

冒烟指数: 大数据监测 互联网金融风险

李崇纲, 许会泉

北京金信网银金融信息服务有限公司, 北京 100101

摘要

提出以冒烟指数为核心技术的大数据监测预警系统。基于大数据金融监管的理念和方法,利用大数据、云计算、人工智能等技术实时监测整合互联网金融多元异构风险信息,并构建以“人、资金、业务”为主线的风险评测指标。该平台可提升并优化互联网金融风险日常监管的手段和效率,为监管机构的决策提供事实依据和数据参考。

关键词

大数据技术 ; 多元异构信息 ; 冒烟指数 ; 互联网金融 ; 风险监测

中图分类号 : F83

文献标识码 : A

doi: 10.11959/j.issn.2096-0271.2018042

Smoke index: big data technologies monitor Internet financial risks

LI Chonggang, XU Huiquan

Beijing Jinxinwangyin Financial Information Service Company, Beijing 100101, China

Abstract

Based on the big data concept and techniques in the financial regulation, smoke index was taken as the core technology to the Internet financial risk monitoring system, and the real-time monitoring and integration of multi-heterogeneous risk information were carried out and a set of “people”, “capital” and “business” risk assessment indexes by using technologies such as big data, cloud computing and artificial intelligence were built. The means and efficiency of routine supervision of Internet financial risks were enhanced and optimized, and factual basis and data reference for the regulatory agencies’ decision-making was provided.

Key words

big data technology, multivariate heterogeneous information, smoke index, Internet finance, risk monitoring

1 大数据监管互联网金融风险的必要性

随着移动互联网等新一代信息技术的迅猛发展,互联网+金融也呈现出爆发式增长的态势,P2P借贷平台、小额贷款公司、股权投资机构、交易场所等各种创新的新兴金融业态不断涌现,并已融入经济社会发展的各个领域。借助国家“互联网+”战略,互联网金融高新技术创新层出不穷,不断冲击着互联网金融行业。金融科技创新更加普惠的金融需求,形成对传统金融的重要补充,是现代金融领域供给侧改革的新动力。

互联网金融是互联网与金融结合的产物,是借助互联网和移动通信技术实现资金融通、支付,并具有信息中介功能的新兴金融模式。金融科技快速发展在带来金融服务和产品不断创新的同时,也由于监管机制和手段的相对滞后,使得互联网金融游离在“一行三会”的金融监管之外,存在的监管漏洞让某些不法分子有机可乘,出现了金融诈骗、风险失控、卷款潜逃之类的问题,影响了金融秩序,对社会稳定、公民财产安全等造成极大的安全隐患。现实中很多事件打着互联网金融和金融创新的旗号冲破监管红线,其涉案金额之大、影响范围之广,成为前所未有的系列涉众非法集资犯罪事件。

金融风险防控是金融业的永恒主题,金融监管理念的转变、监管体制改革和有效化解风险,是关系到我国金融稳定和经济社会平稳发展的重大问题^[1]。互联网金融的统计分析应用是发挥大数据资产自身价值的最有效手段^[2],是构建“智慧政府”“阳光政府”的核心。利用遍布金融监管部门和社会各领域的的数据,在打击互联

网金融风险、融资征信、信用评级、虚假广告监管等领域进行智能分析研究,可以有效实现对重大安全、危机、风险的防范和预警,并可以通过对业务数据的分析,为资源的有效配置、公共安全管理等提供实时的决策支持^[3-5]。

互联网金融风险监测预警的应用针对政府机构内部共享数据、外部舆情数据进行综合风险识别、风险预警、风险排查,是大数据的核心应用之一。围绕互联网金融业务展开全方位的风险识别、风险监测、风险排查以及风险预警等,通过构建风险分析模型量化风险级别,并自动通过短信、移动客户端、邮件、内部协同平台等及时上报给相关的管理部门,辅助金融监管部门开展危机应对,有效支持应急处置工作^[6]。同时,也向公众发布各类风险警示,提供线上消费、投资等日常生活风险提示服务。

金融科技的快速发展使得跨机构、跨行业风险关联度增强。如何进一步完善对金融科技的监管、有效防范金融风险和遏制金融乱象,对监管能力和智慧都是考验。互联网金融的蓬勃发展必然要求各国金融监管当局在监管理念和监管技术层面有所创新^[7-8]。金融的本质就是要解决信息不对称问题,大数据的核心和本质在于应用、算法、数据和平台各个要素的有机结合^[9],大数据监管将成为未来金融监管的主流范式。为防范互联网金融风险和化解潜在隐患,提高金融监管的前瞻性、时效性和精准性,构建以冒烟指数为核心技术的大数据驱动的互联网金融风险监测预警系统,建立“技术驱动型监管创新”思维,在现阶段显得尤其重要和紧迫^[10]。

2 互联网金融风险监测的痛点

传统的监管方式由于人力、机构和信

息化水平等原因的限制不能做到监管到位,始终存在监管视角盲点和隐患,导致监管机构无法准确把握现代金融行业发展的新动向,对其风险和监测的认知与实践相对滞后^[11]。显然,互联网金融的飞速发展与传统监管手段的相对滞后,导致单靠人防的传统监管模式已无法应对当前的严峻形势,亟待通过大数据等技术手段加以预警和监测,将金融风险 and 隐患化解在萌芽状态,切实维护地区金融稳定和人民群众的财产安全^[12-13]。我国互联网金融风险监测的痛点如下。

(1) 互联网金融监管相对滞后

新兴金融业态是传统金融机构与互联网企业利用互联网技术和信息通信技术实现资金融通、支付、投资和信息中介服务的新型金融业务模式。互联网与金融深度融合是大趋势,它将对金融产品、业务、组织和服务等产生更深刻的影响,正成为金融领域研究的热点。但不可否认的是,伴随着互联网金融的迅猛发展,监管机构对其风险和监管的认知与实践却相对滞后,很多新兴金融业态长期游离于正规金融监管之外。从2014年开始,以P2P借贷平台为代表的我国互联网金融产业爆发出许多危机(如借款人失踪、资金不知去向、平台无力支付),导致其无法运营。这些事件使得业界对互联网金融风险的关注加大,对互联网金融监管提出了更紧迫的要求。当前,国家高度重视互联网金融领域的健康稳定发展。2016年4月,国务院组织的力度空前的互联网金融专项整治拉开大幕,而P2P借贷平台由于众多不合规平台导致乱象频出,成为重点整治对象。

(2) 互联网金融监管技术手段亟待加强

由于互联网具有不分地域、快速传播、涉众面广等特性,互联网金融监管相比传统行业监管面临更多的困难。由于人力、机构和信息化水平等原因的限制,不

能做到监管到位,始终存在监管盲点和隐患,导致我国现有的互联网金融监管模式存在发现难、研判难、决策难、控制难、处置难的“五难”现象,缺乏对各类信息汇总风险分析平台。该平台能够获取举报信息,开展企业排查,并具有数据挖掘分析功能,将线上和线下信息进行整合,自动进行高级、复杂的信息处理和数据挖掘分析,结合不同行业的各类模型,配合以强大的人机交互能力,大大提高情报分析的科学性、准确性和时效性。在部分非法集资犯罪活动还处在初期,并未造成重大经济损失和社会危害时,平台可以快速发现、快速定位,从而实现有效的风险防范和事前预测、预警,具有重要的意义。

(3) 跨部门协同不够

当前,互联网金融各类创新应用呈现出多样性特点,各相关政府部门和监管机构仍然存在各管一段、分段自治、信息孤岛等问题,急需建立统一的综合监管协同工作平台,促进信息共享,建立金融监管治理模式。各地区需要建立立体化、社会化、信息化的监测预警体系,充分发挥网格化管理和基层群众自治的经验和优势,群防群治,贴近一线开展预警防范工作。创新工作方法,充分利用互联网、大数据等技术手段加强对非法集资的监测预警,为跨时空的非法金融活动风险超前监测预警工作提供有效的支持。

3 冒烟指数指标体系与建模流程

冒烟指数作为互联网金融监测指标体系,其最初构想来源于“森林着火要冒烟警示”,通过烟与火的形象比拟来推断冒烟指数与集资类企业从事非法集资程度的关系,创新性地把非法集资风险预警同大数据技术结合。冒烟指数在金融风险监管中

运用了大数据、人工智能、机器学习、知识图谱、自然语言处理等技术,应用逻辑主要是导入大量相关数据,利用机器学习形成知识图谱或者建立模型,通过不同算法和神经网络应用预测互联网金融风险,以达到识别风险、量化风险和把控风险的目的。

具体来说,冒烟指数使用工商数据、招聘数据、舆情数据、法院行为信息数据、投诉举报数据、客户提供风险企业数据、监管机构数据和网络爬虫数据等多源异构数据共150个数据项,从中提取320个变量,形成非法性指数、收益率偏高指数、投诉举报指数、传播力指数、特征词命中指数等多角度学习、可增可减的风险分析子模型,最终通过机器学习模型和专家研判模型共同赋权,得到信用风险评分。每一个子模型都从不同的角度预测集资类企业的信用风险状况,克服了传统信用风险评估中单个模型考虑因素的局限性,使预测更为精准。经典的冒烟指数建模流程如图1所示。

冒烟指数风险模型目前已经应用到许多实际项目中,并且获得了金融监管领域的广泛认可。笔者团队在与国家级监管平台以及各级地方监管部门的对接与交流中,对冒烟指数风险模型进行不断的完善与应用,出色地完成了许多大数据监测平台与软件的开发,已经总结出冒烟指数量化金融风险的处理策略依据。

冒烟指数作为量化集资类企业非法集资风险的标准,根据目标企业非法集资风险大小的不同,将冒烟指数分为5个等级,每个等级呈现不同的风险特征。这为监管机构界定其非集界限、辅助决策提供了差异化处理策略的依据(见表1)。

依据互联网金融风险特征,针对不同的业态构建一套完善的互联网金融风险指标体系,是量化评测、监测预警金融风险的关键环节。冒烟指数拟从以下六大方面

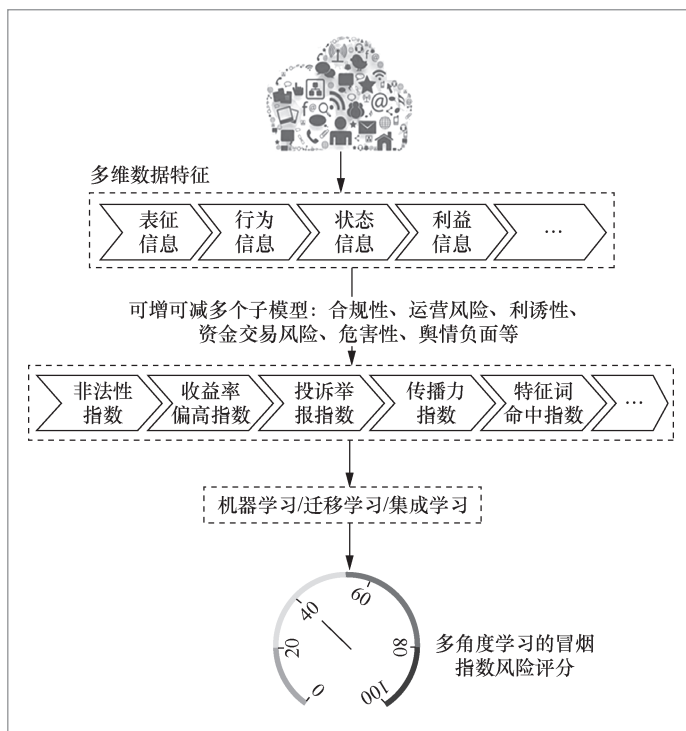


图1 冒烟指数建模流程

挖掘类金融企业风险。

- 重点区域/领域风险排查。利用定量模型,针对重点地区或行业进行大规模风险排查,准确挖掘出高风险企业,同时对主要风险点、企业类别分布等进行可视化展示。

- 集群风险挖掘。针对特定园区、交易所等内部企业成员间的关联关系进行挖掘,可有效识别风险团体及关键人员企业,及时杜绝集群内部的风险传递。

- 关联风险挖掘。通过分析目标企业完整关联图谱,精准锁定核心人员企业,同时根据其关联结构及关联企业类型挖掘其潜在自融、关联担保等风险^[14]。

- 人员分析。通过对受害人(投资人)的投资金额、所在地及职业等进行多维度的统计分析,辅助相关部门完成事后联动处置工作。

- 资金流风险分析。识别企业各类重点核心账户,锁定重点参与人员,同时准确计算

表1 冒烟指数量化金融风险及提供差异化处理策略依据

冒烟指数分值	风险区划分	差异性处置策略	风险特征备注
[80,100]	取缔类	移交线索	经营模式不合规、非法集资行为明显；存在大量负面新闻报道及网络投诉，易导致投资人团体性上访事件等，非法集资风险极高、不可控
[60, 80)	重点关注类	约谈整改	经营模式不合规，存在大量负面新闻报道及网络投诉行为；非法集资风险很高且有可能上升，处于高风险预警区
[40, 60)	重点监测类	规劝改正	非法集资风险较高，经营模式存在一定的不合规行为；媒体负面信息报道较少，网络投诉较多；风险较大但具有可控性
[20, 40)	可疑类	持续监控	存在非法集资的可能性，经营模式有某些不合规行为，但危害程度较小，存在一些网络投诉信息，但不涉及公司经营的根本问题
[0, 20)	正常类	—	没有明显非法行为，经营模式和宣传方式较为合规，收益率处于业内正常范围，网络投诉很少，鲜有负面新闻报道

出实际非法集资总金额及资金最终流向。

- 运营风险分析。对企业集资金额、利率、待还金额等每日变动情况进行实时监测，及时发现企业运营中的异常状况。

4 大数据监测系统思维业务逻辑

基于当前国家有关部门对地方政府落实属地监管责任，加强互联网金融监管的要求，结合各地目前面临的互联网金融发展的形势和传统监管手段存在的不足，以冒烟指数为核心技术的大数据监测互联网金融风险系统解决监管机构面临的“发现难、研判难、决策难、控制难、处置难”的痛点，通过运用互联网、大数据、机器学习等技术手段，构建以“人”“资金”“业务”为要素的风险分析模型，建立风险评价指标，实现发现风险、评估风险、固化证据、判断趋势、及时干预和联合打击等联动工作，推动互联网金融治理由传统被动监管、粗放监管、突发式应对向主动监管、精准监管和协同监管模式转变。

以冒烟指数为核心技术的大数据监测预警系统是以“数据驱动，追根溯源”为指导思想，在加强微观金融数据搜集的基础上，通过统计分析，利用数据挖掘等技术

手段强化对互联网金融风险的识别，并监测跟踪风险产生的根源及传导路径，从而提高风险识别和捕获能力。其系统构架如图2所示。

以冒烟指数为核心技术的一体化大数据系统平台基于统一的云基础设施、云服务支撑平台及大数据管理平台，构建各应用子系统，确保系统的灵活性和可扩展性。

(1) 建立地方金融业态的动态信息数据库

通过建立数据采集与交换系统动态收集各监管部门、新兴金融业机构、投资者、公众等数据信息，重点收集P2P网络借贷平台、小额贷款公司、股权投资机构、交易场所以及非法外汇交易、非法集资、金融诈骗等各类金融基础数据信息，建立全省（市）地方金融信息数据库，掌握新兴金融的基本情况。另外，通过政府购买服务的方式（如腾讯、北京奇虎科技有限公司、网贷之家等相关第三方金融征信信息数据），补充完善本地金融信息数据库。

(2) 构建“人”“资金”“业务”风险分析模型

目前，通过多渠道采集的数据信息，构建以“人”“资金”和“业务”为要素的分析模型，建立模型评价和风险评价指标。



图2 以冒烟指数为核心技术的大数据预警系统构架

发挥“网在看”对互联网金融新业态的主要风险点设置预警指标的作用，实现动态监测。以“人”为主线的模型包括：企业股东和核心高管人员，关联工商、公共安全、税务部门及有关征信机构的个人和机构的信用信息，形成人的全息画像，并建立风险评估模型。以“资金”为主线的模型，即针对大额资金的异动情况，系统与中国人民银行系统对接，或收集各类金融机构定期提供给总部或中国人民银行的系统监测的异常企业和个人名录，掌握大额资金的账户异动情况，收集资金流数据信息，然后通过从相关数据中提取的非法集资行为

资金模型特征，构造资金分析模型。以“业务”为主线的模型包括对公司经营业务数据的文本抓取。通过政府购买服务与自建舆情搜索平台的方式，收集第三方数据机构反馈的业务数据信息和其他资源进行数据处理，根据非法集资犯罪案件的特点分别进行模型建设。

(3) 引入大数据非法集资监测技术

通过各类舆情信息收集、风险指数、企业全息画像、关联方拓扑图等，对本地线上非法集资行为进行监测预警。另外，针对隐蔽度高、监测难度大的“线下”非法集资行为，借鉴北京市金融工作局“随手拍”

软件技术,网格化发动各区、街道办事处、社区进行逐一筛查,同时建立非法集资举报奖励制度,广泛发动社区、群众的积极性,检举揭发非法集资活动,一旦出现风险苗头,根据问题的级别和种类划分,及时将预警信息直接报送给相应职能部门。

(4) 提升并优化日常监管的手段和效率,为政府决策的制定提供统计分析报告

利用信息化手段对P2P网络借贷平台、小额贷款公司和要素交易中心的设立、变更和注销以及日常经营活动等实施日常监管,充实监管力量和手段,进一步优化政府服务效率。依托系统提供的对宏观经济金融形势、金融运行状况、行业发展的统计、分析等,为决策的制定提供参考。

(5) 建立完善的信息披露,不断拓宽系统的对外应用功能

信息的及时有效披露是防范金融风险的重要环节,也是新兴金融业态特别是P2P借贷平台监管的重要政策取向,因此,系统应具备完备、及时的信息披露。此外,本系统不仅仅是政府部门内部使用的系统,也是面向大众的系统,未来将根据本系统的不断完善,拓宽各类应用模块(如投资者教育、宣传等)。

5 冒烟指数监测互金风险的实践

冒烟指数作为互联网金融风险“穿透式监管”的利器,是互联网大数据监测预警平台的核心技术之一,已成为非法集资风险监测领域内的评测标准,并逐步应用于金融监管。

2015年,北京市金融工作局率先应用大数据技术建设了“大数据打击非法集资监测预警云平台”,应用于打击非法集资专项整治行动中,加强对非法集资的监测预警。该平台以冒烟指数为核心技术,

利用互联网思维和先进技术手段对非法集资信息进行高效、全方位的实时监测,快速发现企业的非法集资线索,为政府制定决策提供事实依据和数据参考,目前正在使用并在实践中发挥了良好效用。

截至2017年10月底,预警平台共监测全市金融企业168 387家,风险企业1 775家,分布于全市16个区。其中高风险企业(冒烟指数 ≥ 60 分)共计46家。根据监测数据及结果,制作各类报告共计500余份,在监测预警、风险排查方面发挥了重要作用。

2016年,为有效监测预警新金融业态风险,深圳市人民政府金融发展服务办公室全力开发建设了深圳市非法集资监测预警平台,建立深圳市早发现、早打防的监测预警系统和工作机制,推动深圳金融治理由传统被动监管、粗放监管、突发式监管转变为主动监管、精准监管和协同监管,维护金融稳定。目前,该系统已经成功上线,实时监测深圳市20余万家新金融企业的非法集资风险,及时预警高危企业,辅助监管决策的制定。

6 结束语

冒烟指数基于大数据管理工具和分析方法,能有效构建互联网金融风险监测预警系统,可先验性地识别互联网金融中的潜在风险,从而为监管机构提供互联网金融行业风险监管和治理依据,有利于保障互联网金融行业可持续健康发展和社会经济的稳定。

(1) 解决互联网金融监测的标准化问题

互联网金融监测缺乏权威的、统一的指标体系,不利于互联网金融监测业务走向标准化、规范化、成熟化。推广冒烟指数在P2P网贷行业非法集资风险监测中的应用,构建互联网金融风险监控统一的指标

体系是当前开展相关业务的重中之重,以防出现不同监管机构之间无法共享分析模型、无法共享分析成果、各自为政的局面,从而解决了互联网金融监测的方法论问题。

(2) 实践大数据金融监管成功落地

冒烟指数以大数据为驱动,利用机器学习和数据挖掘等技术实时处理多源风险信息,建立风险预测监测分析模型,以达到及时、有效地识别流动性风险的目的,从而帮助金融机构规避流动性风险,解决了传统金融监管方式和飞速发展的金融服务及产品之间的矛盾。

冒烟指数目前已经成功为数十家省市区级地方监管部门服务,获得了各级政府的认可,有效实现了对互金风险的监管,在未来的金融监管模式发展中走出了新的方向,是一种有效的、及时的和低成本化的大数据金融监管方式。

(3) 解决地方金融监管资源有限的问题

以冒烟指数为核心建设的大数据非法集资监测预警平台在监管报告自动生成、客户或员工资质审核等多方面大幅降低人员工作量,借助了技术手段提升信息采集和分析能力,取代原来耗费更多人力的现场检查 and 书面调查方式,很大程度上缓解了地方监管人手明显不足和专业人才匮乏的困境^[15-16],大大降低了传统监管方式中运营成本的需求。与此同时,以冒烟指数为核心的大数据平台还提高了监管覆盖面和效率,弥补了县级及以下区域出现的严重的监管真空,在很大程度上解决了地方监管部门因行业监管手段严重不足、缺乏必要的执法权、只能采取较多的准入监管和行政性监管的难题。

参考文献:

[1] 陈志武,黄益平,巴曙松.中国金融改革,未来会怎样[M].杭州:浙江大学出版社,2017.

CHEN Z W, HUANG Y P, BA S S. The future of financial reform [M]. Hangzhou: Zhejiang University Press, 2017.

- [2] YANG H, YI D H, XIAO H W. Study on large Internet financial risk early warning based on data analysis [J]. Modern Management Science, 2014(4): 3-5.
- [3] MARKL V. Breaking the chains: on declarative data analysis and data independence in the big data era[J]. Proceedings of the VLDB Endowment, 2014, 7(13): 1730-1733.
- [4] KSHETRI N. Big data's role in expanding access to financial services in China[J]. International Journal of Information Management, 2016(36): 297-308.
- [5] “企业风险防范预警管理有关问题研究”课题组. 基于大数据的企业风险防范预警路径探究[J]. 中国市场监管研究, 2017(7): 63-66. Enterprise Risk Prevention Early Warning Management Related Issues Research Group. Based on big data enterprise risk prevention warning path[J]. China Market Regulatory Research, 2017(7): 63-66.
- [6] 杨扬,周一懋,周宗放. 基于文本大数据的企业信用风险评估[J]. 大数据, 2017(6): 44-50. YANG Y, ZHOU Y M, ZHOU Z F. Credit risk evaluation with text big data from text[J]. Big Data Research, 2017(6): 44-50.
- [7] 王达. 宏观审慎监管的大数据方法: 背景、原理及美国的实践[J]. 国际金融研究, 2015(4): 55-65. WANG D. Big data methods of macro prudential supervision: background, principle and American practice [J]. International Finance Research, 2015 (4): 55-65.
- [8] 杨东. 互联网金融治理新思维[J]. 中国金融, 2016(23): 43-45. YANG D. New thinking on Internet financial governance [J]. China Finance, 2016 (23): 43-45.
- [9] 李涛,曾春秋,周武柏,等. 大数据时代的数据挖掘——从应用的角度看大数据挖掘[J]. 大数据, 2015(4): 1-24. LI T, ZENG C Q, ZHOU W B, et al. Data mining in big data era - big data mining from an application perspective [J]. Big

- Data, 2015(4): 1-24.
- [10] 施水才, 王丽. 探索与发展: 大数据驱动的互联网金融风险监测预警研究[J]. 当代金融家, 2017(6): 45-47.
SHI S C, WANG L. Exploration and development: the big data driven risk monitoring and warning system for Internet finance [J]. Modern Bankers, 2017(6): 45-47.
- [11] 尹海员, 王盼盼. 我国互联网金融监管现状及体系构建[J]. 财经科学, 2015(9): 12-24.
YIN H Y, WANG P P. The current supervisory situation and system formation of China's Internet finance [J]. Finance & Economics, 2015(9): 12-24.
- [12] 霍学文. 关于云金融的思考[J]. 经济学动态, 2013(6): 33-38.
HUO X W. Reflections on cloud finance [J]. New Economics, 2013 (6): 33-38.
- [13] 杨东. 互联网金融监管体系探析[J]. 互联网金融创新与监管, 2014(8): 45-46.
YANG D. Internet financial supervision system [J]. Internet Finance Innovation and Regulation, 2014(8): 45-46.
- [14] 刘新海. 基于企业关联关系的信用风险分析新思路[J]. 征信, 2016(3): 16-20.
LIU X H. A new idea of credit risk analysis based on enterprise relationship [J]. Credit Research, 2016(3): 16-20.
- [15] FCA. Call for input on supporting the development and adopters of RegTech[R]. 2016.
- [16] Deloitte. RegTech is the new FinTech - How agile regulatory technology is helping firms better understand and manage their risks[R]. 2016.

作者简介



李崇纲 (1978-), 男, 北京金信网银金融信息服务有限公司总经理, 中国计算机学会大数据专家委员会委员, 中国互联网金融协会信用建设专业委员会委员, 北京互联网金融行业协会合规专业委员会, 中国首席数据官联盟专家组成员, 中关村大数据产业联盟专家组成员。目前主持开发大数据防控金融风险与智能决策支持系统, 长期致力于推进大数据金融监管科技, 主要研究方向为监管科技、人工智能、大数据、云计算。



许会泉 (1978-), 男, 北京金信网银金融信息服务有限公司副总经理、研发总监, 具有多年一线开发经验及团队管理经验, 研发设计全国第一款大数据监测预警金融风险产品, 建设、搭建监管科技产品生态体系, 应用于全国各省市金融监管机构。主要研究方向为监管科技、人工智能、大数据、云计算。

收稿日期: 2017-12-13