

论数据爬虫行为中的占有保护

——兼论对《数据二十条》中持有权的解释

杨 力

【摘 要】我国司法实践多将《反不正当竞争法》第2条作为裁判依据，判令数据爬虫主体因其不正当竞争行为而承担停止不法侵害及损害赔偿的民事责任，此种裁判模式在法律适用主体上存在局限性，并且将原则条款作为裁判规范基础易造成法官自由裁量权的不当扩张。《数据二十条》中的持有权与传统民法上物之占有存在同质性，均包括体素和心素要件。在司法实践中，可类推适用《民法典》物权编中的占有制度为爬虫行为中的数据提供保护。在物之占有中隐含了公开的排他性要件，在数据占有中需将其加以还原并在司法裁判中进行具体判断。体素、心素与公开的排他性这三个要件是建构爬虫行为合法性边界的基础，有利于保障数据保护和数据流通之间的平衡。占有保护请求权包括占有返还请求权和占有妨害请求权，对数据的适用存在特殊性，为数据权益保护提供了合理且周延的救济路径。

【关键词】数据爬虫 体素 心素 公开的排他性 占有保护

【作者简介】杨力，中国政法大学民商经济法学院博士研究生。

【中图分类号】D923.0 **【文献标识码】**A

【文章编号】2097-1125 (2024) 06-0079-20

一、问题的提出

随着数字经济的兴起和发展，数据爬虫技术被广泛运用于各行各业。同时，围绕数据爬虫产生的数据权益侵害案件层出不穷，这类案件在司法实践中大多经由《中华人民共和国反不正当竞争法》（以下简称《反不正

当竞争法》)进行规制,如在刷宝App不正当竞争纠纷案^①、淘宝公司与安徽美景公司不正当竞争纠纷案^②、新浪诉饭友不正当竞争案^③等案件中,法院均认为原告对数据享有数据权益,被告擅自获取、使用涉案数据的行为构成不正当竞争,故判令被告停止对数据信息的获取和使用行为并赔偿原告一定的经济损失。

此种裁判方式看似实现了在爬虫案件中对数据权益的保护,但存在诸多问题。第一,《反不正当竞争法》的调整对象是生产经营活动中存在竞争关系的经营者,即其调整对象必须符合两个基本条件:一是市场活动中的经营者;二是存在竞争关系,能够扰乱市场竞争秩序。^④因此,对非经营者或不具有竞争关系的其他民事主体在数据爬虫中侵害数据权益的情形,无法诉诸《反不正当竞争法》的保护。质言之,《反不正当竞争法》更多是从保护市场公平竞争秩序的公法角度对数据侵害行为进行规制的,而对其他非竞争关系的民事主体侵害数据权益的行为,《反不正当竞争法》没有适用空间。第二,上述判决均以《反不正当竞争法》第2条为裁判规范基础,但该条是对不正当竞争的概念性规定,系经营者在市场交易活动中的基本行为准则,具有高度的抽象性和概括性,由此使裁判者对“商业道德”“诚实信用”的判断存在模糊性和不确定性。同时,对不正当竞争行为是否成立的问题,裁判并非遵循传统,即立足法条构成要件并在充分辨析行为性质的基础上用三段论的方式对具体条文进行解释和涵摄,而是直接从抽象的定义中得出结论,很容易导致裁判者自由裁量权的过度膨胀。^⑤第三,《中华人民共和国民法典》(以下简称《民法典》)作为私权保护的基本法,包含关于数据和网络虚拟财产保护的指引性规定,虽然目前民法规范缺少对数据权益保护的系统性规定,但对在民商事活动中侵害平等主体权益的行为,民法为行为性质的认定、请求权的多元化供给、损害赔偿的精确计算提供了成熟而体系化的制度基础,绕开《民法典》对民商事主体侵害特定权益的行为进行判决并非定分止争的明智之举。

① 参见北京知识产权法院(2021)京73民终1011号民事判决书。

② 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。

③ 参见北京知识产权法院(2019)京73民终2799号民事判决书。

④ 参见孔祥俊:《反不正当竞争法新原理(总论)》,法律出版社2019年版,第173页。在司法实践中,法院在不正当竞争纠纷案件中也会首先论证原被告之间的竞争关系。例如,深圳市腾讯计算机系统有限公司等诉浙江搜道网络技术有限公司等不正当竞争纠纷案,参见杭州铁路运输法院(2019)浙8601民初1987号民事判决书。

⑤ 在司法实践中,对一般的不正当竞争行为,法院基本依据《反不正当竞争法》第二章“不正当竞争行为”中的具体条文对行为是否属于不正当竞争进行判断;对数据爬虫行为,则是将《反不正当竞争法》第2条的定义性规定作为规范基础。

有鉴于此，本文从《民法典》的相关规定出发，立足当下的司法实践，对数据爬虫行为的违法性标准和裁判路径进行讨论，以期以现行民法制度为依托，为数据爬虫行为寻找统一的、适用于所有民商事主体的合法边界，在现有民法规范的视域下为爬虫行为中的数据权益保护找到合理且可操作的规范指引。

二、违法数据爬虫的裁判进路之比较

所谓数据爬虫，指利用自动化的脚本或程序，通过模拟用户访问网站页面、获取网站特定数据的技术手段，即网络用户通过爬虫软件自动抽取平台数据的行为。数据爬虫表现为“网页抓取”（web scraping）、“屏幕抓取”（screen scraping）等数据爬取形式。^①爬虫技术提高了收集数据的效率，促进了数据的分享和流通，但无差别地允许数据爬虫又会造成对他人数据权益的侵犯。因此，寻找数据爬虫的合法边界在数字时代尤为重要，这既关乎对网络空间资源的开发和利用，又关乎民商事主体的权益保护。可以说，数据爬虫的违法性认定属于数据权益保护的核心问题，如何妥当划定数据爬虫的合法性边界事关个人利益及社会公共利益。

（一）我国司法实践中的裁判进路及其弊端

我国司法实践对数据爬虫的规制主要存在两条路径：第一，平台间的数据爬取行为诉诸被泛化的《反不正当竞争法》第2条暨“一般条款”；第二，个人或单位不当爬取数据构成犯罪的，依照《中华人民共和国刑法》非法获取计算机信息系统数据罪或侵犯公民个人信息罪定罪处罚，如马某某等非法获取计算机信息系统数据案，^②凌某某非法获取计算机信息系统数据、非法控制计算机信息系统案^③等。但是，上述两条规制路径均存在不周延之处。

一方面，如果双方主体并非具有竞争关系的经营者，如其中一方是自然人，如果违法行为达不到刑法入罪的程度，则不能对其定罪论处，也不能依照《反不正当竞争法》判令其承担责任；另一方面，即使刑法已对违法者定罪量刑，但违法者在私法上如何承担责任在实践中也存在疑问。例如，在杨某某非法获取计算机信息系统数据案中，^④杨某某因侵害北京玩

^① 参见 Frank Jennings and John Yates, *Scrapping over Data: Are the Data Scrapers' Days Numbered?*, *Journal of Intellectual Property Law & Practice*, Vol. 4 (2), 2009, pp. 120 - 129.

^② 参见扬州经济技术开发区人民法院（2019）苏1091刑初157号刑事判决书。

^③ 参见广东省东莞市中级人民法院（2021）粤19刑终202号刑事判决书。

^④ 参见北京市海淀区人民法院（2016）京0108刑初1084号刑事附带民事判决书。

蟹科技有限公司的平台数据，使其经济利益受损，被判非法获取计算机信息系统数据罪。但是，在刑事附带民事判决中，法院认为附带民事诉讼的提起条件是被害人因人身权利受到犯罪行为侵犯或者财物被犯罪分子毁坏而遭受物质损害，因杨某某的行为并未造成玩蟹公司的直接损失，故驳回附带民事诉讼中原告的诉讼请求。应该说，此种裁判模式有损法秩序的统一性，虽然刑事裁判认定不法者对特定主体的权益侵害已构成犯罪，但在民事裁判中未能相应地认定其行为构成违法，致使被害人无法获得对损害的救济。

同时，在司法实践中将《反不正当竞争法》第2条作为损害赔偿的裁判规范基础也存在问题。该条具有高度的抽象性和概括性，会导致法官自由裁量权的过度膨胀。从法理上看，《反不正当竞争法》的这一规定更接近于标准性或原则性规定，缺乏能够指引具体裁判的规则刚性。^①例如，在深圳市腾讯计算机系统有限公司等诉浙江搜道网络技术有限公司等案中，^②法院判决被告行为成立不正当竞争的实质性裁判理由为：两被告的行为势必导致微信用户对微信产品丧失应有的安全感及基本信任，进而会恶化两原告既有数据资源的经营生态。显然，前述裁判理由并非基于法条的构成要件进行法律解释，欠缺充分的说服力。损害赔偿的计算标准和赔偿范围也依赖裁判者的自由裁量，破坏了法的安定性。

综上，现有司法裁判路径既存在对爬虫行为中数据权益侵害的规制漏洞，又容易导致法官恣意裁判。为妥善解决该问题，在我国民商合一的立法模式下，有必要选择一条对民事主体和商事主体普遍适用的统一标准，以此尽可能地确保对爬虫行为中数据权益保护的周延性与稳定性。

（二）其他裁判进路的不合理性

对违法侵害他人财产权益的行为，在私法上一般通过侵权法进行规制，但数据权益能否被直接纳入侵权法的保护范围不无疑问。在上文提到的杨某某非法获取计算机信息系统数据案中，附带民事诉讼部分之所以被驳回诉讼请求，在本质上是因为侵权法保护范围的局限性。我国侵权行为的一般条款虽然看似采用法国“一个大概括条款”的立法模式，但司法实践和学界均否认了给予权利和利益同等保护的进路，而是在法国式的一般侵权行为条款基础上发展出德国式的法解释论。^③因此，在我国侵权法的适用

① 参见丁晓东：《论企业数据权益的法律保护——基于数据法律性质的分析》，《法律科学（西北政法大学学报）》2020年第2期，第91页。

② 参见杭州铁路运输法院（2019）浙8601民初1987号民事判决书。

③ 参见于飞：《侵权法中权利与利益的区分方法》，《法学研究》2011年第4期，第106页。

上，侵权行为的构成要件因基于被侵害对象的权利还是利益而存在区别，诉诸侵权法的保护须首先明确被侵害权益的性质。如果裁判者将数据定性为类似绝对权的民事权利，则可以直接适用侵权行为一般条款；如果否认了数据的权利性质，则必须检视该侵害行为是否满足其他权益侵害的构成要件。学界一般认为，数据并不像绝对权一样具备“归属效能、排除效能和社会典型公开性”，^①不能直接被纳入侵权法的保护范围。因此，对上述杨某某案附带民事诉讼部分，法院并未直接判令被告承担侵权损害赔偿责任。

诉诸侵权法的救济途径，势必会因不同法官对数据性质认识不一，而对数据爬虫的违法性认定采用不同的构成要件，最终导致类案判决结果不一致。正如学者指出的，侵权法虽然可以提供对企业数据的额外保护，但这种保护也面临不确定性的问题，仍然依赖法律对数据权属问题的界定。^②此外，在数据权益的性质尚未明确的当下，数据权益被侵害在侵权法的视域中只能被定性为纯粹经济损失，而在认定对纯粹经济损失的侵权行为是否成立时，应当对侵害人的“背俗故意”要件进行证明。但是，在数据时代，平台企业对隐藏在电脑后的侵害行为之“背俗故意”的证明也并非易事，故以侵权法的进路对爬虫行为中的数据权益进行保护存在诸多不合理之处。

有学者提出，数据作为无形资产可由知识产权法进行保护。^③然而，被抓取的企业数据大多都缺乏鲜明的独创性，除了少量涉及商业秘密、技术专利的企业数据，大数据时代的企业数据多是反映客观事实的信息，如交通工具的准时性统计、某客户群体的样本分布、网络购物的用户情况、个人信息等，这些数据不具有独创性，不是智力成果，所以对数据爬虫问题不能简单地以专利权、著作权保护的方式加以解决。^④即便支持采用知识产权进路保护数据的学者也认为，将数据作为独立客体纳入知识产权保护范围的观点，仍停留在正当性论证、路径设计等理论层面探讨的程度，至于如何界定数据知识产权的保护范围、如何保障数据知识产权人的权益、如何弥补数据知识产权人的经济损失等问题，尚需漫长的深入研究和审慎

① 于飞：《侵权法中权利与利益的区分方法》，《法学研究》2011年第4期，第118页。

② 参见丁晓东：《论企业数据权益的法律保护——基于数据法律性质的分析》，《法律科学（西北政法大学学报）》2020年第2期，第91页。

③ 参见单晓光：《数据知识产权中国方案的选择》，《人民论坛·学术前沿》2023年第6期，第38页。

④ 参见龙卫球：《再论企业数据保护的财产权化路径》，《东方法学》2018年第3期，第63页。

的司法实证。^①

商业秘密的保护进路也始终无法得以适用。商业秘密在《反不正当竞争法》第9条中被定义为“不为公众所知悉、具有商业价值并经权利人采取相应保密措施的技术信息、经营信息等商业信息”，“秘密性”、“价值性”和“采取相应保密措施”是商业秘密保护的三个构成要件。然而，在多数情况下，平台数据处于开放的互联网架构中，无法满足“秘密性”和“采取相应保密措施”这两个要件，也无法达到“不为公众所知悉”的要求。^②

综上，司法实践中的主流做法并非对爬虫行为中的数据权益保护诉诸侵权法、知识产权法等方式，但在学说上不乏支持者。然而，数据因不同于有体物而无法被纳入物权保护，因缺乏独创性而无法被纳入著作权保护，因常具有公开性而无法被纳入商业秘密保护，通过侵权法的保护路径或《反不正当竞争法》一般条款的保护路径均存在局限。由此，以上路径都无法为违法数据爬虫行为提供成熟、合理的裁判依据，不能有效解决在我国广泛存在的违法数据爬虫问题。

三、数据占有保护进路之正当性与优越性

对数据权益，目前支持数据确权的学者将其称为“数据权”、“数据产权”或“数据财产权”，^③反对确权的学者虽认可数据上的财产利益，但不赞成从传统产权的角度认定数据上的权利归属，而是主张在数字经济时代对企业数据采取保护和共享相平衡的权益模式。^④虽然学界对赋权问题存在不同主张，但对数据的基本特性和数据流通的基本价值理念已初步达成一致观点，即数据是市场主体的合法财产权益，但这种财产与作为传统所有权客体的物存在区别。传统所有权理论适用于有体物只为一人所所有和排他性使用的情形，其制度安排并不适合数据资源，数据是在不断流通利用中实现其价值

① 参见单晓光：《数据知识产权中国方案的选择》，《人民论坛·学术前沿》2023年第6期，第40页。

② 参见刘志鸿：《企业公开数据法律保护范式选择——从赋权论的证成到否定》，《中国流通经济》2022年第6期，第119页。

③ 申卫星教授借鉴权利分割理论，主张由个人享有数据所有权，平台企业对其合法依规控制的数据享有数据用益权，参见申卫星：《论数据用益权》，《中国社会科学》2020年第11期，第110页；张新宝教授主张应将数据财产权确立为与物权、知识产权并列的第三类具有对世性的财产权利，参见张新宝：《论作为新型财产权的数据财产权》，《中国社会科学》2023年第4期，第144页。

④ 参见梅夏英：《数据的法律属性及其民法定位》，《中国社会科学》2016年第9期，第164页；周樾平：《大数据时代企业数据权益保护论》，《法学》2022年第5期，第165页。

的社会资源，赋予任何单一主体完全排他所有权式的支配利益既不必要，也不利于数据价值的实现。^①

在传统物权法上，作为物权客体的物只能是有体物、特定物，具有一定体积、占据一定空间，能被人控制，这些是物权最基本的特征。^② 数据资源显然不具有物的上述特性，因此很难将之认定为物权的客体。但是，数据的持有行为与传统物权法上物之占有的行为模式存在很大程度的相似性。质言之，权利客体的属性不同并不影响行为模式的同质性，虽然数据不当然受物权法的全面保护，但如果数据持有行为满足占有制度的各项要素，应当认为存在类推适用的制度基础。《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》（以下简称《数据二十条》）中明确提出要“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”，^③“持有权”并非《民法典》中的既有概念，对数据持有权^④的定性，学界也并未达成一致观点。笔者拟从传统物之占有制度的基本理论出发，在剖析数据持有的性质及内涵的基础上，对持有权进行解释。

（一）数据持有权与物之占有的同质性——“体素 + 心素”双重要件

近代大陆法系的占有制度吸收了罗马法上 Possessio 和日耳曼法上 Gewere 的双重营养，是多种制度综合的产物。二者在构成要件上存在差异：Possessio 由体素与心素两个要素构成，缺一不可；Gewere 之构成没有体素与心素之分。^⑤ 德国民法学界在占有的构成要件理论上，经历了从萨维尼的主观说到耶林的客观说再到贝克尔的纯粹客观说三个阶段。德国学者普遍认可的观点是，按照《德国民法典》第 854 条的规定应当坚持客观说立场，即只要存在事实上的支配即可成立占有事实，而不用考虑占有意思。但溯源立法材料可以清楚地看到，《德国民法典》第一草案原来规定占有的成立须以占有意思为要件，后因担心人们误以为占有必须占有人知悉其对物具有管领力

① 参见高富平：《数据持有者的权利配置——数据产权结构性分置的法律实现》，《比较法研究》2023 年第 3 期，第 30 页。

② 参见孙宪忠：《中国物权法总论》，法律出版社 2018 年版，第 95 页。

③ 《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》，https://www.gov.cn/zhengce/2022-12/19/content_5732695.htm，2024 年 1 月 9 日。

④ 本文认为，数据持有权和物之占有的成立具备相同的构成要件，二者的行为模式具有同质性、制度目的存在一致性。因此，数据持有权可类推物之占有的相关规定。二者仅因客体不同而存在概念差别，故在本文中“数据占有”和“数据持有权”具有相同含义。

⑤ 参见杨佳红：《民法占有制度研究》，西南政法大学博士学位论文，2006 年，第 76 页。

才能成立而删除了该要件。^①德国学者、司法判例在解释占有构成要件时也认可成立占有除以事实管领作为体素，还须以占有意思为心素。^②我国学界主流观点同样认为，占有的构成要件包括客观控制事实和主观意图两层要件，即“体素+心素”双重要件。^③

可见，无论是对德国还是我国学界而言，占有的构成要件基本都以萨维尼的占有理论——主客观双重要素为基础。萨维尼在其成名作《论占有》一书中对体素和心素的内涵进行了详细阐述。他认为，体素并不局限于直接的身体接触，而是一种“作用于物的能力”，他在书中举例，一个人对保存在他房屋里的物品存在占有，这并不需要对占有物直接接触为前提。^④由此，占有的体素要件在本质上是一种抽象空间范畴内的管领力，即当客体属于主体支配空间内、具有随时利用可能性时，该客体即被主体占有。随着社会的进步和科学技术的发展，占有的客观要件——“事实上的管领”的成立也日趋观念化、抽象化，学者认为应当依社会一般观念并结合空间关系、时间关系等确定占有是否成立。^⑤可见，空间关系要素和时间关系要素成为占有体素的判断起点。空间关系要素的判断不限于直接的触摸和控制，只要主体和客体在空间上有一定的足以使他人认识的结合关系即可。^⑥这和萨维尼认定体素要件的观点在本质上具有一致性。换言之，体素要件中的空间关系在本质上指向的是一种现实的利用可能性，即只要主体通过外部可辨识的方式能够仅依照自己意志即可实现对客体的利用，便可认为其具备了占有的空间关系要素。时间关系要素的判断，指占有主体和客体在时间上须有相当的继续性。满足空间、时间要素方可成立“事实上的管领”。此外，对占有构成要件中的心素，萨维尼认为它是一种支配意图。我国学者补充强调，占有的主观意思并不是法律行为上的意思，而是一种自然的意思——一种不针对特定物的一般的、潜在的占有意思。^⑦例如，甲在其门外设置牛奶箱，意在取得送奶工放入其中的牛奶，甲是否实际知晓牛奶已投入其中并不影响占

① 参见 Protokolle der Kommission fuer die zweite Lesung des Entwurfs eines buergerlichen Gesetzbuchs fuer das deutsche Reich, Bd. 3, 1899, p. 30.

② 参见王泽鉴：《民法物权》，北京大学出版社2010年版，第419页。

③ 参见李锡鹤：《物的概念和占有的概念》，《华东政法大学学报》2008年第4期，第26页。

④ 参见[德]弗里德里希·卡尔·冯·萨维尼：《论占有》，朱虎、刘智慧译，法律出版社2007年版，第181、172页。

⑤ 参见刘智慧：《占有制度原理》，中国人民大学出版社2007年版，第225~226页。

⑥ 参见谢在全：《民法物权论》下册，中国政法大学出版社2011年版，第1147页；李永军主编：《民法学教程》，中国政法大学出版社2021年版，第242页。

⑦ 参见王泽鉴：《民法物权》，北京大学出版社2010年版，第419页。

有的成立。

数据的持有行为也具备“体素+心素”的双重要件。一方面，在客观层面上，平台企业对其控制利用的数据存在事实上的管领力。虽然数据是虚拟无形的财产，无法像传统有体物一样与平台主体展开现实的、物理的直接接触，但数据的储存以有形的载体如电脑、硬盘、终端存储器等为媒介，将数据存储于实体或虚拟介质上，平台企业基于对网站、终端存储器等介质的空间管领力，能够随时对数据进行处理、加工、使用，这符合现代意义上体素的空间关系要素要求。另一方面，平台企业持有的数据并不是瞬时的，其对存储在媒介上的数据的持有也是连续和长期的，非因本人意志不可剥夺，满足时间要素中继续性的要求。在主观层面上，平台企业利用数据是存在主观支配意图的，属于萨维尼所言的“事实上意图进行的支配”，满足占有中的心素要件。随着占有概念的抽象化、观念化发展，占有人尚可以“命令与服从的从属关系”通过占有辅助人取得对物的占有，在互联网时代，通过网络载体的命令和控制手段同样也成立对储存的数据进行占有的关系。“如果数据持有者能够基于数据存储器的占有获得数据调用能力，那么他就是获得了数据的占有。从对客体的事实支配力角度，数据持有者的地位与物占有人的地位是类似的，因而数据持有者可以类推适用占有规范。”^①有学者认为，数据可以通过载体固定下来且受代码或者技术规则的控制，因而可以被数据处理者支配，从而进行利用、收益、占有和处分。^②因此，平台企业对能随时处理使用的数据持有行为满足传统民法中占有的构成要件。

持有的概念多见于刑法，民法中使用甚少。萨维尼在书中进行了专章论述，他认为持有只是一种事实描述的非法律概念，是一种自然关系，持有之外必须存在自决，即持有的意图，才能成立受法律保护的占有。^③我国也有学者认为，持有者只具有对物的事实控制，即占有体素，但不具备占有的意思，即占有心素。^④笔者认为，持有本身属于法律范畴之外的自然事实，但是《数据二十条》规定的持有权，已经是自然事实进入法律范畴后受到法律保护的权利义务关系，该持有权的本质其实就是民法中的占有，只不过因为传统占有制度限定在有体物上，这里采用持有权的概念，

① 云晋升：《占有理论下数据持有者权的建构》，《暨南学报》（哲学社会科学版）2023年第4期，第54页。

② 参见张新宝：《论作为新型财产权的数据财产权》，《中国社会科学》2023年第4期，第161页。

③ 参见[德]弗里德里希·卡尔·冯·萨维尼：《论占有》，朱虎、刘智慧译，法律出版社2007年版，第79页。

④ 参见赵晓钧：《论占有效力》，法律出版社2010年版，第19页。

系针对不同性质的客体在措辞上进行的区分，并不影响行为的本质属性。申言之，数据持有权和物之占有在本质上属于相同性质的行为，都是持有事实具备其他法律要件后成为受法律保护的行为，只是前者的客体是无形财产的数据，后者的客体则是有体物。可能存在疑问的是，除日本法中明确认定占有的性质是权利，大陆法系的主流观点仍然认为占有属于事实而非权利，但数据持有权明确定性为权利，二者看似存在定性上的矛盾。萨维尼在书中也回复了这个问题，他提到，占有在本质上是一个事实，但与之相联系的法律后果也确凿无疑，因此占有在本质上既是一项权利又是一种事实，即根据其本质是一种事实，就其产生的后果而言则等同于一项权利。^①这与持有权的形态和效果是一样的，即行为外观是一种事实，同时也产生了权利义务的法律效力。综上，《数据二十条》中的数据持有权在本质上就是民法中的占有，与物之占有有一样，二者不仅构成要素相同、具备相同的行为模式，而且均是在事实基础上产生的权利义务关系，占有人（数据持有人）基于持有事实及其他构成要素取得了权利人的地位。因此，数据持有权具备类推适用占有的制度基础，数据持有人也可被称为数据占有人。

（二）公开的排他性亦是物之占有的隐藏要件

当体素和心素要件同时满足时，物之占有即成立。那么，平台企业对数据的持有事实具备“体素+心素”要件后，就能同物之占有一样直接成立对数据的占有吗？笔者认为尚且不能，因为物之占有的成立除了体素和心素两个要件，在实质上还存在一项隐藏要件——公开的排他性，这是基于有体物物理属性的天然要求。有体物是肉眼可见、肢体可触的，占有是一种现实存在的客观事实，因此，两个主体对同一物享有完全相同的排他性支配在物理上是不能成立的。^②有体物之占有人必然拥有对占有物专属的、排他的控制力，并且该控制力对他人有公示性。^③正如有学者所言，要成立物之占有，占有人不仅需要满足“体素+心素”要件，同时该占有状态必须是公开的，而且占有人对物的支配关系应能为一般社会观念判断。^④只有当占有具备了公开的排他性，法律才能基于和平秩序的考虑，不允许其他主体对这种已然占有状态进行侵犯。同理，在数据占有的构成要件中，必须将这项

① 参见[德]弗里德里希·卡尔·冯·萨维尼：《论占有》，朱虎、刘智慧译，法律出版社2007年版，第22页。

② 参见贾婉婷：《罗马物权法——所有权与占有》，中国法制出版社2019年版，第187页。

③ 参见[德]鲍尔、施蒂尔纳：《德国物权法》，张双根译，法律出版社2004年版，第105页；[德]M.沃尔夫：《物权法》，吴越、李大雪译，法律出版社2004年版，第75页。

④ 参见刘智慧：《占有制度原理》，中国人民大学出版社2007年版，第227~228页。

隐藏要件还原，才能真正弥补数据占有与物之占有行为模式之间的鸿沟。换言之，当平台企业对数据的持有同时具备体素、心素和公开的排他性三个要件时，数据持有人方可获得和传统物之占有人相同的法律地位，成立数据占有（数据持有权），可类推适用《民法典》关于占有保护的相关规定，其他经营者或自然人均不能随意对该数据进行爬取。该三要件也构筑起爬虫行为的违法性边界。

公开的排他性要件不仅是民法占有制度本身的隐藏条件，而且彰显了数据资源保护的特殊要求。数据资源保护既要遵循数据本质特性的底层逻辑，如无形性、可复制性、可共享性等，又须维护开放、协作、平等、共享的互联网精神的时代性，即对数据资源流通、共享的现实要求。^①在技术层面，“互联”意味着网络运营者应在“认证”和“兼容”的条件下达成计算机、局域网、万维网之间的互联互通，“端对端”的互联网技术架构实现了海量信息的开放传播。^②在政策层面，数据的自由流通在实现数据驱动的增长和创新方面发挥着重要作用，为此欧盟《通用数据保护条例》（GDPR）和《非个人数据自由流动条例》（Regulation on the Free Row of Non-personal Data）中均明文规定“数据自由流通原则”，《中华人民共和国数据安全法》第11条也明确规定“促进数据跨境安全、自由流动”的基本原则。由于数据具有公共属性，数据的流通既是互联网技术的必然要求，又对经济发展具有积极的促进作用，许多学者都认为数据不应为个人独有，公众对数据的使用不会对数据产生损耗，这种使用权益应当得到维护。^③

在比较法层面，美国联邦最高法院在费斯特案中明确否定了下级法院的“额头汗水”的教义，认为数据本身仍然应当维持公众所有。^④在HiQ诉领英案中，HiQ公司对领英网站实施了网络爬虫行为，但加利福尼亚北区地区法院认为，这种爬虫行为不违反法律，因为领英网站上的数据是公开数据，对公开数据，爬取行为应当是被法律允许的，这就像在白天推开一家未锁门的商店进去看看，不能将其认定为非法侵入。^⑤“数据即言论”（data is speech）的观点在美国已经逐渐成为主流。^⑥

据此，数据资源应当共享、促进数据积极流通的理念已达成共识。如果

① 《数据二十条》中明确提出要“促进数据要素流复兴用”。

② 参见许可：《数据爬取的正当性及其边界》，《中国法学》2021年第2期，第171页。

③ 参见丁晓东：《数据到底属于谁？——从网络爬虫看平台数据权属与数据保护》，《华东政法大学学报》2019年第5期，第75页。

④ 参见 *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U. S. 340, 347 (1991)。

⑤ 参见 *HiQ Labs, Inc. v. LinkedIn Corp.*, No. 17-16783 (2017), paragraph 6。

⑥ 参见 Jane Bambauer, *Is Data Speech?*, *Stanford Law Review*, Vol. 66, 2014, p. 70。

认为只要满足“体素+心素”要件的数据就成立平台企业的独立占有，则会过度保护平台企业对数据的控制，有碍数据的流通和利用。因此，在平台企业持有的海量数据中，只有平台享有公开的排他性占有的数据，才能受到民法中占有制度的无差别保护。这部分数据具有排他性，并且该排他性已然得到公示，即一般社会大众通过外部识别能够确认该数据被个人或平台主体排他地占有，只有他人的随意复制和使用行为才会构成对数据占有的侵害。申言之，为了平衡数据保护和数据资源的开发利用，网络空间中的数据存储在任何一个网页上都应首先被推定为可共享的数据资源，除非部分数据被公开且排他地占有，则应当优先保护数据占有人的利益，未经占有人授权不得随意复制、使用该数据资源。

（三）数据持有与物之占有存在相同的制度目的

数据的占有保护进路符合“相同事物相同处理”的基本法理。按照上文的分析，数据占有与传统民法中的物之占有具有同质性，当数据占有在体素和心素要件之外，增补公开的排他性这项隐藏要件时，数据占有与物之占有具备相同的行为模式，理应获得法律无差别的保护。此外，数据的占有保护进路，符合无涉数据确权的共同认知。目前，学界主张对数据进行确权的学者多从财产权的角度为数据保护设计路径，提出数据产权说、数据所有权和用益权分离说、数据公开传播权说、数据块权利说、新型财产权说等众多学说，这些学说或多或少保留了传统物权的思维模式，即肯认数据权益的“类物性”，正因数据权益具有可支配、可利用、可收益、可处分等特性，所以可以将其纳入权利范围予以保护。基于此，在坚持数据权利化的前提下，数据的占有自不待言。此外，也有不少学者基于数据本身的特点及数据上存在的多元利益主体，不赞成对数据权益进行绝对化的财产权设计，但仍然承认企业对数据的控制利益，“企业对数据的控制并非一项权利，它是一种法律事实……企业对数据的自我控制成为企业做出选择的前提，法律尊重企业对数据的控制状态类似于物权法上的‘占有’保护，故这种数据控制事实不能构成实体权利，只能作为一种防御性的‘法益’存在”。^①可见，无论是否赞成数据确权，将数据权益认定为市场主体在事实上控制和占有的客体的观点都是不存在争议的。

在爬虫行为中以占有进路对数据进行保护，符合占有制度的根本宗旨。占有是对物在事实上的支配，占有一旦成立，即应受保护，以维护社会平和

^① 梅夏英：《企业数据权益原论：从财产到控制》，《中外法学》2021年第5期，第1200～1201页。

与物之秩序，而是否存在本权在所不问，这是占有制度的基本功能。^①关于占有制度的目的，存在物权说、社会秩序维护说、人格权益保护说等理论，目前大陆法系通说认为其是为了维持社会的平和秩序，保护民事主体对物在事实上的管领力。数据权益作为新型的生产要素，需要平衡权益保护与资源的共享共用。^②与传统物权法、侵权法立足个人权利本位的价值基础不同，数据权益保护不仅强调对市场主体数据权益的维护，而且兼顾整个社会层面在爬虫行为中对网络秩序、社会公共秩序之稳定的维护。因此，在数据权益是否确权尚未明确的当下，对爬虫行为中的数据保护以“不问本权、仅看占有”且兼顾社会、市场多方主体利益的理念进行司法裁判，符合数据资源的保护和发展要求。正如学者指出的：“当本权暧昧不明或者有争议时，占有保护的功能能够彰显出来。”^③

（四）数据占有保护进路的优越性

民法上对占有的保护可分为物权法上的保护和债权法上的保护，前者主要指占有保护请求权，后者包括不当得利返还请求权和侵权损害赔偿请求权，救济路径多元。占有保护请求权属于物上请求权类型之一，在爬虫行为中，由于数据的物权属性未明，故无法以物权请求权的保护方式判令数据侵害方停止侵害、返还原物，但占有保护请求权能在未确认权益性质的情况下第一时间终止侵害行为，及时保护受害方的数据权益。同时，借由占有保护的债法规定，可以为损害赔偿的成立及计算提供清晰有效的规范基础，从而为爬虫中数据权益的保护提供周延且具有实操价值的制度保障。例如，在深圳市腾讯计算机系统有限公司等诉浙江搜道网络技术有限公司等案中，法院根据《反不正当竞争法》第2条判决两被告立即停止涉案不正当竞争行为、共同赔偿两原告经济损失260万元人民币，该判决欠缺责任成立和损害赔偿如何计算的规范基础。若以占有保护进路进行救济，首先类推《民法典》第462条，将占有保护请求权作为请求权基础，责令被告停止侵害数据占有利益的行为，由于被告行为造成了原告数据占有利益的损害，损害赔偿范围包括原告所失利益及被告侵占数据权益后获得的收益，在侵权法的范围内对

① 参见王泽鉴：《民法物权》，北京大学出版社2010年版，第426页。

② 国家发展和改革委员会在就《数据二十条》的答问中指出：在数据生产、流通、使用等过程中，个人、企业、社会、国家等相关主体对数据有不同利益诉求，并且呈现复杂共生、相互依存、动态变化等特点，《数据二十条》以解决市场主体遇到的实际问题为导向，聚焦数据使用权流通，构建中国特色数据产权制度体系。参见中华人民共和国国务院办公厅：《国家发展和改革委员会就“数据二十条”答问》，www.scio.gov.cn/xwfb/bwxwfb/gbwfbh/fzggw/202303/t20230320_706715.html，2023年9月2日。

③ 章正璋：《民法典第462条占有保护制度的几个解释与适用问题》，《苏州大学学报》（法学版）2020年第4期，第67页。

行为进行定性及确定赔偿数额，确保裁判有法可依，使权益人从物权法和债权法上获得全方位的保护救济。

对数据爬虫经由占有保护制度进行规范，能弥补《反不正当竞争法》适用主体的局限性。无论是经营者还是非经营者的自然人，不论其是否具有竞争关系，均可一体适用占有保护的相关规定。对数据的侵占往往是通过技术性手段完成的，而数据具有无形性和可复制性，在数据被侵占后一般还存在其他的下游侵占人从前手的侵占人手里继续复制数据资源的情形。占有保护并非保护本权利，而是基于物之秩序保护整个社会平和状态，因此，作为前手的侵占人也可以基于占有对后手的侵占人主张占有保护请求权，这种层层递进的救济方式有助于及时止损，防止数据权益的侵害泛滥，避免合法占有人的损失扩大，^①符合数据本质特性和维护市场秩序的经济目标。

综上，在数据爬虫行为中通过类推适用物之占有规定，数据侵害后的停止侵害、占有妨害的排除、损害赔偿的计算、不当得利的返还等问题均能在《民法典》中找到合适、明确的规范基础，以保障司法实践有法可依及裁判统一。笔者认为，通过传统民法中的占有保护制度对数据爬虫行为中的数据权益进行保护，较其他进路具有更明显的优势。

四、数据占有保护的内涵

数据爬取行为在促进数据流通的同时，也引发了网络流量过载、未经授权收集个人信息、企业数据被不当复制和窃取、破坏安全防护入侵网络等不良后果。因此，必须妥善化解数据爬取权益和拒绝数据爬取权益之间的冲突，^②这就需要为数据爬虫行为找到合法的界限。笔者认为，占有的三要件，即体素、心素、公开的排他性，在实质上建构了数据爬取行为的合法性边界。一般而言，遵循“数据即言论”这一互联网精神，平台数据可以被爬取。但是，当数据持有行为满足三要件时，数据持有就从一项自然事实上升为受法律保护的持有权，即数据占有。此时，基于对平和秩序的优先保护，被占有的数据不得被随意爬取，否则爬虫行为人应当承担相应的责任。在数据持有权的三大构成要件中，体素和心素要件上文已详述，与物之占有

① 在传统物之占有中，以是否存在基础权利为标准，占有可分为有权占有和无权占有。目前，数据权益无涉基础权利的问题，因此本文将有权根据持有数据资源，即基于整合、加工等事实行为或许可协议等法律行为有权对数据进行利用、收益的类型称为合法占有；将不具备合法根据或原因持有数据并进行利用的类型称为非法占有。

② 参见许可：《数据爬取的正当性及其边界》，《中国法学》2021年第2期，第173页。

中的内涵并无本质差别；公开的排他性要件在物之占有中天然内置，不必经外部进行判断，但在数据持有权中，该要件并非不言自明，需由法官进行具体确定。

（一）公开的排他性要件的内涵及认定

何谓公开的排他性？美国著名的马粪案或许可以很好地回答这个问题。该案中的原告请了两个帮工到马路上捡马粪，晚上这两个帮工在马路上堆了18堆马粪后回去取车，准备第二天运走，然而马粪却在第二天早晨被被告拉走。该案法官判决原告胜诉，被告归还马粪。对此，波斯纳法官认为：马粪堆积足以构成对他人的公示——这些马粪不再是抛弃物，如果要求原告为了保护自己的利益，除了将马粪堆积，还必须将它围起来，或者一直看着它，或者预先安排马车在马粪堆放后立即将其拉走，这些只会增加交易成本，而不会产生任何补偿性的利益。^①

根据波斯纳的想法，占有并不必然以实际控制为要件，只要存在外观的合理公示使一般人确信该物上存在他人的排他性占有，这项占有利益便不能被随意侵害，这就是公开的排他性。同理，数据资源的占有也是如此。由于互联网平台的开放性和数据资源的公共性，网络空间中的数据就如马粪案中尚未被堆积的马粪一般，具有一般的可共享性。但是，如果数据持有者通过外在行为对数据的排他性占有进行公示，如通过明显可知的数据加工、处理行为形成的独占性数据资源库，就如已被堆积的马粪一样具备了排他性的公示力，此时应认定加工者占有该数据资源，对其占有的数据可类推民法中的占有保护规则对数据权益的侵害进行救济。

数据占有中的公开的排他性要件具体应当如何判断？概括而言，如上文马粪案依据习惯经验对公示的排他效力进行判断一样，也应当考虑一般社会观念和网络空间的利用习惯，结合多元价值因素进行综合分析。公开的排他性要件必须具体化，才能真正发挥实践价值。案例类型化为法官提供了指导范例，是具体化的最有效途径。应当将实践案例进行系统性的整理，同时提炼出当中与价值有关的标准并剔除其与具体情形的关联，从而对其加以类型化的处理。^② 随着网络的发展和实践案例的丰富，应当从数据的多重属性与场景化特征出发，通过自下而上的实际案例的类型化，最终提炼出公开的排他性要件的抽象标准。

^① 参见 [美] 理查德·艾伦·波斯纳：《萨维尼、霍姆斯和占有的法经济学分析》，于子亮译，谢进杰主编：《中山大学法律评论》第11卷第1辑，广西师范大学出版社2013年版，第155页。

^② 参见 [德] 托马斯·M.J. 默勒斯：《法学方法论》，杜志浩译，北京大学出版社2022年版，第467页。

经过对目前国内外典型案例的整理,本文初步提炼出以下几类认定公开的排他性要件的要害,法官在司法实践中可参考这些案例类型,对是否构成数据的占有进行认定。第一,在数据上存在民事主体的合法权利时,可认定该数据存在公开的排他性。例如,在某项数据上存在自己或他人合法的著作权、专利权、隐私权或个人信息等受法律保护的权利。对数据来源者(个人)而言,具有主体可识别性的个人信息构成其对数据占有的排他性公示。欧盟《通用数据保护条例》第4条将个人数据界定为“任何已识别或可识别的自然人相关的信息”,只要数据的来源者可以被清楚地识别,该项数据上存在的个人信息就同著作权、专利权一样,在该数据上成立公开的排他性。在深圳市腾讯计算机系统有限公司等诉浙江搜道网络技术有限公司等不正当竞争纠纷案中,法院认为,“两原告(平台和个人)主张数据权益的微信平台数据,可以分为两种数据形态:一是数据资源整体,二是单一原始数据个体,网络平台方对于数据资源整体与单一原始数据个体所享有的是不同的数据权益”,^①这符合《数据二十条》对数据分类保护的原则。无论数据来源者是个人还是平台企业,均可以基于公开的排他性对持有的数据主张占有保护。第二,在比较法视野下,对数据库的原创性汇编是否构成公开的排他性要件,还存在认识上的一致。这一点在美国司法中存在争议,^②而欧洲不仅主张对数据库的原创性汇编部分进行保护,而且将数据库作为特殊权利(*sui generis rights*)予以保护。^③笔者认为,只要数据库需要足够的人力、物力、技术等资源投入,这些具有独创性和高技术性的数据库就存在对一般使用人的公开的排他性,此时“额头汗水理论”就应当成为法官重点考虑的因素。这和马粪案的理论一致,如果对马粪只是简单处理,并不能使他人确信存在公开的排他性,很难成立主体对马粪的占有,但如波斯纳所言,原告将马粪堆积成18堆,如此数量使背后的劳动显而易见,存在外观的合理公示使一般人确信存在他人的排他性占有。同理,简单的数据整合行为并不必然彰显平台对数据的专有性和排他性,但具有高技术性、复杂性的数据库如同被堆积的马粪一样公示了平台的劳动投入,按照一般社会观念,存在公开的排他性。法官在裁判时可根据个案情况进行分辨,确定是否构成排他性的公示。第三,Robot协议也是考察公开的排他性要件的一类情形。Robot协议是否有效在各国的司法与法律教义上存在争议,本文不延伸讨论关于

① 杭州铁路运输法院(2019)浙8601民初1987号民事判决书。

② 参见丁晓东:《数据到底属于谁?——从网络爬虫看平台数据权属与数据保护》,《华东政法大学学报》2019年第5期,第76页。

③ 参见 Directive 96/9/EC, of the European Parliament and of the Council of March 11, 1996 on the Legal Protection of Databases, 1996, O. J. (L77) 20, chap 3.

Robot 协议的性质和有效性问题，但由于其存在明示的宣告性，这种宣告可以成为公开的排他性的认定因素。当然，在司法裁判中还需对其合法性和合理性进行综合考量，最终确定公开的排他性要件成立与否。第四，数据是否构成平台的核心竞争资源亦是公开的排他性要件的考虑因素，如因侵犯平台的核心竞争数据引发不当的替代关系，须承担相应的责任。例如，在大众点评网诉百度案中，^① 法院认为百度地图大量使用大众点评网的用户点评，替代之向网络用户提供信息，会导致大众点评网的流量减少，构成对大众点评网的不当替代关系，侵害大众点评网的权益。因此，当数据已构成平台的核心竞争力资源，如大众点评网中的用户点评，一般认为平台对其享有公开的排他性。需要说明的是，能够构成平台核心竞争力资源的数据不能仅是个别单条，单一数据并不能形成竞争力，必须聚合成一定规模成为数据群，该数据群能够直接对平台的流量、经济利益造成影响，才能形成“核心竞争力”，构成公开的排他性要件。第五，非公开的数据具有公开的排他性，一般大众对非公开的数据，可推知平台具有专有性、排他性，构成数据的占有。

（二）数据来源者和数据处理者之间的关系——共同占有

明确数据的占有人是法律适用的前提。众所周知，数据资源涉及多方主体的权益，不仅有数据来源个体的合法权益，而且包括数据处理者对依法持有的数据进行自主管控的权益，还须保障在数据利用中数据要素的共享性和普惠性。^② 虽然数据资源在生产、流通、使用、处分等过程中涉及多方利益，但是数据占有人问题必须回到民法占有制度的基本理论加以探究。按照上文的论述，个人和平台对其实际控制的数据资源均可具备公开的排他性要件，因此，二者均可成为数据占有人。占有制度的立法例和学说普遍认为，存在数个主体的占有形态包括间接占有、占有辅助人与共同占有，前二者都是由一方主体（占有媒介人或占有辅助人）直接控制占有物，另一方主体（间接占有人或占有主人）成立非直接控制的抽象占有。在数据占有中，无论是个人还是平台主体，都是通过自己手中的设备对数据拥有实际的、直接的管领力，因此个人和平台对数据的占有更类似共同占有。共同占有和共有并无必然联系，共有人对共有物并非当然为共同占有人。例如，甲乙共有某二层楼的房屋，二人订立了每人分管一层的契约，此时成立单独占有。^③ 因

^① 参见上海市浦东新区人民法院（2015）浦民三（知）初字第 528 号民事判决书。

^② 《数据二十条》第 3 条提出，要“分别界定数据生产、流通、使用过程中各参与方享有的合法权利，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”。

^③ 参见王泽鉴：《民法物权》，北京大学出版社 2010 年版，第 442 页。

此, 无论是否支持数据确权, 抑或如何确权, 都不影响从事实上认定个人和平台间的共同占有形态, 共同占有只是事实层面、支配行为上的限制, 与权利限制无关。

共同占有可分为重复共同占有与统一共同占有。^① 在数据占有关系中, 无论是个人还是平台均可以在不妨害其他占有人的情形下单独使用数据, 这属于典型的通常共同占有。《民法典》尚未对共同占有进行规定, 但在学说上基本达成共识: 区分单独占有和共同占有的意义就是在对共同占有的内部关系和外部关系建立特殊化的规则。^② 我国台湾地区“民法”第965条规定: “数人共占有一物时, 各占有人就其占有物使用之范围, 不得互相请求占有之保护。” 该条规定有很强的借鉴意义, 个人和平台之间的纠纷可参照共同占有人内部关系的处理机制: 若个人和平台在对共同占有的数据进行利用时存在事前协议, 则须遵照协议要求; 若不存在相关协议, 除非一方完全排除另一方的使用利益, 否则受不利方不得对共同占有人主张占有保护。^③

(三) 数据占有保护请求权的特殊性

数据的占有保护与传统民法物之占有保护的理念一样, 从维护社会秩序平和的目的出发, 防止爬虫行为对占有利益的侵害, 保护占有人对其控制财产的占有利益。物之占有保护在物权法上体现为占有保护请求权, 以占有被侵夺或占有被妨碍为前提, 包括占有返还请求权和占有妨害请求权两类。占有被侵夺指非基于占有人的意思而排除其对物在事实上的管领力, 占有被妨害则指非侵夺占有而妨碍占有人管领其物, 致其使用可能性及利益遭受侵害。^④ 这两项请求权不问原占有人有无占有本权, 均是对不能以所有权或一项物权为支撑的“弱小地位者的法律保护”。^⑤

数据的占有保护请求权因数据特性与物之占有保护请求权存在差异。通说认为, 占有返还请求权仅能对侵占人主张, 《民法典》第462条第1款将“侵占”作为占有返还请求权要件的做法也是遵循了大陆法系民法通例的结

① 参见梁慧星、陈华彬:《物权法》, 法律出版社2020年版, 第411页。

② “区分单独占有与共同占有的意义在于, 既然各个共同占有之物之事实支配均受法律保护, 那么共同占有人之间不存在就各自占有物的支配范围单独请求占有之保护的问题。” 见梁慧星、陈华彬:《物权法》, 法律出版社2015年版, 第284页。

③ 参见王泽鉴:《民法物权》, 北京大学出版社2010年版, 第443页。

④ 参见王泽鉴:《民法物权》, 北京大学出版社2010年版, 第536页。

⑤ 参见[德]鲍尔、施蒂尔纳:《德国物权法》, 张双根译, 法律出版社2004年版, 第160页。

果。^① 数据作为一项受法律保护的财产权益，理应由合法占有人进行控制和利用，如若未经占有人的同意，违法将数据复制或转移控制，则属于数据侵占。由于数据与有体物之显著区别在于数据的无形性和可复制性，因此不同于有体物之侵占行为外观——直接使占有物本身脱离占有人的管领，数据之侵占只是利用技术手段将数据信息加以复制存储，并未使原占有人丧失对数据的占有和利用能力。据此，传统的占有返还请求权能否直接适用到数据的占有保护中不无疑问。

笔者认为，数据的占有从表面上看似似乎只是对原始数据资源这一单一载体通过技术平台实现控制，但在本质上，由于数据在技术上的可无限复制性，占有人实际控制的是该数据资源的原始版本加无限的复制版，即以原始数据 A 为占有的外观，同时包含了对数据复制版 A1、A2、A3、A4 等所有复制版的控制实质。如果其他平台企业未经许可或非基于合理事由获得原始数据的复制版，则可视为使原数据占有人丧失对一份复制品 Ax 的控制，并将占有转移到侵占人的管领之下。此时，丧失数据复制品的原占有人可以类推《民法典》第 462 条的规定，向侵占人主张占有返还请求权。但是，由于该数据内容在原数据占有人手中始终存在，要求侵占人返还、回传复制数据对占有人显然并无实益，出于经济效率的考虑，直接要求侵占人删除数据更合理、高效和安全。因此，基于对数据资源的特性和经济效率的考虑，数据的占有返还请求权在行使方式上与物之占有返还请求权存在区别，以要求侵占人停止使用、删除数据替代数据复制版的返还或回传即可实现占有保护的目，并且更符合《民法典》规定的绿色原则。

占有妨害并非使占有人丧失管领力，而是通过行为干涉、妨碍权益人之占有致使数据资源的使用、收益能力受挫。数据占有人通过技术平台对数据资源进行控制和使用，但在互联网平台的正常运行中存在很多干扰因素，如网络平台的非法运行、储存终端的技术故障、黑客入侵的风险等。数据占有人在使用和处分数据时，如果受到任何外界干扰，可以类推《民法典》第 462 条的规定，对妨害行为主张排除妨害或消除危险。当干扰因素存在公权力行使或有合法权利的事由时除外，如因强制执行、没收、暂扣等执行公务的行为，或具有违法阻却的自助行为等。

综上，由于数据的独特性质，其占有保护请求权的内容和方式不同于传统的占有返还请求权，以停止使用、删除数据代替数据资源的返还请求，同时保留对妨害人主张排除妨害、消除危险的传统权利行使方式。

^① 参见张翔：《论动产占有返还请求权的构造及其要件的扩张解释》，《法律科学（西北政法大学学报）》2021年第5期，第124页。

五、结论

厘清数据爬虫行为的合法性边界具有现实必要性和迫切性，但在实践中，由于数据权益的性质及确权与否尚没有定论，裁判者不得不诉诸《反不正当竞争法》第2条的原则性条款进行裁判，使裁判者在认定行为性质、判决责任承担方式及确定损害赔偿的范围问题上拥有极大的自由裁量权。同时，不构成竞争关系的主体无法依据该条主张权利。因此，寻找对民商事主体一体适用的数据爬虫违法性标准尤为必要。《数据二十条》提出“数据持有者”的概念，与传统民法中物之占有制度类似，可为数据爬虫的违法性认定标准的构建提供现实方案。

在传统民法上物之占有的成立仅包括体素和心素双重要件，但剖析占有制度可看到，由于有体物具有物理属性——现实性、可观性、可触性，故而民事主体对物的占有必然对其他主体产生公示力。质言之，在物之占有中天然隐含了公开的排他性要件。但是，该要件在数据占有中并非不言自明，数据具有无形性，要成立数据占有，其构成要件在体素与心素之外，尚须对公开的排他性要件加以还原和明确，这也符合数据公开性、共享性的现实要求。公开的排他性以马粪案中的理论为基础进行建构，采用类型化的方法进行判断，在实践中对数据持有者（数据占有）是否成立进行确定，可以在数据保护和数据流通间获取最大平衡。

占有保护请求权包括占有返还请求权和占有妨害请求权，数据占有人可以类推《民法典》中物之占有保护的规定对数据进行多元保护。占有保护请求权能在第一时间停止侵害、防止损害扩大，可在数字经济时代网络传输高速化的当下及时止损。民法中的占有还包含着债权性质的损害赔偿请求权和不当得利返还请求权，能够弥补数据占有人的所失利益，以保障合法占有人的权益。因此，类推占有保护制度对数据爬虫问题进行救济，具备周延性、具体性、可操作性等优势，可为数据权益保护提供合理的救济途径。

（责任编辑：方 军）