

从侵权到犯罪：网络爬虫技术刑事立法的边界厘定

魏涵靖

福建农林大学，福建 福州 350002

摘要：网络爬虫技术是大数据时代信息采集中应用较为广泛的一种爬虫技术，但是它的运用边界却引起很多法律争议。对网络爬虫的法律性质进行分析，并对合法应用和潜在侵权以及犯罪行为进行讨论。本文从隐私权保护，知识产权平衡和网络安全等角度出发，对爬虫技术在刑事立法中的规制路径进行研究，提出动态调整策略以满足技术发展需要。本研究旨在明晰网络爬虫技术由侵权向犯罪转化的法律界限，以期对相关刑事立法有所借鉴，从而找到技术发展和法律监管间的平衡点。

关键词：网络爬虫技术；刑事立法；隐私权；知识产权；网络安全

From infringement to crime: boundary determination of criminal legislation on web crawler technology

Wei Hanjing

Fujian Agriculture and Forestry University, Fuzhou, Fujian 350002

Abstract: Web crawler technology is a kind of crawler technology which is widely used in information collection in the era of large data, but its application boundary has caused many legal disputes. The legal nature of web crawlers is analyzed, and the legal application and potential infringement and criminal behavior are discussed. From the perspective of privacy protection, intellectual property balance and network security, this paper studies the regulatory path of crawler technology in criminal legislation, and puts forward dynamic adjustment strategies to meet the needs of technological development. The purpose of this study is to clarify the legal boundary of web crawler technology from infringement to crime, in order to provide reference for relevant criminal legislation, so as to find the balance point between technology development and legal supervision.

Keywords: web crawler technology; criminal legislation; privacy; intellectual property rights; network security

引言

在信息技术飞速发展的今天，网络爬虫技术已经成为了数据采集与分析中的一种重要手段。但它在实践中却频繁引起隐私权侵害，知识产权争议和网络安全威胁等问题。现有法律框架对于网络爬虫的技术规制较为落后，尤其是对于侵权和犯罪行为的区分界限模糊。刑事立法是规制技术运用的重要工具，面临着维护数据权益与推动技术创新的双重挑战。文章聚焦于网络爬虫技术，对网络爬虫的法律性质及潜在风险进行了梳理，提出了对网络爬虫进行刑事立法规制的途径及策略。

一、网络爬虫技术的法律性质

（一）网络爬虫技术的合法应用

网络爬虫技术是互联网领域中合法使用的有效信息获取工具，其应用范围之广、意义之大。其合法性表现为数据公开，技术应用正当。一是网络爬虫技术获取公开数据主要依靠目标网站向外发布信息资源。这类数据一般采用网页的方式显示，其访问权限向所有用户公开。因此，只要爬虫技术在使用过程中未突破技术限制（比如登录权限，验证码的验证等）或违反目标网站明确规定（如 robots.txt 协议），就可认定为符合合法使用的基本

要求。^[1]二是网络爬虫的合理使用关系到社会公共利益。比如搜索引擎公司通过爬虫技术采集网页数据来快速向用户检索信息。这一应用既符合大众获取信息的要求，又促进互联网生态良性发展。另外在科学研究领域中，爬虫技术通常被用来进行数据挖掘和分析以对社会经济和公共政策提供数据支持，这类应用通常被认为是正当的行为。但是应当指出，合法适用以技术操作满足某种行为边界为先决条件。比如爬虫技术开发者与使用者需要清楚地了解数据来源，访问目的与使用情况，以免给目标网站正常工作带来不利影响。法律实践对于爬虫技术正当性的判断更多地取决于技术目的和行为后果两方面的考虑。如果在获取数据过程中

没有损害到任何主体合法权益或者在技术操作上符合行业惯例和法律规范应视为合法使用。

（二）网络爬虫技术可能涉及的侵权行为

网络爬虫技术虽然有着广泛而正当的应用场景，但是它的操作方式与数据使用模式都会带来侵权问题，特别是涉及隐私权，知识产权与数据权利等方面，存在着潜在法律风险。一是隐私权问题是网络爬虫侵权纠纷中的一个主要方面。爬虫技术在打破技术屏障以获取私人数据或者通过大范围数据聚合来揭露用户个人信息时会对个人隐私权造成直接侵犯。^[2] 未经数据主体的同意，此类侵权行为尤其突出。二是知识产权问题又是爬虫技术适用的重要法律风险点。很多网络数据都是有版权属性的，如照片，文章和数据库。爬虫技术抓取这类数据如果不遵守合理利用的原则就会构成著作权侵权。比如把抓取到的东西直接拷贝出来公布于众，就会触犯《著作权法》有关条款。另外，目标网站数据库通常会受到《反不正当竞争法》或者数据库保护有关规定的限制，擅自进行大量抓取也会对数据库的使用形成侵权。再者爬虫技术也会影响目标网站正常运行。比如海量高频访问行为会造成服务器超负荷甚至触发服务中断等。在这一情况中，用户可能会被指控损害商业利益或者构成“不正当竞争”。

（三）网络爬虫技术可能涉及的犯罪行为

当网络爬虫技术由侵权行为上升到严重威胁公共利益或者特定主体权利时，就可能会构成犯罪。刑事责任的确定通常基于行为严重程度以及违法所得，涉及计算机犯罪，侵犯隐私以及非法获取资料等诸多领域。一是网络爬虫技术擅自打破了对目标网站访问的限制，有可能触犯《刑法》关于“非法入侵计算机信息系统罪”。适用该罪的关键是爬虫技术能否避免目标网站采取技术防护措施，如破解验证码，登录认证或者绕开访问限制加密协议等。^[3] 如果技术操作造成了系统瘫痪或者数据丢失等问题，犯罪人可能会受到更为严重的刑事责任追究。二是非法获取，销售或者提供资料的行为有可能违反“非法获取计算机信息系统数据的犯罪”。大数据时代数据这一重要经济资源的法律属性渐趋清晰。网络爬虫技术如果为了盈利而抓取和销售目标网站数据会面临着严重法律后果。比如未经许可在电商平台上大范围抓取用户交易信息用于商业目的就可能会被认定非法使用计算机信息系统。另外网络爬虫技术涉及隐私数据的犯罪性质更清晰。如利用技术手段捕捉用户个人敏感信息进行诈骗，敲诈勒索等犯罪活动不仅会侵害个体隐私而且会构成犯罪。该行为对公共安全和社会秩序构成了严重的威胁，也是法律严厉规制的焦点问题。从刑事立法层面上看，为了有效地防止爬虫技术被误用，应当对相关条款进行进一步的完善。^[4] 如厘清技术行为和犯罪行为之间的边界，在考虑技术发展正当需要的前提下细化非法数据获取与利用的惩罚条款等。

二、从侵权到犯罪：网络爬虫技术刑事立法的边界厘定

（一）网络爬虫技术使用规范的法律界限

以网络爬虫技术使用标准为中心进行合法性判断，但是当前

还存在着许多立法层面上的争议和模糊地带。界定它们的法律边界，需要从技术属性，行为后果，当前法律适用等层面加以讨论。就其技术属性而言，网络爬虫在使用过程中应该遵循公开性原则即只限于抓取互联网中公开可访数据而不能突破对目标站点访问控制机制和技术防护措施。比如一些爬虫绕过登录验证来抓取数据已明显违背法律保护信息安全原则。^[5] 依据《反不正当竞争法》第十二条之规定，擅自取得商业数据可视为不正当竞争，而对高钻爬虫技术进行规制则提供了一个初步法律依据。在行为后果层面上，爬虫技术使用频率，使用对象及使用方式决定着行为合法性。其中最典型的例子就是在2020年，一家搜索引擎使用爬虫技术从竞品网站上抓取数据，并被控给目标网站服务器带来负载压力而影响其正常工作。法院终审认定该行为方系违反“合理使用”技术规范，判其停止抓取行为，赔偿损失。这说明爬虫技术用户需要保证自己的操作不过度侵占目标网站技术资源。另外，现有法律框架在网络爬虫技术应用中也有一定的局限性。例如，《网络安全法》的第41条对于个人信息的合法获取有明确的规定，但是在企业信息、非敏感数据等其他类型的数据采集方面，缺乏明确的边界划定。未来立法应进一步细化爬虫技术行为规范，例如通过技术协议（如 robots.txt）的法律效力认定，构建更加清晰的技术使用规则。

（二）数据采集中的隐私权保护与法律责任

隐私权保护是我国网络爬虫技术相关立法的关键问题。爬虫技术采集数据时可能侵犯个人隐私的事件屡有发生，其核心纠纷是数据采集范围界定和法律责任承担问题。《个人信息保护法》第四条规定：收集个人信息需要征得数据主体的同意。但很多爬虫技术擅自采集用户数据直接侵犯隐私权。比如在2018年，一家电商平台就发现竞争对手利用爬虫技术大范围抓取用户消费记录的行为，随后判定其侵犯隐私权，判赔偿上千万元。^[6] 隐私数据具有敏感性，这决定着隐私数据法律责任之严厉。对于身份信息和财务信息等非授权抓取敏感数据需要行为方承担较高风险。这在欧洲《通用数据保护条例》中表现得尤为突出，GDPR 明确规定未经授权的数据采集行为可能面临高额罚款，甚至刑事责任。我国《个人信息保护法》也有类似条款，如敏感数据获取需要有明确合法性依据等。在实际运行中如何在数据利用和隐私保护之间进行权衡是个难题。比如一些公共服务中需要借助爬虫技术采集数据来完善系统的功能，而这种数据采集对隐私的侵犯与否需要从目的，范围和结果等方面来综合评判。故立法应当在对隐私权予以明确保护的前提下，对合法收集的资料进行适当扩大，对用于具体公益目的资料给予法律豁免。

（三）爬虫技术与知识产权保护的平衡策略

爬虫技术抓取网页内容一般都会涉及《著作权法》所保护的内容，如文章，图片和视频。它的行为是否合法，要看数据的属性，爬取方式和后续使用情况。比如，一个爬虫程序在一个新闻网站上抓取和发布整篇文章，法院就判定其著作权受到侵害。法院认为，这种行为是对网站商业资源的直接拷贝，是以广告牟利，是对权利人经济利益的侵害。在法律上，保护知识产权的界限需要在“合理使用”的原则中加以规定。^[7] 依据《著作权法》

的第24条规定，作品的合理使用并不被视为侵权行为，但必须确保其为非营利目的，并不妨碍权利人的正常使用。比如学术机构用爬虫技术抓数据进行研究而不公开发表的数据一般认为是正当的。但在爬虫技术应用于商业目的特别是对竞争对手核心内容进行大范围抓取的情况下就会形成不正当竞争。比如在2019年，一家电商平台在抓取竞争对手商品信息之后，将其应用于价格比对、商业推广等领域，法院认定其触犯《反不正当竞争法》。为了兼顾技术创新和知识产权保护，应当从立法上厘清爬虫技术行为合规界限。一方面，可通过技术协议（如 robots.txt）的法律效力确认，赋予网站运营者更多的主动保护权；另一方面要加大对侵权行为的跟踪和处罚力度，如利用区块链技术对数据利用路径进行记录。

（四）网络安全视角下爬虫技术的刑事责任

网络爬虫技术会给网络安全领域造成严重法律风险，刑事责任集中体现为非法访问，数据窃取及系统破坏。在爬虫技术变得越来越复杂的今天，爬虫可能带来的危害并不限于特定网站上数据的泄露，甚至会对公共网络系统造成安全威胁。典型的例子就是某个爬虫工具未经授权就对金融机构敏感数据进行了海量抓取并最终造成了系统崩溃。法院以“非法入侵计算机信息系统罪”作为定罪处罚的依据，对犯罪人追究刑事责任。^[9]爬虫技术给网络安全带来的威胁并不仅仅是数据抓取自身，还有高频访问行为会导致系统瘫痪。以2017年某款搜索引擎爬虫技术为例，由于抓取超高频率而造成目标服务器宕机事件，这直接影响到上万名用户正常工作。这类行为由于给网络基础设施带来实质性的破坏，法律定为“破坏计算机信息系统罪等”。另外，非法使用爬虫技术也会造成更大规模网络安全事件。比如通过爬虫技术捕获企业用户登录信息，并且使用这些信息实施网络攻击，这不仅会侵害用户隐私而且会造成大规模经济损失。鉴于上述现状，有必要对《刑法》的有关规定加以细化，厘清非法访问与破坏系统这一特定行为的特点以及刑事处罚标准。

（五）爬虫技术发展与刑事立法的动态调整

伴随着网络爬虫的发展，网络爬虫的应用场景与技术复杂性也在不断增加，刑事立法也需要紧跟时代步伐，迎接科技带来的新的法律挑战。^[9]传统法律体系普遍落后于科技发展，这种情况在爬虫技术治理上体现得尤为突出。比如在传统法律框架下对于大规模数据抓取问题的监管比较模糊，特别是对于自动化爬虫以及深度学习驱动下的高级爬虫技术而言，这些技术可能存在隐私侵害以及系统威胁等问题，因此很难清晰地划分刑事责任。当前，爬虫技术在国际刑事立法中已进行了一些探究。例如，美国的《计算机欺诈与滥用法》已经进行了多次修改，把未经许可的访问和抓取行为纳入了法律的管辖范围。同样地，欧洲的《通用数据保护条例》通过对数据收集和使用的严格规定，对爬虫技术的合法性提出了更高的标准。在我国，《刑法修正案（九）》对计算机犯罪范围作了进一步拓展，主要表现为非法取得计算机信息系统数据、非法入侵计算机信息系统等，这为爬虫技术承担刑事责任奠定了法律依据。^[10]刑事立法的动态调整，需要在权利保护和技术发展中寻求一种平衡。一方面要加强非法行为打击，如明确大范围抓取刑事标准、细化技术规避和合法操作定义等；另一方面要对技术创新给予法律支持，如通过行业标准制定、试验性技术沙盒等方式为爬虫的正当使用营造更加合规的环境。另外，政府与行业应该联合推进技术监测平台建设，实时跟踪与评估爬虫技术在法律边界上面临的挑战并为立法调整提供科学依据。

三、结束语

网络爬虫技术法律边界关涉到错综复杂的权利保护和技术发展难题。合理明晰侵权与犯罪之间的边界既是完善刑事立法的着力点，更是维护数据权益、推动技术创新的关键所在。今后，我们要以动态调整立法策略来迎接技术进步所带来的全新挑战，从而为我国网络爬虫的正当应用与良性发展奠定坚实法律基础。

参考文献

- [1] 李尚林、陈宫、雷勇. 基于 Java 的网络爬虫系统研究与设计 [J]. 新型工业化, 2021, 11(04).
- [2] 郑越. 以宪法为基础构建自媒体背景下隐私权的保护规则 [D]. 云南大学, 2023.
- [3] 解志勇, 李其蔓. 完善网络爬虫行政规制路径提高数字经济治理水平 [J]. 审计观察, 2022, (03): 4-11.
- [4] 何毅平, 黄媛, 湛茂溪, 等. 基于网络爬虫的招聘信息可视化系统设计与实现 [J]. 长江工程职业技术学院学报, 2023, 40(03): 24-28.
- [5] 姚万勤, 宋俊豪. 在技术与法律之间: 网络爬虫刑法规制的边界 [J]. 法治论坛, 2023, (01): 3-17.
- [6] 刘晓魁. 网络爬虫技术与策略分析 [J]. 网络安全技术与应用, 2022, (05): 17-19.
- [7] 郭鹏飞. 利用网络爬虫获取个人信息的刑法规制 [J]. 人民司法, 2022, (10): 21-27.
- [8] 杨瑞虹. 网络爬虫的刑法规制研究 [D]. 贵州师范大学, 2023.
- [9] 陈嵌崧. 网络数据爬取行为刑事规制困境与突破 [D]. 外交学院, 2023.
- [10] 于梓杉. 网络数据爬取行为的刑法规制 [D]. 山东大学, 2022.