Open Journal of Legal Science 法学, 2022, 10(3), 270-280
Published Online May 2022 in Hans. http://www.hanspub.org/journal/ojls
https://doi.org/10.12677/ojls.2022.103035

Hans汉斯

算法黑箱之规制——基于金融法视角

李阳桂

台湾成功大学法律所，台湾 台南

收稿日期：2022年3月25日；录用日期：2022年4月13日；发布日期：2022年5月27日

摘　要

算法作为人工智能的重要组成部分，不断提升人工智能的智能化、自动化、自主化水平，从而使人工智能在网络购物、休闲娱乐、金融理财等领域的应用日益深化。但算法黑箱作为算法的负面效应之一，不仅侵害金融消费者的知情权，为金融机构谋取不正当利益提供便利，而且严重削弱金融监管机构履行正常的监管职能，扰乱正常的金融秩序，给金融市场带来系统性风险。为此，应从立法部门、监管机构、金融机构的维度出发，基于维护金融消费者的合法权益这一目标，构建一个多方参与、系统协同、监管有效、保障有力的算法黑箱应对体系。

关键词

算法黑箱，金融消费者，监管机构

Regulation of Algorithmic Black Box—From the Perspective of Financial Law

Yanggui Li

Department of Law, National Cheng Kung University, Tainan Taiwan

Received: Mar. 25th, 2022; accepted: Apr. 13th, 2022; published: May 27th, 2022

Abstract

As an important part of artificial intelligence, algorithm constantly improves the level of intelligence, automation and autonomy of artificial intelligence, thus deepening the application of artificial intelligence in online shopping, leisure and entertainment, financial planning and other fields. However, as one of the negative effects of algorithms, black boxes not only infringe on financial consumers' right to know and facilitate financial institutions to seek improper interests, but also seriously weaken the performance of normal regulatory functions of financial regulators, disrupt the normal financial order, and bring systemic risks to the financial market. Therefore, from the

文章引用: 李阳桂. 算法黑箱之规制——基于金融法视角[J]. 法学, 2022, 10(3): 270-280.
DOI: 10.12677/ojls.2022.103035

dimension of legislative departments, regulatory agencies and financial institutions, based on the goal of safeguarding the legitimate rights and interests of financial consumers, we should build an algorithm black-box response system with multi-party participation, systematic coordination, effective supervision and strong guarantee.

Keywords

Algorithm Black Box, Financial Consumer, Regulator

1. 引言

人工智能时代，随着机器学习、机器视觉、语言识别、图像识别、自然语言处理、大数据等技术的不断进步，算法作为人工智能的重要载体，已广泛应用于新闻媒体、网络购物、投资理财等领域，人类社会也正式进入“算法社会”。但算法并非既得利益者所宣称的技术中立和价值中立，算法的公正性、客观性、可解释性日益受到社会的广泛重视。国家互联网信息办公室、工业和信息化部、公安部、国家市场监督管理总局联合发布的《互联网信息服务算法推荐管理规定》(以下简称《规定》)，已于 2022 年 3 月 1 日起施行。该《规定》体现了国家对算法负面效应的规制，有利于维护国家利益、社会利益，更有助于保护公民、法人和其他组织的合法权益，防范和化解算法风险，推动国家治理体系和治理能力的现代化。

2. 算法相关概念及成因分析

2.1. 算法相关概念

算法严格来说，是一个数学的概念。为了更好地认识世界，科学家们希望通过数学的方式将原理、理论逻辑化，通过科学的公式证明来实现对客观世界的描述和表达。有的学者认为算法是“为实现某个任务而构造的简单指令集”[1]，是“为了解决一个特定问题或者达成一个明确的结果而采取的一系列步骤”[2]。中国科学技术协会对于算法的定义是：“算法是指解题方案的准确而完整的描述，是一系列解决问题的清晰指令，算法代表着用系统的方法描述解决问题的策略机制”[3]。不同的研究视角，对于算法的界定也各不相同。正如有的学者所说，“从数学意义上说，算法是通过各种步骤得到计算成果的方程式；从生产意义上说，算法是收集处理数据、挖掘数据价值的生产工具，也是人工智能时代平台的架构和运行方式；从社会意义上说，算法成为了支配数据流动的权力”[4]。如果说大数据是人工智能的基石，算法则是人工智能的灵魂。大数据本身并不能体现出价值，只有透过算法的优化、重组，大数据的内在价值才能充分发挥出来。

算法作为人工智能具有决定性的力量，已深刻改变着人类的生产方式、生活方式和研究范式，其对整个人类的影响至今仍然不可估量，并且以惊人的能量不断颠覆者人民的认知。正如著名学者张文显所指出的，“随着信息技术的发展，万事万物、社会运行的每一个部分都可互通互联，并提供海量多样化的数据供智能算法分析处理；智能算法的预测和决策则可以直接控制物理设备，亦可为个人决策、群体决策乃至国家决策提供辅助支撑，带来了智慧家居、智能交通、智能医疗、智能工厂、智慧农业、智慧

城市等诸多领域的发展，为我们描绘出智慧社会的景象，深刻地影响着人们认识世界和改造世界的能力，勾勒着人们生活和社会组织形态的可能边界”[5]。牛津英语大辞典给算法的定义为“一套用于计算和解决问题的程序或规则”，而所谓的“黑箱”作为一种隐喻，它指的是那些不为人知的不能打开、不能从外部直接观察其内部状态的系统[6]。而算法黑箱就是人工智能的数据输入和数据输出之间存在着隐藏层，使人无法获悉其真正的内涵。Jenna Burrell 在其论文《机器如何“思考”：理解机器学习算法中的不透明性》(How the Machine “Thinks”: Understanding Opacity in Machine Learning Algorithms)论述了三种形式的不透明性：因公司商业秘密或者国家秘密而产生的不透明性，因技术文盲而产生的不透明性，以及从机器学习算法的特征以及要求将它们有效适用的测量中产生的不透明性[7]。

2.2. 算法黑箱的成因分析

不可否认，算法为人工智能的应用贡献了巨大的力量。算法不仅使人工智能具有了学习能力，而且算法还使人工智能具有了自主决策能力，使人工智能真正具有了智能性。学者们为了打破算法黑箱，为其增添了透明性和可靠性的要求。算法黑箱在体现了设计者或者算法的用户特定的目的，从而导致算法在很大程度上有利于上述主体的利益。而对于算法的适用对象来说，由于缺乏专业的算法知识，根本不了解算法的具体模型和机制，从而导致信息不对称，或者侵犯使用者的知情权。但事实上，对于算法黑箱的规制远非人们想象的那么简单。

首先，从技术的角度来说，社会公众的算法知识素养不足。由于算法涉及的代码数以万计，不仅普通的公众不太知悉算法的运算规则、决策机制，但随着时间的推移和新数据的不断加入，算法自身也在不断地升级，因此，哪怕算法的专业人士也不一定对所有的代码都能清楚地知道学习模式和决策模式，从而使得算法的透明性和可解释性大打折扣。“在目前的技术水准和社会结构下，编写和阅读代码是一项专业的技能，采用透明性方案将代码公开给一般使用者或普通公众并不能使其获得正确的理解。算法的可解释性方案同样存在技术障碍，由于机器学习算法所采用的对高维度特征的数学优化方法与人类逻辑推理和语义解释的智慧活动在思想和方式上无法匹配，因此很难使用一般的语义表达方法予以说明”[8]。

其次，从知识产权保护的角度来说，算法的透明性和可解释性有侵权之嫌。从算法的性质来看，其满足商业秘密的经济性、秘密性等构成要件，属于商业秘密的范畴。更何况人工智能算法的研发者花费大量的时间、精力、财力设计算法的程序，从而使得算法能够完成特定的工作任务，而且还需要具备一定的智慧性和自动性，完全符合知识产权的保护范畴。在算法的监管中，使用算法的商业机构往往以商业秘密为由拒绝公开其算法，这也是商业机构保护自身知识产权的正当之举。但不少学者对此有所非议，认为其拒绝监管的行为，导致算法处于法律的真空地带。2007 年 3 月，媒体巨头维亚康姆(Viacom)以“厚颜无耻”和“大规模”侵犯版权为由，对 YouTube 提起了 10 亿美元的诉讼，声称谷歌公司旗下的在线视频共享网站在未经许可的情况下向 YouTube 用户提供了约 16 万段维亚康姆旗下的视频。在 Viacom v. YouTube 一案当中，谷歌公司就以算法涉及商业秘密为由拒绝公开算法，并且得到了法庭的支持[9]。

最后，从现行实在法评价角度来说，关于算法方面的立法较为薄弱，法律法规的规定较为模糊。众所周知，人工智能作为前沿技术，尽管其应用范围日益广泛，其影响的行业日渐增多，其所带来的风险日益显现。但我国并无任何一部法律专门针对人工智能所带来的挑战进行立法，更不像美国专门针对算法进行立法，制订了问责法。国务院所制定的《新一代人工智能发展规划》(2017 年)虽然明确提出了要“实现对人工智能算法设计等程序的全流程监管”，但由于该规定过于原则，缺乏可操作性，其效果自是可想而知了。以外卖骑手频繁发生交通事故为例，各大平台利用算法，得出骑手送达物品的最佳时间，从而提高绩效考核标准。外卖骑手为了获取更多的绩效，不得不铤而走险，不断提高汽车速度，甚至故意交通违法，从而导致交通事故频发，造成重大人员伤亡和巨额经济损失。

3. 算法黑箱在金融领域之体现及其风险

3.1. 算法黑箱在金融领域的外在形态

“黑箱”是一个信息控制论里的概念，主要是指研究者不关注系统的内部结构和相关关系，而是从输入和输出的角度来认识和掌握系统的规律性。算法黑箱在人工智能金融领域的应用较为常见，以智慧投顾为例，其表现出的外观是一种程序化交易过程。投资者并不知道该程序的设计原理、运行参数、决策依据、交易机制，而只能看见从数据输入到决策输出的客观事实和结果，这个被暗藏的智慧决策“隐层”，即所谓的大数据金融算法黑箱[10]。算法黑箱在很大程度上反应了金融消费者和金融机构之间的信息不对称，从而导致金融消费者在认知上的不足和理解上的差距。从人的认知能力的角度，人不可能对所有的行业和领域都能达到专业人士的水平，这是不可能也是不现实的，但问题的关键就在于金融机构是否凭借其专业的优势，去诱使金融消费者进行交易，从而谋取不正当利益。对于监管机构来说，如果不能清晰地掌握算法的透明度，对于金融机构侵害金融消费者的信息掌握不够，对金融消费者的保护也就无从谈起，将使金融消费者的权益处于危险的境地。所幸，中国的金融监管部门已经意识到了算法监管的必要性和重要性，在 2018 年 3 月 28 日，中国人民银行、银保监会、证监会、外汇局联合发布的《关于规范金融机构资产管理业务的指导意见》(以下简称《指导意见》)对此进行了规制。“该《指导意见》对人工智能在金融领域的应用进行了规制，从胜任性要求、投资者适当性以及透明披露方面对智能投顾中的算法进行穿透式监管”。从国外对算法的监管来看，美国证监会在 2017 年 2 月发布的《智能投顾监管指引》中将“智能投顾”定义为基于网络算法的程序，利用创新技术为用户提供全权委托的账户管理服务的注册投资顾问。智慧投顾依托互联网，具有委托的特殊性，它为用户提供全权委托的账户管理服务，即在客户和运营者签署全权委托协议的前提下，允许受托人未经客户同意买卖证券[11]，将此种方式将纳入监管的范畴，从而保护投资者的利益。

3.2. 算法黑箱在金融领域的风险

算法作为一种高度智能化的工具，在给金融行业赋能的同时，也给金融行业的不同主体带来不同的挑战。对于金融消费者来说，金融机构或平台出于谋取利益最大化的动机，为了获取足够多的个人信息，其通过网络爬虫技术，过度搜集金融消费者的收入、工作、教育背景、婚姻状况、投资偏好等个人信息，从而使其产品、服务更智能化、自主化和个性化，以提高客户粘度；而对于金融监管机构来说，由于金融机构采用了先进的人工智能技术，传统的监管手段、监管理念、监管方式已落后于时代的发展要求，不能有效履行监管职责，致使正常的金融竞争秩序崩塌瓦解。金融机构或平台，出于自身利益最大化的动机，对于算法的滥用更是肆无忌惮，不仅侵害金融消费者的合法权益，而且破坏正常的金融秩序，逃避金融监管机构的正常监管，使国家面临系统性金融风险的挑战。

3.2.1. 算法黑箱侵害金融消费者合法权

算法黑箱不仅是一个技术问题，而且也是一个关系到金融消费者知情权的法律问题。算法黑箱的出现，有着诸多的原因。从技术上来说，算法毕竟是一个数学和计算器科学的专业术语，有着专业的内涵和数学模型，普通人对于算法的数学函数、决策模型几乎难以洞悉。正如机器学习算法使人难以理解一样，由于一个有效的机器学习算法模型通常包含数百万数据单元和数万行代码，并会随着时间的推移、模型用户的参与、新的数据的输入，自动地调整其内在的决策模型和代码内容。在这种情况下，设计者和运营者很难确切地说明其算法是如何做出决策的，通常只能通过有限的测试用例或其他验证方法给出该模型的预期效果和准确程度[12]。学界对于算法黑箱有不同的争议，有的学者赞成算法公开，理由有：第一，把算法透明当成消费者知情权的组成部分。这种观点主张，因为算法的复杂性和专业性，人工智

能具体应用领域中的信息不对称可能会更加严重，算法透明应是消费者知情权的组成部分。第二，算法透明有助于缓解这种信息不对称。这种观点主张，算法的信息不对称加重不只发生在消费者与算法设计者、用户之间，更发生在人类和机器之间，算法透明有助于缓解这种信息不对称[13]。当然，赞成者的理由还有很多，集中在有助于防止利益冲突、有助于防止信息茧房等方面。反对者的理由认为：一是普通的人们根本不可能知道如此复杂的专业知识，向普通的人们揭露其算法信息，并不具有多少实际意义；二是算法作为技术开发者的知识产权，一旦泄露知识产权，将侵犯技术开发者的知识产权，不利于知识产权的保护。曾经就发生过一个真实的案例，例如在搜索算法中，谷歌(Google)曾经依赖于一种叫作PageRank 的算法确定搜索排序，这种排序方法主要根据 META 卷标、关键词等参数进行排序。当谷歌公开这一算法之后，很多网站就开始利用此类算法，在自己的网页内嵌套符合 PageRank 算法的具有隐藏内容的网页，以此达到提高网站在谷歌搜索结果页面排名靠前的目的。经过此类设计后，一些与搜索内容并不相关的网页也被谷歌搜索结果搜索并排在前面[14]。算法不仅侵害了金融消费者的知情权，而且侵害了其选择权。各大平台利用自己所掌握的资源优势，名义上是基于金融消费者的“知情－同意”搜集其个人信息，但实质上该同意是金融消费者的无奈之举。如果金融消费者选择不同意，则金融消费者就不能使用该平台的软件或产品，给生活、生产带来极大不便。

3.2.2. 算法黑箱导致私人平台控制社会

尤瓦尔·赫拉利曾预言，“权威将从个人转向由算法构成的网络。人类不会再认为自己是自主的个体，不再依据自己的期望度日，而是习惯把人类整体看作一种生化机制的集合体，由电子演算法网路实时监测和指挥”[15]，人类由此进入算法主导的时代。人工智能的“智能性”离不开基于机器学习的算法的加持，如果说大数据是人工智能的躯体，而算法则是人工智能的大脑。算法能否具有智能性，在很大程度上取决于算法的准确性和科学性。从准确性来说，算法进行决策的程序设计，不能出现明显的错误。从算法的科学性来说，算法针对不同的使用场景和适用对象，应适用不同的决策程序。算法影响到社会和个人的机制主要通过两个层次：第一层是算法设计，指设计者编写算法决策代码，并输入数据使算法自主学习，优化决策流程的行为；第二层是算法部署应用，企业在其平台上部署应用算法的行为。这两个层次可能合一，也可能分离[16]。在智能投顾领域，算法通过前期的信息收集，了解个人收入、教育背景、家庭住址、工作职位、年龄、性别等数据，然后建构投资的数学模型，并根据数学模型运行的结果，给出具体的投资建议或者理财方案。在全委托的情况之下，智能投顾就依靠算法，根据委托合同的约定，直接对客户的财产进行资产管理。随着算法在金融领域的不断运用，其深度、广度得以提升，其对社会的影响力、对资源的掌控力、对舆论的引导力不断增强。正如有的学者所担忧的，“因掌握关键信息技术而操纵算法的私人技术公司占据着优势地位，但政府在‘算法社会’中却被边缘化，逐渐失去了对算法关键数据的所有权和控制权，面临着去中心化的挑战”[17]。

3.2.3. 算法黑箱导致金融监管失灵

我国的金融监管机构为“一委一行两会”(“一委”为国务院金融稳定发展委员会，“一行”为中国人民银行，“两会”分别是证监会、银行保险监督委员会)，金融监管机构不仅具有制订和实施国家货币政策的职责，更重要的是在维护金融行业稳定、健康发展的同时，防范和化解金融风险，维护金融安全和国家安全。金融机构或者平台，充分利用算法的智能性、自主性、自动性，大量获取金融消费者的个人信息，通过“精准画像”，为客户推荐个性化、自动化的产品或服务，提高营销的准确性。不仅如此，算法还能根据金融机构的利益需要，在降低投资门槛的同时，向消费者推荐远远大于其自身风险承受能力的产品或服务，误导金融消费者，从而放大投资者的投资能力。一旦市场出现波动，基于算法的量化交易，将使整个市场承受非理智的抛售，冲击整个金融系统的安全和稳定。正如有的学者所指出的，“算

法的技术特征给外部监管带来了巨大的技术挑战，而算法的法律属性又为外部监管增添了新的法律难题，单纯的外部监管无法胜任算法权力治理的重任”[18]。加强对算法的规制，不仅关系到金融消费者的权利保护，关系到金融市场的稳定，而且与社会所有人的利益息息相关，甚至与国家的安全密不可分。

4. 算法黑箱之规制

那对算法黑箱还需要进行法律规制吗？答案是不言而喻的。尽管对算法黑箱进行公开，会有知识产权方面的风险，但笔者并不认为因此就因噎废食，让算法这一人工智能的核心技术游离于法律之外。这只会让作为技术运营者的金融机构利用算法这一技术，肆无忌惮地侵害金融消费者的权利，而不需要承担任何法律后果，这在崇尚法治的当今时代是不能得以允许的。那如何对算法的黑箱问题进行法律规范呢，尤其是涉及到金融领域。笔者认为，对算法黑箱的规制，不仅仅是某一个机构就能解决的，而是一个系统工程，不同的机构，不同的主体，都应齐心协力，将算法纳入法律的轨道。立法部门，通过法律的制订，为算法黑箱的规制提供法律依据，同时设立责任制度，为金融消费者的损害赔偿提供保障；监管部门，成立垮部门、跨专业的技术部门，负责算法的技术监管；金融机构作为技术的运营者，应当建立内部的审核机制，以使开发的技术符合法律和伦理规范。

4.1. 立法部门对算法的规制

人工智能的发展，可以说得上是技术革命引起的社会革命，对人类生产、生活的影响将是各个方面的。正如法国哲学家德布雷(Régis Debray)写道：“归根到底，唯一跳出星球运转的循环意义外的革命不是政治革命而是技术革命，因为只有它们才是不复返的。有了电流后就不再用蜡烛，有了汽轮船就不再用帆船。然而有了十月革命还是回到了东正教……”[19]，面对技术革命对社会带来的革命，立法部门不能无所作为，而应积极响应社会关切，主动作为。遗憾的是迄今为止，大陆还没有一部专门针对人工智能的法律。稍微相关的法律，也只有跟互联网相关的几部法律。更多的仅仅是政府部门颁布的行政规章，跟人工智能相关规范的是 2018 年 5 月 1 日实施的《个人信息安全规范》。然而这只是一部推荐性的国家标准，不具有强制执行力。从法律效力来说，行政规章的法律效力肯定不如立法机构专门制订的法律。同时，对于社会相关主体的利益保护来说，行政规章的法律层级显然太低，不是向司法机关主张法律权利的有效武器。2017 年国务院颁布的《新一代人工智能发展规划》仅仅提到了相关要求，如制定促进人工智能健康发展的法律法规和伦理规范，加强人工智能相关法律、伦理和社会问题研究，建立保障人工智能健康发展的法律法规和伦理道德框架。但该文件只是一部发展规划，还不是部门规章，仅具有倡导作用，并无法律的约束力。

4.1.1. 制订一部由立法专家、法律专家、人工智能专家、伦理专家共同参与的人工智能专门法——《人工智能法》

在我国的规制实践中，算法规制并未被从网络治理和互联网规制中分离出来，保障算法应用安全的主要法律框架仍然是以《网络安全法》、《电子商务法》、《密码法》等法律为代表的互联网法律体系。然而，随着越来越多的算法风险开始来自分散的、独立的算法研发人员或者自行训练算法模型的算法用户，这些通常针对经营者运用的规制手段已经很难全面防范算法风险[20]。在该部法律当中，应当就人工智能使用的基本原则、技术开发者的责任、技术运营者的责任、算法的监管、政府监管机构的职责、行业协会的职责等方面的内容进行规定，同时，吸收欧美国家关于人工智能立法的基本精神和基本制度规定，在结合中国实际情况的基础上，构建有中国特色的关于人工智能的法律体系。正如百度的创始人李彦宏所说：“人工智能技术可能不只是理工科专业人士的领域，法律人士以及其他治理者也需要学习人工智能知识，这对法律人士和其他治理者提出了技术要求。法治管理需要嵌入生产环节，比如对算法处

理的数据或生产性资源进行管理，防止造成消极后果”[21]。

4.1.2. 制订一部专门的《算法问责法》

算法已经以全方位的形式渗透到人们的经济、政治、文化和社会生活当中，深刻改变着人与人、人与社会、人与技术等层面的关系，成为人类不可忽视的重要力量。然而，现有的法律制度无法充分因应商业领域算法应用带来的损害用户利益的问题。这主要由以下两个原因造成：其一，用户进入平台的系统架构需要点击同意使用者协议，而这种所谓的“知情同意”使得算法收集和利用数据的行为合法化。平台可以主张使用者已经通过用户协议知晓和同意算法决策，因此无须为用户受到的算法的不利决策承担法律责任。其二，算法造成的用户利益损害的情形不符合现有法律责任的认定规则。算法的“用户画像”与“个性化价格歧视”是否合法仍有较大争议。技术运营者往往以商业秘密为由拒绝披露算法的运算规则[16]。加强对算法的法律规制，应是因应人工智能时代发展的必然要求。算法专法应当就算法的国家标准、算法的审查、检测、算法的透明性、可解释性、算法的适用领域、算法的法律责任等方面的内容进行规定，从而为算法的规制提供法律保障。一些国家已通过立法对算法的公平和透明进行专门监督。美国参议员在 2019 年 4 月提出的《算法问责法案》要求美国联邦贸易委员会对企业进行算法审查，适用的对象包括年收入超过 5000 万美元的公司，以及拥有超过 100 万消费者数据的数据代理商和企业[1]。

4.1.3. 构建以技术应对技术的立法理念，通过法律化的代码应对算法的挑战

传统的监管理念已不能适应人工智能时代发展的要求，算法的设计者是人，算法的控制者也是人，通过将法律代码化，确保算法符合人类的利益，不得侵犯人类的隐私权、信息权和财产权，更不能侵害人的尊严和人的生命。“通过将被审查的机器学习算法接入算法审查系统中，其在确切条件下所产生的结果分布应当是十分有限的，这一方法通常能够处理一个输出稳定的自动化决策系统。虽然机器学习算法的过程缺乏透明性和可解释性，但通过含有合规性判断的审查算法能够在预先设计的监管框架下一定程度地表明算法系统在运行时的逻辑架构和表现效果”[8]。技术的发展没有止境，而法律的发展往往具有滞后性和安定性，为此，通过技术的手段将算法予以规制，将符合人类利益的价值观和规则植入算法当中，不失为一个较好的应对算法挑战的方法。

4.2. 金融监管部门对算法黑箱的规制

监管部门并不仅仅局限于履行国家监管职责的政府机构，还应包括具有人才优势、技术优势的行业自律组织或者第三方专门的监管机构。作为政府部门来说，掌控者国家行政权力，具有完善的机构设置和高素质的执法人员，能动用国家行政力量要求提供技术开发者或技术运营者公开算法的相关信息。作为行业自律组织来说，其对技术运营者的监管更具专业性、有效性和针对性。作为行业自律组织来说，能够统筹整个行业的权威专家，利用行业协会的力量，加强技术监管，并能通政府监管部门有效沟通，提高监管的效力。政府部门和协会自律组织，作为两种不同的监管方式，具有不同的监管任务，二者协同监管，有助于提升监管的水平。

4.2.1. 明确算法的应用边界与限制

政府机构在审查技术运营者的算法时，重点审查对于人们生命、人格尊严、隐私等构成重大威胁的算法，从而确保金融消费者的利益不受侵害。国外对此早有规制，《欧盟通用数据保护条例》第 22 条(1)规定，如果算法决策对数据主体有法律效力或者重大影响，那么这种决策不应纯粹由算法作出。国内也有学者建议，根据算法应用领域的不同，将算法分为禁止领域、限制领域和自由领域。禁止领域包括危害大范围人群或人类整体健康和生存的技术应用(例如人类基因编辑技术、智能杀伤性武器)，使用机器学

[1]该法案的英文名称为《Algorithmic Accountability Act of 2019》，旨在消除高科技企业当中算法所存在的任何形式的歧视。

习算法将会进一步增加这一领域的不可控性。限制领域包括可能影响一定范围或特定类型人群权利、自由、生命、健康、情感的技术或商业应用。监管机构应当出台禁止应用领域的负面列表，对负面列表上的内容全面地禁止机器学习算法的使用，组织机构在算法影响评估时应当重点判断是否涉及此类领域，只有在得到法律特别授权的情况下才得以使用[22]。

4.2.2. 政府监管机构应当建立算法事前审查、事中监督、事后补救的风险防范机制

算法投入应用之前，应当要求算法设计者进行风险测试和披露责任。正如有的学者所提出的，“算法投入使用前，要求设计者充分测试算法。并且应细化要求算法设计，必须进行必要的迭代测试以达到预测性能的客观视图”[23]。风险测试和信息披露，在很大程度上就是防止算法投入使用后，出现本可以避免的风险。事中监督主要是国家政府机构对算法进行标准化的评估。标准化的评估是国家公权力以公共利益为标准，对算法的运营者有无通过技术手段侵蚀金融消费者的合法权益进行监督。如果算法存在问题，根据问题的大小，则对算法的运营者进行整改、暂停使用、巨额罚款等处罚措施。事后补救，这是算法存在问题，侵害金融消费者权利的补救办法。严格来说，事后补救属于危机之后的规制。国家可以成立算法技术开发者、技术运营者的风险赔偿基金，如算法由于高度的复杂性，现有的科技水平难以发现其缺陷，但又给金融消费者造成了损害的，可以由基金给予补偿，从而最大程度维护金融消费者的利益。

4.2.3. 明确算法透明的具体规则

算法透明是打破算法黑箱最有力的武器，尽管人们呼吁算法透明，但对于如何使算法透明缺乏具体而明确的内容，从而使算法透明沦为一句空话，不具有实际的操作价值。其实，早在 2017 年，美国计算机学会公众政策委员会便公布了 6 项算法治理指导原则：第一项原则是知情原则，即算法设计者、架构师、控制方以及其他利益相关者应该披露算法设计、执行、使用过程中可能存在的偏见以及可能对个人和社会造成的潜在危害；第二项原则是质询和申诉原则，即监管部门应该确保受到算法决策负面影响的个人或组织享有对算法进行质疑并申诉的权力；第三项原则是算法责任认定原则；第四项原则是解释原则，即采用算法自动化决策的机构有义务解释算法运行原理以及算法具体决策结果；第五项原则是数据源披露原则；第六项原则是可审计原则[13]。由政府监管机构承担专业的算法审查，囿于算法的专业性和复杂性，政府监管机构对此恐难胜任。不如由兼具专业知识和人才优势的行业协会或者第三方评估机构担任，不仅具有可行性，而且监管的效果也有保障。第三方机构对算法的数学模型、建构原理、信息收集、决策机制、个人信息保护等内容进行审查，唯有经过审查的算法才能投入应用。国外已有相关的第三方机构从事算法的评鉴，如德国已经出现了由技术专家和资深媒体人挑头成立的名为「监控算法」的非营利组织，宗旨是评估并监控影响公共生活的算法决策过程。具体的监管手段包括审核访问协议的严密性、商定数位管理的道德准则、任命专人监管信息、在线跟踪个人信息再次使用的情况，允许使用者不提供个人资料、为数据访问设置时间轴、未经同意不得将数据转卖给第三方等[6]。尽管该组织是针对媒体进行的监管，但对于金融领域的算法监管来说，仍具有借鉴价值。

4.3. 技术运营者——金融机构对算法黑箱的规制

作为技术运营者的金融机构，对于算法，无论是其自身研发的，还是通过委托合同委托专门的技术公司研发的，金融机构对算法的应用都有高度的注意义务和说明义务，这是信义义务的必然要求。

4.3.1. 设置算法的内部审核机构或人员

金融机构作为算法的技术运营者，内部设置算法审核机构或者审核人员，是履行社会责任的具体体现，也是人工智能时代合法合规经营的必然要求。内部审核机构或者人员应当对算法的准确性、科学性、

风险性、可控性等方面进行审查，同时提供算法审查的分析报告，为监管机构的外部监管提供基础性信息。“算法运营商应当主动解释清楚智能机器人的硬件、软件和外部环境之间的相互作用，如何导致其当下的行为模式，并清晰、直观地就以下事项(包括但不限于)作出完整信息披露：信息挖掘的数据源、典型特征和分类方式；算法程序的运作原理、代码逻辑和预期效果；可能存在的系统偏差、运行故障和矫正机制”[24]。2018 年，中国大陆的金融监管机构，中国人民银行、银保监会、证监会、外汇局联合发布《关于规范金融机构资产管理业务的指导意见》也对金融机构的算法提出了规制的要求，金融机构应当根据智慧投顾的业务特点，建立合理的投资策略和算法模型，充分提示智能投顾算法的固有缺陷和使用风险，为投资者单设智能投顾帐户，明晰交易流程，强化留痕管理，严格监控智能投顾的交易头寸、风险限额、交易种类、价格权限等。针对算法的内部审核方面，欧盟金融工具市场指令(MiF ID II)要求，一家投资公司应该确保其负责算法交易风险和合规的员工具有：1) 充足的算法交易和交易策略知识；2) 跟踪自动警报所提供信息的能力；3) 算法交易造成交易环境紊乱或有疑似市场滥用时，有足够的权力去质疑负责算法交易的员工[13]。

4.3.2. 确立技术运营者对算法的说明义务和信息披露义务

说明义务和信息披露义务，在很大程度上体现了技术运营者与金融消费者之间的信息不对称。而作为技术运营者，很难有主体比其更清楚算法存在的风险和缺陷，也很难有主体比其更清楚其算法的盈利模式。算法的运作方式需要有很大的透明度，才能评估算法的政治、绩效、公平性和治理关系[25]。对于算法的说明和披露义务，美国的《智能投顾升级指导意见》(Investment Management’s Guidance Updates)规定的与算法相关的披露内容包括：管理客户帐户所使用的算法的说明；算法功能的介绍(如通过算法能对客户个人帐户进行投资和重新调整)；算法的假设和限制(如该算法是基于现代投资组合理论，说明背后的假设和该理论的局限性)；对使用算法管理客户帐户所固有的特定风险的描述(例如该算法可能不考虑市场条件而重新调整客户账户，或者进行比客户预期更频繁的调整以及算法可能无应对市场条件的长期变化)；任何可能导致用于管理客户帐户的智能投顾算法重写的状况描述(如智慧投顾可能在紧张的市场状况下停止交易或采取其他临时性防御措施)；关于第三方参与管理客户帐户的算法的开发、管理或所有权的说明，包括对这种安排可能产生的任何冲突利益的解释[26]。

4.3.3. 构建技术运营者算法赔偿基金制度

算法具有高度的复杂性，往往涉及到代码成千上万个，一般的金融消费者自然是难以理解。哪怕就是专业人士，对于机器学习算法这种具有自主决策能力的算法，其最后的运算结果并不一定处于完全掌控之中。如果是算法有缺陷，当然可以根据产品责任追究技术运营者或者技术开发者的产品责任。如果技术运营者或技术开发者因过错导致侵害金融消费者合法权益，可以追究技术开发者或者技术运营者的侵权责任。设立置算法赔偿基金制度的目的不在于对明确的产品责任或者侵权责任承担损害赔偿，而是当算法的缺陷是现有科学技术无法发现，技术开发者或运营者又根据现有法律规定不承担赔偿责任的情况下，如金融消费者因此利益受损，即可以主张由该算法赔偿基金予以补偿，从而实现弥补的正义。正如有的学者所言，人必须是算法的立法者和控制者，法律的算法与算法的法律不应成为一个死循环，它们中间必须有人作为起点和终点。要将社会生活的复杂事实带入一定的法秩序，规范塑造者需要在相关事实和基于规范文本的秩序标准之间保持“目光之往返流转”，能够做到这一点的只有训练有素的法律人[27]。

5. 结语

本文基于文献、比较的研究方法，对算法的内涵、风险进行了探讨，并在此基础上，从立法部门、

金融监管部门、金融机构的维度提出了算法的规制手段。毋庸置疑，算法不仅改变着人类社会的生产方式、生活方式，在给人类的智慧工厂、网络购物、休闲娱乐、投资理财等方面带来巨大便利的同时，算法所带来的算法黑箱的负面效应理应予以重视。尤其是金融领域，由于金融消费者的弱势地位和金融对于整个经济的核心作用，这些年来，金融机构利用算法等先进技术，基于优势地位，屡屡侵害金融消费者的知情权、选择权、信息权、隐私权等合法权益，也使金融监管机构面临难以有效监管的难题。为此，通过立法部门从法律完善的角度出发，扎牢制度的笼子，使得算法黑箱的规制有法可依；监管机构与时俱进，充分应对人工智能技术给传统金融监管带来的挑战；金融机构应主动作为，担负起算法应用的主体责任，维护金融消费者的合法权益，由此构建一个多方参与、系统协同、监管有效、保障有力的算法黑箱应对体系。正如有的学者所呼吁的，要建立算法社会的正义，所需要的不仅是简单肯定算法增益必然高于算法风险，而是要建立对算法的伦理审查机制和风险分析机制，在法律、伦理和技术多层面予以事先预防[28]。

参考文献

[1] 迈克尔·西普赛. 计算理论导论[M]. 段磊, 唐常杰, 等, 译. 北京: 机械工业出版社, 2015: 114.

[2] Diakopoulos, N. (2015) Algorithmic Accountability: Journalistic Investigation of Computational Power Structures. *Digital Journalism*, **3**, 398-415. https://doi.org/10.1080/21670811.2014.976411

[3] 程尧, 蔡一军. 人工智能背景下算法演进的风险及其法律规制——以域外模式为视角[J]. 山西大同大学学报(社会科学版), 2020(2): 33-39.

[4] 张凌寒. 权力之治: 人工智能时代的算法规制[M]. 上海: 上海人民出版社, 2021: 3.

[5] 张文显. 构建智能社会的法律秩序[J]. 东方法学, 2020(5): 4-19.

[6] 张淑玲. 破解黑箱: 智媒时代的算法权力规制与透明实现机制[J]. 中国出版, 2018(7): 49-53.

[7] Burrell, J. (2016) How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms. *Big Data & Society*, **3**, 12 p. https://doi.org/10.1177/2053951715622512

[8] 崔聪聪, 许智鑫. 机器学习算法的法律规制[J]. 上海交通大学学报(哲学社会科学版), 2020(2): 35-47.

[9] Hassanabadi, A. (2011) Viacom v. YouTube-All Eyes Blind: The Limits of the DMCA in a Web 2.0 World. *Berkeley Technology Law Journal*, **26**, 405.

[10] 刘辉. 大数据金融算法的法律规制[J]. 财经理论与实践, 2021, 42(2): 148-154.

[11] 汪庆华. 人工智能的法律规制路径: 一个框架性讨论[J]. 现代法学, 2019, 41(2): 54-63.

[12] Knight, W. (2017) The Dark Secret at the Heart of AI. *Technology Review*, **120**, 54-61.

[13] 徐凤. 人工智能算法黑箱的法律规制——以智能投顾为例展开[J]. 东方法学, 2019(6): 78-86.

[14] Lins Ribeiro, G. (2018) El precio de la palabra: La hegemonía del capitalismoelectrónico-informático y el googleísmo. *Desacatos*, **56**, 16-33. https://doi.org/10.29340/56.1875

[15] 尤瓦尔·赫拉利. 未来简史[M]. 林俊宏, 译. 北京: 中信出版社, 2017: 296

[16] 张凌寒. 算法权力的兴起、异化及法律规制[J]. 法商研究, 2019(4): 63-75.

[17] 谭九生, 范晓韵. 算法"黑箱"的成因、风险及其治理[J]. 湖南科技大学学报(社会科学版), 2020, 23(6): 92-99.

[18] 廖建凯. "大数据杀熟"法律规制的困境与出路——从消费者的权利保护到经营者算法权力治理[J]. 西南政法大学学报, 2020, 22(1): 70-82.

[19] 雷吉斯·德布雷, 赵汀阳. 两面之词: 关于革命问题的通信[M]. 北京: 中信出版社, 2014: 23.

[20] 苏宇. 算法规制的谱系[J]. 中国法学, 2020(3): 165-184.

[21] 李彦宏, 等. 智能革命: 迎接人工智能时代的社会、经济与文化变革[M]. 北京: 中信出版社, 2017: 312.

[22] 张敏, 李倩. 人工智能应用的安全风险及其法律防控[J]. 西北大学学报(哲学社会科学版), 2018(3): 108-115.

[23] 张凌寒. 算法规制的迭代与革新[J]. 法学论坛, 2019(2): 16-26.

[24] 唐林垚. 人工智能时代的算法规制: 责任分层与义务合规[J]. 现代法学, 2020, 42(1): 194-209.

[25] Brauneis, R. and Goodman, E.P. (2018) Algorithmic Transparency for the Smart City. *Yale Journal of Law & Tech-*

nology, **20**, 103. https://doi.org/10.31228/osf.io/fjhw8

[26] 邢会强, 银丹妮. 智能投顾信息披露法律制度的构建[J]. 西北工业大学学报(社会科学版), 2019(1): 82-89.

[27] 郑戈. 算法的法律与法律的算法[J]. 中国法律评论, 2018(2): 66-85.

[28] 李晓辉. 算法商业秘密与算法正义[J]. 比较法研究, 2021(3): 105-121.