
“睿治”数据治理平台 技术白皮书



北京亿信华辰软件有限责任公司
2023 年 9 月

目 录

1 前言	6
1.1 关于本白皮书	6
1.2 数据治理定义	6
1.3 数据治理的重要性	6
1.3.1 企业所面临的问题	6
1.3.2 数据治理可以解决什么问题	8
1.4 数据治理的现状与发展趋势	9
1.4.1 数据治理现状	9
1.4.2 云端崛起	错误！未定义书签。
1.4.3 与人工智能握手	10
1.4.3.1 元数据	10
1.4.3.2 数据质量	10
1.4.3.3 其他应用	11
2 平台简介	13
2.1 睿治涵盖	错误！未定义书签。
2.2 平台价值	14
2.3 平台架构	15
2.4 核心优势	16
2.4.1 各产品模块高度融合	16
2.4.2 先进的产品设计理念	16
2.4.3 专业的服务保障团队	16
2.4.4 丰富的项目实践经验	16
3 平台特色功能	18
3.1 元数据驱动	18
3.1.1 规范的元模型管理	18
3.1.2 多种多样的内置采集适配器	19
3.1.3 强大的元数据管理	23
3.1.4 端到端的自动化采集	24
3.1.5 严谨的元数据变更管理	25

3.1.5.1 元数据变更订阅	26
3.1.5.2 元数据变更查询	26
3.1.6 丰富的元数据分析应用	27
3.1.6.1 血缘分析	28
3.1.6.2 影响分析	29
3.1.6.3 全链分析	29
3.1.6.4 关联度分析	30
3.1.6.5 属性差异分析	31
3.1.6.6 元数据对比分析	31
3.1.6.7 重复元数据分析	32
3.1.7 出色的元数据检核机制	33
3.1.7.1 一致性检核	33
3.1.7.2 属性填充率检核	34
3.1.7.3 组合关系缺失检核	35
3.1.7.4 元数据标准覆盖率检核	35
3.1.7.5 检核例外	35
3.1.7.6 自动检核任务	36
3.1.8 多维度的元数据监控	37
3.2 数据标准建设	37
3.2.1 配置灵活的数据标准属性	37
3.2.2 方式丰富的数据标准录入	38
3.2.3 完备的数据标准审批	40
3.2.4 先进的数据标准落地映射	41
3.2.5 智能精确的数据标准落地评估	42
3.2.6 灵活有效的数据标准监控	43
3.2.7 随时随地掌控标准变化	44
3.3 数据质量控制	45
3.3.1 多角度的规则管理	45
3.3.2 多种方式的新建规则	46
3.3.3 强大的智能修复	46

3.3.4 全面的监控管理	49
3.3.5 详尽的结果管理	50
3.3.6 多样的统计分析	51
3.3.7 细致自主的数据整改	52
3.4 数据整合处理	52
3.4.1 多样的增量捕获和数据装载	53
3.4.2 便捷的拖拽式流程设计器	53
3.4.3 可视化的调试和预装载	56
3.4.4 丰富的数据处理组件	54
3.4.5 “零”表达式清洗组件	55
3.4.6 灵活的调试服务	错误！未定义书签。
3.5 数据资产应用	61
3.5.1 不同视角的数据资产编目	错误！未定义书签。
3.5.2 强大的数据资产检索	错误！未定义书签。
3.5.3 丰富的数据资产服务申请	68
3.5.4 便捷的数据服务开发	69
3.6 数据生命周期监控	71
3.6.1 便捷的数据资产归档	72
3.6.2 灵活的数据资产销毁	73
3.6.3 可视化的数据生命周期监控	74
3.7 数据传输交换	57
3.7.1 跨源数据交换	57
3.7.2 多类型数据交换	58
3.7.3 灵活的任务调度	59
3.7.4 丰富的组件库	60
3.8 数据安全管控	74
3.8.1 多类型的安全规范管理	74
3.8.2 多级别的敏感数据监控	错误！未定义书签。
3.8.3 全方位的数据资产加密	75
3.8.4 多场景的数据资产脱敏	76

3.8.5 严密的权限管理	76
3.9 平台系统管理	76
3.9.1 个性化的门户定义	77
3.9.2 灵活细致的权限管理	77
3.9.3 简便的备份恢复	78
3.9.4 详细的系统日志记录	78
3.9.5 统一的调度管理	79
3.9.6 直观的工作流管理	80
3.9.7 便捷的睿治工作台	81
3.9.8 完善的系统监控体系	错误！未定义书签。
4 产品技术特点	82
4.1 高扩展	89
4.2 自动化	90
4.3 全方位	90
4.4 高性能	91
5 软硬件运行环境	91
5.1 服务器配置推荐	91
5.1.1 推荐客户：小范围试用和用户量较小的客户环境	错误！未定义书签。
5.1.2 推荐客户：大型企业、省级机构	错误！未定义书签。
5.2 客户端配置	94
5.2.1 客户端(建议配置)	94
5.2.2 客户端浏览器	95
6 平台应用案例	91
6.1 金融行业典型案例	96
6.1.1 赣州银行数据标准化项目	96
.....	96
.....	96
6.1.1.1 项目概述	96
6.1.1.2 项目成果	96
6.1.1.3 项目建设内容	97

6.1.1.4 项目展示	98
6.1.2 招商局租赁数据治理项目	98
6.1.2.1 项目概述	98
6.1.2.2 项目价值	98
6.1.2.3 项目建设内容	99
6.1.2.4 项目长远规划	100
6.2 政府项目典型案例	100
6.2.1 禅城区大数据统筹局数据治理项目	100
6.2.1.1 项目概述	100
6.2.1.2 项目价值	100
6.2.1.3 项目建设内容	101
6.2.1.4 项目展示	102
6.2.2 荔湾区政务数据服务平台项目	103
.....	103
6.2.2.1 项目概述	103
6.2.2.2 项目价值	103
6.2.2.3 项目建设内容	103
6.2.2.4 项目展示	104
6.2.3 卫计委-数据质量管控系统	104
.....	104
6.2.3.1 项目概述	104
6.2.3.2 项目价值	105
6.2.3.3 项目建设内容	105
6.2.3.4 项目展示	106
6.3 企业项目典型案例	106
6.3.1 南山集团主数据项目	109
6.3.1.1 项目概述	109
6.3.1.2 项目价值	109
6.3.1.3 项目建设内容	110
6.3.1.4 项目建设思路	111

6.3.2 时代地产数据资产平台项目	111
6.3.2.1 项目概述	111
6.3.2.2 项目建设内容与价值	112
6.3.2.3 项目展示	114
6.4 其他典型案例	111

ESEN/SAFE

1 前言

1.1 关于本白皮书

本白皮书对应产品版本为：睿治 V2.7.1。

最后修订日期：2023 年 9 月。

本白皮书将在阐述数据治理的基础上，详细介绍睿治 V2.7.1 在功能、技术等方面的特点和优势。

1.2 数据治理定义

数据治理（Data Governance）是组织中涉及数据使用的一整套管理行为。由企业数据治理部门发起并推行，关于如何制定和实施针对整个企业内部数据的商业应用和技术管理的一系列政策和流程。

1.3 数据治理的重要性

1.3.1 企业所面临的问题

2023 年，国务院发布《数据十二条》，数据作为新型生产要素，是数字化、网络化、智能化的基础，已快速融入生产、分配、流通、消费和社会服务管理等各环节，深刻改变着生产方式、生活方式和社会治理方式。数据基础制度建设事关国家发展和安全大局。需要加快构建数据基础制度，充分发挥我国海量数据规模和丰富应用场景优势，激活数据要素潜能，做强做优做大数字经济，增强经济发展新动能，构筑国家竞争新优势。但是，当前我国公共数据治理仍面临着资源管理职责不清晰、数据资源供给不顺畅、数据

质量标准不完善、开发应用规则不健全、安全规范监管不到位等痛点堵点，具体问题如下：

问题 1：业务系统之间缺少统一标准

各系统收集上来的数据多样化缺少数据标准，对表字段的命名随意性强，定义混乱，主要表现在：同一英文字段名对应多个中文名，反之也存在着中文的一对多；同一业务元素对应多种数据类型和长度等情况。

问题 2：系统之间数据变更对应的影响分析困难

已存在的历史数据往往庞大复杂，表关系不清晰，如遇到系统改造、表结构变更等情况，对应造成的影响难于甚至无法评估。

对于已然形成的大量数据，系统悠久，找不到接口人，针对其应用困难。

问题 3：数据质量差，错误发现不及时

由于数据质量低下导致统计不准确，许多预期需求无法实现、造成决策失误。数据中的错误发现得不及时，数据处理过程中的问题也未能及时纠正，导致更多的后续错误。

问题 4：数据表和模型繁多，价值未最大释放

系统中表的数量多、相似表数量过多、无效表过多；导致信息获取难，表价值不明。

问题 5：业务系统之间数据资产分散，共享性差

DB、数据模型、应用程序、数据标准、数据质量等信息分散，企业数据资产无法集中管理和展现，共享性差。

问题 6：数据安全得不到保障

数据权限划分不清、敏感数据得不到监控、账号和权限无法追踪、非正常的访问行为无法捕捉等情况都是潜在的数据安全隐患。

问题 7：管理不到位，落地效果差

数据管理体系不完善，权责不明、管理角色分工不清；往往出了问题再补漏，数据管理部门和生产部门相互推脱责任。

同时，企业标准管理系统不完善，已有标准落实差；标准管理软件易用性差，落实效果差。

1.3.2 数据治理可以解决什么问题

政府、企业想要释放数据的强大力量，必须提供准确、可靠、及时的数据。睿治帮助政府和企业有效管理数据，避免因数据价值得不到很好体现而对政府和企业造成负面影响，进而帮助企业提高竞争力，为政府和企业提供更优质、更及时、更完整的数据，让其在政务管理和经营市场中脱颖而出：

1. 帮助政府和企业建设数据标准，制定统一标准。
2. 帮助政府和企业梳理资源，形成数据资产，丰富分析应用全面掌控数据来龙去脉，以获得更多的数据洞察力，进而挖掘出隐藏在资源中的价值。
3. 帮助政府和企业建立数据质量管理体系，对数据质量实时监控，及时整改，全面提升政府和企业数据的完整性、准确性、及时性，减少因数据不可靠导致的决策偏差损失。
4. 帮助政府和企业制定相关流程、政策、标准，保证信息的可用性、可获取性、优质性、一致性以及安全性，提升信息服务水准。

-
5. 帮助政府和企业优化经济治理基础数据库, 推动各地区各部门间数据共享交换。
 6. 帮助政府和企业探索建立统一规范的数据管理制度, 提高数据质量和规范性, 丰富数据产品。
 7. 提升政府和企业数据资产安全性, 并帮助建立相关安全规范和响应机制, 全面保障其数据安全。

1.4 数据治理的现状与发展趋势

1.4.1 数据治理现状

大数据的快速发展, 使它成为 IT 领域的又一大新兴产业。据中央财经大学中国经济管理研究院博士张永力估算, 国外大数据行业约有 1000 亿美元的市场, 而且每年都以 10% 的速度在增长, 增速是软件行业的两倍。数据显示, 2016-2020 年我国大数据产业规模逐渐增长, 2020 年大数据产业规模超 1.1 万亿元。中商产业研究院预测, 2024 年我国大数据产业规模达 1.57 万亿元, 产业发展潜力非常巨大。我国大数据飞速发展的背后存在诸多的问题: 相关利益交织, 协调难; 方案规划容易, 落地困难; 过度依赖技术工具; 对于数据没有明确区分。因此, 数据治理目前可谓是机遇与挑战并存, 有市场, 有需求, 但也有诸多亟待解决的问题。

1.4.2 云集成

近年来, 在大数据和人工智能的推动下, 云计算的应用进入繁荣阶段, 越来越多的企业在商业模式创新过程中都意识到要利用云计算, 迅速扩张云市场。未来将会有更多的企业把工作负载迁移到云端, 数据和系统迁移将面临更大的机遇和挑战。对于越来越多的企业而言, 云的使用将变得更加自然, 混合云将成为企业中最常见的基础架构。云

计算中的数据治理和安全控制，以及实现数据中心和云之间的双向数据传输的混合云管理将成为未来需要改进的一个关键方面。尽管云具有经济性和灵活性等优势，数据安全性(包括数据存储和数据访问)仍然是云的一大问题，需要一个综合全面的数据治理平台。

1.4.3 与人工智能紧密结合

随着数字时代的到来，数据治理和人工智能研究的关系日渐紧密，在未来大数据治理会与人工智能技术深度结合，人工智能技术在大数据治理领域将有如下等应用。

1.4.3.1 元数据

元数据，又称中介数据、中继数据，为描述数据的数据，主要是描述数据属性的信息，用来支持如指示存储位置、历史数据、资源查找、文件记录等功能。元数据算是一种电子式目录，为了达到编制目录的目的，必须在描述并收藏数据的内容或特色，进而达成协助数据检索的目的。

1.4.3.2 数据标准

在数据标准的实施初期，需要对存量系统的数据库字段进行摸底，识别出共有的、重复使用的业务字段，作为建立数据标准的依据。如果完全靠人工梳理，需要协调各业务部门大量人员参与，工作量巨大且容易出错。借助机器学习、自然语言处理技术，可以根据字段业务名快速的整理出高频词根，将可能需要几个月的工作在几天内完成。

1.4.3.3 数据质量

在金融行业中，银行信用卡发卡部门已经开始利用机器学习技术，来识别不合规的申请人、虚假申请信息及可能存在欺诈性的交易行为。此外，既然机器学习可识别信息

系统中的异常数据，那它也可以检测工业制成品中的异常情况。企业可通过将实体物品生产过程数字化，然后使用经过训练的机器学习系统来识别不符合标准或规格的产品数据，挑出异常数据，从而部分替代人类检测员的工作。

1.4.3.4 数据安全

数据安全是指让信息或信息系统免受未经授权的访问、使用、破坏、修改、销毁的过程或状态。人工智能技术可以进行敏感数据的分类分级。应用机器学习、自然语言处理和文本聚类分类技术，能对数据进行基于内容的实时精准分类分级，而数据的分类分级是数据安全治理的核心环节。例如，利用数据分类引擎在邮件内容过滤、保密文件管理、情报分析、反欺诈、数据防泄露等领域明显提升了安全性。

1.4.3.5 其他应用

➤ 数据共享开放

企业可以充分利用人工智能技术，以信息化、自动化方式，共享和开放一部分数据或数据加工结果，对外提供服务，提升企业竞争力。最典型的就是现在各种客户服务机器人，可以使用自然语言处理技术处理回答客户提出的常见问题，并随着时间的推移提高答案的质量。另外一个常见的应用领域就是营销推荐，在大多数商业环境中，将适合的产品投放给恰当的受众对促进商业成功至关重要。机器学习系统可以使用企业收集的用户数据，根据用户过去的购物习惯预测可能喜欢的物品，再将预测结果数据向企业电商系统或销售决策系统开放。

➤ 非结构化数据管理

近年来，利用机器学习的人脸识别系统能力一直在提高，已经大量应用在员工上班

考勤、识别已知的犯罪分子、识别公共场所中超出规范或违法的行为或活动。各类社交网站每天都在产生大量非结构化数据，企业可以利用机器学习技术来实时发现和识别潜在的问题，同时手写识别、语音转写、自然语言处理技术也在不同场景中大量应用，可以提高人们识别、理解和处理非结构化数据的能力。

1.4.4 数据生态到数据要素

随着数字时代的发展，数据的复杂性持续增加，越来越多的企业开始意识到数据的重要性，从被动的响应到开始主动为数据赋能，管理方式从自上而下管理到敏捷协同的一体化管理，管理手段也开始向自动化和智能化转变，构建多元化的数据生态势在必行。全社会数据综合治理发展带来的理论问题、技术问题、管理问题与评价问题，需要建立跨层级、跨行业、跨组织、跨系统的系统性数据治理新框架。全社会数据可信流通需求带来的理论问题、技术问题、管理问题与评价问题，需要建立 标准化、共识化、便捷化、持续化的平台化数据治理新体系。全社会数字化转型提升带来的理论问题、技术问题、管理问题与评价问题，需要建立前 瞻性、落地性、实效性、敏捷性的创新型数据治理新实践。

1.4.5 数据安全步入法制化

大数据的发展加速了社会的信息、资源流动，使社会运转效率更高，同时也隐藏着巨大的隐患。大量看似中性无害的数据收集以后，也可能会产生有害的后果。人们面临的威胁可能包含隐私泄露、信息丢失以及网络安全等等。《中华人民共和国数据安全法》和《个人信息保护法》中均指出需要通过采取必要措施，确保数据处于有效保护和合法利用的状态，以及具备保障持续安全状态的能力。数据作为在数字经济时代的关键生产

要素，其自身具有很大的经济价值，数据安全法的发布，标志着国家鼓励数据依法合理有效利用，保障数据依法有序自由流动，促进以数据为关键要素的数字经济发展。

2 平台简介

2.1 平台定位

亿信睿治平台是北京亿信华辰软件有限责任公司自主研发的由多个产品组成的一整套解决方案，是一款面向实施人员的、智能的、敏捷的数据全生命周期管理应用平台。平台以元数据为基础，所有模块并非串连，而是每个模块都可以单独或与其他模块组合使用，并支持在本地或云上使用。

睿治 V2.7.1 版本采用云原生架构，是一款覆盖数据全生命周期的数据治理平台，它通过对数据从创建到消亡的全过程监控和治理，实现数据的统一管理，为企业保证了业务数据在采集、集中、转换、存储、应用整个过程中的完整性、准确性、一致性和时效性，从而帮助客户建立起符合自身特征的数据架构和数据治理体系。亿信睿治主要涵盖了以下内容：

模块	说明
元数据	元数据是数据治理的基础，是对数据的描述，可总览分析元数据之间的关系和互相之间的影响。
数据标准	数据标准是对数据进行统一的、规范的定义，确保数据在复杂数据环境中保持一致性、规范性，提高数据管控的效率，完善数据治理体系。
数据质量	数据质量包括对数据质量规则的定义、数据质量检查方案的定义，支持定时自动执行质量检查，自动生成质检报告，监控并揭示数据质量问题，提供问题明细查询和质量改进建议。
数据集成	数据处理主要用于对数据进行清洗、转换、整合、模型管理等处理工作，既可以用于对数据问题进行可行的修正，也可以用于为数据应用提供可靠的数

	据模型。
数据资产	数据资产是将全部数据作为资产，进行不同角色的目录化管理，分析数据资产之间的关系，获知资产访问、利用情况等。
数据交换	数据交换用于实现不同机构不同系统之间的数据传输交换，完成基本的数据汇集，确保数据在传输中的完整性、一致性。
数据生命周期	对数据的全生命周期进行管理，主要是根据设置对数据进行自动归档和销毁，对近线数据和离线数据分别管理。
数据安全	数据安全贯穿于数据治理全过程，用于保证数据的安全性，提供对隐私数据的加密、模糊化处理。

2.2 平台价值

企业进行数据治理的最大驱动力来自数据质量，通过提高数据质量实现更多的业务价值。睿治平台主要目的是发现并解决数据问题，通过一系列措施规范数据，减少数据问题发生，整体提高数据的应用价值。将实现业务目标作为数据管理和服务的核心驱动力，优化数据架构，提升数据仓库/信息化管理系统建设，支持管理能力的提高、精细化和决策的科学性。

睿治平台的价值主要体现在：

- 帮助企业降低梳理业务系统人工成本，IT 建设和运维成本；
- 建立统一企业数据标准，规范数据体系，实现统一规范管理；
- 挖掘隐藏的数据关系网络，从而快速了解业务相关内容，提升效率；
- 全面提升企业的数据质量，深度再造与数据相关的业务流程，满足业务拓展需要；
- 健全的数据资产管理，满足风险管理、成本核算、绩效考核等需要；
- 直观精准的分析应用与展示，保障高层决策的合理性和可靠性；
- 保障企业数据安全，避免因数据安全问题造成损失。

2.3 平台架构



数据治理平台架构图

睿治平台从下至上主要分为以下几层：

数据源层，主要用来管理各类数据源，内部数据如：OA、ERP、CRM 等业务系统数据，也支持文本文件上传；外部数据如：社交媒体、互联网等。

数据存储层，通过数据交换将数据输送到 ODS、数据仓库、数据集市进行存储。

数据处理层，主要包括元数据、数据标准、数据质量。元数据可进行元数据采集、血缘分析、影响分析、全链分析，元数据变更管理、绘制数据地图等；数据标准可定义数据标准、标准落地、标准评估监控，及标准版本管理等；数据质量可管理质量规则、数据模型，及数据质量方案。

数据共享层，主要包括数据资产管理。数据资产管理包括数据资产视图、数据资产检索、资产采集更新等功能。

应用层，该层主要通过平台便捷的应用功能，从而改善决策支撑、缩减管理成本、降低数据风险、提升数据价值，并提供数据公开等。

2.4 核心优势

2.4.1 各产品模块高度融合

睿治平台融合了元数据、数据标准、数据质量、数据集成、数据资产、数据服务、数据安全、数据生命周期等 10 大产品，每个模块功能可互相调用，全程可视化操作，打通数据治理各个环节，同时提供各个产品模块任意组合快速解决企业不同的数据治理场景。

2.4.2 先进的产品设计理念

睿治平台采用成熟的体系结构，具有国内先进水平，设计过程充分依照国际规范、标准。丰富的对外接口，使系统具有高度灵活性、扩展性和集成能力。安全方面，平台采用全方位、严密的安全编码加强产品安全性，对不安全的数据进行严格校验。同时，程序具备自检、故障诊断等功能，在出现系统异常时，能及时、快速地进行自我修复，以保证系统具有极佳的稳定性和更长的生命力。

2.4.3 专业的服务保障团队

亿信华辰具有 300+大数据分析师和实施顾问、180+研发专家所组成的专业服务团队，拥有专业的数据治理团队，服务人员遍布全国各省市。服务团队既可对数据治理的工具、方法、模板进行咨询方案设计，也可对数据治理的策略、组织架构、处理流程、支持环境和规章制度等进行合理的规划，为客户提供专业完整的服务和咨询。

2.4.4 丰富的项目实践经验

亿信华辰深耕商务智能和大数据领域十多年，着眼于打造数据全生命周期的智能化

产品线，致力于帮助企业 and 政府一站式解决数据应用难题，在银行、租赁、卫生等行业拥有丰富的数据治理实践经验和完整案例，可根据行业特点和业务组织的数据治理要求，为企业提供咨询和 IT 一体化的综合数据治理解决方案。

2.4.5 云原生架构

睿治数据治理平台 V2.7.1 支持云原生架构。划分了元数据、数据标准、数据集成、数据质量、数据资产、数据服务、数据安全、数据生命周期等服务，服务之间通过 restfull api 通信，实现微服务后产品功能的模块化、高可用、高并发、扩展性强等特点；

支持云上容器化部署，实现高水平资源隔离，适应了当下比较流行的云部署场景；

无状态化，支持弹性扩展，通过多进程承载高并发，根据并发的压力用多个副本共同承担流量；

模块服务化，当模块需要更新时，只需要更新某个服务，其他服务可正常请求访问。

2.4.6 更智能的治理成果展示

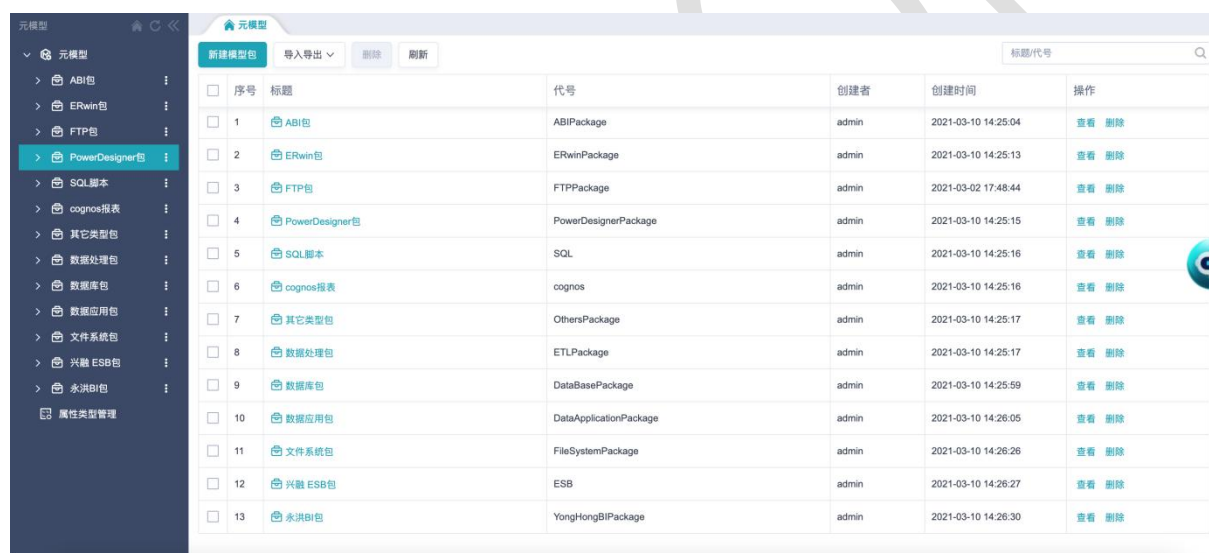
小智博士：利用最新的自然语言处理（NLP）、机器学习、知识图谱等人工智能技术，用语音/文字输入一句话，就能获取最贴切的数据分析图表。智能出图，无需做表，零门槛使用，快速的展现数据治理的成果；

3 平台功能

3.1 元数据驱动

3.1.1 规范的元模型管理

睿治元数据以 Meta Object Facility(MOF)规范为基础，支持 XML 格式的元模型导入和导出，同时内置大量技术元数据、业务元数据的元模型，用户可直接使用。元模型管理对元模型的基本信息、属性、父子关系、依赖关系、组合关系的增删改查操作，内置元模型的内置信息不允许修改或者删除，但可进行新增操作。具体功能界面如下：



序号	标题	代号	创建者	创建时间	操作
1	ABI包	ABIPackage	admin	2021-03-10 14:25:04	查看 删除
2	ERwin包	ERwinPackage	admin	2021-03-10 14:25:13	查看 删除
3	FTP包	FTPPackage	admin	2021-03-02 17:48:44	查看 删除
4	PowerDesigner包	PowerDesignerPackage	admin	2021-03-10 14:25:15	查看 删除
5	SQL脚本	SQL	admin	2021-03-10 14:25:16	查看 删除
6	cognos报表	cognos	admin	2021-03-10 14:25:16	查看 删除
7	其它类型包	OthersPackage	admin	2021-03-10 14:25:17	查看 删除
8	数据处理包	ETLPackage	admin	2021-03-10 14:25:17	查看 删除
9	数据库包	DataBasePackage	admin	2021-03-10 14:25:59	查看 删除
10	数据应用包	DataApplicationPackage	admin	2021-03-10 14:26:05	查看 删除
11	文件系统包	FileSystemPackage	admin	2021-03-10 14:26:26	查看 删除
12	兴融 ESB包	ESB	admin	2021-03-10 14:26:27	查看 删除
13	永洪BI包	YongHongBIPackage	admin	2021-03-10 14:26:30	查看 删除

元模型查看界面

元模型支持发布功能，只有发布后的元模型才可被元数据使用，同时支持查看所有发布版本。

元模型	元模型	数据源包(元模型)	关系型数据源包(数据源包)	元模型详情(表)	元模型历史版本
元模型 AR包 ER包 FTP包 PowerDesigner包 SQL脚本 cognos脚本 其它类型包 数据源包 数据源包 大数据包 关系型数据源包 表 表分区 表空间 表结构 存储过程 存储过程脚本 存储过程子 函数 函数参数 表 检查约束	版本	发布时间	发布人	备注	操作
	v25	2023-09-01 16:04:06	admin	自动更新并发布内置模型	查看
	v24	2023-08-29 10:40:34	admin		查看
	v23	2023-08-28 11:09:25	admin	自动更新并发布内置模型	查看
	v22	2023-02-09 10:39:12	admin	自动更新并发布内置模型	查看

元模型版本查看界面

通过发布过程，将元模型的设计和运用隔离开，元模型只有在发布之后才会生效，使用户在设计完成发布之前，不会影响到元数据的使用。

3.1.2 丰富内置采集适配器

元数据平台总共支持近 60 种元数据适配器，其中数据库类型有 20+种，报表工具有 14 种，etl 工具有 9 种等，满足用户多样的采集需求。

元数据

数据标准

数据集成

数据质量

数据资产

数据服务

数据安全

数据生命周期

智能分析

适配管理

国产数据库(11)

大数据库(7)

关系型数据库(9)

报表工具(8)

建模工具(2)

数据处理工具(6)

其他(10)

适配管理

数据源管理

采集任务配置

采集日志

库串审核

采集模板配置

GBase 8a 采集适配器

工具名称: 南大通用GBase 8a 数据库

工具版本: 8.x

描述: 采集数据库结构及相关对象

GaussDB(DWS) 采集适配器

工具名称: 华为云GaussDB(DWS) 数据库

工具版本: 8.1.x

描述: 采集数据库结构及相关对象

GaussDB(openGauss) 采集适配器

工具名称: 华为云GaussDB(openGauss) 数据库

工具版本: 2.1.0

描述: 采集数据库结构及相关对象

KingbaseES 采集适配器

工具名称: 人大金仓KingbaseES 数据库

工具版本: 8.2, 8.3, 8.6

描述: 采集数据库结构及相关对象

LinkooDB 采集适配器

工具名称: LinkooDB 数据库

工具版本: 2.1.0

描述: 采集数据库结构及相关对象

MaxCompute 采集适配器

工具名称: MaxCompute 数据库

工具版本: 通用

描述: 采集数据库结构及相关对象

PetaBase 采集适配器

工具名称: PetaBase 数据库

工具版本: 2.1

描述: 采集数据库结构及相关对象

优炫采集适配器

工具名称: 优炫数据库

工具版本: 2.1.1.x

描述: 采集数据库结构及相关对象

华宇ArteryBase 采集适配器

工具名称: 华宇ArteryBase 数据库

工具版本: 3.x, 6.x, 7.x

描述: 采集数据库结构及相关对象

星环 Transwarp 采集适配器

工具名称: 星环Transwarp 数据库

工具版本: 5.x, 6.x

描述: 采集数据库结构及相关对象

达梦数据库采集适配器

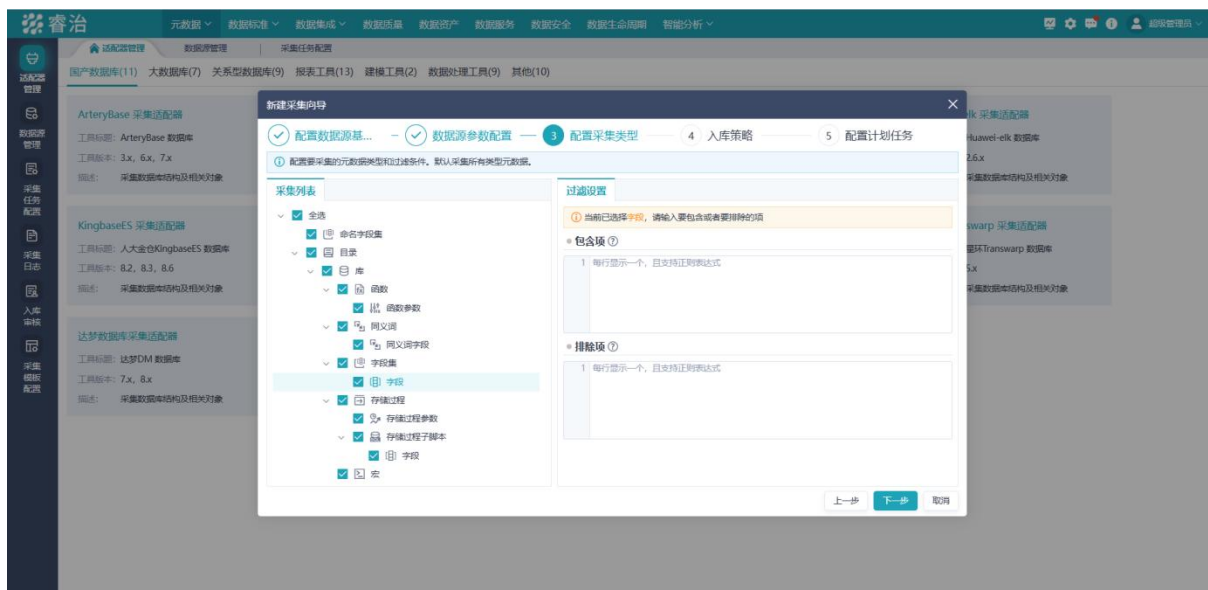
工具名称: 达梦DM 数据库

工具版本: 7.x, 8.x

描述: 采集数据库结构及相关对象

元数据采集适配器

同时采集适配器支持用户自主配置需要采集的元数据类型和过滤条件。满足用户各种采集需求。



自定义配置采集类型

本版本内置的元数据采集适配器：

国产数据库：

- ArteryBase
- Gbase 8a
- GaussDB (DWS、openGauss)
- KingBase8
- MaxCompute
- PetaBase
- 星环 Transwarp
- 优炫数据库

➤ 达梦

➤ LinkoopDB

大数据库：

➤ Elasticsearch

➤ HBase

➤ ClickHouse

➤ HDFS 文件系统

➤ Hive

➤ MongoDB

➤ Vertica

关系型数据库：

➤ Oracle

➤ MySQL

➤ SQL Server

➤ PostgreSQL

➤ Greenplum

➤ DB2

➤ Sybase ASE

➤ Informix

➤ HANA

报表工具:

- 亿信 ABI
- COGNOS
- 亿信 BI
- SSAS
- SSRS
- i@Report
- 指标平台
- 润乾报表

建模工具:

- ERwin
- PowerDesigner

数据处理工具:

- DataStage
- IBM CDC
- Kettle
- ODI
- Solution
- 兴融 ESB
- 数据交换平台

-
- 数据工厂

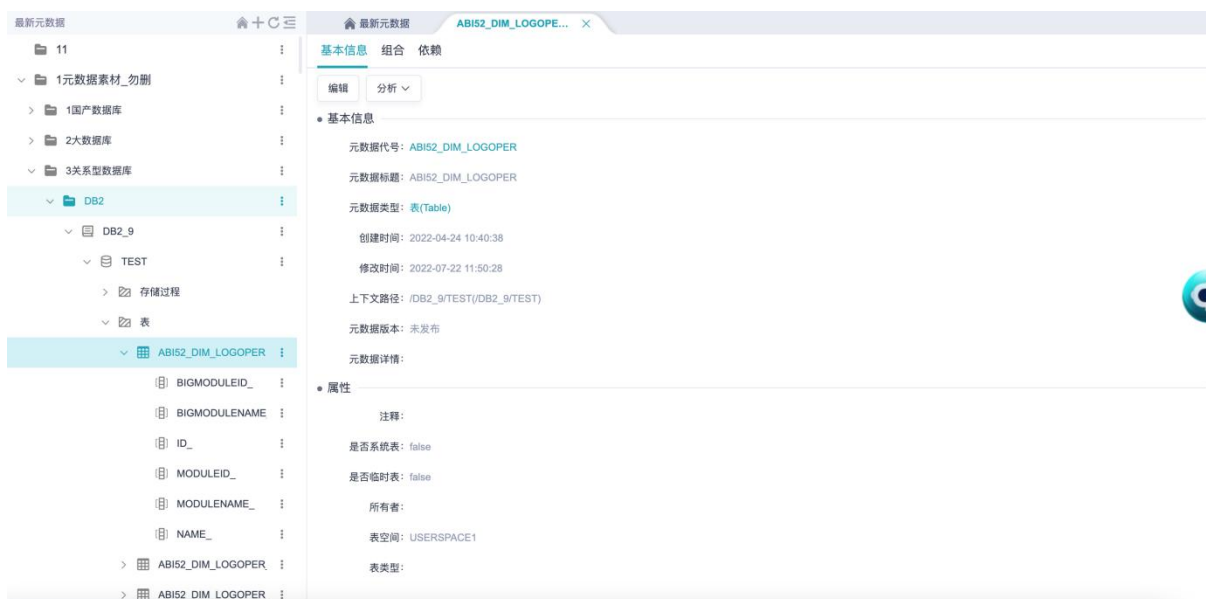
其他:

- Excel 补录
- Excel
- FTP
- JDBC 驱动
- Perl 脚本
- Python 脚本
- SQL 脚本
- 依赖关系
- 库表采集
- 文件系统
-

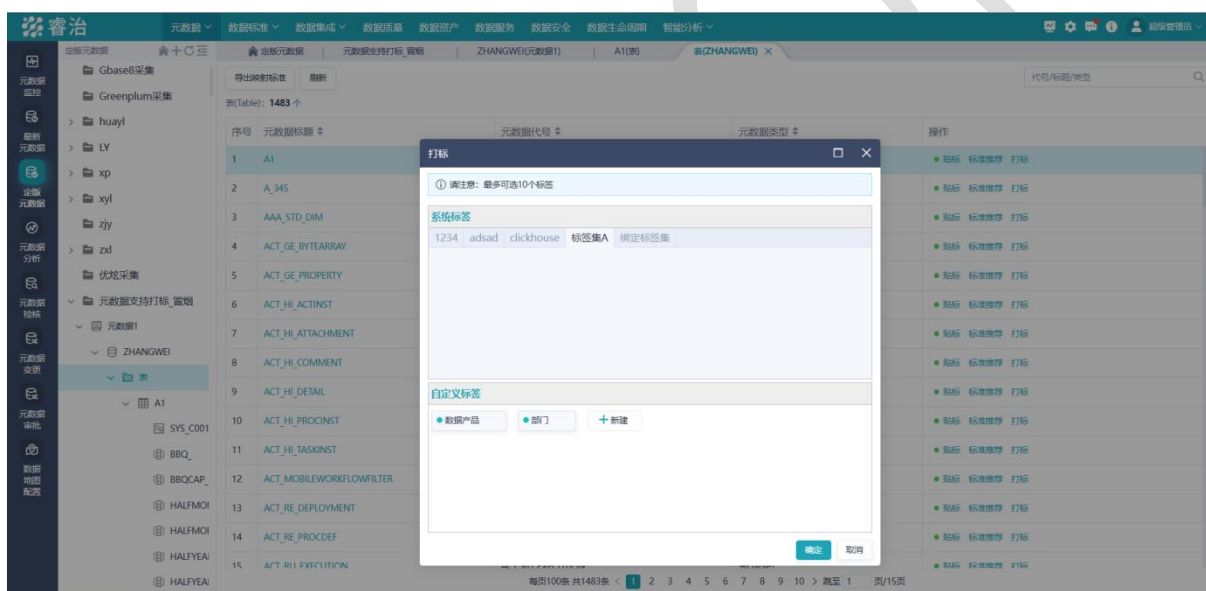
3.1.3 功能全面的元数据管理

元数据管理功能提供各类元数据管理，包括：业务元数据、技术元数据和管理元数据，支持元数据的基本信息、属性、依赖关系、组合关系的增删改查操作。

同时支持对元数据进行打标，支持表级别的数据标签，且打过标签的元数据支持同步到数据资产，实现元数据与数据资产的关联。



元数据管理界面

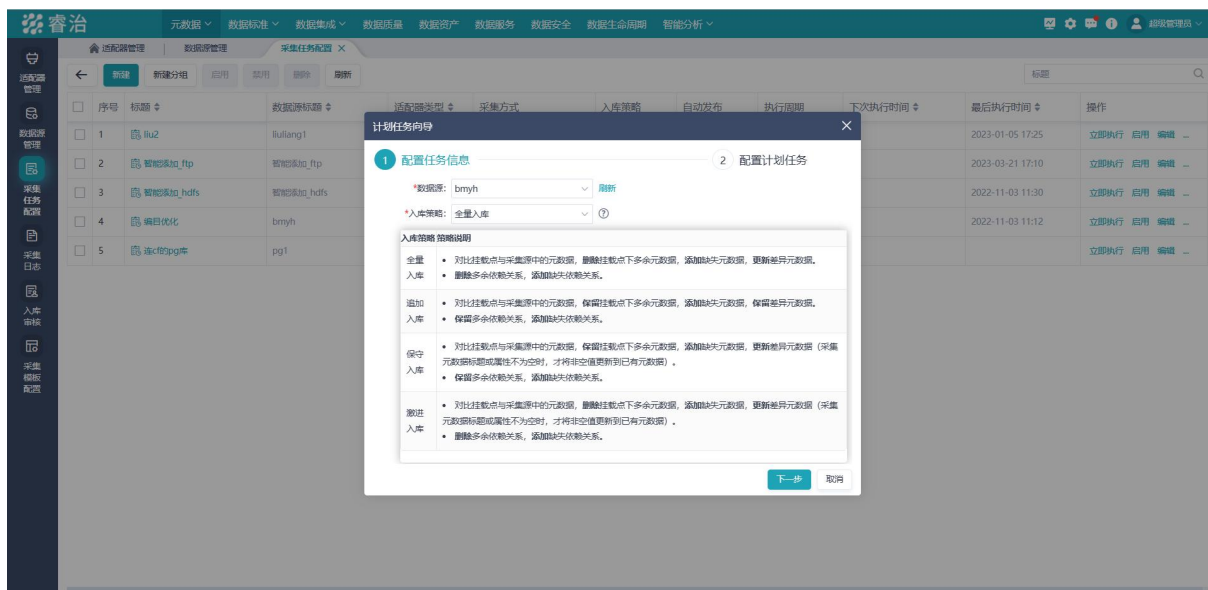


元数据打标界面

3.1.4 端到端的自动化采集

对元数据信息的维护除界面手动操作方式外,元数据管理平台利用内置采集适配器,让用户通过配置数据源参数及定时采集任务,进行自动化采集,实现直连数据源的端到端元数据采集,同时可对不需要采集的元数据进行过滤设置。

同时支持设置四种类型的入库方式,满足用户任何更新的需求场景。



采集任务配置界面

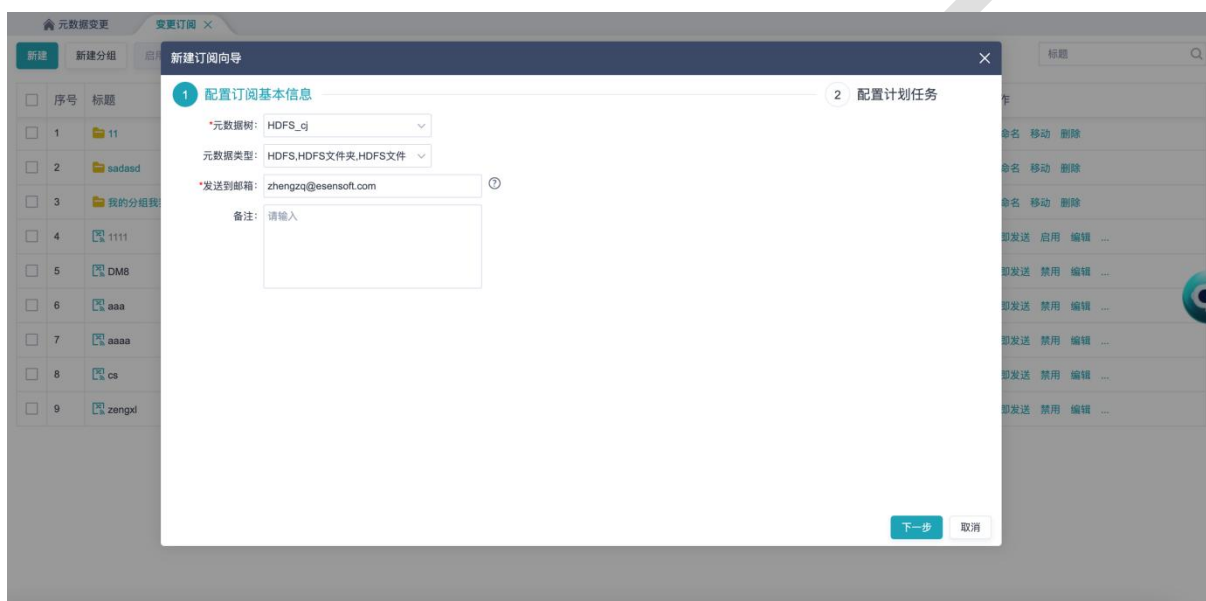
元数据管理系统提供了丰富的内置适配器，来保证自动化采集的同时，还支持对适配器进行扩展。

3.1.5 元数据变更管理

元数据管理平台面向不同场景，提供了严谨的元数据变更管理相关功能。包括元数据变更订阅、元数据变更查询。

3.1.5.1 元数据变更订阅

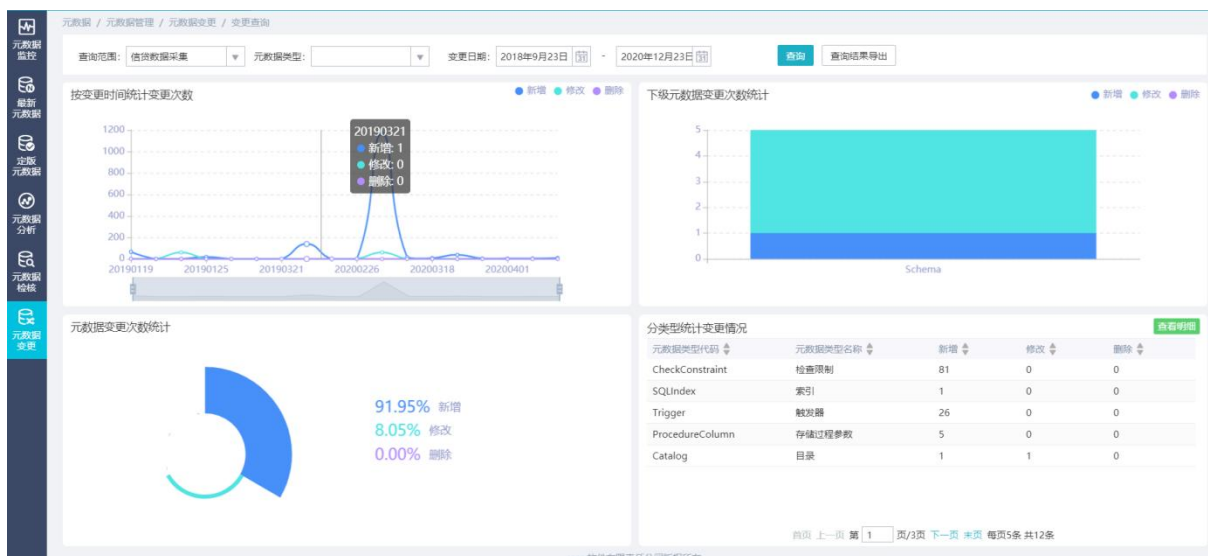
元数据变更订阅是对指定元数据进行订阅，该元数据发生变更，消息将及时推送至短信、邮箱、企业微信、钉钉等即时通讯工具，及时把握元数据变更情况，对变更元数据进行变更跟踪。



元数据变更订阅

3.1.5.2 元数据变更查询

元数据变更查询是通过平台查询元数据变更历史，对最新元数据的每一次变更进行留痕，所有元数据历史信息皆按需可查询。可以选择查询范围、元数据类型、变更日期等多维度条件进行查询。

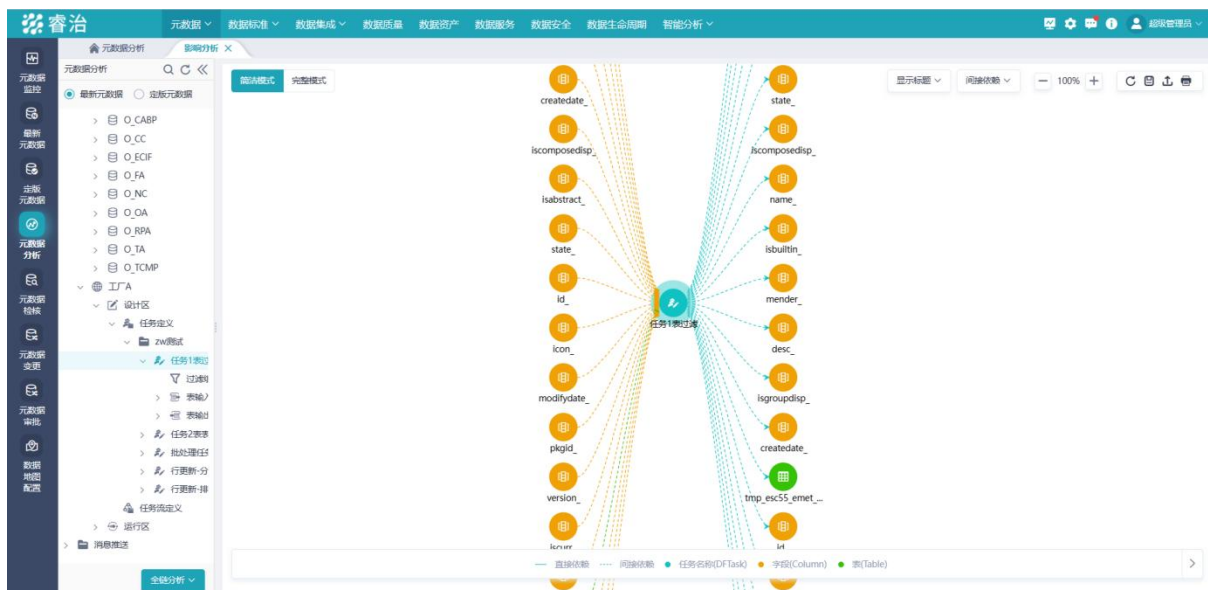


元数据变更查询

3.1.6 元数据分析应用

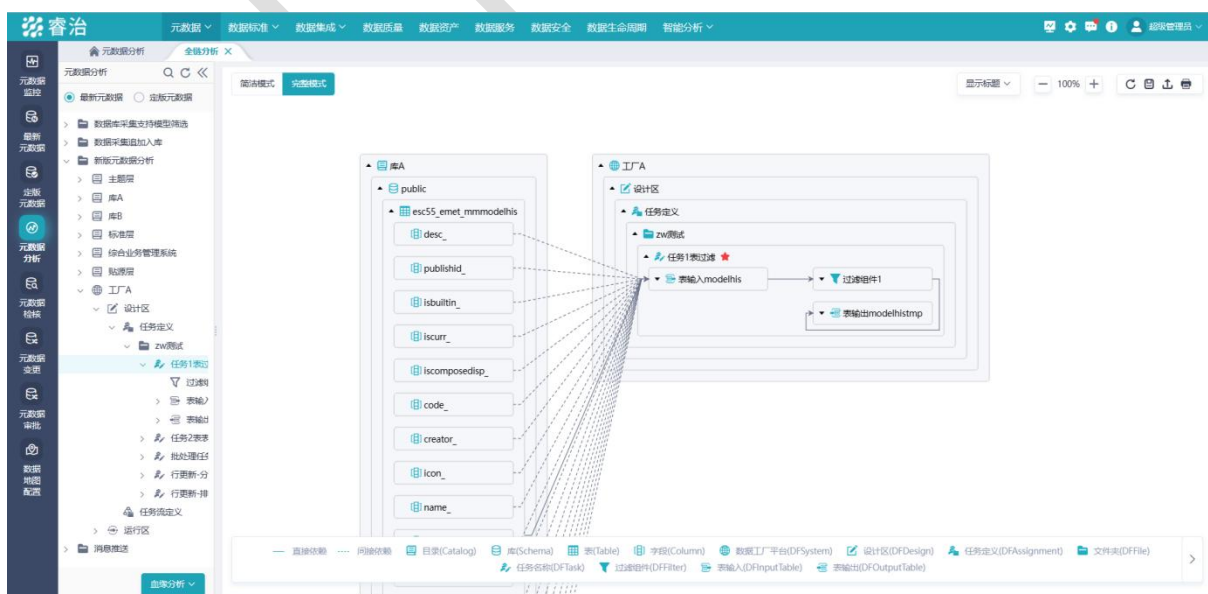
元数据管理平台提供了丰富的分析应用，包括：血缘分析、影响分析、全链分析、关联度分析、属性差异分析、元数据对比分析、重复元数据分析、元数据对比分析、重复元数据分析、同时支持将分析结果进行导出和收藏保存。

同时支持用户进行完整模式/简洁模式的切换，根据不同的需求场景进行切换使用。如用户需要查看完整的数据上下级关系则可使用完整模式，如用户需要查看数据的简单链路则可使用简洁模式。



3.1.6.1 血缘分析

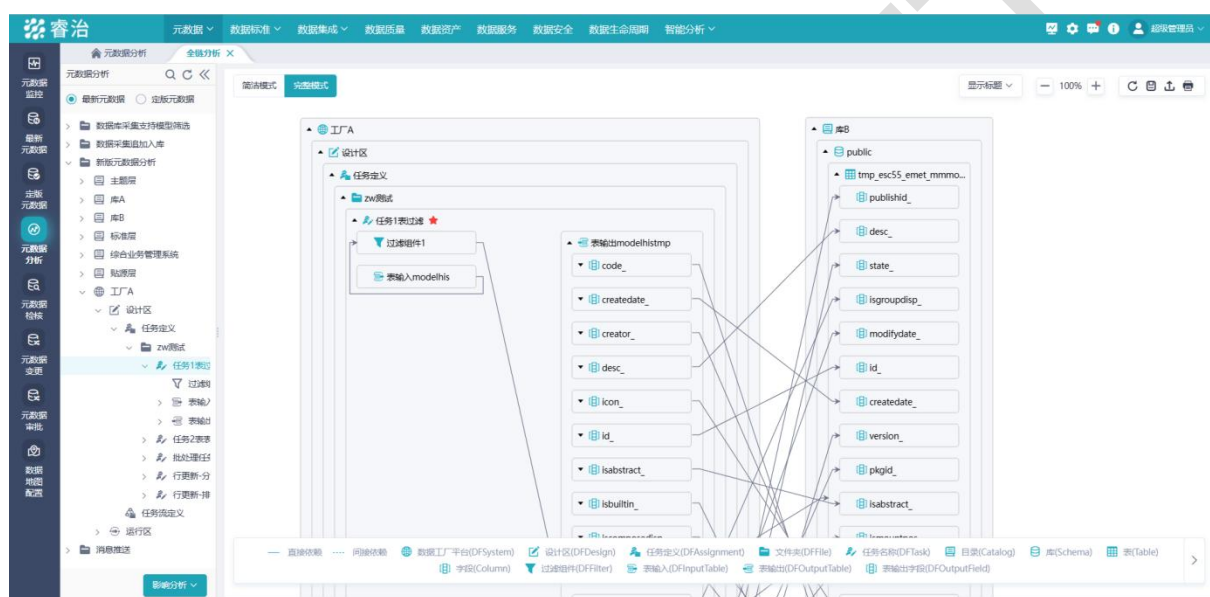
血缘分析是对指定元数据的起源及其推移位置的分析。它反应数据的来源与加工过程，还描述了数据在不同过程中发生的情况。它可以帮助分析信息的使用方式并追踪用于特定用途的关键信息位，支持导出、打印。



血缘分析

3.1.6.2 影响分析

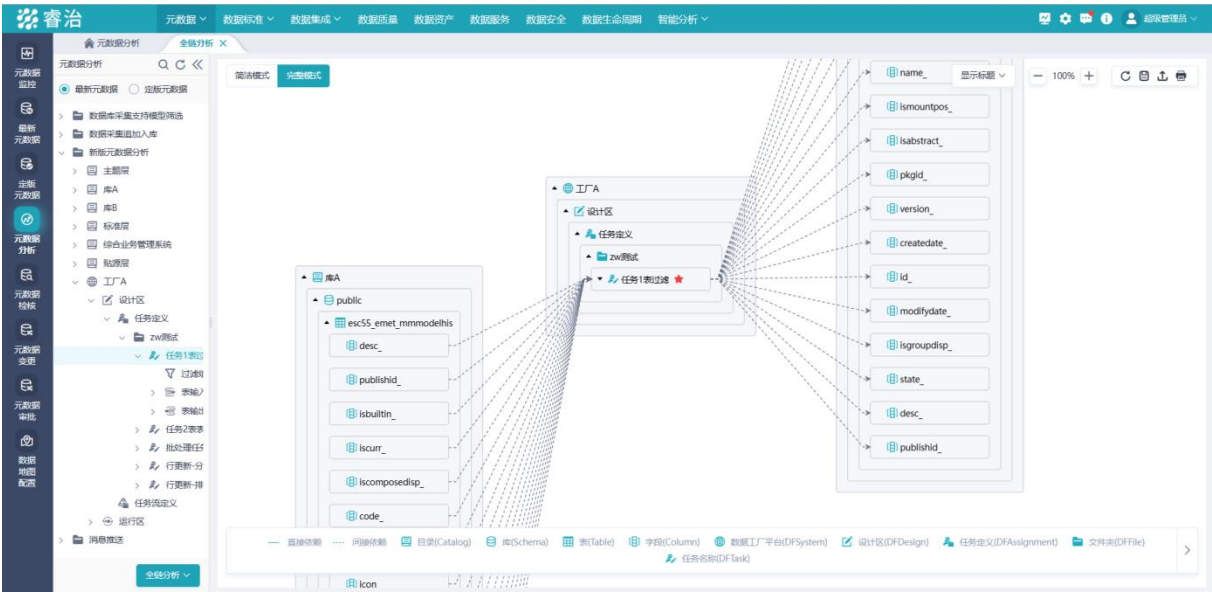
影响分析帮助用户迅速了解分析对象的下游数据信息，快速掌握元数据变更可能造成的影响，以便更有效的评估变化该元数据带来的风险，从而帮助用户高效准确的对数据资产进行清理、维护与使用,支持导出、打印。



影响分析

3.1.6.3 全链分析

全链分析是用来分析指定元数据前后与其有关系的所有元数据，不仅反应了元数据的来源与加工过程，也反应了元数据的使用情况，使用全链分析可清晰的了解该元数据的来龙去脉,支持导出、打印。



全链分析

3.1.6.4 关联度分析

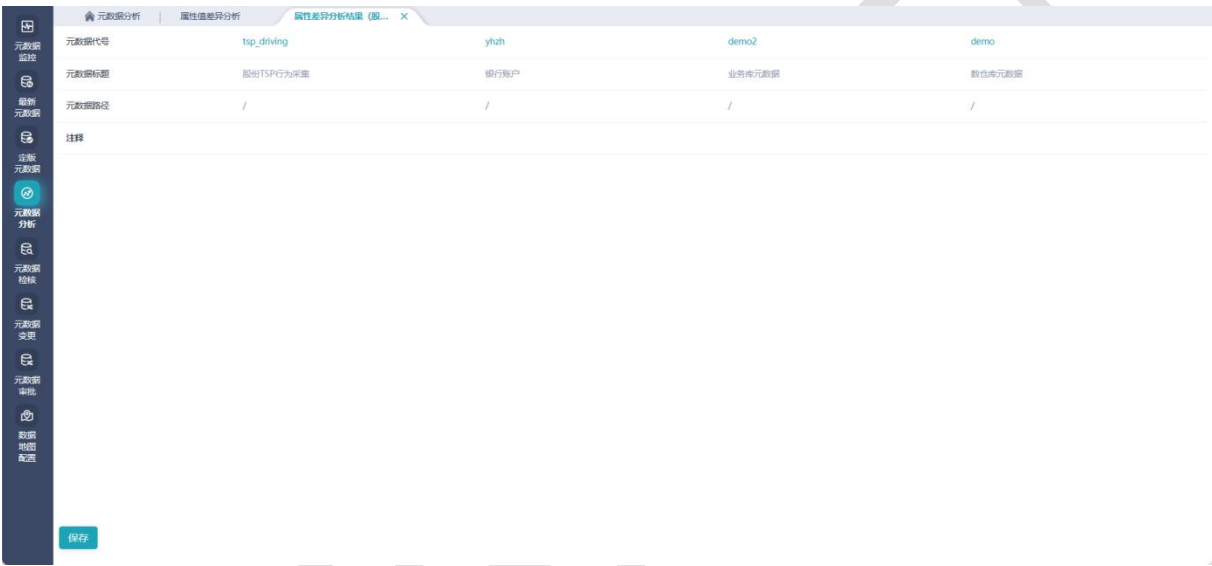
关联度分析是从关系数量的角度对指定元数据进行分析，来体现该元数据在系统中依赖程度的高低，从一定的角度可以反映出该元数据的重要程度,支持导出、打印。



关联度分析

3.1.6.5 属性差异分析

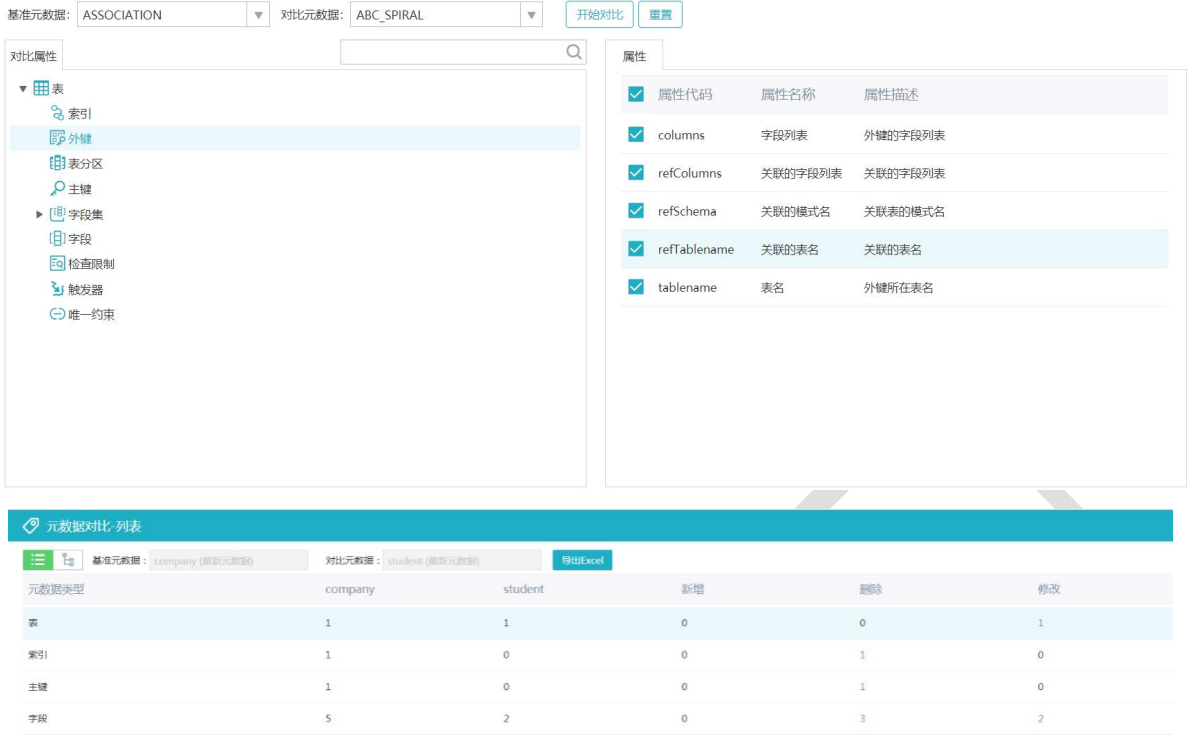
属性差异分析是用来比较同类型元数据之间属性值的差异，方便用户识别相似元数据之间的存在的微小差距，差异分析时可以选择智能推荐元数据功能，选择元数据后，可以根据元数据代码和名称自动推荐相同代码和名称且元数据类型相同的元数据，减少人工比对。



属性差异分析

3.1.6.6 元数据对比分析

元数据对比是用来检查 2 个元数据及其下级元数据之前的属性信息的差异，方便用户识别相似相同类型元数据之间的差异。



元数据对比分析

3.1.6.7 重复元数据分析

重复元数据分析，对元数据进行重复性统计，统计出哪些元数据存在重复或者属性类似的元数据，方便用户对重复的元数据进行后续的去重处理。

元数据分析							
重复元数据分析							
最新元数据		gmhgauss2.1目录gmhgreep		执行		查询结果导出	
序号	元数据代号	元数据标题	元数据路径	重复元数据代号	重复元数据标题	重复元数据路径	重复原因
1	hierarchy_	hierarchy_	/gmhgauss2.1/public/edqu_hanaZh_dc/gmhgauss...	hierarchy_	hierarchy_	/gmhgauss2.1/public/temp_edqu_datacounts/gmh...	标题,代号相同
				hierarchy_	hierarchy_	/gmhgauss2.1/public/edqu_orci2o1_rdl/gmhgauss2...	标题,代号相同
				hierarchy_	hierarchy_	/gmhgreenplum671/public/edqu_greep1_sv/gmhg...	标题,代号相同
				hierarchy_	hierarchy_	/gmhgauss2.1/public/edqu_gaussd_rt/gmhgauss2...	标题,代号相同
				hierarchy_	数据版本	/gmhgauss2.1/public/edqu_orci2o_ex/gmhgauss2...	代号相同
				hierarchy_	hierarchy_	/gmhgreenplum671/public/edqu_greep1_rdl/gmhg...	标题,代号相同
				hierarchy_	hierarchy_	/gmhgauss2.1/public/edqu_orci2o_rdl/gmhgauss2...	标题,代号相同
				hierarchy_	数据版本	/gmhgauss2.1/public/edqu_gaussd_ex/gmhgauss2...	代号相同
				hierarchy_	hierarchy_	/gmhgauss2.1/public/edqu_gaussd_rdl/gmhgauss2...	标题,代号相同
				hierarchy_	hierarchy_	/gmhgauss2.1/public/edqu_orci2o_sv/gmhgauss2...	标题,代号相同
				hierarchy_	hierarchy_	/gmhgreenplum671/public/edqu_hanaZh_rdl/gmh...	标题,代号相同
				hierarchy_	数据版本	/gmhgauss2.1/public/edqu_orci2o_stl/gmhgauss2...	代号相同
				hierarchy_	数据版本	/gmhgreenplum671/public/esc56_edqu_schemeex...	代号相同
				hierarchy_	数据版本	/gmhgauss2.1/public/edqu_orci2o_stl/gmhgauss2.1...	代号相同
				hierarchy_	hierarchy_	/gmhgauss2.1/public/edqu_orci2o1_rtl/gmhgauss2...	标题,代号相同

重复元数据分析

3.1.7 元数据检核机制

由于元数据是很多数据管理活动的基本，所以所有类型的数据中，元数据的质量是最为重要的。元数据管理平台提供元数据质量检核功能，包括一致性检核、属性填充率检核和组合关系检核，是保障元数据质量的重要手段之一。

3.1.7.1 一致性检核

一致性检核用来检验来源系统元数据以及元数据之间信息是否一致的功能，帮助管理人员分析出元数据管理平台与来源系统之间的差异，辅助管理人员更好的维护元数据管理平台。

睿治								
元数据 数据标准 数据集成 数据质量 数据资产 数据服务 数据安全 数据生命周期 智能分析								
元数据稽核 一致性稽核 X								
数据集: 股份TSP行为数据集 执行 稽核结果导出								
序号	元数据类型	元数据集中元数据数量	数据源中元数据数量	元数据一致数量	元数据不一致数量	多余数量	缺失数量	差异数量
1	字段	375	375	0	375	0	0	375
2	库	1	1	0	1	0	0	1
3	目录	1	1	0	1	0	0	1
4	表	1	1	0	1	0	0	1
5	主键	0	720	0	0	0	0	0
6	索引	0	1173	0	0	0	0	0
7	外键	0	9	0	0	0	0	0

一致性稽核

3.1.7.2 属性填充率稽核

属性填充率稽核是用来检验选定范围内元数据属性的填充情况，方便用户了解某些重要元数据的属性的填充情况。

元数据稽核 一致性稽核 属性填充率稽核 X					
稽核范围: 银行客户 稽核元数据类型: 元数据类型 执行 稽核结果导出					
序号	元数据类型	元数据数量	属性总数	未填充属性数量	属性填充率
1	表	1	7	3	57.14%
2	检查限制	41	123	0	100.00%
3	库	1	1	1	0.00%
4	目录	1	1	1	0.00%
5	索引	2	8	0	100.00%
6	主键	1	2	0	100.00%
7	字段	81	729	260	64.33%

属性填充率稽核

3.1.7.3 组合关系缺失检核

组合关系检核根据元模型中设置的组合和被组合关系对元数据进行检查，查找出不符合元模型定义的组合关系的元数据，例如找出未与任何数据库表建立组合关系的字段。

3.1.7.4 元数据标准覆盖率检核

元数据标准覆盖率检核，用于检查元数据哪些进行了标准的落地映射，标准的覆盖率情况，方便用户了解。

序号	元数据类型	元数据数量	映射标准数	标准覆盖率
1	主键	1770	0	0.00%
2	函数	2	0	0.00%
3	唯一约束	43	0	0.00%
4	外键	15	0	0.00%
5	字段	17916	2	0.01%
6	存储过程	139	0	0.00%
7	存储过程参数	552	0	0.00%
8	库	5	0	0.00%
9	检查限制	25	0	0.00%
10	目录	6	1	16.67%
11	索引	625	0	0.00%
12	表	1826	0	0.00%
13	表空间	11	0	0.00%
14	视图	27	0	0.00%
15	触发器	1	0	0.00%

元数据标准覆盖率检核

3.1.7.5 检核例外

对检核出有问题的元数据，如果确实符合实际情况，可以进行例外操作，下次检核后不会列出例外的元数据。

元数据管理

最新元数据

定制元数据

元数据分析

元数据稽核

元数据变更

元数据审批

数据稽核配置

元数据稽核

一致性稽核

组合关系缺失稽核

标准重复稽核

稽核例外管理

元数据标题/元数据代号

取消例外

刷新

序号	稽核类型	元数据标题	元数据代号	元数据路径	未填充属性	操作
1	属性填充率稽核	A	A	/SJK/ZJ/XML_1	精度	取消例外
2	属性填充率稽核	A	A	/SJK/ZJ/XML_061	精度	取消例外
3	属性填充率稽核	A	A	/SJK/ZJ/A55_XML_OUT1	精度	取消例外
4	属性填充率稽核	A	A	/SJK/ZJ/B82	精度	取消例外
5	属性填充率稽核	A	A	/SJK/ZJ/XML_02	精度	取消例外
6	属性填充率稽核	A	A	/SJK/ZJ/CJ_1	精度	取消例外
7	属性填充率稽核	A	A	/SJK/ZJ/XML_07	精度	取消例外
8	属性填充率稽核	A	A	/SJK/ZJ/A55_OUT_1	精度	取消例外
9	属性填充率稽核	A	A	/SJK/ZJ/HTTP_OUT_6	精度	取消例外
10	属性填充率稽核	A	A	/SJK/ZJ/A55_OUT5	精度	取消例外
11	属性填充率稽核	A	A	/SJK/ZJ/JX_04	精度	取消例外
12	属性填充率稽核	A	A	/SJK/ZJ/XML_03	精度	取消例外
13	属性填充率稽核	A	A	/SJK/ZJ/A_A55_10	精度	取消例外
14	属性填充率稽核	A	A	/SJK/ZJ/XML_10	精度	取消例外
15	属性填充率稽核	A	A	/SJK/ZJ/B83	精度	取消例外

每页100条 共108条 < 1 2 > 跳至 1 页/2页

稽核例外管理

3.1.7.6 自动稽核任务

对元数据稽核除界面手动操作方式外，睿治平台让用户通过配置定时计划任务，进行自动化稽核。添加稽核的计划任务，可以按照指定时间执行稽核任务。同时稽核消息将及时推送至短信、邮箱、企业微信、钉钉等即时通讯工具。

元数据管理

最新元数据

定制元数据

元数据分析

元数据稽核

元数据变更

元数据审批

数据稽核配置

元数据稽核

一致性稽核

组合关系缺失稽核

标准重复稽核

稽核例外管理

稽核任务

标题

新建

新建分组

应用

禁用

删除

刷新

序号	标题	创建者	上次稽核时间	下次稽核时间	稽核类型	操作
1	计划任务向导					重命名 移动 删除
2	1 配置稽核条件					重命名 移动 删除
3	2 设置任务执行周期					重命名 移动 删除
4						重命名 移动 删除
5					组合关系缺失稽核	立即执行 应用 编辑

计划任务向导

1 配置稽核条件

2 设置任务执行周期

*稽核类型: 请选择

*数据源: 请选择

通知消息: ☒ 禁用 ☐ 启用

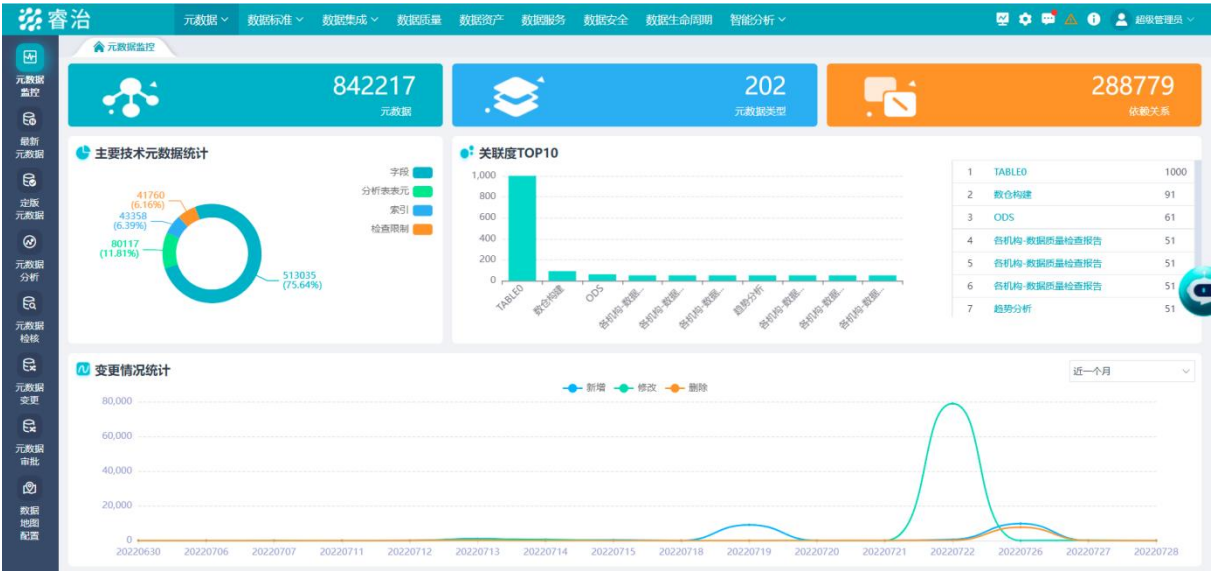
*通知方式: ☐ 邮件 ☐ 企业微信 ☐ 钉钉

下一步 取消

稽核任务管理

3.1.8元数据监控

元数据监控页面，统计了元数据各种数据以及变更和关联情况，可以对元数据的主要信息一览无余，方便用户对已经采集的元数据的情况的监控。



元数据监控

3.2 数据标准建设

睿治数据治理平台提供了一套完整的数据标准管理流程及办法，通过统一的数据标准制定和发布等一系列的活动，结合制度约束、系统控制等手段，实现企业大数据平台数据的完整性、有效性、一致性、规范性、开放性和共享性管理，为后续数据质量检查、数据安全管理等提供标准依据。

3.2.1配置灵活的数据标准属性

定义不同的数据标准可能存在需要录入不同的属性，为了满足不同项目对数据标准的设计，睿治数据治理平台提供了数据标准集管理，内置了业务属性、技术属性、管理属性、质量属性、主数据属性、生命周期属性等供用户选择使用，并支持自定义属性。

编辑标准集属性

基本属性

标准集属性

属性类型

全部属性

业务属性

技术属性

添加属性

添加自定义属性

删除属性

刷新

分类筛选: 请选择

代码/名称

☐	序号	属性代码	属性名称	属性类型	字段类型	字段长度	字段精度	是否允许为空	属性描述	是否在列表显示	默认值	对应元数据	操作
☐	1	code	标准编号	业务属性	字符型	200		否	标准编号	是			编辑 删除
☐	2	cname	标准中文名称	业务属性	字符型	200		否	标准中文名称	是		字段(Column	编辑 删除
☐	3	ename	标准英文名称	业务属性	字符型	200		否	标准英文名称	是			编辑 删除
☐	4	subject	标准主题	业务属性	字符型	200		是	标准主题	是			编辑 删除
☐	5	businessMe...	业务含义	业务属性	大字段型	0		是	完整、准确描...	是			编辑 删除
☐	6	referenceDo...	参考文献	业务属性	字符型	200		是		是			编辑 删除
☐	7	referenceCo...	引用公共代码	业务属性	字符型	200		是		是			编辑 删除
☐	8	dataType	数据类型	技术属性	字符型	100		是	数据库字段的...	是		字段(Column	编辑 删除
☐	9	dataLength	数据长度	技术属性	整型	5		是	数据库字段的...	是		字段(Column	编辑 删除

新建标准集

3.2.2 智能化的数据标准推荐

睿治平台内置大量词根和同义词词典，通过算法可以向用户智能推荐需要制定标准的数据，快速创建数据标准，缩短维护标准的周期。

智能推荐

请勾选想要添加的数据标准

☐	序号	推荐标准	推荐理由	是否采用	字段元数据名称	元数据路径	数据源
☐	1	修改日期	多来源类型	否	修改日期	/oracle/ZHOU/ED...	oracle
					修改日期	/oracle/ZHOU/ED...	oracle
					修改时间	/oracle/ZHOU/ED...	oracle
					修改日期	/oracle/ZHOU/ED...	oracle
					修改时间	/oracle/ZHOU/ED...	oracle
☐	2	昵称	多来源类型	否	名称	/oracle/ZHOU/ED...	oracle
					名称	/oracle/ZHOU/ED...	oracle

首页 上一页 第 1 页/542页 下一页 末页 每页20条 共10834条

下一步 取消

数据标准智能推荐-勾选标准

智能推荐

请确认添加的标准，并可以补充标准信息

序号	标准编号	标准中文名称	标准英文名称	标准主题	业务含义	制定依据	归属管理部门
1	MODIFYDATE_	修改时间	Time_Date				

首页 上一页 第 1 页/1页 下一页 末页 每页20条 共1条

上一步

确认

取消

数据标准智能推荐-补充标准信息

3.2.3 方式丰富的数据标准录入

睿治平台提供方便灵活的操作界面，根据用户选择合适的方式，快速创建数据标准，支持用户手动创建数据标准，同时支持拾取元数据生成数据标准，简化数据标准创建的步骤，同时支持修改、删除等操作。

除了手动创建外，还支持通过导入的方式进行批量创建。通过导出标准集，让用户在线下对数据标准进行整理，将整理完成的数据标准导入到平台后，成为一条可映射、评估的数据标准。

新建标准(部门数据标准)

业务属性

*标准编号:

*标准中文名称:

*标准英文名称:

*一级分类:

*二级分类:

业务含义:

编码规则:

管理属性

归属管理部门:

参考文档:

...

保存

新建标准

从元数据拾取

1 选择元数据

2 与元数据映射

3 从元数据拾取

4 拾取成功

待选元数据

名称

定制元数据

数据地图

数据质量

政务信息

公共数据

DG

函数

COUNT_ROWS

IDCARD15TO18

存储过程

表

视图

融合库元数据

数据治理平台元数据

原始库元数据

治理库元数据

已选元数据

名称

COUNT_ROWS

IDCARD15TO18

清空

下一步

从元数据拾取

3.2.4 完备的数据标准审批

数据标准创建保存，并确认无误后，支持整集发起审批。审批支持单个或者是批量通过、退回、部分退回操作，可采用邮件或任务提醒的方式通知参与审批的用户。同时支持审批列表的搜索，快速定位数据标准。

标准发布审批											
待审批 已审批											
批量发布 批量退回 刷新											
标准集/路径/发布路径											
☐	序号	标准集	路径	发布路径	标准个数	修改标准数	新增标准数	删除标准数	提交人	提交时间	操作
☐	1	CsBug3	/	/	1	1	0	0	admin	2023-09-14 11:17:01	查看 发布 退回
☐	2	链接列表	/zj	/zj	3	0	1	0	admin	2023-09-14 10:32:06	查看 发布 退回
☐	3	CsBug1	/	/	1	1	0	0	admin	2023-09-13 17:51:37	查看 发布 退回
☐	4	CsBug2	/	/	1	0	1	3	admin	2023-09-13 17:51:33	查看 发布 退回

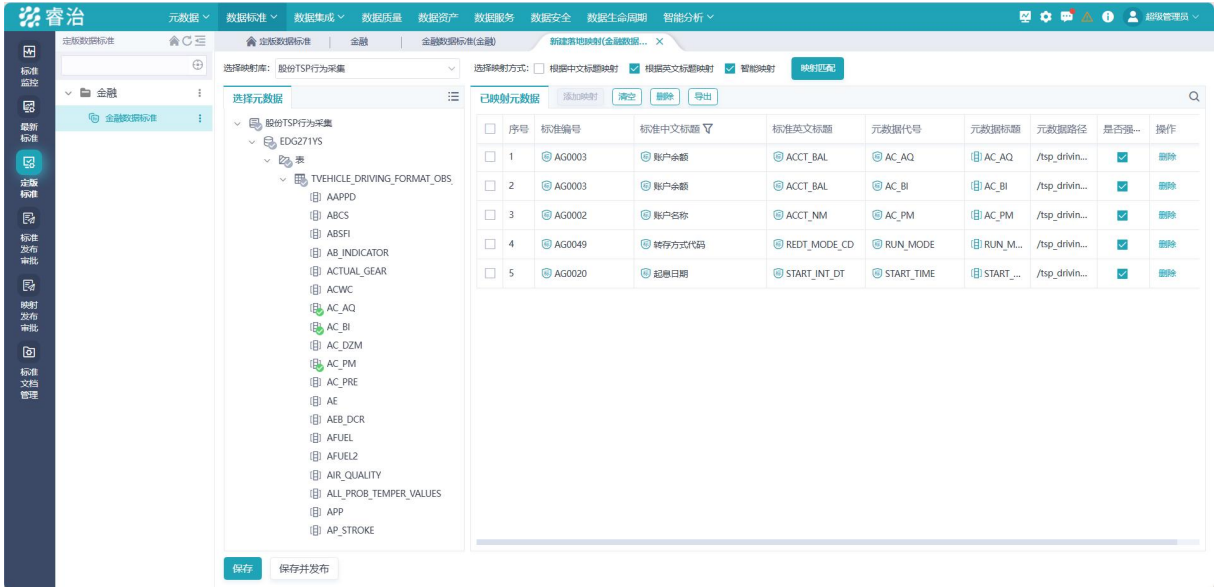
数据标准发布审批

3.2.5 智能的数据标准落地映射

数据标准的设计目的是为了规范各业务系统的数据建设。平台支持对数据标准设置落地映射，一条标准可根据实际业务需求进行多个映射，映射设置细化到实际业务系统对应的元数据上，为后续的落地评估提供依据，设置好的落地映射支持修改、删除。支持映射的手动添加和映射导入。

另外会通过内置算法推荐相应的映射关系来辅助用户快速批量的添加映射关系。

映射过程需要严谨的过程，平台还支持绑定工作流。



智能批量落地映射

3.2.6 精确的数据标准落地评估

为了方便用户检查业务系统是否按照数据标准进行建设，平台提供对数据标准进行落地评估，并支持多种方式评估，包括单条数据标准、标准集进行评估。同时支持通过数据标准和元数据双向评估标准落地情况，并且支持设置非强制性落标，设置了非强制性落标的元数据进行落标时可以例外通过，除界面手动操作评估外，睿治平台让用户通过配置定时计划任务，进行自动化评估。添加评估的计划任务，可以按照指定时间执行评估，不占用系统资源。

落地查询-标准

落地查询-元数据

数据标准集: 数据质量

数据标准:

评估属性:

查询最新评估结果

重新评估

导出

评估时间: 2019-12-10 09:21:07

例外通过

表示通过

表示不通过

数据质量	落地元数据	评估属性	数据标准值	元数据取值	评估结果
• bz1	• BBQ_	不允许为空	true	false	• 例外通过
	• CYW	值域范围	(100,1000)	20	• 例外通过
	• CYW	字符规范	字符	<空>	• 例外通过
	• CZW				

落地评估查询——标准

落地查询-标准

落地查询-元数据

元数据: 表

评估结果:

查询最新评估结果

重新评估

导出

评估时间: 2019-12-10 09:21:07

例外通过

表示通过

表示不通过

元数据名称	数据标准	评估属性	数据标准值	元数据取值	评估结果
BBQ	bz1	不允许为空	true	true	通过
CYW		值域范围	(100,1000)	50	未通过
CYW		字符规范	字符	<空>	未通过
CZW					

Copyright © 2018 ESENSOFT All Rights Reserved 北京亿信华辰软件有限公司版权所有

落地评估查询——元数据

对落地评估的数据可以进行统计查询，多维度的分析。

落地评估统计

落地评估查询

数据标准集: 第二轮

元数据:

强制标准:

查询最新评估结果

重新评估

数据标准总体通过数与占比情况

未通过

未通过: 100%

数据标准总体评估通过率

13% 通过率

数据标准分类统计

元数据总体通过数与占比情况

未通过

总数: 43

通过: 0 未通过: 43

未通过: 100%

数据标准评估统计

标准集标题	评估标准...	创建人	评估时间	可评估属性个数	评估通过个数	评估通过率
有标准集审批流...	1	admin	2022-07-30...	1	0	0%

标准集数据标准评估数与通过数

通过数

标准评估数

1

0.8

0.6

0.4

0.2

0

有标准集...

落地评估查询——数据标准

3.2.7 灵活有效的数据标准监控

数据标准监控对数据标准的 KPI 指标进行监控，包括近 3 个月数据标准发布审批情况、定版标准生效统计情况、各标准集及映射情况、标准集评估通过率、定版标准集标准变更统计情况、落地评估统计、定版标准集标准生效情况等指标。另外，监控表可根

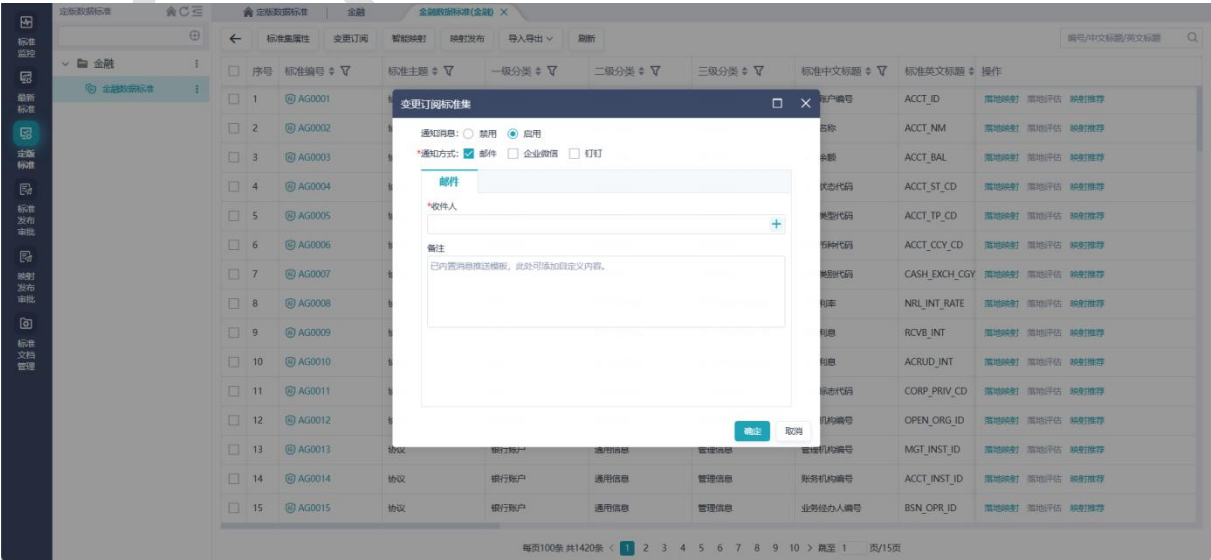
据客户需求进行二次改造。



数据标准监控

3.2.8 随时随地掌控标准变化

可以对用户关心的标准集进行订阅，当标准集发生变化重新发布时系统将及时将消息推送至短信、邮箱、企业微信、钉钉等即时通讯工具。让关注的人随时随地的知道标准的变化情况。



变更订阅

3.3 数据质量严密管控

3.3.1 多角度的规则管理

睿治平台提供数据质量规则的定义和管理，数据质量规则定义数据质量审核的业务逻辑，是数据质量审核和监控管理的基础。

规则定义提供了规则复制和规则模板的功能对规则进行多模型使用，提升了用户在规则定义过程中的便利性。

规则管理支持按评估方法进行分组，也支持用户自定义分组，并可对定义好的规则进行多角度拓扑展现，使用户对所建立的规则一目了然，协助用户建立完整的质量规则体系。

规则支持系统自动检查，也支持用户在建立过程中人工检查，通过实时执行的结果可辅助用户调整规则的定义。

1 设置规则类型

2 规则定义

3 新建成功

*代号:JavaCheck_0

*分组:规则

*实体表:

*Java代码:

更多设置

过滤条件:

问题级别:一般

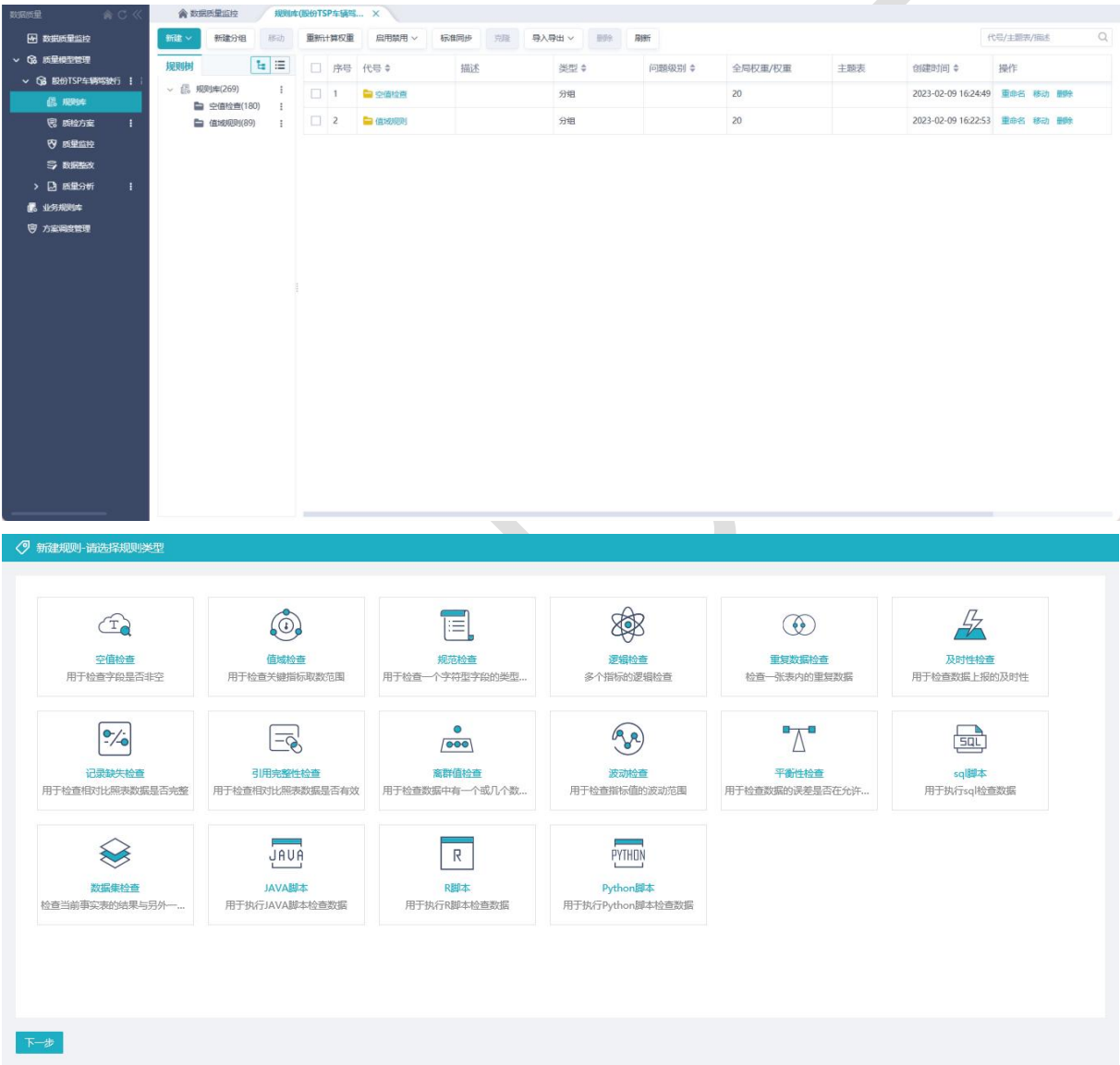
自动启用:☒

权重:10

质量检查方案规则定义

3.3.2 多种方式的新建规则

睿治平台内置 16 种规则模板，用户只需通过简单界面化的方式新建规则。平台支持单条和批量建规则，可以根据已经存在的模板新建规则，可以根据标准来新建规则，满足不同场景创建规则的需求。



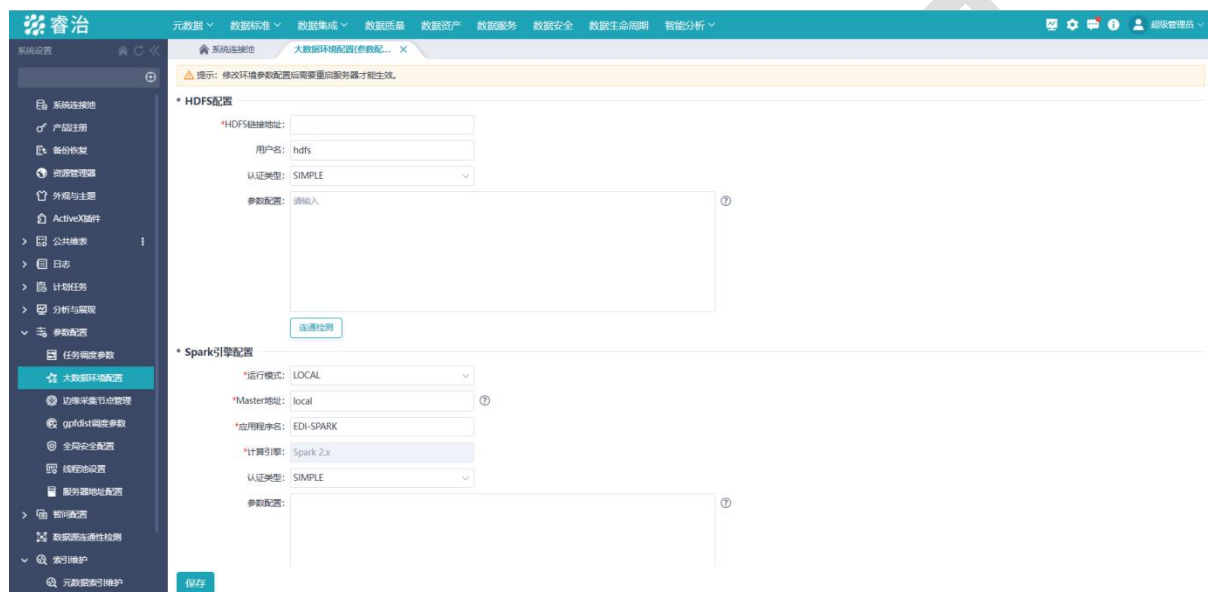
新建规则

3.3.3 海量数据快速质检与写入

支持用户配置 Spark 引擎，利用 Spark 引擎执行数据质量检查，支持 LOACL(单节

点模式)、Standalone（集群模式）、Spark on Yarn(Yarn 模式) 提升质检的效能，达到当用户接入亿级数据进行质检时，质检方案能够正常执行，且在 1 小时内执行质检完毕，实现海量数据快速质检。

当用户接入亿级数据进行质检时，质检结果能够正常存储，并正常查看质检结果，亿级数据 10s 内即可查看质检结果。同时平台能够定期清理质检结果表，释放内存、保证系统运行的效率。



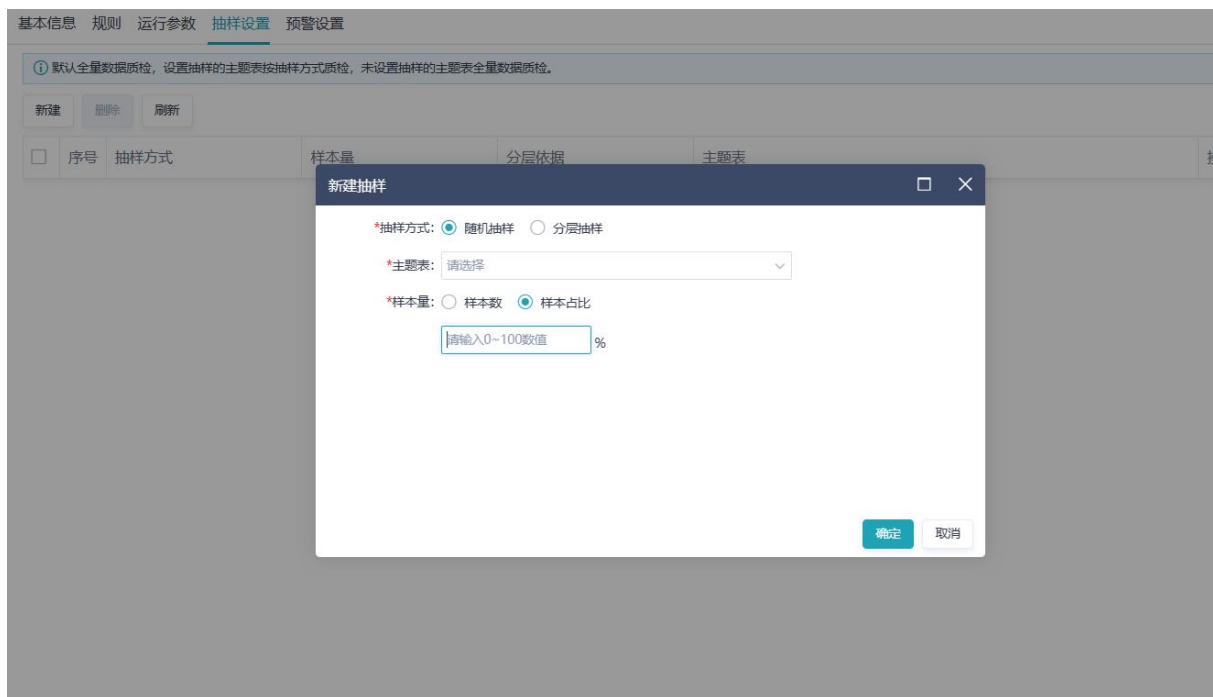
配置 Spark 引擎界面

3.3.4 灵活配置的质检方案

支持根据实际业务场景组合规则形成质检方案，同时支持设置定时、周期执行质检方案，无需人工执行，且执行结果支持预警，当执行出现问题或数据质量低下时，消息将及时推送至短信、邮箱、企业微信、钉钉等即时通讯工具，实时监控保证执行的效率。

支持全量质检或抽样质检，当用户数据量达到亿级时，可设置抽样质检提升质检的效率，支持随机抽样与分层抽样。

同时支持对各种对象进行质检，不仅支持对库表中数据质检，且支持对对象存储服务中的数据质检，例如华为云的 obs。



设置抽样质检

3.3.5 强大的智能修复

睿治平台支持对全量质检或抽样质检检查出来有问题的数据按照一定的规则进行智能修复，可以支持对空值、值域、规范（身份证、日期、全半角）这些规则进行修复，方便用户对数据的整改。

新建规则-空值检查

*代号：NULL_38

*分组：规则

*主题表：

*检查字段设置：

查看库表

查看数据

数字画像

联合不为空：☐ 勾选表示所有的字段不能同时为空，不勾选表示每个字段都不能为空

描述：

更多设置

关联关系：

过滤条件：***

数据显示：***

问题级别：一般

权重：10

公式展示

公式展示

状态：☒ 启用 ☐ 禁用

修复设置

修复策略：无

无

平均值

最大频次值

指定值

重新创建

修复策略编辑

3.3.6 全面的监控管理

睿治平台支持数据质量检查方案的定义和管理，包括检查范围、检查时间、检查规则、评分规则、评估报表等。同时，方案支持人工调度和自动调度。

系统提供了完整性评价、规范性评价、逻辑性评价、及时性评价、重复性评价、外键关联性、波动分析、平衡分析等多种评估方法。各评估方法均采用可视化界面，用户无需编程，即可轻松完成所有规则的建立。



数据质量监控

3.3.7 详尽的结果管理

通过质量监控，会产生和保存质量结果，包括：对象名称、发生时间、违反规则、级别、状态等。系统会自动生成每个质检方案的明细结果表，并允许用户根据分析需要对明细结果表字段进行自定义，从而为用户进行丰富多样的数据质量分析提供数据。

同时当质检的数据量达到亿级影响执行效率与查看结果时，用户能够自行设置保留错误数据量，设置导出错误数据量，保证执行效率与快速查看明细结果。且系统会根据执行的情况清理质检结果表用以释放内存，保持高效率质检。

ee-质量监控						
质检方案: 444	数据源: 无	数据级次: 根节点	主数据: 全部	更多 规则代码		
序号	规则代码	规则描述	数据量	错误合计	例外合计	修复合计
1	NULL_1		3	0	1	0

数据质量结果详情

3.3.8 多样的统计分析

根据数据质量管理及监控需要，对问题数据进行统计分析，系统内置了多种形式的
问题数据分析功能、统计报表功能及数据质量分析报告。

为了满足用户深层次数据质量分析需要，我们支持用户将基于亿信 BI 等第三方业
务系统分析报表地址配置到数据质量系统进行展示。



数据质量趋势分析

3.3.9细致自主的数据整改

睿治平台数据质量管理模块参照 workflow 管理联盟 (WfMC) 标准设计, 用户可自行定义整改流程, 可实现问题分发给数据责任人, 并可实时监督流程处理的状态。同时可根据流程的优先级、最终期限、转派次数等提供多角度查询功能。



数据整改管理

3.4 批流一体化数据处理

3.4.1低门槛的离线数据整合处理

平台提供多种类型离线数据处理组件, 支持用户在数据开发时, 能够从多源异构的数据源中便捷高效的采集、解析、处理数据, 支持按照数据流转逻辑并行或串行调度离线任务, 满足用户在多种大数据场景下, 数据处理速度快和效率高的需求。

同时支持用户配置 Spark 引擎, 利用 Spark 引擎进行计算, 支持 LOACL(单节点模式)、Standalone (集群模式)、Spark on Yarn(Yarn 模式) 提升数据处理效能, 达到当用户接入亿级数据时, 任务可以正常执行。

3.4.2 低延时的流式数据处理

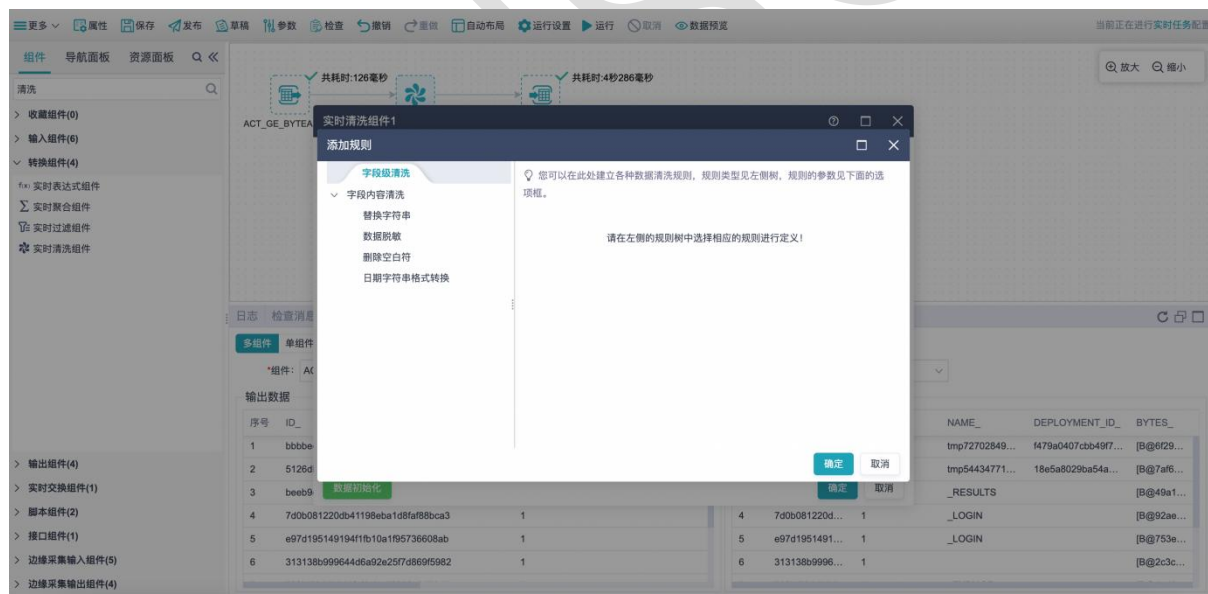
平台内置了实时转换组件、实时脚本组件、实时接口组件、实时表达式组件、实时聚合组件、实时过滤组件、实时清洗组件、实时 python 脚本、实时 Groovy 脚本、实时 CDC 日志增量组件等大量实时处理数据组件，为实时分析场景提供数据支持。

支持基于 Kafka、Spark 技术的实时数据传输处理，可实现结构化、非结构化数据的实时处理入库。

实时数据预览，查看各组件数据在流式里经过处理后的变化。

统一的监控中心，实时监控数据处理流程，统计问题数据处理状况。

可视化预警实时数据问题，系统将问题消息及时推送至邮箱、企业微信、钉钉等即时通讯工具，降低数据安全风险；辅助实时数仓搭建，满足企业实时数据分析需要。

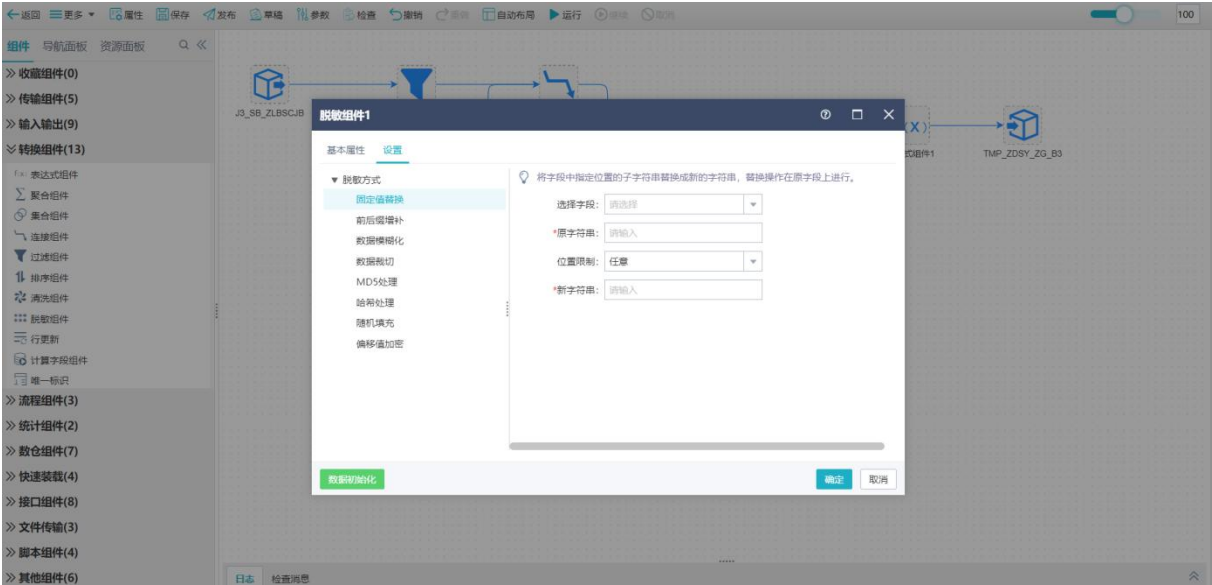


实时清洗组件

3.4.3 便捷的拖拽式流程设计器

全拖拽式流程设计器，用户只需在简单手填数据资源和加工组件就可完成复杂的 ETL 作业和作业流程定义，“零”编码，易操作、易阅读、易维护。

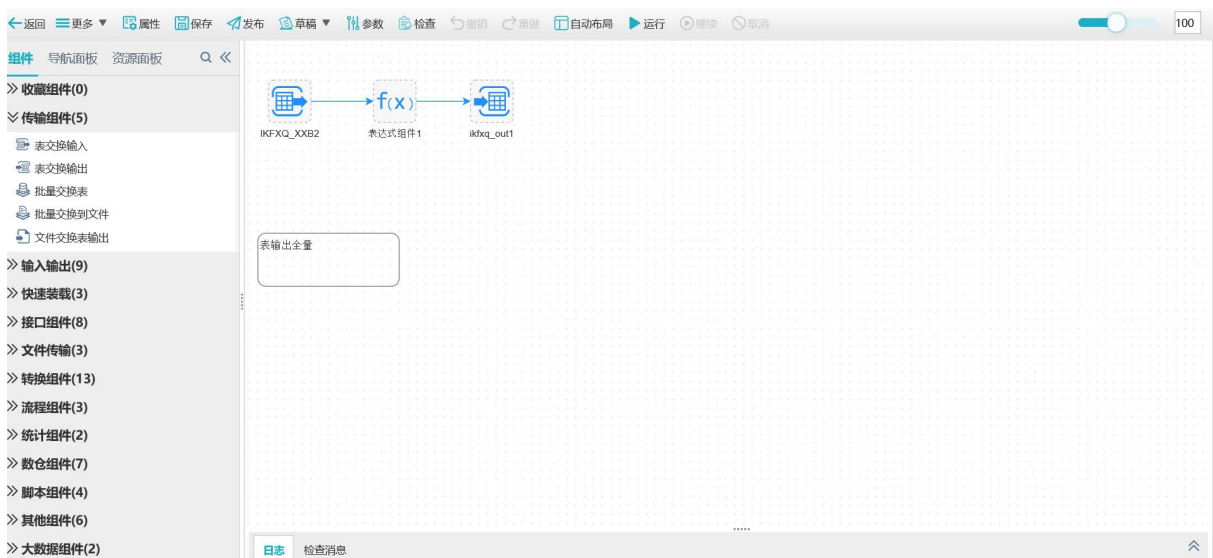
完整图形编辑功能支持，如复制、粘贴、撤销、重做、自动对齐等。



“零”编码设计 ETL

3.4.4 丰富的数据处理组件

数据整合提供了丰富的数据处理组件，并将组件进行合理的分类管理，如：收藏组件、传输组件、输入输出、快速装载、快速卸载、接口组件、文件传输、转换组件、流程组件、统计组件、数仓组件、脚本组件、其他组件、大数据组件等，用于快速完成数据的传输、清洗转换、装载落地等处理过程。同时，平台提供详细的在线帮助手册和案例库，让用户很轻易就能上手使用。



数据处理设计界面

3.4.5 “零”表达式清洗组件

数据整合封装了近 30 种清洗规则，用于完成数据的字段级和记录级清洗转换，省去了复杂的表达式定义，同时清洗组件支持边配置边预览清洗前的数据变化，支持多条规则的自动全关执行，节省性能开销。



添加清洗规则

3.4.6 多样的增量捕获和数据装载

数据整合提供了基于时间戳、MD5、触发器、全表比较等多种方式的非侵入式变化数据捕获机制，对用户透明，只需要简单设置即可完成增量数据获取，提供了数据覆盖、数据追加、数据更新和更新插入等多种数据的落地装载策略，全面覆盖数据落地场景。

3.4.7 可视化的调试和预装载

支持流程调试，如顺序执行、执行到功能，可以查看每步的执行状态、执行时长和执行结果集，同时面向开发人员设置了预装载机制和数据抽样加载，方便开发人员快速验证流程和脚调试。

3.4.8 灵活的设计与运行隔离机制

产品内置两个资源库，分为设计区和运行区，两个资源库相互隔离，可保证运行环境的稳定和可靠，所有作业的修订和更改通过在设计区完成调试和试运行后，通过发布机制，发布到稳定的运行区，在生产环境中可区分模拟运行和正式上线环境，减小系统运行风险。

3.4.9 多方位的运行监控

集中监控服务中心的数据传输流程，方便管理和运维人员掌控全网运行及使用状态，及时发现、分析并解决问题，帮助客户实现数据链路统一管理。监控内容如下：

1. 支持对系统和数据的操作行为审计
2. 利用实时监控模块实现实时数据传输流程的全方位监控，保障传输流程的稳定性并实现异常问题可追溯、可定位。

- 3. 支持变更检测，当任务中输入输出端等结构发生变化时系统能定期检测感知，并发送信息（短信、邮件、站内信、企业微信、钉钉）给平台管理员
- 4. 支持数据源连通性检测，定时检测平台所有数据源是否正常连通。



监控中心

3.5 多源异构数据共享交换

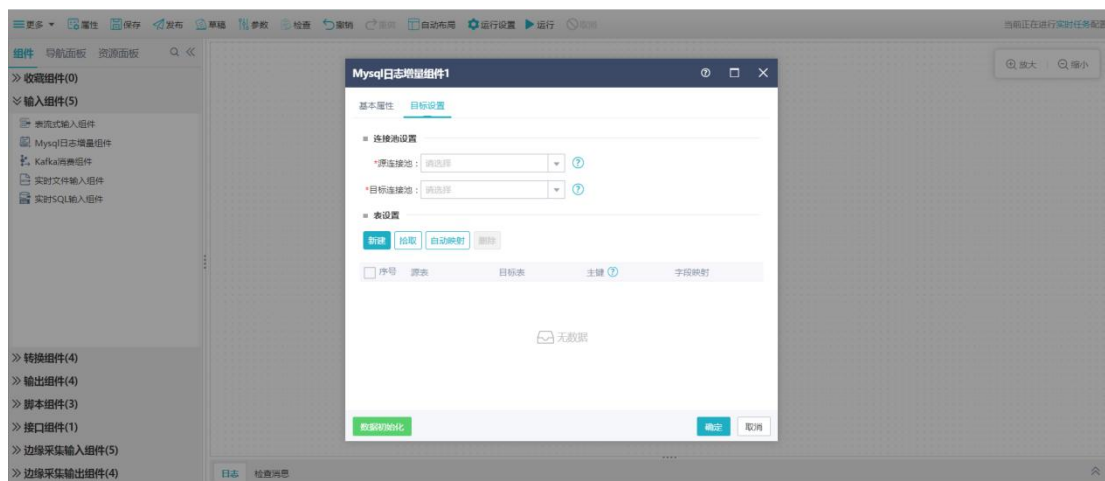
数据传输交换是将分散建设的若干业务系统进行整合，通过平台，可实现若干个业务子系统之间进行数据或者文件的传输和共享，提高信息资源的利用率，保证了分布在异构系统之间的信息的互联互通，完成数据的收集、集中、处理、分发、加载、传输，构造统一的数据及文件的传输交换。

睿治平台支持与各业务子系统直接通过接口进行数据交换，同时可针对不同业务系统的特性，拓展交换接口，从而保证了数据交换的多样性和便捷性。

3.5.1 多源的数据采集

平台提供了输入组件、输出组件、边缘采集输入组件、边缘采集输出组件等多

种采集数据组件，支持对各种数据的采集，并且还可以根据用户配置的边缘节点程序以及采集源的具体信息，采集服务器上产生的日志数据。多样化的组件支撑平台对各种数据源的全方位采集。



Mysql 日志增量组件

3.5.2 跨源数据交换

平台支持多个系统之间的数据交换，在数据源中添加需要进行数据或文件传输的业务系统，并可对这些系统进行编辑、删除等管理操作。

3.5.3 多类型数据交换

平台提供文件到文件、文件到库表、库表到库表、库表到文件等方式的数据交换，还支持批量交换，同时支持对数据进行处理，再对处理后的数据进行交换。

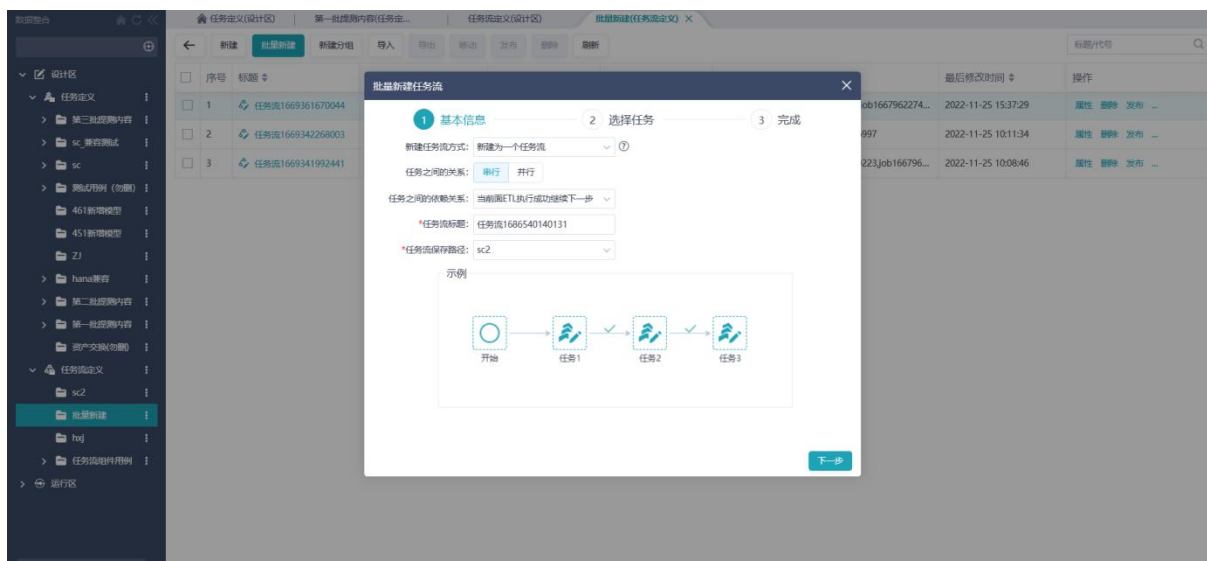


传输组件

3.5.4 灵活的任务调度

任务调度主要用于配置作业的计划执行时间，作业只有在配置的调度时间达到后才会出发作业的执行，在规定开始时间没能启动作业执行时，会根据配置的重调时间间隔和重调次数来进行作业的重新启动，在操作时间窗口和重调次数时，会进行后续的作业运行和处理，用户可以根据需求灵活设置任务执行时间。

同时支持批量配置任务调度、任务流调度，仅需选择需要配置的任务及任务间的依赖关系，即可完成任务流的创建，当用户交换任务过多时，无需重复拖拉拽即可快高效的完成创建。

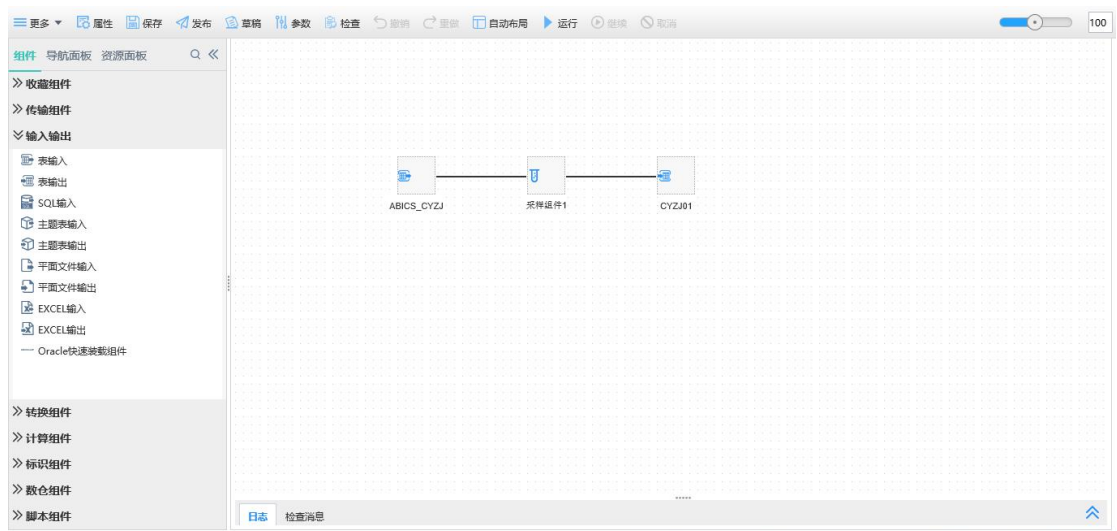


批量新建任务流界面

3.5.5 丰富的组件库

睿治平台内置近百种数据处理组件，支持通过托拽的方式快速构建数据处理任务，提供输入输出组件：表输入、表输出、SQL 输入、主题表输入、主题表输出、平面文件输入、平面文件输出、文件生成组件、Excel 输入、Excel 输出；转换组件：表达式组件、聚合组件、集合组件、连接组件、过滤组件、排序组件、路由组件、分支组件等；计算组件、标识组件、数仓组件、脚本组件等近百种数据处理组件。

平台丰富的组件不仅支持对自有系统中的数据进行处理，还可对交换后的数据进行 ETL 和质量检查，从而保证全面的数据处理和质量把控能力。



ETL 过程编辑

3.6 一站式数据资产运营

睿治数据治理平台为用户提供完整的资产视图，管理者在平台上可概览企业资产，通过合理的方式管理内部数据和提供对外服务。

3.6.1 灵活的数据资产目录定义

平台提供了统一的目录属性库管理，通过添加自定义的目录属性和信息项属性实现对目录模板进行灵活的配置。为后续的数据资源资产和数据产品资产目录提供支持。让用户实现多场景的资产目录配置。

属性库管理											
新建自定义属性 删除 刷新											
▼ 目录属性 +											
基础属性											
权属属性											
共享开放属性											
拓展属性											
信息项属性											
序号 属性代码 属性名称 属性类型 数据类型 数据长度 数据精度 允许为空 属性描述 关联维表 默认值 操作											
1	ASSETDIGEST	数据资产摘要	基础属性	大字段	1000	0	否	数据资产摘要			编辑
2	TYPE	数据资产格式...	基础属性	字符串	100	0	否	数据资产格式分类			编辑
3	FORMAT	数据资产格式	基础属性	字符串	100	0	否	数据资产格式			编辑
4	OTHERTYPE	其他类型数据...	基础属性	字符串	100	0	是	如果默认的类型不能...			编辑
5	AAA	数据资产===	基础属性	字符串	200	0	是				编辑 删除
6	SSS	数据资产	基础属性	字符串	200	0	是				编辑 删除
7	PROVIDER	数据资产提供方	权属属性	字符串	100	0	否	数据资产提供方机构...			编辑
8	SHARETYPE	共享类型	共享开放属性	字符串	50	0	是		共享类型		编辑

属性库管理

3.6.2 全场景的数据资产盘点

平台提供智能添加和手动添加的方式实现企业全场景的数据资产盘点。智能添加可通过对平台元数据实现快速智能编目可帮助企业极大的缩短数据资产编目的时间和提高效率。手动添加可针对丰富的资产类型进行添加，其中包括：数据库、电子文件、图形图像、流媒体、电子表格、其他。实现资产的全覆盖。

数据资产的编目管理，是为了更好地支撑各种数据的应用。睿治平台支持丰富的服务接口拓展，从而实现所管理资产的多渠道应用，实现数据资产价值的最大化释放。

数据资产编目-手动添加

数据资产编目-智能添加

数据资产		数据资产统计	资产编目(数据来源)	手动添加数据资产
元数据 资产打标2		切换版本		
手动添加		智能添加	目录活化配置	申请发布
类别		目录活化配置	退回申请	导入导出
类别一(1)		目录活化配置	列表管理	删除
类别二(6)		目录活化配置	刷新	资产代号/资产标题
类别三(6)		目录活化配置		
序号	数据资产代号	数据资产标题	审批状态	上架状态
1	0001/0015	test1	已通过	已上架
2	0001/0005	zj	草稿	未上架
3	0001/0006	assets	草稿	未上架
4	0001/0004	5-技术方案选型评审指南.docx	草稿	未上架
5	0001/0001	ACT_ID_GROUP	草稿	未上架
6	0001/0002	A	草稿	未上架
		数据资产提供方	最后修改时间	操作
		test	2023-09-13 14:23:57	编辑 申请变更 打标 删除
			2022-11-16 14:19:17	编辑 申请发布 打标 删除
			2022-11-16 14:19:17	编辑 申请发布 打标 删除
			2022-11-16 14:02:08	编辑 申请发布 打标 删除
			2022-11-16 14:01:17	编辑 申请发布 打标 删除
			2022-11-16 14:01:17	编辑 申请发布 打标 删除

数据资产编目-列表

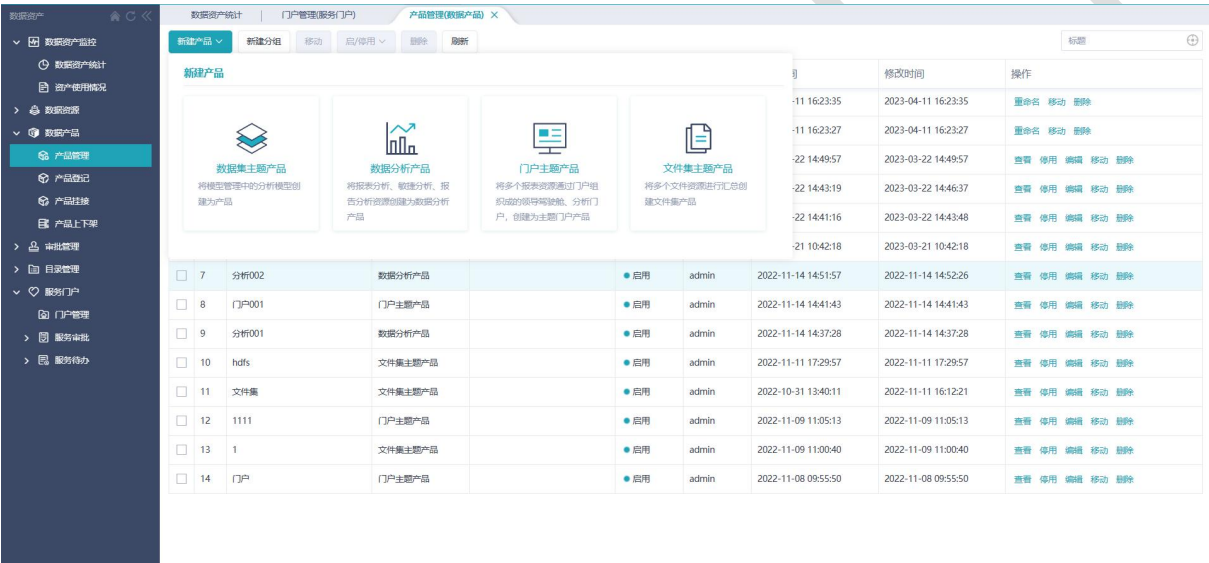
3.6.3 智能的目录活化管理

平台通过可配置的资产目录完整描述数据资源的业务、技术和管理类元信息。资产目录与元数据无缝对接，支持通过元数据批量盘点数据资产，并自动活化资产目录信息，及时响应资源的变更情况。同时系统将及时把活化消息推送至短信、邮箱、企业微信、钉钉等即时通讯工具。

目录活化配置						
刷新						
序号	元数据挂载点代码	元数据挂载点名称	未盘点元数据	待活化数据资产	已活化数据资产	通知方式
1	aaaa	aaaa	37719	0	0	立即同步 编辑 活化日志
2	bigdata	bigdata	40008	0	801	立即同步 编辑 活化日志
3	ftp	ftp	0	1	16	站内信,邮箱 立即同步 编辑 活化日志
4	oracle	oracle	37644	0	0	立即同步 编辑 活化日志
5	oracle7w	oracle7w	79618	0	0	立即同步 编辑 活化日志
6	Oracle	Oracle采集01	5839	0	0	立即同步 编辑 活化日志
7	TEST	TEST	21	0	0	站内信,邮箱 立即同步 编辑 活化日志
8	yyxt_bi	yyxt_bi	3048	0	0	立即同步 编辑 活化日志
9	yyxt_i	yyxt_i	279601	0	0	立即同步 编辑 活化日志
10	zntj	zntj	59	0	0	立即同步 编辑 活化日志
11	xdyw1	信贷业务数据1	572	0	0	立即同步 编辑 活化日志
12	emate_sp	元数据审批	2396	0	1	立即同步 编辑 活化日志
13	jkyria	接口元数据a	2655	0	0	立即同步 编辑 活化日志

3.6.4 丰富的数据资产类型

平台不仅支持多种数据库、文件等结构化和非结构化物理资源资产的定义、盘点和规划。还能够将富含业务知识的分析模型、报表应用、主题文件集合、接口等逻辑资源封装为数据产品，一键挂接，全方位的盘点企业的各类数据生产要素。



支持创建多种类型的产品

3.6.5 直观的数据资产监控

平台提供了健全直观的数据资产监控，通过选择属性版本来查看不同的数据资产统计和资产使用情况，帮助企业快速全面了解全盘资产数据和过往资产使用情况，为其决策性判断提供可靠直观的依据。



数据统计情况



资产使用情况

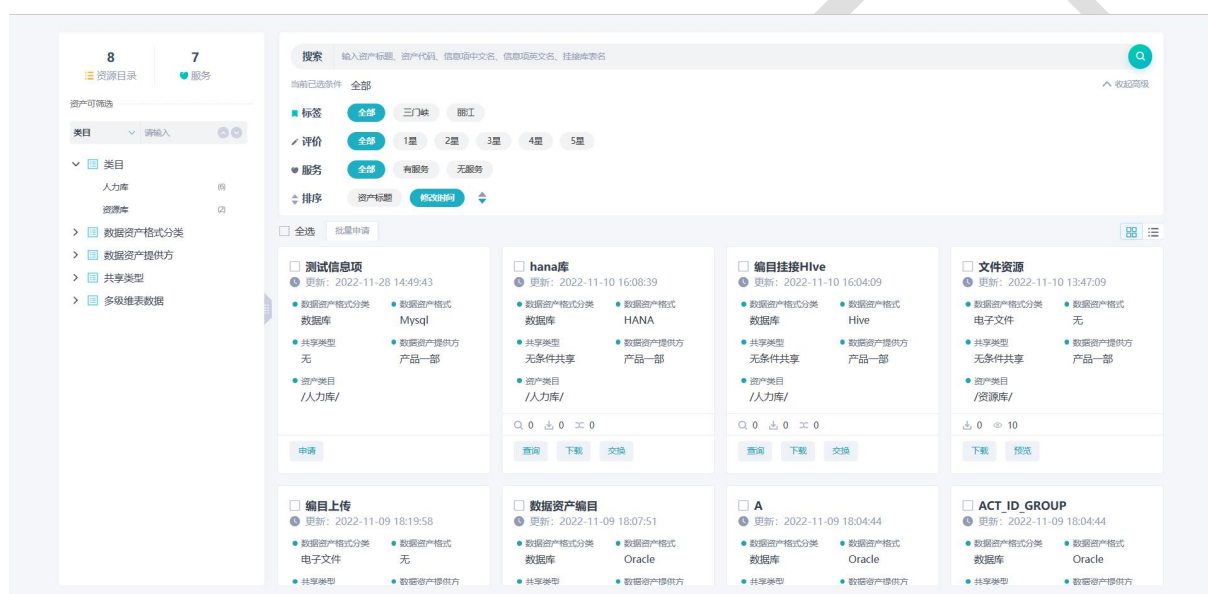
3.6.6 可视化数据资产服务门户

平台盘点的数据资产能够以数据商品的形式在数据资产门户中对业务用户开放，消费者在门户中可以浏览数据资产并且查看详情，对自己关注对资产可发起数据服务的使

用申请。

同时平台提供强大的检索功能，可根据各种维度指标如：资产分类导航、资产标签、评价等级、挂接库表名、关键字检索的方式对平台中各种类型的资产进行检索，以列表的形式显示检索出来的资产，实现数据资产的快速查找定位。

同时门户具有精细化的权限管控。通过简单配置即可实现不同权限的用户查看门户资产时，查看到自己所控制的资产。



数据资产展示与检索



数据资产内置门户界面

3.6.7 多样的数据资产服务方式

平台中管理的数据资源可以快捷开发出多种数据服务类型，包括：数据查询、文件预览、下载、共享交换、API 和敏捷分析服务，并保证消费者符合数据权限管理要求，还能提供分析产品授权使用、基于分析模型报表开发、验证接口调用、主题文件批量下载等高附加值服务。

服务申请

发改委行政许可信息

*申请类型：

在线查询

文件下载

共享交换

敏捷分析

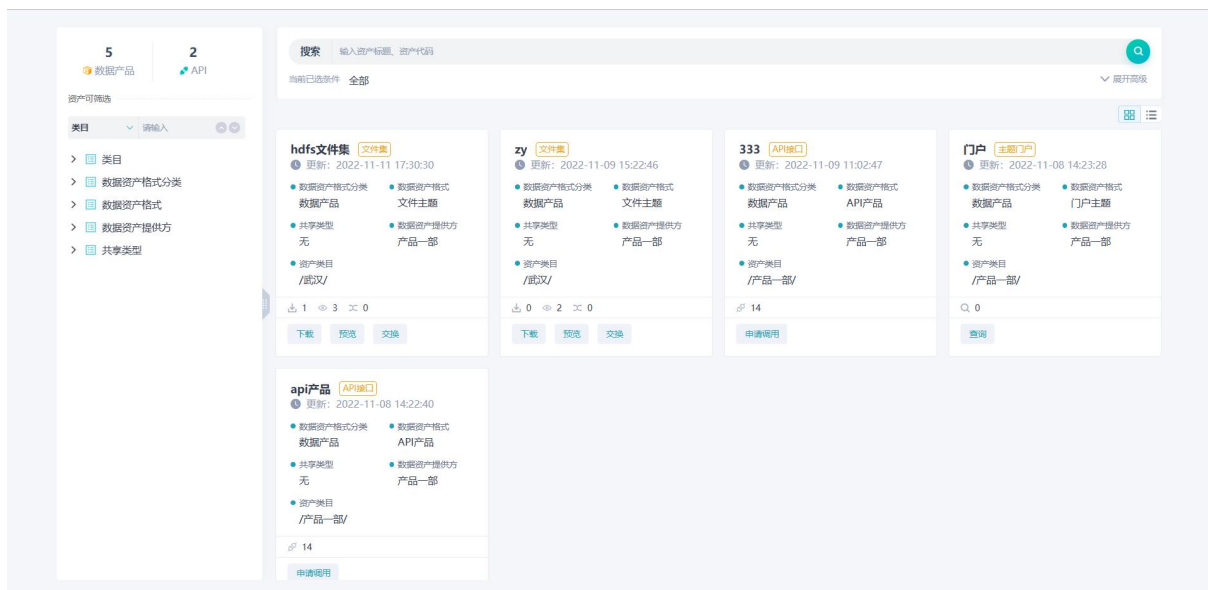
*申请说明：

高级：

禁用

确定

取消



数据产品服务

支持用户根据资产共享情况与企业业务要求，自定义配置服务申请的条件。同时系统将及时把申请情况推送至短信、邮箱、企业微信、钉钉等即时通讯工具。

服务申请

Hive库表数据

*申请类型: ☒ 在线查询 ☐ 文件下载 ☒ 共享交换 ☐ 敏捷分析

*申请说明:

文件型: 界面自查表+v1.0.xlsx

用户型:

维表多选:

*服务格式: ☒ 库表 ☐ 文件

*目标数据源:

*数据库表名: ☐ 追加 ☒ 覆盖

数据处理方式: ☒ 原样数据交换 ☐ 特殊数据处理要求

外部表: ☐

*HDFS地址:

自定义申请条件

3.6.8 多应用的个人中心

平台的用户个人中心包含我的资源、我的收藏、我的分享等功能模块，支持用户对分析结果进行保存方便下次查询，同时也可以将分析结果分享给其它人查看，实现资源

的共享。支持用户对感兴趣的资产进行收藏，实时监控收藏资产的变更情况，随时随地了解资产的变化，同时将数据消费者关心的数据资产、数据服务集合在一起，方便用户使用。支持用户了解各个数据资产服务的状态，对于已经过期和即将过期的服务，可以及时进行服务的申请。多个数据资产的应用场景，让数据资产的价值得到了充分的利用。



序号	数据资产标题	数据资产代码	服务类型	服务来源	资产类型	服务到期时间	服务可用状态	操作
1	加密脱敏	0009/0197	数据交换	我的申请	数据资源	2023-08-12 18:04:21	● 可用	执行交换 日志 收藏 ...
2	zj	0000/0004	明细查询	我的申请	数据资源	2023-08-12 17:26:29	● 可用	明细查询 日志 收藏 ...
3	orcl	0000/0115	数据交换	我的申请	数据资源	2023-08-12 17:07:12	● 可用	执行交换 日志 收藏 ...
4	1	0001/00	授权报表查询	我的申请	数据产品	2023-08-12 14:56:00	● 可用	查看分析表 日志 收藏 ...
5	数据分析产品1-5	0000/0020	授权报表查询	我的申请	数据产品	2023-08-12 14:55:05	● 可用	查看分析表 日志 收藏 ...
6	门户主题产品	0000/0029	授权门户查询	我的申请	数据产品	2023-08-12 14:27:33	● 可用	查看门户 日志 收藏 ...
7	门户主题产品	0000/0029	授权门户查询	我的申请	数据产品	2023-08-12 14:26:04	⊗ 不可用	查看门户 日志 收藏 ...
8	数据产品资产111	0000/0018	授权门户查询	我的申请	数据产品	2023-08-12 14:24:18	● 可用	查看门户 日志 收藏 ...
9	门户主题产品1	0000/0007	授权门户查询	我的申请	数据产品	2023-08-12 14:23:59	● 可用	查看门户 日志 收藏 ...
10	资产1-3	0000/04	ABI分析	我的申请	数据产品	2023-08-10 16:58:12	⊗ 不可用	创建分析表 日志 收藏 ...
11	资产监控数据-1	0000/0000	授权报表查询	我的申请	数据产品	2023-08-08 17:03:16	⊗ 不可用	查看分析表 日志 收藏 ...
12	条件	0000/0179	明细查询	我的申请	数据资源	2023-08-08 10:25:48	⊗ 不可用	明细查询 日志 收藏 ...

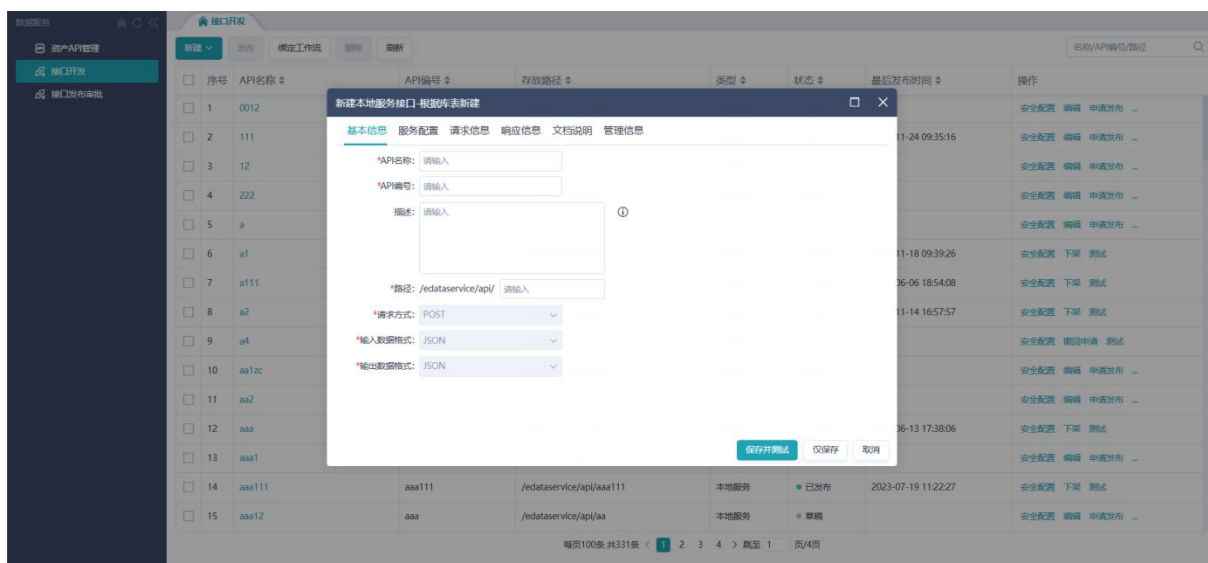
我的服务

3.6.9 便捷的 API 服务开发

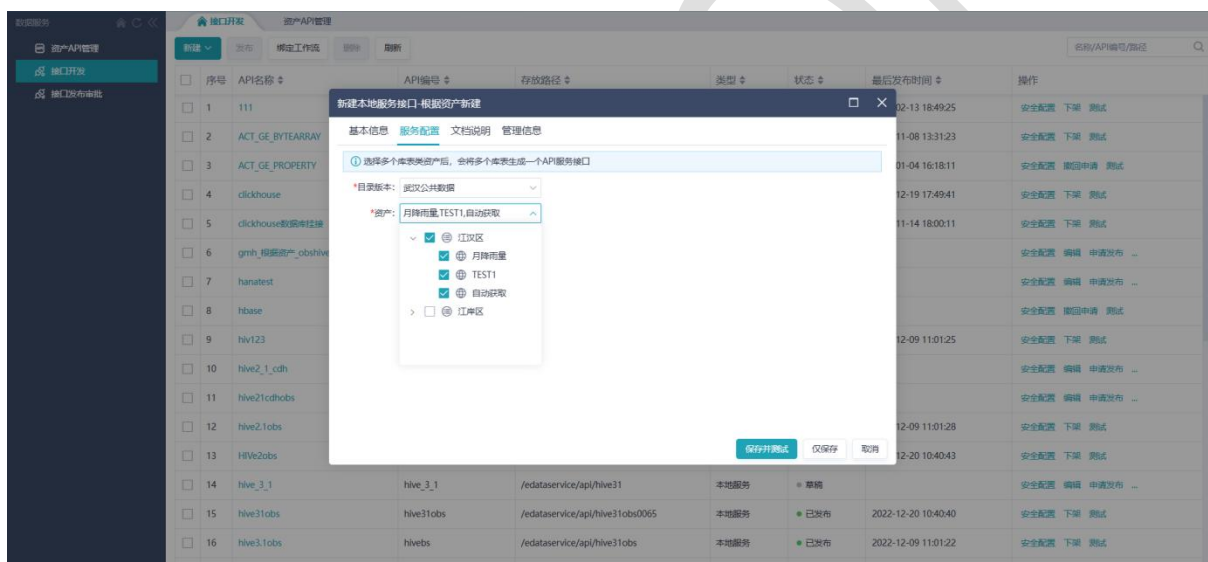
API 服务是对数据进行计算逻辑的封装，包括：数据过滤、多维汇总和关联查询等。平台提供可视化配置方式，实现零编码的 API 服务开发并自动生成规范化的接口调用说明。支持开发本地接口、配置第三方接口，上层应用可以直接对接 API 服务，让数据快速应用到业务场景中。

支持对数据资产直接生成 API 接口对外提供服务，支持单资产单接口与多资产单接口，简单配置即可快速满足用户使用接口的需求。

支持对全局的接口进行安全配置，保障接口使用中的安全。



新建服务接口



根据资产生成接口

3.7 数据生命周期监控

平台支持数据资产全生命周期管理，根据存储周期自动计算每行数据的存储时限，并根据存储时限进行数据自动归档、销毁，为业务系统节省数据空间，提高数据的查询效率。

3.7.1便捷的数据资产归档

平台提供对数据资产进行归档设置功能，可根据数据的产生时间、数据量设置对应的归档策略进行数据近线或离线归档，同时支持批量归档，从而保证业务系统数据的简洁。同时，近线区和离线区支持用户自定义。同时系统将及时将归档消息推送至短信、邮箱、企业微信、钉钉等即时通讯工具。

数据生命周期

数据生命周期监控

归档设置管理(数据归档)

新建

新建分组

删除

移动

刷新

分给标题/数据表

+

☐	序号	归档设置标题	数据表	备注	近线条件	离线条件	创建人	创建时间	操作
☐	16	AJD近线区归档设置	lxqddjgd		最近1年之前的数据		admin	2023-01-31 10:05:37	编辑 删除 移动 详细
☐	17	ALX近线区设置	LXQDJDGD1			最近3条之前的数据	admin	2023-01-03 15:36:52	编辑 删除 移动 详细
☐	18	hana	A2		最近1月之前的数据		admin	2022-11-28 17:03:01	编辑 删除 移动 详细
☐	19	hana_zhangc	A55		最近1月之前的数据	最近1月之前的数据	admin	2023-03-10 11:17:24	编辑 删除 移动 详细
☐	20	Oracle归档Zbase	A5		LEFT(ZCLX,3)='120'		admin	2023-08-17 17:06:45	编辑 删除 移动 详细
☐	21	oracle数据到pg	ACT_GE_BYTEARRAY		ID_='b2fc252e324d...		admin	2023-03-28 11:19:35	编辑 删除 移动 详细
☐	22	SJGD	test		最近12月之前的数据		admin	2023-07-27 16:58:48	编辑 删除 移动 详细
☐	23	xinghuan	a5		最近10条之前的数据		admin	2023-06-29 16:18:29	编辑 删除 移动 详细
☐	24	zcaa	A55_1		最近1月之前的数据	最近1月之前的数据	admin	2023-03-03 14:23:24	编辑 删除 移动 详细
☐	25	Zettabase	DS081811		最近10条之前的数据		admin	2023-08-18 13:59:44	编辑 删除 移动 详细
☐	26	zhangc_click	a5		最近10条之前的数据	最近1月之前的数据	admin	2023-03-10 18:02:33	编辑 删除 移动 详细
☐	27	zz	A55		HY='N'	HY='N'	admin	2023-03-14 15:25:46	编辑 删除 移动 详细
☐	28	zz2	A55		最近1条之前的数据	最近1条之前的数据	admin	2023-03-16 14:14:46	编辑 删除 移动 详细
☐	29	执行进度问题	XX0_MY		最近3月之前的数据		admin	2023-01-13 17:47:33	编辑 删除 移动 详细
☐	30	执行进度问题2	XX0_MY_XXTS		最近1条之前的数据		admin	2023-01-13 17:52:50	编辑 删除 移动 详细
☐	31	迁移	A55_OUT		最近1月之前的数据		admin	2023-08-17 15:48:23	编辑 删除 移动 详细

数据归档

数据生命周期

数据生命周期监控

归档设置管理(数据归档)

新建归档设置

基本信息

* 归档设置标题: 请输入归档设置标题!

* 数据源: 请选择

* 归档数据表: 查看数据

描述: 请输入归档信息!

近线归档

状态: ☒ 启用

归档模式: ☒ 简单归档 ☐ 高级归档

* 依据时间字段: 请选择

* 归档策略: 按时间

* 数据保留时间: 请输入 月

* 近线区: 请选择

* 归档库表: 设置字段映射

高级归档

状态: ☒ 启用

归档模式: ☒ 简单归档 ☐ 高级归档

* 依据时间字段: 请选择

保存 取消

归档设置

3.7.2 灵活的数据资产销毁

平台提供数据销毁操作，用户可根据数据的使用情况进行永久销毁或临时销毁，销毁操作提供审批操作，只有通过审批的数据才能被销毁，临时销毁的数据存放在回收站内，可一键还原。

数据生命周期

数据资产

数据资产管理

数据任务设置

数据回收站

数据销毁

数据设置

执行销毁

销毁审批

回收站

销毁日志

数据生命周期监控

数据设置管理(数据销毁)

新建任务设置

销毁设置(数据销毁)

新建

批量发起

删除

刷新

数据表

Q

☐	序号	销毁设置标题	数据表 #	归档连接池	归档库表	销毁策略	销毁条件	销毁方式	操作
☐	17	kingbaseV8R2高级...	GMH_RV2	DEFAULT_SYS_KIN...	GMHARCHIVE4	自定义销毁条件	销毁SYS_ORIGINAL_="manual L...	回收站	发起销毁 编辑 删除 详细日志
☐	18	kingbaseV8R2高级...	GMH_RV3	DEFAULT_SYS_KIN...	GMHARCHIVE5	按记录数	销毁10条之前的数据	永久删除	发起销毁 编辑 删除 详细日志
☐	19	opengauss2.1简单...	gmh_ry2	DEFAULT_SYS_GAU...	gmharchive1	自定义销毁条件	销毁DATETOSTR(sys_createdat...	永久删除	发起销毁 编辑 删除 详细日志
☐	20	opengauss2.1高级...	xxb_my_1ww761	DEFAULT_SYS_GAU...	gmharchive2	按记录数	销毁5条之前的数据	回收站	发起销毁 编辑 删除 详细日志
☐	21	oracle数据到pg	ACT_GE_BYTEARRAY	DEFAULT_SYS_ABI...	gdb	自定义销毁条件	销毁ID_="b2fc25e324d4e5482...	回收站	发起销毁 编辑 删除 详细日志
☐	22	vertica动态加载	QC_G_762Q38	DEFAULT_SYS_VER...	QCIG_GD	自定义销毁条件	销毁true	回收站	发起销毁 编辑 删除 详细日志
☐	23	zbase_ff1	A4	DEFAULT_SYS_FF	AAA	按记录数	销毁1条之前的数据	回收站	发起销毁 编辑 删除 详细日志
☐	24	zbase_ff2	ABICS_CVZJ	DEFAULT_SYS_FF	BDACC	自定义销毁条件	销毁LEFT(UPIID,3)="200"	回收站	发起销毁 编辑 删除 详细日志
☐	25	zbase_ff3	XXB_DAY_GUID	DEFAULT_SYS_FF	XXB	自定义销毁条件	销毁LEFT(BBQ,6)="20070902"	回收站	发起销毁 编辑 删除 详细日志
☐	26	zbase_ff4	ASS_OUT	DEFAULT_SYS_SL	OOO	按时间	销毁1月之前的数据	回收站	发起销毁 编辑 删除 详细日志
☐	27	Zettabase	D5081811	DEFAULT_SYS_ZET...	ds081811gd	按记录数	销毁8条之前的数据	回收站	发起销毁 编辑 删除 详细日志
☐	28	依格简单1-简单销毁	GMH_RVU717	DEFAULT_SYS_LXD...	gmharchive1	按记录数	销毁5条之前的数据	回收站	发起销毁 编辑 删除 详细日志
☐	29	依格高级2-高级销毁	asdada2	DEFAULT_SYS_LXD...	gmharchive2	自定义销毁条件	销毁e5 <"1998-10-05 00:00:00..."	回收站	发起销毁 编辑 删除 详细日志
☐	30	发播啊	A2	DEFAULT_SYS_HA...	A2_2	按时间	销毁5月之前的数据	回收站	发起销毁 编辑 删除 详细日志
☐	31	简单销毁2	AS2	zj_SIZL	guidang2	按记录数	销毁1条之前的数据	回收站	发起销毁 编辑 删除 详细日志
☐	32	简单销毁3	a53	DEFAULT_SYS_PG	guidang3	按记录数	销毁1条之前的数据	回收站	发起销毁 编辑 删除 详细日志

数据销毁设置

数据生命周期

数据资产

数据资产管理

数据任务设置

数据回收站

数据销毁

数据设置

执行销毁

销毁审批

回收站

销毁日志

数据生命周期监控

销毁审批(数据销毁)

执行销毁(数据销毁)

销毁设置(数据销毁)

待审核

已审核

批量通过

批量退回

刷新

数据表

Q

☐	序号	销毁设置标题	数据表 #	归档连接池	归档库表	销毁条件	销毁方式	审批状态	数据量	操作
☐	1	FYE_XH_7.3	qjgdyb	DEFAULT_SYS_LUC...	qjgdyb	销毁1月之前的数据	回收站	待审批	17	通过 退回

数据销毁审批

3.7.3 可视化的数据生命周期监控

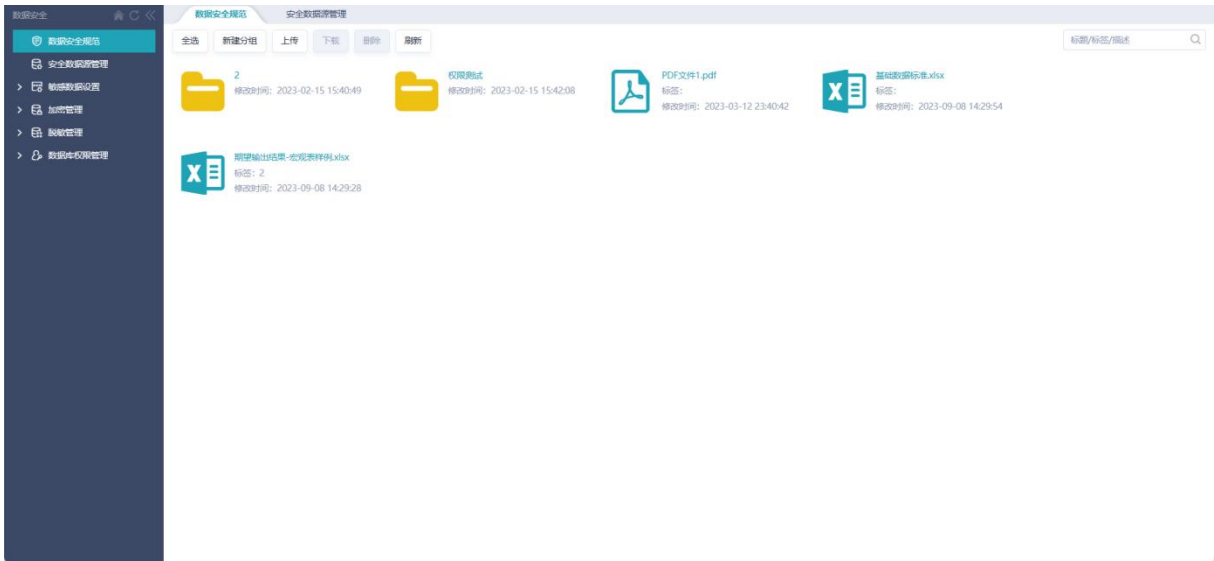
用户可通过数据生命周期监控了解企业系统中在线、近线、离线、销毁的数据量各有多少，可通过不同归档期限统计归档的数据量占比，同时提供归档预警，了解哪些表数据量较大，可能需要归档等统计信息，数据生命周期监控表还提供用户自定义报表，可根据自己的实际业务改造监控表。



3.8 严谨的数据安全管控

3.8.1 多类型的安全规范管理

在数据安全管理过程中往往会形成一些规范性的文档，睿治平台提供对这些文档进行集中的管理、存储和查阅。支持多种文档格式，包括：文本、word、PPT、excel等多种文档的上传，并支持在线预览 PDF 和 DOCX 文档。



数据安全规范管理

3.8.2全方位的数据资产加密

睿治支持对平台自身数据资产进行加密，目前提供：AES-128、AES-192、AES-256、DES、3DES、RC4、RC5 等多种加密方法，同时支持对密钥进行管理，保证了加密的多样性和完整性。

查看资产【机构信息表】

序号	机构ID	地址	主办单位代码	注册资金（万元）	行政区划	机构名称	机构分类管理代码	政府办卫生机构隶属关--	经济类型代码
1	022135485CDZAZB	SM4(b5689c9673ad307e...	1	503	SM4(60f269b686d3e178...	DES(873de8aca3011353...	1	3	11
2	746053418DAAFBF	SM4(1a979ec229028a69...	1	102	SM4(9d3c6161958f2116...	DES(e939110e040505b4...	1	5	11
3	3277ZX417DAZGBA	SM4(7b38cc055035dad8...	1	68	SM4(7449f8aa71c6f658a...	DES(eea0a769939838be...	1	7	11
4	681367420D8ZGZD	SM4(d4578f9facd1d3f85...	2	11	SM4(058454bd5a1c5fcc...	DES(318b67f4cc5272a2...	1		11
5	68260X420D8ZGZD	SM4(0fefb9693654da4...	1	12	SM4(058454bd5a1c5fcc...	DES(41e7d45d928e9f56...	1	7	11
6	501097G72EBBGA	SM4(639aa6954fd5eb92...	2	588	SM4(2fbf5ee18956b7c6...	DES(976aa3580da6d85b...	1		11
7	280049430EBBGB	SM4(5bc48ac8fab35e4d...	1	106	SM4(0ac27a9f8ea4b789...	DES(dcbf2a79a37afe420...	1	4	11
8	890025466C8ZCBD	SM4(151bdea809d0a287...	1	430	SM4(4facb2efbbed9da1f...	DES(cd117aa6be7e4422...	1	7	12
9	114411414BCZDZG	SM4(078f18dcd93978f...	5	88	SM4(0bbb9a139455631...	DES(8a0e9d6d2b48d0b8...	2		13
10	10007XPDYCDZAZD	SM4(d80395da663a2d7f...	2	212	SM4(4785f62b7387f103...	DES(2d764841c9884052...	1		11
11	016984400CDZAAA	SM4(47e9b495fa57a2e7...	3	21	SM4(4f4aac2c6f7b3a1c3...	DES(29c03f97ee079b9b...	1		12
12	000087PDYCCAZBD	SM4(4ecf203bcd710792...	1	13	SM4(d18c8d690937037a...	DES(d43fb54824e416f1...	1	8	12
13	000132PDYCCAZBD	SM4(1a255d7bf100115...	1	9	SM4(d18c8d690937037a...	DES(d43fb54824e416f1...	1	8	12
14	02769X485CDZAAA	SM4(e913c9274165de8a...	1	28	SM4(4f4aac2c6f7b3a1c3...	DES(873de8aca3011353...	1	4	12
15	000204PDYCCAZBD	SM4(c86792870c60b1ff...	1	9	SM4(d18c8d690937037a...	DES(d43fb54824e416f1...	1	8	12

首页 上一页 第 1 页/508页 下一页 末页 每页100条 共50716条

加密数据查看

3.8.3 多场景的数据资产脱敏

支持灵活的对平台数据设置脱敏，除此之外还可以设置脱敏条件，以及脱敏对象，从而实现权限控制，保证脱敏效果能满足多种场景需求。

查看资产【机构财务报表】

序号	净资产	机构ID	检查收入	总收入	总资产	治疗收入	总负债	其他支出	人员支出	总支出
1	*****	507806428DFIZZB	*****	7726.77	1029.83	580.71	311.89	91	478.32	1801.45
2	*****	569059E37AEBHZA	*****	135861.02	114471.49	8174.46	60546	154.88	33387.14	202858.43
3	*****	654638452EACCBF	*****	31871.02	114.02	153.61	7.96	49	107.49	132.47
4	*****	540243452EAAHZB	***	32088.25	675.25	653	26.18	10	458.01	938.47
5	*****	883156427CGZBHC	***	5375.34	163.34	729.74	90.79	56	53.2	486.35
6	*****	000386PDYFEDBZB	*****	23306.63	148.59	24.87	61	42	28.1	89.08
7	*****	380758492CFAACZ	*****	62914.48	42553.48	1131.64	17446.94	864.31	10219.44	53498.26
8	*****	061073457DDAGBA	*****	3670.05	645.05	225.11	394.13	0	230.4	891.82
9	*****	618351739FB8DBF	*****	26134.85	283.85	934.44	90.01	39	162.61	417
10	*****	063626457DDAGZB	***	18477.01	231.98	11	93.25	100	177.36	385
11	*****	141122467CBZEHA	*****	29474.53	1099.52	692.84	333.78	16.09	277.31	1225.65
12	*****	017967732CCZCZB	*****	16529.49	447.44	641.1	88.84	47.93	277.7	324.57
13	****	857660400AAZAF	*****	8346.5	1286.51	121.03	72	11	320.29	165.5
14	*****	646108450EZZCHC	*****	13114.14	646.14	706.09	571.48	39	211.61	714.32
15	*****	13353YA93CE7HHC	***	33016.13	100.76	376	74.11	62	174.76	307.66

页码 上一页 第 1 页/254页 下一页 末页 每页100条 共25359条

脱敏数据查看

3.8.4 严密的权限管理

除此之外，睿治提供系统数据库权限管理功能：

- 数据库权限管理可以统一管理数据库权限，数据库用户新建、数据库权限分配，统一在平台上操作。为防止用户私下在数据库中新增改用户和权限，睿治提供数据库权限校核功能，检查平台上数据库用户和权限与实际数据库用户和权限不一致的情况。

3.9 平台系统管理

平台为上层模块提供一系列具有高复用率的功能，并为各系统的开发提供接口。

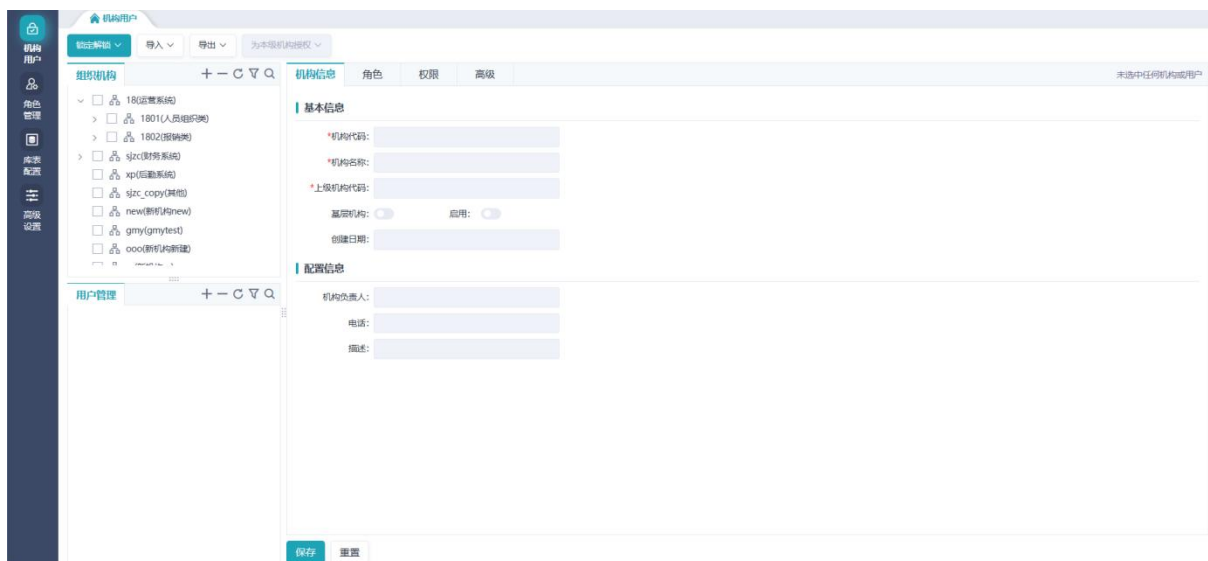
3.9.1 个性化的门户定义

平台提供门户设计器，内置丰富的模板供用户选择或由用户自定义门户样式，用户可以将需要展现的资源集中在门户中。用户只需通过简单的拖拉控件到定义门户区域，就能自己定义个性化的门户，能够简单实现各种复杂的自定义用户视图，无需开发人员以及美工介入。



3.9.2 灵活细致的权限管理

平台提供灵活而细致的用户权限管理机制，既可配置系统功能权限，也可配置数据标准、元数据、数据质量、数据集成、数据资产、数据交换、数据生命周期、数据安全等各模块的资源权限。既可为单个授权，也可批量授权；既可手动授权，也可自动授权。



用户权限

3.9.3 简便的备份恢复

平台提供多级别的备份机制，可对数据标准、元数据、数据质量、数据集成、数据资产、数据交换、数据生命周期、数据安全等各模块进行任意组合备份，同时提供整个服务器备份。备份工作可以人工出发也可以由系统通过定时任务自动完成。

对于备份完成的数据，平台支持异构数据库间的备份与恢复，这对于系统升级、迁移、快速故障恢复等具有特别重要的意义。

3.9.4 详细的系统日志记录

平台提供完备的系统操作日志，记录操作者的 IP、操作类型、操作资源、日志详情等信息，并提供日志查询筛选功能，以后后续对日志进行分析统计。

系统设置

系统连接池

公共库表

系统日志(日志)

系统日志

系统日志

日志设置

详情设置

SLF4J设置

导出

刷新

时间	登录名称	操作	操作耗时	资源ID	资源名称	IP	描述	详情
2023-09-14 13:46:48	超级管理员	发起销毁设置	2秒139毫秒	EDLC\$411\$e196919...	gigdyb	172.21.170.1	发起销毁设置	发起销毁设置成功,销毁表名为...
2023-09-14 13:46:39	超级管理员	查看销毁设置列表	2毫秒	EDLC\$41\$destroyco...	销毁设置	172.21.170.1	查看销毁设置列表	查看销毁设置列表
2023-09-14 13:46:33	超级管理员	查看执行销毁列表	2毫秒	EDLC\$42\$executede...	执行销毁	172.21.170.1	查看执行销毁列表	查看执行销毁列表
2023-09-14 13:46:29	超级管理员	查看数据生命周期监控	2毫秒	EDLC\$2\$datalfecycl...	生命周期监控	172.21.170.1	查看生命周期监控界面	查看生命周期监控界面
2023-09-14 13:46:29	超级管理员	查看销毁审批列表	6毫秒	EDLC\$43\$destroyap...	销毁审批	172.21.170.1	查看销毁审批列表	查看销毁审批列表
2023-09-14 13:46:10	超级管理员	查看连接池	20毫秒	EDSO\$33\$DEFAULT_...	连接池管理	172.21.2.93	查看连接池	查看连接池
2023-09-14 13:46:02	超级管理员	查看数据源系统	1毫秒	EDSO\$0\$server	数据源系统	172.21.2.93	查看系统	查看系统
2023-09-14 13:46:00	超级管理员	查看销毁审批列表	2毫秒	EDLC\$43\$destroyap...	销毁审批	172.21.170.1	查看销毁审批列表	查看销毁审批列表
2023-09-14 13:45:43	超级管理员	查看销毁设置列表	2毫秒	EDLC\$41\$destroyco...	销毁设置	172.21.170.1	查看销毁设置列表	查看销毁设置列表
2023-09-14 13:45:02	超级管理员	查看归档设置列表	2毫秒	EDLC\$31\$archiveco...	归档设置	172.21.170.1	查看归档设置列表	查看归档设置列表
2023-09-14 13:44:55	超级管理员	查看数据生命周期监控	2毫秒	EDLC\$2\$datalfecycl...	生命周期监控	172.21.170.1	查看生命周期监控界面	查看生命周期监控界面
2023-09-14 13:41:24	超级管理员	查看元数据监控	2毫秒	EMET\$81\$monitord	元数据监控	172.21.170.1	查看元数据监控	查看元数据监控
2023-09-14 13:33:17	超级管理员	查看最新数据标准库	0毫秒	EDST\$32\$eec4b94b...	CsBug3	172.21.2.93	查看标准库变化列表	查看标准库【CsBug3】修改的...
2023-09-14 13:33:09	超级管理员	查看最新数据标准库	7毫秒	EDST\$32\$eec4b94b...	CsBug3	172.21.2.93	查看标准库【CsBug3...】	查看标准库【CsBug3】下所有...

每页100条,共446975条 < 1 2 3 4 5 6 7 8 9 10 > 跳至 1 页/4470页

系统日志

3.9.5 统一的调度管理

为方便统一调度管理，实现任务自动化执行，平台提供简单易用的计划任务设置，保证任务的准时性，减少人员工作量。计划任务包括：外部可执行程序、调用其他计划任务、备份服务器、执行 SQL、定时发起工作流实例、计算报表等多种类型的任务。

计划任务向导

×

💡 向导会帮助你制定整个计划任务的流程，请选择计划任务类型

☒ 外部可执行程序

☐ 调用其他计划任务

☐ 备份服务器

☐ 执行sql

☐ 定时发起工作流实例

☐ 计算报表

下一步

取消

计划任务类型选择

计划任务的设置，包括任务类型的选择，任务参数的设置，以及任务的调度执行时间设置。系统默认执行周期为每日 00:00，用户也可根据自身实际场景需求，自定义执行周期，可以按年、月、周，乃至表达式等多种维度进行设置。同时可以设置生效、失效日期，重试次数、时间间隔等，从而保证调度能够完全按设想执行，大大减少人工干预，保证高效准确执行。

计划任务向导

×

💡 请输入该计划任务执行的时间

启用：☒

*名称：

*执行周期：

自定义周期

☐ 每年 ☐ 每月 ☐ 每周 ☒ 每天 ☐ 自定义

每年的第： 月 每周星期：

每月的第： 天 每天的：

请输入cron表达式：

测试 帮助

生效日期：

📅

 失效日期：

📅

重试：☐ （任务执行失败后，进行重试的设置）

重试次数： 间隔时间： 分钟

上一步

完成

取消

新建计划任务执行时间

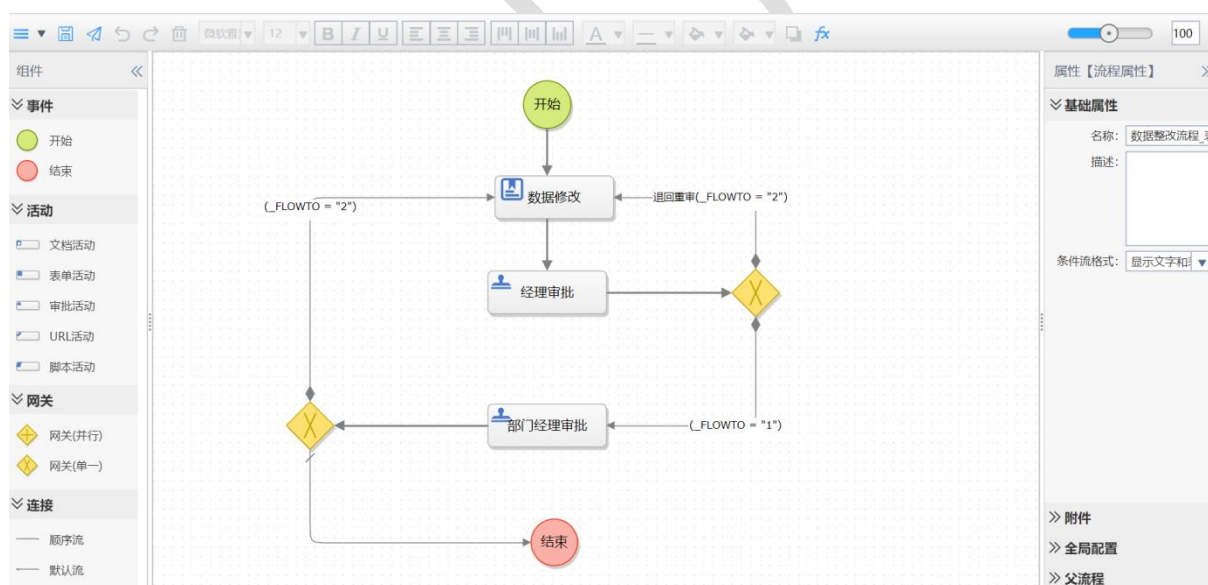
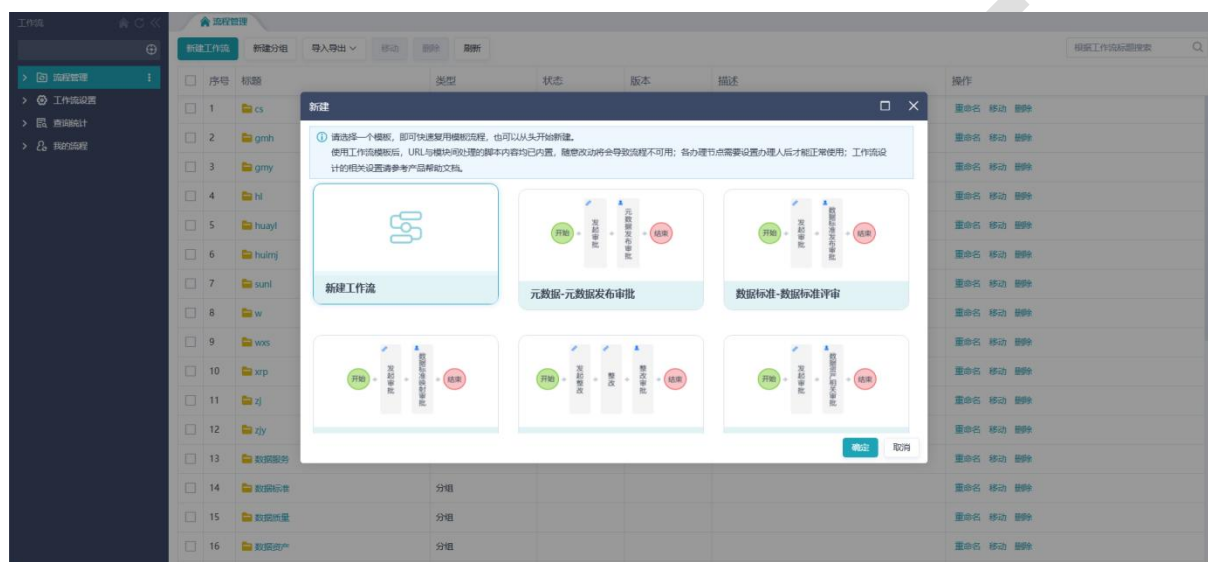
为方便管理员统一管理、了解全局调度情况，睿治提供了调度日志，方便管理员查看和查询日志的执行情况。

3.9.6 直观的工作流管理

睿治作为一款综合数据治理平台，需全面满足不同角色用户的工作需求，因此平台提供了简单好用的流程管理功能。用户可自行新建工作流，通过简单的拖拽操作，就可

把系统内置的各种组件绘制成一个个直观易懂的工作流。同时工作流管理支持流程的导入、导出及发布，并可在我的流程中查看自己的工作流。

同时内置系统常用工作流模板，用户无需编写脚本，简单根据业务要求设置审批人后即可使用工作流。



工作流绘制

3.9.7 便捷的睿治工作台

睿治平台提供了便捷的睿治工作台入口，界面化展现整个数据治理过程与成果，用户可以一目了然的清楚数据治理各模块的成效，可以对数据治理整体效果进行综合评估。

同时提供各种快捷入口，一键跳转到目标界面，如代办事项，直接点击跳转到办理界面，完成代办事项的快速处理。

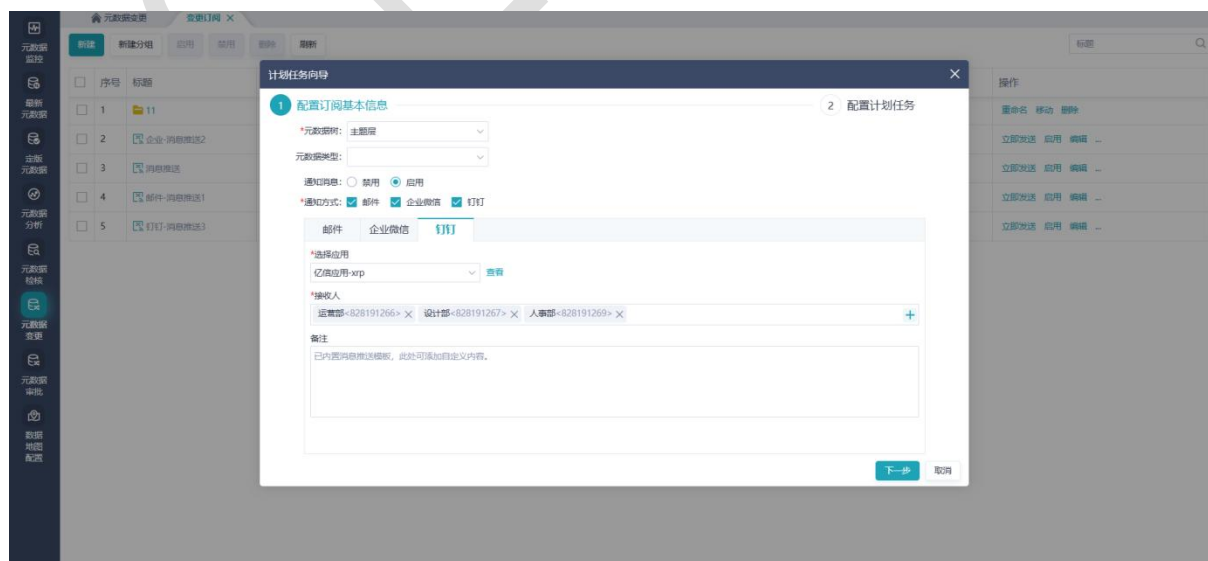


睿治工作台

=

3.9.8 企业信息集成

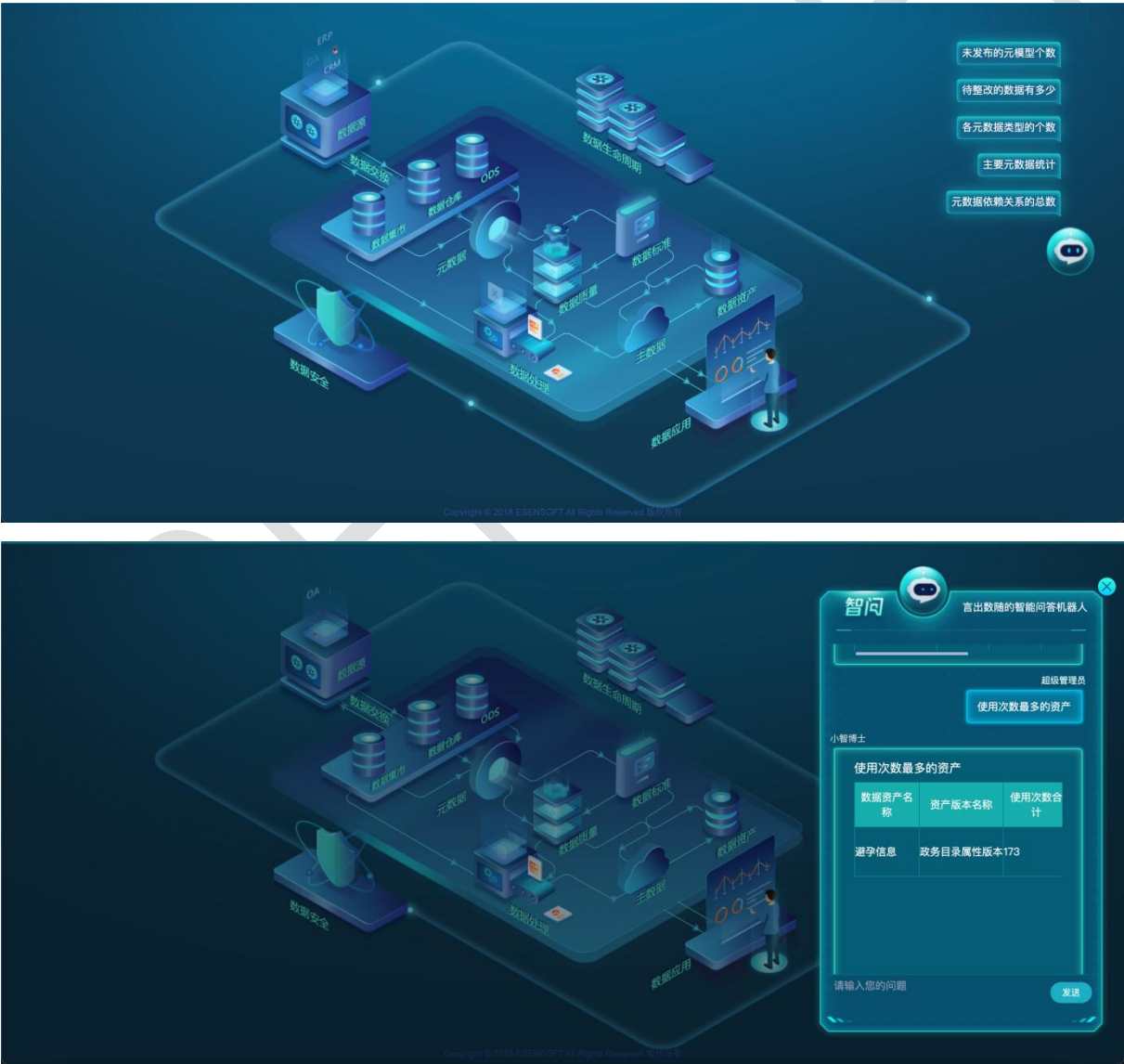
平台集成了企业微信、钉钉等实时通讯平台，支持通过企业微信、钉钉、短信、邮件等方式，仅需简单配置即可将消息及时推送给用户进行数据监控、预警通知等，保证时效性。



3.10言出数随的小智博士

3.10.1 言出数随，智能问答

支持与小智博士进行问答交互，用户采用聊天对话的方式查询数据、分析数据。问题一经提出，图表即刻呈现。



智能助手

3.10.2 理解语义，智能消歧

基于自然语言处理技术，包括文本分词、语义理解、语义消歧等，针对用户提出的问题，能够智能实现自动消歧或引导用户手动消歧，从而探究深层次的数据洞见。



智能消歧

3.10.3 智能绘图

小智博士能够根据问题以及数据类型自动判断最合适的展示图表类型。并支持表格、文字卡、柱状图、条形图、饼图、折线图（曲线图）等大量图表类型，图表类型会根据问题的形式和数据格式动态选择，同时也支持在问题中指定图表类型。



柱状图



线状图



条形图



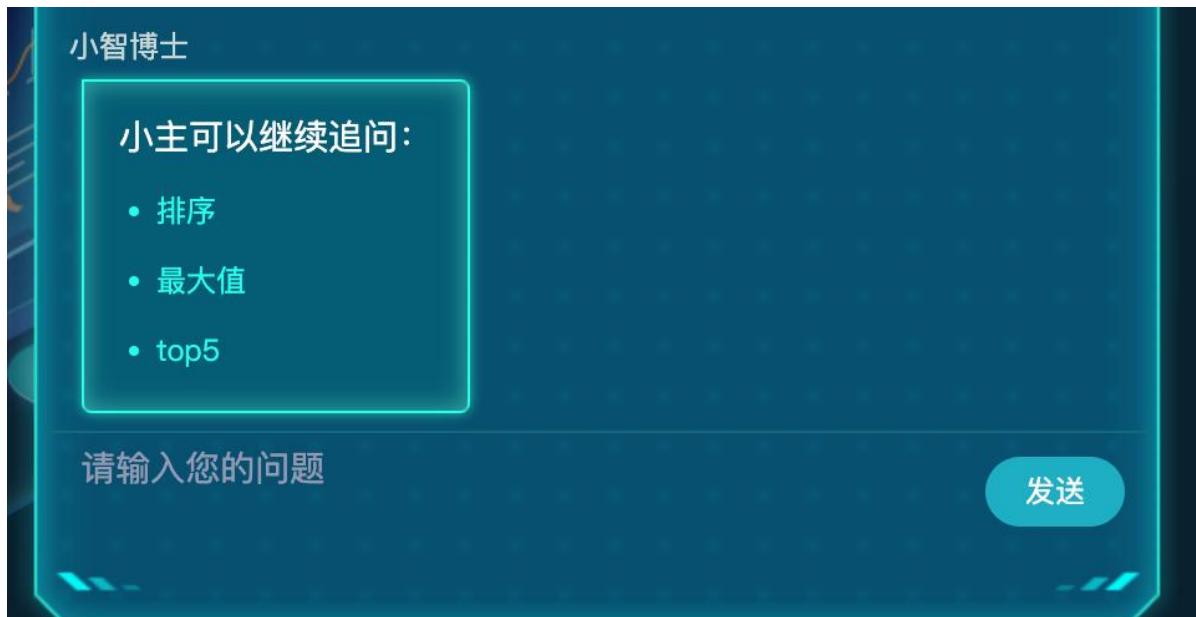
饼图



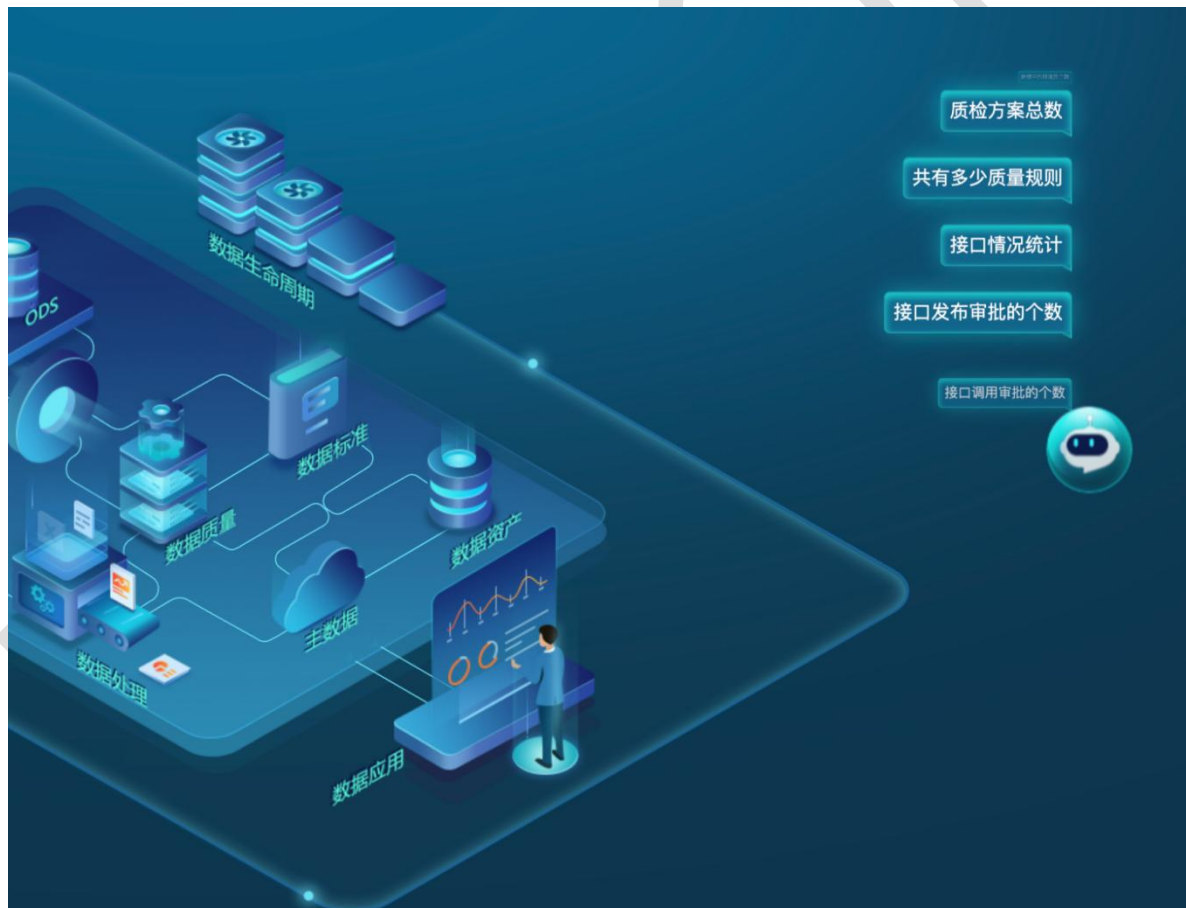
表格

3.10.4 知你所思，智能推荐

小智博士包含了指标和维度推荐、关键词推荐、热门问题推荐、个性化推荐、相关问题推荐、不同问法推荐等多种多样的问题推荐形式，全方位覆盖用户的搜索意图，懂用户所想，知用户所思。



智能推荐



推荐问题

3.10.5 智能学习，知识沉淀

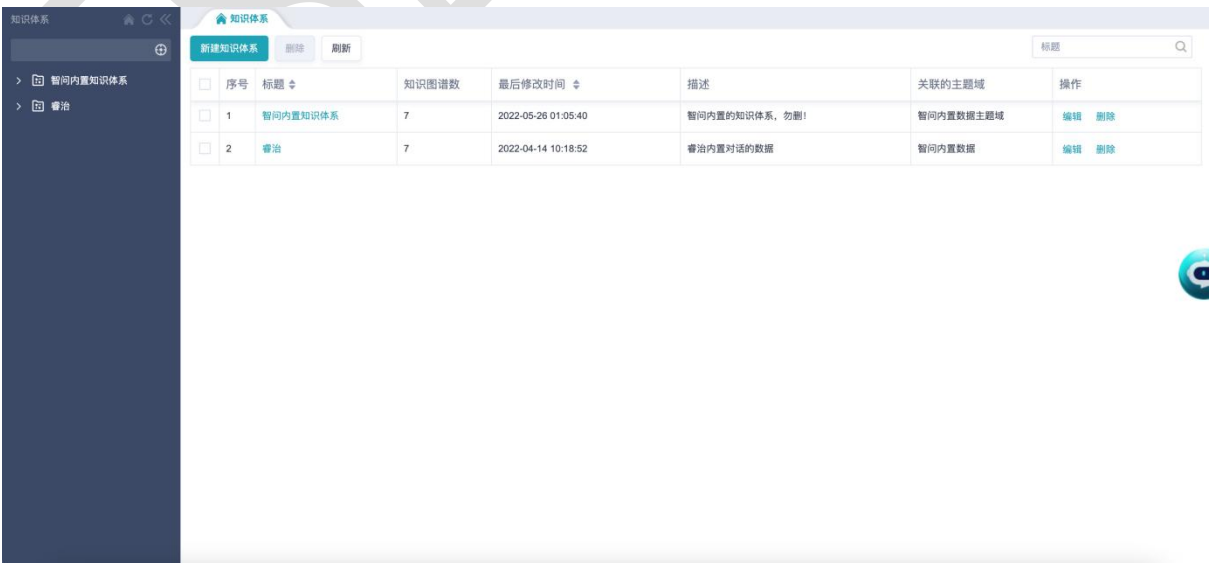
小智博士会自动记录用户在系统中的行为，包括搜索的语句、相关的指标维度等多种信息，并基于记录到的行为进行分析，了解每个用户的分析习惯，将行为进行沉淀，转化为数据分析所需的知识，并进行多种应用。



知识沉淀

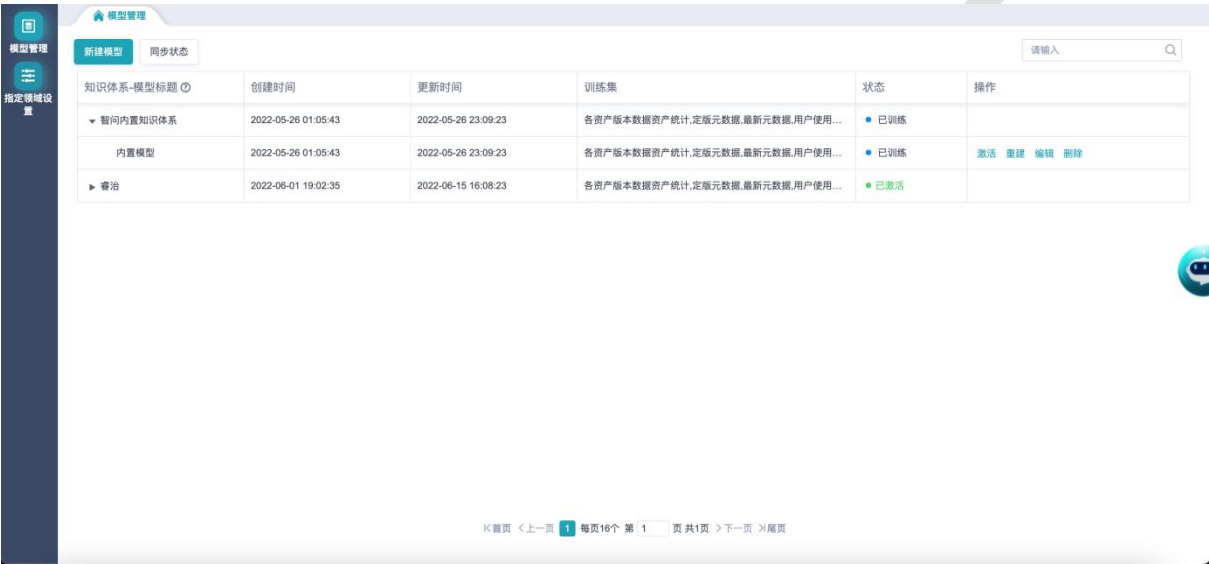
3.10.6 知识图谱

基于知识图谱梳理数据,在保留数据完整性的基础上,增加了对数据间关系的诠释,系统能更智能化的理解数据、知识推理,从而实现智慧搜索与智能交互。



3.10.7 模型管理

选择需要训练的知识图谱，进行 rasa 模型训练，增加分析的准确性，为用户精准提供分析结果。



模型管理

4 产品技术特点

4.1 高扩展

➤ MOF 元模型架构

元数据模块采用 MOF 规范作为元数据管理架构，可同时支持 UML、CWM 标准，可随意扩展添加任意技术元数据，业务元数据和管理元数据的元模型。

➤ 采集适配器

平台采集适配器支持接口扩展，系统内置 10 多种数据源的自动化采集适配器，也可针对新的数据源，在零配置的情况下快速实现并扩展。

-
- 所有功能都是根据可扩展可通用的原则来设计的，可以适应未来未知的需求变化。
 - 提供各种定制接口和调用接口 Web Services 服务，满足第三方服务需要。
 - 内置脚本支持，提供二次开发平台设计，满足未来各种未知需求。

4.2 自动化

- 支持连接各种数据库自动采集元数据
- 支持元数据依赖关系的自动分析和建立关联
- 支持从 SQL 中自动解析元数据和依赖关系
- 支持各种业务系统数据自动交换、处理
- 支持数据质量自动化检查
- 支持图表根据数据维度和数据特征信息自动关联
- 支持多模块的自动监控，预警通知

4.3 全方位

- 遵循全方位、严密的安全编程规范，具有代码级的产品安全性，编码中对不安全的数据进行严格校验。
- 使用专业的安全工具 Fortify、WebScan、AppScan 多重扫描，全面排除 Web 漏洞、SQL 注入、SQL 盲注、隐私侵犯、XSS 攻击等一系列安全问题。
- 支持 AES-128、AES-192、AES-256、DES、3DES、RC4、RC5 等多种加密方法。
- 用户权限认证机制，从源头杜绝数据安全风险源，提供完善的审计日志跟踪和预警通知，防止数据外泄和丢失。
- 拥有自主开发的数据库连接池技术，支持所有主流数据库。

-
- 数据整合内置了多种函数库，提供跨数据库类型和版本的算法兼容性，同时支持自定义，提供了丰富的数据处理组件的方式来保证工具的扩展性。

4.4 高性能

- 睿治平台采用了并行处理技术，执行引擎根据语法解析，实现转换任务的组件和清洗规则的自动化合并，多组件和转换规则并行处理执行，大大提高计算性能。
- 睿治平台使用了 spark 引擎，依赖 spark 引擎的集群、内存计算、并行计算能力，大幅度缩短了大数据量质检以及大数据量交换任务需要的时间，提高了效率，优化了用户体验。
- 元数据分析、检核计算采用线程池和多线程并行技术，大大提高计算性能，减少计算等时间，优化了用户体验。
- 平台可以通过内存计算提高分析表计算效率，跨源数据分析时，将所要分析的多源数据库表缓存至内存后进行分析。
- 合理的 JAVA 内存回收机制，保证系统资源的合理高效利用。
- 平台支持集群部署方式，此环境下，可以提高系统的吞吐量，更好的利用硬件资源。

5 软硬件运行环境

5.1 服务器配置推荐

5.1.1 传统架构部署

5.1.1.1 推荐客户：小范围试用和用户量较小的

业务数据参考				
模块	功能点	数据量	用户数	并发用户数

元数据	元数据检索	十万级	200 户	10 户
数据标准	数据标准检索	十万级	200 户	10 户
数据集成	数据迁移	百万级	200 户	10 户
	数据处理	百万级	200 户	10 户
	调度执行	百万级	200 个	5 个
数据质量	数据质检	百万级	200 户	10 户
数据资产	资产下载	1MB 文件	200 户	10 户
数据服务	API 接口调用	百万级库表	200 户	10 户

资源配置信息	
配置项目	说明
应用服务器*1 台	CPU: 8C 内存: 16G 硬盘: RAID5 500G 网络: 1000M
数据库服务器*1 台	CPU: 16C 内存: 32G 硬盘: RAID1 500G 实际按需 网络: 1000M
推荐操作系统及软件	操作系统: CentOS7 中间件: tomcat-8.5.12 数据库: Oracle11g

5.1.1.2 推荐客户：大型企业、省级机构

业务数据参考				
模块	功能点	数据量	用户数	并发用户数
元数据	元数据检索	百万级	2000 户	100 户
数据标准	数据标准检索	百万级	2000 户	100 户
数据集成	数据迁移	千万级	2000 户	100 户
	数据处理	千万级	2000 户	100 户
	调度执行	千万级	2000 个	50 个
数据质量	数据质检	千万级	2000 户	100 户
数据资产	资产下载	10MB	2000 户	100 户
数据服务	API 接口调用	千万级库表	2000 户	100 户

资源配置信息	
配置项目	说明
应用服务器*2 台	CPU: 48C 内存: 96G 硬盘: RAID5 1T 网络: 1000M

Redis 服务器*3 台	CPU: 4C 内存: 8G
数据库服务器*1 台	CPU: 64C 内存: 256G 硬盘: RAID1 1T 实际按需 网络: 1000M
推荐操作系统及软件	操作系统: CentOS7 中间件: tomcat-8.5.12 数据库: Oracle11g

5.1.2 云原生架构部署

5.1.2.1 推荐客户：小范围试用和用户量较小

业务数据参考				
模块	功能点	数据量	用户数	并发用户数
元数据	元数据检索	十万级	200 户	10 户
数据标准	数据标准检索	十万级	200 户	10 户
数据集成	数据迁移	百万级	200 户	10 户
	数据处理	百万级	200 户	10 户
	调度执行	百万级	200 个	5 个
数据质量	数据质检	百万级	200 户	10 户
数据资产	资产下载	1MB 文件	200 户	10 户
数据服务	API 接口调用	百万级库表	200 户	10 户

资源配置信息	
配置项目	说明
应用服务器节点*1Pod	CPU: 12C 内存: 16G 持久卷: 100G 网络: 1000M
数据库服务器*1 台	CPU: 16C 内存: 32G 硬盘: RAID1 500G 实际按需 网络: 1000M
推荐软件	容器集群管理系统: K8S 数据库: Oracle11g

5.1.2.2 推荐客户：大型企业、省级机构

业务数据参考				
模块	功能点	数据量	用户数	并发用户数
元数据	元数据检索	百万级	2000 户	100 户
数据标准	数据标准检索	百万级	2000 户	100 户
数据集成	数据迁移	千万级	2000 户	100 户
	数据处理	千万级	2000 户	100 户
	调度执行	千万级	2000 个	50 个
数据质量	数据质检	千万级	2000 户	100 户
数据资产	资产下载	10MB 文件	2000 户	100 户
数据服务	API 接口调用	千万级库表	2000 户	100 户

资源配置信息	
配置项目	说明
应用服务器节点*10Pod	CPU：12C 内存：16G 持久卷：100G 网络：1000M
Redis 服务器*3 台	CPU：4C 内存：8G Redis 模式：哨兵模式
数据库服务器*1 台	CPU：64C 内存：256G 硬盘：RAID1 1T 实际按需 网络：1000M
推荐软件	容器集群管理系统：K8S 数据库：Oracle11g

5.2 客户端配置

5.2.1 客户端(建议配置)

所属项	说明
CPU	不低于酷睿 i3
内存	不小于 4G

网络	不低于 100KB/s 传输速度
----	------------------

5.2.2 客户端浏览器

操作系统	浏览器
Windows	Chrome 浏览器, 360 安全浏览器

5.3 信创环境组合

5.3.1 服务器

CPU	操作系统	数据库
飞腾 FT-1500A	银河麒麟 V4.0	达梦数据库 V7.6
龙芯 Loongson-3B4000	UOS 统信	达梦数据库 V8.1
鲲鹏 Kunpeng 920	UOS 统信	达梦数据库 V8.1

5.3.2 客户端

CPU	操作系统	浏览器
龙芯 Loongson-3A4000	中标麒麟 V7.0	360 安全浏览器
鲲鹏 Kunpeng 920	UOS 统信	360 安全浏览器

5.3.3 中间件

中间件	jdk	操作系统	备注
TongWeb6.1.5.8	Open jdk1.8	UOS 统信	

TongWeb7.0	Open jdk1.8	UOS 统信	
TAS2.8.5	Open jdk1.8	银河麒麟 4.02	

6 平台应用案例

6.1 金融行业典型案例

6.1.1 赣州银行数据标准化项目

6.1.1.1 项目概述

本项目是我公司为赣州银行实施的数据治理项目，赣州银行拥有存量系统 61 个，早在 2014 年就已经搭建数仓系统，拥有不错的数据处理水平，信息化已经达到了一定程度。由于没有统一数据管理工具，数据管理很混乱，已经大大不能满足业务的发展需要了。赣州银行于 2019 年正式启动新信贷等系统的建设，趁此机会，同时开展数据治理，用专业的数据治理理念去指导这些新系统的建设。为实现全行所有系统的数据治理积累经验。

6.1.1.2 项目成果

- 1、通过长达 5 个月的元数据管理，实现了新信贷、新总账、新理财、核算等系统的注释填充率到达 100%；每一次元数据变更都是可控，实现了元数据的变更事前管理。
- 2、总计梳理了 8 个主题、1244 条标准，对 7000 多个关键字段进行落地评估，

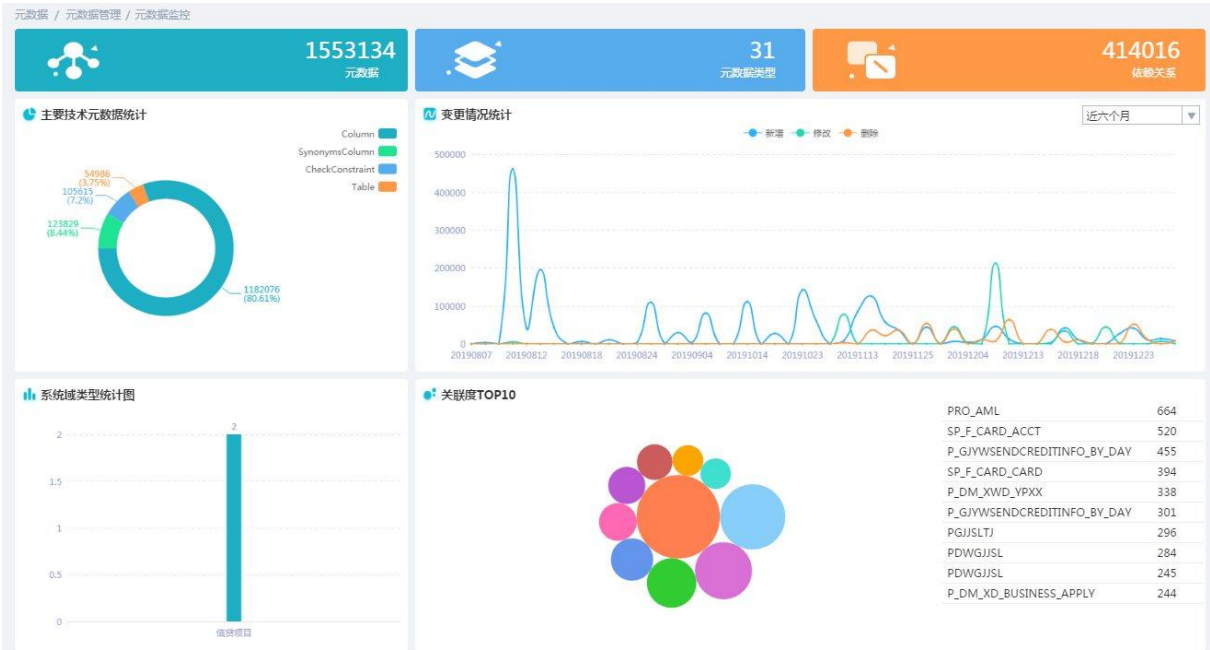
将治理初期 20%的标准通过率提升到现在的 85%的通过率。

- 3、 根据 east4.0 的要求，梳理了 300 余条规则，建立了相应的预警机制，实现了监管报送指标的自动化监管。严格按照 PDCA 策略对数据质量进行持续提升。

6.1.1.3 项目建设内容

- 建立元数据管理体系，建立数据标准管理体系、建立数据质量管理体系，规范全行的数据管理机制。
- 建立管理委员会，让各部门、各项目组积极参与到数据管理中来，提升数据管理水平。
- 制定 8 个主题的数据标准，新建系统全面推进数据标准的落标。
- 搭建数据治理平台，对生产环境、准生产环境、开发环境的数据进行同时治理。
- 确定标准化范围，对重要属性进行标准化；

6.1.1.4 项目展示



元数据监控界面

6.1.2 招商局租赁数据治理项目

6.1.2.1 项目概述

本项目是由我公司为招商局租赁实施的数据治理综合项目，从数据治理体系、数据仓库建设、数据应用等多方面进行建设。项目采用“分阶段、有侧重、体系化推进”等原则进行建设。

6.1.2.2 项目价值

- 梳理了 800 余条数据标准，建立了数据标准管理体系，为企业内部数据标准化持续发力；
- 实现了数仓落标，实现了数据质量的持续提升；
- 提升了数据质量，在数据完整性和准确性上有了明显提升，为业务发展提供了

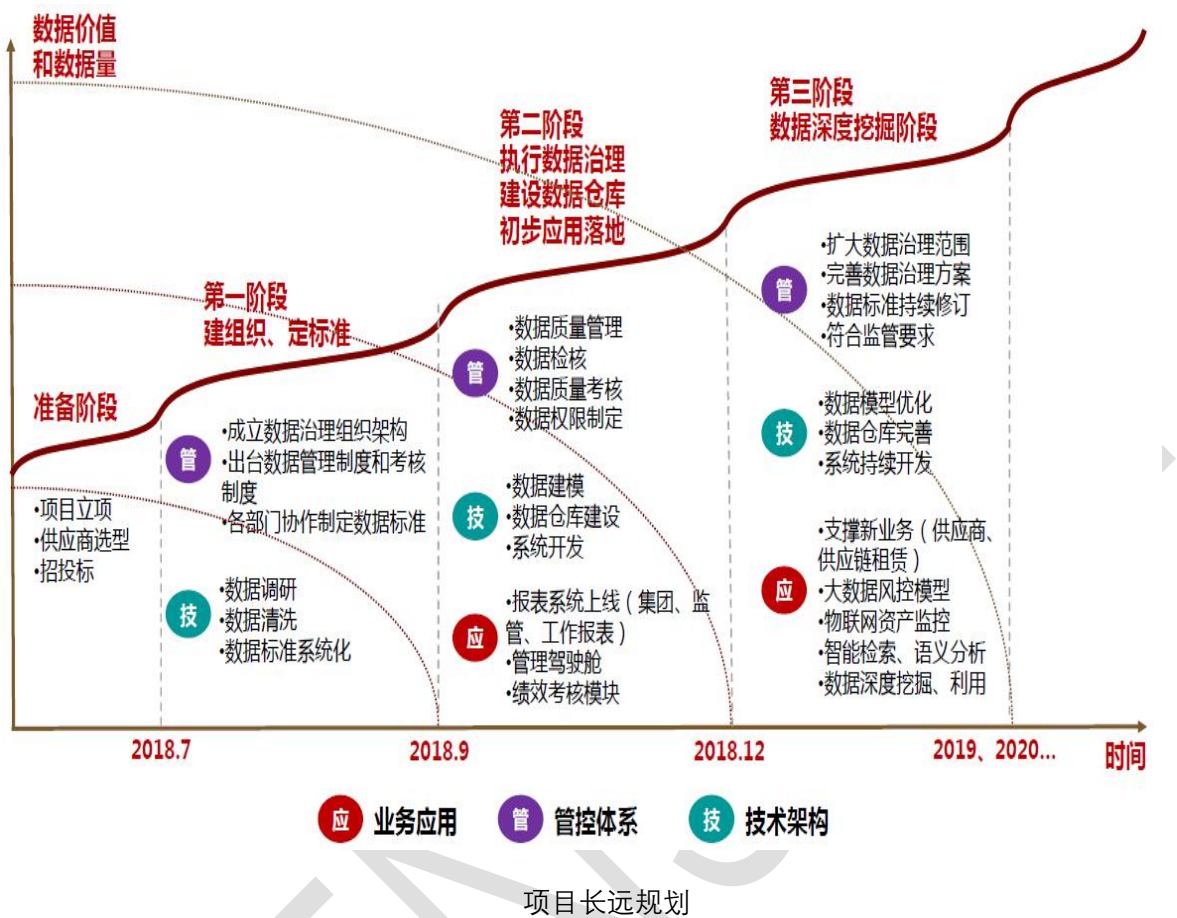
支撑；

- 打破数据孤岛，实现了数据共享，挖掘了数据价值；
- 梳理了监管报表，满足了集团及监管对数据报送的数量和要求的不断提升；

6.1.2.3 项目建设内容

- 数据治理体系：协助建立数据治理组织架构及管理办法；
- 数据标准：梳理 800 多项数据标准，并在数仓落标；
- 元数据：搭建数据治理平台管理元数据，实现数据溯源；
- 数据质量：通过平台实现数据质量自动检核及问题整改，辅助提升数据质量；
- 数据仓库方面，搭建基础数据平台，实现数据标准落地，并依据具体业务领域开发数据集市；
- 数据应用方面，实现管理驾驶舱建设及内部工作报表建设，未来可扩展 1104、人行、East 等监管报送应用。

6.1.2.4 项目长远规划



6.2 政府项目典型案例

6.2.1 禅城区大数据统筹局数据治理项目

6.2.1.1 项目概述

本项目是我公司为禅城区大数据统筹局进行的数据治理项目，帮助禅城区大数据统筹局治理卫计、工商、流管三个部门的数据，推动数据治理工作的发展。

6.2.1.2 项目价值

- 以治理自然人库数据为主，重点治理卫计、工商、流管三个部门的数据

-
- 为部门及数据应用提供高质量数据支撑
 - 支持自助填表系统、教育无纸化两大应用
 - 通过数据比对功能，返回比对结果，提供错误修改建议，提高填报成功率

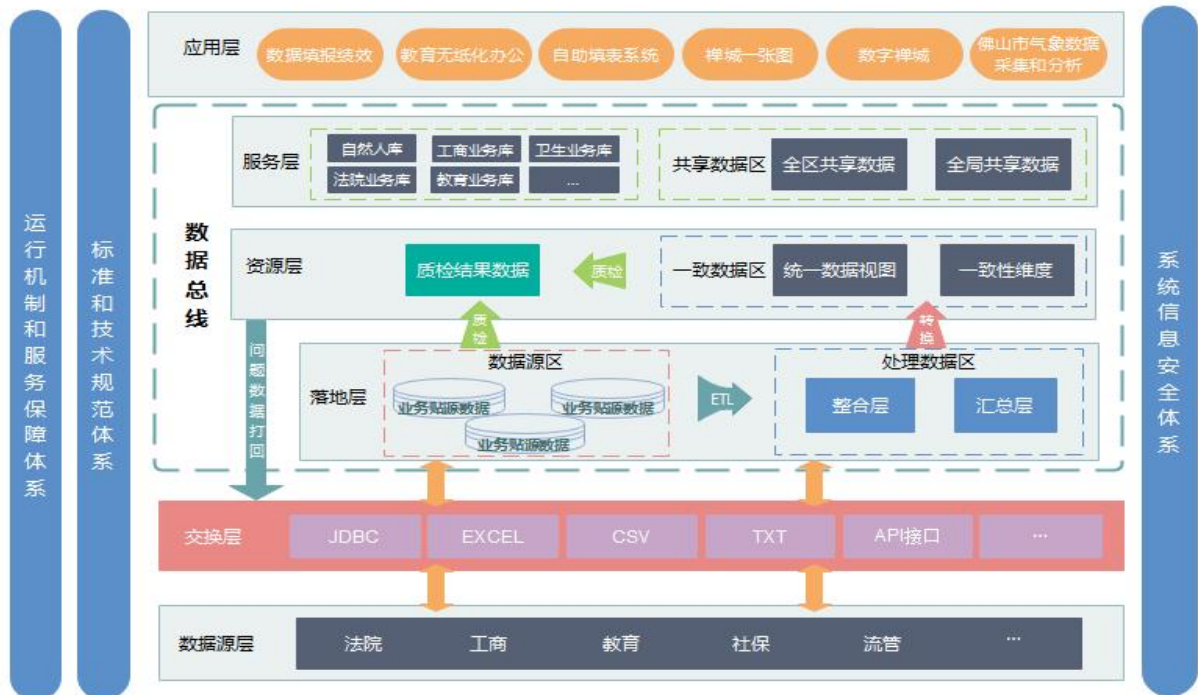
6.2.1.3 项目建设内容

- 盘点数据资产，建设管理驾驶舱，支撑政务数据交换监管报表需求，对历史数据查询
- 初步建立数据管控体系，重点建设数据标准体系
- 进行元数据和数据质量平台搭建
- 建设政务数据模型、建立政务数据交换体系、建立技术管理规范、ODS、集市层建立
- 构建模型、配置规则、人工智能识别等方法对禅城区数据进行数据核查、数据清洗、数据更新及数据修复
- 输出数据分析报告，提出改进建议，为禅城区各个部门数据源提供数据标准规范及改进建议
- 为政府部门、事业单位、企业、市民等数据使用者提供更优质、更高效的数据支持

6.2.1.4 项目展示



资产大屏



功能架构

6.2.2 荔湾区政务数据服务平台项目

6.2.2.1 项目概述

广州市从 2017 年初开始开展“深化平安有序、规范城市管理”专项行动，大力推进“四标四实”（标准作业底图、标准地名地址、标准建筑物编码、标准网格，实有人口、实有房屋、实有单位、实有设施）的建设工作。以此推进城市管理精细化，提升社会治理效能，夯实城市管理和社会治理的基础。

目前，四标四实专项成果应用初显成效，根据全市持续深化平安有序规范城市管理专项行动相关要求，为整合荔湾区政务数据资源，广州市荔湾区城市社区网格化服务管理中心拟建设一个荔湾区统一的四标四实数据平台，响应中央政治局就实施国家大数据战略加快建设数字中国的学习精神和广东省加快推进“数字政府”建设的要求，为大数据时代的政务大数据管理和增值发展提供全面支持。

6.2.2.2 项目价值

借鉴省市及其他城市数据治理管理实施经验，结合荔湾区自身数据运营管理特点，对区内四标四实数据资源进行运营管理和数据资源盘活，有效支撑区内政务服务数据使用需求和各部门业务系统的数据服务需求，以应用驱动事项为引领，对数据共享业务实施管理运营服务，包括建立数据管控体系、实施计划管理、协调监督、绩效考核考评、经验总结归档等，最终形成一整套切实可行的数据标准目录和数据管理运营体系。

6.2.2.3 项目建设内容

- 搭建荔湾区四标四实数据平台

充分利用荔湾区四标四实数据，结合全区政务数据资源，搭建荔湾区四标四实数据平台，对数据进行统一管理，为前端应用提供数据支撑。四标四实数据平台应包含基础模块、管理模块、应用模块三个部分。

➤ 形成四标四实数据示范应用

通过对四标四实数据的接入、处理、组织和管控，构建业务模型和分析模型，结合 5G、数据挖掘、人工智能、物联网、分布式数据存储等技术，为经济、民生及其他政府重点关注热点问题提供辅助决策，形成可靠有效的示范应用，从而提高各项政策制定的科学性和合理性。

6.2.2.4 项目展示



数据资产首页

6.2.3 卫计委-数据质量管控系统

6.2.3.1 项目概述

为用户面向国家部委、三十二个省针对系统的数据质量问题进行分析，形成数据质量评估报告，在数据源头加强对数据质量的控制。

6.2.3.2 项目价值

- 建立了全面的质量规则，使数据无死角，大大提高了数据的数据质量；
- 制定了周密的问题整改计划，让数据问题快速、及时得到有效整制；
- 制作个性化的质量报告，让领导对质量问题一目了然；
- 建立完善质量知识库，有问题可快速查到查关解决办法。

6.2.3.3 项目建设内容

- 建立面向全国卫生统计机构的数据质量管控平台，全面梳理数据质量规则；
- 制定符合业务要求的质检方案，定期自动执行数据质量检查，自动生成描述数据质量情况的报告，提供数据质量问题的详情查询；
- 根据数据质量问题的来源发起整改，从数据源头开始加强数据质量的管控，提升数据质量；
- 将数据质量提升过程中的知识、经验的积累形成数据质量知识库，实现数据质量管控的不断改进。

6.2.3.4 项目展示



质量评估

数据质量		数据期: 2017年	数据级次: 青海省	实体表: 全部				
	规则代码	规则描述	错误合计	例外合计	修复合计	操作		
质量监控	1_1-LOGIC-006	1-1行政区划在县级市,主办单位等于1,隶属关系等于5...	1	0	0	查看		
	1_1-LOGIC-008	1-1主办单位为1,隶属关系为2的机构	10	0	0	查看		
	1_1-LOGIC-010	1-1主办单位为1,隶属关系为3机构名称含区字非直辖...	2	0	0	查看		
	1_1-LOGIC-011	1-1行政区划在县级市,主办单位等于1,隶属关系不等于...	1	0	0	查看		
	1_1-LOGIC-014	1-1医院类主办单位为1,隶属关系为4机构名称含县字...	1	0	0	查看		
	1_1-LOGIC-015	1-1医院类主办单位为1,隶属关系为4机构名称不含县...	1	0	0	查看		
	1_1-LOGIC-016	1-1医院类主办单位为1,隶属关系为4机构名称不含县...	1	0	0	查看		
	1_1-LOGIC-021	1-1机构名称非口腔眼科实有床位小于300三级分支...	2	0	0	查看		
	1_1-LOGIC-025	1-1机构名称含县字三乙医院情况	1	0	0	查看		
	1_1-LOGIC-026	1-1实有床位<100二级综合医院	22	0	0	查看		
	1_1-LOGIC-033	1-1上期实有床位>1本期实有床=0	2	0	0	查看		
	1_1-LOGIC-034	1-1在岗职工数>1000	5	0	0	查看		
	1_1-LOGIC-035	1-1执业医师>500或者执业助理医师>100	3	0	0	查看		
	1_1-LOGIC-037	1-1全科医师>10	2	0	0	查看		
	1_1-LOGIC-038	1-1入院人数>10000,入院人数大于1.5倍出院人数或...	1	0	0	查看		
	1_1-LOGIC-039	1-1入院出院人数和上期入院出院人数>1000,本期入院...	1	0	0	查看		
例外审批	1_1-LOGIC-045	1-1总诊疗人次>500000,<2500000	4	0	0	查看		
	1_1-LOGIC-048	1-1房屋建筑面积>100000或房屋建筑面积>10000时...	6	0	0	查看		

质量规则

6.3 企业项目典型案例

6.3.1 瀚蓝数据资源中心管理项目

6.3.1.1 项目概述

本项目是我公司为瀚蓝集团实施的数据治理项目，瀚蓝环境信息化建设和应用水平

处于行业前列，已建设供水、排水、燃气、固废板块等不少于 15 个业务系统，在公司发展的 20 多年间，信息化的建设大多是以分散建设为主，各板块建设更是以满足单个或局部的业务功能为主，但是还存在一些问题。由于分散建设，各信息系统建设自成一体，数据融合困难，信息孤岛数据共享困难问题突出；数据缺乏标准，系统变更影响分析困难，数据质量有待进一步提升；系统不能互通，数据难以共享和利用，企业数据资产无法集中管理和展现，数据价值难以体现。我公司利用丰富的数据治理经验，为瀚蓝环境制定了合适的数据治理方案，推动整个信息资源中心的建设。

6.3.1.2 项目价值

- 数据互通，为业务赋能，节约成本、提升工作效率，解决了瀚蓝环境在设备维修过程中即时获取设备成本，以便水厂控制生产成本问题
- 数据共享，实现在业务流程中即可快速获取相应数据
- 实现数据可溯源，解决数据来源不明确问题
- 实时获取数据功能，实现数据后台实时调度取数，保证数据的有效性、准确性
- 制定严格的审批流程，减少人工对数据的影响程度，通过系统审核加人员审批两种方式，提升数据质量，为瀚蓝环境后续与水厂实时监测数据、水厂化验数据的对比分析和使用建立基础

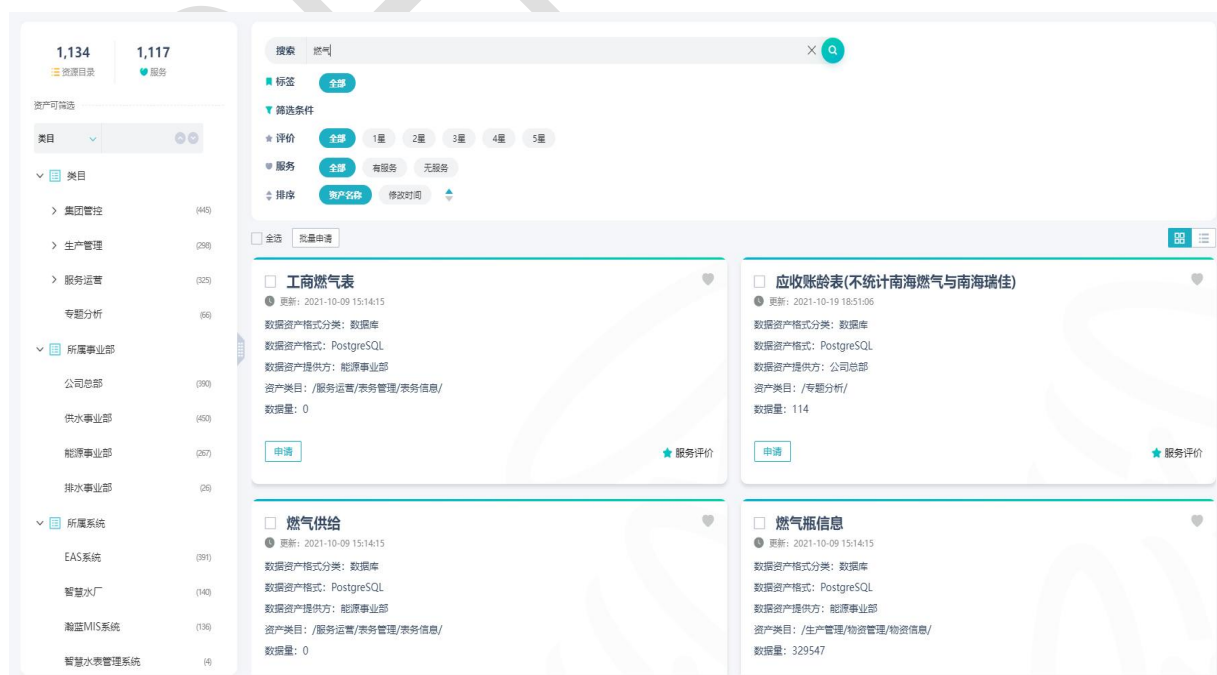
6.3.1.3 项目建设内容

- 进行数据建模，梳理各系统数据，进行分层整合，为瀚蓝环境保证流程不断改进、实时监视数据、控制成本、提高数据质量以及项目后台管控
- 提供实时的数据交换，保障各级监测到的数据为最新数据，且按照统一标准规

范，确保数据准确无误

- 制定规范，对瀚蓝采集回来的数据进行了元数据、数据标准、数据质量、数据资产四个层面的统一管理
- 元数据管理建设，提升数据标准化与数据管控能力，提升共享与重新获取、理解企业信息资产的水平
- 数据标准建设，针对瀚蓝各种重要数据制定的标准，在瀚蓝数据的内外共同使用和交换过程中，可以保障这些数据的一致性和准确性
- 数据质量管理规范，制定规则共有 104 条，覆盖各业务系统及其业务数据表，其规则类型包括但不限于逻辑检查，规范检查
- 资产目录分类，已收录 1128 个资源目录，其中包括资产表 1077 张，集市 41 张，维表 10 张，类目分为“集团管控”，“生产管理”，“服务运营”和“专题分析”。

6.3.1.4 项目展示



资产目录

数据资产 / 数据资源 / 资产目录

瀚蓝数据资产目录

切换版本

手动添加

智能添加

更多操作

列表管理

目录活化配置

刷新

资产代码/资产名称

Q

类别

Q

集团管控(445)

生产管理(298)

服务运营(325)

专题分析(66)

序号	数据资产代码	数据资产名称	审批状态	上架状态	数据资产提供方	操作
1	A02/000755	进场登记类分票	已通过	已上架	公司总部	编辑 申请变更 删除
2	A01002/000735	系统用户表	已通过	已上架	供水事业部	编辑 申请变更 删除
3	A02/000938	合同支付项	已通过	已上架	公司总部	编辑 申请变更 删除
4	A02/000820	供应商分账明细	已通过	已上架	公司总部	编辑 申请变更 删除
5	A02/001002	收款单	已通过	已上架	公司总部	编辑 申请变更 删除
6	A02/001035	工作分解结构	已通过	已上架	公司总部	编辑 申请变更 删除
7	A02/000792	期间成本分解结构台账	已通过	已上架	公司总部	编辑 申请变更 删除
8	A02/000738	借款核销表	已通过	已上架	公司总部	编辑 申请变更 删除
9	A02/000960	申请核销表	已通过	已上架	公司总部	编辑 申请变更 删除
10	A02/000839	费用报销单多收款人分票	已通过	已上架	公司总部	编辑 申请变更 删除
11	A02/000912	分票	已通过	已上架	公司总部	编辑 申请变更 删除
12	A02/000743	费用报销单付款明细分票	已通过	已上架	公司总部	编辑 申请变更 删除
13	A02/000890	费用类型分票	已通过	已上架	公司总部	编辑 申请变更 删除
14	A02/001014	成本分解结构台账	已通过	已上架	公司总部	编辑 申请变更 删除
15	A02/001088	费用类型科目映射	已通过	已上架	公司总部	编辑 申请变更 删除
16	A02/000845	其他工程履约合作合同	已通过	已上架	公司总部	编辑 申请变更 删除

前页 上一頁 第 1 页/5页 下一頁 末页 每页100条 共445条

© 瀚蓝环境 版权所有

目录分类

6.3.2 南山集团主数据项目

6.3.2.1 项目概述

本项目是我公司为南山集团实施的主数据管理平台项目。南山集团已经建立了包括财务管理、资产管理、人力资源管理在内的多个信息系统，且运行多年，积累了一定数量的各类型数据。但过去信息化建设遵循的理念是“业务驱动”而非“数据驱动”，忽略了对数据本身的关注，没有充分挖掘和利用数据的价值。通过项目建设，达到预期项目目标。

6.3.2.2 项目价值

- 建立适合于南山全集团各业务板块都能通用的物资分类体系，以满足供应链上各业务节点对物资分类管理以及管理精细度的需求。

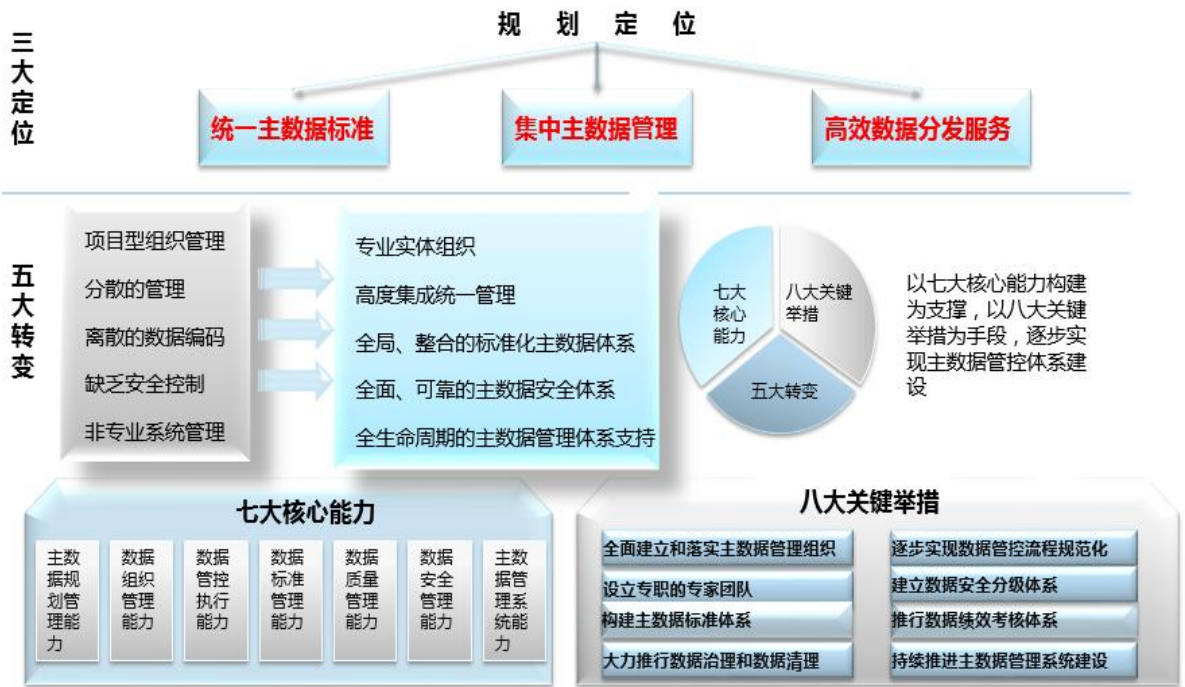
-
- 建立未来物资数据管理的基础能力，设置合适的标准化组织、流程机制与平台固化物资数据管理的工作
 - 按照新制定标准逐步规范统一南山全集团物资数据，具体实施计划，开展试点数据整理的工作
 - 搭建主数据管理平台、实现与核心系统集成

6.3.2.3 项目建设内容

- 六大类主数据建设：人力资源类（组织、人员）、财务类（会计科目、固定资产、往来单位、成本中心、利润中心等）、采购类（物料、供应商）、营销类（客户）、指标类和其他基础类主数据；
- 物料主数据，涉及 40+大类，3000+小类，10 万+实体数据；
- 对接的系统包括：供应链系统（SCM）、销售系统（CRM）、流程审批系统（BPM）、MES（纺织、服饰、轻合金）、ERP（纺织、服饰、轻合金）及其他经营单位系统之间的集成。

6.3.2.4 项目建设思路

总体发展思路：以七大核心能力构建为支撑，以八大关键举措为手段，逐步实现主数据管控体系建设



总体思路

6.3.3 时代地产数据资产管理项目

6.3.3.1 项目概述

本项目是我公司为时代地产集团实施的数据资产管理平台，大数据时代下非结构化数据爆发式增长，对数据加工的复杂度和速度要求更高，传统的数据管理已经无法适应当下企业对数据管理的需求。数据资产管理是对传统数据管理的扩充和升级；通过提高以数据指导管理决策的效率，数据资产管理已然成为大数据时代中企业竞争力的重要来源。

6.3.3.2 项目价值

- 数据资产管理，盘点全集团数据资产，对数据资产进行规范化管理，为业务部门提供数据资产服务；
- 元数据管理平台，搭建集团级的元数据管理平台，发布元数据管理办法，实现集团的元数据统一管控；
- 数据质量管理平台，提升集团业务数据质量，从核心业务出发，逐步提升集团业务数据质量。建立完善的数据质量考核机制，形成数据质量问题库，实现常态化的数据质量管理机制；
- 数据标准管理平台，建立集团级的数据标准体系，发布数据标准管理办法，推动业务系统的数据标准化发展，指导后续业务系统的新建，从根源上解决数据质量问题。

6.3.3.3 项目建设内容

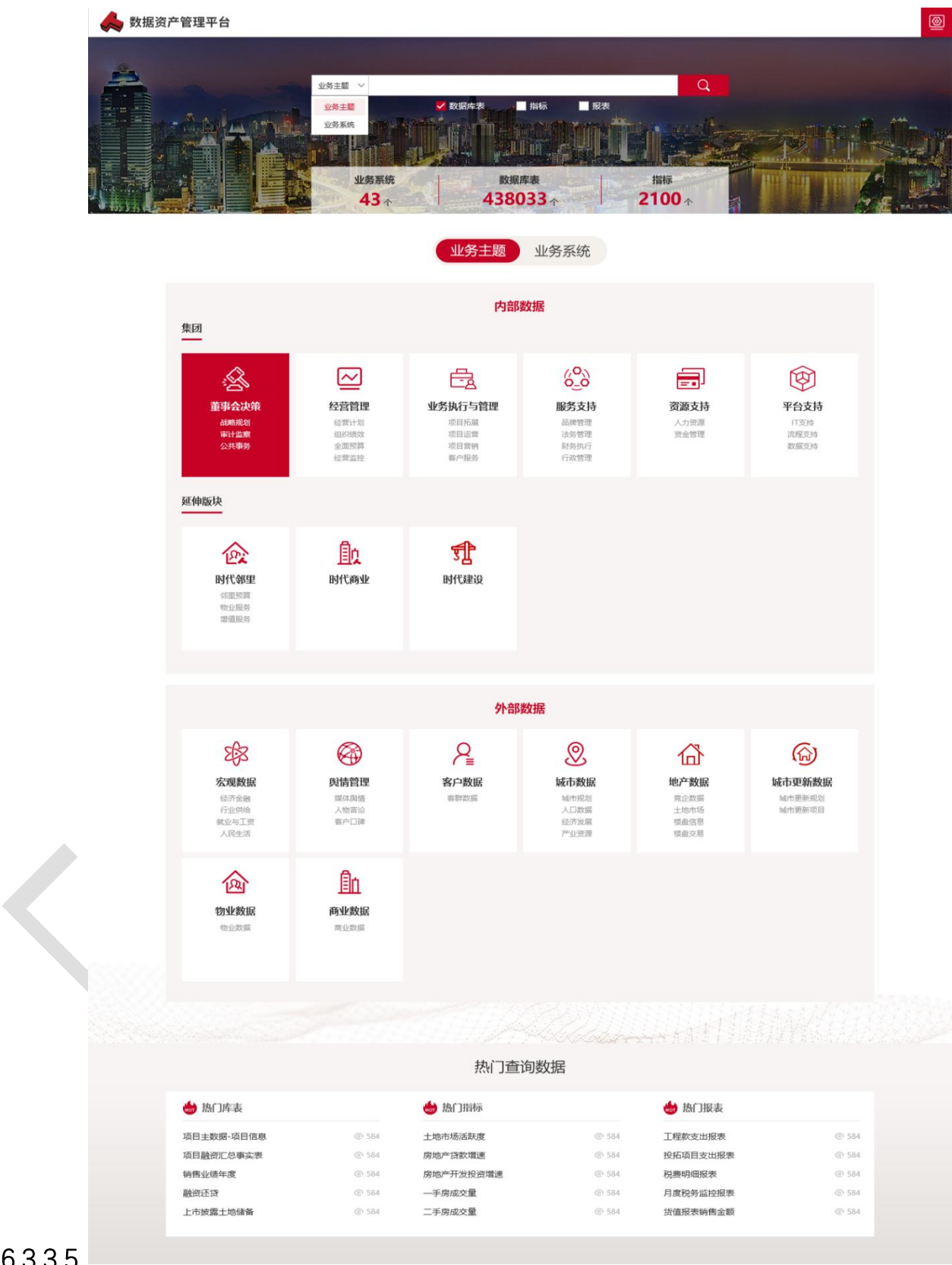
- 完成集团多个系统中的原始数据资产盘点，从源系统、数据平台、数据应用系统等通过元数据、数据标准、数据模型等工具落地数据现状，通过数据质量工具完成数据质量评估和量化，了解集团目前总体数据规模和分布情况。
- 依据数据资产管理需要，定制元数据管理方案。设计元数据模型，根据元模型设计和数据资产管理平台产品需求，制定元数据采集和接入方案并落地实施。
- 建立数据标准体系，导入项目类、客户类、人员组织类、合作伙伴类（供应商）的数据标准近千条，基础类数据标准百余个；实现了人力资源系统等多个系统的落地映射，实现质检规则平台落地。
- 完成从定义质量规则、到建立质检方案、到质检结果分析、再到问题整改的数

据质量管理全流程搭建测试，全流程全方位保障数据质量。

- 打造数据资产管理平台门户，实现了数据分类自动化和数据资产可视化，制定了数据分类维护方案，保证分类的准确性和实施性。建立可视化大屏，让企业数据资产全景展现更直观。

ESENSENSE

6.3.3.4 项目展示



6.3.3.5

6.4 其他典型案例

以下仅列举展示部分案例：

服务单位名称	项目名称	服务单位名称	项目名称
禅城区大数据服务中心	禅城区大数据统筹局数据治理项目	中国人民银行清算中心	数据治理咨询项目
荔湾政数局	荔湾区政数局数据治理项目	中国进出口银行	数据清理项目 外部数据管理平台
江阴大数据局	江阴大数据局数据治理项目	中国农业发展银行	综合报表平台升级项目 数据交换平台项目
国家能源	国家能源指标管理平台项目	锦程消费金融	数据治理项目
广西交通厅	广西交通厅数据共享与交换项目	招商局融资租赁	数据治理及数据仓库项目
陕西卫生厅	陕西卫生区域数据管理平台项目	阜外医院	阜外医院数据治理项目
国网甘肃电力	甘肃电力数据资产管理平台项目	临沂矿业	临沂矿业数据治理平台
甘肃税务局	数据集中管理平台项目	时代地产集团	时代地产数据质量管理平台
凉山州政府	“智慧凉山”项目	南山集团	主数据管理平台建设项目
小龙潭监狱	小龙潭监狱数据治理项目	博威集团	博威集团主数据项目
濮阳环保局	濮阳环保数据治理项目	华致酒行	华致酒行数据治理项目
太平洋保险集团	太平洋保险集团数据治理项目	河南得实	河南得实数据质量管理

	理综合项目		平台项目
赣州银行	赣州银行数据标准化治理项目	上海东方富联	数据质量管理平台项目
长安信托	数据治理平台建设项目	安徽电力物资处	数据资产管理平台项目
洛阳银行	洛阳银行数据治理平台项目	中兴信网	连云港数据治理项目
秦皇岛银行	秦皇岛银行数据治理项目	中油瑞飞	塔里木油田数据治理项目
北矿智能科技	新疆有色金属主数据系统建设项目	长江存储	数据质量管理平台项目
楚昌投资集团	主数据管理系统一期项目	瀚蓝环境	瀚蓝环境数据资源中心项目
航天二院	KJ 研究院 5S 数据中台项目	首钢基金	首钢基金数据治理一期主数据项目
国防科技大学	国防科大西安信息资源目录管理系统项目	中国外贸金融租赁	外贸金租数据治理项目
枣庄矿业	枣庄矿业付村煤业数据治理项目	石油化工研究院	主数据管理平台建设项目
中国电子科技集团	中国电子科技集团公司第三十研究所数据治理与分析项目	神州龙芯智能科技	神州龙芯智能科技主数据项目
浙江顶誉食品公司	顶誉食品公司主数据项目	中铁建电气化局	2020- 中铁建电气化局主数据管理项目
石油化工科学研究院	石油化工科学研究院固定资产数据平台	中交集团	中交主数据咨询项目
云投集团	云投集团主数据管理平台项目	甘肃电力	甘肃电力数据资产 2.0 开发项目

楚昌投资集团	楚昌投资集团主数据项目	新疆有色金属工业	新疆有色金属工业（集团）MDM 主数据系统建设项目
江苏国强	江苏国强主数据项目	中国外贸金融租赁有限公司	外贸金租数据治理项目
雪松控股集团	雪松控股数据安全项目	KJ 研究院	KJ 研究院 5S 数据中台
首钢基金集团	2020 首钢基金数据治理一期主数据项目	国防科大	国防科大西安信息资源目录管理系统
长鑫存储技术有限公司	长鑫存储数据治理项目	浙江吉利新能源商用车集团	吉利商用车 2022 年数据质量监控平台项目