

人工智能作品合理使用困境及其解决

张金平

内容提要:人工智能是收集、整理海量信息并进行计算机分析的技术,因而在一开始就可能复制大量他人享有著作权的作品。这些作品的使用(尤其是商业目的使用),并不必然可以援引现行著作权法的合理使用规定而豁免。如此一来,企业为开展人工智能领域的研究必须对大量作品的复制而支付许可费用。这不仅在实践中不可能完成,也会严重阻碍我国在该领域的科技进步。与此同时,现有著作权法修订草案所提供的合理使用原则性条款也无法为人工智能发展过程中的作品合理使用提供确定性。在美国、日本和欧盟等国家或地区已经提供或正在讨论提供合理使用解决方案的情况下,为扫清人工智能发展的著作权障碍并提供国际相对竞争优势,我国有必要通过修法提供专门的著作权限制与例外。同时,为了防止该例外的扩大化适用对著作权人产生不合理的损害,仍须科学界定该例外的具体适用情形。

关键词:人工智能 著作权 商业目的 合理使用

张金平,中央财经大学法学院博士后。

人工智能在近几年为社会各界所热烈讨论,各国政府也纷纷发布各自的人工智能战略,而在著作权法领域的相关讨论,则集中于人工智能生成物是否构成作品以及机器人是否构成著作权法意义上的作者等问题。^[1] 这种讨论的意义在于追索著作权法的创设初衷和规范逻辑,反思传统著作权法以自然人作者为核心的制度安排,为未来的技术发展提

[1] 参见吴汉东:《人工智能时代的制度安排与法律规制》,《法律科学》2017年第5期,第131-132页;易继明:《人工智能创作物是作品吗?》,《法律科学》2017年第5期,第137-147页;王迁:《论人工智能生成的内容在著作权法中的定性》,《法律科学》2017年第5期,第148-155页;梁志文:《论人工智能创造物的法律保护》,《法律科学》2017年第5期,第156-165页;陶乾:《论著作权法对人工智能生成成果的保护》,《法学》2018年第4期,第3-15页。在国外这方面的讨论则更早并一直持续到现在,具体参见 Timothy L. Butler, Can a Computer be an Author: Copyright Aspects of Artificial Intelligence, 4 *Comm/Ent L. S.* 707 (1981); Kalin Hristov, Artificial Intelligence and the Copyright Dilemma, 57 *IDEA* 431 (2017)。

供适当的激励。^[2] 然而,当我们把所有注意力集中在人工智能输出端生成物的著作权法定性时,我们往往忽视在人工智能的输入端首先是收集和整理海量信息;没有这一过程作为基础,人工智能生成物根本不可能出现。一旦这些海量数据中的部分内容成为著作权法保护的作品,按照著作权法作品使用必先获得授权的原则,那么人工智能对于作品的复制等使用也必须先获得许可,否则即构成著作权侵权。^[3]

然而,在人工智能时代,企业通过人工智能处理的信息中受著作权保护的作品数以百万计,如果都需要一一事先获得许可,我国从事人工智能的个人和组织(尤其是企业)将不堪重负,进而影响我国人工智能技术的研发和广泛应用。与此同时,如果人工智能对于作品的使用构成合理使用,那么传统以公共利益造福社会大众的合理使用制度^[4]就可能容易被采用人工智能的大企业所滥用,反过来损害处于相对弱势的数以万计作者的利益。^[5] 因此,人工智能的作品使用由此陷入合理使用困境,为我国人工智能产业发展蒙上一层阴影。有鉴于此,本文借我国《著作权法》自主修订之机,对这一问题加以探究和明确。

一 人工智能对作品的利用形式

语音识别、图像识别和自动驾驶为大家所熟知,但“人工智能”的概念却尚未统一。人工智能领域最权威教材的两位作者斯图尔特·拉塞尔教授(Stuart Russell)和皮特·诺维格教授(Peter Norvig)将1978年以来有关人工智能的八种定义归纳为四类:像人类一样思考、像人类一样行为、理性思考、理性行为。^[6] 概括而言,人工智能指的是现代计算机模拟人的认知能力自动进行知识学习和对问题的解决。^[7] 其中,人工智能的认知能力并非凭空产生,而是在学习人类现有知识的基础之上产生的。在技术原理上,学习和处理过程大体经历两个环节:第一个环节表现在计算机输入端可以访问或者获取海量的知识和信息,并对这些知识和信息数字化和结构化;第二个环节则对结构化数据进行计算机分析,并在输出端得出分析的结果,从而认识事物发展的模式、趋势、模式与趋势之间的关系,甚至作出决策。^[8] 在这个意义上,人工智能对于知识和信息的处理,又常被称为“文本与数据挖掘”(text and data mining),或者“计算机信息分析”(computer analysis)。^[9] 此外,马克·普朗克创新与竞争研究中心还强调,“尽管文本与数据挖掘非常重要,但仅仅

[2] Adolf Dietz, Copyright in the Modern Technological World: A Mere Industrial Property Right, 39 *J. Copyright Soc'y U. S. A.* 83 (1991).

[3] 参见李明德、许超著:《著作权法》,法律出版社2009年版,第210页。

[4] 参见李明德、许超著:《著作权法》,法律出版社2009年版,第94-95页。

[5] Intellectual Property Office of United Kingdom, *Modern Copyright: A Modern, Robust and Flexible Framework*, 2013, p. 36.

[6] Stuart Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed., New Jersey: Prentice Hall, 2010, pp. 1-2.

[7] 参见腾讯研究院、中国信通院互联网法律研究中心著:《人工智能》,中国人民大学出版社2017年版,第3-4页。

[8] Christophe Geiger, *In-Depth Analysis for European Parliament: The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market-Technical Aspects*, Brussels, 2018.

[9] European Commission, *Proposal for Directive on Copyright in the Digital Single Market*, COM (2016) 593 final, p. 24.

是计算机信息分析的第一个环节。”^[10]

其中,人工智能第一个环节中对知识和信息进行的数字化,指的是将知识和信息转化为计算机可读的数据格式,而且通常选取可兼容的标准数据格式,如最简单的将纸质的书转化为 word 文档或者 pdf 文件。结构化则是在数字化之后,剔除其中多余和错误的信息,并按照一定的结构和模式编排这些数据,便于第二个环节的分析。通常而言,知识和信息的数字化有两类实现方式。第一类方式是实施人工智能的主体针对原本并非数字格式的内容进行数字化,并形成数据格式副本。例如,2005 年以来,谷歌在全世界范围内扫描了大量图书并形成中央集成式的数据库,在此基础上发展出文献搜索、语言模式、智能翻译等方面的人工智能。^[11] 另一类实现方式则由他人先行完成数据格式转换,实施人工智能的主体在此基础上通过授权或者非授权方式访问和获取(如通过 API 端口)该内容。例如,斯坦福大学李飞飞教授为收集足够多的图片发展计算机图像识别的自主学习能力,通过网络爬虫直接抓取并下载互联网上的 10 亿张图片,动用来自 167 个国家的 5 万多个工作人员来清理、分类和标记出 2.2 万种物品的 1500 万照片,并将该数据库免费开放给全球的科研人员。^[12] 值得注意的是,第二类实现过程还可能为了标准化分析而进行数据的格式转化。

从著作权法意义上而言,只要人工智能处理的知识与信息属于作品且仍在保护期之内,这些知识和信息的数据格式生成或者数据格式的转化,都属于著作权法意义上的“复制”。按照我国《著作权法》第 10 条的规定,复制权所规制的行为是“以印刷、复印、拓印、录音、录像、翻录、翻拍等方式将作品制作一份或者多份的权利”。因此,人工智能第一环节的作品数字化虽不在该条明确列举的作品复制方式之中,但只要在结果上制作了一份或者多份复制品则仍为该条的“等”术语所涵盖。

除了对作品的复制之外,人工智能在第二环节还可能对作品进行改编或者演绎等形式的利用。^[13] 我国学界当前热烈讨论的人工智能创作物,如果构成作品,那么该创作物很可能是改编或演绎了他人享有著作权的作品。按照我国当前作品侵权的司法裁判规则“接触+实质性相似”,要满足接触要件,只要证明其数据库中存储有原告作品即可;而要证明实质性相似要件,则通过作品比对可以实现。^[14] 例如,巴黎索尼计算机科学实验室在音乐家约翰·塞巴斯蒂安·巴赫的 352 首合唱歌曲为基础训练其人工智能系统,由此“创作”出的 2503 个作品,在 1600 名专业听众的判断下有超过一半的人以为是巴赫本人

[10] Max Plank Institute for Innovation and Competition, *Position Statement on the Proposal Modernisation of European Copyright Rules*, Part B: Exceptions and Limitations (Art. 3—Text and Data Mining), 2017, para. 6.

[11] 谷歌员工在接受乔治·戴森采访时指出,谷歌扫描图书的目的不是为了人们阅读,而是为了人工智能。参见 George Dyson, *Turing's Cathedral*, <https://perma.cc/EHM4-TLWY>, 最近访问时间[2018-12-10]。

[12] 参见李飞飞:《如何教计算机理解图片》, https://open.163.com/movie/2015/3/Q/R/MAKN9A24M_MAKN9QAQR.html, 最近访问时间[2018-12-10]。

[13] Benjamin Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 *Colum. J. L. & Arts* 45, 64-66 (2017).

[14] 参见最高人民法院(2013)民申字第 1049 号民事裁定书;吴汉东:《试论“实质性相似+接触”的侵权认定规则》,《法学》2015 年第 8 期,第 67-68 页。

的作品。^[15] 如果巴赫的作品还在保护期内,人工智能未经巴赫许可创作的上述“作品”很容易被认定为抄袭。最后,应当强调的是,人工智能第二环节对于作品结构化数据的分析,就好比人们对于作品的阅读与欣赏,本身并不是著作权法意义上的利用形式,无须考虑其著作权法意义上的合法性问题。

二 我国人工智能作品合理使用的障碍

人工智能对于作品的著作权法意义上的利用主要为复制和演绎。不过,鉴于目前国内对于人工智能创作物的作品属性尚缺乏定论,本文并不对该种改编或演绎是否属于合理使用展开分析,而仅针对复制进行论证。表面上而言,人工智能涉及的作品复制是否构成合理使用,在我国法中是一个非常简单的判断问题:我国现行《著作权法》并未明确人工智能场景中作品复制属于著作权法的限制与例外,因而这种行为并不属于法定的合理使用情形,原则上应当事先获得著作权人的许可,否则构成侵权。^[16] 然而,在实践中该判断要复杂得多。

(一) 现行例外与限制的可用性挑战

首先,以科研目的为主的人工智能研究是否属于我国现有的科研目的的合理使用情形存有争议。目前,我国《著作权法》第 22 条规定两种科研目的下的合理使用情形。然而,这两项科研目的的合理使用情形都难以适用于我国开展人工智能科学研究的相关企业。

其中,第一种情形为第 22 条第 1 款第 1 项规定的“为个人学习、研究,使用他人已经发表的作品”。表面上,该例外似乎可以合法化个人开展人工智能方面的学习、研究。然而,“个人学习、研究”中的“个人”通常仅限于“个人或家庭”的范围,并不能当然延及个人的科研团队。^[17] 相比之下,人工智能作为一个系统性工程,单个个人基本上无法承担这类科研的费用,也无法单独完成这样的科研任务。况且,个人学习、研究也限于非商业目的的使用。因此,以“个人研究”名义来合法化科研人员在开展人工智能过程中大量作品复制并不现实。

第二种情形系第 22 条第 1 款第 6 项规定的“为学校课堂教学或者科学研究,翻译或者少量复制已经发表的作品,供教学或者科研人员使用,但不得出版发行”。相对于“个人研究”,该种合理使用情形是专门为科研机构的科学研究而设,因而与科研机构开展人工智能研究看似也非常契合。不过,该例外属于以公共利益为由对著作权进行的限制,因而该情形下的科研机构及其科研人员应当限于国家设立的教育、科研公共事业单位,比如经相关主管部门批准设立的学校、全额财政拨款的国家科研机构等。^[18] 假定李飞飞教授

[15] 参见易继明:《人工智能创作物是作品吗?》,《法律科学》2017年第5期,第139页。

[16] 参见北京市第一中级人民法院(2011)一中民初字第1321号民事判决书、北京市高级人民法院(2013)高民终字第1221号民事判决书。

[17] 参见李明德、管育鹰、唐广良著:《〈著作权法〉专家建议稿说明》,法律出版社2012年版,第248页。

[18] 参见李明德、管育鹰、唐广良著:《〈著作权法〉专家建议稿说明》,法律出版社2012年版,第251-252页。

在中国大学开展图像识别研究,且该项目主要受中国政府资助,那么她在前文提及的 10 亿张照片的复制,不管其中受版权保护的比例多大,她都可援引该例外省去一份巨额的著作权许可费。然而,如果此项研究主要受企业赞助,又或者她直接作为企业的雇员开展此项研究,那么这笔巨额的版权许可费原则上就无法回避。^[19]

(二)在司法实践中创设新合理使用情形的不确定性

在个案中,我国法院能否以合理使用的原则来合法化企业的相关人工智能研究对作品的使用也存在争议。虽然我国《著作权法实施条例》第 21 条将《伯尔尼公约》的三步检测法作了规定,但三步检测法的适用仍限于“依照著作权法有关规定”,即受限于《著作权法》第 22 条规定的情形。^[20] 在实践中,我国法院已经超出《著作权法》第 22 条规定认定新的合理使用情形。其中,针对新形势下与信息网络传播权有关的合理使用,最高人民法院在 2012 年发布的《关于审理侵害信息网络传播权民事纠纷案件适用法律若干问题的规定》第 5 条第 2 款中规定,网络服务提供者“以提供网页快照、缩略图等方式实质替代其他网络服务提供者向公众提供”(属于信息网络传播权规制的行为)相关作品,“不影响相关作品的正常使用,且未不合理损害权利人对该作品的合法权益”,而且“网络服务提供者主张其未侵害信息网络传播权”的,人民法院应予支持。因此,最高人民法院允许下级法院依据三步检测法的第二、三步,创设新的合理使用情形。例如,在 2013 年的丛文辉诉搜狗案中,北京市高级人民法院维持了北京市第一中级人民法院有关搜狗提供网页快照构成合理使用的判决。^[21]

在人工智能领域发生的王莘诉谷歌案中,我国法院虽然认定谷歌将其扫描的图书以片段式向公众提供图书及其关键词的搜索结果构成信息网络传播权方面的合理使用,但并不认为谷歌全书扫描的行为构成复制权方面的合理使用。而且一二审法院在复制权方面的合理使用认定标准上也并非一致:一审法院依据三步检测法第二、三步认为全文扫描原告作品与作品正常利用方式相冲突,且在后果上不合理损害原告的利益;^[22] 而二审法院认为可以参考美国合理使用四要素认定新的合理使用情形,但使用人必须充分举证证明其行为构成复制权方面的合理使用,而谷歌并未提交相关证据,故驳回其上诉。^[23]

在其他个案中,我国法院也常混合使用三步检测法和美国合理使用四要素审查某个行为是否构成合理使用,但在结果上却难以预测。^[24] 即使在个案中被动认定人工智能领域的作品复制构成合理使用,但在并非推行判例制的我国仍然难以合理预测类似复制行

[19] 英国知识产权局明确指出这两种行为都认为不能援引英国版权法下的科研机构的科研目的例外。参见 Intellectual Property Office of United Kingdom, *Exceptions to Copyright: Research*, 2014, p. 9。

[20] 参见冯晓青、胡少波:《互联网挑战传统著作权制度》,《法律科学》2004 年第 6 期,第 124 页;卢海君:《论合理使用制度的立法模式》,《法商研究》2007 年第 3 期,第 28 页;朱理著:《著作权的边界——信息社会著作权的限制与例外研究》,北京大学出版社 2011 年版,第 206-207 页;黄玉焯:《著作权合理使用具体情形立法完善之探讨》,《法商研究》2012 年第 4 期,第 22 页;熊琦:《著作权合理使用司法认定标准释疑》,《法学》2018 年第 1 期,第 192 页。

[21] 参见北京市高级人民法院(2014)高民申字第 1783 号民事裁定书。

[22] 北京市第一中级人民法院(2011)一中民初字第 1321 号民事判决书。

[23] 北京市高级人民法院(2013)高民终字第 1221 号民事判决书。

[24] 参见李琛:《论我国著作权法修订中“合理使用”的立法技术》,《知识产权》2013 年第 1 期,第 16-17 页。

为的合法性,反而容易因创设新的合理使用而被诟病为法官造法。^[25]

三 域外合法化人工智能作品利用的方案

为了解决人工智能领域作品复制合法性方面的不确定性,科研机构和企业可以自行采用市场化解决方案,例如规避使用受著作权法保护的作品或者事先取得著作权人的许可,但这往往会加重这些机构和企业开展人工智能研究的成本,最终减慢科研进度和社会进步。为了给创新者提供法律确定性和减轻创新成本,人工智能产业发展相对较快的美国、日本和欧盟已经采取或者正在通过各自的立法或者司法加以解决,但这些解决方案都有其各自的法律传统和制度考量。

(一) 美国的转化性使用司法裁判和行业自律

美国 1976 年版权法通过立法将司法中形成的合理使用四要素明确下来,但同时强调此举并非要通过立法僵化合理使用理论,而是仍留给法院根据这四个要素去发展合理使用制度。^[26] 其中重要的发展是 1990 在纽约南区地区法院任职的皮埃尔·勒瓦尔法官 (Pierre N. Leval) 针对第一个要素“使用的目的和特点”提出了转化性理论,即“该使用必须是富有成效的,必须以不同于原作的方式或者为了与原作不同的目的而使用他人作品”,并强调这是四要素中最为重要的判断因素,因为著作权法的根本目的是为了激励更多作品的创作。^[27] 四年后,该理论在坎贝尔案中就被美国联邦最高法院所接受。^[28]

正是基于转化性理论,美国法院在司法实践中判决人工智能领域的作品复制构成合理使用。其中,2015 年判决的美国作家协会诉谷歌案恰好由升任美国联邦第二巡回法院法官的勒瓦尔主审。在该案中,谷歌未经作者许可扫描大量图书并向公众提供这些数字化作品的搜索及小片段内容的展示,同时其搜索引擎也允许用户查询某个词或者短语在不同时代的使用频率、语法模式以及在何种主题下的使用。^[29] 作为二审法院的勒瓦尔维持了一审法院有关谷歌的上述作品复制和使用符合转化性理论要求的认定,“谷歌为了批评、评注原作,或者为了提供原作所没有的信息的目的而复制,非常清楚地表明其满足坎贝尔案有关合理使用第一个要素中的转化性目的的要求……在本案中,我们没有理由认为,谷歌的整体盈利目的是否定其虽具有高度说服力的转化性目的但仍不构成合理使用的原因。”^[30] 此外,美国法院在论文查重案、法律数据库案等案中,同样基于其中的作品复制具备转化性功能而认定人工智能第一环节的作品复制行为构成合理使用。^[31]

[25] 参见崔国斌:《知识产权法官造法批判》,《中国法学》2006 年第 1 期,第 146-164 页;熊琦:《著作权合理使用司法认定标准释疑》,《法学》2018 年第 1 期,第 183-186 页。

[26] H. R. Rep. No. 94-1476, p. 66.

[27] Pierre N. Leval, Toward a Fair Use Standard, 103 *Harv. L. Rev.* 1105, 1112 (1990).

[28] *Campbell v. Acuff Rose Music, Inc.*, 510 US 569, 578-585 (1994).

[29] *Authors Guild v. Google, Inc.*, 804 F.3d 202, 209 (2015).

[30] *Authors Guild v. Google, Inc.*, 804 F.3d 202, 219 (2015).

[31] *A. V. v. Iparadigms. LLC.*, 562 F.3d 630, 633-644 (2009); *Edward White v. West and Lexis*, No. 12 Civ. 1340 (JSR) (S. D. N. Y. Jul. 3, 2014).

不仅如此,美国还通过行业自律组织的最佳实践来巩固人工智能领域司法实践的成果。例如,美国图书馆研究协会在 2012 年发布《图书馆教育与研究目的合理使用的最佳实践》,强调图书馆将藏书扫描、形成数据库,并用于学术和参考目的的计算机信息分析,因具有高度转化性而属于合理使用,但同时要求该数据库不得用于个人阅读目的。^[32] 此外,该协会还在 2015 年联合世界上一百七十多个组织发布了《关于数字时代知识挖掘的海牙宣言》,呼吁世界各国的版权法应当允许人工智能和大数据分析中的作品复制和后续分析。^[33]

(二) 日本创设宽泛的计算机分析例外

在世界范围内,日本最早对人工智能所涉及的作品复制提供专门的著作权例外。在 2009 年修法时,《日本著作权法》新增第 47 条之七,允许为了计算机信息分析的目的——“从众多作品或者其他大量信息中筛选出构成该信息的语言、声音、影像或者其他要素,并进行比较、分类或者其他统计分析”——在必要限度内将作品存储在媒介上或者进行改编,但同时强调如果是为了个人阅读目的而制作作品的数据库则不得援引该例外。^[34] 如此可见,日本上述例外并没有限制使用的主体和使用目的,因此企业为了商业目的复制他人作品并用于计算机信息分析,仍然可以援引该例外。^[35] 不过,计算机分析的概念仍限定在“统计分析”范围内。因此,安倍政府认为现有著作权法的计算机信息分析概念并不能完全概括所有新型的人工智能所涉及的所有作品利用方式(如 AI 深度学习中超越统计分析而进行的代数或几何分析),仍可能限制日本人工智能产业的发展。^[36]

在 2018 年 1 月 22 日的著作权法修订中,日本用新的第 47 条之五替代原法第 47 条之七,对计算机分析的界定不再采用原有的“统计分析”的限定,转而选择参考美国的转化性理论,规定“为了提供新的知识或者信息”,开展信息处理的人(包括执行部分任务的人员)可以对他人作品进行必要限度内的复制和向公众提供,并对整理后的作品在实现此目的的必要限度内向公众提供,但同时强调该复制和向公众提供不得合理损害权利人的利益。^[37] 因此,该修订有两大进步之处:一是通过引入美国转化性使用理论扩大了原有比较局限的计算机信息分析范畴,未来不用频繁修法就可以更加灵活适应日本大力发展人工智能产业的需求;二是允许开展人工智能研究的主体可以将其复制他人作品而形成的数据库向公众提供。^[38]

(三) 欧洲相对局限的新例外尝试

英国在 2010 年注意到文本与数据挖掘技术对于版权法的挑战,^[39] 2013 年通过实证

[32] Association of Research Libraries, *Code of Best Practices in Fair Use for Academic and Research Libraries*, 2012, p. 25.

[33] *The Hague Declaration on Knowledge Discovery in the Digital Age*.

[34] 参见《十二国著作权法》翻译组:《十二国著作权法》,清华大学出版社 2011 年版,第 386 页。

[35] Pykov Viktor, *Big Data and IP Law: Risk Assessment and Fostering Data-Driven Transactions*, 2018, p. 61.

[36] 参见日本文部科学省:《著作権法の一部を改正する法律(案文・理由)》,平成 30 年,第 8-10 页。

[37] 参见日本文部科学省:《著作権法の一部を改正する法律(新旧对照条文)》,平成 30 年,第 10-12 页。

[38] 参见内閣官房内閣広報室:《著作権法の一部を改正する法律案:概要説明資料》,平成 30 年,第 7 页。

[39] Ian Hargreaves, *Digital Opportunity: A Review of Intellectual Property and Growth*, 2011, p. 51.

研究发现,“文本挖掘例外将使英国科研人员每年获利 1.24 亿英镑”。^[40] 因此,在 2014 年修订的《版权、设计及专利法》中,英国增加第 29A 条例外,允许科研人员仅为了非商业性的研究目的复制作品并对其进行计算机信息分析,版权人通过合同限制该复制和分析的条款无效,同时设置了两个条件:一是开展该研究的主体本身对相关作品享有合法的访问权,^[41] 不论是通过授权(如购买数据库的访问权,但通常购买数据库时的合同仅准许单个下载文件,不得批量下载)还是其他方式取得访问权;二是不得将复制生成的作品副本向他人提供。到目前为止,英国认为这一例外得到非常好的执行,尚未有争议提交到法院。^[42]

为了快速发展人工智能,欧盟委员会在 2016 年提出《单一数字市场版权指令草案》,在第 3 条规定成员国应当对作品和数据库的复制权提供例外,允许研究机构以科研目的对其原本可以合法访问的作品及数据库进行复制;任何与该例外规定相违背的合同条款无效,但版权人仍可对其作品或者数据库的安全和完整性采取必要措施。^[43] 应当强调的是,上述草案仍受限于欧盟《信息社会版权指令》第 5 条关于复制权例外仅限于非商业目的的限制。因此,草案一经公布,引起轩然大波,版权人普遍反对该立法草案,科研人员普遍支持该草案,但部分欧盟成员国官方、企业代表和学者认为这样的例外规定仍然不利于人工智能的发展。^[44] 其中,马克·普兰克创新与竞争研究所非常委婉地批评到,如果能够将商业目的的科研纳入到该例外中,“将强化欧盟企业对抗那些不受类似限制的他国企业的有利地位”。^[45] 经过两年的争论,欧盟议会终于在 2018 年 9 月 12 日原则通过上述指令草案,但提供成员国两种文本与数据挖掘例外的模式:一种是强制性例外,即保留欧盟委员会的提案第 3 条,但增加了两个可以适用该例外的主体,即教育机构和文化遗产机构;另一种是选择性例外,即允许成员国选择不限制适用文本与数据挖掘例外的主体,任何人在文本与数据挖掘过程中都可以对其合法访问的作品和数据库进行复制,但版权人明确排除使用者进行文本与数据挖掘的除外。^[46] 不过,欧盟议会新增的选择性例外最

[40] Intellectual Property Office of United Kingdom, *Modern Copyright: A Modern, Robust and Flexible Framework*, 2013, p. 37.

[41] 访问权,又被称为“获取权”,并不是严格的著作权法术语,只是指代作品使用权可以合法地接触、获取作品的副本,进而开展学习、阅读、欣赏和分析作品的工作。参见李雨峰:《论著作权的宪法基础》,《法商研究》2006 年第 4 期,第 114 页;冯晓青:《著作权合理使用制度之正当性研究》,《现代法学》2009 年第 4 期,第 33 页;彭学龙:《论著作权语境下的获取权》,《法商研究》2010 年第 4 期,第 116 - 124 页;梁志文:《论版权法上使用者利用的保护》,《法律科学》2013 年第 6 期,第 125 页。

[42] Eleonora Rosati, *Report to European Parliament: The Exception for Text and Data Mining in the Proposed Directive on Copyright in the Digital Single Market-Technical Aspects*, PE 604.942, 2018, p. 8.

[43] European Commission, *Proposal for Directive on Copyright in the Digital Single Market*, COM (2016) 593 final, p. 24.

[44] European Commission, *Impact Assessment on the Modern of EU Copyright Rules*, SWD (2016) 301 final, Part 1/3, pp. 119 - 122.

[45] Max Plank Institute for Innovation and Competition, *Position Statement on the Proposal Modernisation of European Copyright Rules*, Part B: Exceptions and Limitations (Art. 3—Text and Data Mining), 2017, para. 12.

[46] Amendments Adopted by the European Parliament on 12 September 2018 on the Proposal for a Directive on Copyright in the Digital Single Market. 在欧盟议会 2019 年 3 月 26 日一读投票中,新增的选择性条款变为正文第 4 条。See European Parliament legislative Resolution of 26 March 2019 on the proposal for a directive of the European Parliament and of the Council on copyright in the Digital Single Market [COM(2016)0593 - C8 - 0383/2016 - 2016/0280 (COD)]. 该一读结果仍须欧盟理事会批准,并在欧盟官方公报公布才最终生效。

终能否成为法律还需要欧盟理事会的批准,因而仍有待确定。

由此可见,英国和欧盟在文本与数据挖掘例外的设置上,虽然借鉴了日本 2009 年的计算机信息分析例外,但仍然受限于欧盟版权指令的非商业性使用要求,未能像日本那样将此例外扩展适用于商业目的的使用;也不像日本在 2018 年最新立法修订的那样,允许将文本与数据挖掘过程中复制他人作品而形成的数据库向公众提供。

欧洲、美国和日本都是人工智能技术发展较快的国家和地区,但三者因为不同的著作权法传统,在扫除人工智能第一环节作品合法使用的障碍时选择了不同的路径。这三种做法,恰如国际版权法大家保罗·戈德斯坦对世界各国的著作权限制与例外制度的评价,即如果从宽到严画出一个谱系,那么美国处于最宽松的这一端,合理使用(fair use)四要素原则就像美国经典动漫中的燥山姆,拿着左轮枪对兔八哥轮番打击;欧洲大陆法系国家则处于最严格的那一端,著作权的限制与例外就像《双城记》中的德伐日太太,将任何限制与例外都当作仇人的名字编织进她专门制作的织布中;而英国就处于中间,合理交易就像英国管家那样,唯唯诺诺。^[47] 值得注意的是,坚持著作权法体系的日本(一直未引入合理使用原则性条款),在面对人工智能发展前景时非常机智地借鉴美国合理使用制度的优点,用转化性理论来界定计算机信息分析的范畴,但同时又防止转化性理论这个“燥山姆”的狂暴,仅在计算机分析领域引入该理论。

四、增设人工智能合理使用例外

我国政府非常重视发展人工智能产业,但目前尚未明确要为此在著作权法上提供相应的法律环境。鉴于我国现有著作权法下人工智能第一环节的作品合法使用还存在不确定性,借鉴欧洲、美国和日本解决方案,本文认为应当在《著作权法》第三次修订之际做出相应的制度安排,其中专门制定单独的例外机制更为合适。

(一) 改变传统欣赏目的的作品使用方式

正如世界经济论坛创始人克劳斯·施瓦布教授(Klaus Schwab)所指出的,我们的社会已经进入了第四次工业革命,它具有信息处理上不同于传统线性发展的速度、规模和深度,并改变世界各国、各产业和社会整体的系统运行模式等特征,但考虑到能够连接个人的数量、处理能力、存储能力和知识获取能力的无限性,现有的人工智能等技术的发展还只是处于第四次工业革命的初级阶段。^[48] 针对其中关键的人工智能技术,世界各国都纷纷发布相应的发展战略。国务院 2017 年也发布了《新一代人工智能发展规划》,指出在移动互联网、大数据、超级计算、传感网、脑科学等新理论新技术以及经济社会发展强烈需求的共同驱动下,人工智能加速发展,呈现出深度学习、跨界融合、人机协同、群智开放、自主操控等新特征;人工智能也成为了国际经济的新焦点和经济发展的新引擎。面对人工

[47] Paul Goldstein, *The Americanization of Global Copyright Norms*, Lecture at Duke University School of Law, September 17, 2015.

[48] Klaus Schwab, *The Fourth Industrial Revolution*, Geneva: Currency, 2016, pp. 7-8.

智能发展面临的挑战和不确定性,国务院指出要在2020年“初步建立人工智能法律法规、伦理规范和政策体系”。

从著作权法的角度上而言,人工智能对于作品的利用有两大革新:一是改变了传统意义上以人类阅读和欣赏作品的使用方式,进入到了完全不同的计算机检索和分析的使用方式;二是这种使用并不是对于单个作品的利用,而是对于批量甚至海量作品的利用,旨在挖掘出信息安排上的规律和模式,甚至用于未来的预测。不过,在尚不承认人工智能自身具有主体性的框架下,这种作品使用的主体仍然是利用人工智能的个人和组织。

目前,正如前文所指出的,这种使用并未被我国现有著作权法的限制与例外所覆盖。而且,我国也无法像处理临时复制问题那样在著作权法中回避对此问题的规制。在作品数字化的初级阶段,人们利用在网络上发布的作品以及计算机软件都在计算机内存或者缓存中不可避免地产生一份新的作品复制件,个人在主观意志上并无法改变这一事实,所幸的是新的复制件也在使用者关闭计算机时自动消失。^[49]对于作品的这种临时复制,德国等部分国家明确承认其属于复制权规制的行为,但同时规定作品正常利用过程中必然发生的临时复制属于著作权的限制。^[50]我国也曾热议是否需要《著作权法》中加以明确以及规定的方式,但到目前为止都没有对于这种人类在正常使用数字化作品或者计算机软件时必然发生的技术措施列入著作权法的限制与例外当中。^[51]然而,人工智能第一环节对作品的复制并不是人类自身通过计算机正常利用作品所必不可少的技术措施,而是开展人工智能研究的主体必须主动实施的措施。

因此,我国还是无法回避人工智能第一环节作品复制的价值判断:一方面,将其判定为不构成合理使用,那么广大开展人工智能研究的个人与组织(尤其是企业)要合法地利用作品,就得为其中涉及的海量作品付费,或者将其中受保护的作品剔除出数据库,从而大大地加大我国本土企业发展人工智能的成本,限制我国人工智能的发展;另一方面,将其定性为合理使用,虽然能够为我国开展人工智能领域乃至第四次工业革命所涉及的产业中提供相对的竞争性优势,但同时应当将其对著作权人的不良影响降到最低。^[52]

(二) 单纯引入三步检测法或合理使用四要素尚无法提供确定性

我国著作权法很大程度上承袭了大陆法系国家的著作权法律制度和思想,同时也遵守《伯尔尼公约》和《与贸易有关的知识产权协定》(TRIPs)中的三步检测法要求,严格限制著作权的限制与例外。不过,面对日新月异的作品利用方式,我国司法实践中已经创制出超越《著作权法》第22条例外情形的新情形。因此,我国不少学者甚至立法机构都主张在第三次修法时,引入原则性的三步检测法或者合理使用四要素,明确法官可以在个案

[49] 参见王迁著:《知识产权法教程》,中国人民大学出版社2016年版,第127页;崔国斌著:《著作权法:原理与案例》,北京大学出版社2014年版,第384-389页;朱理:《临时复制是否属于著作权法意义上的复制》,《电子知识产权》2007年第1期,第22-25页。

[50] 《德国著作权法》第16条第1款和第44条a款。参见《十二国著作权法》翻译组:《十二国著作权法》,清华大学出版社2011年版,第160页。

[51] 参见《〈著作权法〉修订草案送审稿》第43条。

[52] Benjamin Sobel, Artificial Intelligence's Fair Use Crisis, 41 Colum. J. L. & Arts 45, 85 (2017).

中创设新的合理使用情形。^[53]

然而,三步检测法在《伯尔尼公约》仅明确“本公约成员国法律以及成员国之间现有或将要签订的特别协定”中可以设定具体而特定的著作权限制与例外,并非授权法官在个案中可以创设出著作权人难于预测的新型著作权限制与例外。^[54]而且,TRIPs 第 13 条也并未改变《伯尔尼公约》的这种框架,只是将其扩展适用于所有的著作财产权中。因此,面对美国广泛利用合理使用四要素创制新型例外,不少国际版权法学者认为美国的做法违反 TRIPs 的要求。^[55]即使摆脱三步检测法和美国合理使用四要素这方面的争议,按照前文提及的司法实践,在并非判例制的我国引入原则性的三步检测法或美国合理使用四要素,也无法为人工智能的作品合法使用提供可预测的法律确定性。当然,引入原则性条款产生的不确定性是日本和欧盟等大陆法系国家或地区都无法摆脱的问题。然而,在日本和欧盟都引入相应的例外机制的情况下,我国如果仍停留在这种权宜之计当中,那么我国企业在发展人工智能时要么背负侵权原罪的担忧,要么为此选择规避风险、复制那些不受著作权保护的内容或者进行事先许可,从而无法为我国企业在著作权法律制度方面提供同等的竞争环境。

(三) 引入著作权限制与例外

为了明确人工智能对作品利用的合法性、摆脱三步检测法或合理使用四要素原则性条款的权宜之计对企业产生的不确定性,我国应当引入单独的人工智能领域的著作权例外。作品的法律保护并非自然有之,相反,现代著作权法为作品提供法律保护恰恰是出于鼓励作者创作越来越多的作品,实现社会、经济和文化繁荣的公共政策。与此同时,如《新一代人工智能发展规划》所确认的,人工智能的发展也符合我国社会、经济和文化繁荣等方面的公共政策需求。而三步检测法又恰好是《伯尔尼公约》和 TRIPs 所允许的设定特定的著作权例外的机制。

因此,我国可以根据三步检测法设计人工智能领域的著作权例外。按照三步检测法第一步的“特定具体情况”要求,我们需要限定人工智能中作品使用的情形。否则,按照世界知识产权组织(WIPO)对三步检测法的指南,就没有必要进行第二、三步的分析。^[56]由于人工智能的概念本身具有不确定性,也无法直接体现出其与著作权之间的关系,日本、英国和欧盟在立法或者立法草案中都用新的概念来限定。其中,英国和欧盟所使用的文本与数据挖掘,很大程度上源自于欧盟和英国都规定了数据库的单独保护,所以更适合其本土的法律需要。日本并没有制定专门的数据库法,而将其符合独创性要求的数据库也按照作品来保护,所以日本的立法并不强调数据挖掘,而采用更中立的“计算机分析”,

[53] 参见《〈著作权法〉修订草案送审稿》第 43 条。学者代表性观点参见李明德、管育鹰、唐广良著:《〈著作权法〉专家建议稿说明》,法律出版社 2012 年版,第 260 页;李琛:《论我国著作权法修订中“合理使用”的立法技术》,《知识产权》2013 年第 1 期,第 17-18 页;卢海君:《论合理使用制度的立法模式》,《法商研究》2007 年第 3 期,第 28 页。

[54] 参见《伯尔尼公约》第 9 条第 2 款。

[55] 相关讨论的总结性分析参见 Pamela Samuelson and Kathryn Hashimoto, Is the U. S. Fair Use Doctrine Compatible with Berne and TRIPS Obligations, <https://ssrn.com/abstract=3228052>, 最近访问时间[2019-03-08]。

[56] WIPO, Guide to the Copyright and Related Rights Treaties, 2004, p. 57.

而且考虑到人工智能领域计算机分析不限于统计分析,在2018年修法时不再限定计算机分析是统计分析,而是引入了美国转化性理论将其界定为“为了提供新的知识或信息”而开展的计算机分析。因此,我国在借鉴时,更适合采纳日本的这种定义,尽量囊括未来人工智能发展中可能出现的各种计算机分析方式。应当强调的是,人工智能领域的发展和主导者是科研机构和企业,例如科技部将代表中国人工智能研究水平的百度、阿里巴巴、腾讯和科大讯飞纳入了首批国家新一代人工智能开放创新平台名单。为了充分发展人工智能,该例外的适用主体应当不限于非商业目的的科研机构,而应延及所有开展计算机信息分析的科研机构和企业。

按照三步检测法第二步的要求,人工智能领域的作品使用并“不得与作品的正常利用发生冲突”。WIPO根据1965年政府专家委员会起草的三步检测法文本指出,正常利用指的是规范性意义而非实证意义上的使用,“如果属于作品目前或者未来利用的形式,而且考虑到其重要性,该作品利用可能与作者行使其著作权产生经济上的竞争关系的,构成与‘作品的正常利用相冲突’(质言之,该行为破坏作者及其继承人在市场中的作品商业化利用)”。^[57]作者创作作品获得经济回报是因为人类为了阅读和欣赏作品而购买作品,而仅仅为非人类的“阅读和欣赏”目的的计算机的分析,并不会影响作者的作品正常商业化利用。即使是人工智能领域初级阶段将复制的作品提供人类进行信息定位和模式化分析,从而可能导致部分的片段式内容为人类阅读,也只在微小程度上影响到作者的商业性利用。^[58]例如,谷歌将数字化的图书以片段式提供与搜索者输入的关键词有关的部分内容,可能导致部分搜索者因搜索到这些关键信息之后不再购买该作品,但也可能吸引到此前并未关注到该作品的更多消费者购买该书,至于该片段的大小和提供方式是否不合理损害作者利益,则可以留到三步检测法的第三步分析。

按照三步检测法第三步的要求,人工智能领域的作品使用“不得不合理的损害作者的合法权益”。对于第三步要求,我国学者指出这是争议最多的一步,如果解释不当就容易与第二步要求重合。^[59]事实上,WIPO也坦承1967年的斯德哥尔摩修订大会并没有提出直接而具体的指引。^[60]不过,WIPO还是强调,1965年起草时是要求“不得损害作者的合法权益”,1967年会议讨论阶段英国提出要增加“不得不合理”的限定并最终获得接受,因而第三步的要求是审查作品使用行为对于作者合法权益的影响是否达到不合理的程度,而不是要求审查该使用是否得到社会规范和相关公共政策的支持。^[61]换言之,第三步是在第二步确定特定的作品使用行为并不与作者正常商业化利用产生直接竞争关系下,讨论这个经济影响的程度。人工智能中的作品利用如果复制后纯粹用于计算机的分析,并不会与作者的正常商业利用产生竞争关系,也不会对作者带来不利的经济收益。然而,如果将复制后的作品向公众提供,就会损害作者的经济利益,因此这种向公众提供必

[57] WIPO, Guide to the Copyright and Related Rights Treaties, 2004, p. 59.

[58] 参见内閣官房内閣広報室:《著作権法の一部を改正する法律案:概要説明資料》,平成30年,第7页。

[59] 代表性观点参见熊琦:《著作权合理使用司法认定标准释疑》,《法学》2018年第1期,第192页。

[60] WIPO, Guide to the Copyright and Related Rights Treaties, 2004, p. 59.

[61] WIPO, Guide to the Copyright and Related Rights Treaties, 2004, p. 60.

须限定在合理范围内。例如,仅仅允许提供作品的关键词或者与关键词相关的有限片段。或者,复制后向特定第三方提供,并且仅用于计算机软件本身或者算法本身的检测目的,而不用于个人的搜索和分析,也应当符合第三步的要求。例如,人工智能在发展过程中出现了算法黑箱和算法歧视的不良后果,公众和立法者希望能够检查或审计该算法的合理性。^[62] 由于算法本身是计算机软件保护的范畴,公开算法源代码会严重影响企业的商业利益,目前业界最简单的做法是检验算法输入端的内容是否出现盲点,因此将训练算法的输入端的内容保留下来并提供给第三方仅用于计算机层面的检验和验证目的,^[63] 也应当符合第三步的要求。

综上,本文提议在现有著作权法修订草案基础之上增订专门的计算机信息分析例外,即在《〈著作权法〉修订草案送审稿》第 43 条第 7 项限制之后新增一项著作权限制与例外:“通过计算机信息分析挖掘新的知识或者信息,复制已经合法访问的他人作品,并对复制产生的作品数据库在必要限度内向公众或特定第三方提供。”

[**Abstract**] Artificial intelligence is the latest technology that collects and manages a sea of data. It may also copy a large number of copyrighted works. The copying of these works, especially for commercial purpose, does not necessary fit into the current copyright limitations and exceptions, and therefore may have to be approved by copyright holders. As a result, enterprises must pay a large amount of fees for many such copies in the process of carrying out AI research. However, paying every single work copied in AI research not only is impossible to achieve, but also will crash even the leading IT companies in China. Meanwhile, the proposed general fair use clause would not provide predictable outcome for such copying. Given that the United States, Japan, and the European Union have already provided or are going to provide specific exceptions for such copying, it is necessary for China to also provide specific copyright limitations and exceptions for such copying, so as to clear the copyright obstacle for the development of the AI industry and gain a comparative competition advantage in the global market. However, the concrete application of these new exceptions should be scientifically defined so as to limit its application and prevent it from causing unreasonable harms to the legitimate interests of the author.

(责任编辑:姚 佳)

[62] 参见张吉豫:《人工智能良性创新发展的法制构建思考》,《中国法律评论》2018 年第 2 期,第 108 - 118 页;丁晓东:《算法与歧视》,《中外法学》2017 年第 6 期,第 1609 - 1623 页。

[63] Alex Salkever, Why Your Company May Need an Algorithmic Audit, <https://www.servicenow.com/workflow/algorithmic-audit.html>, 最近访问时间[2019-03-08]。