

人工智能发展的宪法限度

吴璇

扬州大学法学院, 江苏 扬州

收稿日期: 2024年11月7日; 录用日期: 2024年11月19日; 发布日期: 2024年12月23日

摘要

人工智能发展应当给予一定限制, 无序的发展可能会使人工智能走向消失的命途。人工智能研发者有失公正的输入数据行为系“算法歧视”问题产生的根本原因, 衍生而来的“价格歧视”“种族歧视”“性别歧视”问题成为实现全人类平等目标的拦路虎; 无序的人工智能发展对自由、安全等基本价值理念造成巨大冲击, 宪法基本价值根基有所动摇; 侵权认定难现象时有发生, 人工智能成为侵权主体逃避法律责任的“避风港”。由于人工智能发展历史与宪法发展历史存在基本对应关系, 宪法系当下解决人工智能失控所造成的法律问题的最佳手段。因此, 应当将保障人的尊严这一宪法共识贯穿人工智能发展始终, 使人工智能发展符合宪法发展目的——保障人类的主体性地位, 在人工智能研发过程中坚持技术开发类型化的宪法基本原则, 最终形成和平有序的宪法秩序来确保人工智能的有序发展。

关键词

人工智能, 算法歧视, 人的尊严, 宪法共识

Constitutional Limits of AI Development

Xuan Wu

School of Law, Yangzhou University, Yangzhou Jiangsu

Received: Nov. 7th, 2024; accepted: Nov. 19th, 2024; published: Dec. 23rd, 2024

Abstract

The development of artificial intelligence should be given to certain restrictions, and the disorderly development may make artificial intelligence disappear. The unfair input data behavior of AI developers is the root cause of “algorithm discrimination”, and the derived “price discrimination”, “racial discrimination” and “gender discrimination” become the obstacle to the realization of equality; the disorderly development of AI has a huge impact on the basic values such as freedom and safety, and the foundation of the Constitution is shaken, and artificial intelligence becomes the “safe haven” for the infringing subject to evade legal responsibility. Due to the basic correspondence between the development history of artificial intelligence and the development history of the Constitution, the

文章引用: 吴璇. 人工智能发展的宪法限度[J]. 法学, 2024, 12(12): 7132-7141.

DOI: 10.12677/ojls.2024.12121013

Constitution is the best means to solve the legal problems caused by the loss of control of artificial intelligence. Therefore, should safeguard the dignity of the constitutional consensus throughout the artificial intelligence development, artificial intelligence development in the constitutional purpose—safeguard the status of human subjectivity, in the process of artificial intelligence development type of basic principles of the constitution, eventually form a peaceful and orderly constitutional order to ensure the orderly development of artificial intelligence.

Keywords

Artificial Intelligence, Algorithmic Discrimination, Human Dignity, Constitutional Consensus

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目前人工智能行业发展迅猛，智能系统已经渗透进人类生活的方方面面，成为人类生活不可或缺的一部分。但是在赞赏人工智能给人类生活带来“质”的飞跃的同时，人类需要思考人工智能的发展是否具有不利的一面，警惕陷入盲目崇拜人工智能的漩涡，尽量降低人工智能发展过程中产生的非理性风险。

2. 人工智能发展的概述

2.1. 人工智能的定义

科技是决定国与国之间竞争成败的关键，随着第四次工业革命的到来，人工智能成为引领未来科技发展的关键所在。由于技术的跨越式突破和大规模的普及，人工智能已经渗透入人类生活的方方面面，可以说人类的高质量生活已经离不开人工智能的保障。“人工智能”该词初次进入大众视野是源自 2016 年人工智能机器人 AlphaGo 战胜韩国围棋大师李世石这一轰动世界的事件[1]。而当时世界排名第一的中国棋手柯洁对阵升级版的 AlphaGo 三盘皆输的结果更是将大众对人工智能的讨论推上高潮。虽然人工智能已经成为社会的热点话题，但是并没有对其形成统一的定义，学界对人工智能到底是什么议论纷纷。

人工智能在不同的时代有不同的含义，其含义的变化蕴含着科技进步与发展的结果。最早提出“人工智能”一词的计算机科学家约翰·麦肯锡并没有将人工智能限制在人类智力的复制与粘贴上，而是将人工智能扩展到其他智能领域，其主张：“制造智能机器的科学和工程，特别是计算机程序，它与利用计算机来理解人类智能的类似任务有关，但不必自我限制于生物学上可观察的方法”[1]。此时学者对人工智能的研究涉及面窄，多集中于特定领域的某些问题。斯图尔德·罗素和彼得·诺维格在他们的专著《人工智能：一种现代的方法》中强调人工智能的内涵有如下四个方面：像人一样思考、像人一样行为、合理地思考和合理地行动[1]。中国在《人工智能标准化白皮书(2018 版)》从人工智能的实质出发，将人工智能定义为：利用数字计算或数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及其应用系统¹。尽管不同的学者和国家对人工智能的定义不同，但这并不妨碍其对人工智能的研究。目前，人工智能系指研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学[2]。

¹ 中国电子技术标准化研究院：《人工智能标准化白皮书》(2018 年版)，载中国电子技术标准化研究院网 2018 年 1 月 24 日，<http://www.cesi.cn/201801/3545.html>。

2.2. 人工智能的本质与核心特征

数据和算法是人工智能运转的核心基础，其中数据是人工智能分析的资料来源，而算法是人工智能高效运转的关键，换言之算法系人工智能之本质。算法作为社会与技术相结合的技术网络，其有如下三大特征。首先，算法具有有限自主性。“自主性”是指算法可以不依赖人类而做出决策。但算法只能在特定的场景中做出自己的选择，若超出该范围，算法将丧失其“自主性”，即算法的“自主性”是有限的，其自主选择实现有赖于场景、技术等因素，不存在完全自主的情形²。其次，算法具有不透明性，这也是产生“算法黑箱”问题的根本原因。“深度学习”是人工智能在发展过程中形成的一大特征，此时人工智能通过大量训练即可使算法系统对算法主体下达指令进行决策，其中判断过程体现了算法的不透明性。算法研发者虽然可以控制输入程序的数据，但是其很难解释算法的决策系统是如何运行的，那么在此基础上就形成由海量数据组成的异常复杂的“算法黑箱”。最后，算法具有难以归责性。由于算法的不透明性和有限自主性，人工智能使用者难以预测算法产生的危害后果。此外，算法主体的法律地位不明确，这使得传统的规则模式无法用来规制算法主体。当因算法运行而产生危害后果时，算法主体极有可能以危害后果是由算法自主作出的行为产生为由主张免责，这很大程度上妨碍对被侵权者合法权益的保护[3]。

作为人工智能核心特征的机器学习(Machine Learning)是一门涉及统计学、计算机科学、神经网络等诸多领域的交叉学科，其研究计算机是如何模拟以及实现人类的学习行为来获取新的知识或技能并不断改善自身的性能³。人工智能学习包括学习训练和模拟应用两大阶段。学习训练是指由算法研发者输入指令经人工智能深度学习后输出数据的过程，而模拟应用系指人工智能在对输出的数据进行分析与预测的基础上进行判断与决策[1]。不难发现，由于人工智能学习的数据是由算法研发者输入，那么所输入的数据难免带有数据输入者自身的观点与偏见，因此很难保证人工智能所作出选择的正确性，这也是“算法歧视”产生的原因之一。

2.3. 人工智能发展历史概述

“人工智能”这一概念最早由“人工智能之父”约翰·麦肯锡及其同事于1956年在达特茅斯大学举办的计算机技术“侃谈会”上提出，至此人工智能这一学科正式成立，众多科学家在该领域投入大量时间和精力来研究这一新兴话题[3]。在这一时期，由于人工智能刚刚诞生，科学家们对人工智能的研究集中于基于抽象数学推理的可编程数字计算机的问题求解和逻辑推理。然而，在研究过程中科学家们发现很多事物不能进行形式化表达，因此计算机建立的模型存在一定的局限性。与此同时，伴随着计算机研究对象的复杂化，人工智能的发展遭遇挑战。这一时期大致处于20世纪50年代至80年代，是人工智能的初步发展时期。

第二阶段是从20世纪80年代至20世纪90年代末，这是人工智能的实际应用阶段。在这一阶段，专家系统迅猛发展，计算机软件得到重大突破，人工智能逐渐从科学家冰冷的实验室走入多彩的人类生活中，成为提升人类生活质量的重要工具。然而，由于专家系统在知识获取、逻辑推理等方面存在一定局限性，以及人工智能开发成本高等原因，人工智能研究又一次进入低谷期[3]。

自21世纪以来，伴随着算法理论的更新和计算机技术的发展，人工智能发展迎来黄金时代，“人工智能时代”即将到来。在这一阶段，不断涌现人工智能新成果，如“AlphaGo”“Siri”“无人驾驶汽车”以及最近很火的“ChatGPT”等，这些新型人工智能极大地提升人类生活的满意度和舒适度。随着人工智能技术的深入研究，未来一定会出现更多适用于提升人类获得感和幸福感的智能机器人，个人和智能将会深度结合，打造专属于人类自身的人工智能不再是梦想。

²同1。

³同1。

3. 人工智能发展过程中产生的法律问题

斯蒂芬·霍金曾说过：“人工智能对人类来说，可能是好事，也可能是坏事”[2]。如同一个硬币有两面，人工智能发展也具有两面性。一方面，人工智能的发展大幅度提升人类生活质量，进一步促进科学技术的发展进步；另一方面，人工智能发展极大影响和改变人类生活，动摇传统法律基础，其在发展过程中产生了众多法律问题。

3.1. 人工智能发展造成“算法歧视”

由于大数据、机器学习、算法理论的发展，人工智能越来越被应用于社会生活的方方面面，其中许多智能判断与决策均是在算法运行下进行的。由于作为人工智能本质的算法是依赖计算机网络、算法理论等技术而正常运行，所以很多人认为基于算法所作出的判断与决策具有正确性与客观性。然而，算法运行前的输入数据行为不可避免带有算法开发者自身的观点或偏见，也无法保证数据来源本身的正确性，这直接会影响之后经过算法分析所得结果的客观性，“算法歧视”问题就此产生。

在算法背景下所形成的“算法歧视”主要有三种表现形式，即价格歧视、种族歧视和性别歧视。价格歧视是指算法对基于大数据所收集的各类消费者信息进行分析与判断进而了解消费者的消费习惯，并最终根据消费者的消费习惯制定针对性价格[3]。表面上基于算法所形成的价格似乎可以满足各类消费者的消费需求，形成特定化服务，但是本质上无论消费者社会地位、个人收入、家庭背景有多大差异，他们都是商品或服务的消耗者，本应享有同等消费地位，如价格平等、服务平等。但是，“算法价格”客观上造成各类消费者的不平等，不符合社会公平的要求，实质上系“价格歧视”。种族歧视是指算法运行所产生的结果带有种族偏见或种族诋毁色彩。由于算法运行结果有时会体现算法开发者的个人意志，如果算法研发者本人对某一民族具有偏见，很可能在算法运行结果中出现具有种族歧视性观点，这不利于世界各民族和谐相处，破坏了同一地球村上各种族的感情。例如，Northpointe 公司曾推出 COMPAS 软件，该软件是一款预测被告人再犯可能性的犯罪管理软件，其曾预测黑人再犯可能性远高于白人再犯可能性，这种采用算法系统预测再犯可能性所得出的结论并没有科学依据，带有明显的种族歧视色彩，不应该被采纳和使用[4]。性别歧视系指算法运行结果在男女之间存在较大差异，此种歧视主要发生在职场领域。卡内基·梅隆大学和 ICSI 的研究人员使用名为 AdFisher 的软件来模拟互联网用户的浏览活动，但当浏览招聘网站时却发现男性用户会比女性用户有更多机会发现更多高薪工作[3]。原本男女用户有均等机会浏览同等数量的载有高薪工作的界面，但在算法影响下，男性用户却比女性用户更有机会接触到更高层次的工作，这本质上就是“算法歧视”所带来的“性别歧视”结果。

3.2. 人工智能发展对宪法所追求价值造成冲击

人工智能除了造成“算法歧视”，还会对宪法所追求价值造成冲击，如自由、安全等。冲击所带来的危害是无法计量的，因此我们必须重视人工智能发展过程中价值扭曲的现象。

3.2.1. 人工智能对自由的冲击

匈牙利诗人裴多菲·山陀尔曾在《自由与爱情》中写道：“生命诚可贵，爱情价更高，若为自由故，两者皆可抛”[1]。这表明裴多菲·山陀尔认为自由比生命与爱情更为重要，其能够为自由而放弃自己的生命与爱情，表达了诗人对自由的极度渴望与追求之情。在人工智能迅猛发展的今天，作为宪法所追求价值之一的自由遭受着人工智能的巨大冲击与挑战。

在行动自由方面，人工智能通过对海量数据进行分类、归纳与整理能够进一步了解用户的各类习惯与爱好，从而实现定向推送。表面上人工智能可以满足用户的各种需求，实现精准供给，但实际上却阻

断用户接触其他信息的渠道，是限制人民大众进行自由选择的表现。以短视频平台为例，用户应当享有自由选择视频类型的权利，但是由于用户喜好各异，其会根据不同的视频停留不同的时间。此外，点赞、收藏、评论的功能也会为短视频后台收集用户数据，后台智能系统通过分析判断能够获得平台用户的视频喜好从而向其推送感兴趣的视频，由此增加用户的活跃度与黏性[3]。这种用户在浏览视频时所产生的“根本停不下来”的感觉实际上系短视频平台智能系统运行的结果。

在意志自由方面，搜索引擎提供者使用算法系统向用户提供搜索排名，但并不是所有内容均会出现在搜索结果中，有些内容会被优先展示，而有些内容却无法显示，这实际上系人工智能限制搜索内容提供者言论自由的表现[5]。此时不禁产生疑问：为什么不同内容提供者之间会产生如此巨大的差异呢？其实可以用一个词语解释，即“利益”。在商业、技术、政治等利益的驱使下，搜索引擎提供者会利用算法系统将其有利或者给予其利益的搜索内容提供者的言论放置于搜索结果前列，这无形中侵害了其他言论提供者的言论自由权。搜索引擎成为有钱人对普通民众进行思想教化的工具，只要愿意花钱，无论言论内容是否与搜索关键词相关，均可以出现在搜索结果前列位置。如果在用户搜索列表中长期出现这种带有明显偏向性的搜索内容，用户思想将会被慢慢侵蚀，人的知识亦会逐渐碎片化。

3.2.2. 人工智能对安全的冲击

随着信息网络的迅速发展，对安全的威胁主要表现为数据泄露所带来的危害。如今，数据泄露现象屡见不鲜，其中个人隐私数据泄露最为普遍。伴随着人工智能的推广，人类仿佛生活在聚光灯下，随时可能发生个人隐私被泄露的情况。比如，当手机用户安装某一软件时，该软件会在用户界面中发出载有“我同意”和“我拒绝”字样的提示窗口来询问是否允许其访问软件使用者的通讯录、日历以及是否开启定位服务和同意终端用户许可协议等问题，如果软件使用者点击“我拒绝”的按钮，其可能面临着不能正常使用该软件的风险[3]。因此，为了迅速体验完整的软件服务，大多数软件使用者都会同意以上选项，这无形中为软件开发者输送大量的个人隐私数据。此外，由于人工智能拥有强大的数据分析功能，当下许多人都会将包含个人隐私的数据输入人工智能系统中使其为自己提供最佳选择。一方面，采用人工智能系统对事物进行分析会降低出现分析错误情形的概率，进而提高选择效率与正确度；另一方面，人工智能在分析人工智能使用者所提供的包含个人隐私的数据过程中会获取大量个人信息，原本系数据所有者的我们却无法掌握自身所提供的数据，降级成为众多数据提供者中的一员，这无疑增加数据被泄露甚至被不法分子利用的风险。

3.2.3. 人工智能发展造成侵权责任认定困难

虽然现行法律没有承认人工智能的民事主体地位，但其所造成的侵权问题却是实际存在的，因此如何确定侵权责任主体和分配侵权责任是解决人工智能侵权问题的关键所在。在传统侵权责任体系中，以过错责任为法律责任的主要分配方式，无过错责任作为过错责任的补充。对于人类本身来说，确定一个人是否具有过错是相对容易的，可以通过主观方面、行为方式、使用的工具、事后态度等因素进行综合判断，但是对于拥有有限自主意识的人工智能来说判断其是否具有过错十分困难。以无人驾驶为例，汽车的正常行驶不再依靠驾驶员的谨慎操作，而是通过自动驾驶的智能系统即可轻松实现。然而，此种驾驶方式具有一定的风险，如果其中一个环节出错，很可能会引发“多米诺骨牌”效应，牵动其他环节相继出现错误，从而导致交通事故的发生，造成他人人身、财产的损失，侵权行为就此产生。此时汽车的控制主体和对汽车行驶过程中突发情况的判断与决策主体从传统的人类转变为人工智能，但当前立法存在空白，“无人驾驶”汽车的所有权人可能以此为由推脱自身责任，那么人工智能就成为所有权人逃避法律追责的“避风港”[2]。由于确定侵权责任主体存在困难，进而无法对侵权责任进行合理分配，被侵权人的合法权益难以得到及时救济。

4. 宪法规制人工智能发展的正当性

“规制”是指将事物限定在一定范围内使其保持正常状态。任何事物发展都需要设定界限，如果没有对其加以一定限制，任其随意发展，该事物可能会走向失控甚至消失的道路。但是并不是所有事物都需要法律对其加以限制，有时仅仅通过社会伦理道德、行业习惯就可以实现事物的正常发展。然而，如果一个事物在发展的过程中引起了社会恐慌，造成了一系列法律问题，此时我们必须思考是否可以通过法律对事物进行规制。如今，人工智能发展迅猛，算法事业蒸蒸日上，但是不可忽视的是人工智能发展过程中产生的一系列法律问题，其对宪法所追求的价值、宪法本身发展提出挑战，造成“算法歧视”和侵权认定困难等法律难题，因此人工智能的发展应当受到法律规制，尤其是宪法规制。

4.1. 法律规制人工智能发展的正当性

4.1.1. 法律与人工智能关系的法理证明

人工智能本质上是科技发展过程中产生的一门新技术，法律与人工智能的关系即为法律与科技的关系，这是法理学学者长期关注的话题。当代法理学不应仅关注法律与科技之理论关系，更应当将研究视角转向科技发展带来的非理性后果以及如何通过法治降低科技发展可能带来的风险[6]。

首先，应当明确人工智能的发展可能会对法律造成威胁。原本人工智能作为新兴技术的一种，应当具有客观性。然而，人工智能是由人类创造，其运行系统也是由人类设计，因此人工智能无法完全脱离人类控制而正常运行，换言之人工智能是人类智力的延伸。这种人工智能客观性与人类智力的延伸的交界之处所产生的模糊地带难以被法律规制，现行法律可能缺乏对人工智能的明确规定或者在某些方面存在立法空白，这实际上冲击了法律的权威性与稳定性[6]。

其次，警惕人工智能发展可能产生法律问题。没有完美的科技，只有在创造风险和规避风险中完成自我革新的新技术。科技发展具有一定风险性，人工智能亦是如此。人工智能发展有利于人类生活质量的提高和社会进步，但是必须警惕人工智能发展可能带来的法律问题，“算法歧视”“价值冲击”“宪法挑战”“侵权认定难”等一系列法律难题均有待解决。不能神化人工智能，必须认清人工智能发展的双面性，针对人工智能的不利面进行法律规制，确保人工智能的有序发展。

最后，认清法律为人工智能设定界限本质上是对人工智能发展的肯定和保护[7]。任何事物发展都需要设定界限，限制的目的不是限制本身，而是通过限制能够使事物保持有序的发展状态。如果没有对事物发展加以一定限制，任其随意发展，该事物可能会走向失控甚至消失的道路。作为新兴事物的人工智能，其对人民生活、社会管理乃至国家体制均产生积极推动作用，助力国家发展强大，但是无法避免的是人工智能在发展过程中产生了一系列法律问题，人工智能发展界限模糊，滥用人工智能的现象时有发生。因此，需要适用法律来对人工智能发展设定界限，明确使用者使用人工智能的活动范围与禁止事项，设定相应的法律责任来加强对人工智能开发者与使用者的管理，促进人工智能技术的可持续发展。

4.1.2. 人工智能发展的不确定性因素

第一，人工智能自身发展的不确定性因素[6]。首先，与人类综合思考模式不同，人工智能的思考模式系单向思考模式，由此带来单一决策结果，其不会多角度、多方面思考问题从而进行综合决策，只会关注已经输入智能系统的任务有没有被执行，那么这不免会产生决策与执行错误的问题。以扫地机器人为例，一般人发现地上有钻戒等贵重物品会将其拾起并好好保存，而扫地机器人在进行清扫工作时，很可能无法识别地上的钻戒等贵重物品是否为垃圾，其往往因贵重物品的实物性质而将其当作垃圾处理。这就是扫地机器人的局限所在，他们往往无法区分垃圾与贵重物品，只会执行简单的清扫任务。其次，人工智能的决策有时会超出人类控制范围。与人类自主学习不同，理想模式下的人工智能是在对智能系

统研发者所输入的海量数据进行归纳、分析与整理的基础上进行判断并做出最终决策。然而，人工智能系统在运行过程中可能会违反预先设定好的具体规则，从而产生全新的决策结果。此外，在没有提前设定执行指令的情况下，人工智能的决策行为及其结果是难以预测的，其所作出的决策结果很可能在人类控制范围之外，这增加了规制人工智能的难度，一般的制约手段无法对其进行有效限制，需要采取更有效的方法来规范人工智能的发展。最后，人工智能的决策结论可能是错误的。许多人工智能是在分析过去数据的基础上对未来进行预测，数据来源本身的可靠性无法确定，因此依赖以往数据而进行的决策行为所作出的决策结果的准确性是无法保证的。此外，人工智能运行前需要由算法开发者输入相关数据，这种输入数据行为不可避免地带有算法开发者自身的观点或偏见，很可能会出现“算法歧视”问题，那么在“算法歧视”下所作出的决策结论毫无疑问是错误的，不应当被人类采用。

第二，人类生活方式与思维模式的重大改变。人工智能的发展正在潜移默化地影响人类的生活方式与思维模式。随着手机的普及，绝大多数人每天手写的次数大大减少，“提笔忘字”的现象越来越普遍，许多人的手写能力正在逐渐减退。那么，随着人工智能的迅猛发展，人类的某些机能会不会在它的影响下出现退化？答案是显而易见的。当人们需要知晓某个词语的含义时，其通常不会经过分析、推理、对比等思维模式来探究该词语的意思，反而会下意识地打开百度、谷歌等搜索引擎，输入“XXX的含义是什么”等关键词，几秒钟内该词语的含义就会立刻呈现在人类眼前。通过使用人工智能，人类能够迅速获取知识，从而提高学习效率。然而，这种依赖人工智能获取知识的方式会使人类形成“惰性化思维方式”，其逐渐不会通过大脑思考来探究知识的真谛，思考能力可能会有所降低，长久发展下去是不利于人类智力的延伸与进步[2]。

第三，就业环境与择业观念的巨大转变。人工智能的发展会影响整个就业市场，更多工作效率高、不求薪酬、不需要休息、能够从事危险性工作的智能机器人成为同人类一样的“劳动者”，参与到日常劳动中去，取代人类从事技术含量低且简单的工作，就业结构可能会发生改变，人类的就业环境逐渐严峻。如果人类不通过加强自我学习来提升就业竞争力，越来越多的人很可能会陷入失业的沼泽难以自拔。大规模的失业可能会拉大贫富差距，社会矛盾激化，就业环境会进一步恶化。然而，人工智能的发展也会使新兴职业产生。虽然人工智能取代人类从事简单、机械的低端劳动，但是也会促使监管人工智能工作的监督岗位产生，那些被人工智能取代从事机械性劳动的人类有机会转变为监督人工智能工作的监督人员，缓解了部分就业压力[1]。由于技术含量低的工作极易被人工智能取代，这迫使人类在择业时不会考虑那些技术含量低的机械性工作，而是通过提升自我去争取具有更高技术含量且不容易被替代的工作，人类的择业观念有巨大转变。

4.2. 宪法规制人工智能发展的正当性

从宪法地位角度看，宪法作为我国的根本大法，是治国安邦的总章程。既然能够适用法律对人工智能发展加以一定限制，那么具有最高法律地位的宪法毫无疑问也可以对人工智能进行规制。但是，这明显不是宪法能够规制人工智能发展的根本原因，那么最核心的理由是什么呢？

宪法的核心价值和共识是不让人性边缘化，保护人的尊严。但是人工智能的发展无疑会对人的尊严造成冲击，这违背了宪法制定的初衷。因此必须坚持一个基本立场，即人工智能的发展不能以损害人的尊严为代价，必须以保护人的尊严为发展前提[6]。如果人的尊严丧失，那么人类就没有存在的意义，人类文明也不会延续下去。因此，人工智能的发展必须控制在合理的范围内，尽量减少人工智能发展过程中带来的非理性危害，从而降低人工智能发展过程中所带来的风险需要一个制度安排，也即“宪法治理”。

在探究为什么采用“宪法治理”这一模式来规范人工智能发展之前需要明确宪法发展历史与科技发展历史之间存在对应关系。在启蒙运动时期，欧洲掀起了一场“理性”革命，“理性”为众多启蒙运动者

所提倡，这极大解放了人类思想。“理性”的提倡实际上就是早期宪法观念的体现，这为人类基本权利的保护提供了理念支撑，人类从虚无缥缈的神学研究逐渐转向以实验为基础的近代科学研究，为以实验为基础的近代科学的建立和发展奠定思想基础[3]。在第一次工业革命时期，在提高人类生活水平的同时科技也造成了一系列社会问题，如贫富差距拉大、劳动压迫、环境污染等。为了解决工业革命所带来的社会风险，许多国家纷纷制定宪法，如1787年的美国宪法，宪法规制成为解决科技发展所带来的非理性问题的有效手段。随着第二次工业革命的开展，“资本主义”“社会主义”等各种价值观念出现在人类视野中，选择哪种主义的价值观念来统领国家的发展是人类关注的话题。世界上第一个社会主义国家——苏维埃共和国的建立催生了世界上第一部社会主义宪法——1918年苏俄宪法的颁布。“社会主义”的“社会”一词所包含的公平、平等、正义等价值观念就是源自1918年的苏俄宪法。此外，现在很多国家宪法所提倡的社会福利、弱势群体保护、劳动者权利等理念均是来自1919年的魏玛宪法，宪法成为社会价值理念形成的养料。在第三次工业革命期间，第二次世界大战摧毁了原本建立的相对和平的世界秩序，其中化学病毒、原子弹等战争手段的使用给人的尊严造成巨大损害，人类的生存权难以得到有效保障，此时学者们不禁质疑宪法秩序与宪法理论的作用，他们提出当前宪法在面对战争时毫无作用，因此需要对宪法内容进行改革，这很大程度上促进现代宪法的反思与重构[3]。

综上所述，如果没有宪法，科技也不会有如此巨大的进步；如果宪法中不包含对科技发展的反思结果，宪法本身也将失去发展活力。因此，作为科技的人工智能只有依靠宪法，坚守“人的尊严”这条底线，才能在新时代迸发新活力，为提高人类的获得感与幸福感助力。

5. 人工智能发展的宪法限度

人工智能发展具有两面性。一方面人工智能发展引发人类生活方式的变革，进一步提高人类生活质量；另一方面，若人工智能发展失去控制，这会对宪法所追求价值以及宪法本身发展造成冲击，“算法歧视”“侵权认定困难”等一系列法律问题将会产生。因此，为解决由人工智能发展所产生的法律问题，作为控制人工智能发展关键的宪法应当对其设定一定界限，使其在宪法规定的范围内有序发展。

5.1. 达成宪法共识：人的尊严

有学者在分析193个联合国会员国的宪法文本的基础上发现143个国家将“人的尊严”写入宪法[8]，可见“人的尊严”之重要性。然而，“人的尊严”概念的抽象性、复杂性、内涵难以确定性导致学界对“人的尊严”一词内涵争议不断[9]。但是，学者关于人的尊严的争论不是否定人的尊严的价值，而是在认可其价值的基础上论证人的尊严的价值所在。作为人类最高价值的人的尊严是人类享有自由和人权的前提，如果人的尊严没有得到保障，人类的自由与人权将无法实现。如今，保障人的尊严已经成为世界共识，也是宪法所追求的最核心的价值所在，换言之保障人的尊严系宪法共识。人的尊严不可侵犯性不仅是构成宪法基本权利体系的前提，还是确立宪法制度体系的基本原理所在，人的尊严蕴含的价值是一国宪法发展的丰富养料。

确定人的尊严的具体内涵系人的尊严价值实现的前提。首先，人本身就是目的为人的尊严之核心要点。人类发展的目的就是为了人类本身，这是人的目的属性的基本要求，其涉及人与人以及人与国家的关系，是宪法的最高价值所在[8]。无论在哪个时期，保障人的尊严都应当作为人类发展的最高使命，不允许有任何践踏人的尊严的行为出现。其次，保障人的主体性地位是人的尊严的价值目的。关于这一点将放在下文着重论述。最后，尊重人的多样性是人的尊严的价值延伸。其要求尊重独立个体的多样性，绝对禁止歧视行为，不能因为自身人的尊严而阻碍他人的尊严的实现，不允许以已经达成社会共识而否认他人独特个性的存在，每个人都有自由支配自己意志和行为的权利。只有尊重人的多样性才能构建和

谐美好的世界关系，才能在此基础上发现人的更多潜能和扩展人类的发展空间[10]。

然而，人工智能的发展对保障人的尊严这一宪法共识提出挑战，人的尊严之价值根基在人工智能发展失控的冲击下逐渐动摇。人工智能主要依靠算法系统而进行正常运转，在算法系统中，只存在数据，不存在人类，这导致算法运行时并不存在对人的尊严的保障。但是，算法是由人类设计和操作执行的，其中包含着算法研发者对人的尊严的价值判断。“算法歧视”问题出现的主要原因就是算法设计者在输入相关数据时将偏倚的观点混入其中，这导致由归纳、分析、整理数据所得出的结论也带有不公正的色彩，不利于对人的尊严的保护。为此必须重申保障人的尊严是宪法共识这一论断，重新塑造人的尊严的价值内涵，使其与人工智能发展相适应。

5.2. 符合宪法发展目的：人类的主体性地位

保障人类的主体性地位是维护人的尊严这一宪法共识的内涵延伸，是人的尊严的价值目的。其强调任何人都享有在不受他人非法干涉的情形下在法律允许的范围内自由行动的权利，禁止人类被当作物体对待，不允许将人类边缘化、客体化。如果人工智能发展凌驾于人类自身发展之上，忽视人类的主体性地位，那么人类也没有存在的必要，人类文明将不再延续。此外，如果某种人工智能技术的发展超出人类控制的范围，严重损害人的尊严，如开发具有侵害人类生命权、健康权的智能机器人，此时人类需要考虑是否应当禁止该项人工智能技术的发展[6]。因此，必须牢记人工智能只是人类扩展发展空间的辅助工具，人类才是人工智能的主人，其应当对人工智能具有绝对控制力。

当下，过度依赖人工智能而导致人类的主体性地位的削弱现象时有发生[10]。一方面，过度依赖人工智能导致人工智能使用者的决策主体性地位削弱。由于人工智能本身决策的高效性和相对精准性，在人类需要对事物做出判断时其往往会借助人工智能强大的决策功能来获取判断结果。由此，人工智能的发展正在潜移默化地影响人类的思维模式，这种依赖人工智能获取决策结果的模式会使人类形成“惰性化思维方式”，人类逐渐不会通过大脑思考来探究事物的本质，其思考能力可能会有所降低，长久发展下去是不利于人类智力的延伸与进步。另一方面，过度依赖人工智能导致受人工智能决策约束的对象主体性地位削弱。人工智能系统运行时不是把人类当作独立主体对待，而是把其认定为可以被归类、对比以及分析的客体数据。每个独立个体均能够成为人工智能进行判断与决策的数据来源，人工智能在人类所提供数据的基础上对这些数据进行分析与处理，最终形成决策结论。此外，人类并不知晓人工智能运转以及进行决策的工作原理，只能被动接受人工智能所作出的决策结论，这给人类及时应对由人工智能所造成的“算法歧视”等一系列风险设置了天然壁垒，不利于人类基本权利的保护。

因此，人工智能的发展必须符合人的尊严的价值目的，即保障人类的主体性地位。人工智能的发展目的是造福人类，使人类的主体性地位得到重视，不能违背人的尊严的价值目的，必须使人工智能的发展符合宪法发展目的，进而达到保障人的尊严的最终目标。

5.3. 遵循宪法基本原则：技术开发类型化

为了降低甚至杜绝人工智能可能会对人类生命权、健康权等基本权益造成损害的情况发生，国家在进行人工智能研发活动时需要遵循技术开发类型化的宪法基本原则，即并不是所有类型的人工智能技术都可以被开发，只有在特定领域内才可以研发该人工智能以及相关技术[6]。由此，在开展人工智能研发活动时国家应当坚持利益衡量原则，在确认是否需要研发某一领域内的人工智能技术之前首先需要考虑什么领域的人工智能可以被研发，将在该领域开发该人工智能技术的利益与危害进行对比，如果其可能造成的风险并不是涉及人的尊严等基本人格利益，国家可以允许在该领域内进行人工智能技术的研发活动；但如果研发该人工智能技术所带来的危害严重阻碍国家对人的尊严的保障，可能会使人类边缘化、工具化，那么应当绝对禁止该项人工智能技术的研发。对于被允许进行研发活动以及被投放市场应用的

人工智能,国家也应当对其加以严格控制,减轻其给人类带来的相关风险,确保人工智能造福人类的实现。

5.4. 建立宪法秩序: 和平有序

随着全球化的深入发展,世界各国的联系也愈发紧密。通过对人类发展历史的回忆与反思,人类逐渐意识到只有和平有序的国际秩序才可以确保人类文明的延续,宪法也是如此。只有在和平有序的宪法秩序建立后才可以探究对人工智能发展予以一定限制的问题。如果各国宪法秩序混乱,普通民众违背宪法精神,践踏宪法所提倡的价值观,那么研究人工智能的宪法限度这一话题也是没有意义的。因此,建立和平有序的宪法秩序是提倡人工智能发展需要符合保障人类主体性地位的宪法目的和遵循技术开发类型化的宪法基本原则的前提[6]。

参考文献

- [1] 任润泽. 人工智能法律规制问题探析[D]: [硕士学位论文]. 北京: 中国人民公安大学, 2020.
- [2] 郑雅琪. 我国人工智能法律规制问题研究[D]: [硕士学位论文]. 兰州: 兰州大学, 2019.
- [3] 高文杰. 人工智能时代的人权问题研究[D]: [硕士学位论文]. 郑州: 郑州大学, 2021.
- [4] [美]埃德加·博登海默. 法理学[M]. 北京: 中国政法大学出版社, 2017: 85.
- [5] 杨超. 人工智能时代算法权力及其法律规制[J]. 湖北工程学院学报, 2022, 42(1): 89-94.
- [6] 韩大元. 当代科技发展的宪法界限[J]. 法治现代化研究, 2018(5): 1-12.
- [7] 姜野. 算法的法律规制研究[D]: [博士学位论文]. 长春: 吉林大学, 2020.
- [8] 齐延平. “人的尊严”是《世界人权宣言》的基础规范[J]. 现代法学, 2018, 40(5): 22-38.
- [9] 郑玉双. 人的尊严的价值证成与法理构造[J]. 比较法研究, 2019(5): 170-185.
- [10] 洪丹娜. 算法歧视的宪法价值调适: 基于人的尊严[J]. 政治与法律, 2020(8): 27-37.