

情报科学

Information Science

ISSN 1007-7634, CN 22-1264/G2

## 《情报科学》网络首发论文

题目：数字图书馆著作权侵权防范的制度、理论与实践——基于学术与政策文本的关联分析

作者：崔鸿鸣

DOI：10.13833/j.issn.1007-7634.2024.03.015

收稿日期：2023-10-28

网络首发日期：2024-03-15

引用格式：崔鸿鸣. 数字图书馆著作权侵权防范的制度、理论与实践——基于学术与政策文本的关联分析[J/OL]. 情报科学.  
<https://doi.org/10.13833/j.issn.1007-7634.2024.03.015>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 数字图书馆著作权侵权防范的制度、理论与实践

## ——基于学术与政策文本的关联分析

崔鸿鸣

(大连海事大学 法学院, 辽宁 大连 116026)

**摘要:**【目的/意义】通过对数字图书馆著作权侵权防范的制度、理论与司法实践中的文本进行分析,明确防范著作权侵权的核心争议问题与完善路径。【方法/过程】以北大法律信息网、中国知网为数据来源,选取法律法规政策文件、权威学术论文与典型案例裁判文书作为研究样本,综合运用关键词共现、领域学术观点库、主题演化分析、条文比对分析与词嵌入分析等方法进行了关联分析。【结果/结论】通过关联分析和检验,发现有关主体均在不同程度上对数字作品的合理使用存在认识偏差,人工智能为构建具有中国特色的著作权制度体系带来了重要契机,著作权集体管理制度功能的发挥成为著作权侵权防范的关键。【创新/局限】借助领域学术观点库构建方法辅助文本分析,对文本分析方法的单独使用具有提升作用,对完善数字图书馆著作权侵权防范路径的研究具有创新价值。由于关联研究是一个复杂的任务,其结果会受到数据质量、预处理方式以及所选方法的影响。领域学术观点库构建尚须进一步的合理性与可行性验证,可能带来微小的不确定性。

**关键词:**数字图书馆;侵权防范化解;学术观点;政策文本;关联分析

**中图分类号:**G250.2 **DOI:**10.13833/j.issn.1007-7634.2024.03.015

## 0 引言

2023年,中共中央、国务院印发《数字中国建设整体布局规划》(以下简称《规划》)指出,建设数字中国是数字时代推进中国式现代化的重要引擎,是构筑国家竞争新优势的有力支撑<sup>[1]</sup>。作为数字中国的开路先锋,数字图书馆建设已然成为21世纪图书馆的发展方向,尤以数字化典藏与服务为核心内容<sup>[2]</sup>。近年来,数字图书馆著作权侵权案件层出不穷,国家相继出台了相关法规政策,学界亦陆续展开了系统研究。例如,林扬<sup>[3]</sup>、杨家富<sup>[4]</sup>较早地研究了数字图书馆建设与著作权保护问题,杨向明<sup>[5]</sup>较早地研究了数字化作品的知识产权问题,邱均平等<sup>[6]</sup>研究了版权保护及其规避行为的法律对策,徐恺英等人<sup>[7]</sup>专门对著作权保护中的利益关系与对策作出了关联研究,李华伟等人<sup>[8]</sup>以时间为序对数字图书馆知识产权与相关法律问题进行了比较研究。整体而言,以数字图书馆为著作权侵权主体的研究已经取得深入进展,而以其为连带责任主体的数字图书馆用户侵犯著作权问题仍未得到有效控制<sup>[9]</sup>,特别是在“互联网+”背景下,图书馆在数字作品提供过程中的著作权侵权问题仍未得到很好解决<sup>[10]</sup>,而现有研究尚缺乏对于数字图书馆著作权相关政策、理论与实践的系统研究,且缺乏对于学术研究与政策改善的关联分析。鉴于数字图书馆建设的法律困境在于如何实现

数字化利用与防范著作权侵权的平衡,而学术的价值正在于对实践中产生的社会规范加以描述、抽象和法理提炼,从而推动社会规范意义向法律系统传递<sup>[11]</sup>,本研究拟对学术研究主题的演进与政策改善展开关联分析,并通过司法裁判文书分析加以检验,从而在整体上明晰数字图书馆著作权侵权防范的制度结构与功能、理论基础与实践路径,为提出具有结构性张力的法律困境缓和方案提供知识增量。

## 1 研究方法 with 框架

### 1.1 领域学术观点库构建方法

学术观点是学术信息交流的重要内容,通过对其加以分析可以发现领域核心科学问题及其特征。根据信息交流理论,学者间基于论文和专著等形式的交流被称为正式信息交流,有利于克服观点片面性。在学术界,学者首先须对提出的学术观点进行论证,然后须经同行评价质疑并经过学术验证成为知识,并始终满足事实检验,实则在“学术观点——知识——真理”间构成了一条逻辑路线。鉴于学术大数据的发展引发了海量信息与知识相对匮乏的矛盾,有学者提出可通过“领域学术观点库”提高学术交流效率。就数字图书馆著作权侵权防范而言,问题的关键在于以结构化学术观点为研究对象,从逻辑命题理论出发,对不同细粒度的学术文本加

收稿日期:2023-10-28

基金项目:国家社科基金重大专项项目“社会主义核心价值观在民法典中的价值定位与规范配置研究”(18VHJ003)。

作者简介:崔鸿鸣(1992-),男,吉林吉林人,博士,博士后,主要从事财产法、信托法研究。

以结构化分析并挖掘核心观点。通常,学术文本组织可分为词汇、句子与篇章三个层面,分别集中于关键词抽取、创新贡献句识别与篇章对关键词识别效果的提升功能。学术观点的结构化表示可基于联项词、模态词、系词、否定词等的不同而分为不同类型<sup>[12]</sup>。

## 1.2 条文比对分析方法

政策文本比对(Text Compare)通常需要结合文本处理技术、比对算法、版本控制等方法实现法规变迁比对功能。该方法有助于跟踪政策变化历史,准确定位修改点,理解政策的修订意图,研究政策的适用范围,评估政策变动的影响,帮助研究者跟踪政策的变化历史,了解政策随时间的演变过程。这对于分析政策制定的背景、目的以及随着社会变革的发展而做出的相应调整具有重要意义。政策文本包括政府或国家或地区的各级权力或行政机关以文件形式颁布的法律法规等官方文献。政策文本分析方法先后出现了“专家解读—语言学实证研究—量化研究”三个主流研究范式<sup>[13]</sup>。目前,已有很多学者用政策文本分析方法进行分析,例如彭中礼<sup>[14]</sup>和姜国俊等人<sup>[15]</sup>运用共词分析法分别对1979—2016年的有效法律文本中的政策概念和中国环境问责制度的嬗变特征与演进逻辑进行了研究;冯莉<sup>[16]</sup>运用政策文本量化分析方法对流域生态环境规制思路与完善对策进行了分析。同时,闫慧等人<sup>[17]</sup>运用内容分析法对政策条文进行人工编码,从而对政策条文的主题进行了分类。值得注意的是,王锴<sup>[18]</sup>研究发现全国人大的法律与全国人大常委会的法律之间、授权立法与法律之间具有相同位阶。鉴于相同位阶的政策在规范配置语言方面具有相似性而具有比较价值与可行性,故本研究拟主要对法律法规层面的著作权侵权政策文本进行分析。

## 1.3 多元文本分析方法

本研究将采用包括关键词共现、基于狄利克雷分布(Latent Dirichlet Allocation, LDA)的主题演化分析与词嵌入分析(word2vec)在内的多元文本分析方法,相应地对学术研究主题划分及其强度随时间变化情况,以及政策文本上下文展开融贯性的分析。其中,一方面,法律法规文本通常需要根据上下文进行正确解释,运用 Word2Vec 方法,能够将每个单词映射到一个高维向量空间中并保留单词之间的语义关系,使得在向量空间中相似含义的单词会在空间中相互靠近,从而发现不同单词间的相对关系。另一方面,通过对学术研究主题的划分及其演化展开研究,具有识别关键主题、评估研究热点、揭示学科交叉与明晰不同时期的主题变化的独特价值。

## 1.4 裁判文书摘要方法

为克服数据质量、预处理方式以及所选方法的局限,应进一步结合典型裁判文书予以验证,从而检验学术观点与政策文本关联性分析的实际有效性。利用信息处理技术与论

辩理论抽取裁判文书的论证结构,一个关键的技术就是自动摘要。一篇结构完整的裁判文书以首部、判决主文、判决、尾部四大要素构成。学界关于法律裁判文书自动摘要的研究表明,鉴于裁判文书具有结构化程度低等特点,抽取式摘要方案往往存在语义缺失等局限,有学者由此提出了具有融合性的 BASR(Bert Based Attention Seq2Seq Reinforced Model)摘要自动生成模型。实验表明,特别是针对小额诉讼和简易程序,裁判文书中的判决依据和法律依据即使篇幅较短,也能很好地反映案件类型、争议类型等文本自动摘要时的关键信息<sup>[19]</sup>。现有裁判文书数据库能够稳定、准确和相对高效地自动生成案例检索报告,集成法院观点,从而自动完成裁判文书的摘要工作。

## 1.5 研究思路与架构

如图1所示,本研究基于领域学术观点库构建方法、条文比对分析方法、多元文本分析方法与裁判文书摘要方法,分别从制度、理论与司法实践三个维度展开分析论证,以学术研究主题划分及其演化分析,与法律法规变迁及上下文词语相似度分析为基础。其中,学术观点分析的目标在于发现数字图书馆著作权侵权领域的关键科学问题及其演化趋势,法律法规变迁分析的目标在于明晰政策修改或新增背后的目的,进而分别在主体、行为、权利和责任四个维度对二者进行关联分析,最终通过对裁判文书的结构化分析检验前述分析结果的有效性。其中,关联分析和检验的法理逻辑在于,主体通过行为实现权利,因权利受损而产生责任问题,责任的承担回补主体利益的减损。应当指出的是,以上制度、理论与司法实践三个维度各受到相应外部因素的影响,本研究将主要因素概括为制度层面的合法性危机、理论层面的公共与私人领域边界的重塑、实践层面的利益冲突平衡。

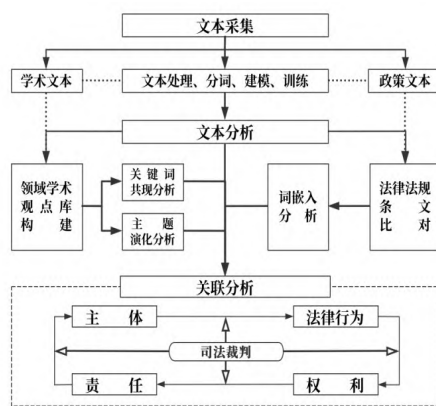


图1 研究框架

Figure 1 Research frameworks

## 2 文本采集与分析

### 2.1 文本采集

本研究关注数字图书馆著作权侵权防范的学术观点与

政策的关联性分析,因此文本的收集、筛选与整理须考虑二者的内在联系。鉴于《中华人民共和国著作权法》(以下简称著作权法)于1990年颁布,分别于2001年、2010年、2020年作出修订,《信息网络传播权保护条例》(以下简称保护条例)于2006年公布施行,《中华人民共和国著作权法实施条例》(以下简称著作权法实施条例)于2002年公布,分别于2011年、2013年作出修订,且1996年IFLA第62届大会后,数字图书馆在我国才引起普遍重视<sup>[20]</sup>,故本研究将以1996年至2023年为时间范围,并在进行相应的政策与学术文本采集时重点关注上述时间节点。

(1)在学术文本采集方面,本研究主要以CSSCI期刊论文作为采集对象,以1996年至2023年为时间范围。一方面,通过中国知网以数字图书馆著作权为主题,仅看有全文的中文学术期刊进行检索,共得到270条Refworks格式文献信息并导出。另一方面,通过中国知网以全文同时包含数字图书馆、著作权、侵权为条件进行检索,共得到2169条论文摘要信息,剔除其中明显不相关、摘要为空的文献后,共获得2130条文献摘要信息,并按发表时间升序排列。

(2)在政策文本采集方面,本研究主要以现行有效的法律法规政策文本为采集对象,以确保政策文本的权威性。通过运用北大法宝数据库的法律法规高级检索功能,分别搜索同一政策全文包含关键词“数字图书馆”与“侵权”、“数字图书馆”与“著作权”“侵犯著作权”的法律法规层面的政策,通过仔细阅读,纳入与主题相关的现行有效政策文件,排除复函、批复、政策解读等相关文件,共获得政策文件10份,按颁布时间升序排列,如表1所示。同时,根据数字图书馆著作权侵权防范命题所涉概念的结构特征,本研究按照主体、行为、权利与责任四个维度对《中华人民共和国民法典》(以下简称民法典)进行主要条文的选取。其中,主体维度关注数字图书馆建设过程中的主体关系,行为维度关注数字图书馆建设过程中主体的典型行为,权利维度关注主体各自权利的范围,责任维度关注主体在著作权侵权中的责任承担模式。

(3)在裁判文书采集方面,本研究运用把手案例(laws-data)裁判文书数据库的高级搜索功能,检索民事裁判文书

“全文”中同时包含“数字作品”与“图书馆”的高院案例22份,普通案例87份,剔除其中重复、明显无关、信息不完整的裁判文书后,共获取裁判文书36份。通过运用把手案例数据库的案例检索报告自动生成功能,获得了上述裁判文书包括序号、案件名称、案由、案号、法院、诉讼请求、公诉机关称、案情简介、本院认为、裁判结果与引用法规在内的数据文档,并按裁判文书案号升序排列。

## 2.2 文本分析

(1)就学术文本而言,针对以数字图书馆著作权为主题进行检索得到的270条文献信息,通过运用VOSviewer软件进行关键词筛选与共现分析,共得到795个关键词,剔除英文与中文关键词的重合部分与并无实际意义的关键词后,得到关键词共现结果(如图2)。初步看来,关于数字图书馆著作权侵权防范的研究,由原来的权利与利益的二元对立分析视角,逐渐演变出融入了合理使用、著作权例外等概念的二元缓和与分析视角。而且更重要的是,该视角的转换表现在从“公众利益”到“公共利益”概念的扩展使用。

针对著作权与公共利益、公共利益间的平衡问题,进一步运用领域学术观点库构建方法,筛选出为学界广泛认可的高被引文献全文,从词汇、句子和篇章三个层面进行学术文本组织研究。因文献数量较少,故人工梳理出核心学术观点句并进行结构化分析,明确其论证路径,并对该问题研究中学术观点的论证结构进行比较,从而进一步提炼出关键科学问题。由表2可知,数字图书馆著作权侵权领域的关键科学问题似应在于如何化解合理使用与侵权之争,其中在数字图书馆身份变化背景下,学界尚缺乏基于区分公众利益与公共利益的著作权侵权防范研究。

针对以全文同时包含数字图书馆、著作权、侵权为条件进行检索最终得到的2130条文献摘要信息,使用Python 3.8 gensim库进行主题演化分析。先对期刊论文摘要语料进行数据清洗,运用Python Jieba Posseg分词并进行词性标注,去除结果中长度小于2的词和非中文字符;综合baidu、hit、scu与常用中文停用词表等去除停用词,另人工添加论文摘要常

表1 数字图书馆著作权侵权领域法律法规

Table 1 Laws and regulations in the field of copyright infringement in digital libraries

序号	颁布时间(年)	法律法规名称	颁布单位	效力层级
1	1990	著作权法(1990)	全国人大常委会	法律
2	2001	著作权法(2001修正)	全国人大常委会	法律
3	2002	著作权法实施条例	国务院	行政法规
4	2006	信息网络传播权保护条例	国务院	行政法规
5	2010	著作权法(2010修正)	全国人大常委会	法律
6	2011	著作权法实施条例(2011修订)	国务院	行政法规
7	2013	著作权法实施条例(2013修订)	国务院	行政法规
8	2013	信息网络传播权保护条例(2013修订)	国务院	行政法规
9	2020	著作权法(2020修正)	全国人大常委会	法律
10	2020	民法典	全国人民代表大会	法律

见无意义停用词,例如“目的”“意义”等。通过计算主题困惑度(Perplexity),发现主题数量超过9个时困惑度下降较少,基于 $u\_mass$ 评估的一致性最高,从而确定最佳主题数量为9(如图3),主题识别结果如表3所示。

从时间脉络上来看,数字图书馆著作权侵权领域的研究主题强度基本上随时间而呈现波动性增加态势,(如图4)其中关于数字图书馆著作权保护的研究主题强度出现了明显

的减弱,与此相对出现明显增强的主题是人工智能与数据,关于作品合理使用的主题强度亦呈现出增加趋势,(如图5)表明学术界关于数字图书馆著作权的整体认识经历了从“强调著作保护”到“保障社会合理使用”的转变。

(2)就政策文本而言,运用北大法宝数据库的条文比对功能,发现相关修订或新增内容亮点(如表4)。此外,民法典第123与第1185条在私权意义上分别确认了民事主体对

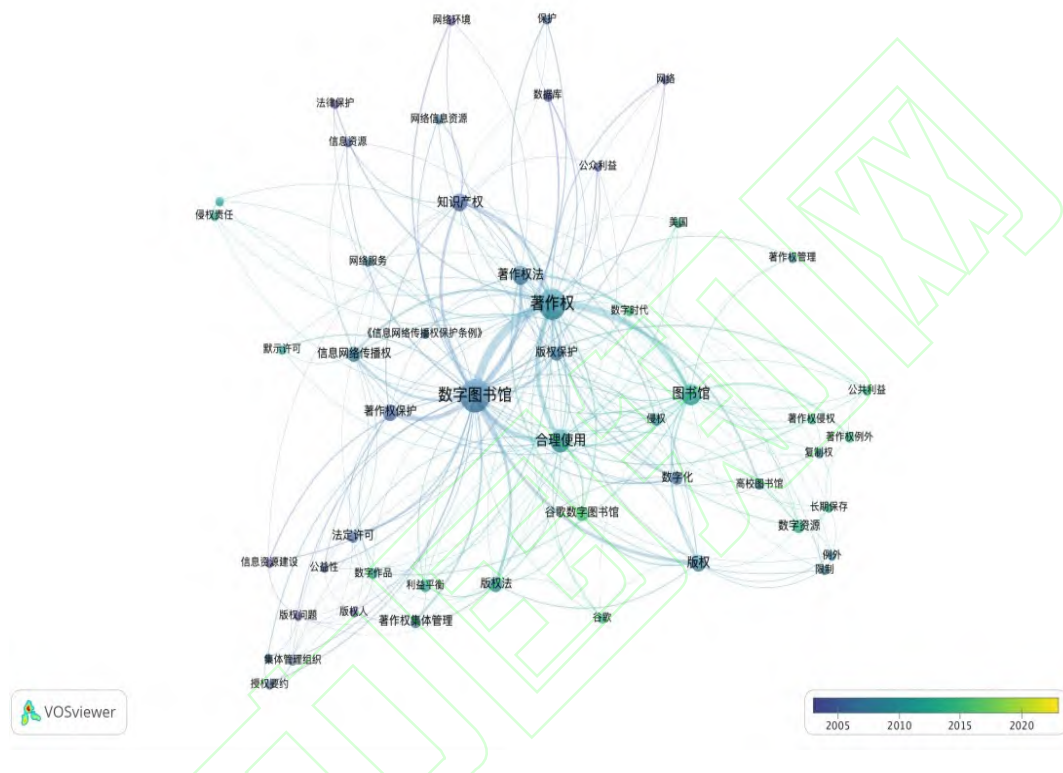


图2 数字图书馆著作权侵权主题研究论文关键词共现图

Figure 2 Keyword co-occurrence map of research papers on the topic of copyright infringement in digital libraries

表2 数字图书馆著作权侵权领域关键科学问题学术观点库(按时间升序排列)

Table 2 Pool of scholarly opinion on key scientific issues in the field of copyright infringement in digital libraries (in ascending chronological order)

序号	核心学术观点句结构化分析	论证路径
1	(1-1,定义型,联项词:是)作为公益性代表机构的图书馆确有维护著作权私权利益、保护社会公众利益、维护著作权人利益和社会公共利益的平衡的责任 <sup>[21]</sup>	著作权私权受保护源于著作权的私权属性==>国家介入因素增强导致著作权私权的公权化因素增强==>图书馆在私权公权化的趋势下负有平衡利益的责任
2	(3-1,定义型,联项词:表现为)知识产权权利冲突的复杂性不仅表现为个体权利与社会公众利益的冲突 <sup>[22]</sup>  (3-2,定义型,联项词:表现为)知识产权权利冲突的复杂性也表现为群体利益与社会公共利益的冲突。	知识产权冲突产生于利益冲突==>知识产权冲突产生的社会基础是权利保护客体逐渐增多、主体制度演进多元化==>应当以保护在先权利与权利共存并重作为知识产权冲突化解机制
3	(3-1,定义型,联项词:是)版权制度是作者私权与社会公共利益平衡的产物 <sup>[23]</sup>	图书电子扫描行为导致合理使用与侵权之争==>数字图书馆身份的变化是主要起因==>公共利益考量不应仅以使用目的有营利性而否定合理使用
4	(4-1,定义型,联项词:是)合理使用制度是著作权保护与公共利益保护的平衡器 <sup>[24]</sup>	商业模式和技术新发展带来著作权利益分配新问题==>转换性使用规则的运用受到限制==>应着力于合理使用规则本身的完善
5	(5-1,定义型,联项词:是)“公共利益”与“其他公民宪法性权利”是我国合理使用制度的宪法性基础 <sup>[25]</sup>	权利限制理论有局限性==>以公共利益作为合理使用的宪法性基础更契合财产权的社会义务==>以其他公民宪法性权利作为补充有助于在整体上涵摄著作权合理使用的种类

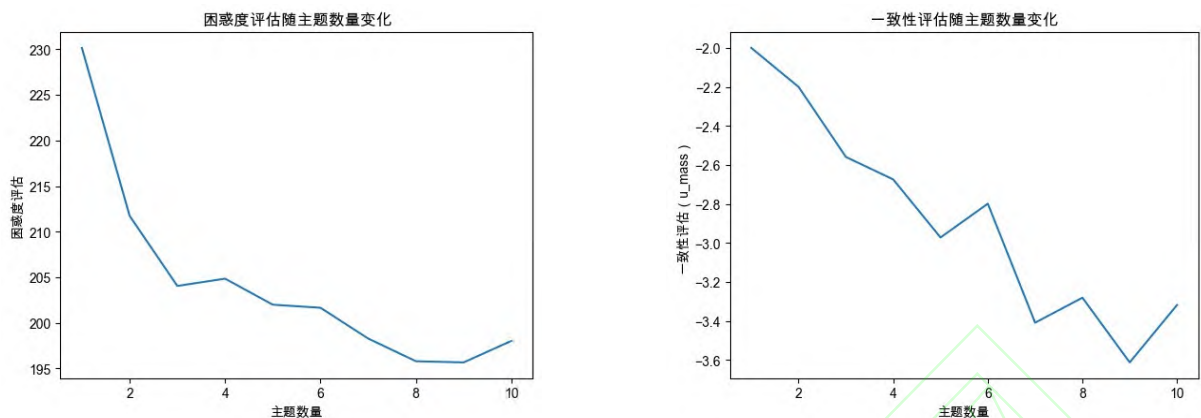


图3 不同主题数的困惑度和一致性分布

Figure 3 Distribution of perplexity and consistency for different number of topics

表3 学术研究主题识别结果

Table 3 Results of academic research theme identification

主题序号	主题词分布	研究主题
Topic 1	服务 建设 高校 科研 高校图书馆 数字化 系统 文献传递 出版	高校图书馆服务建设
Topic 2	知识产权 保护 传统 信息 知识 发展 风险 体系 资源 科技	知识产权保护
Topic 3	数字 出版 借阅 发展 版权 公共 受控 数字版权 模式 人文	数字借阅
Topic 4	数字图书馆 著作权 保护 信息 版权 网络 资源 知识产权 图书馆 建设	著作权保护
Topic 5	资源 数据 图书馆 文献 信息 服务 区块链 风险 数字 领域	数据资源
Topic 6	无障碍 作品 平衡 利益 阅读 授权 条约 格式 实体 马拉喀什	无障碍作品
Topic 7	图书馆 版权 技术 服务 智慧 内容 文本 例外规则 数据挖掘 合理使用	图书馆版权
Topic 8	公共图书馆 作品 合理使用 侵权 著作权 转换 图书馆 实践 司法 著作权法	作品合理使用
Topic 9	人工智能 数据 知识 技术 制度安排 发展战略 文化边界 风控机制 数字化 著作权	人工智能与数据

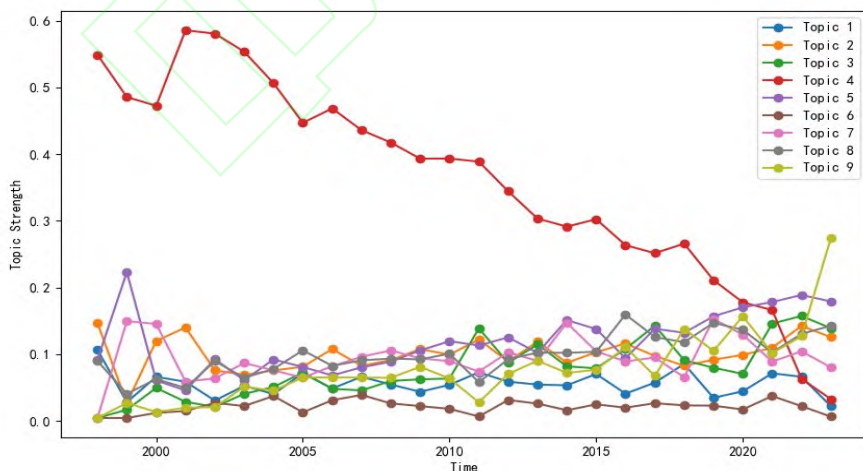


图4 研究主题演化趋势折线图

Figure 4 Line graph of evolutionary trends in research themes

于作品的著作权,以及被侵权人的惩罚性赔偿请求权;第1195条第2款与第1197条分别规定了网络服务提供者未及时采取必要措施的损失扩大连带责任,以及网络服务提供者基于主观过错的连带责任。

将国务院公文主题词表作为自定义词典,对著作权相关

法律法规进行数据预处理,采用word2vec词嵌入模型对著作权法律法规文本进行分析,试图在上下文中找出与著作权侵权最密切相关的问题,结果如图6所示。可见,在数字图书馆著作权侵权领域,问题的关键在于合理使用与侵权行为的界限较为模糊,且经由公共利益问题而与著作权管理及其组



图5 研究主题演化热力图

Figure 5 Heatmap of the evolution of the research theme

表4 数字图书馆著作权侵权领域法律法规重点条文比对

Table 4 Comparison of key provisions of laws and regulations in the field of copyright infringement in digital libraries

法律法规名称	相关修订或新增内容亮点
著作权法(2001修正)	新纳入“信息网络传播权”等著作财产权(第10条)
著作权法(2010修正)	作品范围扩张,不再一概否定违禁作品的著作权,同时进行严格限制(第4条) 权利行使方式扩张,可以著作权出质(第26条)
著作权法实施条例 (2011修订)	依照著作权法有关规定,使用可以不经著作权人许可的已经发表的作品,不得影响该作品的正常使用,也不得不合理地损害著作权人的合法权益(第21条) 有著作权法第四十七条所列侵权行为,同时损害社会公共利益的,著作权行政管理部门可以处非法经营额3倍以下的罚款(第36条)
著作权法实施条例 (2013修订)	有著作权法第四十八条所列侵权行为,同时损害社会公共利益,非法经营额5万元以上的,著作权行政管理部门可处非法经营额1倍以上5倍以下的罚款;没有非法经营额或者非法经营额5万元以下的,著作权行政管理部门根据情节轻重,可处25万元以下的罚款(第36条)
信息网络传播权保护条例 (2013修订)	同时损害公共利益的,可以由著作权行政管理部门责令停止侵权行为,没收违法所得,非法经营额5万元以上的,可处非法经营额1倍以上5倍以下的罚款;没有非法经营额或者非法经营额5万元以下的,根据情节轻重,可处25万元以下的罚款(第18条)
著作权法 (2020修正)	在法律层面引入合理使用判断标准“三步检验法”。(第24条)主体扩张,著作权法保护的主体纳入“与著作权有关的权利人”。(第4条)增加著作权客体的概括性规定、兜底性规定,纳入视听作品概念。(第3条)扩张了著作权内容,在复制权中纳入数字存储过程中的作品复制行为。(第10条)明确了著作权集体管理制度规则。(第8条)新增技术保护措施及其例外。(第49、50条)提高法定赔偿上限和引入惩罚性赔偿(第54条)

织相关联,在数字化技术高度发展和广泛应用的今天,作品的信息网络传播与公众服务紧密相连。

### 3 关联分析

由前述内容分析可知,在学术上,图书馆著作权侵权防范研究的核心在于平衡公共利益与著作权人的私人利益。随着人工智能的发展,学界正试图寻找网络环境下著作权侵权行为与合理使用行为之间新的平衡点。在政策上,相关法律法规的修订或新增内容同样关注到了以信息网络传播权、著作权集体管理为代表的新型法律问题。有鉴于此,为了对学术与政策文本加以充分的关联分析,本研究拟从主体、行为、权利与责任四个方面进行。如图7所示,关联分析冲积

图详尽展示了以上内容间的具体关联。

(1) 主体层面的关联分析。在数字图书馆著作权领域中,相关主体主要包括创作者、授课讲师、视听网站、数字图书馆与用户,在政策层面分别对应著作权人、表演者、录音录像制作者、与著作权有关的权利人、公众。其中,数字图书馆及其合作企业、用户、其他与著作权有关的权利人通常构成著作权侵权的主体。数字图书馆还可能成为教唆、帮助用户实施著作权侵权行为的主体<sup>[26]</sup>。同时,在自建数字资源情形中,图书馆若未获取著作权人许可,则将以内容提供者身份侵犯著作权人的信息网络传播权;在购买与共享数字资源情形中,图书馆或与数据库商等共享主体构成共同侵权,或作为网络服务提供者构成对著作权人的信息网络传播权的间接侵权<sup>[27]</sup>。

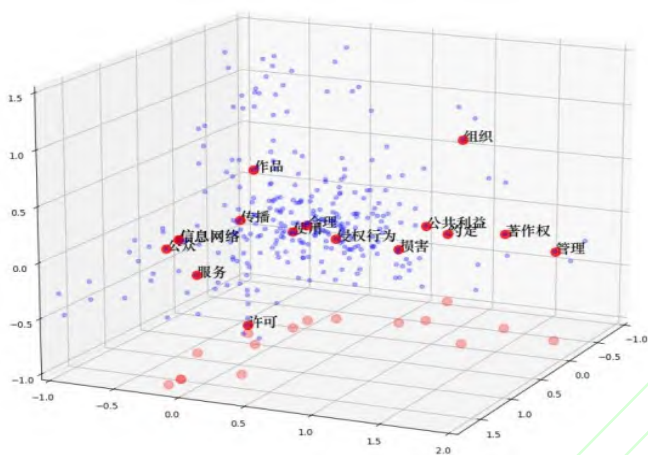


图6 著作权法律法规文本 Word2vec 分析图

Figure 6 Word2vec analysis of the text of copyright laws and regulations

(2)行为层面的关联分析。在数字图书馆建设与服务过程中,无论在事实还是理论层面,作品使用行为的表现形式均呈现出了丰富的样貌,显示出数字技术与人工智能的长足发展。其中,具有代表性的作品使用行为主要包括受控数字借阅(cdl)、带有智能合约的“NFT数字作品”的生成、旨在公益的文本和数据挖掘(tdm)、基于文本和数据挖掘(tdm)的改写,帮助提供定向链接,向用户提供电子图书复制件使其可长期稳定保留作品,在政策层面分别对应著作权许可使用、复制行为,介绍、评论或者说明的使用行为,改编行为、传播行为、发行行为。在理论上,权利通过行为得以实现,同时也因侵权行为而受损。现存争议最大的问题即如何区分侵权行为与合理使用。在人工智能研究主题渐强的趋势下,已经陆续有学者提出应重新审视和界定合理使用行为,如认为“作者中心主义”和严格的“三步检验法”不能适应AI技术变革的需要,著作权法应当采用灵活的、有弹性的合理使用立法模式<sup>[28]</sup>。转换性使用理论是解决问题的核心,但关键在于如何对转换性使用进行界定和适用,司法实践可结合著作权法(2020修正)第24条第(2)项和第(13)项规定,对转换性使用规则进行本土化适用<sup>[29]</sup>。

(3)权利层面的关联分析。自我国著作权法起草以来,始终存在“版权”与“著作权”的法律名称之争。版权侧重于反映对作品复制的权利,著作权则强调对于出版物创作者的权利,特别是能够将书籍图画外的雕刻、模型等美术物纳入权利保护的内容<sup>[30]</sup>。从著作权法三次修正内容来看,现有著作权法条文中仍存在名称混用。此外,在网络环境下,信息网络传播权与发行权的区分历来也备受关注。在NFT数字作品的交易情形下,若参照美国或欧盟将首次交易与转售行为界定为“发行”,则尚须对我国著作权法关于发行行为的构成要件进行调整。实际上,完全可以通过民法典规定的债权转让规则界定交易行为,购买者享有对于“铸造者”的债权,NFT数字作品的后续交易属于债权转让<sup>[31]</sup>。

(4)责任层面的关联分析。在网络著作权侵权行为中,网络服务提供者承担的主要是间接责任,为他人侵权行为负

责的规则是判定网络服务提供者民事责任的基本分析工具<sup>[32]</sup>。数字图书馆间接侵权责任的产生来源,主要包括收藏并外借馆藏资源,给用户查询和打印服务,以连接方式实现资源导航,直接提供作品本身等,特别是在数字图书馆订购数据库资源情形下更易因故意或过失而产生侵权责任<sup>[33]</sup>。此外,在数字图书馆与合作企业共建数据库情形下,也容易因合作企业的不当行为而承担相应责任。

## 4 分析检验

为了检验前述关联分析结果的有效性,本研究将在分析采集到的2008年至2023年的36份民事裁判文书裁判依据的基础上,对数字图书馆著作权侵权的理论及政策问题加深剖析,从而对现有学术观点与政策内容的合理性加以验证,并据此提出本研究的最终结论。如表5所示,从案件裁判依据的分布来看,司法实践中的争议问题聚焦于侵权赔偿、民事责任、未经许可的使用、著作权内容、作者与著作权的归属、合理使用与公共图书馆,涵盖了数字图书馆著作权侵权的主体、行为、权利与责任四个方面,可见前述关联分析所选取的维度与司法实践争议问题相符合。据此,本研究拟选定融合上述四方面因素的典型案例加以深入分析并完成检验工作。

### 4.1 公共利益因素对合理使用认定的影响

公共图书馆作为社会公益组织未经许可使用他人作品是否属于合理使用?在深圳图书馆、读者侵害作品信息网络传播权纠纷案中,深圳图书馆向合作公司订购数字资源平台,以年付费方式远程使用。同时,读者通过互联网进入深圳图书馆的页面,分别点击“数字资源”“中华数字书苑”后的“馆外访问”,通过证号和密码登录,可以检索到并“在线阅读”涉案作品。该涉案作品未获得作者授权。法院基于两点理由否定了深圳图书馆对其行为属于合理使用的说法,可概括为“未经许可使用+公益不等于无限制”,即深圳图书馆作为公共图书馆虽应向公众提供免费的阅读服务,但应事先取



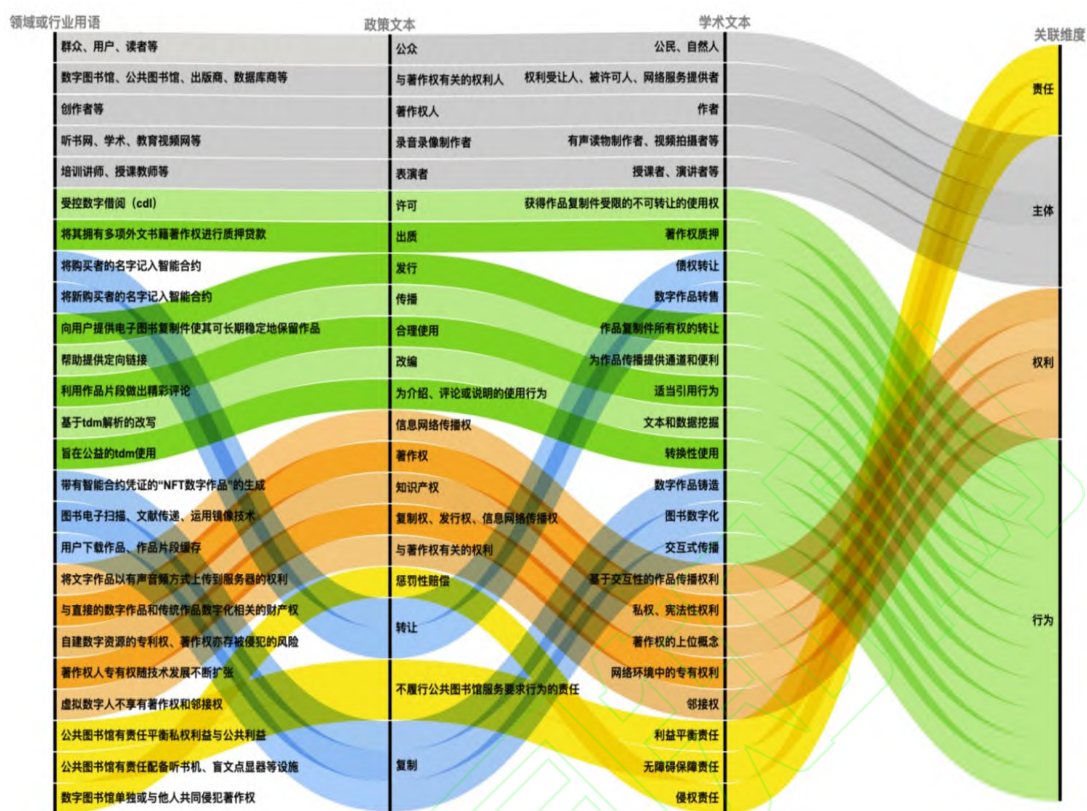


图7 关联分析冲积图

Figure 7 Correlation analysis alluvial map

表5 数字图书馆著作权侵权纠纷案件主要实体裁判依据

Table 5 The main substantive bases of judgement in copyright infringement disputes in digital libraries

裁判依据	主题	案件数量
著作权法(2001修正)第48条	侵权赔偿	30
著作权法(2010修正)第49条		
著作权法(2020修正)第54条		
著作权法(2001修正)第47条第1项	民事责任	27
著作权法(2010修正)第48条		
信息网络传播权保护条例第7条	著作权合理使用	17
信息网络传播权保护条例(2013修订)第7条		
著作权法(2010修正)第22条		
著作权法(2010修正)第10条	著作权内容	10
著作权法(2020修正)第10条		
著作权法(2010修正)第11条	作者与著作权的归属	7
著作权法(2020修正)第12条		
公共图书馆法第2条	公共图书馆	1

得许可并在合理限度内使用作品。同时,深圳图书馆与合作企业共建数据库时,有义务确保作品权利来源合法,读者的阅读和复制行为不能超过必要限度从而保障作者获得报酬的权利,否则应与合作企业承担连带赔偿责任<sup>[34]</sup>。可见,公共利益作为一项重要的影响因素,显然会影响到法院关于数字图书馆行为是否属于合理使用的认定,或者说可以作为数

字图书馆涉诉时的一个重要的抗辩理由,但应该明确的是公益不等于无限制,而是要在著作权人利益与公共利益之间达到合理的平衡。

#### 4.2 复制行为与信息传播行为关系的认定

未经许可复制他人作品并进行信息传播的行为是

否属于合理使用?在王莘与谷翔、谷歌公司侵害著作权纠纷案中,谷歌公司未经作者许可将涉案作品进行电子扫描,谷翔公司对涉案作品进行了信息网络传播。问题在于,虽然未经许可的复制原则上构成侵权,但有些合理使用行为的实施实际上以复制为前提。一审法院认为,谷翔公司的信息网络传播行为是否构成合理作用,与谷歌公司在先的复制行为无关。二审法院对此予以指正,认为专门为合理使用而进行的复制,应与后续使用行为作为整体看待。由于原告虽对谷翔公司信息网络传播行为构成合理使用提出异议,但未提出上诉,所以二审法院不予审理,最终认定谷歌公司实施的复制行为构成侵权<sup>[35]</sup>。由此可见,著作权法原则上认为复制行为构成侵权,但著作权法(2020修正)第24条对于合理使用的规定,旨在对著作权做出合理限制,从而合理地认定在先的复制行为的正当性。

### 4.3 数字作品或作品数字化信息网络传播权侵权的认定

传统图书馆未经许可在网络社交平台使用他人网络数字作品是否构成侵权?在四会市图书馆被诉侵害作品信息网络传播权纠纷案中,作品信息网络传播权人称网络数字作品发表于作者注册的微信公众号,作者将其信息网络传播权转让给了现权利人,后发现四会市图书馆在其官方微信平台“四会市图书馆”上转载了涉案作品,故认为侵害了其信息网络传播权,图书馆则辩称并未直接获得经济利益。法院认为四会市图书馆构成对作品信息网络传播权的侵权,理由在于转载行为实际上是将作品复制供公众阅读,但既未取得授权或许可,也未指明作者姓名,且不属于依法可以不经著作权人许可,不向其支付报酬而使用作品的情形。最终法院判决四会市图书馆应承担包括经济损失及维权合理开支在内的侵权损害赔偿<sup>[36]</sup>。在绵阳市图书馆被诉侵害作品信息网络传播权纠纷案中,合作公司通过镜像技术将电子书完全复制到绵阳市图书馆本地服务器,从而使公众可通过图书馆系统直接下载涉案电子书,使公众可以在选定的时间和地点获得涉案作品,构成对作品信息网络传播权的侵害。同时,由于图书馆与合作公司的合同目的是通过图书馆系统实现电子书的在线阅读或下载,二者具有通过信息网络向公众提供涉案作品的意思联络,在未经授权(授权已过期)的情况下构成共同侵权,应承担连带责任<sup>[37]</sup>。由此,法律实际上对既有权利及经过数字化与数据化后的客体进行了保护<sup>[38]</sup>。

## 5 结 语

本研究基于对学术与政策文本的关联分析与验证,发现现有关于数字图书馆著作权侵权的研究领先于政策制定。通过考察司法裁判文书对前述关联分析进行检验,可以在侵权防范问题上得出以下三个基本结论:

第一,有关主体均在不同程度上对数字作品的合理使用存在认识偏差。数字图书馆之所以频频卷入著作权侵权纠

纷中,一个最根本的原因就是没有充分认识到取得权利人许可或授权的重要性。客观上,数字图书馆的困境在于逐一取得作者的许可或授权成本过高,且有时无法与作者取得联系。但深层原因在于许多公共图书馆认为基于保障公众阅读基本权利就可以随意使用或传播作品,显然缺乏对于著作权作为私权最起码的尊重。同时,作为与数字图书馆合作的相关技术公司往往出于逐利目的而忽视对著作权人的尊重,在很大程度上提高了数字图书馆成为间接侵权责任主体的风险。对此,数字图书馆自身应加强合规建设,重点关注与技术公司合作模式的法律风险,国家或行业协会亦应定期向社会公开发布相关合同示范文本。

第二,人工智能为构建具有中国特色的著作权制度体系带来了重要契机。在人工智能长足发展的当下,出现了许多有代表性的如受控数字借阅(cdl)、带有智能合约的“NFT数字作品”的生成、旨在公益的文本和数据挖掘(idm)、基于文本和数据挖掘(idm)的改写等事实行为,如何对相关行为予以法律评价迫在眉睫。对此,我国不应盲目借鉴美国或欧盟关于NFT数字作品的交易等法律问题中关于相关行为与权利性质的界定,反而应循此契机进一步完善我国基于信息网络传播权的新型著作权法律制度体系。

第三,著作权集体管理制度功能的发挥成为著作权侵权防范的关键。数字图书馆与合作公司以往并不重视作者权益的另一个主要原因在于违法成本过低,现有制度未能对作者提供充分有效的保护。特别是在网络环境下,随着人工智能内容生成技术的发展,如何在利益平衡的基础上促进创作与创新,既需要认可促进作品传播等有利于作者权益的行为的正当性,更需要预防不当行为对著作权的侵害,著作权集体管理制度功能的充分发挥就显得格外重要。本研究认为,主要应从法律关系厘清与降低制度利用成本两大角度予以完善,一个基本的认识就是著作权集体管理属于信托法律关系,对于降低制度利用成本问题,可考虑通过慈善信托的方式予以解决。

## 参考文献

- [1] 中共中央国务院印发《数字中国建设整体布局规划》[N]. 人民日报,2023-02-28(001).
- [2] 程文艳. 数字图书馆——数字中国的开路先锋[J]. 情报科学,2002(6):582-585.
- [3] 林扬. 著作权保护与数字图书馆建设[J]. 情报科学,2002(5):490-492.
- [4] 杨家富. 馆藏资源的数字化建设与著作权法的保护[J]. 情报科学,2002(4):399-401.
- [5] 杨向明. 论数字化作品的知识产权问题[J]. 情报科学,2003(1):60-61,74.
- [6] 邱均平,朱少强. 数字图书馆版权保护技术及其规避行为的法律对策[J]. 情报科学,2006(1):1-7.
- [7] 徐恺英,刘佳. 信息时代著作权保护中的利益关系与对策研究[J]. 情报科学,2007(5):728-732.

- [8] 李华伟,邱奉捷.2008-2010年我国数字图书馆知识产权及相关法律问题研究之比较[J].大学图书馆学报,2012,30(4):50-54,10.
- [9] 刘淑宝,王爱霞.民法视角下数字图书馆用户著作权侵权责任探究[J].图书馆,2017(11):72-75.
- [10] 吉宇宽.图书馆数字阅读服务中的著作权侵权规避策略研究[J].图书馆工作与研究,2020(10):5-11.
- [11] 杨帆.法社会学能处理规范性问题吗?——以法社会学在中国法理学中的角色为视角[J].法学家,2021(6):30-44,191-192.
- [12] 徐健,刘政,桂思思,等.领域学术观点库构建关键问题研究[J].情报理论与实践,2021,44(12):158-164.
- [13] 刘天畅,朱庆华,赵宇翔.中国适老化改造政策的文本分析与演化特征研究[J/OL].情报科学:1-10[2023-09-25].
- [14] 彭中礼.政策概念的法规规范分析——基于1979—2016年现行有效法律文本的解读[J].安徽大学学报(哲学社会科学版),2016,40(3):113-122.
- [15] 美国俊,罗凯方.中国环境问责制度的嬗变特征与演进逻辑——基于政策文本的分析[J].行政论坛,2019,26(4):101-109.
- [16] 冯莉.《黄河保护法》实施背景下流域生态环境规制思路与完善对策——基于法律政策文本量化分析[J].干旱区资源与环境,2023,37(7):190-196.
- [17] 闫慧,贾诗威.信息资源管理学科与国家战略的关系探究——基于学术话语与政策文本的关联分析[J].情报资料工作,2023,44(2):5-13.
- [18] 王锴.法律位阶判断标准的反思与运用[J].中国法学,2022(2):5-25.
- [19] 周蔚,王兆毓,魏斌.面向法律裁判文书的生成式自动摘要模型[J].计算机科学,2021,48(12):331-336.
- [20] 王东艳.中美数字图书馆研究比较[J].情报资料工作,2001(6):32-36.
- [21] 吉宇宽.论著作权私权公权化趋势下图书馆的责任[J].图书馆,2009(4):20-23.
- [22] 谭华霖.知识产权权利冲突的内在机理及化解机制[J].知识产权,2011(2):23-29.
- [23] 阳贤文.从谷歌案析数字图书馆之困境与发展[J].图书馆建设,2015(5):84-87.
- [24] 周贺微.美国转换性使用转型及对我国的借鉴[J].新闻界,2019(4):74-84.
- [25] 易磊.我国著作权合理使用制度的宪法学释义[J].南大法学,2022(1):88-102.
- [26] 王蓉光.网络信息著作权侵权的司法取向[J].现代情报,2004(6):210-212.
- [27] 朱莹.公共图书馆数字化服务中的信息网络传播权侵权问题研究[J].上海法学研究,2022(20):104-110.
- [28] 林秀芹.人工智能时代著作权合理使用制度的重塑[J].法学研究,2021,43(6):170-185.
- [29] 袁锋.元宇宙空间著作权合理使用制度的困境与出路——以转换性使用的界定与适用为视角[J].东方法学,2022(2):44-57.
- [30] 刘春田.中国著作权法三十年(1990-2020)[J].知识产权,2021(3):3-26.
- [31] 王迁.论NFT数字作品交易的法律定性[J].东方法学,2023(1):18-35.
- [32] 吴汉东.论网络服务提供者的著作权侵权责任[J].中国法学,2011(2):38-47.
- [33] 王果,张立彬.面向发展的图书馆著作权侵权诉讼趋势及其风险规避策略研究[J].情报科学,2020,38(3):87-92.
- [34] 广东省高级人民法院.(2018)粤民申11343号深圳图书馆、詹某某侵害作品信息网络传播权纠纷再审审查与审判监督民事裁定书[EB/OL].[2024-01-16].<https://www.lawsdata.com/#/documentDetails?id=5e00ad0e7986893be712fd61&type=1&keyword=WyIyMDE4Iiw57Kk5rCR55SzMTEzNDPlj7ciXQ%3D%3D>.
- [35] 北京市第一中级人民法院.(2011)一中民初字第1321号王莘与北京谷翔信息技术有限公司、谷歌公司侵害著作权纠纷一审民事判决书[EB/OL].[2024-01-16].<https://wenshu.court.gov.cn/website/wenshu/181107ANFZ0BXS4/index.html?docId=i37nDuF0QAJbGiizEGellh-NFirdGPGRHfXzJFclnDYPaj1gYSuGBjpO3qNaLMqsJKAEEXEwiti1WIqQVGvhWczwmZ65MW8p2qoPXci/kks4Vxyp4BvKOBfdZp059GaoW>.
- [36] 天津市和平区人民法院.(2018)津0101民初5405号天津律德文化传播有限公司与四会市图书馆侵害作品信息网络传播权纠纷一审民事判决书[EB/OL].[2024-01-16].<https://susong.tianyancha.com/4d8294bdf3fa4e99b3ff100630666c40>.
- [37] 北京知识产权法院.(2022)京73民终2936号北京世纪超星信息技术发展有限责任公司等与北京大学出版社有限公司侵害作品信息网络传播权纠纷二审民事判决书[EB/OL].[2024-01-16].<https://wenshu.court.gov.cn/website/wenshu/181107ANFZ0BXS4/index.html?docId=q5nmPTFTik/sYXR0TPEr/4gvPhznpxpXyTeNfUEdu-jbaj/1QoZOVXpO3qNaLMqsJKAEXEwiti1WIqQVGvhWczwmZ65MW8p2qoPXci/kks6Q6054HZpdUcPsRzE636pC>.
- [38] 周清华,万丰阁.数据产权的解码与构建——以控制与管理定位[J].中国海商法研究,2023(4):27-37.

(责任编辑:毛秀梅)

---

# Institutions, Theories and Practices of Copyright Infringement Prevention in Digital Libraries

—A Correlation Analysis Based on Academic and Policy Texts

*CUI Hongming*

*(School of Law, Dalian Maritime University, Dalian 116026, China)*

**Abstract:** **【Purpose/significance】**By analyzing the institutions, theory and judicial practice of copyright infringement prevention in digital libraries, the core controversial issues and improvement paths to prevent copyright infringement are clarified. **【Method/process】**Taking chinalawinfo.com and CNKI as data sources, we selected laws and regulations, policy documents, authoritative academic papers and typical case judgments as research samples, and comprehensively applied the methods of Keyword Co-Occurrence, Domain Academic Claim Base System, DTM analysis, Text Comparison analysis and Word2vec analysis to carry out the correlation analysis. **【Result/conclusion】**Through the correlation analysis and test, it is found that the subjects concerned have different degrees of cognitive bias towards the reasonable use of digital works, artificial intelligence brings an important opportunity to build a copyright system with Chinese characteristics, and the function of the copyright collective management system has become the key to the prevention of copyright infringement. **【Innovation/limitation】**With the help of the method of constructing a library of academic views in the field to assist text analysis, it has an upgrading effect on the separate use of text analysis methods, and has an innovative value in the study of perfecting the path of preventing copyright infringement in digital libraries. Since correlation research is a complex task, the results will be affected by the quality of data, pre-processing methods and the chosen methods. The construction of the Domain Academic Claim Base System is subject to further rationality and feasibility verification, which may bring about minor uncertainties.

**Keywords:** digital libraries; infringement prevention and mitigation; scholarly opinion; policy texts; correlation analysis