

网络爬虫的行政法规制

刘宏秀

湖北大学法学院, 湖北 武汉

收稿日期: 2024年10月12日; 录用日期: 2024年11月6日; 发布日期: 2024年11月15日

摘要

互联网爬行器包括了获取信息的技术设备和防止被抓取的方法, 这可能导致多种形式的法律法规问题。针对网络爬虫的法规管理涉及到众多法律领域, 除了民事和刑事责任外, 还需要有相应的行政监管。根据现有的法律规范, 我们可以采取如下几个方向来实施对爬行器的行政管控: 首先, 可以通过行政立法明确定位并划定爬行器的确切合规范围; 其次, 要权衡使用刑罚惩戒和行政处罚的方式; 最后, 还可以采用诸如设立技术准则等方法来指导爬行器技术的正确应用。这种全方位的管理方式能为网络爬行者的法律处理提供无法替代的支持。

关键词

网络爬虫, 个人隐私保护, 网络防御, 数据保护

Administrative Regulation of Web Crawlers

Hongxiu Liu

School of Law, Hubei University, Wuhan Hubei

Received: Oct. 12th, 2024; accepted: Nov. 6th, 2024; published: Nov. 15th, 2024

Abstract

Internet crawler includes technical equipment for obtaining information and methods to prevent being captured, which may lead to various forms of legal and regulatory issues. The regulatory management of web crawlers involves numerous legal fields, and in addition to civil and criminal responsibilities, corresponding administrative supervision is also required. According to existing legal norms, we can take the following directions to implement administrative control over crawlers: firstly, we can clarify and delineate the exact compliance scope of crawlers through administrative legislation; Secondly, it is necessary to balance the use of criminal punishment and administrative

文章引用: 刘宏秀. 网络爬虫的行政法规制[J]. 争议解决, 2024, 10(11): 159-167.

DOI: 10.12677/ds.2024.1011449

punishment; Finally, methods such as establishing technical guidelines can also be used to guide the correct application of crawler technology. This comprehensive management approach can provide irreplaceable support for the legal handling of internet crawlers.

Keywords

Web Crawler, Personal Privacy Protection, Network Defense, Data Protection

Copyright © 2024 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

信息化时代，网络信息的收集和利用已经成为最基本的日常活动之一。几乎所有的网络用户都直接或间接地利用了网络爬虫提供的信息资源。最初，网络爬虫活动主要由各种搜索引擎和通用爬虫来进行，针对全网信息。然而，随着第三方库的发展和计算机程序语言的流行，专注于特定主题信息的聚焦爬虫活动快速增加。爬虫在搜索引擎、舆情分析、大数据挖掘、违法行为监测、犯罪预测、暗网探知等领域的积极作用已经引起了法学界的关注和认可[1]；与此同时，爬虫的负面影响也逐渐显现。一方面，滥用爬取工具可能导致被爬取网站过载甚至瘫痪，进而削弱普通人对网站资源的使用能力；另一方面，突破反爬机制后，爬虫获取的数据有可能侵犯商业秘密、个人信息或个人隐私等权益，非法侵入内网获取政府工作秘密，逃避支付利用网站资源的对价，还可能利用未经同意的爬取数据进行不正当竞争、助力违法放贷和催收活动等[2]。随着爬虫相关风险的不断增加，对爬虫进行法律规制变得日益必要；《个人信息保护法》《数据安全法》等重要法律的制定不仅凸显了对个人信息爬取等爬虫活动进行规制的紧迫性，也为规制爬虫提供了更充分的制度支持。

目前，关于爬虫的研究及其相关法律法规的主要焦点在于民事和刑事法规，但行政法规的相关理论探讨和管理策略同样不容小觑。在国际互联网管理的背景下，大量的行政管控手段被广泛应用以保持良好的商业环境并保障企业的权益和个人权利乃至国家的战略利益。最近颁布的新版《个人隐私保护法》和《数据安全法》中的许多条例都涉及到行政监管问题，这也预示着中国未来的数据使用过程将会紧密地与行政控制联系在一起。因此，依据爬虫行为的基础原则和技术特征，通过有效的法律规定来实施全面且精确的行政管制，这是解决网络爬虫问题的关键步骤。

2. 网络爬虫的基础理论

(一) 网络爬虫概念的再认识

“网络爬虫”这个概念实际上是指一种由计算机模拟人类操作来获取信息的设备。行业内对于它的认识通常有两个主要方向：一是把“网络爬虫”视为自动化收集信息的数据采集软件；二是将其扩展为包含了破解防盗链技术等一系列用于处理数据抓取过程中的反制措施的方法。

如果我们想要深入探究关于网络爬虫的行政管理途径，那么采用更广泛的爬虫定义作为研究目标会更加合适。仅仅通过对狭义的爬虫设定限制，可能会导致脱离实际中的许多公司和个人对于爬虫的使用情况，同时也将大量涉及到爬虫相关的违法行为案例置身于讨论范畴之外，这样一来，制定针对爬虫的管理策略就会出现偏离主题的可能性。所以，本篇文章将会使用一种宽泛的方式来理解并且利用爬虫这个词语。

(二) 网络爬虫和反爬运行机制理论介绍

基本的活动模式在网络爬虫中的爬取工具部分表现为：首先选择一个 URL 作为“种子”，然后下载相关的网页，接着以此网页为起点寻找并提取额外的 URL，通过这些 URL，使用诸如正则表达式的工具来抽取内容并将有用的新信息保存下来，这个过程持续到达到了一定的标准才停止[3]。随着爬虫技术的发展，它不仅推动了强大搜索引擎和方便的互联网的使用，同时也可能对被爬的站点产生严重的负荷和负面效果。例如，元光公司的行为就是个例子——他们大规模地抓取了谷米公司研发的名为“酷米客”的智能公交应用程序的数据，每天大约 300 万到 400 万次，导致谷米公司遭受超过二十万的直接经济损失[4]。所以，为了避免这种情况的发生，网站管理员常常采取一些反爬措施来阻止爬虫(特别是那些专门用于收集或处理特定的数据的聚焦型爬虫)的不合理或者恶意的使用。

一种基本且易于实施的方法是通过网络抓手程序来禁止部分或者所有访问者的动作。然而这种方式不仅缺乏有效的法制约束而且也很难有效地抵御机器人攻击者[5]。一些较大的互联网检索系统可能会因为它们所带来的可能存在的商业价值从而得到目标站点的青睐并遵循相关规定；但是大量的其他的网页浏览器却不一定遵照相关的规则执行。因此许多站点常常采用更先进的技术去检测并且拦截那些来自机械化的页面下载过程的行为。在这个充满竞争的环境里，防盗版措施已然演变成为一套复杂多样的方法体系了。这包括多种类型的防护设备如：(a) 辨别装置——用于确定是否存在人为操控的情况；(b) 限定行动方案——比如需要先登记然后才能开始工作或者是设定每秒钟只能完成一次点击等等这样的做法都是为了减少恶意软件的使用率；(c) 伪装手法——也就是利用各种技巧让对方误以为自己是一个真实的人，而不是一台电脑。

“反反爬措施”一直在适应并响应于反爬机制的变化与发展。以识别机制为例，爬虫开发人员可通过使用机器学习技术或者依赖专门的打码服务平台来规避验证码带来的障碍；同样地，针对限制操作机制，他们可利用代理 IP 池、分散式的爬虫方式、自设频率及数量限制等方式加以对策；至于欺诈机制，通常情况下，爬虫开发人员会采用嵌入式浏览器引擎的方式，收集浏览器界面展示的内容，或是深度解析网站编程结构，寻找出其他可能的漏洞，以此绕过反爬策略或逼迫网站投入更多资源持续更新其反爬策略。

虽然技术方法无法完全避免这些潜在的风险。实际上，任何为了方便网民浏览并设计的数据传输方式都可能被各种工具获取：那些有着严谨且标准化的数据呈现模式的网站，特别容易成为爬虫的目标。无论是采用哪一种防御措施，互联网访客总是可以有意无意地收集和存储来自网站的信息，进而研究其反爬系统，以便找到突破口。然而，网站的管理人员也可以通过频繁更改反爬策略来防止特定的解析与攻击，但是这需要专业的反爬团队和高昂的运营费用，许多小规模的企业、公共机构及政府部门(特别是基层组织)未必能承担得起这样的开销，所以对网络爬虫实施监管也是必要的。

3. 网络爬虫的行政规制需求

网络爬虫不仅具有无法取代的正面影响，还可能引发一连串不容忽视的风险。因此，在法律实践中对于滥用爬虫技术的人承担民事或刑事责任已经司空见惯，但是这仍然不足以有效地管理相关风险，需要行政部门的介入。

(一) 民事途径的不足

针对爬虫的民事诉讼主要由运营网站的企业发起。提起诉讼的企业通常为腾讯、百度、淘宝等大型网络平台企业，这些企业拥有大量有价值的商业数据，也是爬虫侵犯的主要目标。通常只有这些企业才能承担起爬虫侵权诉讼的高昂取证成本。例如，淘宝诉上海载和网络科技有限公司、载信软件(上海)有限公司一案的判决书表明，原告要求两个爬虫侵权案件共支付律师费 10 万元、鉴定费 8 万元、公证费 6 万

元和差旅费 6 万元。¹爬虫类不正当竞争案件的诉讼成本较高,足以阻碍公共机构和许多中小企业寻求民事法律救济。此外,网络平台企业提起的相关诉讼主要集中在侵犯商业秘密和反不正当竞争领域,主要诉讼对象是拥有一定经济赔偿能力的侵权企业,对于大量小企业和个人使用爬虫非法抓取个人信息等行为,可能缺乏起诉的动力。此外,很多公共平台没有设置精准的反爬机制,对于滥用爬虫抓取公共数据的情况,可能既缺乏防范,导致网站超载,也可能实施了过度防御,影响自然人对网站公共资源的正常使用[6]。以最高人民法院的中国裁判文书网为例,从 2018 年 5 月初开始,许多技术公司通过爬虫系统无限制并发访问非法获取裁判文书数据。曾经导致互联网服务器压力增大,众多正常的访客请求被阻塞,使得网页浏览缓慢或者某些页面无法展示等问题,这不仅对司法透明度产生了不利的影响,也妨碍了公众获取判决文件的使用体验。这种由于网络爬虫引起的消极结果是由多个参与者的爬行行为相互作用而产生的,它关系到广泛的社会利益,因此很难仅依靠民事法律途径来解决这个问题。

(二) 刑法的调整范围具有谦抑性

对于非法爬虫行为的主要刑事法律责任在于违反了非法获取计算机信息系统数据和侵犯个人隐私权两项规定,也有可能牵涉到其他如非法入侵计算机信息系统、非法操控计算机信息系统及提供非法进入、操控计算机信息系统的软件/设备等的规定[1]。但是,刑法只能针对那些严重的违规使用情况作出规范,因此许多违法行为不一定直接受到惩罚。首先,从实质角度来看,并不是所有的爬虫活动都能满足相关的犯罪标准。其次,在价值观方面,不是每个需要法律干预的爬虫活动都有足够的必要性和威胁程度以达到刑法介入的要求。通常情况下,爬虫活动是利害共存的,它可以收集大量数据并将这些数据汇集成更有价值的数据资产,这在特定环境下有助于更好地整合分布式数据资源,并且有效地利用数据价值[7]。中国的司法判决案例表明,他们反对因为限制爬虫访问数据导致“互联网上信息的隔离和封闭”的情况发生。由此可见,人工阻碍爬虫访问数据未必合理,如果没有涉及到窃取私人信息或者造成计算机系统损坏等问题,那么它的合法性就应该依据具体的情境来仔细考虑。再者,若爬虫只作为提高收集公共信息的工具,并没有给目标网站带来实际的不利后果,那么直接将其视为犯罪行为似乎并不合适。许多防爬措施的设计主要目的是为了避免网络超负荷,确保互联网的使用顺畅,而且爬虫通常会自己设定访问速度和数量上限。如果爬虫没有在这方面产生显著的影响,受害方也可以采取 IP 封禁等方式来保护自身权益,在这个情况下追究刑法的责任可能并非必须。

(三) 行政法能有效弥补民法与刑法的不足

相较于民法和刑法的方法论,行政法规可以从理论角度实现广泛地涵盖各类非法网络爬行事件,同时也能适当地指引合法的行为,达到一种多维度的监管效应。现代公共管理中有着丰富的调控方法,除了常规的授权、惩罚、强制之外,还有诸如行政指导、奖赏、公开信息、信誉评估、标准制定等多种方式能够被巧妙运用到对于网络爬行的管控当中。根据理论分析,行政控制非常适用于处理像网络爬行这样的新兴科技应用:首先,网络爬行技术的进步和商业模式的发展变化很快,例如专门针对破解验证码问题的“打码平台”(专业的解决网站访问受限问题的方式)、第三方的数据库、仿真浏览器、特定的数据采集软件等等都在不断地升级迭代;其次,这些反击反爬的技术或者第三方服务都有其优缺点,因此需要使用多样化的精确的行政管制策略来对其进行及时而有效的处理和引导。最后,在数字化经济发展阶段,网络爬行已经成为了普遍的现象,越来越多的中小型公司和个人开始参与其中,并且他们也进行了大量的介于合规性和轻度违规之间的事件。现有的民事或是刑法制度都无法有效地调节这种类型的活动。我们必须依赖于政府实施高效的行政管理手段,并联合公司、行业团体、各类社会机构及知识社群等第三方的协助,共同构建出一种全面掌控网络爬行技术的应用与过度使用的危险均衡的能力。这种行政管理

¹ 浙江淘宝网络有限公司与上海载和网络科技有限公司、载信软件(上海)有限公司不正当竞争纠纷案,上海市浦东新区人民法院民事判决书(2015)浦民三(知)初字第 1963 号。

的特性已经在前期的法律制定过程中得到了清晰地阐述。

4. 网络爬虫的行政管理状态

虽然全球范围内并不广泛地实施对网络爬虫的行政管理，但这与其所在国家的行政法规体系及其互联网监管策略有着紧密联系。在中国，因为行政法已然成为了互联网管理的核心法治基石，因此首要任务是审查网络爬虫行政管控的法律根源，然后审视并反思现行的行政管控执行情况，以期深层次探究和发掘潜在的管理方式和手段，构建出一种融合了法律法规与科技的高效管制架构。

(一) 规制网络爬虫的行政法依据

在国内，网络爬虫的民事和刑事法律应用最早被学界所关注，尽管行政法中已经有对此进行规范的法律依据，但到目前为止仍未得到足够的重视。现阶段，与网络爬虫管理相关的行政法律依据主要可以分为以下几个方面：

首先是关于治安管理的法规条例。根据《治安管理处罚法》第二十九条的规定，有几种和爬虫行为相关联的行为类型及相应的惩罚措施被明确定义，尤其是其中的前三个项目与爬虫活动的关联更为紧密。这些都基于“违反国家规定的条件”而设定。这三种情况分别是：“未经授权进入计算机信息系统并导致损害”、“破坏或影响计算机信息系统的功能，使其无法正常运作”以及篡改或添加或删除系统数据的情况。这个规则所维护的利益类似于《刑法》中提到的相关条款。

首先是关于互联网安全的法律法规体系及其与数据库的安全保障法规。主要有国家层次的相关规定如“中华人民共和国网络安全法”“中国人民共和国个人信息保密管理制度”，“电脑资讯系统的保安管理办法”，“等等，”此外还有一些地方法律规则比如广东地区发布的“电子政务网站”的规定；山西地区的“政府部门信息化建设标准”；辽沈两地的“公共服务网点标准化工作实施方案”；宁夏州出台了有关的信息化建设的指导意见等相关政策文件都属于此类范畴内。这些规定的具体内容虽然没有明确针对如何处理违规采集行为的具体条款或者细则但是其对于整个行业的发展有着重要的影响并能为企业提供一定的参考价值。

三是对个人隐私保障的法律法规。《个人信息保护法》的出台使得政府机构能够利用合法的法律基础和灵活的管理策略来应对网络爬虫问题。该法规的第六部分详细列举了负责维护个人隐私权益的各部门应采取的一系列管理手段，而第七部分则规定了相应的行政罚款和信誉惩罚机制，这有助于强化针对爬虫行为的监管力度。特别是条款六十一中的第四点“调查和处理违法行为”，从理论角度看，它能覆盖到所有的非法获取个人信息的行为。此外，这些关于个人隐私保护的法律也通过互联网服务提供商的责任体系起到预防恶意爬虫的作用。相关的规则向网络服务供应商施加了更高的安全责任标准。这些法律依据为防止爬虫风险的相应行政管控提供了有效但适度范围内的权力。在中国国内的实际执行过程中，这种法律应用的方式存在一定差别，从而形成了具有明确重点的行政管控结构。

(二) 网络爬虫的行政执法实践

目前，关于网络爬虫行为的行政惩罚在网络监管执行过程中相对稀缺，同时，能有效规范此类行为的主要法规也并未广泛应用于实际操作中。虽然《治安管理处罚法》第 29 条、《网络安全法》第 27 条和第 63 条等条款在行政执法过程中的使用率较低；而《个人信息保护法》及《数据安全法》则刚出台不久，尚无显著的影响力行政执法实例出现。然而，尽管如此，行政执法依然通过以下三点对网络爬虫相关的行为实施了间接的管理，从而限制了其非法使用的空间：

首先，公安部门依据实施《网络安全法》中的网络级别保护政策所做出的各种行政惩戒。比如，北京山顶洞商业贸易公司的网页包含了代码运行缺陷，被北京市公安局房山区分区认为没有履行网络安全的保障责任，依照《网络安全法》第 21 条和 59 条对其进行了告诫性的处分(京公房行罚决字[2021]50915)。

虽然《网络安全等级保护基本要求》并未明确提出针对爬虫技术的安全防护手段，然而如果公司能严格遵守这些规则，那么至少三级及以上的安全计算环境标准就能有效地抵抗大部分恶意网络爬虫攻击，从而避免国家安全受到威胁或者导致严重的社会秩序和社会公众利益受损；即使是仅达到一级或二级的安全水平，也能够为我们预防爬虫技术的过度使用提供了有益的技术和管理体系支持。

其次是对于市场监管和证监会等单位采取了严格的惩罚措施来打击那些通过爬虫手段进行不正当竞争或散布虚假信息的违规行为。比如，前述案例中，无锡市市场监督管理局就对外卖服务平台使用爬虫工具窃取竞品商家数据并强迫其选择合作伙伴的行为进行了严正处理，罚款金额达到了一百万元人民币[8]。

尽管这种严厉的惩罚并非针对爬虫本身，但却能够对其不当使用产生强大的震慑力量，从而降低了网络爬虫被滥用的风险。

三是对违反规定的爬虫行为由网信、公安及金融监管机构实施监测与警告，以期通过信用体系、区块链等领域来消减非法爬虫的存在环境。比如，当某个区块链商业项目的申报需要备案时，当地公安机关会亲自上门审查，并且特别强调“避免利用爬虫技术获取数据的行为是违法行为”[9]。这种做法并非总是伴随着明确的法律裁决，而更像是一种预先的行政指引或者“推动式”的管理手段，然而它对于防止网络爬虫的不法运用也具有一定的积极影响。

总的来说，目前执行法律的过程中，对于爬虫这一工具持有一种相对平衡的态度，只实施了一些基本的防范手段，并对其违规获取信息的做法进行了惩罚。这种处理方式是符合爬虫本源性质的。由于爬虫只是一种科技或者工程产物，它自身并没有绝对的合法或非法属性，因此，爬虫的相关活动是否构成违法行为，需要依据“反反爬策略”对电脑信息系统的干扰程度及其收集的信息用途来作出判定。鉴于违反爬虫可能轻易达到犯罪标准的情况，政府部门通过向爬取者和潜在受害者两个方向同步展开经济高效且限制适中的安全防护任务，对该项技术持有适当的宽松度和平等对待，这无疑是一个明智的选择。

尽管如此，对于网络爬虫的管理仍存在许多待改进的地方。首先，随着爬虫行为是否合法的问题逐渐明确，我们也需要进一步明晰其违法行为与合法行为间的分野，特别是在区分那些虽然违反了法律法规却尚未达到犯罪程度的行为时尤为重要。接下来，当“合法与非法”和“违法犯罪”的界限变得更清楚后，我们也应该重新审视网络爬虫中刑事与行政责任的不平衡状况。如果能够适当地增强处理网络爬虫行政违规事件的调查权力或者优化由刑事案件转至治安案件的过程，并在两者之间设定合适的过渡阶梯，这有可能使得刑法适用更为贴合克制的原则要求。再者，关于处罚与指导、阻碍与促进的关系也有必要加以调适，因为所有针对网络爬虫的处罚都是一种警示性和负累性质的举措，而行政法的一大优点就在于拥有大量的激励方式，并且已经在形式上发展出了一种名为“引导式管理”的新模式[10]。政府部门可以通过设立通用的爬虫和专注型爬虫的标准，并结合运用各类激励策略来有效地指引爬虫技术的发展方向。此外，政府也可以借助超脱于爬虫之外的支付式管理工具来实现对网络爬虫的间接监管。因此，丰富和完善网络爬虫的行政管控系统，可以在这些领域做出努力，进而构建一套既包括防范、导向和惩治功能的全方位管制方案。

5. 对网络爬虫的行政监管体系进行充实和完善

全球范围内的网络爬虫法律风险主要以民用与刑罚法规为响应方式。然而，尚未形成一套完整的专职管理策略。鉴于我国已经广泛介入互联网行业并且不断创新的历史背景下，面对如网络爬虫监管等问题的新兴挑战，我们无需依赖国外已有的成功案例，而是应根据上述制度优化的需求，务实地制定出针对网络爬虫管理的“中国模式”。预计，未来的网络爬虫控制将会越来越倚重行政管制手段，因此，通过较为灵活且专业的行政立法来健全其法律法规基础和执行流程，将成为提升网络爬虫管控水平的关键所在。

(一) 明确规范性文件的基本原则和合法性边界

立法对爬虫活动的管理(包括行政管理)必须首先确定管理的基本原则, 然后根据这个原则来设定爬虫活动的合法性范围。

当前, 学者们和实践者正逐步达成关于可接受爬取区域和不可接受爬取区域的一致意见。一方面, 随着整个社会的数据资产价值提高, 让公司和个人可以在一定的范畴内使用专门的网络爬虫来深度发掘这些数据的潜在价值似乎已成为一种必经之路; 另一方面, 网络安全法规、数据保密法律法规以及个人信息保护法规日益完备, 这意味着不能容忍爬虫跨越某些网络安全的防护手段, 也不能纵容其违法采集被保护的信息或窃取个人信息的行为发生。此外, 如果我们坚持社会公正的原则, 那么我们也无法容许爬虫侵占普通人在互联网服务和数据资源方面的权益。所以, 在风险管理和价值创造之间寻求最佳平衡点, 通过法律规定明确爬虫行为的合规底线, 鼓励并推动爬虫的使用走向合法化, 同时根据违规行为的严重程度实施适当的比例原则下的限制和处罚, 这是合理的做法。

互联网爬虫的合法范围很大程度上取决于法律法规中的禁令条款, 这对于政府立法来说至关重要。目前, 相关法律条文中包含了诸如未经授权进入他人的网络系统、盗取网络资料、破坏网络安全、违法获取个人身份信息、侵犯私人权益、侵权版权、损害计算机信息系统安全及干扰通讯自由和保密等诸多禁令内容。其中许多具体的条例都源自政府立法, 因此无法仅依赖狭义的法律规则来明确定义所有禁令内容的含义。而行政法规、政策甚至其他的正式文件将在两个方面发挥关键作用: 首先是为网络安全、数据保护和个人隐私等利益提供准确且详细的规定; 其次是以某种形式的“软法”准则来判断“反对反爬行动”在特定情况下的合法性, 特别是在技术标准的指导下, 明确哪些行为属于“威胁网络安全”、“盗窃网络数据”或者“损坏计算机信息系统”等违规行为。这种广泛性的调整虽然可能看起来复杂繁琐, 但有助于更清楚地描绘出互联网爬虫和其他相似在线活动的合法界限, 从而推动更加精准的执法工作、司法审议和公司合规制度的发展, 同时也能降低大量的社会经济开支。

(二) 平衡使用行政处罚和刑事处罚

正如之前提到的那样, 当谈到互联网抓取程序的责任问题时, 我们发现存在一种适用的偏差现象: 虽然许多法律法规赋予了针对这种行为实施惩罚的可能性(例如《个人隐私保密条例》、《公共秩序维护法规》和《信息技术保障规则》等), 但要将其转化为实际执行的过程则需依赖特定的流程来支撑。为了实现两者之间的均衡使用并达到理想的效果, 我们可以通过以下途径加强对于此类违法行为采取行政制裁的方式。

首先引入合法奖励制度。为了充分利用行政惩罚的功能, 必须在刑诉法规制下确定合法奖励空间。合法奖励主要有两类: 行政监督奖赏与刑法奖赏^[11], 这两种方式并非一定要一起使用; 当执行刑法奖励时, 应避免过度或过少的使用行政惩罚, 这可能会对增强刑法奖励效果有所助益。针对那些网络爬虫违法行为较轻的企业, 一旦警方掌握了足够的证据并决定起诉, 可考虑采用合规无罪处理配合行政罚款的方式来实现, 以让这些公司真切感受到法律法规的力量从而迅速调整他们的爬虫行为。

首先, 我们需要构建一套证据准则。公安机关可通过制定相关法规或者其他的规范性文件来设立特定的证据规则, 以接收来自企业的投诉案子与提交的相关资料为基础, 从而决定是否要对违规的网络爬虫行为实施行政惩罚。尽管企业在提供关于违法行为的证据方面可能存在一定的挑战, 但是并非完全无法应对。他们也能够运用一些科技工具去辨别和跟踪侵犯者, 比如网站可以通过强制用户输入账户密码、使用数据标记等方式, 每当有新的访问者登陆时都会对其操作情况做详细纪录和分析, 同时还可以借助最新的互联网数据取证系统来捕捉到爬虫方的或是第三方的盗用数据的直接证据。这些从企业那里收集到的侵权情报根据特定规定有可能被转换成符合行政诉讼要求的证据, 这有助于警方或网络安全管理部

门追责网络爬虫的行政法义务。

三是对合法惩罚范畴进行合理的扩展。例如,根据《网络安全法》来创建特定的执行细则或者实行方案,这样一来,执法部门就可以把打击对象扩大到那些提供非法运营的打码服务平台、代理IP池等等这些对于网络爬虫有关键支持功能的中介服务商。各地区的公安机关可以通过使用《网络安全法》中的“没收非法收益+大额罚款”的责任机制,在侦破与网络爬虫相关的刑事案时,同时也可以对未达到犯罪标准的违法者施加行政制裁,并在适当的时候通过公布行政处分的方式,借助这种行政处分的震慑力量去检查和清除违规行为,以此实现对网络爬虫技术过度使用的有效预防和管理。

(三) 引导合理利用爬虫技术

除了以承担法律责任来阻止非法行动之外,行政监管还可以利用支付和指导方式鼓励个人做出合法并有益于公众的行为选择。从理论角度来看,这种导向型管理策略可以涵盖多种形式,如财政支持、资格认定、奖赏、税收优惠、资源分配等等。然而,针对网络爬虫这一特定领域,因为它具有强烈的商业驱动力,并不一定需要额外使用财务刺激来激发积极性,而应更加关注如何精确把握合法性和盈利的空间。这部分的管理措施主要是基于技术的标准化设定和公开不良信息的发布等方面。

爬虫的合法性范围可以通过技术标准更加明确和精细地界定。技术标准通常包含一些规范性要素,即使是纯粹的技术标准也有助于明确合法性的范围。在司法实践中,技术标准对行政、民事和刑事审判中的事实认定起着重要的实质性作用,帮助司法机关判断合法与非法、罪与非罪的范围。目前只有极少数行业或地方性标准明确规定了与爬虫相关的内容,且这些标准主要针对无害的通用爬虫。根据《数据安全法》第17条的规定,多方主体可以参与制定数据开发利用技术、产品和数据安全相关的标准。在国务院标准化行政主管部门以及网信、公安、工信等相关部门的推动下,爬虫方面的技术标准将有能力为合法的爬虫活动提供专业、明确和精细的全面指南。

对于非违法行为主体的其他公司或个体来说,揭示负面的信息能够产生实际的行为指导效果。在某些条件下,这种方式可以作为一种有效的监管工具,通过提供警示来实现目标。虽然大部分参与互联网爬取活动的小型企业和个人可能会忽视品牌声誉受损和公众舆论的影响,但是政府机构中较高层次的人员可以通过公开网络爬行相关的非法信息,让各方了解合法性的范围,同时也能使得执行人员更加明确并精确地理解新型商业模式下的合规情况。特别是在“反对反爬行动”的信息公布上尤为关键。几乎没有使用“反对反爬行动”的企业级集中式爬取是相对少见的,而各种各样形式丰富的“反对反爬行动”是否会被视为违规,以及违规的严重程度等信息,都是新一代大数据服务行业从业人员非常重要的风险控制信息。如果公安部和网信办能在这一领域不断更新和发布带有一定权威性质的负面案例,那么就能提前预防潜在的法律风险问题,防止那些原本并没有犯罪意图的人陷入法律困境。

参考文献

- [1] 刘艳红. 网络爬虫行为的刑事规制研究——以侵犯公民个人信息犯罪为视角[J]. 政治与法律, 2019(11): 16-29.
- [2] 央视网评:“爬虫”滥觞 窃取数据的行径该整肃了[J]. 电脑知识与技术(经验技巧), 2019(11): 113.
- [3] 周德懋, 李舟军. 高性能网络爬虫: 研究综述[J]. 计算机科学, 2009, 36(8): 26-29, 53.
- [4] 韩轶, 聂晶. 加强规范, 不让网络爬虫变“害虫”[N]. 光明日报, 2021-03-20(07).
- [5] 杨华权. 论爬虫协议对互联网竞争关系的影响[J]. 知识产权, 2014, 24(1): 12-21.
- [6] 苏宇. 网络爬虫的行政法规制[J]. 政法论坛, 2021, 39(6): 41-53.
- [7] 高富平. 数据流通理论数据资源权利配置的基础[J]. 中外法学, 2019, 31(6): 1405-1424.
- [8] 王士浩. 无锡“外卖大战”案例的法律透视[N]. 中国市场监管报, 2019-11-26(007).
- [9] 邢萌. 剖析第二批区块链服务备案名单: 公安网监部门现场核验 一半申报项目被否[N]. 证券日报, 2019-10-

25(B2).

- [10] 苏宇. 区块链治理的政府责任[J]. 法商研究, 2020, 37(4): 59-72.
- [11] 陈瑞华. 论企业合规的中国化问题[J]. 法律科学(西北政法大学学报), 2020, 38(3): 34-48.