Dispute Settlement 争议解决, 2024, 10(2), 1130-1137
Published Online February 2024 in Hans. https://www.hanspub.org/journal/ds
https://doi.org/10.12677/ds.2024.102155

Hans汉斯

论网络爬虫行为的入罪研究

张晓燕，王　硕

上海政法学院，刑事司法学院·纪检监察学院，上海

收稿日期：2024年1月9日；录用日期：2024年2月22日；发布日期：2024年2月29日

摘　要

网络爬虫作为一种抓取数据的自动程序具有极大的技术优势，但是网络爬虫作为一种中立性技术在遭到恶意使用后也具有其危害性。网络爬虫行为从技术角度来说一般分为对计算机系统进行访问、对系统内数据进行爬取、对爬取行为进行存储和利用三个阶段。本文旨在从刑法的视角下对网络爬虫行为进行分析，探析恶意网络爬虫行为入罪的路径，从客观不法与主观恶意两个角度判断其入罪标准。在客观层面，实质判断行为对法益造成的威胁以及侵害是否已经达到可罚的程度；在主观层面，应着重考察行为人是否具有实施犯罪的故意。

关键词

网络爬虫，入罪标准，类型化分析

Research on the Crime of Web Crawler Behavior

Xiaoyan Zhang, Shuo Wang

School of Criminal Justice (School of Discipline Inspection and Supervision), Shanghai University of Political Science and Law, Shanghai

Received: Jan. 9th, 2024; accepted: Feb. 22nd, 2024; published: Feb. 29th, 2024

Abstract

Web crawler as an automatic program to capture data has great technical advantages, but as a neutral technology, web crawler also has its harm after being maliciously used. From the technical point of view, web crawler behavior is generally divided into three stages: access to the computer system, crawling the data in the system, and storing and utilizing the crawling behavior. This paper aims to analyze web crawler behavior from the perspective of criminal law, explore the path of criminalization of malicious web crawler behavior, and judge its criminalization criteria from the

文章引用: 张晓燕, 王硕. 论网络爬虫行为的入罪研究[J]. 争议解决, 2024, 10(2): 1130-1137.
DOI: 10.12677/ds.2024.102155

two perspectives of objective lawlessness and subjective malice. At the objective level, the substance of the judgment on the threat to legal interests caused by the act and whether the infringement has reached a punishable level; at the subjective level, the focus should be on whether the perpetrator has the intention to commit the crime.

Keywords

Web Crawler, Typed Analysis, Crime Standard

1. 网络爬虫行为的性质

1.1. “网络爬虫”技术的原理和优势

网络爬虫(Web Crawler)，也可称为网络蜘蛛(Web Spider)或网络机器人(Web Robot)，属于自动网络浏览程序，根据设定的规则，通过模拟人工点击，自动捕获互联网数据和信息，从而自动高效地读取或收集互联网数据。在大数据时代，网络爬虫已成为互联网抓取公开数据的常用工具之一，可以实现对文本、图片、音频、视频等互联网信息的海量抓取[1]。作为一种中立的计算机技术，“网络爬虫”在法律上并没有被明令禁止。通常情况下的网络爬虫技术是指一般的数据爬取技术，最基础的应用领域就是搜索引擎。从某种意义上而言，只要涉及到信息收集，就一定会使用数据爬取技术，数据爬取技术的高效应用使数据的共享和广泛分析成为可能，形成了现如今网络生态的繁荣景象。

从程序上而言，数据爬取技术可分为数据收集、处理和存储三个环节。大多数“网络爬虫”技术是通过几种基于自动算法的结合实现数据爬取技术，从而在运用中呈现出极大的技术优势。第一，规模性。在应用中对自动算法程序进行设置，使得“网络爬虫”可以将数据爬取的范围提升到最大，尽可能地获取大规模数据；第二，高效性。“网络爬虫”通常只会爬取全新的网页数据，能够避免在抓取、处理和存储数据时产生重复，从而提高爬取效率；最后，精确性。“网络爬虫”能够通过算法对设置范围之外的数据进行过滤，避免提取不相关的信息，由此保证被爬取数据的精准性。

1.2. 恶意网络爬虫行为的危害性

基于前文提到的技术优势，“网络爬虫”已经广泛应用于社会的各个领域。然而，实践中存在滥用网络爬虫技术的现象，给稳定的社会带来了风险。首先，网络爬虫技术的更新往往十分迅速，但是实践中针对不法的网络爬虫技术的反制措施通常是“亡羊补牢”式的，为此，常用的反制措施只能做到提高网络爬虫行为的成本并降低其效率，往往不能达到抑制网络爬虫行为的效果。与此同时，网络爬虫行为的技术门槛较低，对编程技术进行简单的学习掌握之后即可编译网络爬虫软件，进而爬取数据。此外，在大数据时代，获取技术资源方便快捷，其逻辑结构和源代码都可以在网络上自由获取，使得爬取数据的成本较低。根据预防犯罪之原理，犯罪成本越高，行为人实施犯罪行为的可能性越小，反之，行为人实施犯罪的可能性越大。因此，在滥用网络爬虫技术而犯罪成本低的情况下，此类的将犯罪率呈现较高的现象。

恶意网络爬虫作为一种新型的犯罪手段，其犯罪行为往往是由于缺乏对网络爬虫行为的有效监管而

导致的。首先，非法入侵、控制和破坏计算机信息系统。恶意网络爬虫入侵计算机信息系统，获得系统的管理权，可以对计算机信息系统或文件进行删除和篡改，危及计算机信息系统的安全。其次，相关数据泄露，如个人隐私数据、涉及知识产权的数据等。此外，扰乱计算机信息系统的正常运行。如数据抓取过快、过频容易拥塞网络内容服务商，从而对服务器的正常运行造成影响，或者对同一个文件进行反复抓取，进而对服务器资源的消耗造成影响，破坏计算机信息系统的正常使用。

2. 恶意网络爬虫行为入罪探析

2.1. 刑法视角下的恶意网络爬虫行为

现如今，大数据的功能和内涵日趋丰富，其具有“4V”特点，即大量(Volume)、高速(Velocity)、多样(Variety)、真实(Veracity) [2]。科学技术的蓬勃发展，数字技术在社会各个领域遍布丛生。社会交往中的传统法益正随之转化出新型数据法益，而扰乱网络空间安全的方式表现为侵害代表着新型法益的数据。

理论上看，“爬虫”作为一项网络信息搜索技术，具有技术中立性，因而并未被我国现行法律所明令禁止。虽然技术本身具有中立价值，但这并不意味着其必然不受刑法规制。网络爬虫作为一种中立的技术，是一把双刃剑，具有双面性。一方面，网络爬虫技术可以促进数据广泛共享，另一方面，侵害数据信息安全的情况同样存在。虽然国内外在实践中对网络爬虫技术的处理存在“严厉”和“缓和”的不同趋势，但都没有正确认识该技术的特征，对网络数据的类型也没有很好地认识，因此相关法律在界分责任时就出现了问题[3]。对网络爬虫行为进行定性时，需综合考虑行为人利用网络爬虫技术实施了何种行为，是否对刑法所保护的法益造成侵害，并结合行为人的主观心理状态予以判断是否符合犯罪的构成要件。司法实务中，“技术中立原则”的适用也应有边界。如果使用技术的人利用该技术手段实施犯罪行为，进而危害社会，则不因“技术中立原则”而免除刑事责任。从“深圳市谷米科技有限公司与武汉元光科技有限公司反不正当竞争案”[1]，到“全国首例‘爬虫’技术侵入计算机系统犯罪案”[2]，可以看出，在运用爬虫技术侵入系统、提取数据的过程中，轻则可能涉及民事违法，重则可能触及刑事犯罪。在使用网络爬虫技术的过程中，从技术本身的使用行为到抓取数据后的使用、传播行为，可能涉及非法侵入计算机信息系统罪等与计算机信息相关的罪名以及诈骗罪、侵犯著作权罪等。

此外，在适用具体的罪名时仍然需要按照刑法的犯罪构成标准判断行为人的网络爬虫行为是否构罪。根据罪刑法定原则，只有出现刑法所规定的具有侵害法益的情形下才构成犯罪。在适用刑法时首先必须对行为人的行为进行判断，判断该行为是否为危害行为，其次再判断其造成的危害结果对合法权益造成损害的程度，判断该种程度是否应当由刑法来处理，若无需动用刑法则可以通过其他部门法予以规制。通过这种判断可以合理区分网络爬虫行为善意或恶意、罪与非罪的界限，体现出刑法的谦抑性。我国刑法表明了社会危害性是犯罪的本质特征，因此网络爬虫行为只有在具有严重的社会危害性的情况下才能定罪。

2.2. 网络爬虫行为的入罪问题

近几年，我国制定了一系列有关网络空间的法律法规，用以加强对数据的保护和管制，为网络爬虫行为规制了严格的法律界限范围。例如，《网络安全法》侧重于加强网络运行的安全性，规定任何个人或组织不得进行非法侵入他人网络以及盗窃数据信息等行为。与此同时，《刑法》也根据恶意网络爬虫的具体情况，规定了非法侵入计算机系统罪、侵犯知识产权罪、破坏计算机信息系统罪等罪名。然则，

[1]广东省深圳市中级人民法院，深圳市谷米科技有限公司与武汉元光科技有限公司不正当竞争纠纷案，(2017)粤03民初822号。
[2]北京市海淀区人民法院，上海晟品网络科技有限公司等非法获取计算机信息系统数据案，(2017)京0108刑初2384号。

相关法律法规的制定诚然在提升数据维护和控制方面起到了作用，却也在如今社会的数据开放趋势下，限制了数据的流动性。

司法实践中，司法机关往往对网络爬虫行为所获取数据存在属性上的不同未进行详细探究，忽略了数据背后各自代表的不同法益，在裁判中局限于数据的物理属性，导致适用罪名时逐渐口袋化。究其根本，首先，我国现阶段的司法裁判依旧着重保护计算机信息安全，更加关注保护信息数据安全而不是集中在数据保护和管制上；其次，受制于实际操作中存在的一系列问题，在司法裁判中，面对有关网络爬虫技术的案件，源于数据本身的属性问题，裁判者们通常面临着取证困难、被抓取的数据背后隐藏的法益而导致证明标准过高等问题，为了节约人力和成本，疏于对具体受侵的法益进行认定，而直接适用“非法入侵计算机信息系统罪”之类的口袋罪；最后，源于刑法的谦抑性。法律由其制定起就是滞后的。就目前而言，相对于数字技术的迅速发展，社会智能化加快，刑法无法“紧随其后”，在恶意网络爬虫行为是否应当入罪的问题上，依旧无法统一观点，刑事立法明显处于滞后的状态。

2.3. 恶意网络爬虫行为的特征

2.3.1. 规模化

当网络爬虫技术被用于犯罪时，基于其能够高效简便地抓取大规模的数据和无需过多的操作的特征，其犯罪行为也随之呈现规模化特征。随着技术的快速发展，相较于传统的信息搜索，网络爬虫技术的发展使信息的搜索和获取显得更加高效。使用者在程序中写入算法指令就能够自动帮助开发者获取信息，无需人为干预和操作，这种技术能够在较短的时间内从网络中获取海量的信息。网络爬虫技术本是为了克服传统搜索引擎在面对海量化信息时无法实现高效的信息采集的缺陷，而作为传统搜索技术的升级。网络爬虫技术用于合法活动中时能够充分发挥其应有的功能，然而，当这种技术被用于犯罪活动中时，在技术的加持下将导致此类犯罪呈现规模化趋势。

网络爬虫作为一种信息采集技术，技术的使用特性造成了该类犯罪主体的转变，该转变使得犯罪主体不同于传统犯罪的主体，而犯罪主体的转变使得犯罪规模变大。具体情形表现为，在使用这种信息采集技术实施侵犯信息的犯罪行为中，该犯罪行为的实施依靠行为人对这项技术的掌握程度，因此，利用这种技术的犯罪行为排除了很多自然人单独犯罪的可能，取而代之的则是以互联网科技公司成为利用这种技术的犯罪的主体，相较于自然人单独实施的犯罪，互联网科技公司等单位主体凭借其地位优势和掌握的信息资源庞大，一旦实施犯罪行为，其规模必定庞大。

2.3.2. 社会危害性

社会危害性是犯罪的最本质特征，同时社会危害性必须达到一定程度才能构成犯罪，若某一行为不具备社会危害性，或者其社会危害性尚未达到一定程度，那么该种行为就不是犯罪行为。网络爬虫行为构成犯罪同样需要具备社会危害性这一特征，同时其社会危害性也必须达到一定程度。通常情况下，网络爬虫犯罪的社会危害性较大。宏观上网络爬虫技术造成危害对象的规模化，并不是只对公民信息的粗略收集而忽视对特定个体的精细挖掘。与此相反，当该种技术用于个人信息犯罪时，其犯罪特征表现为，宏观上可以对数量巨大的个体进行信息收集，微观上可以针对某个个体进行全方位的挖掘，使特定个体的信息暴露无遗，致使该类犯罪的社会危害性呈现较大的趋势。在过去网络不发达的时代，人人都处于信息孤岛的环境下，信息之间不关联，很难形成针对个人的信息犯罪，此时侵犯信息的犯罪活动危害性并不大。自 2002 年美国《国土安全法》提出通过捕捉信息网络中留下的“数据脚印”来锁定恐怖分子的“万维信息触角计划”，其工作原理从网上碎片化的信息中搜索需要追踪的某个人的所有信息，这项信息采集技术与数据挖掘和分析技术结合后，大到国家安全小到个人信息防护都成为了信息犯罪的对象。

3. 网络爬虫入罪的类型化分析

从网络爬虫的技术角度来看，网络爬虫行为一般包含三个阶段，即对计算机信息系统进行访问、对系统内数据进行爬取以及对所抓取的数据进行存储使用。有鉴于此，应当将网络爬虫的入罪进行类型化分析。

3.1. 非法侵入行为

非法侵入行为指的是未经授权侵入他人计算机系统的行为。这类网络爬虫行为通过非法手段获取系统权限，存在导致个人隐私泄露和计算机系统面临安全隐患的可能性。在法律上，这种侵入行为违反了计算机犯罪相关法律法规。针对这种情况，法律需要明确界定非法侵入的标准，并严厉打击入侵他人计算机系统的行为。

网络爬虫在对数据网站中的数据进行抓取之前，先决条件是有权利进入到该网站，或者得到进入该网站的授权。在没有得到授权的情况下访问有关国家事务、经济建设、政务管理以及涉及商业秘密等领域的计算机内部网站，存在构成非法侵入计算机信息系统罪的可能性。这是由进入该网站之前是否已经得到授权和被入侵的数据网站所包含的数据信息的性质所决定的。本罪的设立是出于对国家事务、经济建设、政务管理以及涉及商业秘密等领域的数据信息安全的特别保护。

3.2. 非法爬取数据行为

非法爬取数据行为是指网络爬虫以未经授权的方式，获取他人网站上的数据。虽然网络爬虫通常是为了搜索、数据分析等合法目的，但当其越过网站的 robots 协议或者网站主体的意愿，获取网站数据时，就可能触犯法律。这种行为侵犯了网站所有者的信息控制权，因此，法律上需要规范网络爬虫的数据获取行为，确保其合法性及道德性。

3.2.1. 爬取具有辨识性的公民个人信息资料，可能构成侵犯公民个人信息罪

在智慧社会中，大多数公民的个人信息以电子数据的形式储存在计算机信息系统或网络中，在便捷人们生活的同时，存在容易受到网络爬虫技术的侵扰的可能性。我国《刑法》第 253 条第 3 款规定了非法获取公民个人信息的行为将构成侵犯公民个人信息罪。2017 年 5 月 8 日，最高人民法院、最高人民检察院颁发的《关于办理侵犯公民个人信息刑事案件适用法律若干问题的解释》之中，明确了个人信息的定义。司法实践中，对于个人信息的认定范围较广，主要包括考生信息、违法犯罪记录、户籍信息等等。这些公民个人信息数据与一般数据的差异在于此类数据与信息主体具有一定的相关性和特异性，可以用于辨识特定的某一个公民，造成公民随时处于个人信息被侵害的危险之中。对于爬取具有辨识性的公民个人信息资料，可能构成侵犯公民个人信息罪。

3.2.2. 非法复制、传播知识产权作品涉嫌构成侵犯著作权罪

非法复制知识产权作品体现为行为人通过网络爬虫从互联网上获取文字和照片等数据信息，并在其服务器上整合这些具有独创性的数据。当用户在搜索引擎查找作品时，搜索引擎通过网络爬虫技术将作品转换为文字样式，同时将作品存储在服务器上，以提高用户的阅读速度。当用户以转码作为触发器访问时，网上的相关作品会被自动存储，随后，通过程序将作品内容转码为手机版或网页版，供用户阅读，并在其中设定广告以获取收益。

非法传播知识产权作品则表现为行为人在影视作品上设置链接，吸引用户点击观看。在网站中，他们会屏蔽原有广告并植入额外广告以牟取利润。行为人通常利用网络爬虫技术将著作权作品形成目录索引以供浏览以及影视作品等设置加框链接。一些学者认为，加框链接损害了著作权人的利益。由于著作

权人难以通过采取自助措施来消除加框链接产生的负面影响，因而著作权法需要对加框链接的行为进行干预。合理的选择为，直接禁止加框链接这一行为，设链者应对加框链接引发的作品传播行为负责。在修订著作权法时，可以采用“实质呈现”标准来改造信息网络传播权，使其包括加框链接引发的作品传播行为[4]。

3.3. 破坏计算机系统正常运行行为

破坏计算机系统正常运行行为是指通过恶意程序或者其他手段，破坏、篡改、删除他人计算机系统的数据，或者影响其正常运行。这种行为不仅损害了数据完整性，还可能导致系统崩溃，给个人、企业带来巨大损失。法律应当明确界定这种破坏行为的法律责任，采取相应的惩罚措施，以维护计算机系统的正常运行和数据安全。

网络爬虫技术遭到恶意使用时存在破坏计算机信息系统的正常运行的可能性。网络爬虫在对某一个系统进行访问和数据抓取时，每一个网络爬虫会使用单独的 IP 对网站进行访问。换言之，当行为人运用数量庞大的网络爬虫在同一时间访问某一个计算机系统或者数据网站时，这些网络爬虫对于网络资源的使用会达到一个惊人的数量，一旦超出系统负载的上限，就会产生计算机系统拥堵乃至崩溃的情况，使得该计算机系统或者数据网站无法正常工作。而在实际生活当中，此种手法经常被犯罪分子滥用。犯罪分子往往利用网络爬虫恶意访问网站，在入侵系统后对计算机保护措施进行暴力破解或者对数据信息进行破坏。如果网络爬虫侵入计算机信息系统后，对系统或数据进行破坏，或者采取暴力手段破解计算机信息系统的安全措施，甚至滥用爬虫技术进行网络攻击等，均可能构成破坏计算机信息系统罪。例如，在“王博一文、黄业兴破坏计算机信息系统一案”[3]中，被告人使用电脑编写了“爬虫”程序，该程序以植入第十三届全运会接待服务系统的方式对该系统进行攻击。他们删除了该系统内大量参赛运动员及技术官员的抵离信息、酒店住宿信息和人员简要身份信息，导致当日天津市全运会组委会接待服务部 39 台计算机无法正常运行接待服务系统。

4. 判断网络爬虫行为入罪的标准

我国《刑法》与网络空间犯罪相关的法律有三条，分别是第 285 条、第 286 条及第 287 条。制定这三条法律是为了维护网络安全以及实现数据的安全性。前者是以网络访问控制、防网络攻击等手段来维护网络边界和安全域、网络入侵防御、网络通信系统或传输安全、网络空间主权；后者则通过加密、减感等手段的数据保护，作为生产性和生产要素的数据的重要性，数据主权，个人隐私保护[5]。这里所称的数据安全不仅包括一般意义上的数据安全，同时包含着在网络上数据化的个人信息及其他涉及知产数据的安全。这一系列法律的制定都要求数据抓取也即网络爬虫行为，在网络中所抓取的数据、信息的界限皆符合法律法规以及行业标准。因此，在客观上，网络爬虫行为入罪主要判断其实施数据抓取行为时是否获得许可或者授权、是否侵害了各类数据安全。在主观上，行为人在实施上述不法行为时是否具有故意。由是观之，网络爬虫行为的入罪标准，在客观上必须具有侵害数据安全法益的不法行为，在主观上需要具有侵害数据安全法益的主观恶意。

4.1. 客观不法

为维护数据安全，数据网站可以通过技术手段对数据设置一定的保护障碍，阻碍他人侵入、获取、使用数据，从而在事实上实现对数据信息的排他性保护，法律也可以保护数据网站的权利，从而实现规范上的排他性保护[6]。为了保证数据的安全性，防止通过不法行为对数据进行抓取，通常会采取以下保

[3] 南开区人民法院，王博一文、黄业兴破坏计算机信息系统案，(2017)津 0104 刑初字第 740 号刑事判决书。

护措施：1) 合约授权，即通过意思表示允许或禁止他人访问、获取数据；2) 技术措施，即通过设置各种技术性手段来监控、防止数据抓取。虽然未经授权就抓取数据或者对技术措施进行突破来抓取数据都代表着对数据法益的侵害，但是上述两种保护措施对数据的保护意愿不同，如合约授权，有一些数据网站对于数据保护采取合约授权措施，这就代表了该网站对数据安全保护的“弱意愿”，此外，未经授权就抓取数据或者对技术措施进行突破来抓取数据这二者在法律上所需要承担的责任也不完全相当。

值得注意的是，根据我国《刑法》第 285 条、《关于办理危害计算机信息系统安全刑事案件应用法律若干问题的解释》第 2 条等规定，可以得出，在判断网络爬虫行为是否入罪时，应当判断该行为是否属于突破数据保护措施的行为。因而，未经许可、授权的网络爬虫行为还应当包括超越许可、授权的行为，从本质意义上而言，超越许可、授权的行为亦属于未经许可、授权，这类网络爬虫行为同样属于突破数据保护措施的行为，具有社会危害性，应当予以规制。

但是，目前网络爬虫技术得以大规模使用，如若仅仅从形式意义上判断网络爬虫行为是否符合入罪的构成要件，存在不当扩大该类行为犯罪圈的可能性。换言之，如果仅仅判断某一网络爬虫行为在形式上是否符合某类犯罪的构成要件，将有可能导致某些不具有社会危害性的行为入罪。例如，如果只是单纯的登入账号，尚未对法益造成威胁或者侵害，那么该类行为就不应该直接入罪。因而，在判断某一网络爬虫行为是否入罪时，应当进行实质判断其行为对法益造成的威胁、侵害是否已经达到可罚的程度。

4.2. 主观恶意

判断一个网络爬虫行为是否应当受到刑法规制，除了从客观上判断有无不法行为以外，还需要从主观上考察是否具有实施犯罪的故意。具体而言，应当从主观上判断该行为是否具有突破技术保护措施的故意。我们通常可以将网络爬虫行为分为善意的网络爬虫行为和恶意的网络爬虫行为。一般而言，善意的网络爬虫行为会遵守 Robots[4] 协议，这类行为不仅能够促进数据流通与资源共享，同时能够为网站提高报告度和流量，实现使用者和网站互利共赢的局面。与此相反，恶意的网络爬虫行为则无视 Robots 协议，采取一系列的技术突破保护措施，造成网站崩溃、数据泄露等危害后果，这些行为可以证明行为人主观上已经认识到其行为将造成一系列的危害后果，仍然继续实施，主观上具有犯罪故意[7]。[5] 比如，2018 年春运期间，12306 (中国铁路网)最高峰时段页面浏览量达 813.4 亿次，1 小时最高点击量 59.3 亿次，平均每秒 164.8 万次，其中，恶意爬虫访问占据了近 90%的流量，破坏了 12306 的正常运行，影响了普通群众正常使用 12306。再如，疫情期间，远程办公、协同办公等软件大量使用，该类软件往往承载着大量的企业数据信息，甚至包含着敏感信息，关系着企业的经营与生存。然而，2022 年，我国乃至全球的大型软件系统被披露存在多个漏洞，且被恶意攻击。其中，我国本土软件钉钉亦受到恶意攻击，行为人通过实施恶意的网络爬虫行为，找出漏洞，诱使受害者点击链接，行为人通过受害者攻击企业内网，抓取企业内部数据，造成企业损失。行为人在实施该类行为时，实施了一系列的诱导措施，进而抓取企业的数据，显然，行为人带有主观恶意，其行为具有严重的社会危害性，对计算机系统、个人或企业信息、社会秩序等等造成了严重影响。

经过上述分析，网络爬虫行为是否入罪不仅需要判断客观上是否为不法行为，同时需要判断其主观上是否具有故意。为此，对于“道德黑客”的行为是否应当纳入刑法予以规制值得商榷。司法实践认为，“道德黑客”是指挖掘计算机系统或者网络平台存在的安全漏洞并且将这些安全漏洞如数告诉官方，帮

[4]robots 协议也称爬虫协议、爬虫规则等，是指网站可建立一个 robots.txt 文件来告诉搜索引擎哪些页面可以抓取，哪些页面不能抓取，而搜索引擎则通过读取 robots.txt 文件来识别这个页面是否允许被抓取。

[5]例如 Baiduspider，其为百度搜索引擎的一个自动程序，它的作用是访问互联网上的网页，建立索引数据库，使用户能在百度搜索引擎中搜索到客户网站上的网页，为用户获取信息提供较大的便利。

助其维护系统安全，且不利用漏洞牟利或从事其他非法行为。“道德黑客”模拟试验数据的程序中，出于运算的需要，从计算机系统或网络平台中爬取或存储一些数据是没有办法避免的，尽可能多的对数据进行运算就越有可能发现安全漏洞。但是，在我国现有的法律框架下，“道德黑客”未经授权入侵网站的行为和对数据的爬取、存储行为，其合法性存在着争议，在实际的司法裁判中易被作为非法获取计算机信息系统数据罪的证据。“道德黑客”未经授权就入侵其他计算机系统及网络平台并对数据进行抓取和存储的行为，从客观而言，该类行为符合刑法中规定的犯罪形式。然而，从主观恶性上来判断“道德黑客”，是否应当对其进行处罚还有待商榷。

5. 结语

本文深入探讨了网络爬虫行为的性质，明确了其在法律体系中的地位。网络爬虫，作为信息获取的工具，本身并不具有恶意性质。然而，在现代社会，随着科技的进步，恶意网络爬虫行为也随之增多。

综合而言，网络爬虫行为的合法性和入罪标准的明晰性，直接影响到网络环境的健康发展和信息安全的保障。通过深入探讨网络爬虫行为的性质、入罪标准、恶意行为的探析以及类型化分析，能够更好地平衡网络爬虫行为的合法性和社会安全之间的关系，为法律体系的进一步完善提供有益的思考和指导。

参考文献

[1] 甘勇, 陶红伟. 大数据导论[M]. 北京: 中国铁道出版社, 2019: 15.

[2] 王玉林. 大数据应用的风控数据监管问题[C]//中国政法大学互联网金融法律研究院. 新时代大数据法治峰会——大数据、新增长点、新动能、新秩序论文集. 北京: 中国政法大学出版社, 2017: 115.

[3] 杨志琼. 数据时代网络爬虫的刑法规制[J]. 比较法研究, 2020(4): 185-200.

[4] 崔国斌. 加框链接的著作权法规制[J]. 政治与法律, 2014(5): 74-93.

[5] 郑云文. 数据安全——架构设计与实战[M]. 北京: 机械工业出版社, 2019.

[6] 纪海龙. 数据的私法定位与保护[J]. 法学研究, 2018, 40(6): 72-91.

[7] 刘艳红, 杨志琼. 网络爬虫的入罪标准与路径研究[J]. 人民检察, 2020(15): 26-31.