关于外部数据集中管理的解决方案

随着我行业务的发展，越来越多外部数据被引入并应用在业务场景当中，通过行内数据与外部数据融合联动，实现强化风险防控管理、拓宽业务渠道、提升客户体验等目标。目前，我行外部数据来源渠道多、数据种类多，外部数据大多散落在各信息系统中，保存较为分散、独立，没有集中管理实现数据共享，导致行内行外数据融合度不足、外部数据利用率低、外部数据保存管理难度大、无法满足快速接入外部数据及时应用的要求。现对我行外部数据集中管理提出解决方案，具体内容如下：

# 一、外部数据现状分析

目前，我行使用外部数据的信息系统约有47个，对接的外联单位约有61个，涉及结构化、半结构化和非结构化等数据类型（详见附件1），相关外部数据大部分在各自信息系统中保存管理。具体情况如下：

## （一）外部数据对接方式。

我行信息系统与外联单位系统对接时，我行是作为参与者的角色，需要遵循对方的接入技术标准和规范，但我行可以选择自身系统与其对接。目前，我行信息系统与外联单位系统对接方式主要有3种：**一是**相关系统开通对外网络权限直接对接外联单位系统获取外部数据，**二是**相关系统通过信贷风险数据集市系统对接外联单位系统获取外部数据，**三是**相关系统通过新中间业务GXP系统对接外联单位系统获取外部数据。外部数据接入方式较为分散，较多的信息系统直接对接外联单位系统，在网络安全方面存在风险隐患。

## （二）外部数据种类及应用情况。

按存量外部数据的用途划分，数据种类主要有以下3种：

**1、查控数据。**根据上级机构防电信诈骗快速查控的要求，我行需要开发相应系统（模块）与其系统对接，我行系统接收对方主动推送的客户、账户信息文件并返回处理结果。上述系统主要是通过GXP系统对接外联单位系统，外联单位包括：法院、人行、公安、国家安全机关等。查控数据在处理结果成功返回后就不会再次被使用，所以该数据不需要作长期保存。

**2、业务数据和对账文件。**在开展支付结算业务、第三方支付、中间业务和基金理财等业务时，我行系统需要对接外部单位系统进行业务数据交互，日终时接收对账文件进行对账处理。在GXP系统建成前，上述业务的系统是直接对接外联单位系统；在GXP系统建成后，上述业务的新建系统主要是通过GXP系统对接外联单位系统；外联单位包括：人行、银联、票交所、供电局、基金公司、保险公司、第三方支付公司等。业务数据一般是以报文的形式交互并在相应业务流水表中保存，对账文件在完成对账处理后就不会再次被使用，所以对账文件不需要作长期保存。

**3、风控数据。**为加强风控管理和提升风控能力，我行持续引入和采购外部风控数据，并应用在FTP系统、交易反欺诈系统、反洗钱黑名单筛选系统、信贷业务系统、信用卡系统、小微金融服务平台等系统中。上述系统通过信贷风险数据集市系统或GXP系统对接外联单位系统获取相关数据，外联单位包括: 人行征信、万得、同盾、道琼斯、工商和法汇网等。外部风控数据可以作为我行客户信息、行为和状态的补充，与行内数据融合运用可以达到强化风险体系的效果。因此，需要对其进行统一保存和管理，深挖数据价值，促进数据的高效应用。

# 二、外部数据集中管理方案

## （一）系统架构规划。

外部数据集中管理旨在实现实时、准实时和批量数据的统一平台架构、统一管理机制和统一服务模式，深挖数据潜在价值，促进风险防控和业务创新发展。

统一平台架构是以GXP系统作为行内对外部数据的统一出入口，实现快速对接外部系统和强化网络安全管理；以大数据应用平台中数据交换平台、数据服务平台和外部数据管理平台作支撑，实现外部数据的统一保存和管理。

统一管理机制是指对外部数据进行规范数据质量管理、规范数据生命周期管理、规范数据清洗、数据挖掘和数据分析等模型设计，实现行内行外数据高度融合和高效应用。

统一服务模式是指提供行内业务系统对外部数据的统一接入标准和统一访问服务，访问方式包括单笔查询接口、批量查询接口、定期订阅接口。

外部数据集中管理系统架构图如下：



## （二）实施计划。

1、建设大数据应用平台。

将外部数据集中管理纳入大数据应用平台项目的落地应用场景，提前考虑数据交换平台、数据服务平台和外部数据管理平台的规划设计，支持对外部数据的保存、管理以及应用。实施计划如下：

1）2018年4-8月，进行公司交流和草拟交流报告；

2）2018年8-9月，进行项目采购；

3）2018年9月-2019年6月，进行项目阶段上线，包括虑数据交换平台、数据服务平台和外部数据管理平台等功能。

4）2019年6月以后按项目计划推进。

2、梳理整合存量外部数据。

梳理存量系统的外部数据对接方式，逐步将直接对接外联单位的系统改为通过GXP系统对接方式，对于有价值的外部数据推送给外部数据管理平台。实施计划如下：

1）2019年1-6月，梳理存量系统的外部数据对接方式，按大数据应用平台的规范标准改造系统。

2）2019年6-12月，分批上线，接入大数据应用平台。

# 三、临时过渡方案

由于大数据应用平台正在公司交流阶段，离系统上线投产仍有一段较长时间，为保障新开展业务不受影响，考虑到目前新增接入的外部数据都是通过GXP系统与外联单位系统对接，数据则在中间业务平台（GAPS）上临时保存并提供查询服务，建议以当前中间业务平台（GAPS）的版本为基准，在生产环境新搭建一套临时的GAPS系统专门用于外部数据的存放和管理，作为临时过渡方案。待大数据应用平台建成后再将该临时系统数据和服务迁移到大数据应用平台，同时停运该临时系统。既可以支持新业务产品开发也可以确保中间业务平台（GAPS）不受持续新增接入外部数据的影响。

## （一）系统架构规划。

充分利用现有中间业务平台系统架构，搭建独立的外部数据集中管理系统，系统架构及具体实现如下：



系统架构图

1、外部通讯

复用现有互联网GXP和专线GXP，由GXP实现通讯格式适配，与外部单位进行通讯。

2、内部通讯

复用现有内联GXP，通过内联GXP与ESB进行通讯。

3、业务逻辑处理

以现有GAPS平台为基础，建立独立于中间业务平台的外部数据集中管理系统，实现数据落地、报文格式转换、数据查询、生命周期管理等功能。

4、数据库管理

独立数据库管理。

5、数仓供数方式

沿用现有方式，通过数据交互系统供数。

6、旧业务迁移。

旧业务迁移范围包括:公积金缴存信息、芝麻信用、电信反欺诈、电信诈骗、国安查控、经侦网络查控、国安网络查控、电信诈骗、网络查控、高法查控等。

7、旧业务迁移方式

由ESB对服务配置进行更改，将服务映射到新系统，服务的迁移对消费者系统透明。

## （二）实施计划

1、新系统建设

以腾讯天御反欺诈对接为契机，8月份完成外部数据集中管理系统建设。

2、旧业务迁移计划

旧业务迁移需根据开发、业务人员实际工作负荷情况逐步进行，不强制要求。

2018年6月25日

附件1：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **系统名称** | **外联单位** | **对接方式** | | |
| **通讯协议** | **实时性** | **文件交互** |
| E兑换 | 外管局 | MQ | 实时 |  |
| 统一支付平台 | CCPC | MQ | 实时 | 文件 |
| 省联社 | MQ | 实时 | 文件 |
| 网银互联系统 | CCPC | MQ | 实时 | 文件 |
| 票据系统 | 票交所 | HTTP/MQ | 实时 |  |
| ACS综合前置系统 | CCPC | MQ | 实时 |  |
| 身份核查系统 | 广州人行 | MQ | 实时 |  |
| 财务系统 | 京东 | https | 实时 |  |
| FTP系统 | 万得 |  | 非实时 |  |
| 交易反欺诈系统 | 同盾科技 | http | 实时 |  |
| 反洗钱黑名单筛选系统 | 道琼斯 | GXP | 非实时 | 文件 |
| 对外风险数据集市 | 工商 | http | 非实时 | 文件 |
| 网络爬虫PIRA | 互联网 | http | 非实时 | 文件 |
| 广东公安查控 | 广东公安 | TCP | 非实时 | 文件 |
| 全国法院查控 | 全国法院 | TCP/IP、FTP | 非实时 | 文件 |
| 防电信诈骗 | 人民银行 | TCP/IP、FTP | 非实时 | 文件 |
| 国家安全机关查控 | 国家安全机关 | TCP/IP、FTP | 非实时 | 文件 |
| 全国公安查控 | 全国公安 | TCP/IP、FTP | 非实时 | 文件 |
| 社区O2O | 腾讯-微信 | JSON接口 | 实时 |  |
| 手机银行 | 讯联收单 | http/json | 实时 |  |
| 网联 | 网联清算平台 | https/xml | 实时 |  |
| 网联 | 网联清算平台 | sftp | 非实时 | 文件 |
| 第三方 | 微信、京东、百付宝、苏宁、银联 | https/xml | 实时 |  |
| 第三方 | 第三方 | sftp | 非实时 | 文件 |
| 支付宝卡通 | 支付宝 | https/xml | 实时 |  |
| 支付宝卡通 | 支付宝 | sftp | 非实时 | 文件 |
| 现管 | 各企事业单位 | https/xml | 实时 | 文件 |
| 新中间业务系统 | 160莞医通、诊中支付系统 | TCP | 实时 |  |
| 新中间业务系统 | 院线通 | TCP | 实时 |  |
| 新中间业务系统 | 翰银支付 | TCP | 实时 |  |
| 新中间业务系统 | 财政局 | TCP | 实时 |  |
| 新中间业务系统 | 迅联国际 | HTTPS | 实时 |  |
| 新中间业务系统 | 广东公安 | TCP | 实时 |  |
| 新中间业务系统 | 工商局 | HTTPS | 实时 |  |
| 新中间业务系统 | 公安 | TCP | 实时 |  |
| 新中间业务系统 | 公积金 | TCP | 实时 |  |
| 新中间业务系统 | 东莞供电局 | TCP | 实时 |  |
| 新中间业务系统 | 地税国税 | MQ | 实时 |  |
| 新中间业务系统 | 道琼斯 | TCP | 实时 |  |
| 新中间业务系统 | 同盾 | HTTP | 实时 |  |
| 新中间业务系统 | 学校 | TCP | 实时 |  |
| 新中间业务系统 | 平安银行 | STD | 实时 |  |
| 新中间业务系统 | 房管局 | TCP | 实时 |  |
| 新中间业务系统 | 联通 | TCP | 实时 |  |
| 新中间业务系统 | 银联 | HTTP | 实时 |  |
| 新中间业务系统 | 网联 | HTTPS | 实时 |  |
| 新中间业务系统 | 京东 | HTTPS | 实时 |  |
| 新中间业务系统 | 微众金 | HTTPS | 实时 |  |
| 旧中间业务系统 | 中国移动 | TCP/FTP | 实时 | 文件 |
| 旧中间业务系统 | 供电局 | TCP/FTP | 实时 | 文件 |
| 旧中间业务系统 | 兴业银行 | TCP/FTP | 实时 | 文件 |
| 旧中间业务系统 | 学院 | TCP/FTP | 实时 | 文件 |
| 旧中间业务系统 | 基金公司 | TCP/FTP | 实时 | 文件 |
| 旧中间业务系统 | 房管局 | TCP | 实时 |  |
| 旧中间业务系统 | 公共资源中心等 | TCP/FTP | 实时 |  |
| 基金代销系统 | 基金公司 | ftp | 非实时 | 文件 |
| 银保通系统 | 保险公司 | TCP/FTP | 实时 |  |
| 短信银行 | 中国移动 | CMPP | 实时 |  |
| 短信银行 | 中国电信 | SMGP | 实时 |  |
| 短信银行 | 中国联通 | SGIP | 实时 |  |
| 贵金属系统 | 金交所 | TCP/FTP | 实时 | 文件 |
| 储蓄国债系统 | 人行、财政部 | MQ | 实时 | 文件 |
| 银联前置 | 银联 | tcp协议socket | 实时 |  |
| 间联前置 | 讯联 | tcp协议socket | 实时 |  |
| TSMP | 银联 | tcp协议socket | 实时 |  |
| TSMP | 东莞通 | tcp协议socket | 实时 |  |
| 信用卡自动化审批系统 | 同盾科技 | https协议 | 实时 |  |
| COMSTAR债劵资金交易系统 | 中国外汇交易中心 | SOAP/HTTP协议 | 实时 |  |
| COMSTAR理财资金交易系统 | 中国外汇交易中心 | SOAP/HTTP协议 | 实时 |  |
| TIPS国库系统处理系统 | 人行TIPS中心 | TCP/IP协议SOCKET | 实时 |  |
| 法院查控系统 | 最高人民法院 | TCP/IP、FTP | 非实时 | 文件 |
| 国安查控系统 | 国安部 | TCP/IP、FTP | 非实时 | 文件 |
| 房维系统 | 东莞市房管局 | TCP/IP协议SOCKET、FTP | 实时 |  |
| 公安查控 | 广东省公安厅 | TCP/IP | 实时 |  |
| 惠州支行预售资金系统 | 惠州房管局 | SOAP/HTTP协议 | 实时 |  |
| 惠州村镇预售资金系统 | 惠州房管局 | SOAP/HTTP协议 | 实时 |  |
| 财政集中支付系统 | 财政局 | TCP/IP协议SOCKET、FTP | 实时 |  |
| 拍卖行保证金系统 | 东莞市拍卖行 | TCP/IP协议SOCKET、FTP | 实时 |  |
| 建设局保证金系统 | 东莞市公共资源交易中心 | TCP/IP协议SOCKET、FTP | 实时 |  |
| 外汇牌价系统 | 中国外汇交易中心 | TCP/IP | 实时 |  |
| 银校通云平台 | 讯联公司 | HTTP | 实时 |  |
| 银校通简易版 | 讯联公司 | HTTP | 实时 |  |