论刑法介入网络数据爬取行为的类型与限度*

□ 姜 瀛

内容提要 网络爬虫是一种互联网信息自动采集技术，单纯利用爬虫技术无法进入后台服务器。突破反爬措施意味着爬取方规避了被爬取方的访问限制，可以伪装成普通用户获得客户端“访问资格”并获取数据，但爬取方并未侵入被爬取方计算机信息系统。只有利用爬虫技术爬取“公民个人信息”等刑法特别保护的数据或是利用其他技术非法侵入计算机信息系统后爬取系统数据的行为，才可构成犯罪。实践部门采取“控制性标准”作为计算机信息系统数据的解释依据，扩大了计算机信息系统数据的认定范围，将突破反爬措施爬取客户端一般数据的行为认定为非法获取计算机信息系统数据罪，有“司法犯罪化”之嫌。针对突破反爬措施爬取数据的行为，刑法应事后介入。在被爬取方先申请法院通过“行为保全”措施来禁止相关数据抓取行为后，若爬取方仍然违反“行为保全”裁定，可以适用“拒不执行判决、裁定罪”予以规制。

关键词 网络爬虫 突破反爬措施 计算机信息系统数据 司法犯罪化

作者姜瀛，法学博士，大连海事大学法学院副教授。（大连 116023）

DOI:10.14167/j.zjss.2021.10.005

一、问题的提出

网络爬虫是一种互联网信息自动采集技术。[①]在“数据为王”的时代，利用爬虫技术爬取数据已经成为直接采集用户数据之外常规的数据挖掘途径。事实上，对于爬虫技术的应用，我们并不陌生，搜索引擎即是爬虫技术的典型例证——在技术上属于针对不特定网站的“通用爬虫”（General Purpose Web Crawler）。实践中，容易引发法律争议问题的是另一类爬虫，即针对特定主题或单一网站进行数据爬取的“聚焦爬虫”（Focused Web Crawler，也称“定向爬虫”或“主题爬虫”），本文的研究对象正在于此。

近年来，利用网络爬虫爬取数据行为引发的争议问题越来越多，“晟品公司利用网络爬虫抓取数据”[②]一案便引发学界与实务界的广泛关注。对于该案，有肯定观点认为：“本案中的网络爬取行为已经超过了合法边界，属于侵入计算机信息系统的手段行为；被告采取突破被害人反爬安全措施的技术手段，未经许可进入计算机系统获取数据，构成非法获取计算机信息系统数据罪。”[③]另有学者指出：“行为人未经许可，强行突破反爬技术，侵入‘国家事务、国防建设、极端科学技术领域’之外的计算机信息系统，并采用爬虫技术获取该系

* 本文为国家社科基金青年项目“刑法立法模式与修改方式研究”（19CFX038）、中央高校基本科研业务费项目“网络数据爬取行为刑法规制问题研究”（3132021287）的研究成果。

统内的数据，其行为可构成非法获取计算机信息系统数据罪。”④当然,也有观点对本案的判决结果提出质疑,主张“被告人破解被害人设置的防抓爬措施并不等同于侵入被害人的计算机系统”。⑤

在“晟品公司利用网络爬虫抓取数据”一案中,“强行突破反爬措施来爬取数据”的刑法定性成为定案的关键问题。进而言之,这一问题又可细化为三个层面。其一,利用爬虫技术爬取的数据具有哪些特征，突破反爬措施后所爬取的数据在性质上是否发生了实质变化;其二,如何把握非法获取计算机信息系统数据罪的法益定位，如何解释该罪所规定的“计算机信息系统数据”;其三,突破反爬措施所爬取的数据与非法获取计算机信息系统数据罪中的“计算机信息系统数据”是否属于同一范畴，可否将网络爬虫爬取的数据解释为计算机信息系统数据。

客观而言,“晟品公司利用网络爬虫抓取数据”一案所涉及的网络爬虫刑法规制问题，既是大数据时代的新兴技术问题，又是一个刑法领域中的常规“论题”,也即实践部门是否将刑法缺乏明文规定的行为在司法上予以犯罪化。事实上,对于突破反爬措施爬取数据的行为,法律人士、互联网行业从业者以及普通民众都可以认识到其危害性,犹如任何人都可以看到食物的外表，但专业的刑法思维却可以透过表面认识食物“内里”,评价某一违法行为是否真正地触碰到刑法所保护的特定法益。⑥基于此,本文尝试对网络爬虫技术进行规范分析，明确网络爬虫应用对象以及突破反爬措施的法律属性。在此基础上,本文将结合典型案例对利用爬虫获取数据的主要类型进行分析，分别评价不同行为类型的刑事违法性，并厘定刑法规制网络爬虫的边界。本文的最终目标是确立刑法介入大数据产业的合理路径。

二、网络爬虫的应用对象及其与计算机信息系统数据的关系

(一)网络爬虫的应用对象及其技术特征分析

从规范层面来讲，网络爬虫的应用对象涉及到不同的数据类型，有研究将之归纳为公民个人信息、商业秘密、涉著作权信息以及一般数据,⑦应当看到，现有研究所采取类型化分析并没有将传统计算机信息系统犯罪的技术性维度考虑在内，具有一定的局限性。在本文看来,对网络爬虫应用对象与反爬措施的技术性解读，乃是探讨刑法规制网络爬虫问题的基本前提。

首先,从技术原理上讲,(聚焦)爬虫按照预先定义的爬取主题在给定初始“统一资源定位符”(Uniform Resource Locator,简称 URL,实际上就是我们所称的“网址”)种子集后,根据一定算法爬取数据并进行分析，并在抓取数据过程中不断将新的 URL 放进待爬行的 URL 队列中。⑧而 URL 作为从互联网上获得资源的位置和访问方法的简洁表示,是完全开放的,爬虫解析 URL 与普通用户访问在技术上并没有区别。⑨网络爬虫实际上是在“客户端”或“客户端与服务器端的接口”进行数据获取操作的，通过模仿普通用户正常发送数据请求，等待服务器向其传输数据后在客户端爬取数据。⑩易言之,网络爬虫只不过是一种可以通过模仿普通用户并且可以更高效地收集并处理客户端数据的技术而已。网络爬虫并不是什么“黑客技术”,不能将之视为非法侵入计算机信息系统的程序或工具。单纯利用爬虫技术不可能获得任何进入后台的权限或机会,也即无法进入服务器端,仅仅凭借爬虫技术并不会触及到服务器的系统数据。当然,行为人若是利用其他技术手段非法侵入服务器端，其同样可以利用爬虫技术获取服务器端的计算机信息系统数据。

其次，反爬措施是指利用某种技术手段阻止他人利用爬虫技术大规模访问自己客户端的方式,“IP 访问量限制、session 访问量限制、User-Agent 限制以及设置登陆验证码”都属于实践中常见的反爬措施。⑪从技术上讲,被爬取方所设置的反爬措施，均属于前端代码，全部运行在爬取方(访问者)自身的计算机上。由于反爬措施属于前端代码,即使突破反爬措施,爬取方也并未进入被爬取方的服务器，因而不会威胁到被爬取方的计算机信息系统安全。突破反爬措施意义在于爬取方规避了“访问限制”,也即伪装成普通用户非法获得了“访问资格”。无论是对于无视网站设置“robots 协议”⑫随意抓取网站数据的行为,还是对于突破反爬措施抓取数据的行为，我们所要判断的仍然是数据的性质。事实上,上述无视“robots 协议”甚至是突破反爬措施而违规爬取数据的行为,其爬取对象仍然是客户端数据。本质上来讲,被爬

方设置反爬措施的目的在于防止大规模的机器访问并获取数据，但对于普通用户（个体）而言，访问网站或 APP 显然是允许的。概言之，仅仅突破了反爬措施，并不能侵入被爬取方的计算机信息系统，所获取的数据也并非是严格意义上的计算机信息系统数据。

最后，爬取刑法特别保护的数据可能导致爬虫技术应用涉嫌犯罪。换言之，抛开爬虫技术是否具有违法性不谈，因数据本身属于刑法特别保护对象，如爬取目标网站中的公民个人信息、目标公司的商业秘密、涉及版权内容的信息或者是非法入侵计算机信息系统后利用爬虫技术获取系统数据等等，获取数据行为本身便可能直接构成犯罪。

（二）“计算机信息系统数据”的规范解读

2009 年 2 月 28 日，全国人大常委会通过的《中华人民共和国刑法修正案（七）》（以下简称《修七》）增设了“非法获取计算机信息系统数据罪”，也即《中华人民共和国刑法》（以下简称《刑法》）第 285 条第二款：“违反国家规定，侵入前款规定以外的计算机信息系统或者采用其他技术手段，获取该计算机信息系统中存储、处理或者传输的数据，或者对该计算机信息系统实施非法控制，情节严重的，处三年以下有期徒刑或者拘役，并处或者单处罚金；情节特别严重的，处三年以上七年以下有期徒刑，并处罚金。”从司法实践情况来看，该罪逐渐成为刑法规制网络爬虫的重点罪名。在对刑法规制网络爬虫的司法实践进行反思的过程中，我们有必要明确非法获取计算机信息系统数据罪的法益定位，尤其要对该罪中“计算机信息系统数据”作准确解读。

一方面，通过考察立法目的，我们可以发现，《修七》之所以增设非法获取计算机信息系统数据罪，是因为一段时间以来不法分子（黑客）大肆非法侵入他人计算机信息系统并非法获取计算机信息系统中储存、处理或传输的数据，具有明显的社会危害性，而我国《刑法》之前并没有将此类行为作为犯罪加以规定。[13]正是为了规制上述行为、弥补法律漏洞，立法机关增补非法获取计算机信息系统数据罪，该罪所针对的对象是使用中的计算机信息系统中存储、处理、传输的数据，这有利于更为全面地保护计算机信息系统安全。事实上，计算机信息系统数据的控制者均会采取特定的安全保障措施，确保计算机信息系统的访问权限仅向特定主体开放，而对其他主体默认为关闭。因此，从违法性角度来讲，是否获得计算机信息系统控制者的权限授予，成为判断进入计算机信息系统并获取数据的行为是否具有违法性阻却事由的主要依据。具体来看，违法获取计算机信息系统数据可包括“侵入计算机信息系统并获取计算机信息系统数据、未经授权‘擅入’或超越权限获取计算机信息系统数据及采用其他技术手段非法获取数据”三种形态。[14]

另一方面，从体系解释角度来讲，计算机信息系统所具有的独特的技术属性以及虚拟空间特有的安全诉求，[15]决定了《刑法》第 285 条专门设置罪名并将之体系化的必要性，而这种体系化整合的可行性源于罪名之间法益定位的一致性。显然，《刑法》第 285 条第二款非法获取计算机信息系统数据罪与该条第一款非法侵入计算机信息系统罪保护的法益应当是相同的，均为计算机信息系统安全；否则，罪名设置的体系安排难以自恰，违背基本的立法逻辑。因此，该罪中非法获取的对象，应当是危及到计算机信息系统安全的数据。换言之，因为非法获取计算机信息系统数据可能会破坏计算机信息系统安全或者是造成了危险隐患，所以立法者才增设了“非法获取计算机信息系统数据罪”予以规制。

此外，还需要说明的是，《刑法》第 285 条第二款所规定的“或者采用其他技术手段”，其更多程度上是一种兜底条款，无论基于何种解释立场，“采用其他技术手段”所能达到的效果也应当是“侵入前款规定以外的计算机信息系统”。如果单纯利用爬虫技术以及突破反爬措施不能非法侵入服务器端的计算机信息系统，那么，爬虫技术以及突破反爬措施便不属于《刑法》第 285 条第二款所规定的“采用其他技术手段”。

（三）将网络爬虫爬取数据解释为“计算机信息系统数据”的路径与困境

通过网络爬虫的技术分析以及对“计算机信息系统数据”的规范解读，我们可以得出两个基本结论。第一，突破反爬措施也仅仅是为了获得访问资格，但无法侵入被爬取方的计算机信息系统，网络爬虫仍然是以客户端数据为爬取对象。第二，非法入侵计算机系统并获取系统数据行为本身即具

有刑事违法性，是否使用爬虫技术来获取系统数据并不会对行为违法性的认定产生实质影响。由此而言,“利用爬虫技术以及突破反爬措施的行为”与“具有刑事违法性的非法侵入计算机信息系统的行为”之间难以产生规范意义上的直接关联。

当然，即使突破反爬措施并不能侵入被爬取方的计算机信息系统,可否通过对“计算机信息系统数据”的解释,来实现将利用网络爬虫爬取的数据涵摄于“计算机信息系统数据”之下的效果呢?笔者认为,对于《刑法》第 285 条第二款所规定的“计算机信息系统数据”，存在着两种不同的解释思路,即“控制性标准”与“技术性标准”。

“控制性标准”以“保护数据控制权”作为解释的核心坐标，侧重数据权利人主观上对数据的控制意愿,该标准可能扩大“计算机信息系统数据”的认定范围。申言之,数据控制权利人设置反爬措施表明了其不希望数据被竞争对手爬取的控制意愿,权利人控制数据的主观态度与立法者对于“计算机信息系统数据”的严格保护意愿在一定程度上是相通的。也即,依托于“控制性标准”,网络爬虫所获取的、由反爬措施所保护的数据,系数据权利人所要控制的数据,因而可以解释为“计算机信息系统数据”；突破反爬措施爬取该数据的行为——即使并未侵入计算机信息系统，将被解释为“获取该计算机信息系统中存储、处理或者传输的数据”,最终被认定为“非法获取计算机信息系统数据罪”。

不过,以“控制性标准”为依据的解释思路面临“技术性标准”的质疑,引发罪刑法定层面的困境。“技术性标准”源于“非法获取计算机信息系统数据罪”的立法背景,强调以计算机信息系统的技术性风险作为解释坐标，侧重在客观上评价计算机信息系统安全是否遭受到威胁,其可以限制“非法获取计算机信息系统数据罪”的适用范围,避免刑法的不当扩张。若是依据“技术性标准”,那些未侵入计算机信息系统而获取的数据或者是未威胁到计算机信息系统安全而获取的数据，在技术上不会对计算机信息系统带来任何危险，并不符合非法获取计算机信息系统数据罪的立法预期与法益定位,因而不应被纳入到该罪的规制范围。由于单纯的突破反爬措施并不能侵入计算机信息系统，爬取的数据也不会给计算机信息系统带来任何危险,难以被解释为“计算机信息系统数据”。

三、网络爬虫的应用类型与刑法规制的边界反思:基于典型判例的分析

(一)爬取违法数据:直接构成犯罪

利用爬虫技术获取数据的行为，只要数据本身属于刑法特别保护对象,爬取行为即可获罪。其中,爬取“公民个人信息”是司法实践中最为常见的涉罪情形。我国《刑法》第 253 条之一第三款规定:“窃取或者以其他方法非法获取公民个人信息的,依照第一款的规定处罚。”在没有得到被爬方或个人用户同意的情况下,行为人利用爬虫技术获取网站公民个人信息,便可能构成犯罪。诸如“马某编写爬虫程序窃取网站用户个人信息”一案[16]、“谢财安等盗取京东商城卖家账号、密码后利用爬虫技术(“smarttool”软件)非法获取用户个人信息”一案[17]、“魏江蒙通过‘网络爬虫’程序下载工商个体户资料”一案[18],都是属于利用爬虫技术非法获取公民个人信息的案例。当然,上述案件与以其他方式非法获取公民个人信息的案件差异性不大,法律适用中争议问题也不多。

不过,在“李威侵犯公民个人信息”一案中,[19]部分案件事实认定与法律适用存在争议。该案判决书指出:“被告人李威在北京某公司任职期间,抓住系统漏洞直接访问服务器端后台,利用‘八爪鱼软件’(一种常见的爬虫软件）获取客户个人信息数据。”对于这一部分事实,判决书认为:“被告人李威作为某公司员工，通过自动化软件收集公民个人信息的行为不具有非法性，并未违反国家有关规定。”因此,法院仅认定了李威(在上述事实之外）实施的在网上以购买和交换等方式非法获取公民个人信息的行为，构成侵犯公民个人信息罪。该案的疑问在于,李威作为公司的高级管理人员,享有获取或保存公司用户个人信息的权限,因而其获取公司用户个人信息的行为并不具备违法性。但李威获取个人信息是利用了系统漏洞进入后台,即非法访问服务器系统,因此,其所获取的数据在性质上又属于服务器后端存储的系统数据，故李威的上述行为符合非法获取计算机信息系统数据罪的构成要件。概言之,本案在核心事实认定上忽视了数据的两面性，即数据本身可能同时具备公民个人信息属性与计算机信息系统数据

属性。此外,需要强调的是,本案中,行为人利用爬虫技术只是提升了数据获取的体量与速度而已,爬虫本身对本案的定罪并没有产生实质影响。易言之,爬虫技术只是高效获取数据的工具而已,并不是非法侵入计算机信息系统的工具。如果行为人利用系统漏洞非法侵入计算机信息系统或者是超越权限访问计算机信息系统,后利用爬虫软件获取计算机信息系统数据,虽然可能构成犯罪,但这与爬虫技术本身无关。

此外,司法实践中还存在一种较为特殊的类型。行为人可能在约定的数据用途范围之外,再次私自使用上述数据并利用爬虫技术从事其他非法牟利活动。事实上,上述违背数据控制者意愿、超越约定用途再次利用计算机信息系统数据的行为也可以解释为"非法"获取控制者所专有的计算机信息系统数据。"北京瑞智华胜公司非法获取计算机信息系统数据"一案就属于此类案例。[20]该案判决书指出:"北京瑞智华胜公司通过其他关联公司与运营商签订精准广告营销协议,获取运营商服务器登录许可,通过部署 SD 程序,瑞智华胜相关人员从运营商服务器抓取采集网络用户的登录 cookie 数据,并将上述数据保存在运营商 redis 数据库中。后利用研发的爬虫软件、加粉软件,远程访问 redis 数据库中的数据,非法登录网络用户的淘宝、微博等账号,进行强制加粉、订单爬取等行为,从中牟利。"

该案中,涉案公司与运营商签订精准广告营销协议,自然就获取运营商服务器登录许可、具有访问计算机信息系统权限,并且可以依照约定使用网络运营商的用户登录 cookie 数据。然而,涉案公司只能基于履行其与网络运营商的约定义务(精准广告营销协议)在特定范围内使用上述用户登录 cookie 数据。即使涉案公司并不存在非法侵入或超越权限访问的实质行为,但其违背数据控制者意愿、超越约定用途再次利用计算机信息系统数据的行为,显然会侵犯到网络运营商对数据的控制权,仍然可以构成非法获取计算机信息系统数据罪。一些学者与实务人员曾经以"利用爬虫加粉软件'打劫'个人信息牟利如何适用法律"为题对该案展开学术研讨。[21]但实际上,该案的关键事实在于行为人违背数据控制者意愿、超越约定用途再次利用计算机信息系统数据,爬虫技术本身并不是其获罪的根本因素。

(二)突破反爬措施爬取一般数据:刑事违法性的具体分析

如前所述,利用爬虫技术爬取数据,无论其是否违背被爬方的意愿——无视 robots 协议甚至是突破反爬措施,所爬取的仍然是客户端数据。进而言之,如果爬取的客户端一般数据本身不属于刑法特别保护的对象,使用代理 IP、使用不同的 User Agent 亦或是其他技术手段来突破反爬措施,都仅仅是一种前端的技术手段,突破反爬措施后爬取数据的行为,也是完全运行在访问者的计算机上,无法进入到被访问者的服务器系统,显然也就不能威胁到计算机信息系统数据安全。由此而言,突破反爬措施本身,并不具有刑法上的独特意义。但对于这一点,司法实务部门可能存在某种误区。

在"晟品公司利用网络爬虫抓取数据"一案中,[22]判决书认为:"被告人在数据抓取的过程中使用伪造 device_id 绕过服务器的身份校验,使用伪造 UA 及 IP 绕过服务器的访问频率限制。在数据采集过程中,被告采取了绕过或突破受害单位反'爬虫'安全措施的技术手段,未经许可进入受害单位的计算机系统,构成非法获取计算机信息系统数据的犯罪行为。"首先,判决书中指出的"使用伪造 device_id、使用伪造 UA 及 IP 绕过服务器的访问频率限制",都是突破反爬措施的常见方式,这些反爬措施运行在访问者服务器端,突破反爬措施仅仅是伪装成普通用户获取访问资格并获取数据,而"突破受害单位反'爬虫'安全措施的技术手段"后并不能进入被访问方的计算机系统。因此,判决书认为行为人"未经许可进入受害单位的计算机系统"的事实认定并不准确。

同时,在"晟品公司利用网络爬虫抓取数据"一案宣判后,该案裁判者曾指出,可基于实质性解释立场,将破坏前置的访问程序限制后实现访问并抓取数据的行为解释为非法获取计算机信息系统数据行为。[23]如前所述,爬虫技术实际上是在客户端进行数据获取操作的,而计算机信息系统数据所指的是服务器端的数据。具有开放性的客户端与强调私密性并且在安全性上具有极高要求的服务器端,系同一位阶下完全对立的范畴,犹如"户外"与"户内"。即使考虑到社会发展的客观情况,这两个对立范畴并没有发生任何变化,难以通

过实质解释来跨越二者之间的技术“鸿沟”。因此,如果严格采取“技术性标准”作为“计算机信息系统数据”的认定依据,那么,突破反爬措施爬取数据的行为虽具有一定的社会危害性,但将其解释为侵入计算机信息系统而非法获取数据,确有“司法犯罪化”之嫌。

此外,“武汉元光科技利用网络爬虫抓取数据”[24]一案也存在类似的法律适用问题。该案判决书指出:“被告人邵凌霜为提高元光公司开发的智能公交 APP‘车来了’在中国市场的用户量及信息查询的准确度,授意陈昂等公司数名员工利用网络爬虫软件获取包括谷米公司在内的竞争对手公司服务器里的公交车行驶信息、到站时间等实时数据,日均 300 万至 400 万条。爬取的数据直接为元光公司所用,使该公司的智能公交 APP‘车来了’准确度提高,造成谷米公司直接经济损失 24.43 万元人民币。”最终,法院认为,上述行为系“违反国家规定,采用其他技术手段获取计算机信息系统中储存的数据”,构成非法获取计算机信息系统数据罪。

事实上,该案同样将突破前端反爬措施后爬取客户端数据的行为认定为非法获取计算机信息系统数据的行为,其中涉及如下问题:第一,“为防止被察觉,元光公司人员不断更换爬虫软件程序内的 IP 地址并利用所设置的不同 IP 地址向酷米客发出数据请求”,表明元光公司实际上只是伪装成一般用户来获取客户端的实时数据,其并未进入到谷米公司后台的服务器系统;第二,判决书同时又认为,“元光公司非法获取公交车行驶信息、到站时间等实时数据是位于谷米公司的服务器数据”,这一理解并不准确。实际上,元光公司人员若是已非法侵入服务器端或计算机信息系统,则根本不需要 IP 地址,也无需发送数据请求,直接就可以非法获取数据;而通过更换 IP 的方式表明,元光公司人员就是为了伪装成一般用户进行访问。若严格采取“技术性标准”作为“计算机信息系统数据”的认定依据,该案同样难以成立非法获取计算机信息系统数据罪。

当然,“武汉元光科技利用网络爬虫抓取数据”一案中所涉及的公交车辆运行路线、运行时间等信息虽属客观事实,但经人工收集、分析、编辑与整合之后,通过商业运行可以带来可观效益,所以实际上已具备无形财产的特征,[25]元光公司人员抓取谷米公司数据的行为本质上属于一种不正当竞争行为。

从“晟品公司利用网络爬虫抓取数据”与“武汉元光科技利用网络爬虫抓取数据”等典型案例来看,在面对信息网络新型违法犯罪问题时,司法机关更倾向基于刑事政策上的考虑,通过刑法的“软性解释”来扩张处罚范围,[26]以彰显刑法及时回应新型问题的社会效果。然而,在突破反爬措施后爬取的数据本身不具有刑事违法性的情况下,我国刑法尚无法直接规制突破反爬措施爬取的数据行为。将“控制性标准”作为刑法规制网络爬虫的实践依据,突破反爬措施爬取的数据便被“软性解释”为计算机信息系统数据,非法获取计算机信息系统数据罪便成为刑法规制爬虫技术的“救火式”罪名。与更为严格的“技术性标准”相比,司法实践部门更愿意采取“控制性标准”,进而实现“将刑法缺乏明文规定的行为在司法上予以犯罪化”的效果。

(三)利用爬虫技术“野蛮”访问:可能构成破坏计算机信息系统罪

从技术层面看,爬虫技术最主要的功能特征在于数据获取的高效性。在单位时间内,网络爬虫“大规模机器访问”的访问数量是普通用户访问(人工访问)无法相比的。不过,大规模机器访问与同等数量的普通用户访问(人工访问)却会引发同样的效果,即增加网络服务提供者的运营负担。

实际上,网络服务器是有承载限度的,频繁的大规模机器访问占用了原本服务器用于向普通用户返回数据的带宽和运算能力,如果不加控制地利用爬虫技术持续访问,实际上会影响到网络服务的正常运行,甚至使网站崩溃而无法访问,这也会影响到普通用户的正常访问。不加控制地利用爬虫技术持续访问的行为,可能构成《刑法》第 286 条规定的“破坏计算机信息系统罪”。[27]在主观方面,行为人具有一种间接故意,也即明知不加控制的爬虫技术可能会危害到网络服务正常运行或其他人的正常访问,而放任这种危害后果的发生;而在客观方面,利用爬虫技术爬取数据的行为确实妨碍到网络服务正常运行,因而可以被解释为“功能性破坏”。其实,不加控制地利用爬虫技术持续访问行为的构罪技术路径与“DDOS 攻击”(Dis-

tributed Denial of Service，分布式拒绝服务）具有相似性，也即发送大量看似合法的访问请求，造成网络阻塞或服务器资源耗尽，从而导致普通用户无法正常访问网络资源。[29]只不过，“DDOS 攻击”主观上是直接故意，访问（攻击）行为是通过大量的“僵尸主机”（间接利用或控制他人主机）完成的；而不加控制地利用爬虫技术持续访问行为在主观上属于间接故意，访问行为是利用爬虫技术实现的，但二者在危害后果上具有相似性。因此，不加控制地利用爬虫技术“野蛮”访问，造成被访问方服务器瘫痪而无法正常运行的，可以被解释为“功能性破坏”，构成破坏计算机信息系统罪。

四、刑法规制网络数据爬取行为的路径重塑

（一）新型网络不正当竞争行为引发的刑法困境

通常而言，企业的数据来源包括“用户数据、全网抓取数据、通过数据众包形式获得、通过合作协议方式获得以及购买数据”等不同类型。为了形成大规模的数据沉淀，经营者往往要投入较高成本、提供免费服务来吸引用户；为了维持用户体量，后期仍然需要稳定投入。[29]行为人付出极高成本挖掘出数据价值，自然会想尽办法确保其数据权利不受他人侵犯。爬虫技术让某些人不劳而获，“抢夺”了本来应该通过支付对价所获取的或者是根本无法获取的数据资源，由此引发了如“武汉元光科技利用网络爬虫抓取数据”的不正当竞争案例。

有学者认为：“正是有了爬虫技术违反民事行为生活准则和道德即构成不正当竞争的判决，才有了不久之后全国首例使用爬虫技术构成刑事犯罪的案件（即“晟品公司利用网络爬虫抓取数据”一案）。从爬虫行为反不正当竞争案到全国首例爬虫行为入罪案，展现了爬虫行为从民事违法到刑事违法的司法认定过程，它充分体现了我国司法实践对爬虫行为的态度。这一过程充满司法理性，因为随着爬虫行为对各大网站数据的暴力爬取、强行爬取等行为的增多，已给网络信息安全以及营运环境造成了极大的破坏。”[30]然而，在本文看来，这种“爬虫行为从民事违法到刑事违法”的“变性”过程，并没有呈现出民事不法状态与刑事违法的实质区别，反而将突破反爬措施误读为民刑分界的标准。客观而论，“晟品公司利用网络爬虫抓取数据”一案反映出司法机关在回应网络新型不正当竞争行为时不当寻求司法犯罪化之“惯习”。这种“惯习”一直存在，在面对“恶意注册”“刷单炒信行为”等网络违法行为时就曾引发质疑。可以预计，在信息技术与社会发展快速融合的过程中，新型网络问题不断涌现出来，其中既包括新型技术性违法问题，也涉及在组织结构或行为方式出现变化的新型网络违法业态，尤其是技术性问题与新型违法业态相互结合，冲击到学者与实务部门的传统认知，挑战现有的刑法体系。

应当承认，在快速变迁、日益复杂的信息时代，刑法解释应当更具目的导向性、实质性与回应性，[31]刑法功能主义或许将成为风险社会与安全国家观背景下刑法解释的新导向。与此同时，回应新型社会问题的过程中必然会出现刑法规范的供给不足，理性的司法犯罪化可以在一定程度上应对日益复杂化的社会状况，缓解刑法规范供给不足的压力。不过，对新型网络违法行为实质内涵的把握以及对其可能侵犯法益的认定均存在较大难度，司法犯罪化中“找法”的解释技术与法治边界必将成为难点问题。[32]

我国网络犯罪可以分为三种类型，即“针对计算机信息系统的犯罪”“利用计算机网络实施传统犯罪”以及“破坏网络业务活动、妨害网络秩序的犯罪”。[33]其中，“破坏网络业务活动、妨害网络秩序的犯罪”的范围与边界是最难把握的。实践中，这一类型的不法行为本身具有很强的“创新性”，常常会引发较为严重的危害结果，但往往又缺乏专门的罪名与之对应，在罪与非罪的认定上极易出现偏差。为了实现某种社会治理效果，司法机关很可能将“针对计算机信息系统犯罪”的罪名或其他传统罪名适用到这一类型的不法行为。当前，司法机关将“利用爬虫技术获取一般数据的行为”犯罪化，不是刑法功能主义引导下刑法解释的理性选择，其本质在于刑法适用中错误地解释了技术性概念，是社会治理过度刑法化的又一例证。

（二）刑法规制网络爬虫的二元结构——直接介入模式与事后介入模式

本文认为，面对突破反爬措施获取数据的行为，应在理论上建构起二元化的规制结构，也即以数据本身的刑事违法性为标准，划分出“刑法直接规制”与“刑法事后介入”两种模式。其中，“刑法直

接规制”,即刑法可以直接适用于利用爬虫技术获取“公民个人信息”行为,对此前文已作充分研讨。这里重点探讨的是“刑法事后介入”的路径选择。

本文所倡导的“刑法事后介入”,是指针对刑法无法直接规制的突破反爬措施爬取客户端一般数据的行为,应尽力避免采取“软性解释”,而应当通过适用《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》)首先寻求民事救济,而这种民事救济的重点在于利用“行为保全”措施(实际上就是“禁令”)禁止相关行为人抓取数据;在此基础上,对于仍然违反“行为保全”裁定而抓取数据的行为,可以适用我国《刑法》第313条“拒不执行判决、裁定罪”予以规制。也即,形成一种“民事措施(行为保全)前置”与“刑法事后介入”相结合的递进模式。

《中华人民共和国民事诉讼法》第100条规定:“人民法院对于可能因当事人一方的行为或者其他原因,使判决难以执行或者造成当事人其他损害的案件,根据对方当事人的申请,可以裁定对其财产进行保全、责令其作出一定行为或者禁止其作出一定行为……”该条确立了我国“行为保全”的制度依据。与此同时,国内首例“涉及爬取数据的行为保全案”[34]为我们提供了理想的实践样本。该案中,法院综合考虑申请人“深圳腾讯公司、腾讯科技公司”的申请理由与被申请人“杭州快亿公司”的答辩意见,裁定“被申请人立即停止提供微信公众号文章信息API、微信订阅号和最新文章API、腾讯滚动新闻API以及辖区内按省市微信公众号及其企业认证信息数据产品的行为”。客观而言,该行为保全裁定的核心内容在于要求被申请人停止在“神箭手”平台上为用户提供数据爬取通道,即API(Application Programming Interface,应用程序接口)。[35]

但可以预计,今后的司法实践完全可能出现“直接要求相关行为人停止爬取数据”的行为保全申请。换言之,面对突破反爬措施爬取数据的不正当竞争行为,我们完全可以凭借行为保全寻求救济。在本文看来,行为保全前置具有两个方面的积极意义,一是可以及时制止违法的抓取数据行为,快速停止损害;二是合理且必要地确立受限行为的范围,避免对合理的数据挖掘行为产生负面影响。

作为一种民事裁定,行为保全应属于《刑法》第313条“拒不执行判决、裁定罪”所保护的文书类型;故意违反行为保全裁定,完全可以构成“拒不执行判决、裁定罪”。只不过,最高人民法院于2015年7月20日发布的《关于审理拒不执行判决裁定刑事案件适用法律若干问题的解释》(法释〔2015〕16号)并没有专门针对违反行为保全裁定后所应达到的“情节严重”入罪标准做出专门规定。考虑到新型网络不正当竞争行为屡禁不止,且变化快速,最高司法机关有必要针对违反“行为保全”裁定的情形,确立相应的入罪门槛。

结　语

2020年3月30日,中共中央、国务院联合发布了《关于构建更加完善的要素市场化配置体制机制的意见》(以下简称《意见》)。《意见》提出了“土地、劳动力、资本、技术、数据”五个要素领域改革的方向,明确了完善要素市场化配置的具体举措。值得注意的是,数据作为一种新型生产要素,首次被写入官方文件。同时,《意见》明确指出,要加快培育数据要素市场。由此而言,数据的开发、权属、流转以及与之相关的规范体系的建构,已经成为大数据产业从业者、监管者以及其他数据利用者共同面对的重要课题,而缺乏规范的产业发展状态为数据要素市场的培育带来了一丝隐忧。

围绕利用网络爬虫技术非法爬取数据的相关问题所展开的刑法思辨,为我们诠释了传统刑法观念与司法逻辑在回应“新生问题”时的挑战。即使在开展数据安全专项整顿的大背景下,[36]我们也不能将爬虫技术过度地“妖魔化”。面对网络爬虫,我们所要评价的对象并不是技术本身,而是爬虫技术所针对的数据。在多数情况下,对于客户端的一般数据,即使行为人突破了数据控制者所设置的反爬措施,该行为也不具有刑法上的违法性。刑法的不当介入,不仅会对日后的司法实践做出错误引导,还会引发大数据行业的恐慌。面对与社会发展相伴相生的新型问题,为了不被错综复杂的案件事实和形形色色的行为手段扰乱思路,司法者需要提炼违法行为的本质属性,也即,首先要准确认知对象问题、做好基础定性,此后方能在治理对策上作出进一步思考,而刑法层面的对策思考应被置于最后。

注释：

①⑪唐松：《Python 网络爬虫从入门到实践》（第 2 版），机械工业出版社 2019 年版，第 1、5 页。

②㉒北京市海淀区人民法院（2017）京 0108 刑初 2384 号刑事判决书。

③㉓游涛、计莉卉：《使用网络爬虫获取数据行为的刑事责任认定——以“晟品公司”非法获取计算机信息系统数据罪为视角》，《法律适用》2019 年第 10 期。

④㉚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》，《政治与法律》2019 年第 11 期。

⑤虞元坚：《爬虫获取数据获刑案件解析及无罪论点探析——以“今日头条案”为例》，http://joint-win.com/Cn/analyse/analyse/id/171/catid/52.html。最后访问时间，2020 年 3 月 19 日。

⑥[日]中村勉：《刑法の基本思考》（改訂版），北樹出版社 2003 年版，第 16 页。

⑦李慧敏、孙佳亮：《论爬虫抓取数据行为的法律边界》，《电子知识产权》2018 年第 12 期。

⑧于娟、刘强：《主题网络爬虫研究综述》，《计算机工程与科学》2015 年第 2 期。

⑨[日]户根勤：《网络是怎样连接起来的》，周自恒译，人民邮电出版社 2017 年版，第 5~7 页。

⑩当然，爬虫技术与普通用户访问也存在某些不同之处。其一，在访问后，多数普通用户对客户端数据并不具有复制需求，而爬虫则会爬取数据并储存。其二，作为计算机程序，利用爬虫技术可以实现自动化高速访问，获取数据的频率和数量要远高于人工获取。

⑫依据中国互联网协会《互联网搜索引擎服务自律公约》第 7 条：“机器人协议（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。”事实上，“robots 协议”，就是技术界为了解决爬取方和被爬取方之间通过计算机程序完成关于爬取的意愿沟通而产生的一种机制。

⑬谢望原：《简评〈刑法修正案（七）〉》，《法学杂志》2009 年第 6 期；皮勇：《我国网络犯罪刑法立法研究——兼论我国刑法修正案（七）中的网络犯罪立法》，《河北法学》2009 年第 6 期。

⑭李遐桢、侯春平：《论非法获取计算机信息系统数据罪的认定——以法解释学为视角》，《河北法学》2014 年第 5 期。

⑮㉝陈兴良：《互联网帐号恶意注册黑色产业的刑法思考》，《清华法学》2019 年第 6 期。

⑯上海市金山区人民法院（2018）沪 0116 刑初 924 号刑事判决书。

⑰北京市大兴区人民法院（2019）京 0115 刑初 570 号刑事判决书。

⑱河南省济源市人民法院（2018）豫 9001 刑初 503 号刑事判决书。

⑲北京市通州区人民法院（2019）京 0112 刑初 62 号刑事判决书。

⑳浙江省绍兴市越城区人民法院（2019）浙 0602 刑初 1143 号刑事判决书。

㉑庄永廉等：《利用爬虫加粉软件“打劫”个人信息牟利如何适用法律》，《人民检察》2019 年第 18 期。

㉔广东省深圳市南山区人民法院（2017）粤 0305 刑初 153 号刑事判决书。

㉕李帅：《网络爬虫行为对数据资产确权的影响》，《财经法学》2020 年第 1 期。

㉖周光权：《刑法软性解释的限制与增设妨害业务罪》，《中外法学》2019 年第 4 期。

㉗《刑法》第 286 条规定：“违反国家规定，对计算机信息系统功能进行删除、修改、增加、干扰，造成计算机信息系统不能正常运行，后果严重的，处五年以下有期徒刑或者拘役；后果特别严重的，处五年以上有期徒刑。”

㉘谢君泽：《从 DDOS 攻击案件看我国网络犯罪立法的体系性缺陷》，《中国信息安全》2016 年第 2 期。

㉙何渊：《大数据战争——人工智能时代不能不说的事》，北京大学出版社 2019 年版，第 87~89 页。

㉛劳东燕：《能动司法与功能主义的刑法解释论》，《法学家》2016 年第 6 期。

㉜王华伟：《网络时代的刑法解释论立场》，《中国法律评论》2020 年第 1 期。

㉞浙江省杭州市铁路运输法院（2019）浙 8601 民初 2435 号之二民事裁定书。

㉟应用程序接口（API），又称应用编程接口，是一组定义、程序及协议的集合，通过 API 接口实现计算机软件之间的相互通信。API 的主要功能是提供通用功能集，为各种不同平台提供数据共享。

㊱近期，公安部门正开展数据安全的专项整顿，诸如“巧达科技有限公司、上海新颜人工智能科技有限公司、杭州魔蝎数据科技有限公司、中国电信控股子公司天翼征信、公信宝、同盾科技有限公司、51 信用卡”等大数据企业都牵涉其中。参见吕笑颜、石丹：《“爬虫”凶猛，大数据风控平台“黑幕”调查》，《商学院》2019 年第 10 期。

责任编辑 陈亚飞

history of political thought, we could see clearly its relationship to sovereign and sovereignty, Enlightenment and counter-Enlightenment, and national-state, and furthermore, understand the later three more.

Key words: Clausewitz; trinity of war; sovereign and sovereignty; enlightenment and counter-enlightenment; national-state

Nietzsche, Bismarck and Great Politics (30)

Yu Mingfeng
(*Department of Philosophy, Tongji University, Shanghai 200092*)

Abstract: After saying goodbye to nationalist aspirations, Nietzsche analyzed and criticized Bismarck with the concept of great politics. Nietzsche's criticism of Bismarck is actually a criticism of modern politics. The later Nietzsche used the great politics in order to overcome the nihilistic essence of modern politics. Nietzsche's great politics is not only the politics of European integration or globalization, but also a kind of value politics or spiritual politics. Whether in the early analysis or in the later application, Nietzsche measured politics with the cultural and value purport beyond politics. Realistic politics does not accept the logic of value war and does not serve culture, which is the real reason for Nietzsche's criticism of Bismarck. Power understood by Bismarck is only a manifestation of the will to power. For Bismarck's Germany, Nietzsche's philosophy of the will to power provides both a path of understanding and a critical perspective. For the current theory, investigating "Nietzsche, Bismarck and great politics" is helpful to reasonably excavate the political meaning of Nietzsche's thought on the one hand, and to understand the great political problems of our own times on the other hand.

Key words: Nietzsche; Bismarck; great politics; culture; nationalism

On the Application Field of Web Crawler and the Boundary of Criminal Law's Involvement in Big Data Industry (37)

Jiang Ying
(*Law School, Dalian Maritime University, Dalian 116023*)

Abstract: Web crawler is a kind of Internet information automatic collection technology applied to the client. It can't enter the background server simply by using the crawler technology. Anti crawling measures belong to client code. Breaking through the anti crawling measures means that the crawling party evades the access restrictions of the crawling party and can pretend that ordinary users can obtain the "access qualification" of the client. However, the crawling party still crawls the client data and does not invade the crawling Party's computer information system. Therefore, only when using crawler technology to crawl the "citizen personal information" and other illegal data, can it constitute a crime; if the crawling party breaks through the anti crawling measures, but only the general data of the client, the behavior does not have criminal illegality. In addition, the server-side data is the data of computer information system, which belongs to the unique legal interest of criminal law protection. Illegal invasion of computer information system and acquisition of such data constitute a crime, which has nothing to do with the use of reptile technology. In view of the behavior of breaking through the anti crawling measures to crawl the general data of the client in practice, the criminal law should intervene afterwards. That is to say, after the crawled party first applies to the court to prohibit the relevant data grabbing behavior through the "behavior preservation" measures, when the crawler still violates the "behavior preservation" ruling and crawls the data, it can be regulated by the "crime of refusing to perform the judgment and ruling".

Key words: web crawler; big data; data of computer information; judicial criminalization

Big Data Evidence and Its Principle on Ascertaining the Facts (46)

Yang Jiwen[1], Fan Yanying[2]
(1. *East China University of Political Science and Law, Shanghai 201620*;
2. *Law School, Southwest University of Finance and Economics, Chengdu 611130*)

Abstract: With the advent of the Big Data revolution, the relationship between evidence and facts needs to be reexamined in China. The urgency and necessity of the reform of the judicial system in the new era are highlighted in the judicial practice as the demand for the new age of evidence at the core of electronic evidence. Big Data evidence is a kind of electronic evidence, which has a special mechanism of fact-finding. The good application of Big Data can improve the ability to identify the facts in the trial system. The change of the new age of evidence, which is characterized by Big Data and electronic evidence, needs to focus on the fact finding of electronic evidence in the context of Intellectual Jurisprudence. Under the objective requirement of fact finding ability, we gradually form the scientific fact finding path from artificial intelligence to intelligence. The good application of big data evidence is beneficial to the improvement of the judicial cognition ability of the fact judge. But we need to pay attention to the scientific limitation and guarantee