

网络爬虫行为对数据资产确权的影响

李 帅*

内容提要：数字经济模式下，商业竞争手段日趋多样。通过爬虫行为获取同业经营者线上数据并作营利用途，除有不正当竞争之嫌外，还构成对传统市场机制的严峻挑战。依据不同标准，可将技术爬取的数据划分为多种类型。受到爬虫不当行为影响的权利在边界范围、主体身份以及法律属性等方面呈现复合特征。从本源看，受损事实的发生多是因相关数据性质不明，从而使其衍生权利的界定面临障碍。以“数据资产”概念的提出为基础，结合公私法域多维视角分析其“权利性”，并最终提出网络数据资产的确权建议，是通过法律途径规制不当数据爬取行为的关键一环。

关键词：网络爬虫 数据资产 不正当竞争 数据安全

数字经济时代，数据作为一种新的生产要素，在预判市场走向、调控交易行为、影响各主体权利义务等方面发挥着重要作用。市场参与者特别是互联网企业，通过收集、存储、挖掘、加密、交易等一系列行为，充分发挥数据作为生产性资源的经济价值，更有企业直接将数据视为一种资产。与传统财产性资产不同，线上数据不具备明显的有体性，而以无形、可共享及可传输为主要特点，^[1]这也就导致商业竞争向新的方向发展，并以争夺数据占有状态为主要形式。以下两则典型案例的核心内容均是一方主体通过线上行为获取另一方数据，但法院却作出截然不同的判决，其原因不仅关涉市场竞争与交易秩序的规范化，更与数据获取行为本身的合法性、数据的法律属性等问题密切相关。

案例一为 hiQ vs. LinkedIn。2017 年，同为职业类信息网站运营者的 hiQ Labs 公司（以下简称“hiQ”）将 LinkedIn 公司诉至法院，起因是后者通过后台程序阻止前者获取其客户数据。据调查，hiQ 将其商业模式基本建立在对 LinkedIn 用户数据的分析上，所依托的技术也以数据的爬取和挖掘为主，直至 LinkedIn 采取反爬措施。对此，法院判决 LinkedIn 不得阻止 hiQ 进入、

* 李帅，北京外国语大学法学院讲师。

[1] 参见刘金瑞：《数据财产保护的权利进路初探》，载《中国信息安全》2017 年第 12 期。

复制并使用其网站的公开信息，亦不得采用技术措施进行阻碍。^{〔2〕} 2019年9月，美国第九巡回上诉法院对该案作出二审判决，认定 hiQ 公司从 LinkedIn 上抓取公开数据的行为不构成违法。案例二为“酷米客”vs.“车来了”。2013年起，深圳市谷米科技有限公司（以下简称“谷米公司”）发布并运营一款名为“酷米客”的实时公交 App，通过与公交集团协议安装的定位系统获取公交车辆的位置信息。2015至2016年间，武汉元光科技有限公司（以下简称“元光公司”）为提高其 App“车来了”的信息查询准确度，利用爬虫软件获取“酷米客”服务器中的实时数据，日均达到300万至400万条。基于此，“酷米客”将“车来了”诉至法院，法庭经审理认为被告“车来了”运营方构成不正当竞争，应立即停止侵权行为并依法进行赔偿。^{〔3〕}

两则案例相反的认定结果，除了因为不同经济体制下法律功能与目标的差异之外，最关键的原因还是爬虫行为在本质、细节等方面存在区别从而使行为外部性的实质不同。以此为背景探讨网络爬虫行为，分析其引发的权利义务变动，能够有效深入问题核心——网络数据的法律属性。

一、何为“网络爬虫行为”

网络爬虫是一种自动获取互联网信息的程序或脚本。可以说在大数据时代，除直接通过用户采集之外，另一大数据来源就是使用网络爬虫采集公开信息。常见的爬虫手段包括有构造合理的 Http 请求头、设置 Cookie、使用代理等。作为一项互联网技术，爬虫行为本身并不违法，甚至很多信息类网站的基础技术就是爬虫程序。然而，当技术中立受到商业偏好的影响时，诸如“同业爬虫”等非正当行为则对存储在云空间中的各类数据形成安全性冲击。实践中，爬虫行为正当性的判定，要在企业竞争保护和数据权利保护的双重思想指导下展开，从而全面实现数据关联利益。

（一）两则案例中的爬虫行为定性

在案例一中，一方面，hiQ 通过爬虫程序获取的数据为 LinkedIn 对外公开的内容，且大部分系由用户直接输入、未经平台加工，所以 LinkedIn 在本质上仅是数据收集者和占有者。此外，hiQ 实施的数据访问行为与一般用户无异，仅在获取的数量上多于后者，但也不足以构成限制其爬虫行为的理由。另一方面，LinkedIn 的主要经营业务是提供职业社交服务，为业内公认的绝对支配者，掌握丰富的职业关系信息；相比之下，hiQ 从事的主要是企业员工行为测评工作，并依托员工职业生涯中的自评或他评信息为雇主及人力资源部门提供“员工离职风险分析”服务，在经营项目和目标群体上与 LinkedIn 均无冲突，不属于同业竞争者，所以除有正当理由外，不应限制 hiQ 作为市场活动参与者的正当数据行为。

在案例二中，“车来了”运营商爬取“酷米客”数据，虽然获取的也是后者实时公开的内容，但与案例一却有着本质上的区别。其一，两款手机 App 的运营商均为用户提供实时公交定位服

〔2〕 See hiQ Labs, Inc. v. LinkedIn Corp., 273 F. Supp. 3d 1099 (N. D. Cal. 2017).

〔3〕 参见深圳市中级人民法院（2017）粤03民初字822号民事判决书。

务，商业模式基本相同，存在明显的竞争关系。本案中，元光公司爬取数据的目的就是得到更精准的信息，扩大市场占有份额，可以判定其实施了不正当竞争行为。其二，在数据性质方面，虽然公交车辆运行路线、运行时间等信息仅系客观事实，但经人工收集、分析、编辑与整合之后，就可为 App 的商业运行带来可观效益，所以实际上已具备无形财产的特征。谷米公司系“酷米客”App 设计者，对相应数据除了享有这种特殊的所有权外还是其著作权权利人，因而对于软件所包含的核心数据，未经权利人同意不得私自爬取并用于经营行为。

（二）网络爬虫不当行为的判定标准

数字经济时代，爬虫行为的存在已是普遍现象。在不违背交易秩序和竞争规则的前提下，市场主体对爬虫基本持默许态度，亦有经营者之间通过内部协议允许互相爬取数据的先例。^{〔4〕}此外，《电子商务法》第 31 条规定平台经营者应当“记录、保存商品和服务信息、交易信息，并确保信息的完整性、保密性、可用性”，亦是强调“保密”与“可用”并不矛盾的精神，侧面反映出监管的目标既在于维护网络数据安全，又关注数据多维价值的实现。因此，以“爬虫行为正当性高度依赖数据价值属性”这一特征为基础，可将判定网络爬虫不当行为的考量因素总结为以下两个方面：

1. 数据来源及目标用途

基于商业目的，爬虫技术的应用不仅局限在获取已公开的数据，更多是为了得到对方未开放的内部数据或后台数据，在数字经济过程中则是爬取企业运营所依赖的基础信息。例如，金融爬虫重点获取用户的真实姓名、消费金额记录、信用借贷记录等既关涉个人隐私、又属于商业秘密的信息内容，这与搜索引擎爬取网站界面数据的行为完全不同，其负面性根源在于信息渠道的非正当。

但是，并非所有未涉及内部数据的爬虫行为都是正当的，亦即，合理来源要与合理用途相结合。具体到实践中，判断爬虫行为正当与否，应当考虑以下因素：第一，爬虫获取的数据被用于同业竞争还是非同业经营。实践案例除了前文提及的“酷米客”与“车来了”纠纷之外，还有商业银行信用卡中心之间、互联网现金贷 App 之间相互爬取后台数据，基于对方的放款额度或授信额度，确定本商业实体对自身用户的信贷金额，这些都属于利用同业经营者的数据资源从事相同经营活动，且客观上未履行自身作为金融机构对用户的资格审查职责，故应当被定性为爬虫不当行为。第二，爬虫行为的社会效果是利大于弊，还是弊大于利。当前，大量互联网抢票软件利用爬虫程序在铁路购票及航空购票网站持续点击，实时抢购平台余票。这一操作虽然可以实现加价抢票效果，但实际却影响到网站使用流畅度和所有消费者公平购票的权利，因而属于爬虫不当行为。

2. 授权获取方式及数据使用范围

《网络安全法》第 41 条规定：“网络运营者收集、使用个人信息，应当遵循合法、正当、必要的原则，公开收集、使用规则，明示收集、使用信息的目的、方式和范围，并经被收集者同

〔4〕 例如，诸多互联网平台与百度、Google 等搜索引擎运营公司签订内部协议，允许后者爬取其数据作为用户搜索结果，并链接至原网站平台。

意。”基于法律对数据来源主体的权利保护规则，虽然网络爬虫并非直接从用户手中获取数据，但理论上同样应在事前取得授权。

当前共享经济与电子商务协同发展，互联网实时线上数据体量巨大，用户关注的重心往往与购买商品、享受服务以及获取资讯相关，而对平台上的消费者协议内容一带而过，并通过勾选的方式将“同意”与“授权”一次性作出。实践中，有网络运营商采用各种方式扩大用户信息的获取和使用范围，很大程度上侵害了用户个人信息安全。常规互联网平台在获取授权过程中尚如此，爬虫程序获取用户授权则更具“表述模糊性”与“方式隐蔽性”。例如在商业竞争中，有经营者希望获得更多同业商家持有的用户信息，则要求自己平台上的用户提供在其他同类平台注册时的用户名和密码，以便获得更好的服务。这种做法虽然得到了用户授权，但大部分授权人并不知晓实际获取信息的主体是爬虫程序，被爬数据的真实范围远大于用户知情范围，且数据的最终用途也与获取授权时告知的内容不符，这就使得数据来源主体的知情权和依托数据产生的其他权利受到严重侵害。

二、数据权利内涵新解及其与爬虫行为的相互关系

在网络爬虫行为中，线上数据作为一系列活动的标的，其所承载的内涵与传统数据相比实现了量级化突破。这种行为对象上的拓展，直接导致相应权利客体发生变化，受爬虫行为影响的权利种类和性质亦有一定改变。因此，对数据权利属性进行研究，既有利于厘清爬虫行为中受损的权利内容，为相关制度的完善提供依据，同时更是制定数据资源确权、开放、流通、交易相关制度的起点和基石，有利于完善数据产权保护制度。^{〔5〕}

（一）以网络为媒介的数据种类界分

传统经济时代，数据的分类标准主要包括产生时间、涵盖范围、权利主体等，相应地，其类别有当前数据与历史数据、宏观数据与微观数据、个体数据与群体数据等。而现今，网络技术丰富了数据的产生、流转、利用与交易途径，数据分类方式和分类标准不再局限于二维层面，其复合属性日益明显。

首先，以数据的生成是否需要网络平台加工为标准，可以将线上数据分为用户个人信息、实时位置信息等原始数据，以及交易记录信息、活动轨迹信息等加工数据。其中，加工数据涉及多重主体的贡献迭代问题，与原始数据相比附加了更多的社会价值，因而利用规则也相对复杂。具体而言，实践中对加工数据的获取与应用，要额外考量数据处理的经济成本，并以加工主体的意愿为决定因素之一。

其次，以数据的存在形式为标准，可分为文本数据、状态数据与新模态数据。文本数据指的是可经记载、描摹或复制，将内容以有形形式存储的信息，包括但不限于各种文字、符号、图像及其组合。状态数据是以非文本形态生成并存储，具有较强时效性的信息，例如实时地理位置信息等。人工智能技术背景下，电子身份识别所需的指纹、瞳孔、声音等信息具备了高度代码化特

〔5〕 参见李爱君：《数据权利属性与法律特征》，载《东方法学》2018年第3期。

征，可独立成类，即为新模态数据。实践中，数据的特殊形态常被用作不当获取行为的抗辩理由，其原因主要就是信息类型化的不甚完善。所以，扩展数据的传统认知范畴，能够有效划定爬虫行为的对象领域。

再次，以数据内容的重要性程度和生成环节为标准，可分为核心数据与周边数据。此分类标准下，数据子类包含的具体内容并不确定，需要结合实际情况，分析相关数据是否体现核心权利义务关系，以及该数据是否系数据生成、流转或存储行为的附属产物。具体言之，作为数据行为的目标产物，且蕴含社会、经济价值的数字通常被认定为核心数据，爬取时需要得到数据主体的知情同意和授权；除此之外，数据流通过程中产生的程序信息、档案信息或其他信息，不具有明确的使用价值或对行为人权益无实质性影响，通常在确认不涉密之后则可供爬取。

最后，以数据主体的身份差别为标准，可以分为用户数据与商家数据。前者系网络用户基于自愿，主动向平台提供的有关个人身份、行为习惯等方面的信息；后者为平台经营者实施商业运营行为产生的新信息，或者从事数据加工行为产生的增值信息等。因为与主体权益密切相关，所以规制此类数据时需要以主体身份的明确为前提，在此基础上深入分析权利种类，将主体与权利相匹配，最终为数据爬取行为的实施者提供意见征询对象及征询内容。

（二）数据权利内涵的时代性拓展

伴随数据形态和内容的丰富，依托其生成的数据权利也面临着因应性变革，主要趋势就是权利内涵的进一步充实，以及多种属性权利的并行存在。理论研究进程中，邻接权客体与财产权客体的争议至今尚存，〔6〕而蕴含着人格权内涵的“财产权客体说”作为新生理论，在网络社会中日益显现更适于保障数据主体权益的能力。

当前实践中，立法对数据权利的定性尚未明确，世界范围内亦不存在一部调整各类数据行为的法律规范，仅在个人信息或隐私保护的维度上有所突破。以2018年欧盟出台的《一般数据保护条例》（GDPR）为例，该法案以强化数据主体对个人数据控制为导向，但未能明确划分数据权利项下人格权与财产权的边界，一定程度上掣肘了立法目标的实现。因此有学者认为保护公民的隐私权益，应当更多采取公法风险规制与消费者法保护的框架，而不是寻求一种具有确定性边界的隐私权或个人信息权。〔7〕推及互联网背景下的所有数据范畴，这一理念同样具有其合理性：其一，大量互联网数据系从用户个人处收集所得，且主要来源于用户的消费行为，可界定为消费数据的组成部分，因而以消费者法的思路对相关信息进行保护具有客体的适格性。其二，存储在网络数据库中的信息并不局限于原始的用户信息，还有因流转和加工而不断充实的信息，所以坚持以一种确定的权利名称界定互联网数据，不论对权利属性的全面认知还是合法权利的有效保障而言，都缺乏科学合理的指导意义。

综上，选择具有综合意义的“财产权”概念诠释新时代的数据权利，赋予其人格、身份、财

〔6〕 持“邻接权客体说”的学者主要有林华、秦珂等，参见林华：《大数据的法律保护》，载《电子知识产权》2014年第8期；秦珂：《大数据法律保护摭谈》，载《图书馆学研究》2015年第12期。持“财产权客体说”的学者主要有高富平、齐爱民等，参见高富平：《信息财产——数字内容产业的法律基础》，法律出版社2009年版；齐爱民：《捍卫信息社会的财产——信息财产法原理》，北京大学出版社2009年版。

〔7〕 参见丁晓东：《什么是数据权利？——从欧洲〈一般数据保护条例〉看数据隐私的保护》，载《华东政法大学学报》2018年第4期。

产乃至知识产权等多重内涵，将是数据时代规范爬虫行为的权利设定基础。正如龙卫球教授所言，一旦承认了用户具有数据财产权，那么就会迫使数据使用者主动与数据主体进行商议，如此则改变了用户在数据市场被忽视的境地，使得用户获得了一定的议价能力。^{〔8〕}这在互联网商家竞相爬取用户数据的当下，可以说为最核心的数据主体——个人用户——提供了权益保障的良性思路：既关注了用户在互联网数据活动中的数据来源者地位，认可其对自身数据享有一定的决定权与处分权，又洞悉了用户参与互联网活动的关键意图是成为消费者，所以承认并保障其包括知情权、选择权等在内的消费者权益，将进一步拓展数据权利的内涵和外延。

（三）受爬虫行为影响的具体数据权利

以爬虫时代为背景，当数据获取行为日渐成为互联网主体的常规活动时，窃取他人数据谋取经济利益的行为就可能混同于其他数据获取行为中，给用户、商家、平台的数据权益造成不利影响。

从不同主体在数据来源、应用及流转中的身份差异来看，可将受到网络爬虫不当行为影响的具体数据权益做如下划分：第一，对互联网平台经营者而言，商家之间的数据爬虫不当行为破坏了平台正常的运行规则，在影响数据安全的同时，还可能使用户对平台本身的中立地位产生质疑，从而侵害到平台运营方的经营权，给其商业信誉带来负面影响。当然，如果利用爬虫程序不当获取数据的行为系平台方所为，则前述不当后果应由平台自行承担。第二，对在平台上从事经济活动的商家来说，如果是被爬取方，则其依托相关数据而享有的商业秘密、知识产权将会因此受到侵害；如果是爬取方，则在利用爬取数据进行后续经营活动时，极易因资质欠缺或使用不当而造成用户损失，从而违反现行法律的相关规定、影响自身商誉，严重者还可能构成犯罪。第三，对于互联网用户群体而言，爬虫不当行为的直接对象就是其个人数据，即使这些数据关涉的内容主要是商业活动，但通过大数据技术深入挖掘，很容易精确定位个人身份并形成专属于个体的消费习惯信息，对隐私权、自主选择权等权利造成损害。

以上提及的“合法权益”，其界定思路仍然以传统部门法学视角下的分类标准为主，即在民事权利项下划分具体的权利种类，并按照不同部门法的规定对侵权行为进行处理。这种方式在实践领域沿用已久，虽具备较强的可操作性，但在处理实效和便捷度方面却存在不可否认的缺陷。因此，以互联网爬虫不当行为的频繁出现为背景，法学实践和理论研究都开始寻求一种更为简捷的权益保障路径，这也对一种更具概括性的权利种类提出了明确需求。综上，传统法学研究开始从单纯的“原理型”转向同时侧重原理和对策的多重价值型，并以切实解决当前数据时代的新型问题为主要任务。这在应对爬虫不当行为时，主要表现为对已有权利种类的丰富和对传统法律保护方式的更新。

三、数据“资产型保护”路径的确立

（一）从“泛资产”到“大资产”的概念演变

在金融领域，“泛资产”概念的出现基本始于2010年前后，标志性事件即投资基金逃离传统

〔8〕 参见龙卫球：《数据新型财产权构建及其体系研究》，载《政法论坛》2017年第4期。

实体经济，流向农产品、艺术品等具有资产性质的实物或者影视一类具有资产意义的行业，从而使这些非典型财产获得了资产的属性，被称为“泛资产”。伴随互联网布局的扩大和大数据技术的发展，泛资产的概念亦无法与实践的发展保持完全同步，诸如信息、数据以及由此衍生出的各类电子资料，以及表现为声音、图像、指纹、虹膜等多种模态的生物档案也日益呈现出可供挖掘与利用的经济价值。因此，以泛资产时代的出现为参考和借鉴，笔者提出“大资产”概念，用以指称上文提及的信息和数据等新型财产客体。

之所以称其为“大资产”，而没有在“泛资产”的概念上进行外延扩充，一方面是因为“泛资产”的用法源自金融领域，并且在一定程度上已经成为具有特指意义的术语，所以法学研究需要结合社会需求和学科特征，创设相对新型的术语。另一方面则是基于当前“大健康”“大文娱”等概念的出现，社会公众对此类表述已形成一定的认知和理解，因此用“大资产”一词来统摄现存具有资产性质的事物、行业或产业，具有现实基础。在法学研究中，因大资产而派生出的权利主要是与其相对应的“资产权”，当然，与传统法学研究中学者曾提出的“资产权”概念相比，其在权利的主体、客体、权利义务关系等方面均呈现出较大差异。

传统研究中，资产权主体多为企业或国家，因而问题的探讨较多集中于国有土地使用权以及PPP项目中的权利转让等领域，^[9]并未将其视为公民个体法定权利的一种。因而本节提出构建“资产型”的权利保护路径，首先就要突破现有的理论认知，将资产权主体范围拓展至个人，并赋予其综合性权利的属性，具体来说，综合的主要是物权、知识产权的特性，并且在特定情况下还可能包括人格权属性。

（二）爬虫时代 Robots 协议的失灵现象

Robots 协议，又称“爬虫协议”或“机器人协议”，^[10]是技术界为了实现爬取方和被爬取方之间的意愿沟通而设定的一种计算机程序。^[11]根据中国互联网协会《互联网搜索引擎服务自律公约》^[12]第7条的定义，Robots 协议是指互联网站所有者使用 robots.txt 文件，向网络机器人（Web Robots）给出网站指令的协议。具体而言，Robots 协议的关键内涵在于提示网络机器人哪些网页不应被抓取，而哪些网页可以抓取。^[13]关于此处 Robots 协议的法律属性，我国目前尚无立法明确规定，可供适用的调整型规范仅有前文述及的《互联网搜索引擎服务自律公约》。^[14]鉴于该公约的非强制性特征，实践中商业主体为获取经济利益而违背协议的现象频频

[9] 相关研究例如：占一熙、陈芬辉：《农村集体资产权属的几点思考》，载《农村经济管理》2016年第8期；赵琰：《PPP项目资产权属及会计核算方法》，载《财会学习》2017年第22期；郭滨辉：《PPP项目的资产权属、会计核算及税务处理》，载《财会月刊》2019年第3期。

[10] Robots 协议诞生于1994年2月，由荷兰软件工程师 Martijn Koster 创建，据称起因是 Koster 的服务器遭受到不良爬虫的爬取而致使服务堵塞。不久该协议即成为现存及未来的网络爬虫都被期望遵守的行业惯例。参见瞿森、杜承彦、谭晓明：《数据之争：网络爬虫涉及的法律问题（二）》，载微信公众号“知产力”。

[11] 参见罗刚：《网络爬虫全解析：技术、原理与实践》，电子工业出版社2017年版，第45页。

[12] 《互联网搜索引擎服务自律公约》系2012年11月由百度、即刻搜索、盘古搜索、奇虎360、盛大文学、搜狗、腾讯、网易、新浪、宜搜、易查无限、中搜12家企业联合发起并签署的行业自律公约。该公约共4章22条，适用于中国互联网协会会员单位和自愿加入的互联网从业单位，并且倡议其他从业单位积极遵守。

[13] 参见瞿森、焦晨恩：《数据之争：网络爬虫涉及的法律问题（一）》，载微信公众号“金杜研究院”。

[14] 《互联网搜索引擎服务自律公约》第7条第1款规定：“遵循国际通行的行业惯例与商业规则，遵守机器人协议（Robots 协议）。”

发生。

在日趋强烈的经济利益导向下，仅有的行业性规范很容易被突破，Robots 协议面临严重的失灵问题：爬虫行为不仅大批量获取其他商业主体的后台信息，还会强行突破对方网站设置的技术措施以实时、无限制地抓取信息。形成这一现象的原因，除了经济效益的驱使之外，主要还在于 Robots 协议并不构成具有法律效力的协议，当前在我国仅被视为业内普遍遵守的“公认的商业道德”，缺乏国家强制力约束，对 Robots 协议性质的界定和详细论述可参见 2013 年百度诉奇虎 360 案判决书。^[15] 但需要注意的是，即使协议不具备强制约束力，上述行为同样可能违反其他现行法律的规定，例如构成《反不正当竞争法》第 12 条中“利用技术手段，实施妨碍、破坏其他经营者合法提供的网络产品或者服务正常运行的行为”。此外，对于强行突破某些特定被爬取方技术措施的，还可能构成刑事犯罪。这里对违法抑或犯罪行为的处理思路，仍然是将被侵犯的权利进行分离，结合上文所述两则案例来看，即司法机关在裁判过程中并未明确界定受损数据的权利性质，只是通过所有权或著作权模式进行保障，尚不存在一个相对复杂、综合的保障方式。

（三）新型数据保护思路的确立

从目前已发生的数据纠纷情况来看，其化解方案主要还是从既有法律体系中寻求解决办法，最多援引的法律规范即为《合同法》《反不正当竞争法》《侵权责任法》以及知识产权法律。例如，数据合同法路径救济的前提是争议主体之间存在预先的合同安排，但是我们发现这种合同安排无论如何周密也只是一种债的保护，本身不具有排他性，无法对抗来自第三人的数据加害，而在现实中数据加害往往就是来自企业数据合同关系之外的第三人侵入或者非法利用。^[16] 与此同时，相对松散的 Robots 协议无法起到绝对的约束作用，这就更加呼吁一种具有排他性，且在权利保护顺位上占据优先级别的保护方式。

在法学领域确立“大资产”概念并通过资产权路径保护互联网用户、商家及平台的数据权利，成为当下新型权利保护的可行思路。具体而言，就是应当同时关注社会的风险规制和消费者的选择与预期，^[17] 综合运用公法和私法理念。落实到实践中，即需要设计一种兼具物权与知识产权属性，并同时体现人格权与财产权价值的权益，建议采用“资产权”的表述，其保护路径应适当借鉴欧盟和日本等国家（地区）的“放松收集环节，加强使用监管”的数据处理规则及相应立法例，强化数据加工处理阶段的保护要求。

四、数据资产确权背景下相应法律保护措施的展开

（一）推进统一的综合性数据资产保护立法进程

互联网数字经济时代，爬虫不当行为对数据安全造成的负面影响实际上只是冰山一角，更多行为诸如数据交易、数据不当应用等也对存储于网络数据库中的海量信息产生安全性威胁。因

[15] 参见北京市第一中级人民法院（2013）一中民初字第 2668 号民事判决书。

[16] 参见龙卫球：《再论企业数据保护的财产权化路径》，载《东方法学》2018 年第 3 期。

[17] 参见前引 [7]，丁晓东文。

此，在国家层面制定统一的数据保护类立法，并以各类个人数据、商业数据等为主要保护对象，成为当前该领域内的立法发展趋势。在 2018 年 9 月公布的“十三届全国人大常委会立法规划”中，《数据安全法》已被列为“条件比较成熟、任期内拟提请审议的法律草案”，标志着具有一般性法律规范性质的数据法制定工作已经提上日程。

从法律名称来看，《数据安全法》仅调整与数据安全有关的法律关系，对数据的财产权益等其他法律关系并不涉及，^{〔18〕}但结合实践需求分析，此次立法除了应对社会公共层面的数据安全问题外，还应着力解决网络数据权属不明、保障措施不力的现状。因此，制定中的《数据安全法》应充分考虑不同主体对数据立法的差异化需求，明确在“财产权”概念下保护数据资产的方式：第一，清晰界定受本法保护的“数据”范畴，明确用户、商家及平台享有的数据财产权在范围、时限上的差异；第二，规定各数据主体的责任和数据处理、利用规则；第三，授权设立专门的数据保护监管机构，监督数据的收集、使用与流转；第四，明确侵犯数据财产权所应承担的责任。

（二）区别不同阶段的数据保护和处理规则

数字经济下，大数据的多维度、多层次应用给传统网络法中的同意原则、自主可控原则、透明度原则、匿名原则、最小化原则等法律原则带来严重挑战。首先，在数据收集环节，如今基本可以做到只要接触网络就会留有痕迹，即凡是“触网”的数据必然遭到收集。所以，要求数据的收集必须经权利人同意，甚至必须通过特定程序获取同意，已然不适应于当前社会的发展。在此建议适当放松对数据收集环节的管控和限制，规定只要经合理的明示方式履行了告知义务，即可收集与法定经营权限相关的用户信息。其次，在数据加工处理阶段，商家或平台通常会利用各种分析工具对数据进行挖掘，获取成倍于数据表层含义的信息量。对此，应当严格规范数据的加工、处理与应用环节，本着权利人同意为前提的基本原则，在获得授权的范围内实施数据处理行为，并将数据加工后所得形态、数据流转的去向以及未来可能适用的场域均告知权利人。

可见，在数据行为的不同阶段，法律的保护力度和限制条件应当有所差别，这既是法治发展适应社会需求的基本要求，也反映了法律制度在变革中不断趋于科学、理性的过程。

（三）强化侵犯数据保护的责任追究

数据财产权保护模式下，电子化的个人信息、商业信息具备了产权性质，排他地从属于权利主体，具有直接支配性。换言之，用户、商家的个人数据、消费数据或经营数据等内容，不再只是受到非法侵犯后才能获得救济式的保障，而是在日常生活中就享有与所有权相似的“第一性”或“充分性”，通过法律文本对权利的确认而获得主动式的保障。

在这一变化背景下，应当合理配置数据保护责任体系，强化侵犯数据保护的责任追究机制。一方面，建议以我国现行《物权法》第一编的立法体例为蓝本制定《数据安全法》总则部分，转变传统数据保护的侵权法思路，明确规定数据财产权的实现与保护路径。另一方面，划分数据违法行为的类型与损害程度，规定不同的法定责任承担方式及担责范围，并加强对违法提供、使用

〔18〕 参见史宇航：《今年将推进立法的〈数据安全法〉，有哪些看点？》，载 <https://baijiahao.baidu.com/s?id=1627780949714211002&wfr=spider&for=pc>，最后访问时间：2019 年 8 月 20 日。

或侵犯个人信息和商业数据者的责任追究。

五、结 语

网络爬虫不当行为的日益泛滥，给数字经济的发展和互联网整体环境的优化都带来严重的负面影响。传统侵权保护模式只能起到有限的私权恢复作用，而无法实现公法意义上的行为规制与风险防范目标。以《数据安全法》的立法推进为契机，确立数据资产概念并开拓资产保护路径，是行政法与民法的一次交流与合作，亦是大数据时代法治不断成熟与完备的必然进路。

Abstract: In internet economy era, the means of commercial competition have become increasingly diverse. Obtaining online data from peer operators through the crawler program and using it for profit-making purposes, are largely suspected of unfair competition, meanwhile constitute a serious challenge to traditional market mechanisms. According to different standards, the content of technology crawling can be divided into raw data and processing data, text data and status data, core data and peripheral data, as well as merchant data and user data, etc., which makes the rights affected by crawler misconduct shows composited features in boundary scope dependent subject and legal attribute. From the root cause, the facts of damage happening are mostly due to the unclear nature of the corresponding data, which leads to the obstacles in the definition of derivative rights. Therefore, based on the concept of data assets, combined with the multi-dimensional perspective of public and private jurisdictions to analyze its rights' status, and finally put forward the recommendations for the determination of online data assets, is a key link to regulate improper data crawling behavior through legal channels.

Key Words: web crawler, data assets, unfair competition, data security

(责任编辑: 刘 权 赵建蕊)