



## 智能审判技术下的责任主体研究

成梦露

中南财经政法大学刑事司法学院，武汉

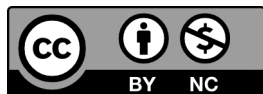
**摘 要** | 在人工智能介入司法裁判过程中，由于系统本身存在的技术限制以及其运行逻辑，人工智能系统应被定位为人类裁判的辅助工具。目前，人工智能作为独立的法律人格体的条件并不成熟。因此，在审判责任的承担主体方面，只能是人类承担，并且由人工智能的设计者与运营者来承担责任缺乏一定的理论支撑。案件的真正审理者是法官，法官在满足构成要件时应当承担相应的审判责任。

**关键词** | 人工智能；司法裁判；审判责任

Copyright © 2023 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



人工智能是对全人类的智能行为加以研究，把握其规律，并在此基础上构造出一整套系统，此系统在完成工作时具有一定的智慧能力。<sup>[1]</sup>自从人工智能被提出以来，随着科学技术的发展，人工智能在不断扩大，已经延伸到诸多专业领域。我国在司法领域也取得了一定的成效，例如北京法院的“睿法官”系统、河北省高院的智审系统、苏州法院的智慧审判模式以及上海法院的“206”办案系统等模式，为人工智能的司法应用提供了良好的经验指导。可以说，人工智能革命正在如火如荼地开展，但在给司法带来便利的同时，也催生了许多困难和挑战。因此，人工智能在司法领域的适用问题还应当进行慎重审视。

### 一、问题的提出

当人工智能系统介入到司法裁判中，如果系统在操作过程中产生了失灵的情况，或者是由于自身的原始算法存在错误，在案件的审判过程中向法官提出了错误的法律意见，这时该如何确定该案审判责任的承担主体呢？<sup>[2]</sup>审判责任主体的界定是人工智能系统在司法适用过程中不可忽视的关键

[1] [美] 卢克·多梅尔. 人工智能：改变世界，重建未来 [M]. 赛迪研究专家组，译. 北京：中信出版社，2016：1.

[2] 苗苗. 人工智能在司法裁判中的应用研究 [D]. 保定：河北大学，2021.

问题。

从国外实践来看,已经出现了这样的案例,例如美国威斯康星州诉艾瑞克·卢米斯案。在该案中,法官采纳了一款累犯风险评估系统(COMPAS)提供的裁判意见,但由于将评估过程公开存在着泄露商业秘密的风险,因此该系统的研发公司仅将最终评估结果提交给法庭。尽管该研发公司的服务从自身发展的角度考量具有一定合理性,但也给司法服务带来了挑战。在对卢米斯最终的评估结果中,报告显示卢米斯仍然具有再犯的可能性,法院依据这一风险提示对其进行定罪量刑。<sup>[1]</sup>在该案件的审理过程中,法院的裁判依靠人工智能系统,根据电脑评估报告做出判决。然而,在结案后发现该系统可能存在技术漏洞,并且找到了新的证据证明系统给出的裁决存在错误。此时,谁应为裁判错误承担责任成了不可忽视的问题。

人工智能系统介入司法裁判时,其技术功能总体可以归为下述三类。首先是为裁判提供技术支撑的功能,即人工智能可以在最基本的层次发挥作用,在此层面人工智能系统通过其数据库与技术向参与裁判的人员提供案件数据信息以及法律信息;其次是人工智能系统可以替代普通的技术人员与司法人员,之所以将人工智能的此类技术称为替代性技术,是因为人工智能系统的此类功能所执行的任务活动较为简单,并不具有极强的专业性和专属性,在这些活动层面,人工智能可以取代人类;人工智能的最后一项功能是对传统审判模式的颠覆,这类技术与司法人员之前的工作方式有很大的不同,在此层面,人工智能系统所提供的是一种崭新的裁判方法,这种裁判路径的形成不同于以往之经验,是一种全新的挑战。<sup>[2]</sup>前两类技术的共通性体现为为人类审理案件提供便利。而此三类技术中,讨论最多、争议最大的当属最后一类颠覆性技术。这也是本研究关注的重点。在最后一项技术中,人工智能借助一定的算法和程序已具有独立生成裁判结果的表象。适用人工智能系统生成对案件的裁判也属于颠覆性技术的范畴。因此,在法官采用智能审判技术下的裁判结果审理案件,在此情况下其责任主体何为,此乃亟需明确的问题。

## 二、人工智能司法判断能力的缺憾

### (一) 算法黑箱问题违反公平正义的内涵

在卢米斯一案中,报告人向审判法院提交了一份缓解动议以示其对裁判结果之拒绝,被告人卢米斯认为,自己对控告的相关信息具有知晓的权利,这是宪法第六修正案对公民所赋予的正当权利,然而在审判过程中,法院仅依据人工智能系统的评估测算结果作出判决,这种判决缺乏相应的控告理由侵犯了其知情权。卢米斯此项动议也引发了对算法黑箱问题的讨论,算法黑箱的产生背景极为复杂,其中最关键的两个因素如下,其一乃人工智能技术本身存在一定的复杂性,其研发和操作具有一定的专业性,技术壁垒难以破除;其二是研发公司以及相关利益公司的排他性,从保护知识产权以及维护自身发展的角度而言其推出的商业政策也无法做到将相关过程公开分享,以上问题的存在让人工智能的算法系统像“黑箱”一般存在着——用户无从知晓相关信息,又何谈对算法流程的评判分析与监管监督?<sup>[3]</sup>在司法裁判逐渐人工智能化的现实背景下,对案件作出裁决的过程,如同工业时代流水作业中加工制造产品一般程序化,司法裁判人员只消将案件的相关信息按照相应的程序进行录入,即可生成相应的裁判结果。而至于人工智能系统内部如何进行推理论断、价值选择与监督监管,这是裁判中所无法解决的问题。<sup>[4]</sup>由于算法黑箱的存在,在技术使用过程中,许多学者开始质疑人工智能是否具有应有的中立性。司法案件大数据透明性的缺乏和算法黑箱技术壁垒难以破除,这些问题的存在导致审判的公开性原则无从体现。而审判公开是维护法律之尊严权威中的必然要义,也是

[1] 朱体正. 人工智能辅助刑事裁判的不确定性风险及其防范: 美国威斯康星州诉卢米斯案的启示[J]. 浙江社会科学, 2018(6): 76-85, 157.

[2] 黄琪. 司法实践中人工智能应用的现状与定位思考[J]. 边缘法学论坛, 2019(2): 8-14.

[3] 王聪. “共同善”维度下的算法规制[J]. 法学, 2019(12): 66-77.

[4] 黄琪. 司法实践中人工智能应用的现状与定位思考[J]. 边缘法学论坛, 2019(2): 8-14.

实现司法公正不容忽视的典型范式。在司法裁判中,公平正义乃其核心价值所在。完全依赖人工智能系统进行裁判是对公平正义的忽视。因此,司法的核心价值决定了即使在裁判人工智能化的背景趋势下,也应当坚持人类裁判的本质内涵。

## (二) 个殊化判决背离自由裁量的本质

在卢米斯一案中,被告在动议中主张法官侵犯了其公民权利,其并未获得一份基于“个殊化判决(individualized sentence)”原则的而形成的裁判结果。<sup>[1]</sup>被告上述主张的理由在于人工智能系统在产出报告时所依赖数据为有限的特定群体,法院所作出的裁决并非基于被告人卢米斯自身的特定信息,而是基于群体信息的产物。

一方面,从推理方法这一角度进行对比,人工智能系统自动生成裁判和法官裁判的推理逻辑存在差异,演绎推理和归纳推理这两种推理方法存在很大的不同。在庭审中,法官对案件进行裁决时运用的是演绎推理,即通过已知的大前提和小前提,最终得出判决结论,这是一种纵向的思维模式;而人工智能技术运用的是归纳推理的推理方式,即法官在对案件进行审理时,系统会自动给法官推送与本案相似的一些案件以及相关的一些法律法规,将其进行分类整理后呈现在法官面前供其参考。换言之,人工智能系统是通过对比类案的大数据评估进而给出裁判意见,这是一种基于数据分析的横向思维模式,而非关注案件自身的纵向思维模式。这种推理方法也决定了人工智能系统在技术层面无法做到对被告人的个殊化考量。因此,完全依赖人工智能系统的裁判实则是法官忽视了个殊化判决的应有之义。

从另一方面分析,人工智能无法发挥自由裁量的判决功能。法院拥有自由裁量的权力,这是人类裁判所独有的特征。法律作为一门社会科学,其本身存在一定的抽象性与不确定性,加之社会快速的发展变动,使得法律应对纷繁芜杂的社会现实时呈现出一定的滞后性,这就需要法官充分行使自由裁量权对上述限制因素进行调节,以尽可能作出公正并且令人信服的裁决。而人工智能的运行原理在于,其是通过一定的算法逻辑加以运作,进而产出预期的相应结果。人工智能这种程序化的运行模式是刚性的,与自由裁量这

种柔性裁判模式存在一定的对立。法官在行使自由裁量权时考虑了诸多因素,这些因素包括但不限于裁判结果的社会效应、裁判过程中的价值衡量与取舍以及其他任何与之相关的环境等。<sup>[2]</sup>而人工智能裁判则是按照一种预先设定的程序模式,其对相关信息的选择也已被进行范式化的处理。虽然这种程序设定有可能达到形式正义,但却在实现实质正义层面存在瓶颈。缺乏自由裁量的人工智能,其所作出的裁决可能无法令人理解与信服,甚至可能存在出现激进或者不准确结果的情况。自由裁量乃审判中不可或缺的关键步骤,这是人类法官在审理案件时所拥有的无可比拟的优越性。而人工智能在审判过程中无法模拟法官的内心心证,因而不假思索地全然接受人工智能的裁判结果,是对法官自由裁量权的干扰和侵犯。

## 三、人工智能司法责任能力之否认

### (一) 人工智能并非独立的法律人格体

在法律上,人工智能介入司法裁判会导致审判责任主体不明确,对其原因进行深入分析,可以发现问题之症结在于,人工智能系统是否具有独立的法律人格地位尚未形成明确的定论。<sup>[3]</sup>换言之,要想让人工智能系统承担审判责任,其前提便是人工智能系统需以法律人格体的身份独立存在,这是人工智能承担审判责任的必要前提。承担法律责任不仅要求存在特定的违法或违约行为,还要求违法或违约行为的实施者乃适格主体。如果人工智能并非法律上的适格主体,那么在此谈论人工智能的责任问题便毫无意义。

对于人工智能系统是否具有法律人格?理论界与实务界存在诸多观点。有的国家所持有的态度比

[1] 朱体正. 人工智能辅助刑事裁判的不确定性风险及其防范: 美国威斯康星州诉卢米斯案的启示[J]. 浙江社会科学, 2018(6): 76-85, 157.

[2] 李叙明. 法官自由裁量权的伦理规制研究[J]. 湖南大学学报(社会科学版), 2013, 27(5): 111-115.

[3] 尹国金, 梅娟. 人工智能法律拟制人格问题研究[J]. 法制与经济, 2019(4): 20-21.

较超前,认为人工智能系统是可以介入到司法裁判中进而产生相应法律责任,如沙特阿拉伯政府赋予人工智能以公民的身份,人工智能系统在权利与义务层面与普通公民相当。<sup>[1]</sup>而多数国家对此采取了较为保守的做法,即并没有在法律上对人工智能系统的审判责任之主体地位予以明确。就我国当前的情况来看,关于人工智能系统介入司法裁判中导致的失职问题是否需要承担责任,法学界并未得出一致结论。<sup>[2]</sup>学者们对此问题的态度,概括而言主要分为以下三类。

(1) 否定由人工智能承担责任的学者主张,人工智能系统并非法律上的人格主体,创建人工智能系统的目的是为了让其更好地发挥对人类服务性的功能。正如季卫东教授所强调的,“大数据、云计算、信息技术、人工智能都只是实现合法正义的辅助手段,切不可本末倒置这是我们始终应该铭记的一条基本原则”<sup>[3]</sup>。即人工智能系统所起的只是辅助性裁判或者可替代性的服务功能,但该系统本身在法律上并不具有独立的人格地位,因而也不能独立承担审判责任。

(2) 支持方则认为,从理论上而言是可以赋予人工智能以法律上的人格。其理由在于人工智能的技术进展速度令人瞠目,人工智能发展到一定阶段就可以自主作出决定。其具有独立裁判的潜力,相应地也具有独立承担责任的能力。并且如果在法律上确认了人工智能系统具有主体资格,就可以进一步充分挖掘和利用人工智能系统所具有的多样化功能,可以更好地服务于社会、造福人类。<sup>[4]</sup>

(3) 第三种观点认为,人工智能系统是否应当承担责任应该视情况进行分类讨论。当案件的审判结果出现瑕疵或者错误时,应当根据人工智能系统所起的作用分而视之。如果案件是由弱人工智能进行辅助审判的,此时应由人类承担相应的责任。而如果案件是由强人工智能独自进行审判的,审判责任就由智能系统自身负责。

但有学者对此观点提出质疑,认为第三种方式视人工智能的强弱程度对审判责任之主体进行划分,虽然在形式上具有一定的合理性,但这种观点仍然存在问题。如果视情况进行讨论,那么案件的关联主体并不确定,对人工智能系统的处

罚也会缺乏明确性。<sup>[5]</sup>

能否将人工智能视为独立的法律主体并由其承担责任?这一问题早在2016年就得到了欧洲国家的关注。欧洲议会在其报告中指出,赋予人工智能以法律上的人格并由其承担责任是时代发展的必然选择,报告创新性地提出电子人格的概念,认为应当由功能复杂的机器人系统为已然损害负责。虽然在责任的认定上,此份报告已经走在时代的前列,但遗憾的是对责任的具体落实上,报告并没有提出具体且明确的操作方案。

承袭欧洲报告的思路,我们不妨大胆假设人工智能系统是一个独立的法律人格体,暂且让人工智能系统为其裁判错误负责。那么,人工智能系统又可以以怎样的方式承担民事、刑事责任呢?如果人工智能可以以合理的方式承担责任,那么智能审判模式的责任主体问题或可迎刃而解。但如果人工智能无法在现有的责任形式范围内接受约束,那将人工智能视为独立的法律主体并由其承担责任的观点或是一种无法实现的美好愿景。本文接下来将对此进行分析,探明究竟。

首先,无法让人工智能系统承担赔礼道歉的民事责任和自由刑的刑事责任,因为此类责任具有人身专属性,对人工智能系统并不适用。赔礼道歉是一种侵权责任形式,这源于道德责任的要求,其目的在于通过行为人对错误行为的产生愧疚、表达歉意,以缓解被害人的精神痛苦。但人工智能却无法产生真实情感的悔过之心,赔礼道歉实乃荒谬。而对人工智能系统施以自由刑不仅无法发挥刑罚的威慑作用,也无助于对人工智能系统的后期规制和约束。

其次,就赔偿责任而言,无论是对人工智能系

[1] 郑戈. 人工智能与法律的未来[J]. 探索与争鸣, 2017(10): 78-84.

[2] 苗苗. 人工智能在司法裁判中的应用研究[D]. 保定: 河北大学, 2021.

[3] 季卫东. 司法人工智能不可本末倒置[N]. 北京日报, 2019-09-16.

[4] 姚万勤. 人工智能影响现行法律制度前瞻[N]. 人民法院报, 2017-11-25(2).

[5] 程凡卿. 我国司法人工智能建设的问题与应对[J]. 东方法学, 2018(3): 119-130.



统施以罚金刑——由其支付赔偿金，还是采纳最为传统的“缴出赔偿”原则——即把人工智能系统的所有权缴出，由受害者或其家属对其进行处置，上述责任承担形式都无法适用于人工智能裁判系统。因为对于前者而言，人工智能系统首先自身并无财产，更何谈财产的支配权，让人工智能承担罚金刑的责任就有着进一步转嫁风险。对于后者而言，暂且不论人工智能的权力归属问题，即使能把人工智能系统交由受害方处置，这对受害方也并无多大的实质性意义。

最后，是否可以在此参照有关学者的提议，将人工智能系统类比法人从而让其承担审判责任呢？笔者认为，无论采用法人拟制说还是法人实在说，都需要关注到法人是具有自己独立的意志，并且法人行为来源于人的组织体的意志。但人工智能系统是否也具有独立的意志是一个值得思索的话题。笔者认为，人工智能系统所采取的是一种既定的程序模式，这和人的自主的思维、思想存在本质的差异，人工智能系统作出裁判的行为并不直接来源于人的意志，其是将人类意志进行程序化设定的结果，故而将人工智能系统类比法人承担责任的设想并不妥当。<sup>[1]</sup>

#### （二）人工智能无法承受司法责任之重

人工智能介入司法裁判，严格来说，其更像是一种技术层面的裁决符号，而非法律意义上的司法活动。虽然人工智能系统有着独立生成裁判结果的表象，但其裁判的依据是对案件信息进行转化后呈现的数据代码，从一定程度上而言，人工智能系统的程序应用本质上是对代码、数据等符号的转化处理，而非以司法视角裁判案件、定纷止争。在对案件进行裁判的过程中，只有人类法官才能识别出纠纷所涉及的事实与法律问题，并在把握司法内在精神基础上对裁判结果权衡利弊，以尽可能地制作出一份兼顾法律与社会效果的裁决，而不是机械地制作出一份毫无人文情怀的法律文书。

人工智能系统无法承担司法责任的重担，由系统本身担责对于维护社会公平正义、稳定社会秩序并没有实质上的作用。让人工智能系统对错误的司法裁判负责，这与当事人的预期相去甚远，在缓解当事人对立情绪、修复司法裂痕上收效甚微。在某种程度上，人工智能裁判不仅不会

更好的定纷止争，反而会产生意想不到的负面效果。<sup>[2]</sup>因此，人工智能应被定位为人类裁判的辅助工具。

上述分析证明人工智能作为独立的法律人格体的条件并不成熟。人工智能在司法应用过程中承担责任的主体只能是人类。而且人工智能审判系统既无可惩罚的肉体也无可惩罚的灵魂，容易引发责任的真空，从而架空司法权的权威性与公信力。<sup>[3]</sup>无论对人工智能系统施以何种形式的责任，最终的责任承担者都是人类本身，从这种角度而言，人工智能系统的“法律人格”则显得多余和毫无必要。<sup>[4]</sup>

### 四、人工智能的审判责任主体——法官

#### （一）人工智能设计者不能作为责任承担的主体

人工智能在司法应用过程中承担责任的主体只能是人，但是否意味着人工智能的设计者也需要承担责任？笔者以为答案是否定的。原因如下。

首先，人工智能的设计者和运营者在设计之初、运营之时是一种价值中立的态度。尽管人工智能系统的智库建设过程中不可避免的存在对一些数据的取舍和筛选，但设计者对数据的筛选是基于客观因素的考量，而非针对具体案件中特定的主体。人工智能的设计者并不存在对未来特定案件的预测，其对案件的审理也不存在价值倾向性问题。因此，设计者和运营者在主观上并不具有错误裁判和非法裁判的意图，这是否定其责任的原因之一。

其次，人工智能技术在应用过程中存在一定的风险，但这种风险应当被社会所允许。如果让人工

[1] 郭纪萍：论人工智能刑事主体地位确立之理论建构与路径设想[J]，互联网天地，2020（5）：8-13。

[2] 黄琪：司法实践中人工智能应用的现状与定位思考[J]，边缘法学论坛，2019（2）：8-14。

[3] 韩世鹏，秦勇：人工智能司法的法理逻辑与决策限度[J]，淮阴工学院学报，2022，31（2）：40-45。

[4] 郑戈：人工智能与法律的未来[J]，探索与争鸣，2017（10）：78-84。

智能的设计者和运用者承担严格责任,这种责任分配对于设计主体和运营主体过于严苛,会在一定程度上阻碍技术的进步。如果人工智能的设计者不是别有用心地恶意利用人工智能进行错误裁判,那么让其为正常的技术研发活动承担责任无疑为科技的发展进步设置了巨大的障碍。我们需要的是最小化责任的技术,不宜过于扩大系统设计者和使用者的责任。<sup>[1]</sup>

## (二) 法官承担审判责任之理论依据

通常意义上的审判可以进一步划分为审理与裁判两部分,审判权也应当包括以下两种权力,一是对案件进行审理的权力,二乃对案件进行裁判的权力。<sup>[2]</sup>因此,“让审理者裁判,由裁判者负责”。在人工智能系统介入司法裁判的过程中,案件真正的审判主体实际上是法官,因而承担责任的主体也应当是法官。原因在于:首先,在对证据的识别应用上,人工智能系统无法判断证据的证明力大小,也不具备对证据加以取舍的能力。相比之下,法官在对证据的界定上可以充分发挥人类智慧,行使自由裁量权对证据进行界定。其次,人工智能难以复制人类法官的审判经验。审判经验是法官在长期的司法工作中、基于大量的实务案件形成的知识积累。纵使人工智能可以通过归纳推理为裁判提供一定的数据支撑,但这与人类法官基于经验而积累的人类知识有很大的不同,人工智能系统在审判过程中缺乏法官所拥有的灵活性,案件真正的审理者是法官,人工智能在司法审判中的准确性应为人类裁判之辅助。

## (三) 审判责任的构成要件

在人工智能介入司法裁判时,法官承担审判责任需要进一步明确构成要件,这可以起到限制审判责任范围的作用,其乃责任承担的核心要素。所谓构成要件,是指在何种前提条件下可以发生相应的法律效果,也即法官在满足何种条件时才需要承担审判责任。责任范畴的内涵包括主客观两方面要素,在司法裁判中,审判责任之构成要件也需从以上两方面加以考虑。

审判责任的客观方面指的是司法裁判人员违反法律规定或导致裁判错误的客观行为。具体而言,是指法官在审判过程中所对应的违规裁判和违法裁判的行为,例如输入错误的身份信息(如年龄、精

神状况)、案件信息(与定罪量刑有关的案件事实和证据)等信息,过分依赖人工智能系统生成的参考意见而完全忽视审判常识,或者对人工智能系统所提供的一些重要因素视而不见等等。审判责任的主观方面,则是指审判者在审判时的主观心态乃故意或者过失,即明知其审判行为的性质及可能产生的后果,但仍然对此采取希望或者放任的态度,或者因其疏忽而对错误的审判行为听之任之。<sup>[2]</sup>只有在审判人员同时满足上述两种条件,才可以要求其承担审判责任。

## 五、小结

在卢米斯案件中,应用人工智能系统对被告人进行定罪量刑。其量刑前的调查报告需要包含5项书面警示。这些警示的发布目的是为了通过明确的书面报告,向审判案件的法官传递正确的理念——时刻保持怀疑。即使拥有强大数据和高超算法的人工智能系统介入司法裁判,也要对其准确性保持普遍的怀疑态度,尤其是人工智能系统通过大数据分析,其所关注的乃定罪与量刑的重点与普遍因素,不可避免的存在着对特殊情况的忽视。法官作为审判责任的承担主体,应当始终坚持人工智能辅助介入司法裁判时的工具性价值定位<sup>[3]</sup>无论人工智能系统的技术发展到何种程度,服膺于司法裁判的本质规律始终是不变的价值准则。因此,在司法裁判中,应尊重法官的职业裁断,并由法官为错误的司法裁判负责。

(责任编辑:郭志姣)

[1] 马长山. 司法人工智能的重塑效应及其限度[J]. 法学研究, 2020, 42(4): 23-40.

[2] 陈卫东. 司法责任制改革研究[J]. 法学杂志, 2017, 38(8): 31-41.

[3] 王晓飞. 对人工智能辅助司法裁判的冷思考[J]. 武汉交通职业学院学报, 2019, 21(4): 28-32, 73.

## Research on Responsible Subjects under Intelligent Trial Technology

Cheng Menglu

*School of Criminal Justice, Zhongnan University of Economics and Law, Wuhan*

**Abstract:** Artificial intelligence in the process of intervening in judicial adjudication, due to the technical limitations of the system itself and its operating logic, the artificial intelligence system should be positioned as an auxiliary tool for human adjudication. The conditions for artificial intelligence as an independent legal personality are not mature, the subject of trial responsibility can only be humans, and the designer and operator of artificial intelligence bear responsibility lack of certain theoretical support. The real adjudicator of the case is the judge, and the judge should bear the corresponding trial responsibility when meeting the constituent elements.

**Key words:** Artificial intelligence; Judicial adjudication; Trial responsibility