

# 算法维度下非个人数据确权的反向实现

韩旭至

**【内容摘要】** 正向确立以非个人数据为客体的数据权利存在极大的理论争议，非个人数据确权困境现实存在，数据利用所产生的社会问题亦日益突出。通过规制手段反向考察数据权利尤为必要。对数据价值生成与数据权利生产进行分析可知，算法是其中的核心要素。一方面，在数据价值的形成中，算法始终处于中心地位，使从低价值的原始数据到高价值的衍生数据的转变成为可能。另一方面，伴随着数据价值的增加，数据制造者亦提出了相应的权利诉求。在应对这些诉求的过程中，算法规制的内容为数据行为划定了边界。由此可见，运用算法规制在正向实现数据正义的同时可反向实现数据确权。具体而言，应以合法性为中心构建算法规制原则，保护合法的数据；以用户为中心设置具体的算法规制规则，塑造数据权利的边界。

**【关键词】** 数据 非个人数据 数据权利 算法规制 信息社会

**【作者】** 韩旭至，华东政法大学法律学院特聘副研究员、师资博士后。（上海 201620）

**【基金项目】** 中国法学会青年调研项目“信息社会新型权利的生产机制及其法律规制”（CLS2018Y07）；中国博士后科学基金第65批面向资助项目（2019M651448）；华东政法大学科学研究项目“实现数字正义的算法规制路径”（19HZK002）

## 引 言

当前，被称为“新石油”的数据已显现出极大的财产价值。数据成为平台企业与人工智能产业的生产资料。数据交易开始日益频繁，众多数据交易所亦开始涌现，甚至出现了数据入股、并购数据企业等事例。明确作为交易客体的非个人数据<sup>①</sup>属性成为迫切需要解决的现实问题。

然而，关于数据权属众说纷纭。存在公有说、合同说、物权说、知识产权说、商业秘密说、不正当竞争一般条款说、数据库权利说、混合权利说、新型权利说等多种理论。尤其是，司法实





践习惯采取激活《反不正当竞争法》第2条释义性条款的特殊路径,以实现企业数据保护。然而,这些理论与实际均未能从数据价值与数据新型权利生产机制的角度对数据确权进行解读。尤其是依附于既有权利框架的“旧瓶装新酒”式理论模型,实为“头痛医头脚痛医脚”,且各个解释之间往往又发生新的矛盾。即便是试图正向实现数据赋权的新型权利理论,亦脱离了数据权利的生产机制,其不仅难以确定基本的权利主体和客体,更面临理论资源匮乏的困境,以致极难形成通说。

从法经济学上的规则菜单可见,对某一客体进行确权不仅可正向进行还可反向实现。即除明确赋予拥有者定价权的财产规则外,法律还可以选择根据公共利益施加具体标准的管制规则,或设定“买断”或“卖断”价格的责任规制,甚至是设定禁止转让的禁易规则。同时,数据不仅涉及私主体私益,更涉及他人其他权益、社会公共利益的保护。数据采集、存储、加工、利用更可能产生负外部性效益。大数据时代所特有的“数据孤岛”“数据鸿沟”“数据垄断”“数字霸权”“算法黑箱”“算法歧视”等现象已引起社会的广泛关注。由此可见,通过规制的手段反向实现数据确权既有可能亦有必要。

诚如哈贝马斯所指,“权利毕竟既不是一支枪,也不是一台独角戏。它是一种关系、一种社会惯例”<sup>②</sup>。反向实现的数据确权亦必须回到新型权利的社会关系中进行考察,而非从学者自身朴素价值观出发,进行天马行空的创造。数据确权的研究应是一个权利发现的过程,在价值合理性的基础上证成权利合理性。通过对数据价值生成与数据权利生产的分析可知,智能算法是其中最为核心的因素,同时也是网络时代重要的规制对象。因此,反向实现数据确权应依算法规制模式展开。

## 算法构成数据价值的中心

在数据价值的形成中,算法始终处于中心地位。虽然生活中的各类数据长期存在,但仅在智能算法出现之后,数据被抓取、处理、提炼才产生的基于数据价值的数字经济。也正因为数据价值的产生,才有了数据确权的需求。事实上,无论是原始数据的抓取、海量数据的处理,还是衍生数据的生成,均有赖于背后的智能算法。正是算法使从低价值的原始数据到高价值的衍生数据的转变成为可能。“淘宝诉美景案”<sup>③</sup>中一审法院即指出,单个原始数据的确不具有财产性权益,但网络经营者对经算法处理形成的大数据产品具有财产性权益。具体而言,算法起到架构设计的作用,决定了数据的价值。

### (一) 以算法作为架构设计

20年前,莱斯格提出了“法律—社群规范—市场—架构”的经典网络规制框架。在这四个规制要素之中,他尤其指出代码是架构的相似物,且重申“代码即法律”。<sup>④</sup>此处的代码应属作为算法的代码,而非代码层意义上的数据。用以描述通过特定计算过程实现特定的输入/输出关系的算法,符合作为架构的代码所强调内在结构的设计与制造的属性。

事实上,算法与数据之间的紧密关系早已引起了广泛关注。大数据处理算法与数据交易算法即被认为是赛博经济系统中两种重要算法。人工智能算法服务,亦是大数据交易平台的主要产品之一。算力、算法、数据、产品与服务更被认为是人工智能应用的四大层面。以传感器与芯片为基础的硬件设备确保了算力与数据获取,而算法则在软件层面上决定了数据产品的最终结构。随着硬件的发展,算力得到了极大提升,依赖于数据的产品和服务更多受到算法的影响。深度学习、

神经网络算法、遗传算法等多种算法的发展,便使得数据应用与数字经济得以长足发展。

值得注意的是,有学者认为网络架构不仅包括算法。如胡凌认为,结合物理层、代码层、内容层的区分,以及“账户—数据—评分”的框架,算法只是赛博空间架构内的一个组成部分,而平台企业、基础服务、资源、网络、数据、算法又共同形成了平台经济的核心要素。<sup>⑤</sup>然而,具体到数据价值的视野来看,物理性的基础要件与外部条件只是数据价值生成的必要不充分条件,有相关硬件并不必然能产生数据价值。而作为架构的算法通过直接作用于代码层以决定了内容层的构造。其不仅作用于数据抓取后的匿名化与数据分析等处理环节,而且直接通过数据抓取规则的设定作用于包括个人信息与其他数据在内的全样本数据获取,以及通过应用性规则生成大数据与数据分析结果。因此,算法是架构的最佳对应物。

## (二) 以算法实现数据价值

作为架构的算法不仅与数据市场共同作为规制数据的规制要素,在一定意义上算法还决定了数字经济。算法贯彻数字经济的全部流程,决定了数据的资源的获取、处理与应用,因此数字经济即被称为“算法定义经济”<sup>⑥</sup>。互联网经济的实质亦在于通过算法进行的数据分析实现了各类生产要素的精确配置。这一精准配置过程,即数据价值的实现与利用过程。

通过算法的处理,数据的价值密度实现了从低到高的转变。尤其是算法处理形成的大数据已具有了“4V”(容量大、类型多、存取速度快、应用价值高)的特性,其价值得到进一步放大。具体而言,通过数据挖掘、预测性分析等特定算法的处理,原始数据转变为衍生数据,可通过定性营销、信用评价、变量分析等直接作用于相关市场行为,带来巨额利润。如上海数据交易中心即通过“营销应用与征信应用”两大应用板块,以实现数据衍生品的流通。可见,在数据市场中算法处于决定性地位,“‘无形的手’已经被一只‘数字化的手’所取代”<sup>⑦</sup>。

值得注意的是,实践中常脱离算法对数据的价值进行判断,机械地将数据价值限于大数据之中。目前我国法院仅承认大数据的财产价值。我国《促进大数据发展行动纲要》《贵州省大数据发展应用促进条例》等文件更是采取了大数据的“4V”定义,直接承认了大数据的价值。经济合作与发展组织亦将大数据列为企业的核心资产。无疑,经过数据应用算法处理得出的大数据具有巨大价值,但不能因此否定“小数据”的价值。大与小是相对的概念,数据权利的客体更不应以价值大小作为限制,否则将变成对算法技术以及数据规模的无益审查。只要经过算法处理的非个人数据,不管是处于何种处理阶段,均具有一定的价值,是数据权利的客体。

此外,由于算法通过代码层决定内容层,应明确数据价值亦体现在该两个层次之中。可以肯定的是,数据价值并不以物理层为体现。数据的物理层强调的是数据的底层物理支撑,如硬件设备或云端存储。而数据与书面文字不同,其能轻易实现与物理设备的剥离,数据交易亦非通过交付底层物理设备进行。于代码层与内容层的区分而言,其于数据价值的判断并无益处。虽然,根据国际标准化组织的定义,信息与数据是特定知识内容与表现形式的关系。“淘宝诉美景案”一审法院即采取了这一定义。纪海龙更将内容层上的数据称为数据信息,将数据权利客体限于代码层上的数据文件。然而,一串特定代码固然是数据,但亦可能展现特定的内容,且往往内容的展现才是数据的价值之所在。大数据、数据分析报告的巨大价值,并不在于其代码优势,而在于能利用其内容指导市场活动。以“大众点评诉百度案”<sup>⑧</sup>与“大众点评诉爱帮案”<sup>⑨</sup>为例,案争的点评数据价值亦在于相关内容。信息与数据的区别只说明其对应的不同规制手段以及可能存在责任竞合,不可因此否定数据内容的价值及其权利客体地位。



## 算法奠定数据权利的基础

由数据权利生产机制可见,从价值生成到权利诉求再到权利维护,算法均在其中起关键作用。以企业为代表的制造者通过算法进行数据挖掘和分析,投入了大量成本并产出了巨大价值,由此产生了数据保护诉求。不难判断数据权利的主体应是数据制造者。为了维护自身利益,这些制造者已经运用算法防止数据未经授权被他人利用。同时,域外相关法规制的内容亦为数据抓取和利用划定了边界。我国后发的法规制立法,正是数据权利危机的深层原因。

### (一) 以算法逻辑判断劳动赋权

由于忽视算法在数据价值与数据权利中的基础性地位,关于何为数据劳动以及数据权利的主体问题存在争议。一方面,有学者根据洛克的劳动赋权理论指出,由于以企业为代表的制造者投入相关成本,因此制造者应获得相应权利。<sup>⑩</sup>也有学者从法经济学角度论证应将权利赋予制造者。如纪海龙认为,应将数据权利“分配给可让其发挥最大效用的人”,而将产权赋予制造者恰能激励创造数据。<sup>⑪</sup>另一方面,有学者认为数据源于个人信息采集,理应将产权赋予用户。卡尔即以“现代的佃农制度”对现代数据业态提出严厉批判。<sup>⑫</sup>也有学者主张“数据作为劳动”,认为用户才是数据生产中的劳动者。<sup>⑬</sup>

从上文所分析的数据价值生成路径可知,数据价值生成的中心在于制造者的算法投入。个人信息只是与其他数据共同构成了原始数据的原材料,通过挖掘算法采集后方构成原始数据。当且仅当原始数据中的个人信息被匿名化处理,产生的非个人数据方为数据确权对象。虽然个人信息保护制度约束着数据价值生成与利用的全过程,但当合法遵循相关规范,并将相关信息转化为非个人数据时,数据便具有了独立的价值。由此可见,信息主体从未作为生产者存在。是否获得信息主体知情同意、有无遵循数据最小化原则、是否符合数据安全标准等,均是个人信息保护制度考察的范畴。

从数据权利生产机制亦可见,正是以企业为代表的制造者对数据提出了相关权利诉求。而制造者所付出的成本使其对数据享有权益,则是数据权利诉求中的常见论点。这一论点已为司法实践所普遍接受。早在“阳光数据诉霸才数据案”<sup>⑭</sup>中,法院即接受了企业对数据的投资应受保护的逻辑。“大众点评诉爱帮案”中一审法院更是明确指出,“汉涛公司为此付出了人力、财力、物力和时间等经营成本,由此产生的利益应受法律保护”。“淘宝诉美景案”中一审法院进一步直接认定“数据产品系淘宝公司的劳动成果”。

然而,制造者的投入,尤其是通过算法使数据价值倍增的客观现实,只能说明制造者是数据权利诉求的提出者,而不能直接证成数据权利。诚如“大众点评诉百度案”二审法院所指,我国并不存在所谓“劳动成果权”,由于“模仿自由”,使用他人劳动成果也不能与“搭便车”和“不劳而获”画等号。数据权利必须在司法实践与制度框架中考察。

### (二) 以法规制生成权利空间

司法实践中,制造者通过验证、加密、协议等算法技术手段维护数据利益获得了法院支持。典型如网页运营者常通过 Robots 协议禁止或限制一定的数据抓取行为。在“百度诉 360 案”中,法院即认为 360 公司抓取相关数据“没有遵守百度网站的 Robots 协议,其行为明显不当,应当承担相应的不利后果”<sup>⑮</sup>。此外,网站针对特定算法的限制声明亦有助于维权。在多起案件中,美国法院均指出若禁止爬虫的网页声明足够明显,则不得对网页数据进行爬取。然而,相关算法

技术手段限制并不必然导致数据权利的产生。在“大众点评诉百度案”中，法院认为 Robots 协议只是相关行业准则，而不能据此判断数据抓取行为是否合法。在“hiQ 诉领英案”中，美国法院更是认为限制访问措施可能构成不正当竞争。<sup>⑧</sup>因此，归根到底仍需在制度中探寻数据权利空间。

如格伦顿所指，“就法律目的而言，一项权利只是一个‘预言’，它预示着公共强制将适用于那些违反它的人”<sup>⑨</sup>。霍菲尔德亦通过其著名的权利理论，指出了权利与无权利的相对关系及权利与义务的相关关系。<sup>⑩</sup>也就是说，通过对义务内容，尤其是公共强制的内容，可反向推出权利。而算法作为数据价值生成的中心，处于架构地位，无疑应作为公共强制的对象。当下，法律更被认为走向“死亡”并“被代码/算法取代”。<sup>⑪</sup>法律必须在注重算法规制的同时实现帕加罗所指的“法律作为元技术”<sup>⑫</sup>，即将制度设计深入到算法层面。

从域外的相关实践来看，虽然数据权利并未被普遍承认，但基于算法的规制使数据行为的边界得以明晰。当前，域外已就算法规制的整体制度提出了一定的方案。如2017年纽约市议会通过《算法责任法案》，建议成立工作组对公共部门所采用的算法进行评估，以判断是否根据年龄、种族、信仰、性别、残疾、性取向进行歧视。2019年，欧盟《可信人工智能伦理指引》强调人工智能算法的合法性及合伦理性，并提出了算法开发、部署和使用中应遵循“人类作用和接受监督，技术稳健性和安全性，隐私和数据治理，透明度，多样性、非歧视和公平，环境和社会福祉，问责制”7个关键要求。具体而言，域外算法规制手段更常见于数据采集、存储、加工中的个人信息保护、非个人数据利用以及相关经济监管几个层面。

第一，虽然数据权利客体限于非个人数据，但是由于原始数据很可能来源于个人信息，此时数据算法规制与个人信息保护存在交叉。一方面，算法运行的结果必须符合个人信息保护的规定及原则。如欧盟《一般数据保护条例》第5条所列举的合法、公平和透明原则、目的限制原则、数据最小化原则、数据准确原则、存储限制原则、完整性和保密性原则、责任原则，均是衡量算法结果的重要标准。另一方面，应将相关规范内化到算法设计之中，即“自设计保护隐私”（privacy by design）。欧盟《大数据时代个人信息处理指引》即指出，应根据风险评估程序，在算法设计中分阶段设计个人信息保护方案并运行相关测试。美国《健康保险流通与责任法案》关于“名称、地理区域、日期、电话号码、电子邮件地址等”多项识别符删除的匿名信息标准规定，亦为数据处理中的匿名算法设计订立了标准。

第二，随着个人信息保护范围的扩张，某些个人信息保护的规范亦被借鉴到非个人数据规制之中。随着《一般数据保护条例》而广为人知的算法解释权与数据可携带权对数据权利构建亦有重要影响。一方面，虽然《一般数据保护条例》只在序言部分提到对自动化决策的解释权，并未实质确立算法解释权，但在自动化决策中数据主体有权获得一个可理解的解释已为包括29条工作组与英国信息专员公署在内的权威机构所认可。虽然本质上算法解释权仍属个人信息保护的范畴，但其亦能直接约束涉及个人权益的衍生数据利用。另一方面，数据可携带权亦为欧盟《非个人数据自由流动框架》第6条所规定，通用标准格式的数据、数据迁移的便利性以及数据质量等要求已扩张到非个人数据之中。

第三，在经济监管层面，“算法合谋”“大数据杀熟”“算法交易”等现象已引起了监管部门重视。2015年美国司法部即注意到了部分互联网企业利用算法收集数据实现操纵市场价格的行为，并提出了相应指控。以“电子产品动态定价”进行的“大数据杀熟”亦引起了监管部门的注意。尤为值得注意的是，欧盟通过《金融工具市场指令2》明确对高频算法进行规制，要求建立有效的系

统和风险控制，并要求向监管部门披露相关算法交易系统的信息。

在上述算法规制框架下，若企业没有遵循相关规范，其不仅无法主张数据权利，而且需承担相应法律后果。如2014年“阅后即焚”公司即因未根据承诺及时销毁相关数据而被联邦贸易委员会处罚。与之相比，我国算法规制的制度资源明显不足，这正是我国数据权利困境的重要原因。虽然，我国《网络安全法》《电子商务法》《网络产品和服务安全审查办法（试行）》等规范亦涉及个人信息保护、关键信息基础设施安全、网络安全审查制度等内容，但均为具体制度层面上相对零散的规范，尚无直接针对算法规制的内容，只能在极其有限的意义上为数据确权服务。如“淘宝诉美景案”中法院即将数据合法作为数据权益的论证前提。“新浪诉脉脉案”<sup>④</sup>中法院更是结合个人信息保护相关规定，在第三方数据利用上确立了著名的“用户授权”+“平台授权”+“用户授权”的“三重授权原则”。但这些具体案例，只是通过激活《反不正当竞争法》第2条，以竞争法法益保护数据权益。其具有一定的不确定性，且无法回应非市场竞争下的数据权属争议问题，亦无法实现数据赋权。可见，唯有完善算法规制、明确数据运行空间，方可实现数据确权。

## 算法确立数据权利的规制框架

由算法在数据价值生成与数据权利生产中的核心地位可知，算法规制是反向实现数据确权的必由之路。在数据治理中，数据时代特有的数据利用负外部性问题，亦与算法问题密不可分。因此，在架构的意义上，运用算法这个“新的权力代理人”的权力，可在塑造数据权利的同时回应数据利用所带来的社会问题。具体而言，应以合法性为中心构建算法规制原则，保护合法的数据；以用户为中心设置具体的算法规制规则，塑造数据权利的边界。

### （一）以规制实现赋权：合法性原则的提出

算法规制可通过合法性原则实现数据赋权。即确认数据制造者对合法数据具有财产性权利，而数据的合法性又体现在算法主导的数据行为合法性之上。司法实践中，合法性亦是法院衡量的关键因素。虽然自“新浪诉脉脉案”以来，在数据权益纠纷中法院常将“合法、正当、必要的原则”并列作为判决标准。然而，正当与必要实际上亦是合法的判定因素。如“淘宝诉美景案”中一审法院即分别通过数据获取渠道是否合法，数据使用目的、方式和范围是否合法两个方面考察算法的正当性与必要性。基于数据行为合法与数据收益价值，“淘宝诉美景案”二审法院进一步明确指出，案涉数据产品“无疑属于竞争法意义上的财产权益”。

从数据自由流动的角度，赋予合法数据以权利既是数据自由流动的前提，也是数据自由流动的要求。一方面，当前的数据交易实践即以合法性作为数据流动的前提。如上海市数据交易中心《流通数据处理准则》即明确规定只保护合法的数据。另一方面，数据只有在流动中才能发挥最大价值。2018年欧盟出台《非个人数据自由流动框架》即明确指出，“数字经济以透明性和交互性原则为基础”，“除非根据比例原则以公共安全为正当事由”限制，数据应在欧盟内自由流动。

从历史的发展来看，正是互联网的开放共享的特性促进了网络产业的迅速发展。1984年乐维于《黑客》一书中即提出互联网无限制、不付费、去中心的观点。1999年奥莱理在《开源革命之声》中进一步提出了开放源代码的倡议。随后，以开放获取为中心的“网络中立性原则”亦被确立。虽然“互联网无政府主义”早已宣告破产，“网络不是法外之地”的观念也已深入人心，甚至2017年美国联邦通讯委员会更废除了“网络中立性原则”，使得网络服务提供商可以对数据流

通速度进行限制。然而，数据自由流动的价值从未被否定，甚至产生了以获取更多链接、追求更大数据的“数据主义”思潮。“大众点评诉百度案”中，二审法院亦明确提出“需要考虑产业发展和互联网环境所具有信息共享、互联互通的特点”。

即便是数据抓取算法与数据处理算法可能涉及的个人信息保护问题，亦以合法性为原则，并服务于数据自由流动。长期以来，源于隐私保护的知情同意被误认为是个人信息保护的首要原则。然而，在网络点击合同与浏览合同的框架内，用户缺乏菜单式选择，只能一揽子同意。由于使用的需要，用户又不得不选择同意，从而形成“强制同意”问题。事实上，知情同意只是个人信息保护的一种例外规定。从经济合作与发展组织《隐私保护和格式数据跨境流通指南》到欧盟《一般数据保护条例》，均未将知情同意作为一项原则进行规定。根据《一般数据保护条例》第6条的规定，数据主体的同意只是判断数据处理的合法性的情形之一，除此之外还存在为履行与数据主体之间的合同、为履行法定义务、为保护数据主体或第三人的重大利益、为履行涉及公共利益职责等情形。可见，合法性才是个人信息保护的首要原则。同时，根据《一般数据保护条例》第1条第3款的规定，数据自由流动亦是个人信息保护的目标之一。

因此，涉及个人信息的数据抓取算法与数据处理算法，亦不能简单以是否获得授权而对合法性作出判断。当前，“新浪诉脉脉案”所确立的“三重授权原则”虽为我国司法实践所广泛接受，却存在一定的缺陷。一方面，对于已取得用户同意但未取得企业同意的信息收集行为，是否按照该原则而具有违法性不无疑问。如2017年华为曾未经腾讯授权而通过用户点击同意的方式收集微信聊天记录以实现定向推送。另一方面，其亦未对用户同意的范围与权限做进一步限制。在德国“WhatsApp案”中，法院即认为WhatsApp获取用户联系人信息不仅需要用户的有效同意，更需要所有联系人的书面同意。<sup>②</sup>事实上，无论是“三重授权原则”还是“合法、正当、必要的原则”，归根结底均旨在考察算法是否合法。

值得注意的是，在合法性的判断中，必须排除技术中立的抗辩。在很长一段时间内，技术在伦理上被认为是价值中立的。1984年美国联邦最高法院通过著名的“索尼案”确立了“实质非侵权用途”标准，更被认为是技术中立的体现。<sup>③</sup>然而，至少在算法设计的意义上而言，从来没有一种算法是完全中立的，算法设计必定受数据制造者所追求的功能目标、价值偏好、商业利益所决定。同时，数据利用的负外部性亦说明，算法的运行确有侵害他人权益的可能。实际上，在互联网技术发展达到一定阶段之时，司法机关即将算法纳入调整对象。此时，价值中立只是一个价值分析的概念，而无损于法律责任的承担。2005年美国联邦最高法院即在著名的“Grokster案”中对“实质非侵权用途”标准进行补充。判决认为，在P2P文件分享模式中，企业明知侵权行为的发生而采取放任态度并从中获利，属于引诱侵权。<sup>④</sup>此外，在“大众点评诉爱帮案”与“大众点评诉百度案”中我国法院亦指出，垂直搜索算法应受到限制，价值中立并不能作为豁免法律责任的依据。

具体而言，以合法性为原则实现数据赋权需要从算法上考察数据行为是否符合相关标准，其体现了对交易行为予以限制的管制规则。与合法数据相对，必须明确建立违法数据的禁易规则，禁止个人信息、国防机密等不合法流通的数据交易。如上海市数据交易中心即将危害国家安全和社会稳定的数据、侮辱诽谤及损害名誉的数据、未经授权的个人信息、未经授权的企业数据等违法数据列入《数据流通禁止清单》。同时，以合法性原则确立数据权利亦可能对数据自由流动构成限制。由于产权的确立，分散独立的数据可能形成“数据孤岛”。因此，在涉及公共利益的特定情形下，为了进一步打破“数据壁垒”，责任规则也将予以适用。即为数据的强制流动设定一

定的对价。如美国高速公路安全管理局即规定，企业开展自动驾驶道路测试的条件之一是应将有关测试数据提供给当局。

## （二）以规制塑造权限：以用户为中心规则的构建

具体到管制规则层面，必须构建算法的具体约束性规则，以塑造数据权利的权限。具体规则的确定，又必须从回应型法的角度“考虑在其所处环境中各种新的力量”<sup>⑤</sup>。人工智能时代下，数据正义日益重要。数据制造者滥用数据权利将对用户产生极其严重的伤害。数据可能被用于进行市场价格操纵，通过“算法合谋”，形成垄断行为；可能被用于“产品动态定价”进行“大数据杀熟”，有害于消费者利益保护。同时，数据利用还可能出错。将普通人错误标注为犯罪嫌疑人，将旅客错误标注为恐怖分子，将“无家可归者”错误标记为“乞讨者”等大数据应用中的算法致害案例更是比比皆是。在这个意义上，算法也被称为“数字杀伤性武器”<sup>⑥</sup>。对此，以用户为中心的数据规制规则至少应在以下几个方面构建：

第一，风险最小化的算法设计。通过“自设计保护隐私”，风险最小化算法设计可解决个人信息保护与数据权利的界分问题。有学者认为，应以权利位阶原则与比例原则解决数据权利与个人信息保护的冲突。<sup>⑦</sup>“大众点评诉百度案”中二审法院亦强调“兼顾信息获取者、信息使用者和社会公众三方的利益”。然而，由于数据权利客体限于非个人信息，数据权利与个人信息实际上是互相排斥的。个人信息保护的要求，只能内化为数据处理的规制标准。而匿名算法是否有效则是其中的关键。“朱烨诉百度案”<sup>⑧</sup>中，相关数据因匿名而不受个人信息保护规范约束，便是二审改判的重要理由。客观而言，在一定条件下匿名数据也有被还原的风险。在分类经济中，即便是匿名数据也可能对个人造成威胁。因此，必须通过“自设计保护隐私”在算法中采取个人信息保护设计。

参照欧盟《大数据时代个人信息处理指引》，数据制造者应针对数据处理的不同阶段采用相应的解决方案，仔细考虑算法设计，尽量减少冗余数据存在，避免潜在的数据偏见或歧视。且在算法使用之前对其采取充分的风险评估，尽量减轻潜在的负面影响。同时，风险最小化的算法设计应被进一步广泛适用到所有与用户权利有关的风险评估中，而不应限于个人信息保护领域。也就是说，必须在算法设计中贯彻用户权利保障优先的理念，将监管与责任延伸至算法设计之中，并对算法进行相应的风险评估。通过加强算法保密性、构建安全防护措施、设立应急系统等具体技术方法减少算法安全风险。

第二，算法可解释权的构建。数据处理中，用户极易受到来自算法的伤害，“算法偏差”与“算法歧视”屡见不鲜。而数据的背后却存在一个“算法黑箱”。一方面企业常拒绝公布算法，另一方面算法复杂性的“技术鸿沟”也为理解算法决策带来障碍。算法解释权源于欧盟个人信息保护制度中的反对自动化决策权与知情权。虽然《一般数据保护条例》并未明文确立算法解释权，但却引起了算法解释权的重要讨论。一般认为，算法解释权可便于用户理解自动决策及其原因，帮助其反对这一决定，使其可改变未来的行为以获取更好的评估结果。

借鉴欧盟《监管目的的自动化个人决策和分析指南》的规定，用户应有权要求数据制造者对相关自动决策进行解释。数据制造者无须对算法内部机制进行解释，只需要根据“若非A则无B”的反事实(counterfactual)，以一个“没有受过教育的门外汉”能理解的标准，通过简洁、透明、易懂和容易取得的形式进行解释。需要注意的是，解释的简洁易懂不等于模糊不清。在多因素共同作用的情形下，应对各种因素的影响力作出说明；在相关性替代因果性的大数据应用中，亦应

对不同因素的相关性作出解释,否则将无益于用户理解。

第三,数据迁移标准的制定。就实现数据自由流动而言,唯有连通“数据孤岛”方可发挥数据的最大效用。韦思岸即指出,“数据自由迁徙的权利”如人身自由迁徙的权利一样重要。除数据权利的限制外,非个人数据本身的流动并无法律的禁止性规定。<sup>②</sup>然而,由于数据制造者所采取的算法不尽相同,数据的格式亦千差万别,为数据迁移造成了人为的障碍。增强数据的可流动性,必须在算法中运用标准传输格式。

参考欧盟《非个人数据自由流动框架》第6条的规定,相应的数据格式标准应是结构化的、常用的且机器可读的,同时应在订立数据服务合同前为专业用户提供关于数据迁移的必要信息,并建立相关数据质量与信息安全管理规范。

第四,算法审计制度的建立。复杂算法和海量数据要求算法审计必须由专门部门与专业人员进行。应在数据制造者内部设置数据顾问,在监管层面设置数据监管部门,二者作为算法审计的主体,实现良性互动。2017年美国加州的“算法审计”法案,即指出应建立专门机构对公共决策中算法可靠性、可能存在的“算法歧视”、可能导致的不利影响、算法验证性等方面进行审计。实际上,上文所分析的算法风险、算法可解释性、数据可迁移性都是算法审计的内容。此外,“算法合谋”“大数据杀熟”等可能对用户产生的不利影响也是算法审计的重要内容。一方面,“算法合谋”以算法作为隐性的垄断协议,虽规避了《反垄断法》第13条之规定,却实际上形成了价格操纵的效果。另一方面,“大数据杀熟”更是涉嫌违反《电子商务法》第18条关于定向推送算法应尊重消费者合法权益之规定。通过对算法各个方面的审计,方可落实用户权益保障的要求。

必须指出,以用户为中心的数据规制规则不以上述四点内容为限,算法的具体约束性规则亦绝不限于用户权益保障规则。以上述四项内容为代表,提出构建以用户为中心的数据规制规则,旨在回应数据权利可能带来的社会问题,并说明算法规制的具体构建模式,并未否认其他规则的重要性。除用户权益保障规则外,数据权利至少还需受公共利益和国家利益的限制。这些限制又内化在合法性原则的要求之中。

## 结 语

技术的发展呼唤法律体系的革新。就连“历史终结论”提出者福山亦认为,“除非科学终结,否则历史不会终结”<sup>③</sup>。正是在硬件技术以“摩尔定理”几何倍增的基础条件上,以算法技术为核心的数字经济得以形成。原始数据抓取算法、海量数据的处理算法、衍生数据的生成算法合力使得数据价值成为可能。数据制造者通过算法对数据生产进行投入,又以算法对数据利益进行维护,产生了数据权利诉求。而通过算法的规制手段,又进一步明确了数据权利行使的空间。

从算法的角度对数据权利进行考察可知:利用算法进行数据行为的数据制造者是数据权利主体;经过算法处理的非个人数据,无论表现为代码层的数据还是内容层的数据,均是数据权利客体;在算法规制的角度,以合法性原则与用户为中心的具体规则考察数据行为的合法性,又划定了数据权利合法行使的界限。由此,算法规制虽未以设立权利的方式正向确立数据权利,却通过规制手段反向实现了数据确权。

注释：

- ① 需要特别指出的是，由于《刑法》第253条之一的禁止性规定，作为交易客体的数据必须是非个人数据。同时，基于《民法总则》第111条与第127条对个人信息与数据的区分，本文所研究的数据确权与数据权利的对象仅限于非个人数据，仅在部分问题的论述中涉及与个人信息保护的交叉。
- ② 哈贝马斯：《在事实与规范之间：关于法律和民主法治国的商谈理论》，童世骏译，北京：生活·读书·新知三联书店，2014年，第111页。
- ③ (2018)浙01民终7312号《民事判决书》。
- ④ 劳伦斯·莱斯格：《代码2.0：网络空间中的法律》，李旭、沈伟伟译，北京：清华大学出版社，2018年，第1、135—136页。
- ⑤ 胡凌：《从开放资源到基础服务：平台监管的新视角》，《学术月刊》2019年第2期。
- ⑥ 徐格、李沁：《算法统治世界：智能经济的隐形秩序》，北京：清华大学出版社，2017年，第18页。
- ⑦ 阿里尔·扎拉奇、莫里斯·E.斯图克：《算法的陷阱：超级平台、算法垄断与场景欺骗》，余潇译，北京：中信出版社，2018年，第39页。
- ⑧ (2016)沪73民终242号《民事判决书》。
- ⑨ (2011)一中民终字第7512号《民事判决书》。
- ⑩ 龙卫球：《再论企业数据保护的财产权化路径》，《东方法学》2018年第3期。
- ⑪ 纪海龙：《数据的私法定位与保护》，《法学研究》2018年第6期。
- ⑫ 尼古拉斯·卡尔：《数字乌托邦》，姜忠伟译，北京：中信出版社，2018年，第38页。
- ⑬ Imanol Arrieta Ibarra, Leonard Goff, Diego Jiménez Hernández, Jaron Lanier, E. Glen Weyl, "Should We Treat Data as Labor? Moving Beyond 'Free,'" *American Economic Association Papers & Proceedings*, vol. 1, no.1, 2017, pp.1-2.
- ⑭ (1997)高知终第66号《民事判决书》。
- ⑮ 北京市第一中级人民法院：《百度诉奇虎360违反Robots协议案一审宣判》，<http://bj1zy.chinacourt.gov.cn/article/detail/2014/09/id/1446252.shtml>, 2019-05-01。
- ⑯ hiQLabs, Inc. v. LinkedIn Corp., 273 F. Supp.3d 1099

(N.D. Cal. 2017).

- ⑰ 格伦顿：《权利话语：穷途末路的政治言辞》，周威译，北京：北京大学出版社，2006年，第115页。
- ⑱ 程燎原、王人博：《权利论》，桂林：广西师范大学出版社，2014年，第14页。
- ⑲ 余成峰：《法律的“死亡”：人工智能时代的法律功能危机》，《华东政法大学学报》2018年第2期。
- ⑳ 乌戈·帕加罗：《谁为机器人的行为负责？》，张卉林、王黎黎译，上海：上海人民出版社，2018年，第153页。
- ㉑ (2016)京73民终588号《民事判决书》。
- ㉒ Vgl. AG Bad Hersfeld, Beschluss vom 15. Mai 2017 - F 120/17 EASO.
- ㉓ Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417 (1984).
- ㉔ MGM Studios, Inc. v. Grokster, Ltd., 545 U.S. 913 (2005).
- ㉕ P. 诺内特、P. 塞尔兹尼克：《转变中的法律与社会：迈向回应型法》，张志铭译，北京：中国政法大学出版社，2004年，第85页。
- ㉖ 凯西·奥尼尔：《算法霸权：数字杀伤性武器的威胁》，马青玲译，北京：中信出版社，2018年，第21页。
- ㉗ 许可：《数据保护的三重进路——评新浪微博诉脉脉不正当竞争案》，《上海大学学报（社会科学版）》2017年第6期。
- ㉘ (2014)宁民终字第5028号《民事判决书》。
- ㉙ 安德雷斯·韦思岸：《大数据和我们》，胡小锐、李凯平译，北京：中信出版社，2016年，第233页。
- ㉚ 弗朗西斯·福山：《我们的后人类未来》，黄立志译，桂林：广西师范大学出版社，2016年，第1页。

编辑 杜运泉