财经法学 No. 6，2026
pp. 105 - 121

数据的结构类型及其保护规则

梁志文*

内容提要：数据是符号意义上的信息代码，解封代码下封存的语义可获得不同的信息，这是数据复用性的体现，亦为数字经济发展的关键所在。数据产权客体非指单个数据，亦非原始数据或预处理数据，而是数据产品，即数据集合和基于数据洞见而提供的信息服务。信息服务的经济价值与实时性、动态性密切相关，契合禁止盗用原则，可为《反不正当竞争法》“数据条款”所涵摄。来源合法的数据集合，其保护路径应区分公开数据与秘密数据。前者受《反不正当竞争法》“数据条款”保护，以实质性投入、实质性替代与行为不正当为判断要件；后者可能同时符合商业秘密保护的条件，从而形成规范竞合。数据产权的限制规则主要体现为数据开放义务。该义务不是指数据向社会公开，而是在法定条件下与特定主体共享数据，包括语义层的用户权利、代码层的维护公平竞争以及公共利益等。

关键词：数据代码层　数据语义层　数据条款　数据产品　禁止盗用

一、引　言

数据是数字经济发展的石油或黄金。这一广为传播的比喻揭示了新技术时代的淘金热。从智能汽车、物联网、在线平台到人工智能各项应用，无处不在的各种传感设备正源源不断地收集、生成数据，数据也得以流通。数据已成为企业的关键资产，从产品开发到市场营销，它都发挥了基础性的作用，甚至被视为支撑数字经济的基础设施。〔1〕但是，如何具体构建数据产权制度，理论和制度上仍然处于探索阶段。

在理论上，数据赋权保护模式以确权为基础而主张数据法益的权利化，但数据权利排他性的

* 梁志文，华东政法大学知识产权学院教授。

〔1〕 See Ivan Stepanov, *Introducing a Property Right over Data in the EU: The Data Producer's Right-An Evaluation*, 34 International Review of Law, Computers & Technology 65, 65 (2020).

强弱各不相同。而行为规制模式则以数据利益的动态性为依据，主张规制数据市场相关主体的竞争行为，反不正当竞争法是其制度依据。〔2〕在制度建设上，前者开展了以确权为目标的数据权利登记试点，但是，作为确立新型财产权的法律，数据赋权立法殊为困难，故以行为规制为中心的反不正当竞争法成为法院处理数据纠纷的基本依据。〔3〕2025 年修订的《反不正当竞争法》第 13 条新增第 3 款数据保护的规定（以下简称“数据条款”），该款规定：“经营者不得以欺诈、胁迫、避开或者破坏技术管理措施等不正当方式，获取、使用其他经营者合法持有的数据，损害其他经营者的合法权益，扰乱市场竞争秩序。”

毫无疑问，“数据条款”属于行为规制范畴，以竞争者扰乱市场竞争秩序的不正当行为为规制对象，但它并未终结数据财产保护的制度路径之争。因为数据产权客体为何、怎样配置权利、如何保障交易等核心问题仍未得到明确解决。此外，数据具有非竞争性，为“促进数据要素流通复用”〔4〕，还需要解决第三方如何合法利用等问题。上述问题带来的不确定性将对数字经济的健康发展产生负面影响，难以契合“数据产权、流通、交易、使用、分配、治理、安全等基本规律”〔5〕的要求。

本文试图为解决上述核心问题提供理论方案，而所有方案的出发点都必须从数据这一基本点出发。为此，本文将回到人们习以为常的数据概念。数据的结构类型，即数据的代码层与语义层结构，是解开这些难题的钥匙。因为它揭露了“数据产权”这一术语所遮蔽并进一步由《反不正当竞争法》“数据条款”所强化的假象：财产权的客体即数据。但单个数据及其所承载的语义不是数据财产；单个数据可能涉及个人信息权、消费者权益、物权等。数字经济独特的价值创造来自数据规模及其预测结果，这才是数据财产的客体——数据产品。因此，从数据产品的代码层与语义层出发，应根据其不同的法律属性而确定其保护规则，并据此完善其权利限制规则，促进数字经济的健康、快速发展。

二、作为数据产权客体的数据结构类型

数据产权制度的首要问题是保护客体，客体即数据，这似乎是常识。但何谓数据，民事法上尚无明确定义。《民法典》第 127 条沿袭《民法总则》（已失效）的规定，概括性地规定“法律对数据……的保护有规定的，依照其规定”。此前，《网络安全法》（2016 年）第 76 条第 4 项界定了“网络数据”，即“通过网络收集、存储、传输、处理和产生的各种电子数据”。四年后通过的《数据安全法》第 3 条第 1 款规定：“数据，是指任何以电子或者其他方式对信息的记录。”据此，数据不等于信息，而是对信息的记录。欧盟《数据法》在其序言（recital）第 15 段也区分信息与数据，前者是从数据中推断或派生出来的结果，即“对数据赋值或获取洞见的额外投资结果，特

〔2〕 国内学者近年来发表了大量数据财产保护的研究成果。较全面的述评，参见陈志君：《数据知识产权司法保护进路探析》，载《中国应用法学》2025 年第 5 期。

〔3〕 参见陈志君：《数据知识产权司法保护进路探析》，载《中国应用法学》2025 年第 5 期，第 113－115 页。

〔4〕《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》（2022 年 12 月 2 日发布）第二部分第（七）节。

〔5〕 同上注，第一部分第（二）节。

别是通过专有复杂算法（包括作为专有软件一部分的算法）获得的结果”〔6〕。

（一）数据的代码层与语义层

数据是“对信息的记录”，〔7〕它是符号意义上的信息编码，在其意义被处理或分析之前，以比特等符号形式存在，是表示行为、事实或事件属性的未解释的符号。信息则属于符号的语义层面，是经过处理的、已经具有意义或可用性的文本、图片或视频等。〔8〕这如同语言文字与文字作品的关系，文字是符号代码，文字传递的人类可理解的意义是作品。因此，数据具有如下双层结构：以符号形式呈现的代码层和封装了人类可理解的语义层。被代码层封装的内容可以通过解释而赋予数据意义，从而在代码与语义之间存在线性流动，为数据的使用者、持有者带来附加价值，即形成知识。

1. 数据价值的增值来自代码层的数据复用性

数据之所以被类比为石油，是因为它具有解封代码层下的语义而获得进一步的信息、形成新知识之能力，即数据具有复用性。以消费者经常购物的超市地理位置数据为例：导航软件可以预测交通拥堵程度，可以推荐附近加油站或餐馆；银行可以分析客户消费习惯并用于信用卡欺诈预警，如将异地刷卡判断为异常；其他商家可据此投放精准定向广告；主管部门在突发传染病疫情时可用以确定密切接触者。〔9〕数据复用性体现了数据的非竞争性特征，它可为任意数量的使用者同时利用而不会产生损耗。例如，任意数量的机器学习算法可同时利用人口普查数据、汽车行驶数据进行训练，而不会产生数据损耗。

数据规模被认为是提高人工智能输出质量的重要因素，因为洞察本质、解决问题和提高效率往往基于对数据大量汇集后潜在相关性的揭示，如导航软件拥堵预警需要综合权衡高精度地图数据、传感器数据、海量实时路况等交通数据，通过算法依据一定规则进行预测。规模化的数据可以分析出小数据中未能展现的相关性。例如，医生为很多患者开出的药物很安全，却在某一患者身上出现严重副作用，这通常难以解释原因。当汇集大量患者数据而特定群体出现同样副作用时，则可以发现出现副作用的原因，如遗传因素或与其他药物产生相互作用。〔10〕

数据生产即数据收集、处理和使用，是一种独特的价值创造形式，这也使得数据具有重要的交换价值。例如，诊断皮肤癌的人工智能，其算法通过病理报告和癌症结果标注数据对潜在癌症图像进行训练。如果医院只能运用其拥有的患者标注图像来训练算法，与可以使用全世界所有患者图像进行训练的竞争算法相比，基于更大样本的后者将具有更高的诊断质量，能更好地治疗患

〔6〕 Regulation（EU）2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation（EU）2017/2394 and Directive（EU）2020/1828，OJ L 2023/2854（Data Act，DA）.

〔7〕 有学者将数据的代码层与语义层分别称为数据文件与数据信息。参见纪海龙：《数据的私法定位与保护》，载《法学研究》2018年第6期，第72页。但也有学者认为数据与信息在法律概念上并没有严格区分的必要。参见梅夏英：《信息和数据概念区分的法律意义》，载《比较法研究》2020年第6期，第151页。

〔8〕 See Josef Drexl，*Designing Competitive Markets for Industrial Data-between Propertisation and Access*，8 Journal of Intellectual Property，Information Technology and Electronic Commerce Law 257，263（2017）.

〔9〕 See William Magnuson，*A Unified Theory of Data*，58 Harvard Journal on Legislation 23，39－40（2021）.

〔10〕 参见〔荷〕玛农·奥斯特芬：《数据的边界：隐私与个人数据保护》，曹博译，上海人民出版社2020年版，第49页。

者并挽救生命。〔11〕因此，数据交易既是数据复用性的内在需要，也是大数据汇聚的重要保障。在这个意义上，单个数据，无论是来自自然人还是物联网生成，仅具有极小的经济价值，“小得可以忽略不计”〔12〕。当然，这并不意味着单个数据在特定应用场景下不具有价值，如前述消费者经常购物的超市地理位置数据可以成为定向精准广告的依据，既可能丰富消费者的选择，也可能产生消费者歧视等后果。这表明，数据的非竞争性特征在具体应用场景中不仅产生正外部性，也产生负外部性。

2. 数据（代码）与信息（语义）的区别

数据与信息的不同之处在于，前者作为机器能够执行计算功能的编码，是输入端的“原料”；后者作为有意义的内容层面，则是输出端的结果。以智驾系统为例，汽车通过摄像头、激光雷达、GPS 等各种传感器获取的路面数据和交通数据，连同专业司机在面对各种交通状况时的操作数据，这些参数被算法运行以生成尽可能准确的预测。成功的智驾算法本质上是通过训练数据和实况数据，权衡非线性模型参数而形成预测的规则。〔13〕

数据与信息的不同之处还在于其排他性程度不同。从技术层面来看，数据传输便捷，“只需按下按钮，数据就能通过互联网发送出去”〔14〕。它通常以非常低的边际成本进行转移或复制。信息也具有非竞争性特征，但因其负载人类可理解的意义，需要投入资源才能实现其转移，如人类理解信息时的成本。此外，数据可以被加密，加密数据很难被获得者破解，更容易被监控，具有更高的排他性。例如，机器学习算法的训练数据通常是保密的，开发者自行收集训练数据，并形成自己的竞争优势。即使是开源人工智能算法，也不公开训练数据，“训练框架尤其是训练数据的开源在业界少有先例”〔15〕。例如，DeepSeek 开源了模型架构、训练方法与应用生态，但未完全公开训练数据。但信息不同，获得信息的工程师在离职时会带走知识，因为人们的自由迁徙和交流，最终会导致信息的传播。

（二）数据产权客体的结构类型

从数据的代码层与语义层结构出发，数据产权的客体存在不同学说。有观点认为，“数据产权的保护客体乃是位于句法层的数据，而非语义层的信息”〔16〕；并强调数据本体与信息的“共生统一体”关系。但是，代码层（即句法层）的数据与语义层的信息并非简单的一体二面关系，因为数据集合的更多价值在于其复用性所产生的新知识或信息，此类新信息往往超出了“共生统一体”的范畴，但同样具有极其重要的经济、社会价值。与此相反的观点认为，数据产权的客体既不是数据，也不是数据集合，而是“运用大数据分析技术对数据集合加工而成的关于特定信息内

〔11〕 See Charles I. Jones & Christopher Tonetti, *Nonrivalry and the Economics of Data*, 110 American Economic Review 2819, 2820 (2020).

〔12〕 Eli Freedman, *Data Unions: The Need for Informational Democracy*, 111 California Law Review 657, 660 (2023).

〔13〕 See Charles I. Jones & Christopher Tonetti, *Nonrivalry and the Economics of Data*, 110 American Economic Review 2819, 2821 (2020).

〔14〕 Ibid.

〔15〕《DeepSeek 到底是真开源还是假开源？业内：已经有足够多细节学习》，载网易号 DeepTech 深科技 2025 年 2 月 2 日，https://m.163.com/dy/article_v5/JNDVS2MI05119734.html，2026 年 9 月 4 日访问。

〔16〕 申卫星：《论数据产权制度的层级性：“三三制”数据确权法》，载《中国法学》2023 年第 4 期，第 32 页。

容的产品，其能够用于解决特定问题、满足特定商业需求”〔17〕。然而，它局限于静态数据的语义价值，未能考虑到数据集合本身比特定知识或信息具有更多经济价值，即数据复用性。代码层的数据集合与语义层的预测数据的关系，更恰当的比喻是食材与菜品的关系，就如同烹饪时食材的不同组合会形成不同风味的菜品，食材提供者与菜品制作者都贡献了各自的价值。因此，数据产权保护客体不仅应该包括食材（数据集合），也应该包括菜品（预测数据或信息），它们可统称为数据产品。

数据产品是对原始数据与预处理数据进一步予以实质性投入进而形成的作为数据分析对象的数据集合，或者形成的可理解或可用的信息。它包括如下两个类型：

其一是代码层面的数据集合，这是数据分析与数据复用性的重要载体。它与简单的数据收集不同，需要权利人予以实质性投入。实质性投入的判断具有伴随数据体量变化的连续性特质。以国家知识产权局2025年公布的数据知识产权登记典型案例为例，属于数据集合类型的有“医疗大模型预训练数据集”“多模态成年人群脑影像数据”等。〔18〕数据集合不同于通用性质的数据预处理，它需要根据特定目的对原始数据、预处理数据进行整理、提炼、标注，有些情况下还需要对数据重新校准和去噪。因此，“数据集合具有更高的应用价值和市场价值，具有规模化、结构化、标准化的显著优势”〔19〕。数据集合的本质是数据的代码层，可通过解封其语义而获得新的信息、知识。特定目的之数据处理可能消除原始数据封装的部分信息，如果存在元数据的配套协同，原始数据或经过通用处理的数据具有更大的潜力来满足各种用途并产生多样化的结果。

其二是狭义上的数据产品，它是从数据或数据集合中推断或派生出来的成果，即语义层的预测数据。它是权利人基于数据分析所获得的信息、知识，为特定目的向相对人提供的服务或实物。它主要包括两类。其一为有形实物，如安装智驾系统的汽车。其二为数据（信息）服务，如在高速公路与市政道路实时交通数据基础上提供的导航服务、基于消费者平台购物记录形成的畅销产品信息等。前述数据知识产权登记典型案例中的“电力看‘双碳’监测数据集”“船舶实时报位信息数据”等，〔20〕它们名为数据集，但实际上是通过实时数据分析而提供的服务，如“双碳”监测、船舶实时报位。

数据产品要实现作为驱动经济发展的石油资源之作用，就需要从数据分析中创造经济价值，实现其使用价值与交换价值。其经济价值的变现通常有三种方式。前两种是对外变现，包括直接的数据集合销售，以及基于数据分析形成的预测、知识（简称为“预测数据”）驱动的产品和服务销售；第三种是内部变现，即基于数据洞见而提升自身竞争优势，包括明智决策、优化运营、提升质量、创新产品、识别趋势、洞察客户行为等等。〔21〕数据产权制度的核心在于保护投资者避免因数据的非竞争性特征而失去竞争优势和数据控制。因此，其制度构造主要体现为外部变

〔17〕 刘维：《论数据产品的权利配置》，载《中外法学》2023年第6期，第1584页。

〔18〕 参见《数据知识产权登记十大案例》，载《中国知识产权报》2025年5月24日，第3版。

〔19〕 申卫星：《论数据产权制度的层级性：“三三制”数据确权法》，载《中国法学》2023年第4期，第32页。

〔20〕 参见《数据知识产权登记十大案例》，载《中国知识产权报》2025年5月24日，第3版。

〔21〕 See Samiksha Shukla et. al., *Data Economy in the Digital Age*, Springer, 2023, pp. 37－43.

现，即前两种数据价值实现方式，包括数据集合交易与预测数据服务。

（三）数据产品不包括未有实质性投入的数据

依前所述，数据产权的客体不是单个数据。有些单个数据属于权利人开展正常业务的附随产物，如消费者购物记录或在线社交数据；有些无需额外投入资金。欧盟《数据法》序言第 15 段不仅区分信息与数据，而且还细化了数据的类型，界定了“原始数据”“预处理数据”“元数据”等概念，并将这些数据都统称为“未进行实质性修改的数据”。〔22〕

原始数据是指产品使用或提供服务时对产品或用户行为的数字化记录。它包括产品传感器自动生成的数据以及嵌入式应用程序记录的数据、如指示硬件状态和故障的应用程序所记录的数据、由模拟信号转换为数字信号获得的数据，以及初始电子信号记录。例如，传感器将物理环境中温度的模拟信号转换成电子信号，并以数字化形式（如 101，11）记录。然而，原始数据不是直接的物理值/量，它不是摄氏度计量的温度值，与可理解单位（如摄氏度）之间的关系可能并不清晰。因此，需要将传感器记录的原始数据转换为可理解的测量单位，如代码 101 等于 43 摄氏度。〔23〕这一过程被称为“数据预处理”，它主要包括校准与去噪两步。

校准是将原始数据与传感器校准规则进行比较，从而获得可理解的数据。用来进行校准的数据是元数据，即描述数据的数据，它是对数据内容或使用情况的结构化描述，以方便数据的发现或使用。〔24〕不同厂家的传感器具有不同的规格，需要使用相应的校准公式。来源庞杂的数据集合可能混合有不同传感器所生成的数据，用元数据进行校准极为重要。以火星轨道器事件为例，该轨道器旨在研究火星气候，经历多年准备，投资达数十亿美元，但当其进入火星轨道时，因过于接近火星，最终在大气中烧毁。事后分析发现，原因是数据库使用了两种不同的轨迹标准，两种不同规格的数据组合导致了轨迹计算错误。〔25〕

去噪是消除数据误差。与系统误差不同，噪声是随机误差，它与被测量的物理现象无相关性，因此无法通过校准消除。〔26〕由于传感器首先产生模拟信号并最终转换为电子信号，任何干扰传感器或电子设备的环境因素都可能在信号中产生误差。一般使用算法去噪，去噪算法的原理是运用精密度衡量数据。因为没有噪声干扰的物理对象所测量的数据值应该保持平滑，通过测量传感器数据的波动程度，可以有效滤除噪声数据。无论是原始数据还是校准后数据，都可以使用去噪算法，因为在原始数据转换为校准数据的过程中噪声不会减少。

经过校准、去噪后形成的数据被称为预处理数据。为便于后续处理和分析，数据交易的对象应经过预处理；当然，如果伴随有足够的元数据，原始数据也是可用的。当目的一致时，获取特

〔22〕 See Recital 15 of EU Data Act.

〔23〕 See Daria Kim & Man Wai Kwok, *Data Usability as a Parameter of Rights and Obligations under the EU Data Act*, 15 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 139, 142 (2024).

〔24〕 See Article 2 (2) of EU Data Act.

〔25〕 See Michal S. Gal & Daniel L. Rubinfeld, *Data Standardization*, 94 New York University Law Review 737, 747 - 748 (2019).

〔26〕 See Daria Kim & Man Wai Kwok, *Data Usability as a Parameter of Rights and Obligations under the EU Data Act*, 15 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 139, 142 (2024).

定用途的处理数据是性价比较优的选择，但即便是细微的差异，也可能导致数据接收方更倾向于自行进行预处理。通用性质的数据预处理主要通过自动化实现，通常涉及微不足道的成本，主要取决于校准设备与去噪算法的成本，如去噪方法的选择、去噪算法的复杂性和数量。

三、数据产品中语义层的保护规则

数据产品中的语义层，其本质是解封数据的语义内容，向相对人提供对数据进行特定目的处理之预测结果，属于可直接理解与产生意义的内容，即信息。例如，平台对“双十一”期间商品交易的记录，如详细购物清单的支付方式、支付时间、收货地址、物流状态等，属于数据，但统计分析后得出的某品类畅销产品型号、销量、价格、消费者群体等内容，如淘宝公司“生意参谋”产品，则属于信息。在司法实践中，部分法院将其称为衍生数据、数据产品。〔27〕预测数据属于平台对用户浏览、交易等行为痕迹数据予以特定目的处理后的结果，与用户行为痕迹等原始数据不具有直接关联性。例如，平台基于用户行为痕迹数据而提供的定向广告服务，系通过对用户偏好及其社交数据进行挖掘而形成潜在消费者之预测数据，进而将之销售给广告主。上述例子揭示了预测数据的两种形式。“生意参谋”属于多项信息的集合，而定向广告则属于单一内容的信息。一些案件中被认定的数据集合，本质上属于多项信息的集合，如通过统计分析所形成的金属价格走势图与每日最高价、最低价、均价及涨跌等。〔28〕

（一）预测信息保护的可能路径

部分信息可能构成著作权法保护的作品。多项信息的集合可能是具有独创性的汇编作品，也可能构成特定类型的作品，例如，地物、地貌、路网等信息及其地理坐标汇集成的地理信息数据库，在软件的运行下最终形成的电子导航地图，属于受保护的地图作品。〔29〕部分信息具有秘密性，在权利人采取了相应保密措施的情况下，可以作为商业秘密得到保护。还有部分信息是算法分析的结果，算法的软件著作权和由计算机实施的算法专利权可以间接保护信息，如分析实时交通数据而生成的拥堵指数与绕行导航。“拥堵延时指数”由算法对多种基础数据进行加工、处理和计算而自动生成，独立于用户 GPS 轨迹数据；〔30〕试图通过基础数据的盗用而提供同样的指数服务，可能侵犯算法专利权。

但是，大部分信息无法获得知识产权的保护。首先，有些信息是对事实的描述，如拥堵指数、生意参谋等，因其表达的方式极其有限，依据合并原则而不具有独创性，或者属于算法自动生成而不具有创作性，不构成作品。其次，信息本身既不属于专利保护的产品，也不属于方法专利所延及保护的产品。此外，信息一旦公开发布，就不具有商业秘密的地位。

〔27〕 参见浙江省杭州市中级人民法院（2018）浙 01 民终 7312 号民事判决书。
〔28〕 参见福建省高级人民法院（2024）闽民终 933 号民事判决书。
〔29〕 参见广东省高级人民法院（2008）粤高法民三终字第 290 号民事判决书。
〔30〕 参见北京知识产权法院（2024）京 73 民终 1761 号民事判决书。

（二）预测信息的反不正当竞争法保护

无论是否赞同信息产权理论，〔31〕受到排他权保护的信息必须具备特定的条件。基于预测数据的信息服务受排他权保护的首要障碍在于，客体确定性、价值性与信息实时性、动态性之间的矛盾。信息服务需要实时数据来支持信息的精确性，其价值也与实时性数据的收集、分析密切相关。例如，对微博关键词数据进行统计分析、加权计算形成指数，它是微博热点趋势或舆情变化趋势的量化指标，用户可以作为市场运营、舆情分析、学术研究等活动的参考，〔32〕但其决策价值明显具有时效性。尽管历时性数据可以提供某种程度的预测，但它明显劣于建立在实时数据基础上的信息服务。例如，提供公交车到站查询服务，需要依据公交车实时运行路线、运行时间、公交车 GPS 精确定位等数据，运用算法形成精确的到站预报，它仅在某一时刻提供的到站信息服务是具有价值的，这也是它在同类软件中的竞争优势所在。〔33〕

上述具有实时性的信息服务，与美国法上禁止盗用规则的适用范围极为类似。该规则肇始于美国联邦最高法院于 1918 年审理的一起涉及未经许可而系统性复制新闻报道的案件。版权法不保护时事类消息新闻，但新闻报道需要耗费大量的人力物力，极强的时效性是其竞争优势的体现，因此可被视为准财产而受到保护，盗用新闻的不劳而获行为构成不正当竞争。〔34〕自此，禁止盗用规则为具体知识产权之外各类准财产保护提供了测试标准，“这在整个二十世纪都不是一个新现象”〔35〕。此后，美国判例法逐渐完善了盗用侵权规则测试标准的五项因素：原告为信息的收集付出了成本；该信息具有时效性；被告使用该信息构成不劳而获的搭便车行为；被告提供的产品或服务与原告构成直接竞争；被告搭便车行为将减少原告提供产品或服务的动机，使其产品或质量受到实质性威胁。〔36〕

禁止盗用规则属于典型的行为规制模式，属于反不正当竞争法的调整范围，它聚焦于竞争关系中的竞争行为。〔37〕我国法院也大抵从上述要件论证数据权益是否应受保护。例如，有法院认为，拥堵延时指数是原告付出了技术、资金、人力、物力等的实质性投入合法生成的，可对城市拥堵状态及其趋势以精确、简练、直观、量化的方式呈现，在原始地图数据上增加了新的商业价值，是受保护的竞争性权益，被告在未经许可的情况下通过技术手段获取并在其产品中实时提供上述信息，违背了数据服务行业的商业道德，属于不劳而获的行为。〔38〕

除了盗用行为外，数据干扰行为也是一类新型的不正当竞争行为。基于数据洞见而开展的信息服务需要建立在高质量的数据集合基础上，否则会产生“垃圾（数据）进，垃圾（结果）出”的后果。“数据投毒”是一种恶意干扰信息服务的行为。所谓“数据投毒”，是指攻击者将少量恶意数据或已篡改的已有数据注入训练数据，使模型在训练或者微调中出现错误模式，从而输出错

〔31〕 参见李琛等：《中国知识产权法学术史》，中国人民大学出版社 2025 年版，第 170－174 页。
〔32〕 参见广东省高级人民法院（2022）粤民终 4541 号民事判决书。
〔33〕 参见广东省深圳市中级人民法院（2017）粤 03 民初 822 号民事判决书。
〔34〕 参见李明德：《美国知识产权法》（第 2 版），法律出版社 2014 年版，第 675－690 页。
〔35〕 梁志文：《从“山寨现象”看禁止盗用原则及其应用》，载《法学》2009 年第 7 期，第 101 页。
〔36〕 See National Basketball Association *v*. Motorola，105 F. 3d 841（2d Cir. 1997）.
〔37〕 参见梁志文：《从“山寨现象”看禁止盗用原则及其应用》，载《法学》2009 年第 7 期，第 106－107 页。
〔38〕 参见北京知识产权法院（2024）京 73 民终 1761 号民事判决书。

误结果。〔39〕例如，有些社交平台使用用户数据进行内容推荐或热点趋势预测，攻击者通过批量注册用户和添加虚假评分，操纵系统来推举或压制特定对象。攻击者有意采取的数据投毒行为破坏了数据要素价值，使数据的采集、存储、流转、利用等全生命周期都面临被污染的风险，造成信息服务的准确性、可信度下降。

相比于排他权的保护，以行为规制为核心的反不正当竞争法保护既符合信息服务的实时性、动态性特征，也为竞争者合法挖掘数据、独立分析数据提供了制度保障，因为其并不限制代码层数据的合法流通，进而保障数据的可复用性。《反不正当竞争法》“数据条款”具有极大的弹性空间，该款中“经营者合法持有的数据”足以覆盖预测数据的语义层，即信息，故法院无须再适用其一般条款。

四、数据产品中代码层的保护规则

《反不正当竞争法》“数据条款”中“经营者合法持有的数据”当然包括数据集合。数据集合与预测数据具有不同的法律特征。首先，预测数据所适用的禁止盗用原则通常具有时效性，仅在实时情境下才具有竞争优势，如导航软件提供的交通拥堵指数及其绕行规划。历时的拥堵指数具有一定预测价值，但不再具有同等竞争力。而数据集合并不具有时效性。其次，两者可能产生的竞争损害也不同。信息服务或实物交易即可获得经济价值，它不需要对据以分析的数据集合进行交易，如微指数具有预测热点趋势或舆情变化等价值，微指数的交易并不需要微博关键词数据集的交易。与之不同的是，直接的数据集合交易让权利人不仅可能失去对数据的控制，还可能失去基于数据洞见而具有的竞争优势。因此，“数据条款”对数据集合的保护应作与预测数据不同的适用性解释。

（一）数据来源合法性与数据分类

数据集合集成了经营者合法采集的大量数据。尽管单个数据对于数据经济而言微不足道，但它可能涉及个人信息等人格性利益，也可能涉及用户作为财产所有人的合法利益，如汽车购买者对行车数据的利益。因此，数据的采集、存储、流转、利用等全生命周期均应严格合规，满足合法性条件。例如，敏感数据予以脱敏处置是数据集合合法性的重要条件之一。故单个数据的法律属性也影响数据集合的保护。在学理上，数据有很多分类，如《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》按照来源将数据分为个人数据、企业数据与公共数据，该分类偏重于来源合规与数据开放等问题。

在各种学理分类中，公开数据与秘密数据的区分对于数据集合保护具有重要意义。一般而言，物联网生成的数据大都属于机器交互时产生的数据，如汽车激光雷达等传感器捕捉车辆周围环境数据并传输至服务器，它无需向公众展示，属于事实上的保密控制，构成秘密数据。平台数据与之有所不同，它既有公开数据，也有秘密数据，后者是存储于服务器、不向公众展示的后台数据。

〔39〕 参见马润艺：《“数据投毒”的风险与规制路径》，载《科技管理研究》2025 年第 8 期，第 158 页。

（二）公开数据的保护

由公开数据组成的数据集合受保护的条件，除了要求数据来源合法之外，实质性投入是其决定性条件。实质性投入包括实质性的智力投入与财力投入，它是数据产品与原始数据、预处理数据的重要区分。数据的简单汇总不足以构成实质性投入，因为它们大都属于正常业务的附随物。但如果运用专有算法对数据进行清洗，如处理缺失值、去除重复数据、修正数据错误、统一数据格式、清除噪声数据等，则属于付出实质性的投入。对原始数据或通用预处理数据添加说明、解释、分类或编码，是数据标注产业的主要业务。它主要通过人工或半自动的方式进行，形成具有特定格式的结构化数据，是数据从潜在资源向可交易资产转变的关键环节。在这一过程中，人工标注的工作不可替代，甚至被称为“有多少人工就有多少智能”〔40〕。作为数字经济的重要组成部分，同时作为人工智能产业的支撑行业，数据标注产业的健康、高速发展可赋能各种应用场景。

对数据集合的保护不延及公开数据本身。公开数据处于自由流通状态，他人可以通过合理、正当手段进行收集。当然，它并不等同于公共数据，不代表其他经营者可以不受任何限制地获取和使用这些数据。〔41〕例如，爬虫技术在网络环境中被广泛使用，大部分流通数据来自公开数据的爬取；〔42〕爬取公开数据往往是正当的，但绕过技术措施爬取则不具有正当性。《反不正当竞争法》第 13 条第 3 款具体列举了“欺诈”“胁迫”“避开或者破坏技术管理措施”等不正当方式。此外，从他人的数据集合中合法爬取公开数据，除了采用合理、正当手段外，还不得对权利人的产品或服务构成实质性替代。最高人民法院第 262 号指导案例认为，原告实质性投入人力、物力、财力而形成的由用户信息、短视频、用户评论等数据汇聚的集合，具有独立于单一视频的经济价值，被告抓取搬运了案涉数据集合中的大量用户信息、短视频和评论等数据并向公众提供，足以实质性替代原告的产品和服务，构成不正当竞争行为。〔43〕因此，数据来源合法与实质性投入成为公开数据集合受保护的基本条件，而不正当行为与实质性替代的损害后果则是侵权行为判断的核心要件。

最后还要指出的是，采取开源策略的数据集合不是公共数据或公开数据。开源许可证本质是权利人附条件的单方许可，授予所有人依许可证条件使用。如果使用者违反了许可证限制性条件，则因丧失授权条件而构成侵权。〔44〕例如，在某一案例中，采集、录制全国范围内数千人员的口语化句子，并经专业标注而形成的语音数据集，符合实质性投入的条件；其采集的数据不能识别特定自然人，具有合法性。该数据集对数据的选择或编排不具有独创性贡献，不构成汇编作品，其以开源形式在官网提供下载，不具有秘密性，不受商业秘密保护，但其符合实质性投入与来源合法的条件而受到保护。被诉行为违反了涉案开源协议的非商业目的使用规则，被法院认定

〔40〕 粟瑜：《人工智能时代数据标注众包劳动的法律保护》，载《暨南学报（哲学社会科学版）》2023 年第 1 期，第 65 页。

〔41〕 参见浙江省高级人民法院（2023）浙民终 1113 号民事判决书。

〔42〕 参见许可：《数据爬取的正当性及其边界》，载《中国法学》2021 年第 2 期，第 172 页。

〔43〕 参见最高人民法院指导案例 262 号（2025 年）。该指导案例发布的裁判理由特别指明，该案应适用 2025 年 10 月 15 日生效的《反不正当竞争法》第 13 条第 3 款的相关规定。

〔44〕 参见〔美〕希瑟·米克：《开源软件许可实用指南》（第 3 版），刘伟译，人民邮电出版社 2023 年版，第 63－64 页。

构成不正当竞争。〔45〕

（三）秘密数据的保护

与公开数据不同，秘密数据因其本身处于秘密状态，导致第三人获取、使用数据行为的正当性高度存疑。未经授权获取秘密数据的行为通常与规避技术措施、违反合同约定等不当行为相关联。如果通过技术手段爬取了平台的后台数据，如不在网页上公开展示的数据、正常通过浏览器无法获取的数据，则可推定其爬取手段是不正当的。〔46〕平台对于授权获取数据的用户设置了限定特定用途及保密等义务，违反保密义务而获取、使用或允许他人使用的行为也难谓正当。

大量数据处于事实上的保密状态。汽车、智能家居、可穿戴设备等物联网数据大多数属于机器之间交互形成的数据，存储在终端设备与远端服务器上。有些可为设备的物权人获取，如智能手表记录的血压、血氧、运动轨迹等数据，但不向公众开放。还有些数据为设备的生产商或服务商所控制，不向社会公开，甚至也不向设备的物权人公开。例如，汽车通过雷达、相机、GPS、轮速传感器、温度传感器等一系列装置收集交通环境数据（如障碍物、行人等），以及车辆性能数据（如发动机、变速器温度等），并将其记录于车载数据记录器和云端数据记录器。这些数据涉及智驾性能等核心竞争优势，也涉及交通事故责任、车辆质量责任等法律争议，数据控制者常常以商业秘密为由不予公开。〔47〕还有一类合成数据，即通过算法生成的数据，处于合成者的事实控制下，在公开前具有秘密属性。

由秘密数据组成的数据集合既是“数据条款”语义涵摄的范围，也可能符合商业秘密保护的条件，属于规范竞合而须择一适用的情形。《反不正当竞争法》第10条第4款将商业秘密的客体界定为“商业信息”，其范围似乎难以涵盖代码层的数据，但“商业信息”的含义非常宽泛。首先，“商业”一词不能作狭义解释，它包括市场竞争能够利用的所有信息。故它仅仅排除了非工商业领域的竞争行为。〔48〕其次，“信息”一词的含义也同样宽泛。例如，公开上映前的电影、游戏安装包被拆包而非法获得的游戏隐藏角色、上市前的产品包装设计等等，都受商业秘密保护。〔49〕《最高人民法院关于审理侵犯商业秘密民事纠纷案件适用法律若干问题的规定》第1条对技术信息、经营信息的界定，就明确列举了与技术及经营活动“有关的……数据……等信息”。欧盟《商业秘密指令》序言第2段也明确指出商业秘密保护的信息“不限于技术知识，还包括客户和供应商信息、商业计划以及市场调研和战略等商业数据”。

数据是否受商业秘密保护的争议还有秘密性要件。〔50〕采取保密措施的数据集合并不当然符合秘密性要件。秘密性不是新颖性，也不是著作权法上的独创性或专利法上的创造性，但它仍需要体现为一定程度的智力投入或投资。需要权利人付出实质性投入所获得的数据，就不属于“普

〔45〕 参见北京知识产权法院（2024）京73民终546号民事判决书。

〔46〕 参见广东省高级人民法院（2022）粤民终4541号民事判决书。

〔47〕 参见湖北省武汉市江岸区人民法院（2025）鄂0102民初8068号民事判决书。

〔48〕 参见梁志文：《知识产权法教程》，南京师范大学出版社2024年版，第296页。

〔49〕 参见陶乾：《商业秘密保护法的规范构造研究》，法律出版社2022年版，第53页。

〔50〕 代表性的反对意见，参见芮文彪等：《数据信息的知识产权保护模式探析》，载《电子知识产权》2015年第4期；赞同意见，参见崔国斌：《新酒入旧瓶：企业数据保护的商业秘密路径》，载《政治与法律》2023年第11期。

遍知悉和容易获得”的情形，满足秘密性条件。〔51〕同样，负有法定或约定保密义务的人知悉的信息，不影响其秘密性。此外，公开数据的集合也可能符合秘密性条件。将公共信息进行整理、改进、加工之后形成的新信息，在学理上被称为“混合式商业秘密”，其典型类型是客户信息，它通过汇集客户名册，经过特定处理而增加某些“独特的”“新”的内容，形成了受保护的深度信息或广度信息。〔52〕

五、基于数据开放的数据产权限制规则

反不正当竞争法提供的防御性权利对数据本身并不赋予排他权，“对于不属于国家秘密、个人信息和商业秘密的数据信息，应允许自由流动，非因法定事由不应过度管控，以防止形成‘数据壁垒’‘信息封闭’”〔53〕。对公开数据，他人可依法采集并对其加工形成数据产品。商业秘密的保护也不影响他人通过独立研究或反向工程而合法获得相关数据。例如，其他汽车厂家自行采集的交通数据，并不侵犯相关数据集合的商业秘密。因此，反不正当竞争法保护数据财产的防御属性既维护了权利人的竞争利益，又保障了数据的合法流动，即数据开放。必须指出的是，数据开放不等同于数据向社会公开，更多场景下的数据开放是指在法定条件下与特定主体共享数据。

（一）语义层的数据开放：用户权利

对于物联网设备和平台生成的数据，权利人付出了实质性投入，但用户也对数据的生成具有重要贡献，也享有合法权益。其中，有些属于其个人信息的记录，有些属于物权人、消费者或知识产权人（如短视频作者）合法权益的载体。对于用户而言，这些数据往往是承载具体意义的信息，属于数据语义层的利用，它涉及数据集合中的部分数据或单个数据。

第一，基于个人信息保护的数据开放。有些数据能够直接或间接识别到特定自然人。电子商务平台消费记录中的各项数据，如收货地址、支付方式等，构成个人信息。即使单个数据不能识别，但通过数据集合可识别到特定自然人，也属于个人信息。依据《个人信息保护法》的规定，作为数据主体的用户享有数据处理的知情决定（第 44 条）、查询复制（第 45 条第 1、2 款）、更正补充（第 46 条）、删除（第 47 条）以及数据可携带（第 45 条第 3 款）等权利。其中，数据可携带权是指数据主体有权要求个人信息处理者提供结构化、常用且机器可读格式的个人数据，不受阻碍地将这些数据传输至其他个人信息处理者。数据可携带权使得数据主体便捷地改变其服务提供者，从而促成数据在不同竞争主体之间的流动，增强市场竞争。〔54〕

第二，基于消费者权益保护的数据开放。有些数据涉及物权人或消费者的合法权益。数字经济以使用权或有限所有权而非所有权为主的市场结构冲击着传统物权观念，也逐渐消解了消费者

〔51〕 参见崔国斌：《新酒入旧瓶：企业数据保护的商业秘密路径》，载《政治与法律》2023 年第 11 期，第 13 页。

〔52〕 参见梁志文：《知识产权法教程》，南京师范大学出版社 2024 年版，第 295－296 页。

〔53〕 最高人民法院指导案例 264 号（2025 年）。

〔54〕 参见〔波兰〕马里厄斯·克里奇斯托弗克：《欧盟个人数据保护制度：〈一般数据保护条例〉》，张韬略译，商务印书馆 2023 年版，第 241 页。

权益保护。例如，消费者购买的某些品牌汽车安装有加热座椅，但只有支付定制费后才能使用。〔55〕有些新能源汽车搭载的电池并未配套出售，而是租赁给消费者。这些侵蚀物权或消费者权益的做法之所以能够实现，是因为操作上述功能的相关数据由制造商实际控制，如汽车各种传感器生成的数据大部分存储在由制造商控制的云服务器上。这不仅让汽车制造商有能力监控车辆的使用，如依据车辆行驶里程或使用时间而锁定电池，也有可能让其规避相应的责任承担，如交通事故或汽车故障分析需要使用车辆性能、行驶轨迹以及驾驶员操作记录等数据，还有可能让其锁定消费者。例如，车主传统上可以自主选择汽车维修商，甚至自行维修，但现在却只能选择特许维修商，因为汽车维修需要调用的相关后台数据由制造商专供其指定的维修商。

物联网时代，产品制造商可能会采用硬件手段来锁定用户，如故意使用特制螺母和螺栓或者在主板和其他组件上使用焊接，也可能会采用限制维修手册、诊断软件和工具的获取而使维修变得困难，还可能使用数字手段来阻碍未经授权的维修，如使用加密、远程服务器和其他数字工具来保护其专有软件，防止未经授权的维修。产品购买不再是买卖关系的终点，而是开始，从而确保消费者向制造商提供源源不断的现金流。〔56〕上述维修手册、软件和各种数据都因制造商采取保密措施而实现事实上的私有控制，并以商业秘密保护为由不予提供。但是，消费者有合法权益要求制造商提供产品的相关数据，它属于《消费者权益保护法》所保障的消费者知情权（第8条）、自主选择权（第9条）、获取信息权（第13条）等，具有强制执行的效力。而且，有权要求数据提供的主体，不应局限于消费者，而必然要涵盖辅助消费者的经营者，如汽车维修商和配件提供商。

当然，上述数据也确实属于制造商合法的商业秘密，向消费者或物权人强制开放的数据，并不能因此而丧失保护。这构成了秘密数据接收者（即消费者及必要辅助者）的法定义务。即，对合法获取的数据，接收者不得违反法律规定或合同约定而披露、使用或者允许他人使用。秘密数据的接收者应履行如下义务：其一，数据用途目的合法，即服务于法律保护的合法利益，如消费者权利。它也包括为这些合法权益提供服务的行为，从而保障第三方为消费者提供竞争性服务，促成竞争性市场结构的形成。然而，超出前述目的而使用数据的行为，如创建竞争产品，或为此目的与其他人共享，则不再具有合法性。其二，遵守权利人采取的技术保护措施，这是商业秘密保护的基本要求。其三，删除义务，即在实现合法目的之后，不得保留前述数据。另外，数据权利人具有法定拒绝权。权利人可要求数据接收者满足必要条件才向其披露，必要的技术和组织措施包括保密协议、访问协议、技术标准以及行为准则等。〔57〕

第三，基于自主经营权保护的数据开放。作为经营者的用户也享有与个人信息可携带权类似的合法权益。例如，短视频制作者将其作品转换到其他视频分享平台，商家将其商品或服务投放于不同的电子商务平台，都涉及相关数据的流动。欧盟《数据法》规定，非个人数据的可携带权

〔55〕 See Joao Marinotti, *Defragging Ownership: How Corporations Sliced, Diced, and Sold the Bundle*, 100 Indiana Law Journal 1931, 1938 (2025).

〔56〕 Ibid, p. 1956.

〔57〕 See Ulla-Maija Mylly, *Trade Secrets and the Data Act*, 55 IIC-International Review of Intellectual Property and Competition Law 368, 385 (2024).

也属于不可放弃的权利，〔58〕并且还规定欧盟数据库指令下的特殊权利保护不适用于物联网产品或相关服务下获取或生成的数据。〔59〕正如最高人民法院第 263 号指导案例裁判要点指出的，竞争性平台向用户提供关联账号服务，经用户授权后转移其关联平台上获取的数据，“为用户在合理范围内处理该数据提供便利，未扰乱市场竞争秩序的，不构成不正当竞争行为”〔60〕。但是，要明确指出的是，应区分竞争性平台使用的两类不同数据：用户账户及用户发布的各项信息（如文章、图片、音乐等）数据，消费者评价、转发、点击数量等平台提供服务时生成的数据。后者体现了平台的实质性投入，属于受法律保护的经营性利益，不可自由抓取或使用。在第 262 号指导性案例中，虽然被告以案涉短视频由用户自行上传为抗辩事由，但该观点并未得到法院的支持，其重要原因在于其未经许可抓取、搬运了大量的用户评论等数据，实质性替代了原告的产品和服务。〔61〕

（二）代码层的数据开放：维护正当竞争秩序与公共利益

数据具有规模效应，其规模回报递增效应推动了超大型企业的增长，因为它们能够从规模数据中提取更多的价值。〔62〕一方面，数据的网络效应和规模经济保护那些占主导地位的企业免受竞争，使它们能够获取市场产生的大部分剩余价值；另一方面，数据的非竞争性催生了企业间活跃的数据市场，它们通过向其他企业提供数据服务来盈利。〔63〕数据交易不是为了使竞争对手模仿自己的产品或服务，而是增加开发新产品和服务的机会。例如，车辆行驶轨迹数据不仅是导航服务的重要组成部分，也能为保险公司制定保费标准提供更准确的风险状况。从这一角度来看，数据洞见解封数据代码的语义，但它需要经由代码层的数据交易才可实现，故数据开放对促进数字经济健康、快速发展与维护公共利益极其重要。

第一，公开数据集合保护中的数据开放。数据复用性的特点使得数据财产的权利配置不同于专利权等排他权机制。未经授权而广泛地使用专利技术，是以与专利权人相竞争的方式来开展业务，必将侵蚀专利权人的利润回报。然而，数据产品的利用将产生新的数据洞见，增加开发新产品和服务的机会，这就需要在制度上维护一定程度的数据开放空间。以禁止不正当竞争与商业秘密保护的模式来保护数据产品，其本质属于赋予数据产品防御性权利，为数据开放提供了一定的自由空间。防御性权利本身并不禁止数据的合法获取、使用，而是针对“竞争损害”后果。因此，划定开放空间的关键在于如何界定“竞争损害”。

“竞争损害”在我国司法实践中常常用“实质性替代”来表达，但应该区分两类不同的数据产品权利人。其一是专职型的数据权利人，其主要业务是收集、清洗、管理、交易数据，属于数据标注产业从业者。其二是辅助主要业务而生成数据的权利人，其数据产品是电子商务平台、视频分享平台、社交平台等主要业务的辅助或附随产物。假设权利人甲和使用者乙不是直接竞争

〔58〕 See Art. 7 (2) of Data Act. 该条规定：“任何对用户不利，排除适用、降低或改变用户在本章下权利效力的合同条款，均不具有约束力。”

〔59〕 See Art. 43 of Data Act.

〔60〕 最高人民法院指导案例 263 号（2025 年）。

〔61〕 参见最高人民法院指导案例 262 号（2025 年）。

〔62〕 See Maryam Farboodi& Laura Veldkamp, *Data and Markets*, 15 Annual Review of Economics 23, 29 (2023).

〔63〕 Ibid, p. 32.

者，乙利用该数据产品而开发出新的产品或服务，不与甲在同一市场竞争。表面上看，这属于甲对其数据产品预期的许可收入损失如何定性的问题，如果甲可以合法主张完全的预期收入损失，则它与排他权规则并无二致。〔64〕这一假设案例应依据不同情形而有所不同。例如，某时尚品牌基于社交平台的热点话题数据而设计出畅销的潮牌服饰，其是否属于竞争损害，应该依数据产品提供者不同而有所区别。如果数据提供者是社交平台，则服饰设计的用途并非其主要业务，故并无直接的竞争损害；如果数据提供者属于数据标注产业从业者，则因许可收入属于其合法竞争利益而构成竞争损害。

第二，秘密数据集合保护中的数据开放。不同于公开数据集合，竞争损害要件往往并非秘密数据集合纠纷的争议焦点，关键之处是数据获取行为的合法性。法律允许的未经授权而合法获取商业秘密的事由主要有二：反向工程与独立研发。然而，这两者均无益于秘密数据的获取。因为反向工程需要以合法获取产品为前提，而从信息服务（如“微指数”）中无法通过技术手段复原数据。类似困境也存在于独立研发。因此，基于非同业竞争者的数据开放考量，秘密数据保护有必要确立新的法律规则，即强制许可。

专利法等法律已经建立了强制许可制度，但现行商业秘密法尚缺乏相应规则。新冠疫情引发的公共健康危机引发了人们对药品相关商业秘密颁发强制许可的讨论。〔65〕欧盟《数据法》第4、5条的规定类似于商业秘密强制许可机制，其前提是数据接收者须采取保护秘密的合理措施。〔66〕作为反垄断和权利滥用的矫正手段，强制许可应该适用于超大型平台或优势企业的数据拒绝许可等权利滥用情形。如前所述，数字经济时代的锁定效应严重影响相关市场的竞争结构，如汽车厂家对车辆性能、故障等数据采取保密措施而不对独立维修商开放，具有锁定消费者的效应。下游市场的经营者需要获取、使用设备制造商持有的数据来开发其他产品或提供服务，强制许可机制将增强数据的可交易性，有助于释放机器生成的秘密数据。〔67〕当然，强制许可机制不得侵犯数据涉及的他人合法权益，如构成个人敏感信息的健康数据，权利人可以依法拒绝其共享。〔68〕

欧盟《数据法》还对数据许可合同可能存在的不公平条款进行了干预，确保数据开放的条件对双方公平。类似于标准必要专利，其第8条规定数据开放协议“应在公平、合理且非歧视条件下进行”。鉴于合同通常处于保密状态，数据接收者难以评估获取数据的条款是否公平，该法还要求权利人证明，在缺乏合同条件信息的情况下，合同条款不具有歧视性。〔69〕当一方处于更强势地位时，其可能利用这种地位对数据合同中的另一方造成不利影响，这会导致数据访问变得难

〔64〕 See Daria Kim, *No One's Ownership as the Status Quo and a Possible Way Forward: A note on the public consultation on Building a European Data Economy*, 13 Journal of Intellectual Property Law & Practice 154, 162 (2018).

〔65〕 See David S. Levine & Joshua D. Sarnoff, *Compelling Trade Secret Sharing*, 74 Hastings Law Journal 987 (2023).

〔66〕 See Tanya Aplin, *The Data Act and Trade Secrets: An Experiment in Compulsory Licensing*, in A. Sattler & H. Zech eds., The Data Act: First Assessments, Universität Trier, 2024, p. 101. 需要指出的是，欧盟《数据法》的数据共享规则并不适用于处理后的数据。See Tanya Aplin, *The Data Act and Trade Secrets: An Experiment in Compulsory Licensing*, in A. Sattler and H. Zech eds., The Data Act: First Assessments, Universität Trier, 2024, p. 91.

〔67〕 See Daria Kim, *No One's Ownership as the Status Quo and a Possible Way Forward: A note on the public consultation on Building a European Data Economy*, 13 Journal of Intellectual Property Law & Practice 154, 162 (2018).

〔68〕 参见李世刚、郝智鑫：《个人健康数据共享的法益平衡与边界厘定》，载《甘肃政法大学学报》2025年第4期，第32页。

〔69〕 See Rec. 38 - 41 of Data Act.

以实现或根本无法实现。因此，某些不公平的合同条款不应具有约束力。〔70〕

第三，企业与消费者、公共部门间的数据开放。不同于企业间的数据市场，消费者与企业间的数据市场体现了数据的正、负外部性特征。一方面，如果企业利用数据改进其产品和服务，或更有效地为消费者提供服务，则部分价值剩余归消费者所有并提高了其福利。另一方面，如果数据赋予企业过大的市场权力，或相对于消费者的过度优势，数据可能会损害消费者（如数据歧视）甚至社会的整体福利。〔71〕在缺乏具体技术支持时，消费者权益保护法的路径难以消解数据的负外部性，它迫切地需要公共部门的监管来审查消费者数据市场的公平性。除此之外，传染病防治等公共健康事件、自然灾害等突发事件都涉及公共利益，公共部门有使用某些数据的特殊需求，以履行其公共利益下的法定职责。〔72〕因此，需要综合性的数字立法改革进程，来促进企业与公共部门、社会公益组织之间的数据开放，即规定企业与政府数据共享义务。〔73〕

六、结　论

数据是对信息的记录，是符号意义上的信息编码，封装了人类可以理解的语义，其因非竞争性特征而具有共享的属性。大数据可以通过解释、赋予意义而形成新的信息，进而产生新的知识。它超越了单个数据所承载的意义，是数据外部性及经济价值的体现。从这一点来看，将数据与石油、黄金进行类比并非完美比喻。尽管它们都是经济发展的基础性资源，但它们在物理或法律属性方面区分明显，包括是否属于可再生资源、是否具有非竞争性以及经济价值来源等。这些差异意味着黄金、石油的财产权立法经验不能直接套用于数据产品。这也是数据财产能够在现有法律框架下得到承认或保护的同时，又给现行法律规则的适用带来困惑之原因。

数据产权制度有三大核心支柱。首先，从权利客体来看，数据集合或者从数据集合中推断、派生出来的成果，构成了受保护的、分属代码层和语义层的数据产品。《反不正当竞争法》“数据条款”中“经营者合法持有的数据”，不是指单个数据。其次，从权利保护路径来看，应依数据产品的属性不同而提供不同的保护。语义层数据产品之信息服务具有实时性、动态性特点，契合反不正当竞争法上的禁止盗用原则。代码层数据集合是数据复用性的重要载体，也是数字经济发展的核心驱动力，其法律保护则依公开数据与秘密数据不同而有所不同。反不正当竞争法的行为规制落脚于“竞争损害”，较好地平衡了数据集合的私有控制与复用性之间存在的张力关系。最后，作为数字经济发展的关键支柱和黏合剂，数据集合意味着规模性和上下文依赖。它不是数字市场上交换的信息流，而更类似于电缆与路由器，具有基础设施的属性。〔74〕因此，数据开放属于数据产权制度的关键组件，它体现为语义层用户权利的保障以及代码层正当竞争与公共利益的

〔70〕 欧盟《数据法》第 13 条详细规定了不公平合同条款的审查规则。

〔71〕 See Maryam Farboodi & Laura Veldkamp, *Data and Markets*, 15 Annual Review of Economics 23, 32 (2023).

〔72〕 参见梁志文：《数据慈善的基本范畴与法律框架》，载《上海交通大学学报（哲学社会科学版）》2022 年第 2 期，第 63 页。

〔73〕 欧盟《数据法》第五章即第 14—22 条对此作出了详细规定。

〔74〕 See Elettra Bietti, *Data Is Infrastructure*, 26 Theoretical Inquiries in Law 55, 56 (2025).

维护两大方面。

立法应如何回应数据产权与传统财产权制度的复杂关系？制定数据产权保护的综合性法律是一项极具诱惑的选择。然而，数据产品的不同属性具有不同的保护需求，并大体能够在现有的法律框架下得以实现；即使是数据产权的综合性立法，也不得不分别指向相应的现有法律框架。《反不正当竞争法》“数据条款”的增订是制度上恰当的拾遗补阙路径，建立统一的数据财产登记制度也是其中重要的一环。商业秘密保护法、消费者权益保护法、反垄断法、个人信息保护法、突发事件应对法、突发公共卫生事件应对法等法律均需要综合性的数字立法改革进程，在数据保护与合理流通之间权衡，最终促进数字经济健康发展。

Abstract: Data is the encoding of information in the syntactic sense, and unlocking the code to release stored semantics can yield different information, which reflects the reusability of data. The reusability of data is the key factor in the development of the digital economy. The object of data property rights is not individual data, nor raw data or pre-processed data, but a data product, namely data sets or information services provided based on insights from the data sets. The economic value of information services is closely related to real-time and dynamic nature, and is protected by the Unfair Competition Law through the principle of prohibition against misappropriation. For legally sourced data sets, how to protect the right should be distinguished between public data and secret data. The former is protected under the data provision of the Unfair Competition Law, based on criteria such as substantial investment, substantive substitution, and illegitimacy of conduct; the latter may simultaneously satisfy the requirements for trade secret protection, resulting in normative concurrence. Restriction rules on data property rights are mainly embodied in the data opening obligation. It does not mean making data publicly available to society, but sharing data with specific subjects under statutory conditions, including user rights at the semantic level, maintenance of fair competition at the code level, and public interests.

Key Words: the syntactic level of data, the semantic level of data, data provision of the Unfair Competition Law, data product, prohibition against misappropriation

（责任编辑：王慧群）