

司法大数据赋能的非法集资风险识别 及监管路径优化研究*

○王洪¹ 冯晨² 贺娅洁¹ 李志勇²

1. 中国司法大数据研究院 北京 100043

2. 西南财经大学金融学院 四川成都 611130

摘要：非法集资严重干扰正常的经济金融秩序，影响社会稳定。由于互联网传播提速、集资手段层出不穷、资金跨域流动加剧，非法集资风险呈现出涉众性扩散、隐蔽性升级、跨域性增强等特征。传统以事后打击为主的监管模式在风险识别、预警防控及协同治理等方面逐渐暴露出制度性不足，在大数据背景下如何利用先进的数据技术进行事前监管成为破局关键。本文在系统梳理非法集资主要模式、风险特征及相关政策法规的基础上，深入分析当前非法集资风险监管在数据整合、风险识别机制及监管协同等方面面临的现实困境；同时，从司法大数据的视角出发，采用民事案件司法大数据，揭示非法集资风险演化规律，构建了一个非法集资风险识别与评级的技术框架；在此基础上提出通过完善风险信息整合机制、推进分级分类监管以及健全多主体协同治理体系，优化我国非法集资风险监管路径，以期为非法集资的前端防控与长效治理提供理论支持与制度参考。

关键词：大数据；非法集资；金融监管；数字化转型；数据安全；风险防控；新质生产力；监管科技

中图分类号：F832 文献标识码：A 文章编号：1009-4350-2026(03)-0029-16

基金项目：本文系国家重点研发计划项目“基于隐蔽通道数据的新颖非法集资线索及主体挖掘技术研究”(2023YFC3305302)、国家自然科学基金青年项目“面向市场监管行政执法的知识图谱关键技术研究”(62306341)、国家自然科学基金面上项目“基于多模态融合的上市公司退市风险预测研究”(72471196)的阶段性成果。

作者简介：王洪，男，高级工程师，中国司法大数据研究院金融事业部总经理，研究方向：多模态大数据分析、大规模知识处理；冯晨，女，西南财经大学金融学院博士，研究方向：金融风险管理和监管科技；贺娅洁，女，博士、博士后，中国司法大数据研究院研究员，研究方向：智慧司法、监管科技；李志勇（通讯作者），男，博士，西南财经大学金融学院教授、博士生导师，研究方向：金融安全、智能风控、监管科技。

引言

近年来，我国非法集资活动呈现出涉众性扩散、隐蔽性升级与跨域性增强的复杂特征，风险结构日趋复杂。一方面，实体经济下行压力加大、中小企业融资约束持续存在，部分市场主体在“融资难、融资贵”的现实环境中，转向游离于正规金融体

系之外的非规范融资渠道；另一方面，居民财富积累与投资需求快速增长，在金融知识相对不足、风险偏好错配的情况下，为非法集资活动提供了社会土壤。根据2025年全国两会《最高人民法院工作报告》，2024年全年持续推进重大非法集资案件审判执行工作，审结非法集资、洗钱等犯罪案件2.5万

件,判处罪犯4.8万人,同比增长5.3%。在此背景下,非法集资已不再局限于传统意义上的“高息揽储”“熟人集资”,而是深度嵌入私募基金、互联网金融、虚拟资产、养老服务等新兴领域,实现跨行业、跨区域、跨平台蔓延(Liu et al., 2018; 曹源芳、蒋志芬, 2019; 石奎, 2023)。大量司法案件表明,非法集资往往涉及人数多、资金规模大,极易引发群体性事件,冲击区域金融秩序和社会稳定,具有明显的滞后性和放大效应(Lai et al., 2022; 叶良芳, 2022)。

面对当前严峻复杂的金融风险形势,中央金融工作会议提出“要切实提高金融监管有效性,依法将所有金融活动全部纳入监管,消除监管空白和盲区,严厉打击非法金融活动”,并强调以数字化手段实现风险的“早识别、早预警、早暴露、早处置”。2021年,国务院颁布《防范和处置非法集资条例》,从监测预警、行政处置、资金清退、法律责任等多个维度对非法集资的防范处置进行规范。目前《全国非法集资监测预警体系建设规划(2020—2022年)》重点任务已基本完成,但现行治理模式仍面临数据孤岛突出、误判率较高、跨部门协同不足等现实约束。在数字经济背景下,数据要素作为新型生产要素正深刻重塑经济结构与风险形态(黄旭等, 2026; 王天雨、卢宇, 2025),司法大数据因其覆盖违法行为全生命周期的独特优势而成为识别和治理非法金融活动的重要信息基础。然而,现有非法集资监管仍主要依赖行政监管数据与金融交易数据,司法数据的治理价值尚未得到充分挖掘,跨域数据融合与风险建模能力相对不足。基于专家经验的识别模型难以有效刻画跨区域集资、虚拟货币交易等新型非法金融活动特征,导致监管实践中仍以事后处置和刑事打击为主,风险识别多依赖线索举报或案件暴露后的被动介入,前置防控与动态监测能力亟待提升。尤其是在非法集资活动日益“合法化包装”“专业化运作”的背景下,传统以行政审批、行业监管为核心的监管工具难以及时捕捉风险演化信号。如何在风险尚未全面爆发之前实现有效识别与干预,已成为当前非法集资治理亟须回应的核心问题。

本文遵循“现状描述—问题诊断—风险评估—

治理路径优化”的研究思路展开。首先,从司法实践和监管现实出发,系统梳理我国非法集资的风险特征;其次,分析现行监管体系在运行机制中的突出困境;再次,引入全周期视角对非法集资风险的阶段性跃迁进行理论分析,并进一步在司法大数据赋能下构建包含Schema风险体系、UIE信息抽取技术(Unified Information Extraction)、“BERT+双向注意力机制”语义互促解读在内的非法集资风险评估及预警技术框架;最后,提出完善我国非法集资监管体系的优化路径与政策建议。

一、文献综述

(一)非法集资的界定及影响

非法集资是法学和金融学领域的重要研究问题,根据《防范和处置非法集资条例》,非法集资是指“未经国务院金融管理部门依法许可或者违反国家金融管理规定,以许诺还本付息或者给予其他投资回报等方式,向不特定对象吸收资金的行为”。我国对非法集资的认定主要依据刑法及相关司法解释,通常围绕“未经依法批准”“向社会不特定对象吸收资金”“承诺还本付息或给予其他投资回报”“扰乱金融秩序”等要件展开。这一认定体系强调非法集资犯罪具有“非法性、公开性、利诱性、社会性”这四个关键特征(刘伟, 2020)。刑法中与非法集资活动相关的犯罪,包括非法吸收公众存款罪、集资诈骗罪、虚假广告罪以及擅自发行股票或者公司、企业债券罪。从行政犯角度,非法集资行为的界定应当以“国家规定”为核心,将案件分为合法集资、行政违法的非法集资、行政犯型非法集资犯罪和刑事犯型非法集资犯罪四类,强调司法机关在认定其“非法性”时应独立于地方政府的行政认定(胡彦涛、刘莉, 2018)。从集资者的主观目的与资金投向角度,叶良芳(2022)将非法集资划分为由于生产经营需要引发的生产型集资、基于资金运转的投资型集资和诈骗型集资三大类。

法理层面研究重点在于非法集资行为的“非法性”认定标准及其在刑民交叉语境下的法律效力。刘伟(2020)提出,对非法集资犯罪的认定必须坚持法治原则,并以此作为“非法性”判断的唯一准据,

防止行政权力对司法判断的过度干扰。胡彦涛和刘莉(2018)主张司法机关应独立认定集资行为是否满足行政犯的前提条件,明确“地方规定”不能作为定罪依据。在刑事与民事的法律交叠处理上,夏昊晗(2025)对传统的“当然无效说”进行了批驳,提倡“综合权衡论”,认为刑事构成犯罪并不必然导致民事合同损害公共利益,主张应视集资参与人的主观状态区分处理。

部分学者从金融治理的角度对非法集资展开研究。非法集资被视为游离于正规金融体系之外的“影子金融”活动,通常伴随着高杠杆运作、期限错配和信息披露不足等问题,主要依据《中华人民共和国反洗钱法》打击违法犯罪活动(石奎等,2025)。治理重点在于通过明确法律边界来保护普惠金融等新兴业态,在防范系统性金融风险的同时,筑牢社会治理的底线(胡彦涛、刘莉,2018)。从制度经济学角度,非法集资是金融体制二元化与金融压抑的产物,其根源在于正规金融供给不足与民间资本投资渠道匮乏之间的矛盾(潘淑娟、瞿苇,2008)。刘志伟(2016)认为应从单一的“行政控制”转向“市场导向与法律规制并重”,通过完善普惠金融体系来实现从强制压制向制度引导的路径转换。叶良芳(2022)基于总体国家安全观提出“分类治理”策略:严惩诈骗型集资,必要惩治脱实向虚的投资型集资;对服务实体经济的“生产型集资”,应基于鼓励创新与稳定就业的需要,探索其“除罪化”路径。非法集资案件的资产清退直接关系到受害群众的切身利益,必须建立标准化的处置程序挽回公众损失并维护社会秩序(谢勇、陈小杉,2019)。

(二) 非法集资的监管与治理

学界对非法集资监管现状的讨论多从制度结构与监管逻辑出发,揭示非法集资滋生的深层原因。伴随互联网技术的深度嵌入,非法集资活动呈现出更强的隐蔽性与跨地域流动性,传统以线下排查和事后惩治为主的监管方式难以及时应对,监管滞后性成为风险持续累积的重要诱因(魏景茹、程万宝,2021),亟须推动监管范式向“穿透式”监管转型(詹子友,2017)。进一步地,季洁(2020)基于长三角地区的实证分析发现,非法集资的区域性高发

并非偶然,其根源在于金融制度供给与实体经济融资需求之间的结构性错配,应通过地方性法规试点降低跨区域监管协调成本,推动行政监管与司法裁判在制度层面的有效衔接。另外,金融科技的发展一方面为非法集资活动提供先进的技术工具,可能加剧非法集资活动的发生频率(Lai et al., 2022),另一方面通过提升金融知识水平提高识别非法集资的水平,从而降低非法集资风险(张博等,2025)。

针对监管困境,学术界正推动从“事后打击”向“事前防控”模式转变(童德华,2024)。石奎(2025)认为针对非法集资应当采用超脱于“五大监管”模式之外的“预防监管”模式,但“预防监管”与当前《中华人民共和国个人信息保护法》《中华人民共和国消费者权益保护法》等法律存在一定的矛盾,有必要建立“安全港”规则。曹源芳和蒋志芬(2019)、康建华(2023)等研究认为,大数据技术在非法集资侦查与风险预警中起到关键作用,通过对资金流向、人员关联网络和互联网行为轨迹的深度挖掘,可以有效提升风险识别的前瞻性,缓解因信息不对称形成的监管盲区。一些研究提出了识别非法集资风险的具体方法,如Peng et al.(2019)基于用户交易行为特征构建了卷积神经网络;陈钢(2023)采用门限循环神经网络和注意力机制进行风险识别;蔡尚宣和朱丹青(2024)针对吸收资金账户与返利账户进行区分研究,引入PU-learning方法改善样本标注质量,并将阈值判定机制与机器学习模型相融合;黎若然等(2025)构建指标体系,采用层次分析法将非法集资风险量化为0—100分的风险指数,为监管部门提供了可操作的研判工具等。

(三) 文献评述

现有研究在非法集资的风险特征、识别技术以及行刑衔接理论方面积累了丰富的成果。然而,结合当前司法治理需求,仍存在以下不足:首先,风险识别机制研究尚不深入。数字化转型背景下,风险表征难、追踪难、评估难的问题依然突出,现有技术在对新型隐蔽风险时存在滞后性。其次,对司法数据的系统性讨论有限。虽然有研究针对长三角或东北地区进行了区域性分析,但缺乏基于全国跨省域司法大数据的实时监测与深度关联。最后,风险

分级与监管衔接研究薄弱。尤其是金融机构监测机制与公安机关立案追诉标准之间的实质性衔接研究仍处于起步阶段,难以为实务工作提供充分的证据支撑与决策依据。

二、非法集资现状及困境

(一)非法集资风险现状分析

2008年后互联网金融兴起,非法集资一度呈现高发态势,对金融秩序和社会稳定构成持续冲击。随着防范和打击金融违法犯罪政策体系的不断完善,以及金融监管、司法执法协同力度的持续加强,非法集资治理取得阶段性成效。司法大数据统计显示,2021—2025年与非法集资相关的民事案件与刑事案件数量总体呈现逐年下降趋势(如图1所示),表明高压打击与制度整治在遏制非法集资活动扩张方面已发挥显著作用。但从案件结构与犯罪样态的演变趋势来看,非法集资风险并未随案件数量下降而同步消解。在强监管背景下,犯罪手段不断翻新,呈现出涉众性扩散、隐蔽性升级、跨域性增强的新特征(康建华,2023;石奎,2023)。

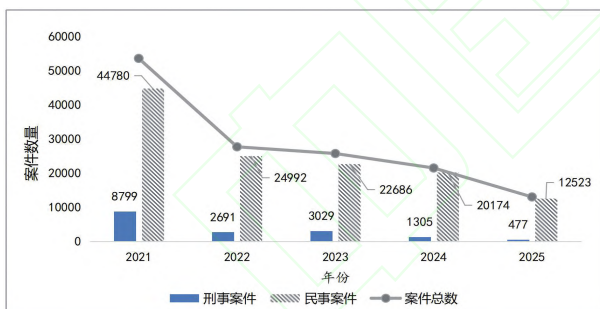


图1 2021—2025年非法集资案件数量

涉众性扩散。非法集资已从传统的“熟人圈层”传播转向“全民渗透”模式,受害群体除缺乏金融知识的老年人、农村居民等传统高危群体以外(曹源芳、蒋志芬,2019),公务员、金融从业者、高校毕业生等具备一定风险认知能力的群体占比逐年上升(苏如飞,2025),已超过60%。单案参与人数从数百人向数万人甚至数十万人突破,如2023年“小牛资本非法集资案”涉及投资者超130万人,涉案金额超1000亿元,“中植系”案件关联投资者覆盖多个省市,形成“一人参与、多户牵连”的结果。根据司法大数据统计,自2021年以来,涉嫌非法集

资案件的诉讼原告方为自然人的占比超过70%。

隐蔽性升级。非法集资手段持续迭代,依托合法经营外壳与新兴概念进行包装伪装。犯罪主体通过注册正规企业、取得普通经营资质,将非法集资行为与正常业务混合运作;利用云存储、虚拟货币、第三方支付平台等技术实现资金流转的“去中心化”与“匿名化”,通过境外服务器搭建集资平台、利用虚拟货币完成资金转移,规避传统资金监测;在宣传推广中采用异形字、黑话来替代敏感词汇,通过AI生成虚假项目视频、伪造投资者盈利截图,逃避监管部门的舆情监测与线索排查。因此,也衍生出多个类别的案由,2020—2024年非法集资案例的案由主要围绕“教育培训服务合同纠纷”“买卖合同纠纷”“民间借贷纠纷”展开,2025年案由类别更加细化,“不当得利纠纷”“委托合同纠纷”“人事争议”等新类别涌现。

跨域性增强。随着资金流动网络化与产业组织平台化,非法集资活动呈现出“跨区域布局、跨行业嵌套、跨市场传导”的复合跨域特征。在空间维度,犯罪主体往往采取注册地、运营地、资金归集地相分离的布局策略,在监管相对宽松或准入成本较低地区注册企业主体,在经济发达地区开展营销宣传与资金募集,并通过异地账户或第三方支付渠道实现资金归集与转移,借助行政区划分割与部门信息壁垒形成监管套利。例如,e租宝注册于安徽,主要面向北京、上海、广东等多地投资者开展线上线下集资,资金通过多地账户集中归集与层层转移,形成跨省(市)市运作链条。在行业维度上,非法集资活动突破传统金融与房地产领域边界,加速向养老服务、健康医疗、教育培训、农业投资、文旅等民生与实体产业领域渗透,通过产业项目包装与金融收益承诺相结合的方式实现跨行业嵌套运作。非法募集资金通过购买信托受益权、私募基金份额或委托贷款等方式嵌入正规金融体系,形成跨行业、跨市场的风险扩散链条。

(二)非法集资监管现状分析

1. 相关法律政策及监管体系建设。我国禁止任何形式的非法集资,对非法集资坚持“防范为主、打早打小、综合治理、稳妥处置”的原则。非法集资

监管体系历经多年发展并逐步完善。自1993年首次提出非法集资概念,到1998年国务院颁布《非法金融机构和非法金融业务活动取缔办法》,再到2021年《防范和处置非法集资条例》实施,我国的非法集资治理进入了行政处置与刑事打击并重的新阶段。非法集资监管相关政策见表1。根据《防范和处置非法集资条例》第八条,地方各级人民政府应当建立非法集资监测预警机制,并将其纳入社会治安综合治理体系。

当前监管体系以部际联席会议为统筹核心,省级政府主导建立监测预警机制,各地陆续建成大数据监测平台,如2015年北京市上线以冒烟指数作为测度指标的“北京打非监测预警云平台”,重庆市2018年上线非法集资监测预警、信息共享、有奖举报和指尖打非四平台系统。深圳市2018年通过“深圳市金融风险监测预警平台”“深圳市地方金融监管信息平台”和“灵鲲金融安全大数据平台”三大平台整合数据信息,识别具有非法集资风险的企业。2025年中国证监会在“处置非法集资部际联席会议”上进一步指出要从健全完善法规制度体系、非法集资风险预警、加强政策宣传和投资者教育、加强检查执法四个方面防范和打击非法集资活动。

2. 非法集资监管模式。非法集资活动形式不断变化,单纯依赖事后惩治或传统监管手段已难以

有效应对相关风险。我国的非法集资监管逐步从打击处置转向风险预防,形成了多元监管模式并行的格局。从监管理念层面进行划分,主要包括机构监管、功能监管、穿透式监管。机构监管以金融牌照和行政隶属关系为依据,对银行、证券、保险等持牌金融机构实施分类监管,在规范传统金融业务方面具有明确的责任边界。功能监管以金融行为和风险特征为核心,强调看行为而非看主体,通过关注资金募集方式、收益承诺特征及资金运用路径,对不同业态下的类集资行为进行统一识别,突破行业分割形成的监管盲区。穿透式监管在于突破形式审查和表层合规的限制,对集资行为的实质风险进行全面识别,强调对多层嵌套主体结构的逐级拆解,重点审查实际控制人、关联交易安排以及资金的最终流向,防止非法集资主体通过复杂的法律结构和业务包装规避监管。通过穿透“保本高收益”“委托代投”“消费返利”等表象,监管部门能够准确把握集资活动的真实性质,有效应对风险扩散。

监管技术层面,大数据智能监管和冒烟指数是主流的监管模式。大数据智能监管通过整合工商登记、税务、司法、银行账户及互联网平台等多源数据,构建资金链、人员关联网络和信息传播路径,对非法集资活动进行持续监测和动态识别,尤其适用于应对网络化、隐蔽化和高频化的非法集资行为。

表1 非法集资监管相关政策

时间	政策内容
1993年	《批转中国人民银行关于集中信贷资金保证当前经济发展重点需要意见的通知》首次提出要“要坚决制止和纠正违章拆借、非法集资”
1997年	与非法集资活动相关的非法吸收公众存款罪、集资诈骗罪、欺诈发行股票、债券罪和擅自发行股票、公司、企业债券罪四类犯罪被纳入刑法
1998年	国务院发布《非法金融机构和非法金融业务活动取缔办法》,明确人民银行为行政处置机关,首次界定非法吸收公众存款等非法金融业务的基本形式
2007年	建立由原银监会牵头的处置非法集资部际联席会议制度
2008年	国务院发布《处置非法集资工作操作流程(试行)》,首次将非法集资的处置工作标准化,详细规定了线索发现、初步调查、性质认定、立案侦查等各个环节的操作规程
2015年	国务院发布《关于进一步做好防范和处置非法集资工作的意见》,强调“以防为主”的监测预警思路,确立了“省级人民政府对本行政区域内防非处非工作负总责”的机制
2020年	处置非法集资部际联席会议发布《全国非法集资监测预警体系建设规划(2020—2022年)》,明确了打造“全国一张网”的建设要求。强调利用大数据、人工智能技术,通过对资金流、信息流的穿透分析,实现非法集资的主动发现和自动预警,推动监管由“事后处置”向“事前防范”前移
2021年	国务院发布《防范和处置非法集资条例》,赋予地方政府牵头部门查封、扣押、查询资金等强力行政调查手段,从系统源头切断非法集资渠道

冒烟指数是建立在大数据基础上的前瞻性预警模式,其通过综合考察异常高收益承诺、刚性兑付暗示、投诉举报数量、网络舆情波动及资金流动异常等指标,构建动态风险指数。当指标组合达到预警阈值即进入“冒烟”状态,监管部门可据此提前核查,防止局部异常演变为系统性事件。该模式突出了风险导向和前端治理,大幅提升了监管灵敏度和响应速度。

监管体制层面,常见的模式包括监管沙盒和属地协同。监管沙盒在非法集资防范中主要发挥风险筛查和边界厘清的功能,通过将新型金融业务限定在特定范围、对象和风险敞口内进行试点运行,使监管部门能够在业务扩散前及时获取关键数据并评估其潜在风险属性。属地协同治理是我国非法集资防范与处置的制度基础,其核心在于明确地方政府在风险防控中的主体责任,并通过多部门协同实现综合治理。在该模式下,地方政府统筹金融监管、公安、市场监管和司法机关等力量,实现行政处罚、刑事打击与司法救济之间的有效衔接。

(三)非法集资风险监管困境

1. 风险线索汇聚不畅,跨部门数据融合度低。首先,部门间数据割裂与异构性突出,非法集资的风险信号往往散落在工商登记、税务申报、资金流水及司法诉讼等不同系统之中。由于行政监管与司法机关分属不同的职能体系,双方在数据采集标准、风险定义及存储格式上存在显著的语义鸿沟。其次,数据共享面临法律合规的制度性掣肘。随着《中华人民共和国个人信息保护法》与《中华人民共和国数据安全法》的实施,数据隐私保护要求日益严格。在缺乏明确的“公共利益优先”豁免条款或数据共享安全港规则的情况下,监管部门与金融机构在跨机构调取涉嫌违规主体的资金流向、通信记录等敏感数据时,面临极大的合规不确定性,这种数据割裂状态导致监管部门无法对企业的资金流向、关联关系及经营异常进行全景式的预警监控。

2. 识别研判机制滞后,缺乏前置性预警机制。传统监管模式过分依赖事后处置与行政核查,缺乏对风险演化规律的动态捕捉能力。非法集资犯

罪具有显著的演化过程,往往先表现为民事违约或借贷纠纷,最终才转化为刑事案件。对于辖区内短期激增的涉众型民事诉讼、同质化借贷纠纷缺乏系统性监测,错失了在资金链彻底断裂前介入的窗口期。随着金融伪创新的泛滥,非法集资主体普遍通过复杂的合同架构、虚构的贸易背景或伪装的科技概念进行合法化包装。基于形式合规审查的规则引擎难以识别文本背后的隐性诱导话术与违背商业逻辑的收益承诺。由于缺乏基于算法驱动的深度穿透手段,监管部门往往被表面的合法形式所迷惑,难以从交易实质层面锁定潜在的非非法集资行为。

3. 处置手段缺乏分级分类治理体系。现行监管在处置策略上存在僵化与资源错配的问题,监管实践多依赖人工经验排查,缺乏基于多维数据的量化评估模型,难以对风险主体的危害程度进行精准分级,往往出现对低风险违规行为过度干预,而对高风险隐患的监管穿透力不足的结构性矛盾。同时,对涉险主体处置路径单一,缺乏柔性退出机制。在处置非法集资案件时,往往重刑事打击轻资产保全,重事后惩治轻事前化解。对处于风险初期、主观恶性较小的违规主体,缺乏矫正机会或市场化退出路径;对已爆发风险的主体,往往因行刑衔接不畅,导致资产冻结滞后追赃挽损难度大。

4. 治理架构碎片化,跨域协同与社会共治乏力。非法集资风险的跨区域流动性与网络化特征,与当前属地化、条块分割的治理架构形成冲突。非法集资活动往往利用行政区划分割进行监管套利,呈现出注册地、经营地、资金池所在地分离的特征。属地监管原则下,地方监管部门往往缺乏跨省协查与联合执法的抓手,尽管属地协同监管被认为是有效的监管方式,但目前监管政策落地仍然存在现实困难。另外,社会共治体系尚未有效激活。实践中金融机构、第三方支付平台及专业中介机构的责任压实不够,缺乏将内部风控数据向监管部门主动推送的激励机制,社会公众监督多停留在事后维权阶段,缺乏便捷化、数字化且具有隐私保护的举报反馈渠道,未能形成政府监管、行业自律与社会监督良性互动的共治生态。

三、基于司法大数据的非法集资风险评估及预警

(一)理论分析:非法集资风险的生成及演化

非法集资已由孤立的欺诈事件演变为具有高度涉众性的复杂风险,本文引入全生命周期视角,对非法集资风险从风险孕育到最终爆发的跃迁过程进行深度分析,构建了“潜伏诱发—扩张传染—质变爆发—风险外溢”的演化理论分析框架(如图2所示),旨在为非法集资风险前瞻预警与精准治理提供理论支撑。

1. 潜伏诱发:风险隐蔽与初始累积。初始阶段,非法集资风险往往以高度伪装的形态处于潜伏状态,违法主体通过概念包装与信息操控,人为制造显著信息不对称,以获取目标群体的初始信任。正规金融信任通常依托牌照资质、监管披露与长期声誉;非法主体缺乏合法信用基础,通过包装虚构信用将普通甚至虚构资产包装为契合政策导向或技术前沿的创新项目,利用公众在新兴领域中的认知盲区实施诱导。例如,以“生态农业”“新能源开发”“养老服务”等实体名义,或在数字经济语境下转化为“区块链节点”“元宇宙资产”“算力共享”等概念,配合高档办公场所租赁、虚假专家背书等构建貌似完备的商业逻辑,并以低风险、高收益或保本付息承诺迎合投资者逐利预期。

2. 扩张传染:风险扩散与结构性累积。首批资金进入后,非法集资突破初始信任门槛转入加速扩张阶段。风险传播由主体单向宣传转化为依托社

会网络与利益激励的自发扩散。为快速吸筹,组织者普遍设置高额“拉新返利”或多层级佣金机制,并在早期严格履约、按期兑付,以营造项目高收益假象,形成显著示范效应与正向强化。早期参与者因获得超常收益而降低风险警觉,在返佣诱导下向亲友同事持续推介,借助熟人网络的信任资本实现扩散。资金池规模与参与人数呈几何式增长,系统脆弱性同步累积,高成本吸资导致未来兑付义务快速膨胀,风险结构趋于失衡。

3. 质变爆发:风险显化与利益链条断裂。随着规模扩张,非法集资不可避免触及经济与资金约束边界,其本质为缺乏真实资产支撑的庞氏结构:向投资者支付的收益主要来源于后续投资者本金,具有内在不可持续性。系统运转依赖新增资金流入持续覆盖到期本息流出,一旦新增资金增速因投资者饱和、流动性收紧或监管介入而边际递减,当期募集不足以兑付当期本息,资金循环即发生断裂。当提现需求集中释放并演化为挤兑时,资金链迅速断裂,主体失联或转移资产,风险完成由量变到质变的转化并集中爆发。

4. 风险外溢:风险转化与社会冲击。资金链断裂后,非法集资进入外溢阶段,风险由个体经济损失演化为公共治理事件,对区域金融生态与社会稳定产生复合冲击。其一,投资者财富大幅损毁并伴随强烈心理冲击,弱势群体受损,家庭层面风险集中暴露。其二,区域信用环境受损,公众对金融机构及地方治理的信任度下降,当存在资金通道或违规担保等关联情形时,甚至可能引发局部金融波动并外溢至实体融资环境。其三,经济纠纷迅速向司法与社会治理领域扩散,集中报案与诉讼导致司法资源承压,可能诱发跨区域群体性事件与维稳压力。

甚至可能引发局部金融波动并外溢至实体融资环境。其三,经济纠纷迅速向司法与社会治理领域扩散,集中报案与诉讼导致司法资源承压,可能诱发跨区域群体性事件与维稳压力。

(二)司法大数据在破解非法集资监管难题中的作用机制

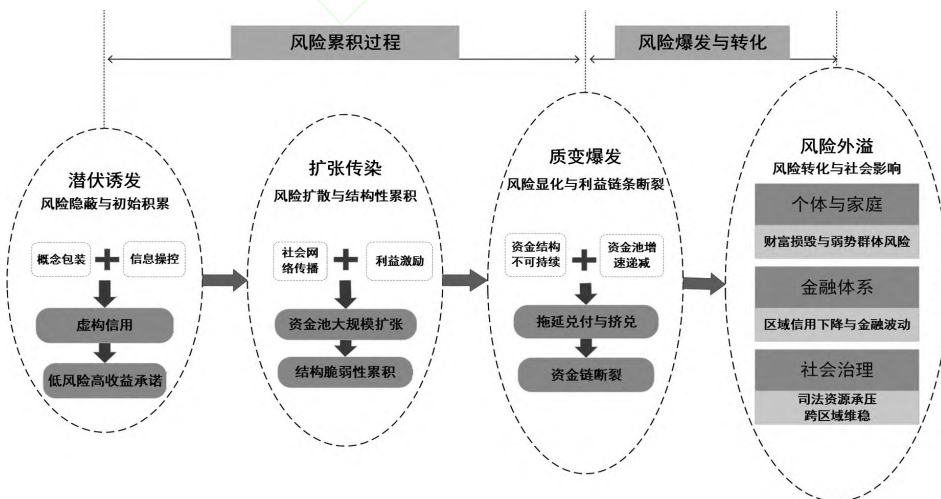


图2 全周期视角的非法集资风险跃迁

基于上述风险演化分析可以发现,非法集资从潜伏诱发到外溢呈现出明显的阶段性演进特征,不同阶段在行为模式、资金流动与参与主体网络上均会留下可识别的客观痕迹。这些痕迹既体现在资金往来、合同关系与组织架构等交易层面,也体现在案件立案、审理裁判与执行处置等司法活动之中,构成了风险运行的过程记录。

1. 法定权威属性,强化监管措施的司法协同效力。司法大数据由人民法院在法定程序中依法生成,并经过审判活动与文书制作的多重审核,天然具备高度的权威性、规范性和证据效力,为跨部门监管协同提供了可靠且可核验的数据基础。在行政监管与处置环节,司法裁判文书中“资金流向账户”“主要责任主体认定”等核心信息,可依法直接转化为账户冻结、资产查封与追缴等行政执法措施的依据,有效缓解行政监管中长期存在的取证成本高、证据效力不足等现实难题。在刑事打击阶段,既有民事裁判中形成的“关联企业交易记录”等事实认定,能够作为刑事侦查的重要补充材料,延伸和夯实刑事证据链条。例如,在某非法集资案件中,执法机关通过系统整合17份民事调解书,成功还原资金跨主体转移路径,进而锁定涉及8家壳公司的资金通道结构。通过构建“行政监管—司法裁判—刑事打击”相互衔接的数据闭环,不仅避免了多部门重复取证所带来的资源浪费,也借助司法裁判所蕴含的强制性与惩戒性,显著提升了监管措施的整体震慑力与执行效果。

2. 数据全流程覆盖,摆脱监管数据碎片化困境。司法大数据系统整合了刑事诉讼与民事诉讼的全流程信息,构建起涵盖案件受理、证据调查与认定、裁判结果形成以及执行反馈在内的完整数据链条,实现了对违法行为从发生、暴露到处置全过程的动态刻画,具有样本规模大、时间跨度长和类型结构完整等显著优势。相关数据特征不仅包括被告人类型、罪名分布、案件地域等基础信息,还进一步延伸至资金流向结构、参与主体行为模式、合同条款特征,以及裁判文书中的事实认定与说理逻辑等深层内容。相较于行政监管数据普遍存在的“重结果、轻过程”问题,这种覆盖司法审理全流程

的数据结构有助于还原非法集资行为的形成路径和演化机制。例如,通过对民事裁判文书中“借款合同纠纷”与“年化收益率显著高于合理区间(如超过15%)”等要素的关联分析,可以在案件进入刑事立案之前,提前识别“以合法借贷形式掩盖非法集资实质”的隐蔽模式。司法大数据全流程、跨类型的数据整合能力有效缓解了非法集资活动因跨区域布局、跨行业渗透所引发的监管信息割裂问题,在提升传统行政排查针对性与前瞻性的同时,也显著提高了风险识别与处置的整体效率。

3. 民事纠纷预警,构建前瞻性刑事风险防控机制。在司法实践中,民事诉讼数据是社会矛盾与市场风险的最早显性化载体,其所反映的异常行为模式和纠纷结构,对刑事犯罪的提前识别与风险防控具有不可替代的前置预警价值。非法集资类犯罪通常并非瞬时爆发,而是经历由民事借贷纠纷频发到非法吸收资金规模化,再到刑事犯罪定性的渐进式演化过程。司法大数据分析表明,多数涉案主体在刑事立案之前,普遍存在持续2至3年的民事纠纷高发期,且纠纷类型高度集中于借款合同纠纷、委托理财纠纷等资金密集型领域。从风险传导与关联机理看,民事诉讼数据中所呈现的异常特征,往往对应着刑事风险形成过程中的关键结构性要素。例如,在部分企业相关民事案件中出现的“高频次诉讼+相同诉讼代理人+固定标的额”组合特征(如某针纺企业涉及142件追索劳动报酬纠纷,均由职工自行代理且单笔标的额集中于2000—3000元区间),可能反映企业通过虚假或程序性诉讼实现资产转移与责任规避的潜在动机。在非法集资监管语境下,对民事案件与刑事案件之间转化规律的系统识别,使民事司法数据得以转化为刑事风险预警的重要工具。据此,可将监管响应节点由传统的刑事立案之后前移至民事纠纷高发阶段,显著提升风险防控的前瞻性与主动性。

4. 多维度关联耦合分析,实现穿透式与动态化监管。司法大数据天然具备类型多样、结构复杂的特征,既包含案件数量、案由分布、主体类型等结构化信息,也涵盖裁判文书、庭审记录等高度依赖语义理解的非结构化文本数据。通过引入自然语言

处理技术,对民事裁判文书中的事实认定与裁判说理进行语义解析,可以有效提取“保本付息承诺”“货币交易”等具有显著风险指向性的关键词与情境标签,并在此基础上构建风险要素集合。进一步结合知识图谱技术,将企业主体、自然人、高频账户、涉案地区及交易行为等信息进行关联建模,有助于实现对跨区域、跨主体案件的穿透式分析。例如,司法数据关联分析揭示涉案企业通过分布于3个省份的12家关联公司实施“甲地注册、乙地集资、丙地洗钱”的跨区域犯罪模式,而传统基于属地或单一部门视角的监管方式难以及时识别该类复杂结构。司法数据在案件新增、裁判生效及执行反馈等环节的持续更新,使监管部门能够根据风险态势变化动态调整监管策略,推动形成风险识别、监管处置、效果反馈、模型修正的闭环优化机制,不断提升非法集资风险防控体系的适应性与有效性。

(三)非法集资风险识别及预警技术框架

传统监管模式过度依赖刑事打击的定性反馈,存在明显的时滞效应。根据前文分析,非法集资在演变为确凿的刑事案件前,会经历一段时间的风险演化过程,通常伴随着民事诉讼领域风险的异常集聚。将民事案件中蕴含的异常交易线索转化为行政监管的有效抓手,能够推动非法集资治理从单纯的“事后司法惩治”向“刑民协同的穿透式监管”转型。基于这一思路,本文基于民事诉讼线索提出了非法集资风险的识别及预警技术框架(如图3所示),主要包括文本信息抽取、风险特征互促解读、风险分级预警三部分。

1. 文本信息抽取。

(1)Schema 风险特征体系构建。Schema 体系是指通过对对象概念、属性特征及其相互关系进行统一、结构化和语义化定义的规范框架,其作用在于降低信息复杂性、提升跨主体理解一致性,并支撑数据共享和知识表达。由于非法集资监管涉及银行监管、市场监管、司法审判与公安执法等多个部门,不同部门在风险认定标准、证据采信范围和程序启动条件等方面存在明显差异,客观上形成了规则衔接不畅与信息传导断层。传统命名实体识别模型主要聚焦人名、机构名称等通用实体,难以

表2 基于司法文书的Schema风险特征体系

类别	风险特征
资质合规性	未经金融监管部门批准或超出经营范围
筹资方式与对象	向社会不特定对象吸收资金
	向社会公开宣传的方式
	集资人群
诱导机制	承诺固定收益利诱出资
	集资噱头
	虚拟标的物
资金表现	资金去向异常
	预存资金的表现

有效识别“利诱出资”“集资噱头”等具有显著行业特征的违规行为。为破解上述难题,本文系统梳理《防范和处置非法集资条例》与现行金融监管规则,融合司法裁判实践中的风险认定逻辑,构建了面向金融领域非法集资治理的基于司法文书的特定领域Schema体系(见表2)。

该Schema体系涵盖了对非法集资四类共计九种核心风险标签的系统挖掘,非法集资的核心构成要件首先表现为“未经金融监管部门批准或超出经营范围”,即相关主体在没有取得相应金融业务许可或在许可范围之外开展吸收资金活动,从主体资质层面缺乏合法性与合规性。在筹资方式与对象方面,非法集资通常以“向社会不特定对象吸收资金”的方式进行筹资;筹资过程可能依赖“向社会公开宣传的方式”,即通过网络平台、媒体广告、线下推介等公开渠道发布信息,使筹资信息快速传播并引发广泛关注;而“集资人群”则指向包括金融知识薄弱、收益需求强烈或受高收益承诺与宣传策略影响较大的社会群体。在诱导机制上,非法集资常以“承诺固定收益利诱出资”作为核心吸引手段,通过承诺稳定、确定的回报来降低投资者对风险的警惕,并通过“集资噱头”制造情绪化和故事化的吸引力;“虚拟标的物”则指其所宣称的投资标的可能是虚构或被夸大包装的资产与项目,缺乏真实可验证的经济基础,从而在本质上形成无实物、无项目、无收益来源的投资幻象。在资金运作层面,“预存资金的表现”即通过预付、预存、预充值等方式提前收取资金,但并未进入真实的投资运营体系,属于先

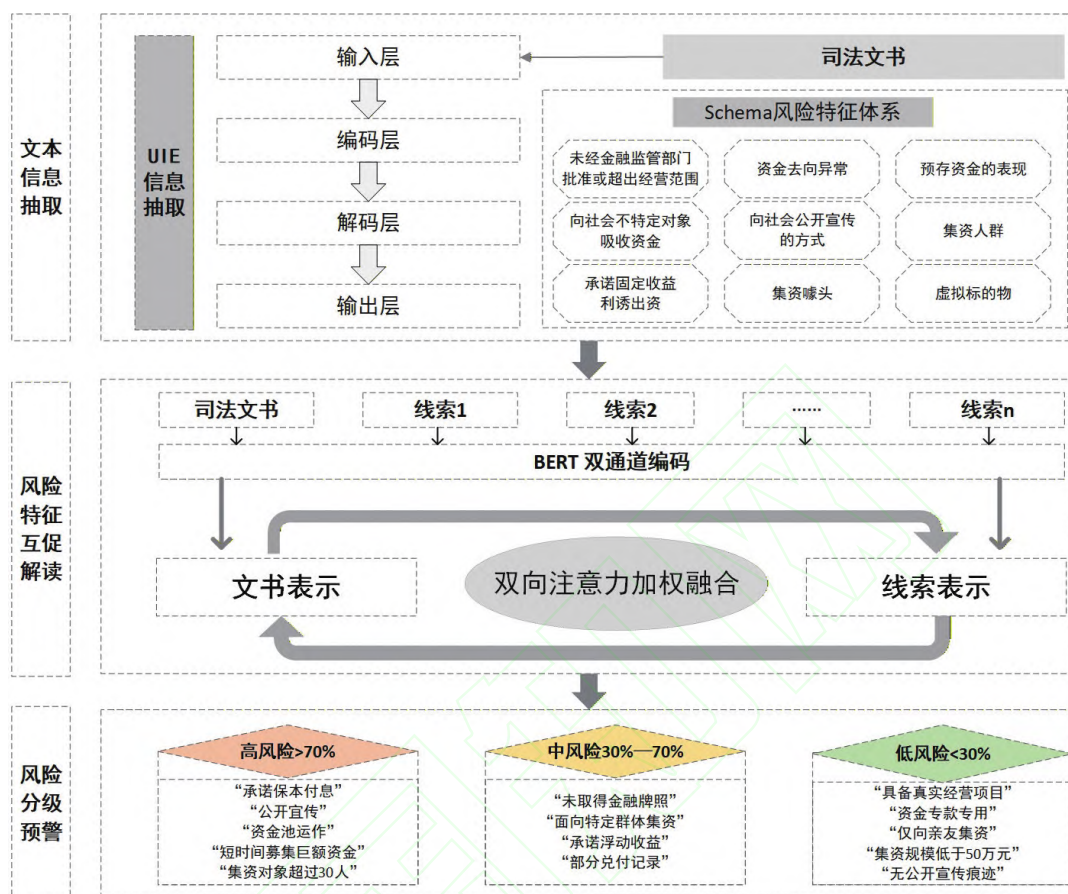


图3 非法集资风险识别及预警技术框架

收钱后承诺的结构;“资金去向异常”则指资金流向不透明、用途与宣传不一致,常见于资金被挪用于偿付早期投资者、关联方输送或高消费等非经营用途,导致资金链断裂风险加剧。

(2)UIE信息抽取与标签识别。司法大数据的特征提取与关联耦合分析需要从海量非结构化文本中精准抽取关键实体、关系及事件信息,UIE作为一种通用信息抽取框架,能够统一处理实体抽取、关系抽取、事件抽取等多种信息抽取任务,具有良好的迁移泛化能力。本阶段的关键是关系抽取与事件抽取,关系抽取通过提取案件中实体之间的关系,如“原告A与被告B约定借款利息和借款期限”等;事件抽取则是从文本中抽取完整的事件描述,包括事件的时间、地点、结果等。在非法集资监管场景中,UIE从司法文书等文本中精准提取“承诺年化收益率”“集资期限”等核心表述,同时识别“面向社会公众”“保本付息承诺”等违规特征;从结构化数据维度可提取参与人数、单笔集资金额、资

金累计规模等量化指标。这些提取结果既包含文本中的语义风险点,又涵盖可直接用于风险计算的结构化特征,为后续BERT双通道编码提供高质量的输入素材。

本文采用的UIE是基于ERNIE 3.0的信息抽取模型分层架构,风险特征提取包括输入层、编码层、解码层和输出层四部分。输入层将任务需求与待抽取文本以统一格式输入模型,形成序列化输入。编码层由多层Transformer构成,通过深层的自注意力结构同时捕捉文本中局部与全局的语义关联,如司法文书中跨段落的因果关系、实体的前后照应等。指针网络位于编码层与解码层之间,能够天然支持重叠实体、嵌套结构、多实例抽取等复杂的司法文本场景,为每个文本片段构建起始位置与结束位置的概率分布。解码层通过“双指针”机制,从指针网络的概率分布中选取可置信度最高的起止位置,形成具体的抽取结果。对于每个抽取任务,解码器会同时抽取多个同类型实体,并结合任务指令确保

抽取的文本片段与指定的属性或槽位语义相一致。输出层对解码层产生的结果根据置信度、文本一致性、任务语义、实体边界合理性等多维度规则,进行筛选、去重、冲突处理和结构化输出。

2. 风险特征互促解读。完成风险信息提取和标签化识别后,需要构建模型对文本语义进行深度解读以实现风险的分级。黄辉等(2021)在智慧法院的建设中首先提出基于BERT阅读理解框架的司法要素抽取方法,通过引入要素标签信息和法律先验知识构造相应的辅助问句,建立辅助问句和文本之间的语义联系。本文在此基础上使用多头注意力融合架构,通过构建“BERT+多头注意力机制”实现信息交互与整合,使得两类信息能够基于彼此的特征进行针对性的关注与提取,是实现“文书与风险特征互促解读”的基础结构。

具体而言,利用预训练语言模型对线索片段与司法文书分别进行深度语义编码,形成两类文本相对独立但结构一致的向量化表征。在文本编码的基础上,模型引入双向的语义交互机制,使线索片段与司法文书能够相互强化:一方面,司法文书的整体上下文能够帮助模型对线索片段中可能出现的歧义或偏差进行校正;另一方面,线索片段能够引导模型聚焦司法文书中与任务密切相关的关键内容,使长文本中的重要细节得以突出。这一互补性融合过程依托多层次的表示调整与特征重组,将线索信息与司法文书的全局语义逐步整合为统一的高阶表示。模型将融合后的语义表达输入分类模块,生成相应的低、中、高风险等级判断结果。

3. 风险等级预警。高风险对应的概率区间为70%以上,该结果的生成与司法文书中“承诺保本付息”“公开宣传”等高频风险特征高度关联。双向注意力机制作用下,文书语义中“短时间募集巨额资金”“无法兑付本息”等表述与刑事风险标签的匹配度超过90%;“风险特征→文书”通道进一步强化了“集资对象超过30人”“资金用途不明”等特征对文本风险等级的决定性影响。特征重要性加权结果显示,资金规模、兑付能力、宣传方式三

项特征的权重占比合计超过60%,成为高风险结果的核心支撑。

中风险对应的概率区间为30%—70%,其显著特征是风险特征与文本语义呈现混合性关联。一方面,文书中存在“未取得金融牌照”“面向特定群体集资”等违规表述,与行政违法风险特征的匹配度达60%至80%;另一方面,“承诺浮动收益”“部分兑付记录”等内容降低了刑事风险特征的关联强度。

低风险对应的概率区间为低于30%,其生成逻辑体现为风险特征与文本语义的弱关联性。司法文书中“具备真实经营项目”“资金专款专用”“仅向亲友集资”等表述在双向注意力机制中与低风险标签形成强映射。“风险特征→文书”通道显示,结构化风险特征中“集资规模低于50万元”“无公开宣传痕迹”的权重占比不足20%,且残差归一化融合后,此类特征与文书语义的冲突度低于10%,最终通过联合表示分类器输出低风险概率。

上述三个风险等级结果的区分度还体现在特征交互的动态变化中:高风险结果中特征间呈现正向叠加效应,单一高权重特征可显著提升整体风险概率;中风险结果存在特征对冲现象,部分合规特征与违规特征的相互抵消导致风险等级波动;低风险结果则表现为特征协同,多项低风险特征共同强化了风险等级的稳定性。

(四)非法集资风险预警效能验证

本文采用2021年至2025年6月30日前全国法院审结的非法集资刑事、民事案件数据(共24159件),涵盖4572件刑事案件和19587件民事案件的案件判决书及裁定书。为了排除与非法集资无关的案件对计算资源产生的消耗,从中选取涉嫌经济犯罪及刑事犯罪且被法院驳回起诉并移交公安机关的民事案件作为模型的数据集,并整合民事裁判文书中的关联诉讼数据,数据样本共9223条。通过对真实的非法集资案例进行人工风险分级,并采用五折交叉验证精确率和召回率^①的均值形成模型效能评估。Schema体系中不同的风险特征结果见

^①精确率是真实结果和预测结果一致的实体占预测实体的比例,召回率评价风险样本识别能力,即模型识别的风险样本数与实际风险样本数之比。

表3 Schema风险特征预测效果

风险特征	精确率	召回率
承诺固定收益利诱出资	98.47%	96.37%
向社会不特定对象吸收资金	87.90%	84.50%
未经金融监管部门批准或超出经营范围	80.77%	91.30%
向社会公开宣传的方式	91.18%	93.94%
集资人群	97.01%	98.98%
集资噱头	93.46%	97.28%
资金去向异常的表现	90.15%	83.94%
预存资金的表现	94.86%	93.18%
虚拟标的物	100.00%	89.47%
风险识别整体效果	94.65%	93.19%

表4 非法集资风险分级模型评估

风险类型	精确率	召回率
高风险	88.89%	96.55%
中风险	95.23%	64.52%
低风险	98.81%	93.26%

表3,与非法集资本质高度契合的行为特征在识别效果上表现尤为突出。其中,“承诺固定收益利诱出资”与“集资人群”两项特征标签的精确率分别达到98.47%和97.01%,召回率分别达到96.37%和98.98%,表明此类特征在区分非法集资与一般金融活动方面具有显著优势,既能够准确识别高风险样本,又能有效覆盖潜在违法行为,是非法集资风险识别中的核心判别要素。与此同时,“虚拟标的物”特征的精确率达到100.00%,显示出极强的判别稳定性,但其召回率相对较低(89.47%),一定程度上反映出该类特征虽具有高度排他性,在实际案件中并非所有非法集资行为均以虚拟标的为载体。在资金流向与运作方式相关特征方面,“资金去向异常的表现”与“预存资金的表现”均取得较为均衡的预测效果,表明资金层面的异常特征在识别非法集资中具有较强的解释力,但受限于资金路径复杂化和隐蔽化趋势,其召回效果仍存在一定提升空间。一些制度合规性与传播方式相关特征在预测效果上呈现出一定差异。例如,“未经金融监管部门批准或超出经营范围”的精确率为80.77%,但召回率达到91.30%,说明该特征在覆盖非法集资样本方

面较为充分,但由于部分合法与非法行为在形式上存在交叉,其区分精度相对有限;而“向社会公开宣传的方式”则表现出较高的精确率和召回率(分别为91.18%和93.94%),反映出传播方式仍是识别涉众型非法集资的重要外显信号。分级模型预测效果见表4。非法集资风险分级预警模型对低风险的识别精确率及召回率均达90%以上;高风险的识别精确率达88.89%,召回率达96.55%,说明本文提出的非法集资风险识别及预警框架能够有效地对风险进行早识别和早预警。低风险精确率比高、中风险精确率高的原因在于,部分文书中高/中风险的边界较模糊,部分高风险与中风险标签易混淆,但高/中风险与低风险的边界相对明显。

在单个案例应用中,本文所提出的非法集资风险识别及预警框架也取得了较好的应用成效。例如,犯罪团伙注册海南某能源科技公司以“共享充电桩众筹”为噱头,宣称投资者购买充电桩后可获“周返利2%,年收益近100%”,实际仅销售充电桩200余台,却通过App虚标数万台吸收资金超11亿元,涉及2000余名投资人。该公司无金融牌照却超经营范围吸收资金,通过“层级返佣”模式(团队队长提成15%—20%)使资金流分散隐匿,并通过伪造充电桩生产照片、购买政策文件包装合法性,资金通过App直连第三方支付规避对公账户预警。在穿透式监管应用中,首先由风险识别模型通过挖掘该案例司法文书中的非法集资行为线索,并融合文书中的案件事实描述,可提前实现非法集资风险识别。由风险分级模型动态量化案件风险等级,触发差异化处置机制,实现高风险非法集资案件精准拦截、中低风险主动干预。在案例中实现提前4个月预警风险、精准冻结涉案资金2.1亿元(占拟募集总额19%)、识别充电桩重复质押率43%,实证显示高风险预警准确率达91%,资金拦截时效提升300%。

四、我国非法集资监管优化路径

在当前金融科技迅速迭代与金融市场日益复杂的双重背景下,传统的监管模式面临着数据孤岛、“一刀切”处置以及部门协同不畅等现实困境。

基于前文对非法集资风险生成机理、监管现状、风险跃迁过程以及模型预测结果的分析,本文认为优化非法集资监管路径应当从信息整合机制的完善、风险的分级分类处置以及协同治理体系的健全三个维度展开,旨在构建一个全方位、立体化、智能化的风险防控生态系统。

(一)完善信息整合机制,构建全息风险映射

一是重视社会源头信息的收集与反馈。非法集资活动往往发生在社区、楼宇和网络群组中,应建立标准化、便捷化的举报奖励信息系统,通过开发手机App、微信小程序等端口,鼓励公众随手拍、随手报。社会举报信息不应仅仅停留在信访部门,而应通过数据接口直接导入风险监测数据库,与后台的资金数据、工商数据进行交叉比对验证。若某一区域在短时间内收到大量关于同一主体的咨询或投诉,系统应自动提升该主体的风险等级。同时,建立监管部门与互联网头部企业的合作机制,利用搜索引擎、社交平台的算法优势,对涉及“高收益”“零风险”等敏感词汇的推广信息进行定向捕捉,形成“线上+线下”“官方+民间”的全息化信息情报网。

二是明确可用非法集资风险信息边界,确立法律适用优先原则。在现行法框架下,《个人信息保护法》《数据安全法》等法律对数据处理提出严格限制,但在涉众性金融风险治理场景中,单纯强调数据保护可能对公共安全构成制度性掣肘。应通过金融监管立法明确“公共利益优先”的冲突规则,在涉众性非法集资防范领域构建数据共享的“安全港”制度,界定监管部门与司法机关在风险监测、案件溯源、证据收集等场景下的合法数据使用边界,并设定比例原则与最小必要原则作为合规底线。通过确立明确的免责规则,降低监管主体的数据合规不确定性成本,从而推动跨部门数据共享由“被动合规”向“主动协同”转型。

三是建立基于Schema体系的跨部门风险语义对齐标准。当前司法裁判文书、行政监管系统与金融机构内部风控系统在概念体系与风险标签上存在显著异构性,导致数据难以互操作。本文基于Schema体系构建跨风险语义框架,将司法裁判中

的法律事实与构成要件抽象为机器可识别的风险特征标签,随着非法集资风险的演变特征变化,Schema特征标签有望实现动态生成以捕捉复杂的风险变化,生成企业的“风险热力图”。这种语义对齐不仅提升数据互操作性,也为跨域智能监管模型提供统一的知识底座。

四是打破部门壁垒,构建国家级与地方级互联互通的金融风险大数据监测平台,形成多源异构数据验证闭环。非法集资往往披着“科技创新”“金融互助”等合法外衣,其资金流、信息流散落在银行、第三方支付机构、工商登记部门、互联网平台及各类社交媒体之中。监管部门应当依法获得授权,整合市场监管部门的企业注册信息、税务部门的纳税申报数据、金融管理部门的大额可疑交易报告以及公安机关的情报信息。通过API接口技术实现这些异构数据的实时抓取与清洗,将原本割裂的静态数据转化为动态的风险图谱。在风险识别实践中,单一行政数据源易受到噪声与策略性行为干扰,导致误报率与漏报率并存。应构建以行政数据为前端感知、以司法数据为权威校验的多源验证闭环机制。当工商、税务、市场监管等系统发现企业存在异常经营行为时,自动触发司法数据关联检索,通过对涉诉频次、执行标的、群体性案件等指标进行交叉验证,实现从“行为异常”到“法律风险”的证据跃迁。

(二)优化研判逻辑,确立从民事案件入手的风险传导预警机制

一是确立民事诉讼作为刑事风险的预警指标。传统金融监管主要依赖行政检查与事后执法,难以及时捕捉刑事风险的早期信号。在非法集资演化路径中,民事纠纷往往先于刑事犯罪暴露,呈现出由合同纠纷向刑事诈骗的风险传导轨迹。应构建民事诉讼风险阈值模型,对类金融机构涉诉数量、诉讼结构与诉讼行为模式进行动态监测。一旦出现诉讼集中爆发、代理人高度集中或标的额高度同质化等异常模式,即触发刑事风险预警标识。

二是针对非法集资主体通过合同形式掩盖犯罪本质的策略性行为,应从形式审查转向交易实质审查,建立基于异常交易结构的实质穿透研判框架。监管重点应放在识别民事案件中是否存在脱

离实体经济基础的高额回报承诺、不符合风险收益对称原则的保本承诺、虚构债权债务关系等违背商业常理的交易特征。同时,可通过比较民事裁判所认定的资金成本与行业平均利润率之间的偏离度,构建量化偏离指标体系,从而识别利用民事司法程序对非法集资行为进行“合法化包装”的制度套利路径,并从法律关系性质层面锁定潜在刑事犯罪嫌疑人。在技术层面,可考虑加大部署基于算法驱动的非结构化文本深度穿透。非法集资活动高度依赖营销文本、合同话术与宣传材料进行风险掩饰,传统规则引擎难以识别语义隐蔽性。应将深度学习模型嵌入监管系统,对多源文本数据进行语义嵌入与注意力权重分析,识别隐性诱导话术与收益承诺模式,并结合金融业务结构进行合规性推断。

三是建立基于“异常诉讼特征组合”的深度画像机制。从聚焦“单一案件事实”转向识别“群体同质化模式”,是利用司法大数据提升非法集资识别能力的关键路径。相较于一般商业纠纷,非法集资相关民事案件往往呈现出高度组织化与标准化特征,具有明显的模式化诉讼行为轨迹。因此,监管研判应基于司法大数据进一步构建“异常诉讼特征组合”的群体识别框架,重点监测具备“高频次诉讼—代理人高度集中—标的额和计息方式高度同质化”等复合特征的案件集群。当系统监测到大量民间借贷纠纷案件由同一或少数代理人集中代理,且诉讼请求中的本金规模、利率结构或计息规则呈现高度一致性时,通常意味着背后存在标准化集资合同文本与组织化法律策略,反映出通过民事诉讼程序进行风险转嫁与责任规避的制度性安排。通过对上述微观诉讼特征的深度挖掘与跨案件关联画像,可以实现从个案层面向组织层面的风险结构识别,有效剥离隐藏在批量民事诉讼表象之下的系统性非法集资风险网络。

(三)优化处置策略,实施基于量化评级的风险评价和穿透治理

一是构建算法量化评级与制度分级响应的双重驱动识别机制。应改变传统监管依赖经验判断的主观性与滞后性,确立以数据驱动为核心的风险量化评估体系。依托智能模型对涉案资金规模、资

金流向结构、涉众广度、诉讼行为特征和舆情扩散强度等关键指标进行多维特征提取与加权计算,形成客观的风险概率输出。在此基础上,建立模型评级与监管制度之间的映射关系,形成动态化的高风险、中风险与低风险三级预警管理体系。该机制的核心在于推动监管模式由静态合规检查向动态风险感知转型,通过模型结果与监管强度的结构性匹配,实现监管资源向高风险领域的精准配置,缓解低风险领域过度干预与高风险领域监管不足并存的结构性矛盾。

二是完善差异化监管,优化分类处置体系。对高风险主体,应实施阻断型治理路径,强化行政执法与刑事司法的衔接机制。当模型监测触发高风险阈值时,立即启动跨部门应急联动程序,依托司法大数据的证据转化功能,实现监管部门与公安机关的提前介入,并同步采取账户冻结、资产保全等诉前控制措施,以切断资金外逃路径,从源头上遏制风险扩散,实现对重大金融犯罪的前置打击与损失最小化。对中风险主体,应实施矫正型治理路径,引入监管沙盒理念开展动态观察与结构性纠偏。通过高频次非现场监测、监管约谈与穿透式检查的组合手段,抑制其业务过度扩张行为,并引入第三方审计与法律评估机制,推动企业剥离违规业务、化解存量风险,引导其向合规金融活动转型或有序退出市场,从制度层面阻断一般违规行为向刑事犯罪演化的风险路径。对低风险主体,应实施免疫型治理策略,坚持审慎包容与底线监管相结合,避免过度行政干预对金融创新与实体经济活力形成抑制。监管重心转向合规边界划定、风险提示与投资者教育,通过典型案例发布与风险预警机制,提高市场主体的风险识别能力与制度免疫能力,从源头压缩非法集资活动的生存空间。

三是强化风险柔性化解与社会心理治理的系统性稳定机制。风险处置不应仅停留在行政处罚层面,还应兼顾社会关系修复与金融秩序稳定。对于因经营失序导致流动性困难的违规主体,应优先采用市场化与法治化路径进行风险化解,通过资产重组、债务重构与引入战略投资者等方式推动风险平稳出清,政府应主要发挥制度协调与规则引导功

能,而非直接兜底。同时,应将投资者适当性管理延伸至风险处置阶段。一方面,加强对老年人等弱势群体的心理疏导与社会帮扶,防止经济风险向社会风险乃至政治风险转化;另一方面,对高风险偏好型投机参与者,应严格落实卖者尽责与买者自负原则,强化风险认知与责任意识教育,避免道德风险引致的逆向选择与制度套利行为。

(四)深化跨域协同与社会共治格局

一是深化纵横联动的跨域协同机制,构建“一地确证、全网联防”的风险阻断网络。针对非法集资跨行业、跨地域的异质性特征,在横向协同层面,应强化“功能监管”与“穿透式监管”理念,建立由金融管理部门牵头,公安、市场监管、网信等多部门参与机制,明确各方在准入把关、广告清理、情报导侦与联合执法中的法定职责,消除监管真空地带。特别是依托全国性监管规则库,推行“一地确证、全网联防”机制:一旦某地确证新型非法集资模式(如虚拟资产包装、跨境资金盘),即刻通过系统生成标准化的风险识别模板与算法规则,并在全国范围内实现特征共享与同步预警,防止风险的空间转移。在纵向贯通层面,压实“中央统筹、省负总责、市县落实”的垂直责任链条,加强中央对跨省大案的指挥协调,解决异地取证与资产处置难题,同时推动监管重心下沉,利用基层优势对商务楼宇与居民社区实施地毯式排查,打通监管的“最后一公里”。

二是建立多源数据验证与行政司法衔接闭环处置通道。为破解监管数据孤岛与行刑衔接不畅的难题,应探索建立基于多源异构数据的全流程证据链机制。首先,构建多源异构数据验证闭环。通过API接口技术打通市场监管、税务、金融及公安等部门的数据壁垒,将原本割裂的静态数据转化为动态风险图谱。其次,探索行政认定与司法确认的要素化衔接。对行政机关已完成初步事实调查与违法性认定的案件,建立证据格式化移交与司法快速确认程序,降低制度性摩擦成本与交易成本,从而提高违法成本的即时兑现率与监管威慑效应。

三是打造“政府主导+专业赋能+社会参与”的共治生态共同体。非法集资治理不仅是行政与司法的单向管控,更需构建多中心协同的社会治理结

构。一方面,要强化关键主体责任。通过立法明确银行业金融机构、第三方支付机构的反洗钱与风险审查义务,要求其将司法风险指标嵌入业务系统,对异常资金链路实施“预警阻断”;同时,引入律师事务所、会计师事务所等第三方专业力量,建立行业自律公约与执业报告免责制度,在审计与法律服务中发挥监测作用。另一方面,完善数字化的社会监督机制。开发低成本、高效率的公众举报数字平台,并建立隐私保护与数据激励制度;结合常态化的国民金融素养教育,将社会监督线索实时纳入智能监管模型与风险评分体系。通过政府监管、行业自律与社会监督的协同演化,实现非法集资风险治理从“被动式打击”向“内生性防控”的根本性转变。

参考文献:

- [1]蔡尚宣,朱丹青.基于Pu-learning的智能化非法集资异常交易监测[J].中国金融电脑,2024(10).
- [2]曹源芳,蒋志芬.非法集资风险的防范化解[J].中国金融,2019(4).
- [3]陈钢.基于混合神经网络的非法集资风险预测模型[J].计算机与数字工程,2023,51(3).
- [4]胡彦涛,刘莉.非法集资行为的“国家规定”与司法判断标准——以行政犯相关理论为视角[J].东北大学学报(社会科学版),2018,20(4).
- [5]黄辉,秦永彬,陈艳平,黄瑞章.基于BERT阅读理解框架的司法要素抽取方法[J].大数据,2021,7(6).
- [6]黄旭,洪美玲,叶俊,董志强.数据要素助力新质生产力发展的机理与提升路径[J].西南金融,2026(1).
- [7]季洁.区域性非法集资对金融监管的反哺式启发[J].现代经济探讨,2020(6).
- [8]康建华.大数据背景下非法集资犯罪侦查的难点与对策[J].四川警察学院学报,2023,35(6).
- [9]黎若然,张阳,徐莹.基于大数据技术的非法集资监测预警模型构建[J].产业与科技论坛,2025,24(7).
- [10]刘伟.非法集资犯罪“非法性”标准的重拾与展开[J].学海,2020(3).
- [11]刘志伟.非法集资行为的法律规制:理念检视与路径转换[J].江西财经大学学报,2016(1).
- [12]潘淑娟,瞿苇.从制度经济学角度分析非法集资产生的原因[J].新金融,2008(8).
- [13]石奎.金融安全视域下非法集资风险的预防监管[J].

财经科学,2023(5).

[14]石奎,陈凤玲,于子芮.非法集资风险预防监管的合法性审视与安全港规则构建[J].经济法论坛,2025,34(1).

[15]苏如飞.反欺诈国际金融监管动向及启示[J].中国金融,2025(4).

[16]童德华.非法集资犯罪的管制:制度变迁及其有效性分析[J].南京大学学报,2024,61(6).

[17]王天雨,卢宇.数据要素赋能新质生产力发展与新业态培育的耦合关系、内在机制和发展建议[J].西南金融,2025(8).

[18]魏景茹,程万宝.新型非法集资风险成因与防范[J].当代金融家,2021(12).

[19]夏昊晗.刑民交叉案件中合同效力的认定[J].中国法律评论,2025(3).

[20]谢勇,陈小杉.非法集资案件财物处置刑民交叉的规

范路径[J].湘潭大学学报,2019,43(2).

[21]叶良芳.总体国家安全观视域下非法集资的刑法治理检视[J].政治与法律,2022(2).

[22]詹子友.e租宝非法集资的监管视角[J].中国金融,2017(17).

[23]张博,张晓帆,蔡子阳.金融科技、金融普惠与非法集资风险[J].金融研究,2025(1).

[24]Lai J T, Xie J, Cao S, Zhang H. Digital financial inclusion and illegal fundraising in China[J]. Applied Economics, 2022, 54(48).

[25]Liu X, Flora Huang, Yeung H. The regulation of illegal fundraising in China[J]. Asia Pacific Law Review, 2018, 26(1).

责任编辑:冯晶

A Study on the Identification of Illegal Fund-Raising Risks and the Optimization of Regulatory Approaches Empowered by Judicial Big Data

Abstract: Illegal fund-raising severely disrupts financial stability and social order. Due to the acceleration of information dissemination on the Internet, the proliferation of fund-raising methods, and the intensification of cross-regional capital flows, the risks associated with illegal fund-raising have become increasingly widespread, covert, and transnational. The traditional regulatory model, which primarily relies on ex post enforcement, has gradually revealed systemic shortcomings in risk identification, early warning and prevention, and collaborative governance. Against the backdrop of big data, leveraging advanced data technologies to implement ex ante supervision has become the key to breaking this deadlock. Based on a systematic review of the primary models, risk characteristics, and relevant policies and regulations governing illegal fund-raising, this paper provides an in-depth analysis of the practical challenges currently facing the regulation of illegal fund-raising risks in areas such as data integration, risk identification mechanisms, and regulatory coordination. At the same time, from the perspective of judicial big data, this paper utilizes judicial big data on civil cases to uncover the patterns of risk evolution in illegal fund-raising and establishes a technical framework for identifying and rating such risks. Building on this foundation, the paper proposes optimizing China's regulatory approach to illegal fund-raising risks by improving risk information integration mechanisms, advancing tiered and categorized supervision, and strengthening a multi-stakeholder collaborative governance system, so as to providing theoretical support and institutional references for the early prevention and long-term governance of illegal fund-raising.

Keywords: big data; illegal fund-raising; financial regulation; digital transformation; data security; risk prevention and control; new quality productive forces; regulatory technology