味着信息所标识的主体与该信息所生成的数据之间存在着不可忽视的相关性。因此,信息所标识的主体应当属于数据之上的相关主体之一。鉴于“信息所标识的主体”这一表述过于冗长,有部分学者将其概括为“信息主体”[3],有部分学者采用“数据主体”[4]。这两种命名方式分别从信息和数据两个不同的层面对“信息所标识的主体”进行归纳,称谓不同,但殊途同归。然两相比较,使用“数据主体”一词指代“信息所标识的主体”存在些许不足。首先,“数据主体”一词则显得用语不明,有语义过泛之嫌。按照通常解释,数据主体是指一切与数据相关的自然人、法人和非法人组织。然而,在与数据相关的主体之中,不仅仅涵盖着数据的来源主体,同时还囊括了数据的生产主体。显而易见,信息所标识的主体所要概括的对象本应当仅限于数据的来源主体,不应当包含数据的生产主体。因此,“数据主体”一词含义过于宽泛;另一方面,即使对语义进行了限缩,也容易与“数据采集者”、“数据分析者”、“数据加工者”等数据相关主体相混淆,徒增理解之困难。与“数据主体”一词相比,“信息主体”一词更能精准和妥帖地反映出主体与数据的本质关系。“信息主体”是主体基于对信息享有的利益而涉及到该信息所生成的数据之上,因此成为数据的相关主体。故采用“信息主体”一词来指代“信息所标识的主体”则更加恰当。下文中除了引用专家学者的原文之外,其余部分也统一采用“信息主体”一词来指代“信息所标识的主体”。

2.2.2. 数据控制者

在数字时代,信息主体被裹挟在数字化的浪潮中,自身的信息被转化为数据以维系数字社会的运转。因此,可以得出数据来源于信息这一论断,但更严谨地表述应当为,原始数据的直接来源是信息。信息主体主动或被动地将自身相关的信息录入到数字设备之中,并以互联网设备作为中枢进行社会活动。这些信息一经录入到数字设备之中即转化为数据,这些数据经过网络传输被收集并存储到服务器之中。此时的数据相较于后续经过对数据进行分析、加工、融合等一些系列行为之后所得到的数据不同。这种数据是直接由信息生成的,是未经过任何人为性地增添或删减的数据,更加原始和天然,因此称之为“原

始数据”。通过合法地分析、加工、融合等方式对原始数据进行数据处理之后,此时得到的数据具有高附加值且富含边际效应,称之为“次生数据”。概言之,在数据的整个生命周期中存在着两次飞跃,第一次飞跃是由信息转化为原始数据,是数据的诞生;第二次飞跃是由原始数据晋升为次生数据,是数据的升华。在数据的第二次飞跃中,需要经过收集、存储、清洗、加工、传输、提供、公开等一系列数据处理环节,其中的每一个环节的主体可能同一,也可能不一,但相较于作为数据相关主体的一方——信息主体而言,这些数据处理环节中的数据相关主体可以被统称为数据控制者(通常也称为数据处理者) [5] [6]。

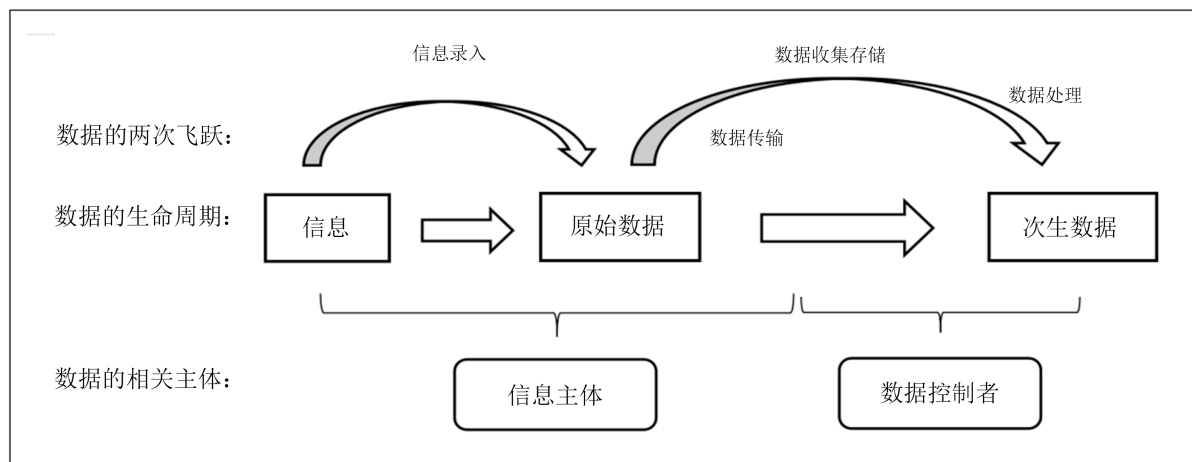


Figure 1. Life cycle of data

图 1. 数据的生命周期

综上观之,纯粹从信息主体的角度对数据进行观察,数据更侧重自身的功能属性,以体现其对信息的承载功能,在此种状态之下,信息主体基于对信息享有的权益而似乎延及到承载该数据之上;然而,一旦切换到数据控制者的视角,数据似乎更专注自身的价值属性。在一个完整的数据生命周期(见图 1)中,数据处理者必然会经过从信息→原始数据→次生数据三个阶段的数据处理工作,最终实现数据的“两次飞跃”。事实上,数据的每一次“飞跃”都意味着数据的性质和涉及的利益发生了巨大的改变。鉴此,对数据确权的研究,必须遵循数据的生命周期,分别聚焦于“信息→原始数据”和“原始数据→次生数据”这两个阶段,将数据确权问题一分为二,拆分为原始数据的确权问题和次生数据的确权问题依次进行研究和探讨。

3. 数据确权之迷思：原始数据与信息混同

数据确权的意义在于,通过法律来认可和保护数据利害关系人对数据享有的权益。基于以上对数据基本内涵的厘定,首先可以明确数据利害关系人有两方:一方为信息主体,一方为数据处理者。申言之,在“信息→原始数据”阶段,探讨数据确权问题,实质上就是对信息主体和数据处理者各自享有何种权益进行考量。然而,在考量双方对数据享有的权益的过程中,始终横亘着数据确权研究之迷思。

3.1. 数据确权迷思的根源

在“信息→原始数据”这一阶段,信息主体基于对自身信息所享有隐私权、人格权益、商业秘密、商业信誉等信息权益而涉及到该信息所生成的原始数据之上,同时数据处理者通过对原始数据的采集和存储,并期待对该原始数据享有开发利益。换言之,信息主体和数据处理者在该阶段对原始数据享有的权利(或权益)处于并行和交织的状态,如何厘清两方主体在原始数据之上的权利(权益)边界,是原始数据

确权研究中最核心的问题。权利(权益)边界的厘定,需要对信息权益与数据开发利益进行具体的考量。虽然信息主体对自身信息享有法律规定的信息权益,但是当这些信息被转化为原始数据之后,信息主体先前对于自身信息享有的权益究竟会在何种程度上追及到原始数据之上?是完全追及?还是部分追及?亦或是根本没有追及性?概言之,数据之上是否存在信息利益?显而易见,这一前置性问题的答案将影响着上述核心问题的解决。在事实上,当前学界在数据确权研究中的争论也发端于此。

3.2. 数据之上是否存在信息利益

在既有数据确权研究之中,部分学者认为数据之上存在信息利益,这些学说或倾向于保护信息主体的利益,或注重信息主体与数据处理者的利益平衡。例如,在倍受推崇的数据信托理论[7]中,数据处理者基于信息主体的信任和委托对原始数据进行收集、分析、处理和加工。数据处理者的数据管理运用权限包括但不限于访问控制、访问审核以及数据的匿名化处置等重要内容,以此平衡信息主体的隐私保护与数据可交易价值之间的紧张与冲突[8]。鉴于信托理论根植于信息主体和数据处理者之间达成的数据信托关系,该理论在事实上默认信息主体为原始数据的权利人,这也意味着,信息主体在原始数据之上享有绝对性、完整性的信息利益。相较于数据信托理论对信息利益的完全接纳,双方赋权学说着眼于信息利益与数据开发利益的衡平,对数据之上的信息利益的范围进行深入研究。有学者提出,对于用户,同时配置人格权益和财产权益;对于数据经营者(企业),应分别配置数据经营权和数据资产权[9]。就信息利益与数据开发利益的位阶而言,“个人数据赋权对数据控制者与处理者施加责任,要求数据控制者与处理者满足个人的一系列数据权利,承担维护个人数据安全与数据质量等责任”[10]。由此观之,双方赋权说也认为数据之上存在信息利益,数据开发利益应当以信息利益的维护为前提。

然而,并非所有学者都认为数据之上存在信息利益。持此种观点的学者认为,数据确权的意义在于发挥数据的利用价值,他们将研究重心聚焦在数据处理者的数据开发利益之上。例如,有学者提出,“在现有理论与实践基础上,应搁置相应企业数据产权争议,并设立企业数据的利用与分享准则”[11]。甚至有学者认为,应在数据之上创建具备有限排他性的准财产权,使数据控制者得以据此对抗特定类别主体和特定类别行为[12]。

通观上述学者的观点,认为原始数据之上存在信息利益的学者,倾向于将信息与原始数据等同,进而认为信息主体对于自身信息享有的权益可以完全追及到原始数据之上;认为原始数据之上部分存在或不存在信息利益的学者侧重于将对信息与原始数据进行区分,认为信息主体对于自身信息享有的权益不能追及或部分到原始数据之上。鉴此,破除原始数据确权之迷思,在根本上需要明确信息与原始数据之间的真正关系。

4. 数据确权之桎梏:当前确权学说的固有缺陷

数据处理者对原始数据进行处理并筛除冗余数据,使得产出的次生数据具有极高的经济价值和资产属性。因此,对于次生数据的确权意义非常重大,不仅关系到次生数据的资产化,也关系到数据交易和数据要素的市场资源配置。由于鲜有学者明确提出将数据确权问题区分为二次确权进行研究,导致确权问题涉及的利益和价值复杂化,不便于问题分析和解决。好在部分学者侧重于对次生数据的确权,即便其中不免会掺杂着原始数据的确权问题,但这些学者在提及到次生数据确权的进路和思考却非常值得思量,代表性的观点主要有:算法说、知识产权说、场景化说和权利结构说。这些学说代表了当前确权研究的主要进路和方向,但也存在着不妥之处。

4.1. 算法说之局限

算法说认为,应当通过算法规制来明确数据权利归属。具体而言,原始数据到次生数据的嬗变需要

依靠算法, 同时算法也关系着次生数据的合法性, 既然算法决定着次生数据的价值性与合法性, 那么自然就可以认为算法对次生数据起着决定性的作用。鉴于算法是由数据处理者投入的, 数据处理者应当成为次生数据的权利人。有学者进一步认为, “即便是匿名的非个人数据在数据采集、数据处理时仍可能涉及个人信息保护问题, 在数据清洗、数据还原时亦可能显现出具有识别性的信息。因此, 需以用户为中心规则塑造算法规制的具体规则” [13]。

诚然, 算法确实对于次生数据具有重要作用, 通过对算法的规制也确实可以实现对次生数据的规制。但是将数据权利依附于算法之上, 即使存在法律对算法的规制, 数据处理者作为算法的设计者和使用者仍然会尽可能的将算法的权利扩张而形成“算法霸权” [14]。在事实上, 数据权利应当独立于算法权利。譬如, 确定一桶水的权利归属, 应当去寻找打这桶水的人, 即付出打水劳动的人, 而不是去寻找这个水桶所有人。算法在数据处理的关键作用不能忽视, 但是也应当承认, 它在数据处理的过程之中始终扮演着“桶”的角色, 具有很强的工具属性。所以不能因为数据处理者是算法的设计者和使用者, 就理所当然地认为, 通过算法这一工具挖掘出来的次生数据归属于数据处理者所有。可以明确的是, 对算法的规制可以更好地规范数据处理行为, 但这仅局限于其本身的工具属性。因此, 对算法的设计和应用进行规制, 并不能在根本上作为数据确权的理论支撑。

4.2. 知识产权说之错位

相较于算法说对技术层面的关注, 知识产权说更侧重研究数据处理过程中的创造性。该学说认为, 按照著作权法的规定, 汇编作品是在选择或者编排中体现独创性的作品集合或其它信息集合。当数据处理的结果, 即次生数据表现为具有独创性的数据库数据集, 那么数据处理者就享有著作权。当次生数据表现为不具有独创性的数据库数据集时, 虽然不满足汇编作品的独创性, 但是其收集、选择和编排过程中付出了一定的劳动 [15], 数据处理者也可以对其享有邻接权。当次生数据没有形成数据库数据集的, 次生数据拥有者对数据享有财产权。

按照知识产权说的观点, 具有独创性的数据库可以被纳入汇编作品的范畴而受到著作权法保护。客观上, 相较于数据库的独创性来讲, 数据的真正价值更多的体现在数据本身而并非其独创性的编排方式与结构, 数据本身才是数据处理的真正动机, 而数据库的编排方式与结构归根到底只是为了更好地进行数据处理, 该学说基于数据库的独创性而将其归入汇编作品中, 此时著作权法保护的也只是该数据库的结构独特性, 而并非数据本身。侵权人如果将数据库的编排方式加以调整使其丧失独创性, 就可以轻易规避侵权责任, 从而使著作权法的保护失去意义 [16]。另外, 相较于独创性的数据而言, 非独创性的数据更为常见, 该学说认为非独创性的数据库可以采用邻接权的方式进行保护, 根据《著作权法》规定, 邻接权特指表演者对其表演活动, 录音录像制作者对其制作的录音录像, 广播组织对其播出的广播信号, 以及出版者对其版式设计所享有的专有权利。当前《著作权法》对邻接权的类型只限于以上四种情形, 根据权利法定原则, 如果数据库按邻接权的方式保护, 必然需要通过立法将数据库创设为邻接权的权利客体。邻接权作为广义著作权的一部分, 对其权利类型的扩张必然也会引起整个著作权法内部体系和理论逻辑的嬗变。最后, 该学说认为数据处理者对不能形成数据库数据集的数据享有财产权。由此可见, 该学说是基于次生数据作为数据处理者的财产这一基本观点而来, 即次生数据不以数据库形态呈现时, 作为数据处理者的财产以财产权的方式保护; 以数据库的形式表现出来的, 可以通过邻接权或者著作权进行保护。笔者对次生数据以财产权的方式进行保护并无异议, 但是基于数据是否为数据库亦或者数据库是否具有独创性, 而决定采用著作权或者邻接权进行保护的观点似乎偏离了数据确权的方向。客观上讲, 该观点的确权研究对象并非数据, 而是数据库, 无论是著作权或者是邻接权, 保护的主体也并非数据本身, 而是数据的编排形式。

4.3. 场景化说之宽泛

场景化学说并未直接将数据确权作为研究数据产权的目标,而是以场景化的方式聚焦在数据的各方利益之上,通过对数据之上利益的区分而实现对数据产权的界定。支持该观点的学者认为,场景化方式给予了数据上利益各方在不同场景下充分实现利益诉求的机会,不武断地排除或者忽略某一方对数据享有的合理利益诉求,给予了个人、企业、行业、产业、社会、国家甚至整个人类命运共同体等多元利益主体在具体场景中细化权益诉求的公平机会,提高了各方参与数据治理的积极性[17]。

场景化思考的方式确实有助于数据确权问题的研究。如果采取“一刀切”的方式对数据一概而论,反而会导致研究无法具体化。原始数据与次生数据之上本就存在着不同的利益主体,各方利益主体也有不同的利益诉求,“具体问题具体分析”是场景化学说在研究数据确权问题之上最值得研究者借鉴之处。但凡事过犹不及,该学说认为需要针对不同数据特性、不同主体需求、不同数据利用场景采取分步骤、分阶段、分类分级的方法进行研究,这难免会导致数据确权复杂化和成本过高,也不利于数据确权问题的解决。应当承认的是,场景化的规则设定固然是纠纷化解之道,但由于没有普遍的秩序立场和理论框架可依托,相关规则不具有应对未来问题的拓展空间,也因此戕害了法的安定性[18]。

4.4. 权利结构说之缺憾

总体上来讲,权利结构说是在权利构造的层面对数据的确权问题进行研究和探索,但具体到细分层面,又基于所持的权利构造观点的不同,分化出“权利球”、“权利束”和“权利块”三种截然不同的数据权利形态观点。“权利球”是基于传统民法的“主客二元”理论,将数据权利构造为类似于具有弹力的球状结构,在权利受到侵害或存在侵害危险之时,数据权利人可以通过请求停止侵害、消除危险、排除妨害、赔偿损失等方式,使得“权利球”臻于圆满;“权利束”则是进一步地突破了“权利球”在“客体”之上的局限,着眼于权利的对人性,并基于主体或者利益的多元将数据之上的权利进行分解,随即将一个个彼此独立的权利木棍扎成一捆权利束,这样的数据结构形态相比“权利球”形态更具有灵活性和相对性。有学者认为,数据权利中表现出的“束点”“边界”的权利构成,利益性、享有性与排他性的价值内涵都与“权利束”概念相契合[19],鉴此,他试图以数据权利束为切入点数据权利进行研究;“权利块”虽然也是以“权利束”的相对性为前提条件,但相较于“权利束”对权利分解的无限制性,“权利块”对权利的拆解更强调以独立的功能为导向,将权利分解为一个个相互独立的模块权利,即权利结构的模块化,最后这些模块化权利通过标准的“接口”组合成一个完整的权利结构形态。亨利·史密斯用乐高积木对此做了形象的比喻:“一旦人们精通了连接积木的接口,就不需要去担心个别积木的形状和颜色,以及它们为何如此。”[20]对于数据权利块的构造,有学者基于数据的公私二元利益,认为数据权利可分解为服务公共利益目的的“公共数据(Public Data)权利模块”和以私主体利益为依归的“私人数据(Private Data)权利模块”。

通观权利结构说中的三种观点,不难发现“权利球”结构形态执迷于客体的“物性”,侧重主体对客体的支配关系,这显然与数据本身的无体性相背离,在事实上,数据的价值在于其天然的流动性,过度强化主体对数据的支配反而会掣肘数据的流动。因此,传统物权法的“权利球”结构与数据权利的性质缺乏内在契合性,无法为数据权利结构提供理论性的支撑。相较于“权利球”而言,“权利束”突破了“客体”的藩篱,更突出强调主体之间的关系,这种数据结构模式似乎能够承载数据之上多元的主体和复杂的利益。然而,高度的灵活性会导致“权利束”分解的无限制性和极度开放化,权利在过度肢解的同时也意味着其自身内容的空洞化,难免有损数据法律体系的稳定性。“权利块”通过权利模块化的方式对权利的分解进行限制,既可以满足数据,尤其是次生数据流通交易的诉求;又能避免“权利束”的不确定性。但需要指出的是,当前对“权利块”的数据权利研究尚处于公私二元相对宏观的层面,充

分发挥次生数据的流动价值, 还需要进一步深入到私法领域探求数据权利的建构。

5. 数据确权困境之破除

5.1. 原始数据与信息的区分

信息在经历第一次飞跃之后意味着原始数据的诞生。部分学者认为, 在这次飞跃中, 原始数据仅仅是转变了形式的信息, 由此得出数据与信息等同的结论。不可否认的是, 原始数据确实是信息的一种载体, 也恰恰是因为原始数据之中包含着信息而具有极大的可利用价值。但值得注意的是, 基于原始数据包含着信息这一客观事实, 并不能当然地认为原始数据等同于信息。原因在于, 原始数据与信息在法律属性、控制效力两个方面存在着明显分殊。

5.1.1. 法律属性的区分

信息具有人格属性。从私法意义上讲, 信息主体包括自然人和法人。与自然人相关的信息会涉及到隐私权和个人信息权益, 隐私权和个人信息权益属于人格权的范畴, 本质上是一种人格利益; 与法人、非法人组织相关的信息会涉及商业秘密、商业信誉、非公开性质的经营信息等, 无论是商业秘密还是商业信誉亦或是非公开性质的经营信息在实质上都属于法人的具有防御性质的人格利益, 一旦这种信息遭到泄漏或非法使用将会严重侵害法人组织的财产利益。

原始数据兼具财产属性与人格属性。当具有人格属性的信息被转化为原始数据之后, 人格属性仍然存在, 只是由于数据处理者的劳动和贡献进一步地增添了财产属性。数据时代的标志是数据成为社会基础资源、经济活动的要素[21]。数据处理是实现数据要素化的关键。数据处理者通过投入技术、劳动、资金等生产要素, 以收集、汇总、存储等方式将信息转化为原始数据。因此, 数据处理者是原始数据的生产者, 也是数据要素化的贡献者。原始数据也由于数据处理者的劳动和贡献具有财产价值。

5.1.2. 控制效力的区分

信息主体对信息享有绝对的控制力。信息主体基于人格尊严和自由对于自身的信息享有自决权, 有权决定自己的信息的处理主体、处理方式和处理范围。数据处理者必须通过告知的方式将数据处理的告知, 并取得信息主体的同意, 才能够进行数据处理。同时, 信息主体也可以随时不受任何限制的撤回自己的同意, 以终止数据处理。因此, 信息主体对于自身的信息享有绝对的控制力, 在不违反法律地强制性规定的前提下, 其对于信息的控制不受到任何主体的干涉和限制。

信息主体对原始数据具有相对控制力。如前所述, 从信息转变为原始数据, 实质上是在人格属性的基础上添附了财产属性。信息主体对于自身信息享有的控制权延及到原始数据之上, 必然要尊重数据处理者的劳动和贡献。一方面, 相对的控制力在于互惠。信息主体和数据控制者之间的关系类似于“花与蜂”的互惠关系, 作为“花”的信息主体通过扮演“蜂”的数据控制者所搭建的数字通讯网络进行交流、生产、生活, “蜂”在传递“花粉”——信息的过程中, 需要在不伤害“花”的基础之上, 对“花粉”进行加工酿成高价值的“蜜”——次生数据。另一方面, 相对的控制力在于防御。信息主体对于信息的控制力延伸到原始数据之上, 其对于原始数据的控制在于防御数据控制者的非法利用。当数据控制者在处理原始数据时未尽到安全保障义务或是故意篡改数据导致信息主体的信息利益受损时, 信息主体依然能够基于自身享有的控制力, 请求数据控制者恢复原状, 赔偿损害。

5.1.3. 小结

厘定原始数据与信息的关系, 是数据确权研究中的重要命题, 是信息主体与数据处理者之间权利(或权益)划分的理论依据, 也是原始数据确权的理论基石。如果将数据与信息混同, 信息主体享有的信息权

益必然会完整地追及到对应的原始数据之上。此时信息主体对原始数据享有的权益与其自身的信息权益在范围和性质上并无二致, 具有相同的绝对性和人格性。即信息主体对原始数据享有人格权益, 并且这种权益具有绝对性。在此背景下, 信息主体的同意许可必然会成为数据主体进行数据处理的合法性基础。然而, 这最终会落入到与 GDPR 相同的困境: “宽泛合法性基础未实现个人信息流通” [22]。对此有学者指出, “必须通过某种方式, 使数据的财产权属性与人格权属性相分离, 而这种方式, 正是数据确权所必需的技术因素” [23]。然而, 厘定信息与原始数据混同的方式并非在于技术, 而在于认知。一直以来, 我们深深地陷入了对原始数据确权的迷思之中不能自拔。明确数据与信息区分, 是跳脱出这一迷思的关键。综上所述, 原始数据与信息是不能完全等同的, 应当坚持和尊重彼此的法律属性和控制效力。

5.2. 数据确权桎梏之破除: 重塑数据权利结构

5.2.1. 数据确权目标: 数据主体之间的权利归属

数据主体之间的权利归属, 是数据确权的目標。从整个数据的整个生命周期来看, 数据权利会随着“信息→原始数据→次生数据”的变化而发生调整。数据处理的第一步就是对信息进行收集。信息经过合法收集而成为原始数据, 数据处理者对合法收集的原始数据享有数据处理的排他权, 依据排他权而享有对原始数据的处理权益, 并可以排除其他数据处理者的妨害。在数据处理过程中必然需要依据《数据安全法》第八条的规定, 最终才能得到合法的次生数据。在这种情形下, 数据处理者作为次生数据的合法生产者, 应当对次生数据享有财产权益。原因在于, 次生数据是由数据处理者对享有排他权的原始数据进行使用、加工、传输、提供、公开等一系列处理措施而生成的数据, 尤其在数据的处理环节, 数据处理者需要投入大量的人力、物力和技术手段才能保证得到的次生数据在质量、价值、法律等层面符合市场或业务需求。次生数据包括数据资源和数据产品。数据资源, 是数据控制者以特定的目的, 对原始数据进行加工, 生产出具有特定利用价值的原始数据集。数据产品, 是数据使用者对数据资源进行利用, 生产出相应的算法模型、决策支持等数据产品。如果将原始数据比作原料, 数据处理者通过劳动投入将原始数据加工而成的产品就是数据资源。数据使用者通过对数据资源进行分析使用得到的产品就是数据产品。

在“信息→原始数据”这一阶段, 应当聚焦于信息主体与数据控制者的权利归属。由于原始数据与信息存在着区分关系, 这意味着信息主体的权益应当仅限于对信息的控制性和完整性保护, 并不能主动地追及到原始数据之上, 其对于原始数据只享有被动防御的权利; 数据处理者对于原始数据的处理权利属于财产性权利, 该权利来源于信息主体的同意。但由于原始数据具有溯源性, 可以被还原成信息, 这意味着数据处理者也必然需要承担保障原始数据安全, 防止数据泄露而损害信息主体的信息权益的义务。

在“原始数据→次生数据”这一阶段, 应当聚焦于数据处理者内部的权利归属。数据处理者必须依据相应的法律法规与数据安全能力建设框架对原始数据进行处理, 使得次生数据所承载的信息达到去标识化和匿名化的标准, 即次生数据无法回溯到特定的信息主体, 这也意味着数据处理者必须将原始数据之上附着的信息利益予以剔除, 此时得到的次生数据之上也仅存在数据处理者内部的权利纷争。易言之, 通过对原始数据进行合法地数据处理而产生的次生数据, 在理论上不应当涉及信息权益。数据处理者对原始数据进行处理并筛除冗余数据, 使得产出的次生数据具有极高的经济价值和资产属性。因此, 对于次生数据的确权意义非常重大, 不仅仅关系到次生数据的资产化, 也关系到数据交易和数据要素的市场资源配置。在这一阶段中应当聚焦于数据处理的内部过程(即原始数据的使用、加工、传输、提供、公开等各个环节), 将数据处理者之间相关权利的厘定作为该阶段数据确权的目標。

5.2.2. 数据确权路径: 构建数据权利结构

纵观学界对于数据确权的观点, 可以发现, 除了数据结构说聚焦于研究数据的本体, 其他的学说似

乎更集中于借助其他因素研究数据确权问题。具体而言, 算法说将算法视为确权的决定因素, 知识产权说将数据的编排形式作为确权的核​​心, 场景化说着重于对场景中的利益关系进行厘定, 但仍然是将场景划分作为确权的前置条件。一言以蔽之, 工具层面的数据确权研究倾向于将数据本体的外部要素作为数据确权的依据。客观的讲, 这种研究范式在很大程度上丰富了数据确权探究的视野。但就研究态势而言, 这类研究进路意图借助其他因素确权以规避数据作为无形财产确权的本质问题, 不可避免的会游离在数据确权的核​​心要素——数据本体之外。“不是从现有制度框架和理论逻辑出发来论证‘数据/信息确权’的正当性、合法性, 而应基于人的生物/数字双重属性、虚实同构的行为规律、算法决策的秩序机制等来重新界定权利形态、阐释权利理论” [24]。因此, 就研究进路而言, 将数据权利本体构造作为研究核​​心的权利结构说对数据的确权问题更具有研究前景和理论价值。

综上所述, 数据权利结构的构建应当分为两个阶段。首先, 对于“信息→原始数据”阶段的权利进行配置。在这一阶段中, 信息主体与数据控制者对原始数据享有相对控制权。其中, 信息主体享有对原始数据的防御性控制权, 该权利在于防范数据控制者对于自身信息的滥用。数据控制者对原始数据享有利用性控制权, 该权利在于激励数据控制者对于原始数据的开发利用, 提升数据价值、释放数据红利; 其次, 对于“原始数据→次生数据”阶段进行权利配置。这一阶段的权利构建应当以数据控制者享有的利用性控制权为起点, 依次分置为原始数据加工权与数据资源使用权两项基本权利。原始数据的加工在于生成数据资源。数据控制者根据特定的目的和需求, 对原始数据进行分类、筛选、汇集, 形成具有特定利用价值的原始数据集合, 即数据资源。原始数据的使用在于数据控制者许可他人(使用者)对自己生产的数据资源进行利用, 数据使用者通过对数据资源进行分析利用生产出数据产品。数据使用者作为数据产品的生产者, 享有对数据产品的所有权。

6. 结语

数据确权面临着双重困境, 意味着很难从整体上对数据权利进行配置。为纾解这一难题, 笔者基于数据的生命周期, 提出“二次确权”的分析框架, 聚焦于数据的具体形态之上, 分别对原始数据和次生数据的权利进行分别配置。原始数据的确权, 应当以信息与原始数据的区分作为基点, 明确信息主体与数据控制者对于原始数据享有的相对控制权。其中, 信息主体享有对原始数据的防御性控制权, 数据控制者对原始数据享有利用性控制权; 次生数据的确权, 应当以数据控制者享有的利用性控制权为起点, 依次分置为原始数据加工权与原始数据使用权, 以原始数据的加工使用来判定次生数据的归属。

基金项目

本文系浙江省教育厅科研项目“基于‘二次确权’的数据产权法律制度构建研究”(Y202250136)研究成果。

参考文献

- [1] 维克托·迈尔-舍恩伯格, 肯尼思·库克耶. 大数据时代: 生活、工作与思维的大变革[M]. 盛杨燕, 周涛, 译. 杭州: 浙江人民出版社, 2013: 104.
- [2] Information Technology—Metadata Registries (MDR), ISO/IEC 11179-1: 2015. <https://www.iso.org/obp/ui/#iso:std:iso-iec:11179:-1:ed-3:v1:en>
- [3] 李爱君. 数据要素市场培育法律制度构建[J]. 法学杂志, 2021, 42(9): 17-28.
- [4] 王雪乔. 论欧盟 GDPR 中个人数据保护与“同意”细分[J]. 政法论丛, 2019(4): 136-146.
- [5] 何渊. 数据法学[M]. 北京: 北京大学出版社, 2020: 101.
- [6] 库勒. 欧洲数据保护法: 公司遵守与管制[M]. 旷野, 杨会永, 等, 译. 北京: 法律出版社, 2008.

- [7] 翟志勇. 论数据信托: 一种数据治理的新方案[J]. 东方法学, 2021(4): 61-76.
- [8] 席月民. 数据信托的功能与制度建构[J]. 民主与法制, 2021(3): 54-55.
- [9] 龙卫球. 数据新型财产权构建及其体系研究[J]. 政法论坛, 2017, 35(4): 63-77.
- [10] 丁晓东. 论算法的法律规制[J]. 中国社会科学, 2020(12): 138-159+203.
- [11] 姚佳. 企业数据的利用准则[J]. 清华法学, 2019, 13(3): 114-125.
- [12] 杨翱宇. 数据财产权益的私法规范路径[J]. 法律科学(西北政法大学学报), 2020, 38(2): 65-78.
- [13] 韩旭至. 数据确权的困境及破解之道[J]. 东方法学, 2020(1): 97-107.
- [14] 魏远山. 算法解释请求权及其权利范畴研究[J]. 甘肃政法学院学报, 2020(1): 143-156.
- [15] 王渊, 黄道丽, 杨松儒. 数据权的权利性质及其归属研究[J]. 科学管理研究, 2017, 35(5): 37-40+55.
- [16] 郝思洋. 知识产权视角下数据财产的制度选项[J]. 知识产权, 2019(9): 45-60.
- [17] 秦天雄. 对场景化界定数据产权的思考[J]. 上海法学研究, 2021, 5(1): 1-13.
- [18] 许可. 数据权利: 范式统合与规范分殊[J]. 政法论坛, 2021, 39(4): 86-96.
- [19] 闫立东. 以“权利束”视角探究数据权利[J]. 东方法学, 2019(2): 57-67.
- [20] Smith, H.E. (2012) Property as the Law of Things. *Harvard Law Review*, **125**, 1691-1726.
- [21] 高富平. 数据生产理论——数据资源权利配置的基础理论[J]. 交大法学, 2019(4): 5-19.
- [22] 高富平. GDPR 的制度缺陷及其对我国《个人信息保护法》实施的警示[J]. 法治研究, 2022(3): 17-30.
- [23] 陈肇新. 要素驱动的数据确权之法理证成[J]. 上海政法学院学报(法治论丛), 2021, 36(4): 131-145.
- [24] 马长山. 数字法学的理论表达[J]. 中国法学, 2022(3): 119-144.