专家洞察

箭在弦上
步步为营

强监管、强合规下的金融机构数据治理应对方略

IBM 商业价值研究院

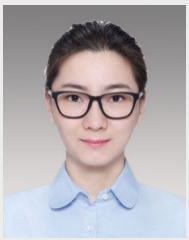
IBM

主题专家

吴大维
IBM GBS CBDS 团队
副合伙人
wudavid@cn.ibm.com

张玉明
IBM GBS CBDS 团队
首席数据架构师
zhangyum@cn.ibm.com

周茜
IBM GBS CBDS 团队
数据治理高级顾问
bjzhouxi@cn.ibm.com



韩玲玲
IBM GBS CBDS 团队
数据治理资深顾问
hllhanbj@cn.ibm.com

陈晓熹
IBM GBS CAI 团队
数据治理高级顾问
bjcxchen@cn.ibm.com

王莉
IBM 商业价值研究院
高级咨询经理
gbswangl@cn.ibm.com

扫码关注 IBM 商业价值研究院

官网　微博　微信　微信小程序

谈话要点

监管单位对数据治理提出新要求

银保监会明确提出数据质量专项治理要求，不断加强对金融机构的业务开展合规问题的监管以及对银行各类业务风险问题的监控。

专项数据治理应当聚焦 3W1H

专项数据治理工作的开展需要关注在什么时点、由哪些人、做哪些动作、交付什么成果物。

数据治理不是项目，而是旅程

数据质量提升不是一蹴而就的。金融机构需要构建常态化的工作机制，开展组织保障制度、质量规则设计、问责通报等一系列举措。

—

道高一丈

从 1104 报表到 EAST 报送，从监管统计报表到基础数据报送，银行的“表哥”“表姐”们持续与 Excel 相爱相杀，标准不明确、口径不统一、数据找不到、问责压力大，数据不完整、不准确，报送不及时等的数据质量问题让多少银行人焦头烂额。

2020 年 5 月 20 日，银保监会下发了《关于开展监管数据质量专项治理工作的通知》（以下简称《通知》），[1] 明确银行业机构需要就 1104、EAST 和客户风险报送的数据开展数据质量的专项治理。

事实上这已经成为近年来监管单位的常规动作了，随着金融市场乱象问题的层出不穷，监管当局不断加强对金融机构的业务开展合规问题的监管以及对银行各类业务风险问题的监控。不仅仅是 1104 报表、EAST 报送，还包括反洗钱、反欺诈、关联交易等，都已成为监管单位的重点核查对象。

正所谓“魔高一尺，道高一丈”， 当下的金融风险呈现着各类业务交叉性强、隐蔽性高、传染性广的特点，监管当局在数据治理上的铁拳动作体现了其“强监管、强合规”，“深化整治银行市场乱象”的决心。

虽然监管体制在 2017 年、2018 年进行了改革，但是依靠现有的人力与监管模式很显然难以应对如此繁重的监管任务。传统的监管模式主要依靠监管报表与监管指标以及监管当局现场检查，不论是监管的细度、还是监管的覆盖面以及穿透性都难以满足。此外，从监管的经验看，监管往往滞后于金融创新，且往往只能进行金融风险事后监管。金融行业发展速度快，主体多、跨市场风险传染性高，监管人力资源极其有限。因而监管当局亟需以科技改善金融监管，利用大数据、人工智能技术监控与防范金融风险，提高监管的主动性和权威性。

站在商业银行视角，银行经营的无非两类资产，一类是“资金流”，一类是“数据流”。资金流最后又体现为数据流。大部分商业银行都已经意识到数据是战略资源，数据是生产要素，并且逐步加大在数据治理和数据应用方面的投入。各类数据基础设施逐步成熟，具备数据管理经验和能力的人才队伍逐渐壮大。无形中这也为监管当局开展以数据为核心的穿透式监管提供了必要的条件。

EAST 数据覆盖面广、数据之间关联性好、数据粒度细、穿透性强，透过 EAST 可以全面分析银行各类业务，精准定位银行业务违规问题，评估银行各类风险，在银保监对银行“市场准入、非现场检查、现场检查”三驾马车都发挥重要作用。对于银行而言，EAST 为银行合规管理提供了监管视角的思路，EAST 大量的数据也是银行重要的数据资产，甚至有些银行基于 EAST 数据规范来制定本行的数据标准规范，用于指导应用系统的开发建设，所以银行完全可以利用 EAST 开展合规管理和数据治理提升工作。

监管单位已经“亮剑”，但是商业银行并未完全做好准备。不少商业银行负责监管报送的人员，在接到《通知》这项任务时，不禁会发问:

— 监管数据质量专项治理做什么，对监管数据质量规范、口径和要求理解透彻了么?

— 监管数据质量专项治理应该如何开展，选择哪些专题，这些专题间有何联系?

— 专项治理之后银行的数据质量会有多大程度的改良，能解决“表哥”“表姐”们的困扰吗?

— 这一年专项治理工作结束后，应该建立什么样的常态化的机制来保障银行数据质量的持续提升?

对于上述几个困惑，IBM 希望能够带您拨云见日，步步为营。

规圆矩方，准绳嘉量

我们先来看看监管的思路和重点，以及企业所面临的挑战。

监管动态

2018 年 5 月 21 日，银保监会下发《银行业金融机构数据治理指引》（以下简称“《指引》”），[2] 其中，对于监管数据，明确要求银行业金融机构应当将监管数据纳入数据治理，建立工作机制和流程，确保监管数据报送工作有效组织开展，监管数据质量持续提升，法定代表人或主要负责人对监管数据质量承担最终责任。

2020 年 5 月 20 日，银保监会又下发了《关于开展监管数据质量专项治理工作的通知》，通知中明确给出数据质量专项治理的工作方案，同时对数据范围和数据质量提出了更高要求。

从 2018 年 5.21《指引》推出到 2020 年 5.20《通知》的发布， EAST 报送已进入了 4.0 时代，共包括十个监管主题域、66 张报表、1852 个数据项和新增了 897 条监控规则，可以看出监管对报送的数据范围进一步扩大，要求进一步精细化，也意味着监管数据体系逐步迈向成熟，正在往更标准化、效率化的方向发展（见图 1）。

监管不断发出对于数据治理强监管的信号，且相关规范更明确、更落地，甚至更严厉。2020 年 5 月初就 EAST 开具罚单，涉及 8 家银行，共计 1770 万元。[4] 数据报送不满足监管要求被监管开具罚单并不少见，而此次 EAST 的罚款，恰恰是在《通知》发布之前，这一次的信号更像一次警钟：数据不治理，那金融机构就会被监管治理。这也更符合《指引》的要求，数据治理纳入公司治理，并与监管评级挂钩。

监管思路解读

我们认为，可以从以下三个方面来解读本次监管的思路：

监管脉络：《指引》是宏观的，从组织、机制、流程等方面对数据治理提出的要求，例如，第四章用整个章节来明确数据质量控制机制，包含了源头、监控、整改等，而《通知》是微观的，以数据为抓手，问题为导向，追根溯源。因此二者结合，管理上强化，数据上控制，有针对性的逐步提高改善，使得数据质量的提升呈现常态化、长效化趋势；

监管重点：覆盖报告报表和源头数据，覆盖指标数据和基础数据，多监管视角对数据质量提出更高要求。从发文要求来看，重点排查核心监管指标，那对于核心监管指标与 1104 报表中指标的关联，核心监管指标拆解后的基础数据与 EAST 数据的关联，核心监管指标与客户风险报送数据的关联，这些也许都会作为数据真实性、准确性检查的重点。

从处罚情况来看，银保监会此次开出的罚单，均与监管标准化数据报送质量有关，处罚事由主要包括分户账信息、关键且应报字段信息、资金交易信息漏报错报等问题，一方面体现出监管对信息报送的重视度，另一方面也反映出银行对信息申报的重视程度有待提高。

基于本次监管处罚情况，分析出以下四个方面的重点问题：

第一，分账户明细数据漏报很严重。有 8 家银行存在“分户账明细数据记录应报未报”的问题，7 家银行存在“分户账账户数据应报未报”的问题，均收到监管公开行政处罚。

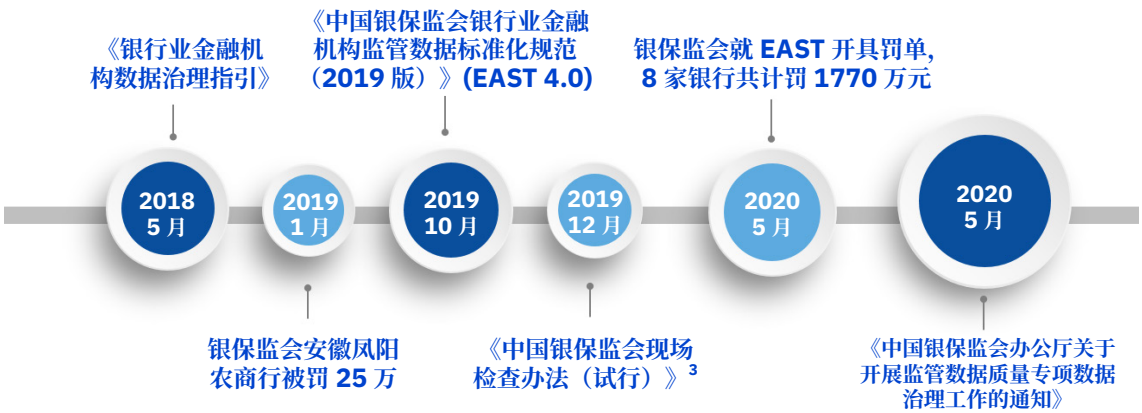
第二，关键且应报字段漏报错报较为突出。关键字段的缺失、错误会直接或间接反映出金融机构的数据质量问题。

第三，资金交易、贸易融资交易、信贷资产转让和贷款核销业务也存在漏报错报情况。而“理财业务”作为新增报送的数据报表，4 家银行机构因“理财产品数量信息漏报”被监管处罚。

第四，信贷业务数据质量尤为不高。信贷业务是业务活动重中之重，也是监管关注的风险重点。

监管手段：自查自评和监管现场检查相结合，《指引》中已经明确要求每年银行业金融机构要开展数据治理的自评估，而此次有针对性的数据质量检查，数据范围、机构范围覆盖之广还是首次，对于发现问题的整改具体到时间，落实到人，管理形成闭环，对于银行保险机构来说，是一次巨大的挑战。

图 1
近两年数据治理方面的监管动态



企业面临的挑战

面对监管提出的新要求，企业面临一系列的挑战。

数据治理落地困难：根据《指引》要求，各家商业银行均需建立完备的数据治理体系框架，但在落地执行过程中仍存在不足，包括数据标准内容更新不及时；数据质量问题处理机制未形成闭环，效率低；主数据及数据生命周期管理有待加强等问题；

系统竖井、数据割裂：受限于法规、监管要求及企业的管理现状，不同业务部门间的数据并未完全连通，尤其在取数、用数过程中存在壁垒。同时，企业级架构不规范、不成熟导致信息系统建设存在一定程度的割裂，系统竖井现象突出，数据关联关系构建和还原需要耗费大量手工工作；

职责不清、意识不足：业务部门更多关注于业务的经营和发展，数据意识不足，对数据管理并不敏感，尤其对监管报送工作重视程度不足，遇到数据问题时，往往将责任推脱给信息系统建设部门，并未承担起数据治理的相应职责；

数据治理人才匮乏：虽然金融机构日益加大对金融科技人才建设和培养方面的投入，但是数据治理领域专业人才的匮乏仍是不争的事实。大部分商业银行建立了专职的监管报送团队，但是人员数量极其有限，承担了超负荷的工作量，数据治理又是一项专业性和技术性极强的工作，既要有扎实的银行保险实务基础，也需要有高超的分析问题和解决问题的能力。专业人才的匮乏也制约了监管数据治理工作的开展。

专项治理，箭在弦上

长期以来，银行保险机构的业务开展、数据分析等严重依赖IT。但其实很多银行和保险机构（尤其是地方中小银行）的业务和管理粗放、数据散乱、数据质量不高。导致数据价值无法完全体现，而数据资产管理和数据变现也大多停留在蓝图规划阶段。

大多数银行和保险机构近年来在业务快速发展的过程中，积累了客户数据、交易数据、外部数据等海量数据。数据质量的主要问题大多体现在数据准确性、完整性的欠缺，数据时效性、适应性的不足，进而导致监管报送不准确、不规范等结果。

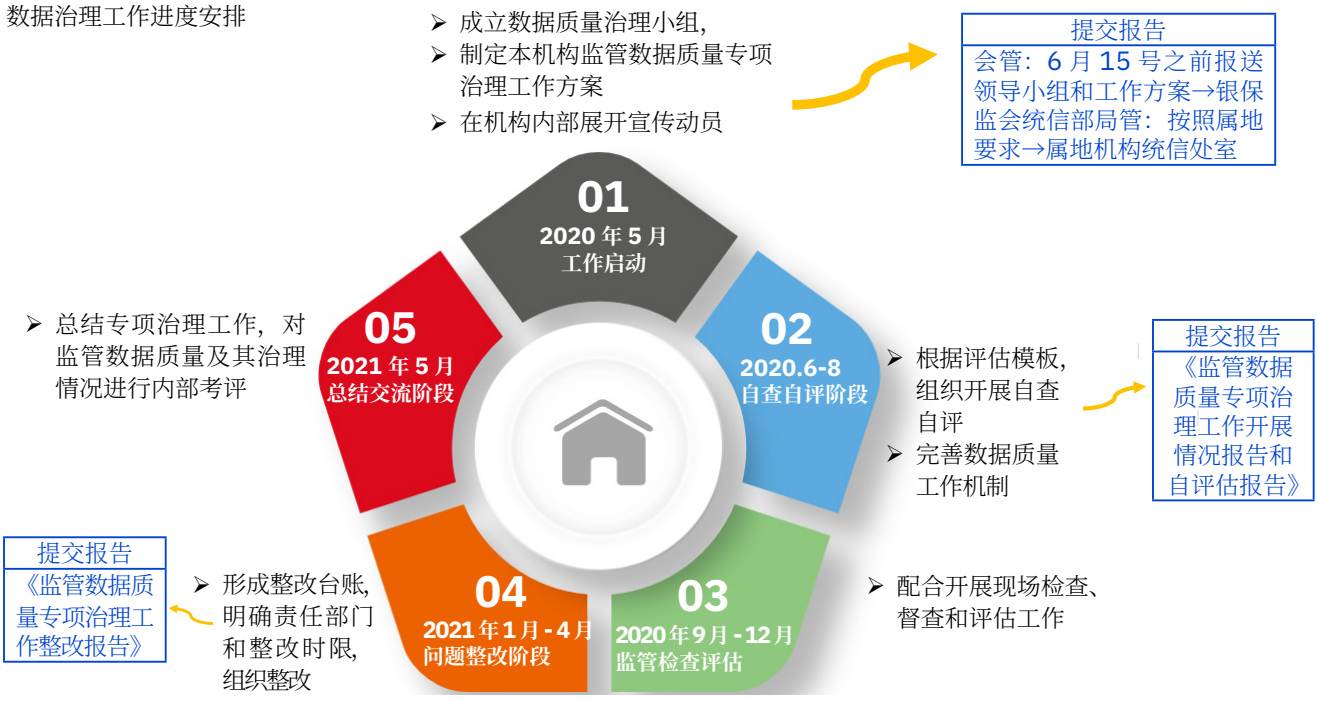
本次监管数据质量专项治理提出了“提高认识、压实责任；突出重点、标本兼治；强化整改、完善机制”的总体要求，覆盖所有银行与保险类金融机构，数据范围包括监管数据及相关源头数据，要求通过为期 1 年的数据治理专项工作（时间从 2020 年 5 月开始到 2021 年 5 月底结束），以监管数据质量问题为导向，通过银行保险机构开展专项自查自评和监管检查评估双向驱动，切实提升监管报送数据和相关源头数据的质量，以优质的数据信息服务于监管工作，并促进银行业保险业高质量发展。具体内容如下：

总体工作要求

监管数据质量专项治理的总体工作要求包括以下三个方面：

图 2

数据治理工作进度安排



提高认识，压实责任：监管数据治理是一项基础性、系统性工作，银行保险机构应充分认识提升监管数据质量的重要意义，把数据质量作为有效监管的基础。银行保险机构应严格落实监管数据质量的主体责任，确保监管相关数据治理要求落实到各环节；

突出重点，标本兼治：应以发现的监管数据质量问题为抓手，确定重点监督和检查的机构范围、数据范围和时间区间。同时注重追根溯源，强化相关源头数据质量治理，夯实管理基础，补齐组织、制度、机制、系统等方面的工作短板；

强化整改，完善机制：应针对自查自评和监管检查评估中发现的问题和薄弱环节，明确整改分工和时间表，确保问题整改按计划和要求落实到位，建立全面提升监管数据质量的长效工作机制。各银保监局要督促切实落实整改措施，强化监管问责，进一步加强数据质量监管机制建设。

专项治理范围

监管数据质量专项治理范围包括机构范围和数据范围，通知中要求覆盖所有银行和保险机构，包括大型银行、股份制银行、保险集团、保险公司等；数据范围包括监管数据及相关源头数据，其中，监管数据包括：非现场监管（1104）、客户风险、监管数据标准化（EAST）、保险统计信息等。

工作进度安排

根据银保监会的通知要求，[1, 5] 对于银行保险机构来说，有以下几个关键时间点（见图 2）：

— 2020 年 5 月：成立数据质量治理领导小组，制定监管数据质量专项治理工作方案；

— 2020 年 6-8 月：开展自查自评，完成数据质量工作机制；

— 2020 年 9-12 月：银保监进行数据质量现场检查、督查和评估工作；

— 2021 年 1-4 月：形成整改台账，落实责任制，上报整改报告；

— 2021 年 5 月：对监管数据质量专项治理情况进行内部考评。

结合银保监本次监管要求来看，对监管数据颗粒度越来越细，覆盖范围越来越广，基本覆盖银行保险各类业务和各个业务系统，从监管角度出发，监管合规数据报送质量很大程度上会有所改观，而对于银行保险机构而言，借监管东风、将监管报送数据作为切入点，可以有效的推进整体的数据治理工作，促进数据质量的提升。

数据治理，全面提升数据质量

某省级农商银行自成立以来，积累了大量客户数据和交易数据。由于未进行系统化的数据治理，导致经营分析数据不准确，对客户 360 视图分析、客户精准营销难以有效进行，并且在 EAST 和反洗钱等报送中时常发现数据问题，如数据填报不准确、获取不到数据等。该银行与 IBM 合作，全面启动数据治理工作，重点从日常和专项两个方面提升数据质量。

该项目围绕 IBM“八横四纵”数据治理方法论，从数据治理的组织、流程、制度及技术支撑构建入手，搭建数据质量的事前防范、事中监控、事后改善管理保障体系。基于外部监管要求和内部用数需求，收集并梳理全行各类数据问题，开展对数据问题的探查、分析和提升方案制定。针对问题暴露较多的“客户”信息，开展了客户识别要素的专项治理，分别从存量数据和增量数据两大范围，抽取不同时间段 2 个核心业务系统、10 余张表展开数据根因分析，按不同的维度分析出无效客户数量、有效客户数量、实际客户数量等，并提出相应的问题解决方案，实现了客户主题数据质量的有效提升。

通过对全行客户数据进行全面的质量摸底，对近 20 年来积累的客户信息从三要素到十五要素进行探查和分析，识别出大量无效的、重复的、错误的客户数据。该项目还在业务制度建立、批量数据清洗、增加系统校验、线下人工走访等领域，开展了有针对性的解决方案落地实施，并同步推动业务流程规范化和系统功能优化，为客户信息经营分析、精准营销和监管报送数据的准确性打下了良好基础。

多方参与，场景结合

某政策性银行 2009 年行内遵循“数据治理，标准先行”的IT发展战略，全面启动了数据标准体系的规划设计工作，并同步支持企业级数据仓库的建设。随着数据标准发布及数据仓库、统一报表和高管驾驶舱等系统的上线，该行建设覆盖统计报送、经营分析的数据应用体系初见成效。但由于工作重心侧重于 IT 建设，也同时面临一部分数据管理滞后的问题，基于相关管理提升诉求，该行于 2013 年启动了常态化的数据治理项目。

在与 IBM 的长期深入合作中，该银行完成数据管控与应用的专项规划，完善数据治理的管理制度、建立数据质量考核和问题通报机制，全面设计数据质量检核规则，进行关键领域的重点数据质量提升，开展常态化的数据质量考核，提升了全行的数据质量和数据治理水平。

通过多期项目执行中关键问题的解决，总结出以下两方面经验，对有效开展数据治理工作具有重要意义：

- 数据治理工作的开展需要多方的参与，特别是业务管理部门的参与，不但包括组织上的人员参与，还包括认责、业务指导方面的工作参与。业务管理部门作为业务的权威解释方，应对数据标准和数据质量规范进行定义，作为数据管理的依据。在复杂的数据质量问题处理中，业务部门还应承担起问题根因分析的职责，参与到解决方案的制定过程中；

- 数据治理规则的梳理应与数据的应用场景和问题发现情况结合起来。孤立地从系统视角梳理数据质量规则往往产出较多，但存在数据质量规则与业务需求、监管要求存在差异等问题。数据质量的需求采集工作应由业务驱动或问题驱动，以保证每个正式执行的数据质量检核规则都具备充分的业务价值。

分解落实，足履实地

针对此次数据治理监管要求，我们建议银行和保险机构可以将动作拆解开来，分步实施，并且做好闭环管理。

动作分解

此次为期一年的监管数据质量专项治理工作，从银保监会的发文来看，明确且有针对性，对于银行保险业金融机构来说，相关工作可以拆解为五个阶段，提交一个方案、两个报告。对于此项工作，IBM 建议在实操中应注意以下三方面：

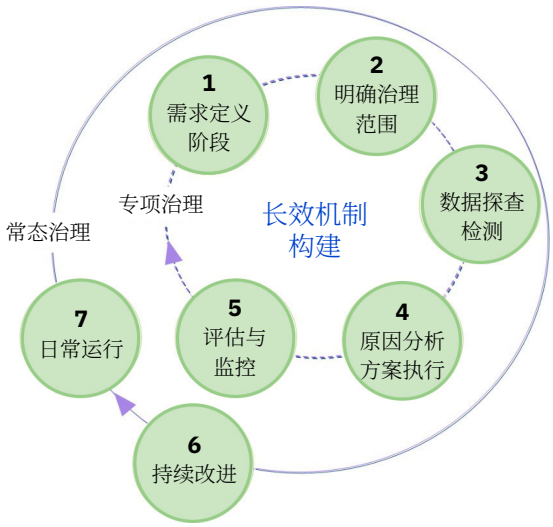
宣传动员：一个为期一年数据质量提升治理项目，覆盖从监管数据到源头数据，机构范围不仅有总行层面，分支机构还需要满足属地银监局的要求，因此一个从全行层面展开的宣传动员非常有必要，无论是一线的数据录入部门还是二线的管理部门，“数据质量，人人有责”的数据文化应该被宣扬；

责任分工：监管数据本质上覆盖银行的所有数据，所以对于监管数据的专项治理，并不是某个归口部门或者几个报表部门就能独立完成的，一个良好的责任分工，才能保障这项工作的有序展开。尤其此次强调源头数据，数据的录入部门、定义部门、系统的开发部门、报表的管理部门等都应该参与且承担相应的责任；

机制优化：整个监管数据质量管理自查自评和总结的过程，其实也是对整个机构数据质量管理的一次大摸底。制度上的不足、流程中的缺陷、职责的模糊不清等一系列管理上的问题都会通过数据问题暴露出来，相关机制的改善也会更有针对性和落地性。在更规范和更水土相符的制度体系下，整个数据治理的工作才会更有效果。

图 3

专项治理和常态治理的闭环管理



闭环管理

此次银保监会的专项治理，其实就是一次从监管层面引导银行业和保险机构建立数据质量闭环管理的过程，这与 IBM 对于数据质量闭环管理的思路不谋而合（见图 3）。

IBM 始终强调，数据治理是“旅程”而不是“项目”，不可期望一蹴而就，因此需要构建“长效机制”，具体工作体现在两个层面：

第一，是针对关键数据项（首批用例或者治理服务化需求）的专项治理过程；

第二，是专项治理过程转化为日常工作，并且进行持续改进的常态治理过程。

此次监管就是一个针对“监管数据”的专项治理。而随着数据的不断产生，“监管数据”的专项治理也需常态化。而对于其他的业务领域，按照机构的内部管理要求，也可以展开专项治理，确保从方案的制定到最终问题的整改到常态化的运转形成闭环。

精耕细作，步步为营

监管数据质量治理是银行保险机构数据治理的重要组成部分，是一项基础性、系统性的工作，因此只有精耕细作，才能步步为营。我们将整个数据质量提升的方法归纳为“4321”，涵盖机制、方案、规则和整改，具体如下：

四大保障机制

四大保障机制助力数据治理工作的有效开展：

制度流程：《银行业金融机构数据治理指引》第十九条明确，银行业金融机构应当制定与监管数据相关的监管统计制度和业务制度，及时发布并定期评价和更新；从制度层面明确各方职责、报送流程环节及系统支持等，固化报送工作机制，确保有章法可循；

数据认责：建立数据认责机制，落实数据的责任部门，才能保证数据报送时能找到管理部门，问题发生时能找到主管部门，问题追责时能定位责任部门，降低沟通成本，提高管理效率；

图 4

事前、事中、事后的数据质量提升方案