

# 金融强国目标下数字金融平台算法黑箱的系统治理机制

程雪军

(同济大学法学院,上海 200092)

**摘要:**在金融强国目标与人工智能技术的驱动下,数字金融平台快速兴起并将算法技术深度应用于传统金融行业。数字金融平台是数字金融与平台经济的融合产物,其通过输出技术或提供场景与金融机构开展合作活动,具有天然的技术偏好与金融属性,有效促进金融产品与服务的数字化与普惠性。然而,任何事物都具有双面性,不是非此即彼的结果,而是利弊同在的产物。通过采用案例分析方法发现,算法技术深度应用于数字金融平台总是存在着一定的误差或测不准(算法黑箱),具有不同于传统企业的太大不能倒、太多不能倒、太强不能倒的系统性风险特征,并在三方系统(平台端、用户端与监管端)形成了算法黑箱风险;算法黑箱具有平台迷雾风险,阻碍平台端的稳定发展;算法黑箱具有技术遮蔽风险,加剧用户端的风险缺陷;算法黑箱具有监管真空风险,制约监管端的公权治理。因此,为实现金融强国的目标,防范化解数字金融平台算法黑箱的系统性风险,基于三方控制理论与国际算法黑箱风险的系统治理经验,我国应当从三方系统(平台端、用户端、监管端)构建数字金融平台算法黑箱的系统治理机制。

**关键词:**金融强国;数字金融;平台;算法黑箱;技术风险;算法透明机制;系统治理机制

**中图分类号:**D922.28 **文献标志码:**A **文章编号:**1671-4970(2024)02-0107-14

## 一、问题之缘起

作为人工智能时代最为先进的技术之一,算法在数字金融运行中发挥着重要作用,提升了金融业态的综合效率和金融服务的实际效果。然而,任何新生事物兼具促使社会经济创新发展的正外部性与衍生系列风险的负外部性:其一,算法技术有利于提升数字金融行业的效率,具有显著的正外部性。数字金融平台通过深度应用算法技术,大幅促进了金融产品、服务的数字化与智慧化发展。其二,数字金融,尤

其是算法技术具有螺旋式上升的技术特征,当数字金融平台将算法技术广泛应用于数字金融产品与服务并存在算法不当或者违法应用时(比如算法黑箱),容易滋生不可预见的系统风险,带来明显的负外部性。众所周知,法律制度,尤其是成文法律制度具有明显的时代滞后性,通常可能出现当前法律或治理制度难以跟上新兴科技的创新步伐,这便导致治理体系真空下的数字金融平台凭借算法技术从事金融活动容易触发相应的金融风险。为促进金融强国目标的实现,保障数字金融平台的创新向善,防

**引用本文:**程雪军.金融强国目标下数字金融平台算法黑箱的系统治理机制[J].河海大学学报(哲学社会科学版),2024,26(2):107-120.

**基金项目:**教育部人文社会科学研究青年基金项目(23YJC820005);上海市哲学社会科学规划课题(2022EFX001);中央高校基本科研业务费专项资金资助项目(22120230353)

**作者简介:**程雪军(1988—),男,讲师,博士,主要从事经济法、金融法与数字法研究。E-mail:chengxuejun222@126.com

范算法黑箱的风险扩大,有必要建立健全数字金融平台算法黑箱的风险治理机制。

在算法技术与平台经济的快速发展背景下,国内外数字金融平台算法黑箱的风险治理研究主要集中于以下方面,且并未形成系统的风险治理机制观点:第一,对数字金融平台算法黑箱风险的技术端研究。有些学者强调对金融算法施以代码或科技规制<sup>[1]</sup>,但算法代码可能涉及商业秘密等知识产权问题,如果要求数字金融平台完全披露其算法系统代码,那么可能会侵犯相应的知识产权保护,所以对金融算法尤其是算法黑箱的规制重点,并不是要求算法代码完全披露,而是在于算法的解释说明<sup>[2]</sup>。第二,对数字金融平台算法黑箱风险的用户端研究,侧重于算法技术侵害金融消费者权益及其保护研究<sup>[3]</sup>。有些学者专门从细分的数字金融平台(如智能投顾平台)视角切入,探讨算法技术深度嵌入细分金融领域后可能引发信义义务的扩张,从而对金融消费者权益产生影响<sup>[4]</sup>。第三,有些学者着眼于数字金融平台算法黑箱风险的平台端研究,即从算法技术兴起探讨了算法权力(比如数字金融平台的算法权力),认为这种算法权力对金融权力造成了压制作用,不利于金融公平的实现,因此需要从权力约束与权利补强的角度完善法律应对<sup>[5]</sup>。第四,对数字金融平台算法黑箱风险的监管端研究,主要强调法律治理与权益保护。其中,对于算法黑箱风险的法律监管重点在于算法透明与问责等<sup>[6]</sup>。相较于我国算法黑箱风险的监管实践,当前学术研究依然处于滞后阶段,而且没有形成系统性的风险治理方案。近年来,为加强数字平台等算法黑箱的风险治理,我国陆续颁布多部法律规范予以调整与治理:2021年,《中华人民共和国民法典》通过专章设置隐私权与个人信息保护;同年,《中华人民共和国个人信息保护法》通过专条明确个人信息处理者利用个人信息开展自动化决策的基本原则,即透明度、结果公平公正以及无差别待遇。2022年,《互联网信息服务算法推荐管理规定》构建了全面的算法审核机制,即通过主动评估、算法备案、算法公示等程序,有效防范算法黑箱风险;同年,《中华人民共和国反垄断法》规定

经营者不得利用数据、算法、技术与资本优势等从事垄断行为,并基于该条款对某些数字金融平台从事金融垄断行为开展高额行政处罚。由此可见,在算法社会背景下,对数字金融平台算法黑箱风险开展系统治理具有时代的紧迫性及其法理的必要性。

近年来,我国陆续提出金融强国的战略目标,扎实推动金融高质量发展。2023年中央金融工作会议明确提出加快建设金融强国的宏伟目标,做好以数字金融为代表的五篇大文章,全面加强金融监管,有效防范化解金融风险<sup>[7]</sup>;2024年,习近平总书记在省部级主要领导干部推动金融高质量发展专题研讨班开班式上发表重要讲话,强调坚定不移走中国特色金融发展之路,推动我国金融高质量发展<sup>[8]</sup>。由此可见,数字金融是通往金融强国目标的关键路径,现代金融监管与治理是实现金融强国目标的重要手段。唯有做好以数字金融为代表的五篇大文章,完善现代金融监管与治理,方可促进我国从金融大国迈向金融强国。然而,当前国内外研究并没有从金融强国目标下的数字金融平台角度切入,深度思考算法黑箱的系统性风险特征及其风险治理的必要性,而仅仅是从某个或某几个方面强调对科技本身的治理,且大都锁定于用户端,并未真正弥合平台端、用户端与监管端等之间的技术鸿沟与信息壁垒,尚未形成完善的算法黑箱的系统治理机制。因此,本研究从数字金融平台案例透视算法黑箱的形成机理与系统性风险特征,基于系统性风险防范与金融强国目标的视角,深度探寻数字金融平台算法黑箱风险的系统治理逻辑,从而构建我国数字金融平台算法黑箱的本土化系统治理机制。

## 二、从数字金融平台案例解构 算法黑箱的形成机理

随着金融混业经营的加深和市场效率需求的强化,为保持经营规模与行业竞争地位,金融机构争相发展数字金融,为科技布局金融业创造了无限的空间。作为汇聚数字金融的实体,数字金融平台将算法等技术在金融市场全面铺开,并给传统金融业态带来了不可逆的质变影响。通过采用案例研究方法,从国内某知名数

字消费金融平台案例切入发现,数字金融平台深化算法技术应用,掌握了大数据、算法等各类先进技术,促进了金融行业的创新发展,但数字金融平台算法技术应用具有技术不可知性、信息不对称性特征,同样衍生了算法黑箱。

1. 从数字金融平台案例透视算法技术的深度应用

在算法社会背景下,什么是算法?从语义学角度看,算法是一种“计算的方法”;从计算机学角度看,算法一般被视为在有限且明确的操作步骤内,将形式规范的输入值转化为输出值,高效解决某特定问题的模式化过程<sup>[9]</sup>。什么是数字金融平台?不同时期数字金融平台因信息技术的深化而呈现不同的发展形态,从互联网信息技术下的互联网金融平台发展至大数据技术下的金融科技平台,并进一步发展至人工智能算法下的数字金融平台。无论数字技术如何深化演进,数字金融平台的本质都是数字金融与平台经济的融合产物,其通过输出技术或提供场景与金融机构开展合作活动,具有天然的技术偏好与金融属性,能有效促进金融产品与服务的数字化与普惠性。其一,数字金融的技术基石是大数据、人工智能等信息技术。数字金融最早源于美国,并在英国、中国、日本等快速发展。其中,英国早在2011年便将数字金融(尤其是金融科技)作为国家发展战略,金融行为监管局(FCA)于2015年推出“监管沙

盒”政策,旨在激发数字金融创新与合规发展。其二,数字金融平台作为连接人与资金的互联网平台经济模式,充分发挥投融资功能与作用,深度嵌入各式各样的金融服务行业,具有天然的技术偏好与金融属性,形成了诸如海外Paypal、Stripe与国内腾讯、蚂蚁等数字金融平台。质言之,数字金融平台是掌握先进的信息技术以及雄厚的金融资源,通过输出技术或提供场景与金融机构开展合作活动的平台实体。

为何数字金融平台广泛应用算法技术于金融场景?这是因为算法技术可从技术必要性与可行性两个层面促进数字金融平台的市场竞争与创新发展,助力数字金融平台的权力兴起:一方面,从算法技术必要性看,数字金融时代下的算法技术具有先进性,代表了当下最先进的技术生产力。数字金融平台的本质在于技术驱动的金融产品与服务创新,与技术具有天然的深度耦合性,而算法技术是人工智能时代最具有代表性的科技方式,是社会权力运行中的重要变量<sup>[10]</sup>,因此数字金融平台有必要深度应用算法技术提升其金融创新效率与服务效果。另一方面,从算法技术可行性看,凭借数据、算力与算法优势,通过深度应用算法技术(如贝叶斯、随机森林、人工神经网络等)于更广更深的金融场景,数字金融平台可根据“输入→执行指令→输出”运行模式提升综合效率。如图1所示,以国内某知名的数字金融平台为例(该数

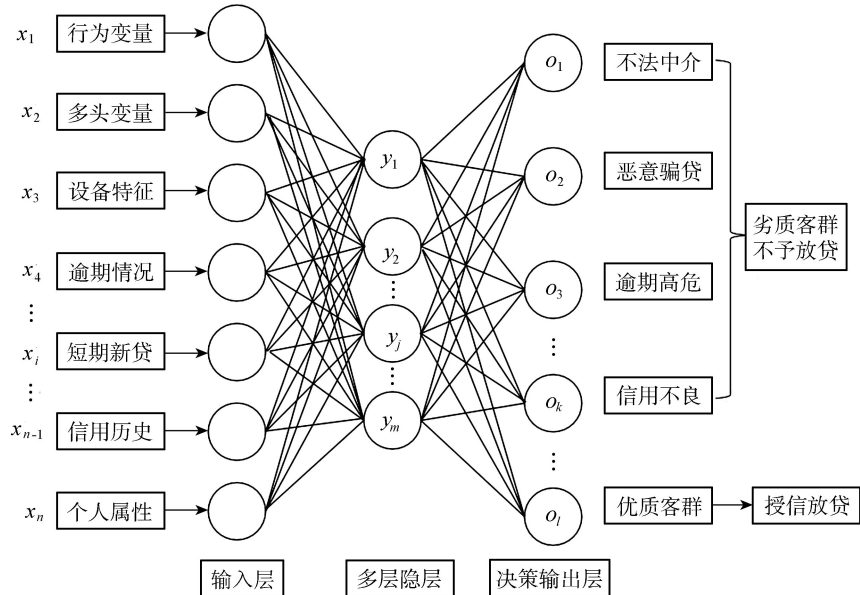


图1 从数字消费金融平台透视算法黑箱的形成机理



字金融平台的主营业务是消费金融业务,可称之为数字消费金融平台),此数字消费金融平台通常根据消费者的基本信息、金融信息及其他信息等输入客户属性,以人工神经网络、深度学习等算法为技术基础研发算法决策系统并开展消费金融活动,输出客户类别并判定是否对消费者授予信用贷款及授予多少额度与利率的信用贷款。最终,在算法技术的驱动下,数字消费金融平台的运营效率是传统消费金融公司的数倍。

2. 从数字金融平台案例透视算法黑箱的形成机理

从数字金融平台算法模型的形成机制看,其是根据已知的输入与中间运行程序完成一系列操作并进行结果输出的机制。但中间运行程序如何运算却难以被用户端甚至平台端获悉,输出的结果也难以被他人控制和预料,天然具有较强的黑箱特质,容易衍生算法黑箱风险。一旦开发者(如数字金融平台)通过算法与数据建构、测试与改进模型,甚至在算法黑箱中增加过度的商业追求、价值偏好与不公正因素,那么凭借算法侵蚀平台端的其他经营者与消费者的情况将变得日益普遍。

同理,从数字消费金融平台实际案例看,其算法运行一般包括三步流程(图1):第一步,数字消费金融平台通过人工神经网络算法输入客户属性,深度挖掘与分析消费者的海量数据,这些数据包括行为变量、多头(信贷)变量、设备特征、逾期情况、信用历史及个人属性等。此时,通过合法合规的方式获取消费者的大量数据可视为输入层,其向量可表示为  $\mathbf{X}=(x_1, x_2, \cdots, x_p, \cdots, x_n)^T$ 。第二步,数字金融平台驱动计算机自动编写代码程序并开展算法决策,从而研判是否应当给予消费者消费信贷。该阶段采用的数据驱动型算法较多,包括人工神经网络、随机森林、贝叶斯等,本案例采用的人工神经网络算法可被视为隐层,其输出向量为  $\mathbf{Y}=(y_1, y_2, \cdots, y_j, \cdots, y_m)^T$ ,隐层数量  $m$  可以不受限制的扩展<sup>[11]</sup>。第三步,数字金融平台通过人工神经网络算法完成金融活动,体现了强大的机器学习功能,并输出相应的客户类别,包括不法中介、恶意骗贷、逾期高危、信用不良、优质客群

等,唯有被认定的优质客群才具有消费金融授信资格,其他类型的客户类别则不具有消费金融授信资格。在数字消费金融平台,输出层向量为  $\mathbf{O}=(o_1, o_2, \cdots, o_k, \cdots, o_l)^T$ ,其中,样本容量  $n>m>l$ <sup>[12]</sup>。在算法程序的运行流程中,由于中间执行环节往往难为外部所知情,所以输入层向量  $\mathbf{X}$  与输出层向量  $\mathbf{O}$  之间总是存在着一定的测不准,因而中间运行层被命名为算法黑箱,而输出层所输出的客户类别常常具有不准确性或误差,即真正的优质客群没有通过数字金融平台的审核,而劣质客群却被数字金融平台划入优质客群并且对其授信放贷,从而加大了数字消费金融平台的风险。质言之,算法社会下的数字金融平台算法黑箱反映了信息悖论,即数据快速延伸,但只有局内人才可获悉重要信息,可能导致黑箱社会<sup>[13]</sup>。

三、金融强国目标下数字金融平台  
算法黑箱的系统治理逻辑

在全面建设中国特色金融发展之路,提出加快建设金融强国,有助于我国从金融大国向金融强国迈进,提高全要素生产率并适应现代市场经济发展<sup>[14]</sup>。毋庸置疑的是,数字金融平台是促进金融强国目标实现的关键手段:一方面,数字金融平台具有建设性创新的作用,即其利用算法技术实施创新活动,有助于全面提升金融产品与服务的综合效率;另一方面,数字金融平台具有破坏性创新的作用,即其具有平台信息不透明性、技术不成熟性及监管不完善性等特性,从系统层面无形中放大平台端、技术端与监管端的风险,进而可能衍生相关的算法黑箱风险(平台迷雾风险、技术遮蔽风险、监管真空风险)。因此,基于系统性风险防范以及金融强国目标的要求,我国有必要从数字金融平台算法透明、技术风险防范、推进公权治理等层面,加强数字金融平台算法黑箱的系统治理,从而矫正数字金融市场无序行为并祛魅平台算法黑箱风险问题(图2)。

1. 数字金融平台算法黑箱的系统性风险特征

数字金融平台是数字金融与平台经济的融合产物,是数字金融驱动下的创新平台,其基础



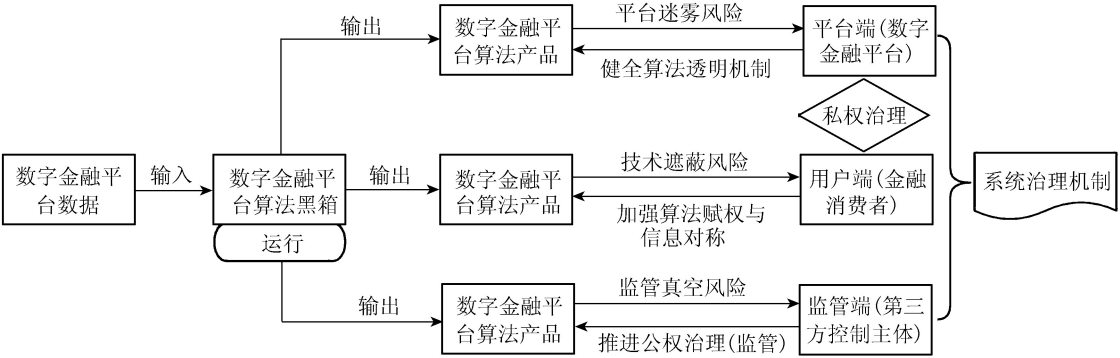


图 2 数字金融平台算法黑箱风险的综合治理逻辑

构成是虚拟世界中的比特,而非物理世界中的原子。与原子相比,比特具有得天独厚的技术优势:在原子构成的产品与服务中,其生产、运输以及其他成本很高;在比特构成的产品与服务中,其可以低成本快速复制,以光速快速传播并永不变质<sup>[15]</sup>。数字金融最早源于美国,并在英国、中国、日本等国家快速发展,随着机器学习、深度学习为代表的算法广泛应用于金融服务行业,金融服务市场逐步形成了数字金融平台。

根据《互联网平台分类分级指南(征求意见稿)》,我国将平台分为超级平台、大型平台与中小平台<sup>[16]</sup>。同理,参照平台分级标准,可将数字金融平台划分为三类平台:超级数字金融平台、大型数字金融平台与中小型数字金融平台。尤其是,超级数字金融平台作为兼具超大用户规模、超广业务种类、超高经济体量、超强限制能力(超强的限制商户接触消费者的能力)的数字金融平台,具有内在的风险巨大性,容易滋生系统性风险,并显现出3个系统性风险特征(表1)。第一,太大不能倒,即用户规模与经济体量太大而不能倒下的风险。数字金融平台是典型的平台经济模式,具有双边市场甚至多边市场的特性,有效连接众多的商户与消

费者,所以大部分数字金融平台的用户规模都比较庞大,并基于庞大的用户规模、先进的算法技术、雄厚的数据资源等优势,发展出了较高的经济体量。其中,超级数字金融平台的年活跃用户数不得低于5亿人,市值或估值不得低于10000亿元,如海外的Paypal与国内的腾讯金融等,由此都易滋生因为平台太大而不能倒下的风险。第二,太多不能倒,即业务太多而不能倒下的风险。数字金融平台通过深度应用算法技术于金融产品与服务的方方面面,驱动各式各样的金融业务混合化发展,形成了相当数量的核心金融业务。此外,众多金融业务都是因为技术驱动而形成的创新型金融业务,如算法信贷业务、量化基金业务、网络支付业务等。其中,超级数字金融平台具有两类甚至更多种类不同的核心金融业务,业务种类众多且混业化发展,从而不太可能轻易倒下或破产。第三,太强不能倒,即算法技术与市场能力太强而不能倒下的风险。不同于传统市场经营者,数字金融平台兼具雄厚金融与先进科技等优势,而且这两种优势可以完美地融合在一起,因为金融与科技的实质都是数字,所以数字金融平台在市场竞争中具有天生的优势,通过不断积聚资源形成强大的算法技术能力与市场垄断能力,

表 1 数字金融平台分类及其主要特征

| 平台分级     | 平台代表          | 分级依据及其具体标准             |                       |                 |                         |
|----------|---------------|------------------------|-----------------------|-----------------|-------------------------|
|          |               | 用户规模                   | 经济体量                  | 业务种类            | 限制能力                    |
| 超级数字金融平台 | Paypal、腾讯金融   | 超大用户规模<br>(上年度活跃用户超5亿) | 上年底市值(估值)<br>超10000亿元 | 核心业务种类<br>两类以上  | 超强地限制商户接触<br>消费者(用户)的能力 |
| 大型数字金融平台 | Stripe、美团金融   | 较大用户规模<br>(活跃用户超0.5亿)  | 上年底市值<br>(估值)超1000亿元  | 具有表现突出的<br>主营业务 | 较强的限制商户接触<br>消费者(用户)的能力 |
| 中小数字金融平台 | Ripple、建信数字金融 | 国内具有一定的<br>年活跃用户       | 具有一定的市值(估值)           | 具有一定业务          | 一定的限制商户接触<br>消费者(用户)的能力 |

限制其他市场经营者进入相关数字金融市场,从而排除限制市场的公平竞争。质言之,基于自身的资源优势,数字金融平台形成了太大不能倒、太多不能倒、太强不能倒的风险特征。参照梅特卡夫定律,基于互联网技术而创生的数字金融平台风险可以表述为  $Risk = \alpha n^2 + \beta$  (其中  $Risk$  是风险,  $n$  是用户数量,  $\alpha$  与  $\beta$  为常数),这样的风险是指数量型放大而非非线性增长的,具有风险巨大性与传导性<sup>[17]</sup>,容易诱发系统性风险。

## 2. 数字金融平台算法黑箱的三方系统风险及其治理逻辑

当数字金融平台深度应用算法技术时,算法的建设性创新属性有利于数字金融平台的创新发展,但算法技术的破坏性创新属性同样衍生数字金融平台的算法黑箱风险,集中体现为平台迷雾风险、技术遮蔽风险及其监管真空风险。因此,为防范数字金融平台算法黑箱风险可能衍生的系统性风险,有必要对其加强系统的风险治理。

### (1) 算法黑箱具有平台迷雾风险,阻碍平台端的稳定发展

在信息技术与市场需求的双轮驱动下,传统金融逐步发展至数字金融阶段,数字金融平台的平台端、技术端与用户端等快速兴起。算法黑箱将算法技术应用中的复杂模型与代码等隐藏起来,导致接收者无法获取产生重要影响的信息,只能直接进入决策执行环节获取简要结果。尽管算法金融应用模型看似完美,但因为缺少必要的、制衡性的治理机制,导致某些数字金融平台试图利用算法黑箱从事非法金融活动。与主动投资策略不同,量化交易策略普遍不依赖于个人判断,而是通过复杂的数学模型预测投资标的未来价格,但是人为算法程序的主观性与偏好性常常滋生相应的算法金融风险事件。2023年美国量化巨头 Two Sigma 透露公司高级副总裁在未经授权的情况下擅自调整交易模型,其目的是让自己获得更高的奖金收入,却导致了部分基金发生 1.5 亿美元的额外损失,而这些基金的损失主要由用户承担。同样地,中国也发生过多起平台算法金融风险事件,比如因为算法失误导致 A 股证券市场发生 6%

左右剧烈震荡的光大证券乌龙指事件、因为期货公司巨额平仓交易导致的中证 500 股指期货合约(1C1706)剧烈震荡的期货乌龙指事件等<sup>[18]</sup>。

围绕算法技术而形成的算法黑箱具有平台不透明(迷雾)风险,阻碍平台端的稳定发展:一方面,作为算法的开发者兼运营者,数字金融平台可以对算法技术及其程序施加某种主观性偏见,利用算法黑箱拉大平台端与用户端的信息差距。其中,最典型的主观算法偏见是算法价格歧视,与传统价格歧视仅限于简单事实判断不同,算法价格歧视因为算法黑箱的不透明性与难理解性而变得更加隐蔽且复杂,它是以海量数据与算法模型为基础而表现的主观歧视性算法行为,所以数字金融平台更关注个人信息收集以及算法模型优化,从而提升数字金融平台的算法精准度。另一方面,数字金融平台算法技术的复杂性在客观上隐藏了施加行为或算法机械性缺陷,导致金融类的违法犯罪活动在黑箱庇护下乱象丛生,而且难以被金融治理或监管机构发现与治理。质言之,在算法治理体系尚未系统化与完备化的情况下,一旦算法黑箱掩盖下的行为触发了金融安全风险,长尾用户的敏感度与应对措施具有明显的滞后性,监管部门在事后迅速难以通过非禁止措施纠偏,可能对平台端的稳定发展带来严重的负面影响。

### (2) 算法黑箱具有技术遮蔽风险,加剧用户端的风险缺陷

金融市场一直都是科技驱动型的市场,从信息技术驱动的传统金融到互联网技术驱动的互联网金融,并进一步发展至数字金融驱动的算法金融,金融发展每个阶段的核心生产要素都是技术。诚如柯布-道格拉斯生产函数所揭示的,技术在生产函数中发挥着至关重要的作用,即  $Y = AF(K, L) = AK^\alpha L^{(1-\alpha)}$  ( $0 < \alpha < 1$ ),产量( $Y$ )与技术要素投入量( $A$ )呈正相关关系,同时与劳动要素投入量( $L$ )与资本要素投入量( $K$ )正相关,而且  $\alpha$  表示劳动所得在总产量中所占的份额<sup>[19]</sup>。不过,在金融市场逐步达到全面科技化的阶段,技术的创新发展速度往往超越了大众认知程度,尤其是算法技术的深度开发和

多场景应用加剧了认知差异。对于算法技术不成熟的风险问题,构建完备的风险治理体系可谓势在必行。

其一,数字金融平台算法黑箱的技术遮蔽性,导致金融风险的爆发更加频繁。技术要素的不成熟与不可知通常会成为引发金融风险的导火索,从2000年的互联网泡沫危机到2008年的次贷危机,无一不是因技术不成熟所引发的金融危机。同理,在当前算法技术的助推下,数字金融平台得以快速地创新迭代,但是算法技术的先进引领性与社会大众的认知滞后性具有显著的矛盾与差距,进而导致数字金融平台的风险发生更具不成熟性。传统金融风险论认为金融风险是相关权益蒙受损失的可能性,收益与风险具有相关关系,享有相应的“风险溢价”,即 $E(r) = wE(r_f) + (1-w)r_f = rf + w[E(r_f) - r_f]$ ,其中, $E(r)$ 为投资组合的期望收益率,无风险资产的收益率为 $r_f$ ,风险资产的期望收益率 $E(r_f)$ ,投资于风险资产的比例为 $w$ <sup>[20]</sup>。在数字金融平台场景下,算法黑盒模型使得用户端无从得知做出算法决策的依据性以及执行效果,加剧了数字金融市场运行的不成熟性并削弱了金融风险的可衡量性。因此,基于数字金融平台的破坏性创新以及算法黑箱导致金融风险的发生更具有不成熟性,非常有必要对数字金融平台算法黑箱风险采取系统化治理,以防范技术不成熟性所导致的金融风险蔓延。

其二,算法黑箱具有信息不对称风险,损害用户端的合法权益。在假设的虚拟世界中,任何商品与服务的信息皆具有完全性,其获取的成本较低。但是在现实世界中,信息和普通商品一样都是有价值的资源,且现实世界并不是虚拟世界中的完全竞争模型,市场的供求双方并不具备完全的信息,双方获取商品与服务的信息可能花费很高的代价,届时信息的不完全与不对称性将会引起相应的逆向选择与道德风险等问题,致使市场机制难以有效发挥作用<sup>[21]</sup>。在数字金融平台的背景下,由于算法黑箱的遮蔽性所带来的信息不透明,致使黑箱场域内的算法决策实际上具有严重的信息不对称,各种算法黑箱与算法歧视等问题屡见不鲜<sup>[22]</sup>。不同于传统金融业态下信息不对称所

引发的市场失灵(即一方占有较大优势从而导致交易阻碍),算法黑箱对传统金融信息不对称边界进行了扩张,促使平台端与用户端之间的信息差越拉越大,届时数字金融平台可以实施金融监管套利空间愈来愈大。在算法技术背景下,平台端的数字金融平台比用户端的金融消费者具有更加充分与完整的信息,容易发生损害用户端合法权益的行径:一方面,数字金融平台可能利用算法的复杂性隐藏定价规则,对不同用户群体开展“千人千面”的画像,从而基于不同的标签实施差别定价;另一方面,数字金融平台可能利用双方之间的信息不对称性,虚构交易误导和欺骗用户,损害金融公平与正义。无论在传统金融时代抑或数字金融时代,信息不对称始终是金融市场的永恒话题,因为其直接关涉数字金融平台端与用户端的信息归属与地位情况。数字金融平台对算法技术的违法或不当使用滋生了算法黑箱风险,严重加剧了信息不对称,导致金融风险逐步累积并偏离预期。数字金融平台应当以信息对称理念为核心,通过构建可解释算法模型等KYC(认知你的客户)路径,逐步完善数字金融平台算法黑箱的风险治理,破解平台端和用户端之间的信息不对称,方可更好地促进金融稳定与保护用户权益。

(3)算法黑箱具有监管真空风险,制约监管端的公权治理

任何新生事物的快速发展都需要具备核心的发展要素,监管或者说制度建设,对于新生事物的发展起着至关重要的作用。然而,数字金融平台算法技术创新与应用的速度过快,导致相应的监管制度并没有及时跟上,引发了一系列因监管不完善而带来的算法黑箱风险,严重影响了数字金融平台以及数字金融市场的稳健发展。由于市场结构(包括数字金融平台在内的市场)往往不是完全竞争的,因此企业行为存在低效率现象,此时以政府监管与社会监督为主的监管(这里的监管主要是指除私权治理之外的公权监管或治理,主要包括社会公众、社会组织与政府机构等第三方控制主体)就是纠正企业低效率行为的一种方式<sup>[23]</sup>。

为防范算法黑箱的监管不完善风险,我国多个监管机构发布了相关监管指引。其一,陆



续出台多部法律规范性文件。如,中国人民银行就陆续发布多项规范及政策;2021年发布《人工智能算法金融应用评价规范》,提出建立人工智能金融应用的算法评价框架,从可解释性、精准性和安全性等方面提出具体要求;2022年印发《金融科技发展规划(2022—2025年)》,强调数字金融的应用及其数据问题的监管要求,为数字金融平台算法黑箱风险的监管路径提供了原则性的指引方向;2023年发布《银行业普惠金融业务数字化模式规范》,明确提出应用深度学习、神经网络、时序分析等算法开发欺诈风险监测模型,从而防范相应的算法金融风险。其二,逐步制定数字金融算法伦理指引。中国人民银行于2022年颁布的《金融领域科技伦理指引》在两大方面提到了算法伦理的具体要求:一是强调包容普惠,防止不公平歧视,本指引明确提出降低数据驱动、算法驱动决策导致的不公平结果……坚决抵制利用技术优势从事算法歧视等不当行为;二是强调公平竞争,重点在于公平公正使用智能算法。同年,中共中央办公厅和国务院办公厅发布的《关于加强科技伦理治理的意见》提出增进人类福祉、尊重生命权利、坚持公平公正、合理控制风险、保持公开透明等共同构成科技伦理原则。然而,这种主要由政府监管机构颁布的原则性指引并没有构建起完善的协同监管体系,亦没有出台明确的实施细则,普遍缺乏系统化的公权监管或治理机制,因此难以破解数字金融平台算法应用中的黑箱风险问题。

#### 四、构建金融强国目标下我国数字金融平台算法黑箱的系统治理机制

为实现金融强国目标,从金融大国向金融强国迈进,我国有必要从防范数字金融平台算法黑箱的系统性风险切入,平衡算法黑箱内外部之间必要的信息流通,有效保障数字金融产品交易公平以及适配有限的金融服务资源,满足金融消费者的切实需求及权益保护。具言之,我国可以从三方控制理论出发,从平台端(第一方控制)、用户端(第二方控制)与监管端(第三方控制)这3个层面构建本土化数字金融平台算法黑箱风险的系统治理机制。

##### 1. 从三方控制理论重塑数字金融平台算法黑箱的系统治理理念

关于如何构建全面的系统治理,三方控制主体(五种控制者)可作为制定规则并实施制裁手段的社会控制主体,从而形成全面社会控制体系:第一方控制是行动者通过自律规则实施自律监管,其主体是强调自我制裁与自我控制的行动者(它是一种体现为私权治理的主体)。第二方控制是根据合约的行动者通过合同约定实行合同双方或多方的控制,其主体是强调有约必守的个人自助合约者<sup>[24]</sup>(私权治理主体)。第三方控制是未经行动者同意而实施的社会控制,依据社会控制规则与手段的差异性,第三方控制主体包括三类社会控制者,分别为社会公众、社会组织与政府机构,因为社会公众与社会组织可归类为社会主体层面,可以统一为社会监督层面的治理<sup>[25]</sup>,所以它们与政府机构共同构筑为第三方社会控制主体(公权治理主体)。其中,社会公众与社会组织分别依据社会规范与组织规则实行社会约束和组织执法,政府机构通过颁布法律规则(表现为立法、执法与司法权的行使)实施社会治理,集中体现为采取政府执法的治理手段<sup>[26]</sup>。

目前,海外发达经济体对数字金融平台算法黑箱的风险治理经验比较成熟,集中体现为3种个性化的风险治理范式,但普遍呈现出三方控制理论下共性化的风险治理格局:第一,欧洲地区强调数据治理框架下的个体赋权体系,采取个体赋权与数据处理流程控制的方式<sup>[27]</sup>。欧盟通过颁布《通用数据保护条例》强调对数据主体的用户端赋权以获得选择空间,围绕数据和算法等构建新兴主体权利(如知情权、访问权、更正权、删除权、数据可携权等),助力数据主体获得影响与控制,从而对自动化算法决策施加干预的行动空间<sup>[28]</sup>。在数字金融平台算法黑箱风险的治理主体上,欧盟注重多元治理或协同治理,其治理主体包括欧洲行政机构、金融监管机构(银行业管理局、证券和市场管理局、保险业监督机构)以及其他专业部门与专家机构等。第二,美国强调公共治理框架下的外部问责体系<sup>[29]</sup>,重视协同共治体系。算法黑箱系统几乎没有透明度可言,并且常常以技

术为借口逃避对算法黑箱输出结果的责任承担。数字金融平台使用复杂而秘密的算法黑箱,将数十亿用户的个人信息数据转变为数千万亿美元的利润,尽管算法系统为社会生活带来巨大的便捷,但是人们对其如何创设与运作的理解较为狭隘<sup>[30]</sup>。因此,美国 2022 年通过《算法责任法案》强调对自动化决策系统的过程公开与透明,其关键透明度措施有望改变算法黑箱状况。在算法治理主体层面,美国强调第一方控制(企业自治)、第二方控制(用户与消费者等)及第三方控制(社会组织、第三方力量、G20 与世界银行等国际组织)等主体共同构筑的协同治理。第三,其他国家和地区融合欧美发达经济体对算法黑箱风险的治理经验,初步提出了综合化治理原则。其中,新加坡金融管理局提出了公平、道德、问责与透明原则<sup>[31]</sup>;2023 年中美等 28 个国家正式发表《布莱切利宣言》,同意通过国际合作构建人工智能监管办法,呼吁和倡导以人为本,希望科研机构、企业等以负责任的方式,设计、开发和使用人工智能算法。因此,针对数字金融平台算法黑箱风险的治理,我国可以从第一方控制(平台端)、第二方控制(用户端)与第三方控制(监管端)层面切入,推动形成安全可控、开放协作、包容共享、共担责任的算法黑箱治理的新格局(图 3)。

2. 构建金融强国目标下数字金融平台算法黑箱的系统治理路径

从基础理论到实践场景的落地,需要穿越双层系统性迷雾:一是如何从基础理论提炼基本原则的可落实性,从而引导实践工作;二是如何从横向历史与纵向国际经验抽象基本路径的

可行性,通过法律继承或者法律移植等方式指导本土化实践工作。为防范数字金融平台算法黑箱风险,我国应当立足于基础理论(三方控制理论)与国内外算法风险的系统治理经验,从基本原则与进路层面构建本土化数字金融平台算法黑箱的系统治理路径。

(1)秉承公开透明原则,从平台端健全数字金融平台的算法透明机制

对于大多数算法系统而言,其普遍是不透明的,这意味着其可以规避审查。而且,大多数平台端(包括数字金融平台在内)并不愿意公开算法系统的源代码,导致用户端无法看清算法规则、提出异议与参与决策<sup>[32]</sup>,他们面对的是一个黑箱般的算法系统。因此,对于算法技术深度应用而引发的算法黑箱风险,作为平台生态体系中的强势法律主体,数字金融平台应当遵守法律法规、章程和伦理准则,开展数字金融平台内部的企业自治,尤其是需要秉承公开透明原则,从算法解释与算法评估等角度健全算法透明机制。

一方面,数字金融平台需要增强算法解释义务。数字金融平台端作为算法决策的实施者与行动者,理应对作为算法决策相对人的数字金融平台用户端做出解释说明;后者同样有权知晓算法决策者的金融算法实施机制,要求数字金融平台端履行算法解释义务,从而防范算法黑箱风险。一是数字金融平台应当对算法系统进行说明与解释。《中国人民银行金融消费者权益保护实施办法》指出金融机构对金融产品和服务进行信息披露时,应当使用有利于金融消费者接受、理解的方式……对其中的关键的专业术语进行解释说明。对收集与使用用户端

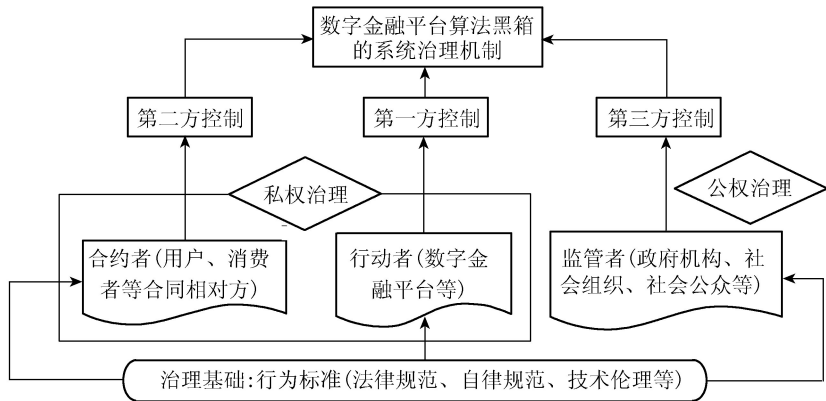


图 3 三方控制理论下数字金融平台算法黑箱的风险治理体系

数据的范围、内容以及用途等方面,数字金融平台有必要开展解释说明,通过借鉴已有成熟互联网平台的经济模式,提升数字金融平台算法使用的透明性与责任<sup>[33]</sup>。二是金融治理部门应当要求数字金融平台对算法系统进行说明与解释。作为算法的设计者与应用者,数字金融平台从业的底线是必须遵循算法相应的法律规范与道德、科技伦理等,不能通过算法形成巨大的技术优势后,从而以算法从事算法黑箱、算法歧视、算法权力等不公平、不正当行为损害数字金融算法产品与服务相对人的合法权益。针对数字金融平台算法黑箱风险,金融治理部门可以制定相应的解释标准规范,有权要求数字金融平台对算法所遵循的程序和决策全面解释,有效提升数字金融平台算法解释的综合水平。

另一方面,对于数字金融平台算法黑箱的风险治理,需要增强数字金融平台的算法黑箱的算法评估。对于数字金融平台算法评估而言,其有利于构筑一种良性的环境,旨在建立较少“有毒”或“黑箱”化的自动化算法决策系统,有必要鼓励数字金融平台根据企业客观情况,制定契合企业发展的算法标准与管理制度,对算法黑箱研究、生产或服务实施风险评估,强化企业内部层面的合规管理,抓好企业内部的“三道红线”建设,防范潜在的算法黑箱风险。在算法技术背景下,域外发达经济体普遍强调对包括算法黑箱风险在内的算法风险实施分级分类治理。其中,欧盟议会《算法责任与透明治理框架》提出建立算法影响评估(AIA)机制;欧盟委员会于2021年发布了首个欧洲AI法律框架提案,基于AI系统功能与具体用途等视角,将AI应用风险分为四级风险(不可接受的风险、高风险、有限风险和低风险),并强调分级分类治理<sup>[34]</sup>;德国算法评估方案的核心设想在于构建数字服务企业(包括数字金融平台在内)使用数据或算法的五级风险评级制度(具有较低潜在危害的系统、具有潜在危害的系统、具有一般或明显危害的系统、具有相当潜在风险的系统、具有潜在不合理危险的系统),对不同风险类型的数字服务企业采取不同的治理措施<sup>[35]</sup>。质言之,在数字金融算法应用初期,需要审慎评估算法系统对金融市场的影响,防范

系统性金融风险<sup>[36]</sup>。当前,我国需要从国家与地方层面出台相关的法律规范,强调对人工智能算法实施算法评估与分级治理:由国家互联网信息办公室等于2023年颁布的《生成式人工智能服务管理暂行办法》在第3条、第6条与第16条,多次提及对生成式人工智能(AIGC)实行包容审慎和分级治理;《上海市促进人工智能产业发展条例》在第65条明确提出了探索分级治理,即对高风险的人工智能产品和服务实行清单式管理,对中低风险的AI产品和服务采用事前披露和事后控制的治理模式。因此,针对数字金融平台算法黑箱的风险治理,我国可对高中低风险等级采取差异化的算法评估与分级治理:一是对潜在风险较高、敏感程度较高的算法金融产品应用赋予高风险定位,采取事前评估和风险预警的治理模式,但应遵循必要、正当与可控等原则开展合规审查,避免发生运动式治理;二是对潜在风险较低、敏感程度较低的算法金融产品应用赋予中低风险定位,采取事前审批和事后控制的治理模式,在创新发展和法治规范相结合的发展理念下,通过采用监管沙盒等鼓励数字金融平台算法技术的先行先试。

(2)坚持技术治理原则,从用户端加强数字金融平台的算法赋权与信息对称

数字金融平台算法黑箱风险的本质原因在于技术遮蔽风险,算法黑箱可以被刺破,但需要注意的是技术细节。对此,我国可以基于法律治理尤其是技术治理视角,从用户端加强数字金融平台用户的算法赋权(集中体现于算法知情权与同意权),完善数字金融平台算法系统的信息对称,规避算法黑箱的用户端风险。

当强势的平台端与弱势的用户端开展金融交易活动时,用户端常常处于不利的位置,主要原因在于双方之间的信息不对称。在现实社会生活中,各类数字金融平台通过算法技术获取用户信息,并向用户精准推送数字金融产品与服务,导致用户被困在算法系统中,对用户端(金融消费者)的合法权益造成实际侵害。以某些数字技术驱动的数字金融平台为例,它们通过算法技术与模型吸引用户并开展综合的算法金融业务,提高用户的依赖性与高转换率,绑定用户的算法金融需求,但因为违反公司治理、



消费者保护等而对用户合法权益带来较大侵害,因而被金融监管部门高额处罚。不过,从全国范围来看,我国对于数字金融平台通过算法技术与模型推荐产品与服务的治理力度较轻,亟待增强相应的风险治理。一方面,从用户端赋予金融消费者的算法知情权,包括可视化技术等方式。基于金融知识的差别、金融服务者的服务操守等因素,金融交易的风险远高于普通商品交易<sup>[37]</sup>。数字金融平台作为金融发展较高阶段的平台实体,同时也是数字金融算法产品的开发者与运营者,应当在用户使用数字金融算法产品之前,向用户告知金融算法产品与服务的具体信息情况。欧盟 GDPR 第 13 条规定数据主体有权知悉其信息被处理的主体身份、使用目的以及法律依据等情况,GDPR 在背景引言中对信息披露提出数据控制者应当以易见、易懂、易读的方式(包括可视化技术等方式)提供数据与算法影响类型示例<sup>[38]</sup>。《中国人民银行金融消费者权益保护实施办法》指出,金融机构应当依据金融产品或者服务的特性,及时、真实、准确、全面地向金融消费者披露重要内容。算法知情权应当包括社会公众从任何合法渠道获取信息的权利,包括政府机构和非政府机构的渠道。在数字金融平台背景下,我国应当同样赋予金融消费者相应的算法知情权,避免因信息不对称而导致其他不必要的社会交易成本,方可防范数字金融平台算法黑箱的信息不对称风险,保护金融消费者的合法权益。另一方面,从用户端赋予金融消费者的算法同意权。GDPR 第四条规定数据主体通过一个声明或某项清晰的确信行动而自由做出的、充分知悉的、不含混的、表明同意对其相关个人信息进行处理的意愿<sup>[39]</sup>。《中华人民共和国个人信息保护法》第十三条规定,个人信息处理者在取得个人同意的情形下可以处理个人信息,此时用户具有同意或者不同意的自由选择的权利。因此,数字金融平台的用户端应当对金融算法决策过程具有算法知情权并做出自由选择的权利(算法同意权)。

(3)恪守协同监管原则,从监管端推进数字金融平台算法黑箱的公权治理

金融监管有狭义和广义之分,前者是指金

融监管机构依据法律规范对金融业实施的监督管理,后者还包括金融机构的内部自治、同业自律性组织的监管、社会中介组织的监管等内容。其中,第一方控制者(平台端,如金融机构)对算法黑箱风险的内部自治以及第二方控制者(用户端,如金融消费者)对算法黑箱风险的依约治理都属于私权治理,其他第三方控制者(政府监管机构、社会公众、社会组织等)对算法黑箱风险的监督管理属于公权治理。

首先,在政府监管层面,我国需对数字金融平台实施算法行为采取包容审慎的金融监管理念,通过创新金融监管模式与金融监管工具,比如采用试点监管、沙盒监管等监管工具,不断完善数字金融平台算法决策系统的政府监管,从而在政府监管层面防范算法黑箱风险。其中,沙盒监管是针对数字金融平台应用的算法技术进行深度测试的监管机制,引导数字金融平台发现算法问题、创新算法技术、降低算法黑箱风险,实现数字金融平台算法技术创新与风险防范的动态平衡。此外,还需要从政府监管层面强化数字金融平台算法黑箱的科技伦理建设。参考科技部的监管政策,我国可在全国与地方层面设立人工智能伦理委员会,明确人工智能伦理委员会的职责,加强伦理委员会对数字金融平台算法黑箱风险的统筹规范和指导协调,推动构建覆盖公平公正、可控可信、规范有序等的人工智能伦理治理规则。

其次,在行业自律层面,数字金融行业尤其是金融算法行业组织应当依照法律规范和章程规定,积极开展行业自律规范:一方面,我国需要在全国范围内鼓励数字金融行业组织制定算法相关标准、技术指南、设计准则等自律规范,建立适应算法产业发展的标准体系,在人脸识别、金融服务、金融安全等重点领域启动标准编制工作,并且有效组织宣传培训活动,引导相关数字金融平台依法合规经营。另一方面,我国需要在地方范围内根据国家层面的行业组织结合本地发展情况建设配套的行业自律组织,由其主持基础共性、技术标准、行业应用标准、安全伦理等领域的规则制定和管理工作。行业自律规范的有效试验与稳健运行,有利于为数字金融平台算法黑箱提供丰富的风险监管经验。

最后,在社会监督层面,鼓励社会组织和公众对数字金融平台算法研发与应用开展社会监督。社会公众作为数字金融平台产品与服务的潜在消费者,应当知情必要的算法基本常识与负外部性,既要通过合理知情权的方式更多地获悉数字金融平台的算法情况以破除算法黑箱,也要通过合法合理地保护个人信息以防范算法黑箱可能带来的潜在风险。其中,为了保证对数字金融平台滥用数据与算法资源所形成的“算法权力”监督,可从以下层面对算法研发与应用夯实社会监督建设,提升对算法黑箱风险的第三方监督:一方面,充分发挥新闻媒体的算法监督职能。在人工智能技术的驱动下,社会经济逐步进入到算法时代,“信息茧房”与“信息孤岛”等层出不穷,桎梏了信息的自由流动<sup>[40]</sup>,人们更容易陷入“算法蒙蔽”的困境,更需要充分发挥专业新闻媒体机构的专业能力,在合法合理的范围内,提升新闻媒体机构对数字金融平台算法使用情况的调研与核查,勇于对数字金融平台滥用算法导致的黑箱、歧视与错误进行提问、调查和追责,从而充分发挥新闻媒体作为事实真相提供者与社会监督者的算法监督职能。另一方面,建立公平、公正、可信赖的算法伦理委员会等组织,全面充实第三方非营利组织的监督力量。数字金融平台滥用算法黑箱形成优势实施金融监管套利等行为,不仅是个法律合规问题,也是道德伦理的问题,自此有必要设立专业可信赖的算法伦理委员会等第三方监督组织,对数字金融平台滥用算法黑箱行为开展全面监督。

## 五、结 语

在金融强国的目标下,数字金融场景化程度日益加深,各类数字化金融服务广泛应用于日常生活,成为用户严重依赖的应用程序,如支付宝、微信支付等的用户场景深度连接了超10亿用户及超亿商家。但是,在算法技术与市场需求的双轮驱动下,数字金融平台快速发展并衍生了算法黑箱,这为掩盖其非法意图与监管套利提供了实施机会。从微软负责任的算法伦理框架,到《布莱切利宣言》以及欧美发达经济体的算法治理规范,算法决策系统的安全可控

与风险防范充分彰显了重要性。倘若缺乏系统的风险治理机制,无人治理的算法技术进化方法,通常容易因为技术不可知与信息不对称而出现算法黑箱与算法失控的局面,那么通过算法技术促进数字金融平台发展的“解药”,可能成为摧毁数字金融平台基础的“毒药”。

当前,尽管我国在数字金融平台的实践运用与算法治理方面取得一定效果,但是现代信息技术的快速演变导致数字金融平台算法创新及其黑箱风险的矛盾突出,包括平台迷雾风险、技术遮蔽风险与监管真空风险,严重阻碍金融强国目标的实现。如果我国依然采取传统工业时代下的风险治理机制,完全依赖技术规制或法律规范进行解决,那么这可能因为无法应对数字金融平台系统性风险特征,从而制约数字金融平台算法创新以及金融强国。面对数字金融平台算法黑箱而引发的潜在系统性风险,基于三方控制理论以及国内外数字金融平台算法黑箱风险的系统治理经验,我国可以从平台端(第一方控制)、用户端(第二方控制)与监管端(第三方控制)重塑数字金融平台算法黑箱的系统治理理念,全面构建本土化数字金融平台算法黑箱风险的系统治理路径,穿越系统性风险迷雾,从金融大国迈向金融强国:其一,秉承公开透明原则,从平台端健全数字金融平台的算法透明机制;其二,坚持技术治理原则,从用户端加强数字金融平台的算法赋权与信息对称;其三,恪守协同监管原则,从监管端推进数字金融平台算法黑箱的公权治理。

## 参考文献:

[1] 王怀勇. 金融科技的算法风险及其法律规制[J]. 政法论丛, 2021(1): 105-116.

[2] 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.

[3] 黄博文. 算法不完备性及其治理——以互联网金融消费者保护为中心[J]. 西南金融, 2018(8): 49-56.

[4] 刘烽. 算法语境下信义义务在金融领域的扩张[J]. 江西财经大学学报, 2020(6): 125-137.

[5] 王怀勇, 邓若翰. 算法时代金融公平的实现困境与法律应对[J]. 中南大学学报(社会科学版), 2021, 27(3): 1-14.

- [6] 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.
- [7] 中央金融委员会办公室, 中央金融工作委员会. 锚定建设金融强国目标 扎实推动金融高质量发展(深入学习贯彻习近平新时代中国特色社会主义思想)[N]. 人民日报, 2024-02-20(09).
- [8] 习近平在省部级主要领导干部推动金融高质量发展专题研讨班开班式上发表重要讲话强调 坚定不移走中国特色金融发展之路 推动我国金融高质量发展 赵乐际王沪宁丁薛祥李希韩正出席 蔡奇主持[N]. 人民日报, 2024-01-17(01).
- [9] 刘友华. 算法偏见及其规制路径研究[J]. 法学杂志, 2019, 40(6): 55-66.
- [10] 张爱军, 李圆. 人工智能时代的算法权力: 逻辑、风险及规制[J]. 河海大学学报(哲学社会科学版), 2019, 21(6): 18-24.
- [11] 何哲. 人工智能时代的治理转型: 挑战、变革与未来[M]. 北京: 知识产权出版社, 2021: 37-38.
- [12] 杨楠. 科技型小微企业信用评价模型研究[J]. 武汉金融, 2014(3): 56-58.
- [13] 弗兰克·帕斯奎尔. 黑箱社会: 控制金钱和信息的数据法则[M]. 赵亚男, 译. 北京: 中信出版社, 2015: 259.
- [14] 何德旭, 郭晓婧. 加快建设金融强国的理论内涵及实现路径[J]. 行政管理改革, 2024(1): 4-13.
- [15] 哈尔·R·范里安. 微观经济学: 现代观点(第八版)[M]. 费方域, 译. 上海: 格致出版社, 上海人民出版社, 2011: 543.
- [16] 单勇. 数字看门人与超大平台的犯罪治理[J]. 法律科学(西北政法大学学报), 2022, 40(2): 74-88.
- [17] 程雪军. 互联网金融控股公司的风险特性与监管建构[J]. 银行家, 2019(2): 44-46.
- [18] 曹中铭. 期指再现“乌龙指”健全风控机制迫在眉睫[EB/OL]. (2022-11-10) [2024-04-06]. <http://www.nbd.com.cn/articles/2022-11-10/2546999.html>.
- [19] 罗伯特·J·巴罗. 宏观经济学: 现代观点[M]. 沈志彦, 陈利贤, 译. 上海: 格致出版社, 上海人民出版社, 2008: 69-70.
- [20] 彭兴韵. 金融市场学[M]. 上海: 格致出版社, 上海人民出版社, 2018: 121-122.
- [21] 高鸿业. 西方经济学(微观部分)[M]. 北京: 中国人民大学出版社, 2021: 435-437.
- [22] 程雪军. 金融科技公司算法黑箱风险的生成机制与多维治理[J]. 金融论坛, 2023, 28(8): 25-34.
- [23] 陈钊. 信息与激励经济学(第二版)[M]. 上海: 格致出版社, 上海人民出版社, 2010: 23-24.
- [24] 王伟. 数字经济治理体系的运行逻辑——以合作治理为视角的考察[J]. 电子政务, 2023(10): 14-27.
- [25] 程雪军. 金融科技公司算法风险的体系化治理: 欧美比较治理视角[J]. 经济社会体制比较, 2023(6): 99-108.
- [26] 罗伯特·C·埃里克森. 无需法律的秩序: 相邻者如何解决纠纷[M]. 苏力, 译. 北京: 中国政法大学出版社, 2016: 132-139.
- [27] 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: 348-353.
- [28] European Commission. Guidelines on automated individual decision-making and profiting for the purpose of regulation 2016/679 [EB/OL]. (2018-8-22). <https://ec.europa.eu/newsroom/article29/items/612053>.
- [29] 张欣. 人工智能时代的算法治理机制与方案[M]. 北京: 法律出版社, 2022: 137-138.
- [30] WYDEN R. Support for the Algorithmic Accountability Act of 2022 [R/OL]. (2022-02-23). <https://www.wyden.senate.gov/download/support-for-the-algorithmic-accountability-act-of-2022-pdf>.
- [31] Monetary Authority of Singapore. Principles to promote fairness, ethics, accountability and transparency (FEAT) in the use of artificial intelligence and data analytics in Singapore's financial sector [EB/OL]. (2018-11-12). <https://www.mas.gov.sg/publications/monographs-or-information-paper/2018/feat>.
- [32] 马长山. 迈向数字社会的法律[M]. 北京: 法律出版社, 2021: 166-167.
- [33] 徐智华, 解彩霞. 平台经济算法用工的挑战与规制研究[J]. 宁夏社会科学, 2022(3): 98-107.
- [34] 邱惠君, 梁冬晗, 李凯. 欧盟人工智能立法提案的核心思想及未来影响分析[J]. 信息安全与通信保密, 2021(9): 92-99.
- [35] 张凌寒. 算法评估制度如何在平台问责中发挥作用[J]. 上海政法学院学报(法治论丛), 2021, 36(3): 45-57.
- [36] 姜海燕, 吴长风. 智能投顾的发展现状及监管建议[J]. 证券市场导报, 2016(12): 4-10.
- [37] 赵园园. 金融科技背景下信用法律制度完善研究



[M]. 北京:法律出版社,2021:204-205.

[38] 张永亮,林盛浩,张洁玉. 算法权力之规制:算法影响评估制度的生成及展开[J]. 证券法苑, 2022,35(1):227-246.

[39] TIKKINEN-PIRI C, ROHUNEN A, MARKKULA J. EU general data protection regulation: changes and implications for personal data collecting companies[J]. Computer Law & Security Review, 2018,34(1):134-153.

[40] 龚莉红. 基于“信息茧房”理论的意识形态话语权研究[J]. 河海大学学报(哲学社会科学版), 2019,21(5):35-40.

(编辑:高虹)

**The System Governance Mechanism of Algorithm Black Box in Digital Financial Platforms under the Goal of Financial Power**/CHENG Xuejun ( Law School, Tongji University, Shanghai 200092, China)

**Abstract:** Driven by the goal of becoming a financial power and artificial intelligence technology, digital financial platforms have rapidly emerged and deeply applied algorithm technology to the traditional financial industry. As a fusion product of digital financial and platform economy, the digital financial platform engages in collaborative activities with financial institutions by exporting technology or providing scenarios, and has natural technological preferences and financial attributes, so it effectively promotes the digitalization and inclusiveness of financial products and services. However, everything has two sides, not the result of either one or the other, but the product of both advantages and disadvantages. Through the use of case analysis methods, it is found that when algorithm technology is deeply applied to digital financial platforms, there is always a certain degree of error or uncertainty (algorithmic black box), which has systemic risk characteristics different from traditional enterprises, such as being too big to fail, too much to fail, and too strong to fail. In addition, algorithmic black box risk is formed in the third-party systems (platform side, user side, and regulatory side), including the risk of platform fog which hinders the stable development of the platform end, technical masking risk which aggravates risk defects in the user end, and regulatory vacuum risk which constrains the governance of public authority at the regulatory end. Therefore, in order to achieve the goal of becoming a financial power and prevent and resolve the systemic risks of the algorithm black box in digital financial platforms, based on the theory of tripartite control and the system governance experience of international algorithmic black box risk, China should build a system governance mechanism for the algorithm black box in digital financial platforms from the tripartite system (platform, user, and regulatory end).

**Key words:** financial power; digital finance; platform; algorithm black box; technical risks; algorithm transparency mechanism; system governance mechanism