

大数据时代 网络新闻聚合平台的版权保护研究^{*}

□文 | 彭 辉

[摘 要] 大数据时代,网络新闻聚合平台的版权保护面临良好发展机遇,同时也面临更为严峻的挑战。针对智能推送行为版权困境,一方面应与传统媒体达成版权合作,支付转让费用,实现内容授权使用;另一方面引进默示许可制度,以应对网络海量授权。针对智能写作行为版权困境,可在一定前提基础上,将人工智能出版物纳入《中华人民共和国著作权法》保护范畴之内;针对智能社交行为版权困境,可建立健全基于区块链的数字版权治理体系。

[关键词] 大数据 网络新闻聚合平台 版权保护

在大数据时代,随着智能产品的全面普及,以数据技术和网络技术为主要支撑的网络新闻聚合平台获得了迅猛发展,其独特优势在于作品传播途径方式发生了巨大的变迁,将分散于网络中的内容集中于一个平台,对于用户而言,其可以一次性阅读、使用或者下载由数字出版者采集的他人作品或自己创作的作品,而不需辗转登录多个网站;而对网络平台服务提供商而言,这种基于优化用户体验的一站式内容供给服务可以增加用户黏性,从而获得丰厚的经济利益。然而这种聚合内容的商业模式却时常引发版权纠纷和争议。^[1] 本文梳理网络新闻聚合平台运作规律与版权保护的内在关联,聚焦有关智能推送行为、智能写作行为、智能社交行为所引发的版权困境与瓶颈,提出有针对性的化解对策,为新闻聚合平台的健康发展提供参鉴。

成信息仓库,按照不同特征打上标签,并形成最终的信息分类,最终实现个性化产品推送。从版权角度来看,这是“版权灵感→版权创造→版权开发→版权管理→版权消费”的过程,也是版权价值不断得以实现和增值的过程。版权保护在整个网络新闻聚合平台价值传播链中担当引领、保障、聚合角色,并且根据平台在各个阶段运作机理特点,采取版权保护的侧重点和落足点也各有千秋,独具特色(见图1)。

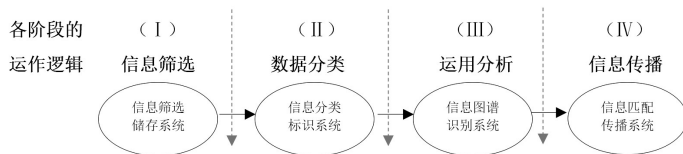


图1 大数据时代网络新闻聚合平台各阶段版权保护重点

一、网络新闻聚合平台运作逻辑

笔者研究认为,在网络新闻聚合平台的数字出版传播过程中,各阶段版权的核心内容分别围绕着阶段Ⅰ(信息筛选),阶段Ⅱ(数据分类),阶段Ⅲ(运用分析)及阶段Ⅳ(信息传播)展开。通过“网络爬虫”的智能算法机器人,在浩如烟海的新闻资讯网站中提取超链接,并储存在自身客户端中,形

由于网络新闻聚合平台以大数据形式存在的各种网络基础信息包含了每一个网络使用者浏览、使用互联网的全部记录,包括阅读的时长、转发、收藏频率,以及用户在微信、微博等社交账号上的各类资料(年龄、地理位置、性别、关注人群、好友评论),甚至是手机型号,均可作为网络新闻聚合平台用户兴趣标签的依据。通过进一步分析用户的阅读习惯,可以达到丰富用户兴趣图谱、建立用户

^{*} 本文系国家重点研究计划“全流程管控的精细化执行技术及装备研究”(2018YFC0830400)、国家社会科学基金重大项目“大数据与审判体系和审判能力现代化研究”(17ZDA130)、国家社科基金专项课题“深化司法体制综合配套改革研究”(18VSJ078)的阶段性研究成果

画像的目的,帮助经营者准确了解阅读者的最新需求,充分挖掘作品的市场价值(见图2)。

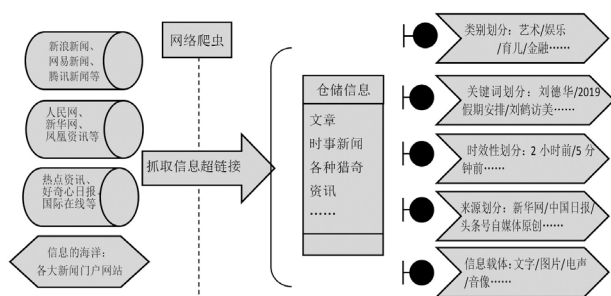


图2 网络新闻聚合平台运作机理

二、网络新闻聚合平台版权瓶颈及化解

大数据时代,网络新闻聚合平台持续加深用户对其依赖性,触动了著作权人与传统内容提供商的利益,进而引发各类纠纷争议,具体表现为以下三类。

1. 智能推送行为版权困境及应对

与传统搜索引擎相比,大数据时代的数据服务提供商不再仅仅局限于被动接受用户指令,而是在抓取大量用户行为数据的基础上主动通过“链接推送”等方式满足用户的个性化需求,传播的方式也逐渐从过去全面铺开的“广撒网式”向突出个性的“点对点式”转变,从而达到“量身定制”的效果。一般而言,传统搜索引擎的运作模式只是一种媒介服务,往往不涉及版权问题。但“链接推送”则有所不同,如果将个性化的推送行为认定为类似“作品创造”的加工行为,就需要对所推送的内容和对象进行版权甄别。当推送行为所使用的内容涉及未经授权的原创性作品时,就有可能涉嫌版权侵权或不正当竞争。

笔者研究认为,版权制度的基准在于鼓励新闻内容的生产者创造出更多有价值、高质量的原创作品。否则,如果新闻成果经常被数字出版商随意抓取和推送,势必会打击创作者的热情,到时便会出现空有大数据手段却无有价值内容的窘境。基于此,应对之策可从两个维度展开,一是短期之策:与传统媒体达成版权合作,支付转让费用,实现内容授权使用。在全网维权、维权赔付、版权登记、侵权投诉等方面强化版权保护措施,以平衡双方利益,促进作品的传播,进而达到繁荣文化的最终目

的。据新华网报道,今日头条已与1万家媒体达成版权合作。^[2]二是长期之策:引进默示许可制度以应对网络海量授权。版权默示许可制度是指作品使用人虽然没有得到版权人的明示授权,但是通过版权人行为可以推定版权人不排斥他人对其作品进行利用,作为一种补偿,使用人应当向版权人支付报酬的一种版权许可使用方式。^[3]大数据时代,网络聚合要涉及对各类媒体的文章的转载,如果都要经过被转载者的授权许可,那互联网的传播效力就会大打折扣。因此,应当在修订《中华人民共和国著作权法》(以下简称《著作权法》)时引入默示许可制度。

2. 智能写作行为版权困境及应对

在大数据时代,平台通过网络爬虫技术,全自动全天候采集信息的程序或者脚本,通过不同维度自动抽取新闻线索内容特征的计算规则,再通过事件发展曲线图进行多渠道验证,过滤掉无效内容和判断消息的真伪性,进而快速准确地对数字出版内容进行加工和制作,减少了出版从业者前往现场获取线索需要消耗的时间,提升了出版内容生产的时效性。目前,国内外网络新闻聚合平台纷纷推出各自“写作机器人”。

表1 国内外主要新闻写作机器人一览

国内					国外				
名称	所属机构	上线时间	领域	功能	名称	所属机构	上线时间	领域	功能
梦幻写手	腾讯	2015.9	财经	写稿	地震机器人	洛杉矶时报	2014.3	地震预报	写稿
快笔小新	新华社	2015.11	财经/体育	写稿	词语大师	美联社	2014.7	财经/体育	写稿
DT稿王	第一财经	2016.5	财经	写稿	赏花机器人	纽约时报	2015.5	新媒体	编辑
张小明	今日头条	2016.8	体育	写稿	回光仪	华盛顿邮报	2016.8	体育	写稿

人工智能技术给网络新闻聚合平台带来生产传播便利的同时,也对现有版权制度体系形成挑战。例如,人工智能出版物是否属于作品?产权和责任主体如何划分和厘清?《著作权法》如何对其实施有效规制和保护?等等。对此,学术界大致可以分为“不属于作品不予保护论”和“未来立法保护论”两大派。持前者观点的代表学者王迁教授认为,人工智能出版物是应用算法、规则和模板的结果,不

符合独创性要求,不能成为作品。^[4]持后者观点的代表学者熊琦教授认为,在没有明确标明来源的前提下,人类创作的作品与人工智能出版物的区分度难以准确分辨。既然无法从表面上区分两者之间的差异,这就意味着人工智能出版物可以被认定为作品。^[5]

笔者认为,目前人工智能出版物不能认定为作品,不能享受版权保护。理由有二:一是人工智能出版物不属于人的智能成果。目前人工智能出版物是根据特定格式,或者写作模板来传递信息,表达意思,如作为跨平台的人工智能机器人微软六代小冰,撰写诗歌也是按照“意象抽取→灵感激发→文学风格模型构思→试写第一句→第一句迭代一百次→完成全篇→文字质量自评→尝试不同篇幅→完成”的固定程式展开,尽管所做诗歌已经达到了与人类创作难以区分的水平,但其本质上还是通过算法机械式拼凑、组合而成,不能反映创作者的情感和思想,不属于思想或者情感的表达。二是人工智能出版物缺乏独创性。符合版权作品的前提条件是独创性,要求具有鲜明的个性和思想高度,人工智能出版物,即便是按照“深度神经网络”规则生成内容的新颖性明显,但依旧是依据特定预设算法、规则和模板进行编排处理的结果,其差异性并不显著。例如,体育赛事写作机器人“张小明”拟人化程度更高,但其最大的瓶颈问题在于缺乏对事实和数据的概括提炼能力,整篇文稿的用词造句甚至结构都几乎相同,雷同性高,缺乏个性化色彩,使得读者无法获得高质量阅读体验与“沉浸式”阅读感受。

由于人工智能技术处于不断迭代优化探索过程中,笔者认为,为了促进该产业健康持续发展,可在一定前提基础之上,将人工智能出版物纳入《著作权法》保护范畴之内。前提条件就是写作机器人能够自行判断、收集和学习新的数据,摆脱既定的算法规则设定,脱离设计者在数据和算法规则上的参与,从而能够解决新问题,生产出新的内容。当且仅当满足上述条件时,我们认为未来版权立法可将人工智能出版物视为作品。

3. 智能社交行为版权困境及应对

网络新闻聚合平台往社交方向发展能带来更多想象空间,其重要策略是平台更多层面更宽领域赋

能创作者,体现了一个平台的价值倾向和价值选择,这个过程的核心要义在于强化版权保护。

在网络技术极其发达的大数据时代,数字版权保护主要面临“四难”:一是打击盗版难。数字内容传播具有复制成本低、盗版传播快等特点,同时侵权行为还具有天然的隐蔽性和难追查性,这使得数字内容侵权越来越容易发生,加大了惩治难度。二是版权确权难。传统版权登记因其中心化管理、信息透明度较差、历史不可追溯的体制机制,造成版权登记成本高、耗时长,这显然无法匹配网络新闻聚合平台作品“产量多、传播快”诉求,给确权、维权、交易都带来了不同程度的掣肘。三是版权交易难。传统的版权交易一般遵循“买方阅读作品→寻找作品出品方或版权所有方→提出购买意向→双方细致磋商→签订版权交易协议”的线下交易流程运作,整个交易程序复杂、耗时耗力,且整个过程无标准、透明度低,导致版权交易市场乱象。四是版权维权难。在网络上获得非法复制的数字化版权作品并投入使用的侵权者,往往都是分散的个人用户,版权人在维权时遭遇的首要问题是侵权者无法准确锁定抑或规模化,或是维权成本与收益差距过于悬殊等难题。

笔者认为,化解上述瓶颈问题的关键在于建立健全基于区块链的数字版权治理体系。区块链具有的去中心化、难以篡改、扩展性大、可追溯特性、灵活性强等技术特征非常适应数字内容版权保护的核心诉求,^[6]这些特性使得版权保护与版权交易成为区块链技术最具吸引力的应用场景之一。对于版权确权而言,通过区块链技术写入具有唯一性的作品摘要信息,从而无法伪造和篡改数字作品内容,达到数字登记和确权要求;对于创作者利益维护而言,基于区块链技术以及智能合约的约束机制,数字版权的交易与流转得以在安全可靠的环境中顺利完成,保护文化创新原创者权益;对于版权执法而言,执法机构直接采信这一具有公信力的区块链技术作为证据支撑,简化了维权手续,缩减了取证时间,降低了维权成本,有利于版权执法部门高效执法。

国内外越来越多的公司意识到区块链对于版权保护与交易裨益良多,纷纷将较为成熟的区块链技术引入版权保护与交易业务领域。

表2 国内外在数字版权领域运用区块链技术的代表性企业

企业	领域	国家	企业	领域	国家
巴比特	版权确权、保护	美国	阿德列尔	数字版权管理	俄罗斯
启示者	音乐版权交易、推广、版税结算	以色列	监护入	版权保护、交易	南非
天鸽座	数字资产管理、数字版权记录	以色列	正派	数字版权管理	瑞士
觅宝	版权保护	美国	根源	版权确权、保护	德国
音乐公约	音乐版权确权、交易	英国	存在证明	文件时间戳、证明所有权	阿根廷
灵感	版权交易、版权费用管理	丹麦	咬痕	知识产权、学历证书公证等	美国
图姆斯	音乐版权共享及分销	澳大利亚	视野	图片管理、图片版权保护	加拿大
声田	正版流媒体音乐服务平台	瑞典	趣链科技	版权保护	中国
到达科技	数字出版、众筹问题	中国	纸贵版权	数字版权确权、交易问题	中国

基于区块链技术的版权保护体系可以为原创者和权利人提供多元化的保障服务，其具体框架主要包括三个维度：①版权权利证明服务。使用区块链进行版权确权可以节省大量的人力物力及系统的运营成本，并且整个过程几乎是瞬时的。②“智能合约式”版权交易。一是版权服务商可以在作品创作的最初产生灵感到最终作品出炉的所有状态变化过程中的任一阶段随时参与进来，提升交易效率；二是集中处理传统版权交易的碎片化信息，原创者与版权服务商可以有更多选项的交易方式和成交策略；三是版权交易过程中，可生成真实的交易原始数据，从而方便版权准确定价，并在实现版权交易的同时，也完成版税结算服务流程。③侵权检测服务。目前原本、版权家、麦片网等国内版权登记平台，都承诺构建可探测到全网络每一个角落的监测系统，让每一篇认证作品实时反馈它的全网传播路径，提供侵权行为预警、侵权存证服务、场景化维

权服务、版权诉讼信息及案例查询参考等服务，从而使得维权趋于标准化、工具化、大幅降低维权门槛，激发原创者和权利人的维权热情，从而持续净化和优化版权生态，促进国家创新型经济发展转型。

《最高人民法院关于互联网审理案件若干问题的规定》第11条规定明确指出：“当事人提交的电子数据，通过电子签名、可信时间戳、哈希值校验、区块链等证据收集、固定和防篡改的技术手段或者通过电子取证存证平台认证，能够证明其真实性的，互联网法院应当确认。”这是最高人民法院首次以法律解释的行为对以区块链技术进行存证的电子数据真实性做出认可，^[7]据此，区块链存证法律效力取得了实质性进展和突破。

三、结语与展望

在网络技术的强势冲击下，传统印刷时代版权人、出版商及社会公众之间形成的相对稳定的利益格局被打破。基于版权制度的效率和公正的价值目标，一方面信息技术、网络新闻聚合平台与版权保护在大数据时代维持适度均衡，另一方面原创者、网络出版商、社会公众三者主体利益之间也应当保持动态平衡。网络新闻聚合平台是我国出版业在大数据时代的主要发展方向，当下，平台市场活跃度逐渐显现，侵权行为频繁发生。数字版权治理体系构建的基准在于快速版权确权、安全有效交易、实现作品价值、原创者权益保障、维护版权市场秩序、创新型版权经济转型等，这将是大数据时代网络新闻聚合平台的版权保护研究进一步展开的逻辑起点和主要方向。

（作者单位：上海社会科学院法学研究所）

参考文献：

- [1] 卢海君，张雨潇．试论大数据时代的版权保护[J]．中国出版，2015（3）
- [2] 易潇．华西都市报与今日头条实现全面合作[N]．新华网，2017-11-20
- [3] 冯晓青，邓永泽．数字网络环境下著作权默示许可制度研究[J]．南都学坛，2014（5）
- [4] 王迁．论人工智能生成的内容在著作权法中的定性[J]．法律科学，2017（5）
- [5] 熊琦．人工智能生成内容的著作权认定[J]．知识产权，2017（3）
- [6] Kiyomoto S, Rahman M S, Basu A. On Blockchain-Based Anonymized Dataset Distribution Platform[C]. IEEE, International Conference on Software Engineering Research, Management and Applications. IEEE, 2017: 85-92
- [7] 宋戈．最高法院出台司法解释 认可区块链固定证据的“真实性”[N]．每日经济新闻，2018-09-10