



# 基于微服务的监管数据集市和 监管报送系统建设实践

## 湖北农信监管数字化转型之路

付毅 | 湖北省农村信用社联合社信息科技部重大项目专班 经理



1

建设目标与思路

2

总体架构设计

3

技术创新及优势

4

实施过程与推广

5

运营成效与指标

6

数据及报送质量提升

7

项目实施经验总结

8




监管合规分析与自查

1




## 建设目标与思路

# 传统模式下的痛点与新形势下的刚需

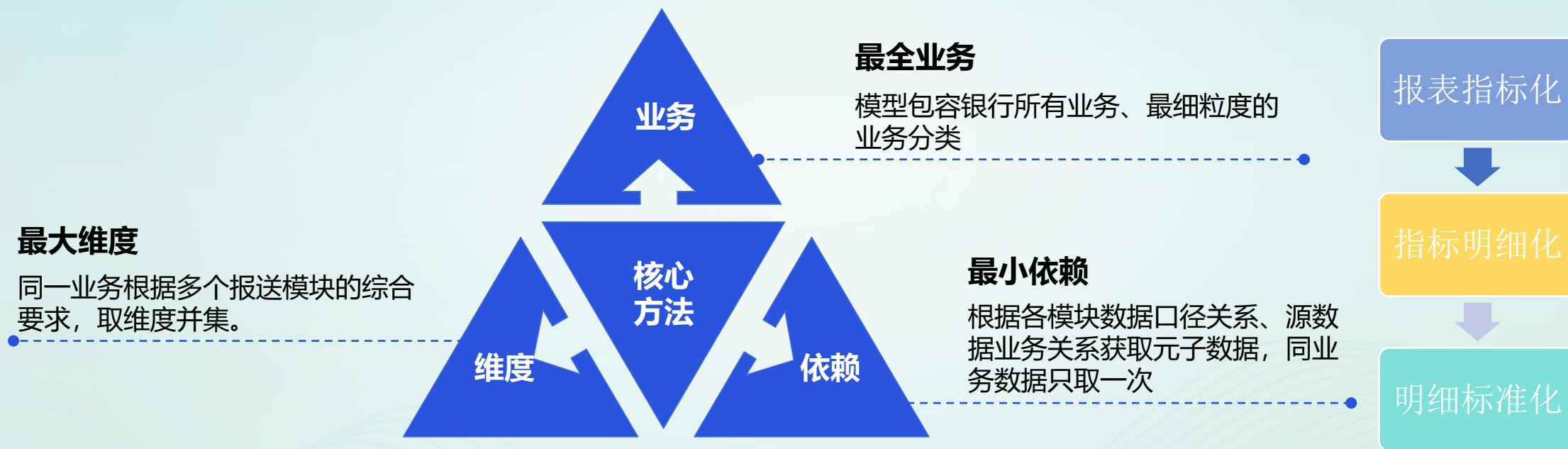
## 传统痛点

-  “烟囱式”系统：数据多头抽取，口径不一，质量追溯困难。
-  高度依赖人工：手工补录与核对工作量大，错报、迟报风险高。
-  缺乏交叉核验：难以实现跨报表、跨系统的数据校验。

## 新形势下的刚需

-  监管要求趋严：制度更迭频繁，对自动化、明细化要求日益增强。
-  跨域整合需求：需要整合跨监管、跨部门、跨业务的复杂信息。
-  统一化目标：统一各项监管报送数据源头、统一监管报送数据标准、统一数据加工处理流程

核心目标：建立统一数据底座，实现质量问题标本兼治与报送质效双提升。



### 实施思路

架构原则：微服务架构，敏捷迭代

推进策略：分层建模、先急后缓

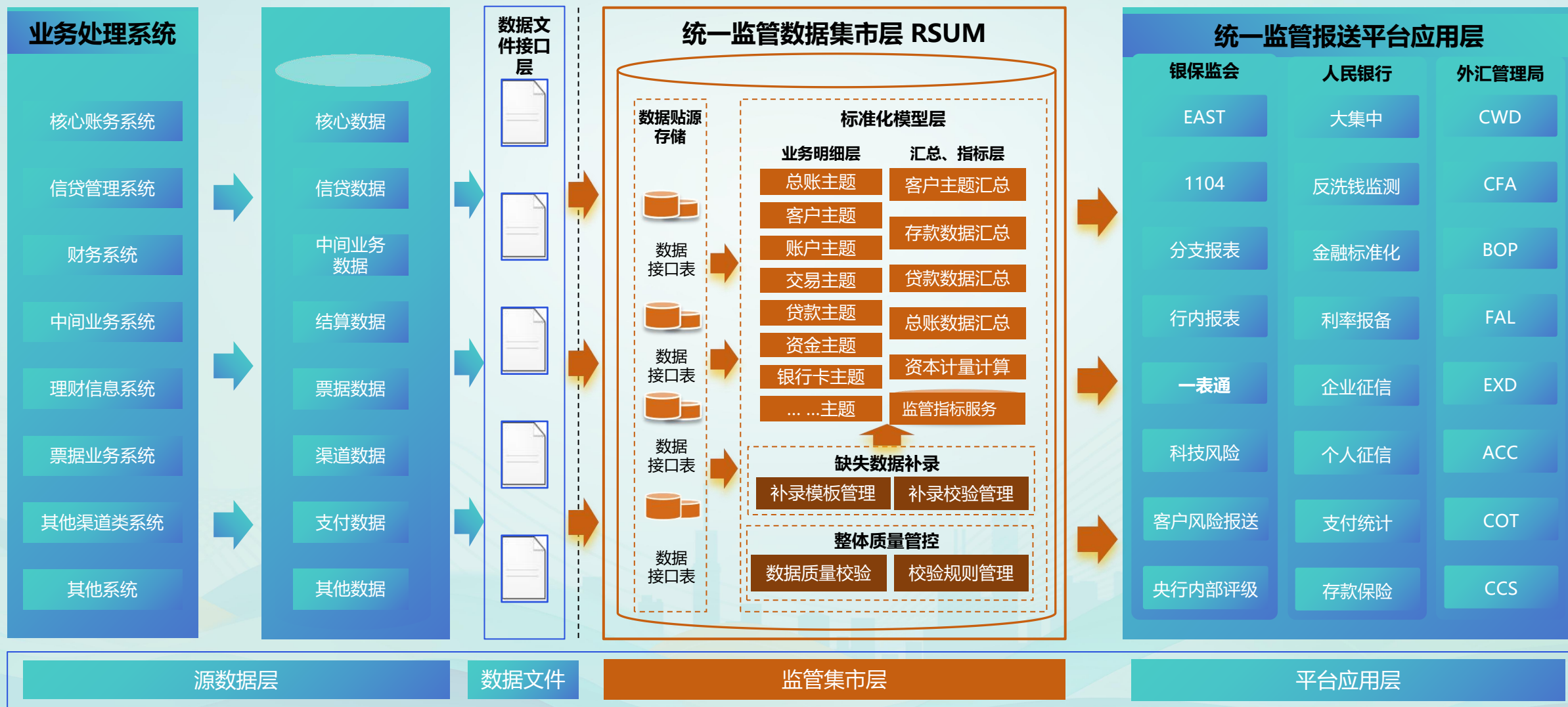
数据原则：数据同源、口径统一

最终目标：可持续演进的监管数据集市和报送平台



2

## 总体架构设计





## 监管报送平台

实现监管报送数据的数据初始化, 汇总计算、数据校验、数据审核、数据打包等功能。提供基于报送数据的数据加工逻辑溯源功能; 提供报送相关的数据查询及报送考核功能。

## 数据补录平台

实现基于补录模板的数据补录、任务审核、补录数据查看、补录进度查看及补录质量看板等功能。

## 数据检核平台

实现基于全平台各业务场景的校验集中预警管理, 可基于补录数据、模型数据、报送数据的校验结果进行检核分析, 也可针对全平台的校验情况进行多维的预警分析。

## 数据管理平台

提供基于全平台的统一数据管理, 包括统一指标管理、统一表单管理、统一规则管理、统一计划任务管理等, 形成了基于监管报送平台的数据资产管理。

## 数据开发平台

提供模型层加工逻辑的配置功能, 可依据访谈的结果, 配置各数据加工层的处理逻辑, 可为数据溯源提供基础。

## 基础配置平台

提供统一监管报送平台内的机构、用户以及权限管理, 并提供各报送模块相关的报送配置管理。

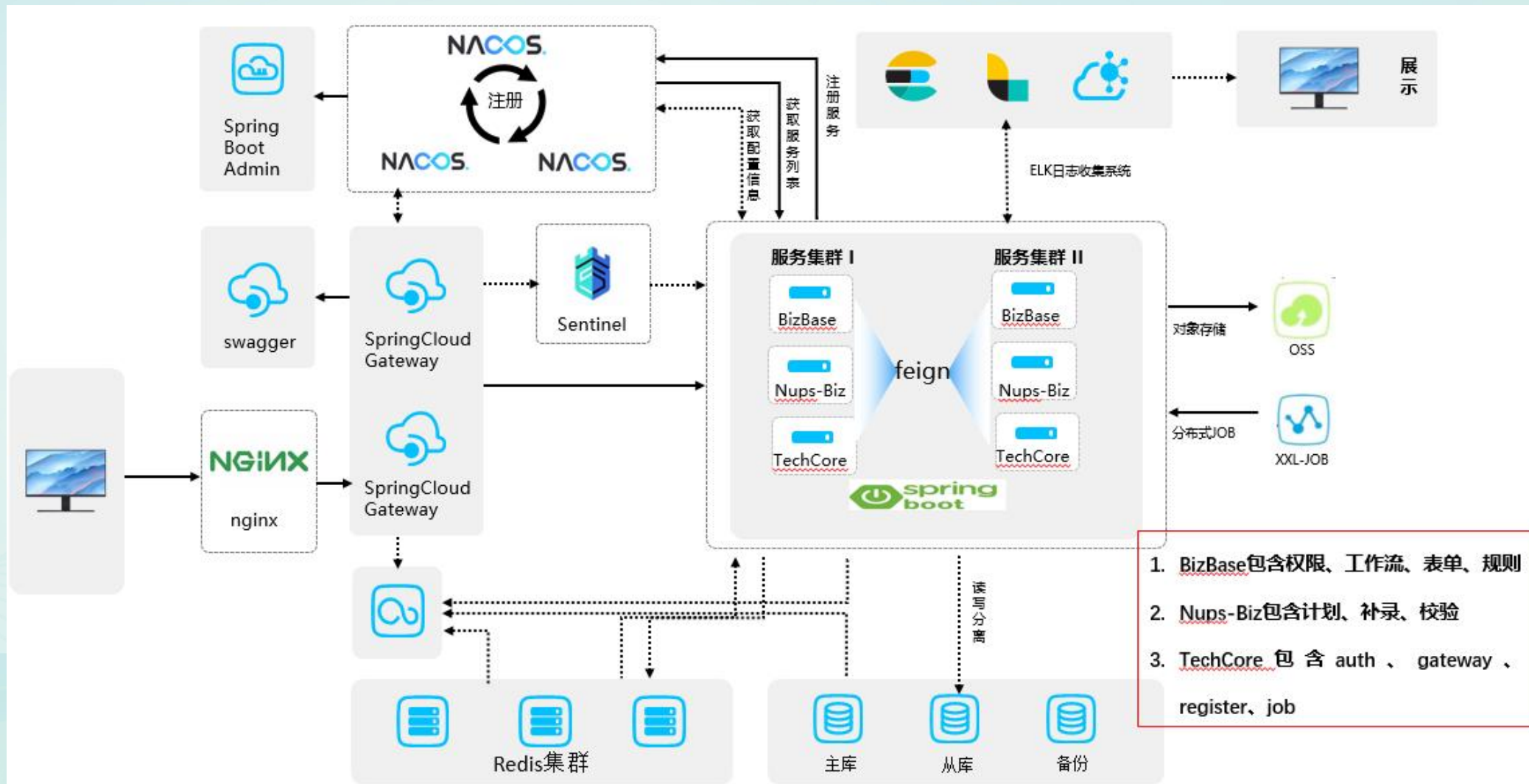
# 2-总体设计-产品功能架构

系统总体架构采用分层设计，包括基础设施层、数据层、平台层、应用层以及运维管理五个方面，总体架构如图所示：



## 2-总体设计-产品功能架构

系统平台基于流行的J2EE技术栈，应用稳定的Spring Cloud微服务框架进行构建；采用微服务架构设计，系统平台基础模块的每一个应用功能都使用微服务完成。每个服务都有一个清晰的功能边界，每一个后台服务开放REST API，服务本身也采用了其它服务提供的API。通过API Gateway来统一API暴露，通过Nacos配置中心来统一管理平台服务配置。每个微服务可以选择独立部署和弹性资源配置。





3

## 技术创新及优势

系统平台采用的分布式数据库最核心的创新在于其多模型统一架构；在处理需要联合多种数据的复杂业务时，开发人员无需在不同系统间进行繁琐的数据导出导入，也无需分别编写Cypher、SQL和RESTful API等多段代码，只需一句标准的SQL即可完成跨模型关联查询，从而大幅简化开发流程，并从根本上保障数据的一致性和实时性。

在监管数据集市的数据底座上，创新性地采用了行列混合存储结构。对于需要频繁更新少量数据的场景，利用行式存储的速度优势；对于需要大规模扫描和聚合分析的场景，则发挥列式存储的查询性能优势，从而构建高效稳定的监管数据集市。



新一代的监管数据集市和监管报送系统进行了全栈信创改造，从底层的信创芯片服务器，到中间件、数据库，都采用了国产产品，这大大降低了对外部技术的依赖。

# 3-技术创新及优势-统一监管数据集市

过去“烟囱式”、“孤岛式”的报送系统建设导致数据口径不一。湖北农信结合监管数据集市和监管报送系统的项目提出了建设统一的监管数据集市的思路。将分散在各业务系统的数据在集市中进行整合、清洗和标准化加工，纳入统一的监管数据集中，从源头上保障了报送数据的一致性、准确性和可复用性，为实现“同标同源”报送打下坚实基础。

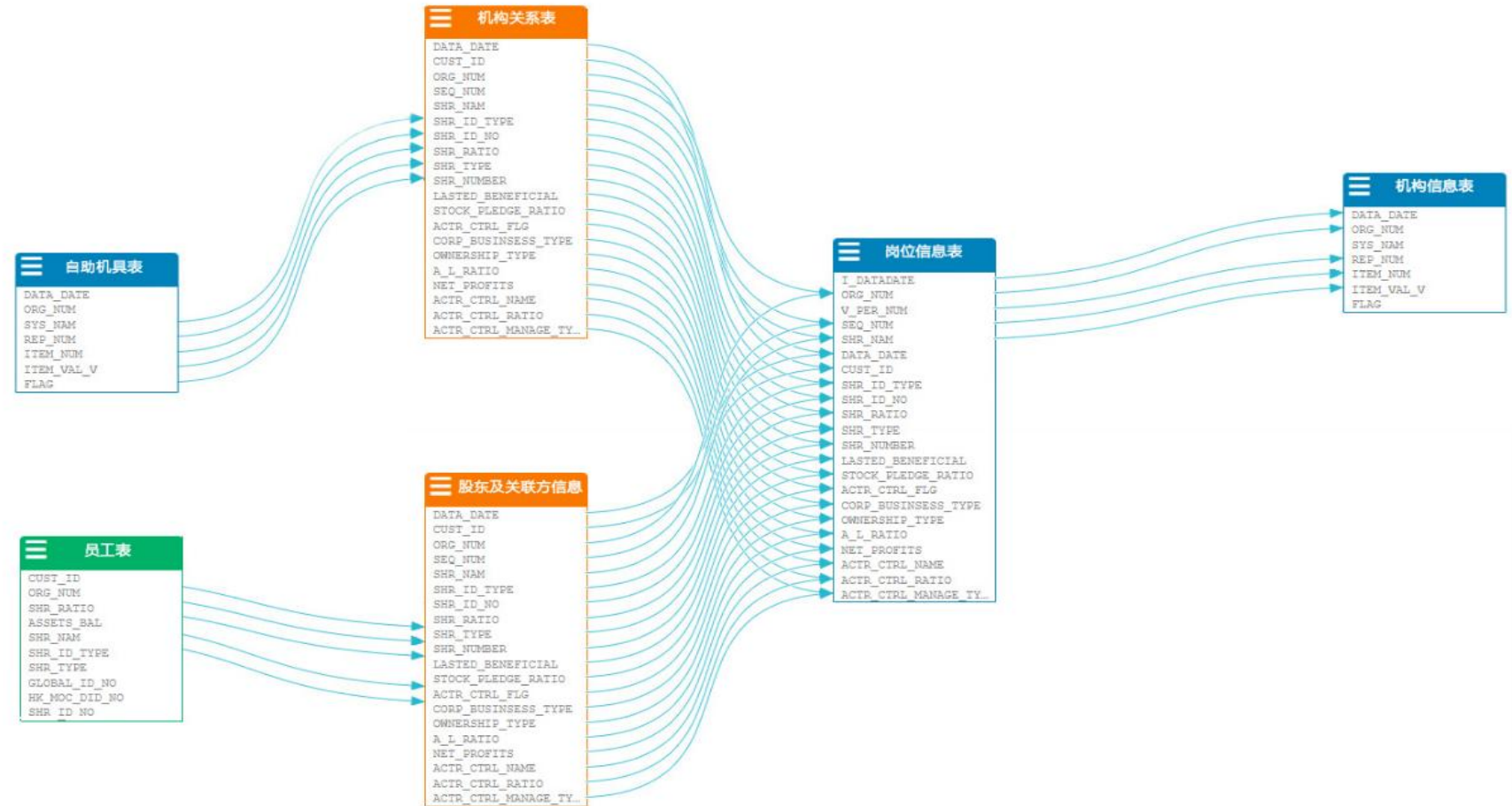


10个一级业务主题；67个二级业务主题；264个业务主题表  
业务主题表数量、结构稳定，是报表数据加工的基础，专人专职持续维护

当指标或报送数据出现异动时，快速定位问题至关重要。在项目建设中，监管数据集市建立了覆盖全报送范围的数据血缘追溯管理体系。业务及填报人员可以在系统页面上，对报表指标进行正向的血缘分析（追溯其来源）和反向的影响分析（评估其变动的影响），使数据的来龙去脉一目了然，极大提升了问题排查效率。

### 智能数据血缘追溯

血缘关系图



### 低代码与模块化配置

通过低代码或模块化的方式提升灵活性。监管报送系统的报表设计器支持通过拖拉拽、导入报表模板或制度模板等实现快速、高效开发新报表或临时报表。降低了对专业开发能力的依赖，使业务及报送人员也能参与配置，加快了对新规需求的响应速度。

### 多法人架构支持

农信体系通常包含省联社和多家市、县级法人机构。监管报送系统在设计上就支持多级法人管理，不同法人机构可以在统一的平台下进行独立的参数配置、数据查询、报送操作及权限控制等，既满足了统一管理的要求，又尊重了各法人的独立性。



应用模式  
体验创新

# 3-技术创新及优势-业务赋能与价值创新

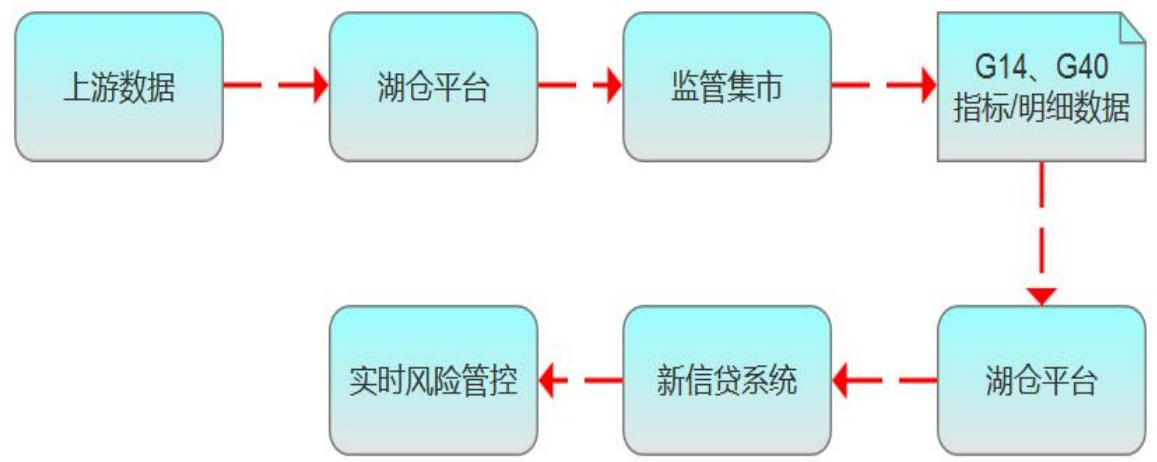
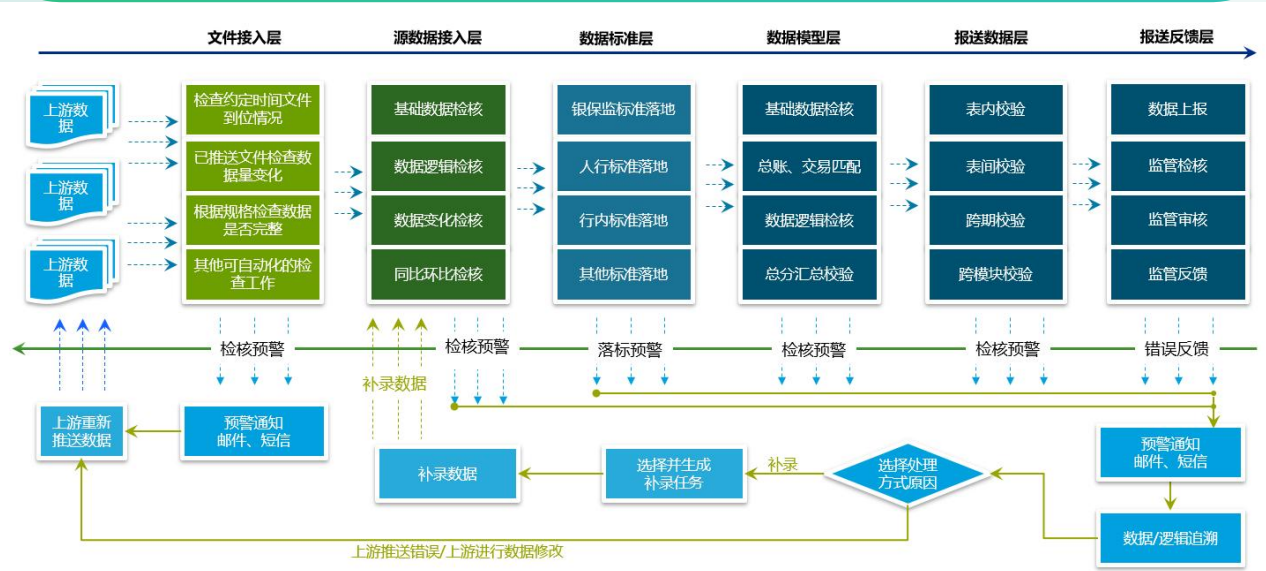
# 2025农村金融科技创新与共享发展会议

## 数据治理与闭环管理

监管报送系统基于数据质量检核模块将数据质量监控体系拆分为三层，分别为**贴源层监控、模型层监控、报表层监控**，逐层解决数据质量问题，建立科学规范的标准体系，有效降低信息系统的复杂性和管理难度，提升金融机构信息科技应用的前瞻性和主动性，也才能更好地规范系统开发、运行和维护升级，提高应变能力和运营效率。

## 从“合规报送”到“价值反哺”

基于内部数据挖掘可将监管报送数据应用于客户画像营销、利率定价、贷前审批等多个业务领域，可通过复用高质量的监管数据来构建企业征信模型，反哺小微贷款业务，让合规数据产生更大的商业价值；截止目前，已将**大额风险暴漏**等相关数据应用到新信贷系统中，支撑信贷系统在授信审批、放款环节的实时风险管控。





## 实施过程与推广

监管数据集市和报送系统项目的实施周期通常呈现“分阶段、递进式”的特点，项目实施建设总体采用“统筹规划、分步实施、敏捷迭代”的原则：

### 第一阶段： 规划与基础构建

- 周期：24年02月-24年03月
- 1) 技术平台选型与部署：主要涉及硬件服务器、高性能分布式数据库、分布式应用系统、搭建及部署开发、测试、验证环境等。
- 2) 数据标准制定：根据监管要求，对全行的客户、账户、合同等核心数据元素制定统一的标准和业务口径，从根本上解决“数据多源、口径不一”的问题。

### 第二阶段： 系统设计及开发阶段

- 周期：24年03月-24年10月
- 监管集市模型设计：按照“明细层(LDM) -> 汇总层(ADM) -> 报文层(ADM)”的数据架构，构建企业级的监管数据模型，确保数据可追溯。
- 报送系统口径调研与访谈：按照调研访谈计划，采用不同形式的调研方式，将全行级业务及报表等进行访谈与调研，编制并确认ITMapping访谈底稿。
- 监管报送数据迁移：监管报送业务迁移到新平台上，核心模块优先，通常优先实施监管要求最迫切、最通用的系统，如1104、大集中、EAST5.0、金融基础数据、客户风险等。

### 第三阶段： 全面推广与深度应用

- 周期：24年11月-25年07月
- 平稳迁移：采用“平行双跑”策略，即新旧系统并行运行一段时间，进行数据比对和业务验证，确保无误后再逐步切换，这对于金融系统至关重要。
- 全面推广：在系统稳定运行后，项目进入扩展和深化阶段，目标是最大化数据价值和应用广度。
- 深度应用：覆盖更多报送系统，将支付统计、利率报备、二代征信等更多报送模块集成到统一平台中。





05

## 运营成效与指标

### 高并发处理能力

基于分布式架构，系统能够轻松应对日报、月末、季末等报送高峰期的压力，批量数据处理任务完成时间从过去的**20多小时缩短到6小时左右**，报表生成与提交效率显著提升。

### 报送系统自动化准确率

**1104报表**：由项目建设前的**不到20%**，提升到现在的：**96%**；**EAST数据质量**的校验错误率从**30%降低至现在的0.1%**。

### 降低合规风险

通过搭建的监管数据检核体系，实现数据填报前、中、后三层的“体检”，做到提前发现问题，跨模块交叉核验。

### “数据同源”成效显著

统一的监管数据集市从根本上解决了过去数据口径不一、多头取数的难题。通过内置的上千条数据质量检核规则，在数据入库、加工、报送等环节自动进行校验，**有效将数据差错率降低了99%以上**。

### 自动化水平大幅提升

报表自动生成、一键报送、低代码平台等功能将业务人员从繁琐的手工操作中解放出来，报表自动化生成与报送率普遍达到**99%以上**。

### 高可用性保障

采用集群化部署和负载均衡技术，系统实现了**99.9%以上**的高可用性，有效避免了因单点故障导致的服务中断，满足了金融机构业务连续性的严格要求。

### 智能化运维

集成了监报告警、短信通知、日志分析等功能的统一运维平台，使运维人员能够实时掌握系统健康度，实现问题的快速发现与定位，降低了运维复杂度。

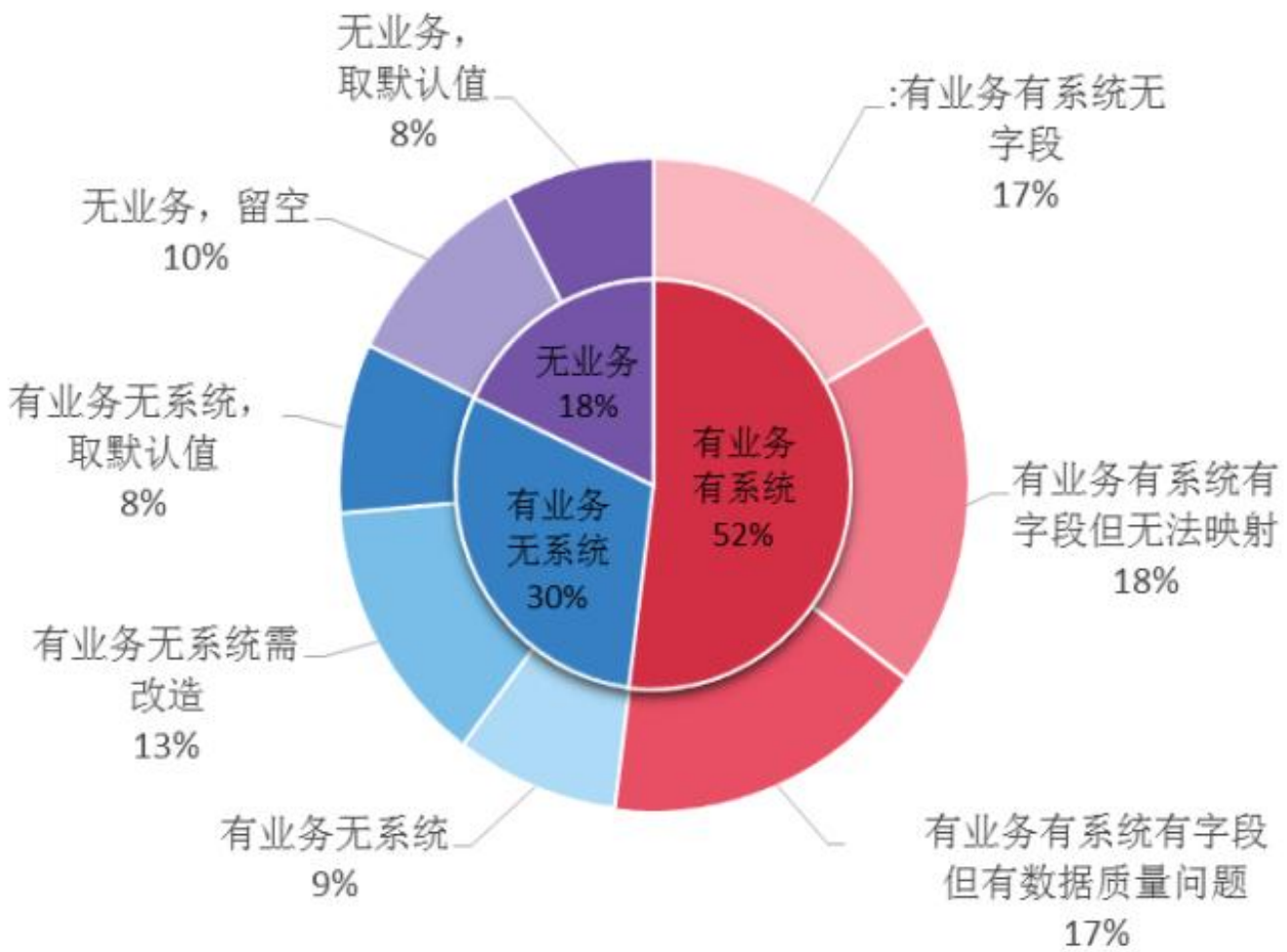


## 数据及报送质量提升

# 6-数据及报送质量提升-问题分类及定义

代码	一级分类	代码	二级分类	解释	举例
A	有业务 有系统	A1	满足	无问题	无
		A2	有业务有系统无字段	模型中字段在行内已开展业务，系统中存在字段缺失，模型无法取到数。	例如模型表《贷款借据信息表》中的“贷款资金使用位置”，用于区分贷款投向是境内或境外，在信贷系统中找不到对应字段，也无法通过其他字段判断加工得出， <b>则标记为A2类问题。</b>
		A3	有业务有系统有字段但无法映射	模型中字段在行内已开展业务，系统虽有该字段但码值无法转换成监管定义的口径，或达不到监管定义的粒度，模型无法进行码值映射	例如模型表《贷款借据信息表》中的“还款方式2”，行内分类为等额本金和等额本息，模型码值为按月、按季、按年、利随本清...无法转换成模型码值， <b>因此标记为A3类问题。</b>
B	有业务 无系统	B1	有业务无系统（系统待建，不满足访谈条件）	模型中字段在行内已开展业务，但无系统支撑，不存在访谈条件，模型暂不取数。	例如：模型表《理财账户信息表》目前行内的理财系统正在待建中，不存在访谈条件，模型暂不取数， <b>因此标记为B1类问题。</b>
		B2	有业务无系统需改造	模型中字段在行内已开展业务，但现行系统正面临改造重建，不存在访谈条件，模型暂不取数。	例如模型表《贷款借据信息表》中的“贷款资金使用位置”，用于区分境内或境外，在信贷系统中找不到对应字段，也无法通过其他字段判断加工得出， <b>则标记为B2类问题。</b>
C	无业务	C1	无业务，留空	模型中字段在行内暂时还未开展此类业务，系统中没有字段对应，模型取数采取留空方式的字段。此类问题可作为未来如果开展此类业务时的参考	例如模型表《贷款借据信息表》中的“投贷联动类型”字段，用于标识此笔贷款借据是否属于投贷联动，如果银行没有开展投贷联动业务，留空 <b>则标记为C1类问题。</b>

### 数据满足度分析



**统一数据口径：**整合分散在各业务系统的监管指标，消除数据孤岛；

**提升数据质量：**解决源系统数据缺失、逻辑冲突等问题；

**推动源头治理：**反向驱动上游系统改造，形成数据闭环；

### 字段二级问题分布





7

## 实施经验与总结

### 数据质量

#### 规划先行、平滑演进

坚持“统筹规划、标准先行”：在项目启动前制定清晰的顶层设计和演进路线图。

#### 同源同标、质量至上

构建自动化的数据比对和校验工具，确保信创集市中的数据与源系统、传统平台的数据在口径和内容上完全一致。

### 变化应对

#### 制定数据规范

在建设数据集市之初，就同步开展数据标准治理，明确监管指标的口径、来源和加工规则，这是保证数据质量、实现“数据同源”的基石。

#### 赋能业务、构建生态

建立由总行牵头，业务部门（合规、风险、数管、计财等）、科技部门（数据、架构、运维等）和厂商共同参与的常态化协同机制。

### 工作效率

#### 平行双跑、逐步切换

采用“平行双跑、逐步切换”的策略。选择非核心、复杂度适中的报送业务（如金数、EAST5.0等）作为首个信创化试点。有助于验证全栈信创环境的稳定性，并积累初步经验。

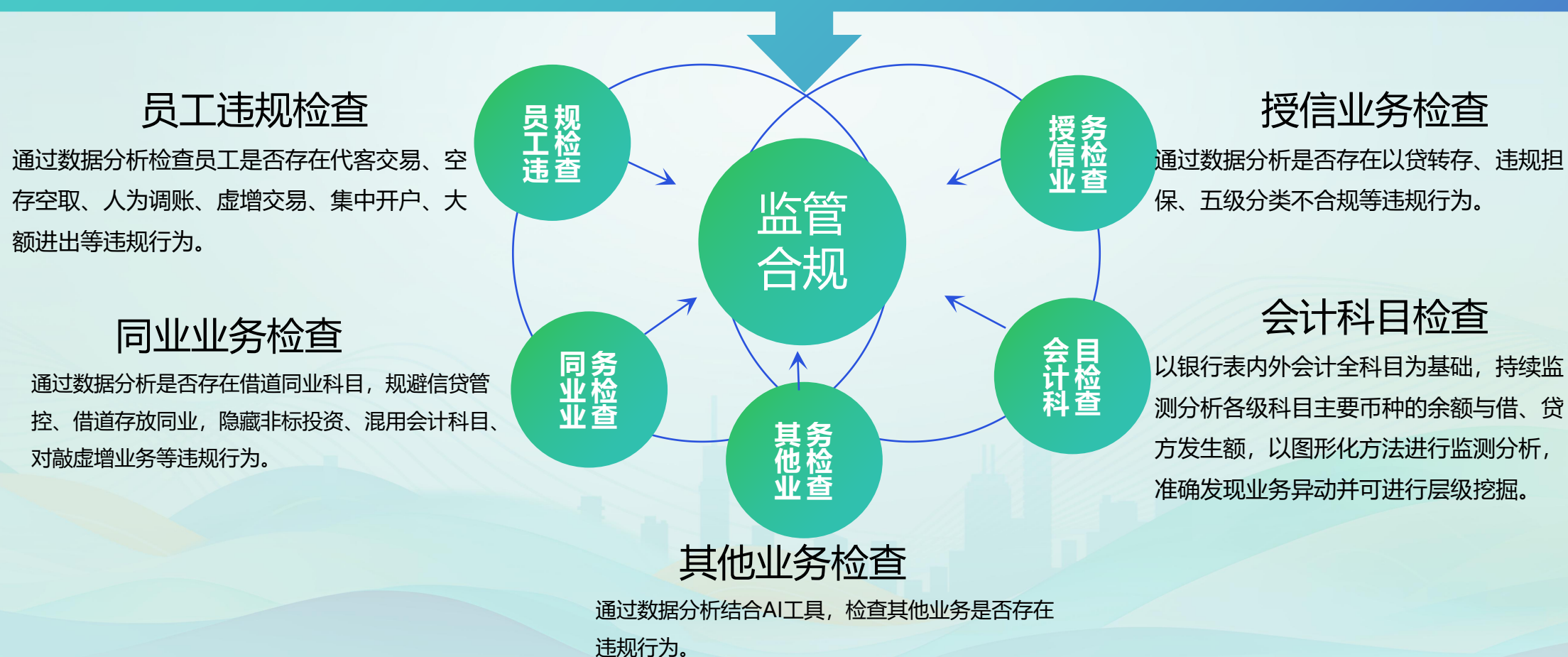
#### 携手共赢、可持续发展

主动与信创生态伙伴（芯片、操作系统、数据库厂商）建立深度合作与联合攻关机制。共同解决兼容性、性能优化等难题。



## 监管合规分析与自查

以EAST数据为基础的合规自检及违规业务识别，未来扩展到整套监管集市



截止目前，已完成监管合规模型设计：**103个**

16个

机构治理

股东和股权管理：13个  
关联交易：3个

9个

票据业务

票据业务：9个

35个

信贷业务

房地产金融：4个  
个人贷款：6个  
授信管理：19个  
一般性对公贷款：6个

19个

信用卡

风险监控和贷后管理：19个  
信用卡其他：1个  
信用卡授信管理：3个

24个

员工行为及案防

经商办企业或兼职：1个  
人员管理：5个  
员工违规贷款：2个  
员工异常行为：3个  
账户异常：13个

谢谢观看