

数据治理服务说明

Data Focus
Value Creation

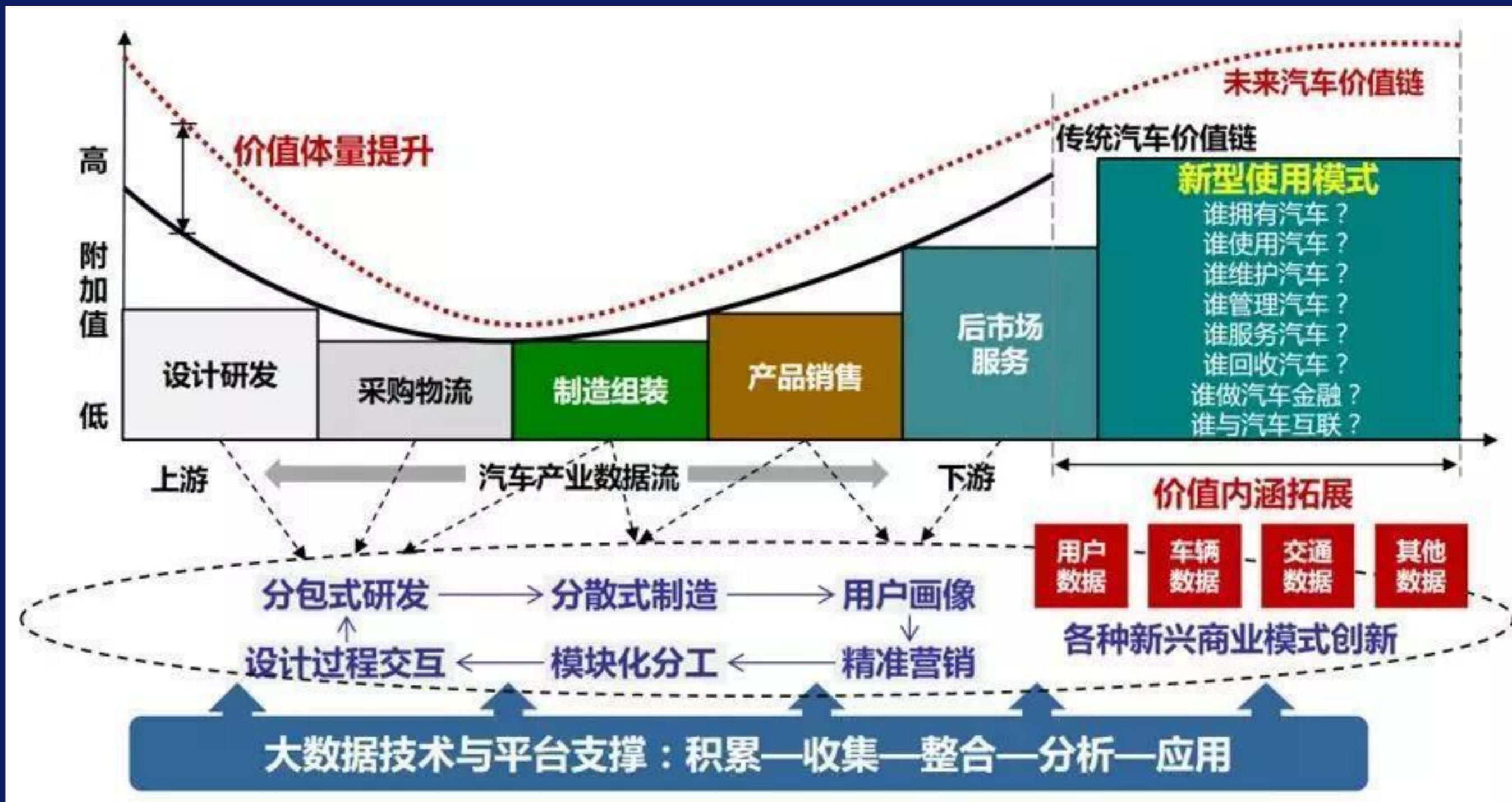
专注数据 创造价值

NASDAQ:GSUM

价值链及价值主张



数据分析和治理支撑价值体量提升，内涵扩展



方案架构



数据分析与治理平台的业务价值

现状 / 痛点

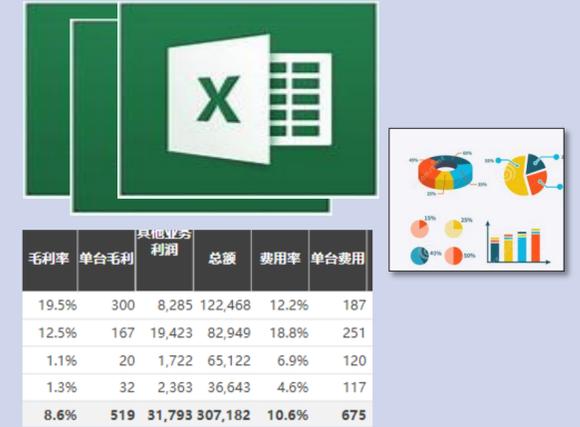
业务价值

示例

敏捷分析

- 报表只能找科技部，项目周期长
- 相似需求不断重复
- 制作日常报表费时耗力
- 缺乏图形化展示，表格展示不直观

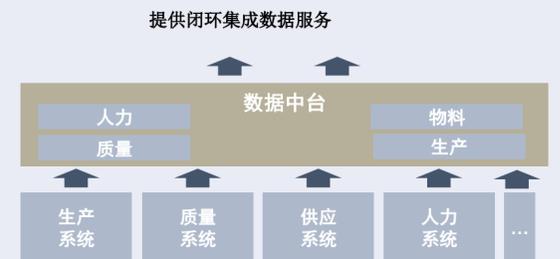
- 实现自助式数据分析
员工在平台自主进行数据分析，减少简单需求的立项开发时间
- 可定制化的报表
相似需求合并在同一报表，用户灵活选择展示数据与筛选条件，无需重复开发相似内容
- 自动的数据汇总与整合
实现公司主要业务数据的打通与汇总，无需人工收数，整合，减少数据处理时间
- 可视化、直观的结果展示
图形化的结果展示更直观的凸显问题，无需在表里内寻找答案



数据汇聚

- 系统众多，大多数系统间无接口，数据孤立，不闭环
- 数据需求需要跨部门、跨系统收集基础信息，需要较长时间得到可用的数据，并进行Excel操作进行整理

- 将核心系统通过数据层面进行打通，实现业务数据层面的闭环
- 可实现数据的全链路追踪，例如生产问题的跨系统，跨部门定位



数据治理

- 无数据标准定义，各系统间数据格式不统一
- 各系统间的数据整合难度较大，需要进行数据清洗、匹配，整理

- 在数据中台对各系统间的数据标准进行统一，关联数据自动匹配，夯实数据服务的基础，减少基础与重复的数据清理，提高效率，减少错误，提高员工满意度

设备	运行时间	状态
传输	20分	正常
M23	120	OK

设备：焊点 M230
运行时间：120分
运行状态：正常

快速响应

- 科技部对数据需求需要进行跨系统、跨平台的分析，通过多个异构系统的数据提取，清洗，整合，开发后才能提供满足需求的数据

- 通过跨系统，跨平台的数据整合，数据中台的自动化数据整合，清洗能力，缩减数据需求的响应时间。科技部同事只需要分析业务需求，提供数据，消除了重复的跨系统按需数据整合。



数据分析与治理方案架构



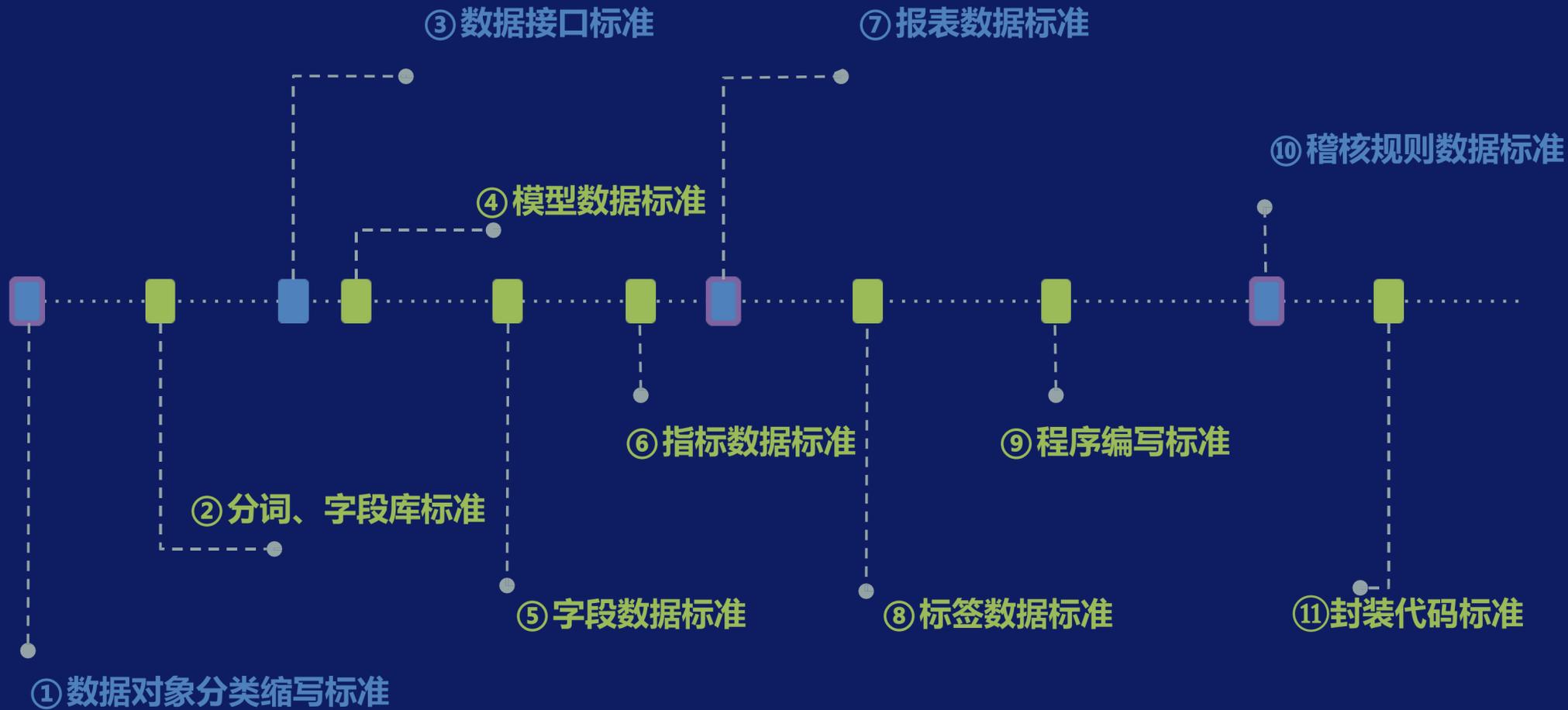
数据入仓



数据治理



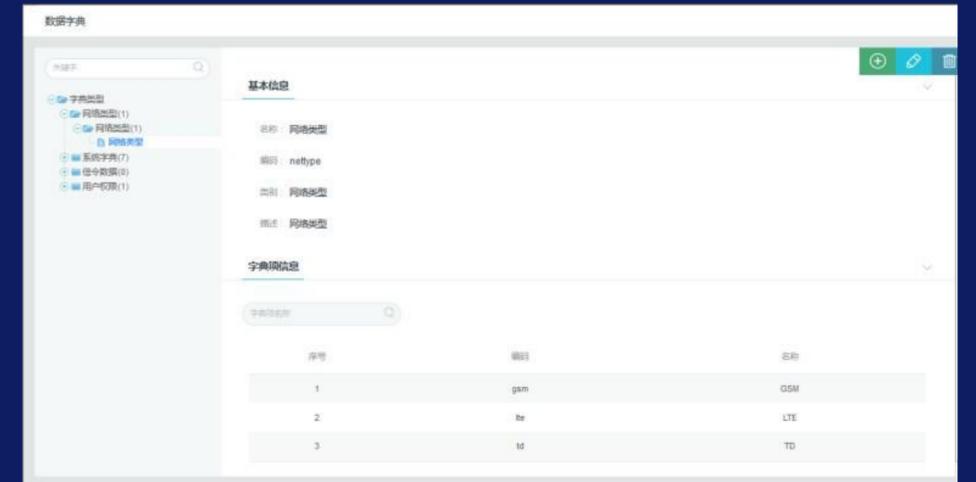
数据标准管理



数据标准体系

- 01.大数据平台数对象分类缩写标准>
- 02.大数据平台数分词、字段库>
- 03.大数据平台数接口标准V1.0.doc
- 04.大数据平台模型数据标准v1.2.doc
- 05.大数据平台字段数据标准V1.0.doc
- 06.大数据平台指标数据标准v2.3.doc
- 07.大数据平台报表数据标准v1.0.doc
- 08.大数据平台标签数据标准v2.0.doc
- 09.大数据平台程序编写标准v1.0.doc
- 10.大数据平台稽核规则数据标准.doc
- 11.大数据平台封装代码标准v2.0.doc

数据标准管理



字典管理



标准管理

数据质量管理

统一的数据质量监控

质量告警监控

数据监控系统-数据质量告警

告警列表

告警状态	操作	告警级别	编号	初次告警时间	最近告警时间	对象类型	对象	告警信息	逻辑
未处理		四级告警	6190258	2014-04-11 09:42:55	2014-04-11 09:42:55	RFI	Attach成功率	[2014-03-31 00:00:00]错误记录数占比:3.0,超过...	通
未处理		四级告警	6190257	2014-04-10 17:28:47	2014-04-10 17:28:47	RFI	覆盖区域	[2014-04-09 00:00:00]错误记录数占比:76.0,超过...	通
未处理		四级告警	6190256	2014-04-10 17:22:01	2014-04-10 17:22:01	RFI	PDF激活成功率	[2014-03-31 23:00:00]错误记录数占比:48.0,超过...	通
未处理		四级告警	6190255	2014-04-10 17:22:01	2014-04-10 17:22:01	RFI	PDF激活成功率	[2014-03-31 22:00:00]错误记录数占比:48.0,超过...	通
未处理		四级告警	6190254	2014-04-10 17:22:01	2014-04-10 17:22:01	RFI	PDF激活成功率	[2014-03-31 21:00:00]错误记录数占比:44.0,超过...	通
未处理		四级告警	6190253	2014-04-10 17:22:01	2014-04-10 17:22:01	RFI	PDF激活成功率	[2014-03-31 20:00:00]错误记录数占比:40.0,超过...	通
未处理		四级告警	6190252	2014-04-10 17:22:00	2014-04-10 17:22:00	RFI	PDF激活成功率	[2014-03-31 19:00:00]错误记录数占比:38.0,超过...	通
未处理		四级告警	6190251	2014-04-10 17:22:00	2014-04-10 17:22:00	RFI	PDF激活成功率	[2014-03-31 18:00:00]错误记录数占比:36.0,超过...	通
未处理		四级告警	6190250	2014-04-10 17:22:00	2014-04-10 17:22:00	RFI	PDF激活成功率	[2014-03-31 17:00:00]错误记录数占比:37.0,超过...	通
未处理		四级告警	6190249	2014-04-10 17:21:59	2014-04-10 17:21:59	RFI	PDF激活成功率	[2014-03-31 16:00:00]错误记录数占比:34.0,超过...	通
未处理		四级告警	6190248	2014-04-10 17:21:59	2014-04-10 17:21:59	RFI	PDF激活成功率	[2014-03-31 15:00:00]错误记录数占比:35.0,超过...	通
未处理		四级告警	6190247	2014-04-10 17:21:59	2014-04-10 17:21:59	RFI	PDF激活成功率	[2014-03-31 14:00:00]错误记录数占比:36.0,超过...	通
未处理		四级告警	6190246	2014-04-10 17:21:59	2014-04-10 17:21:59	RFI	PDF激活成功率	[2014-03-31 13:00:00]错误记录数占比:35.0,超过...	通
未处理		四级告警	6190245	2014-04-10 17:21:58	2014-04-10 17:21:58	RFI	PDF激活成功率	[2014-03-31 12:00:00]错误记录数占比:35.0,超过...	通

质量问题呈现

数据质量管理系统

服务器监控

数据库监控

FTP接口监控

ETL进程监控

数据源监控

多方式质量问题分析

知识库管理与查询

告警列表

知识库

编号	预警状态	逻辑分类	检测对象	逻辑子类	完整性
2988449	未处理	数据时间	2013-05-30 00:00:00	清除时间	
2	预警时间	2013-06-26 09:05:00	数据时间	2013-05-30 00:00:00	清除时间
TABLE	对象	CMF_G_C_CELL_HISTORY_BJ	网元类型	BSC	网元
R0726					

规则名称: 规则获取

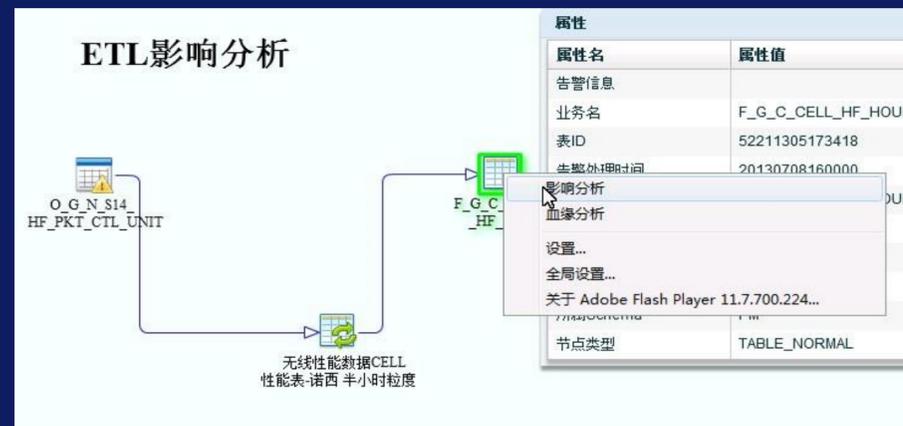
预警信息: 数据时间:2013-05-30 00:00:00,BSC:R0726,数据完整率:96.42%

工单ID: 处理人: 葛利平

处理说明: 填写如何处理的

处理人: 处理时间: 2013-06-26 09:04:52

血统/影响分析帮助问题定位



灵活规则设置

质量规则的统一管理

规则管理

查询结果

ID	规则名称	检查时间	检查对象	告警级别	是否启用	是否退单	处理人
1060	Attach成功率	0 30 T 0/1 * ?	Attach成功率:	一级告警	未启用	不退单	INSPUR
1061	传输成功率	0 5 0/1 * * ?	传输成功率:	二级告警	未启用	不退单	INSPUR
1062	会话成功率	0 5 0/1 * * ?	会话成功率:	三级告警	未启用	不退单	INSPUR
1063	覆盖区域	0 30 T 0/1 * ?	覆盖区域:	四级告警	未启用	不退单	INSPUR
1064	小区KEY	0 30 T 0/1 * ?	小区KEY:	四级告警	未启用	不退单	INSPUR
1065	PDF激活成功率	0 5 0/1 * * ?	PDF激活成功率:	三级告警	未启用	不退单	INSPUR
1077	指标核查规则测试	0 0 0/1 * * ?	PDF激活成功率得分:	一级告警	未启用	不退单	root1

灵活的规则设置

规则详情-KPI

基本信息

规则名称: Attach成功率

检查对象: Attach成功率

检查时间: 0 30 T 0/1 * ?

告警级别: 一级告警

处理人: INSPUR

是否启用: 否

是否派发工单: 否

标准规则:

- 同环比波动检查: (当前值-前 周期值) / 当前值 大于 %
- 空值记录数: 总记录数 * 100 大于 %
- 不唯一记录数: 总记录数 * 100 大于 %
- 超出正常取值范围的记录数: 总记录数 * 100 大于 %

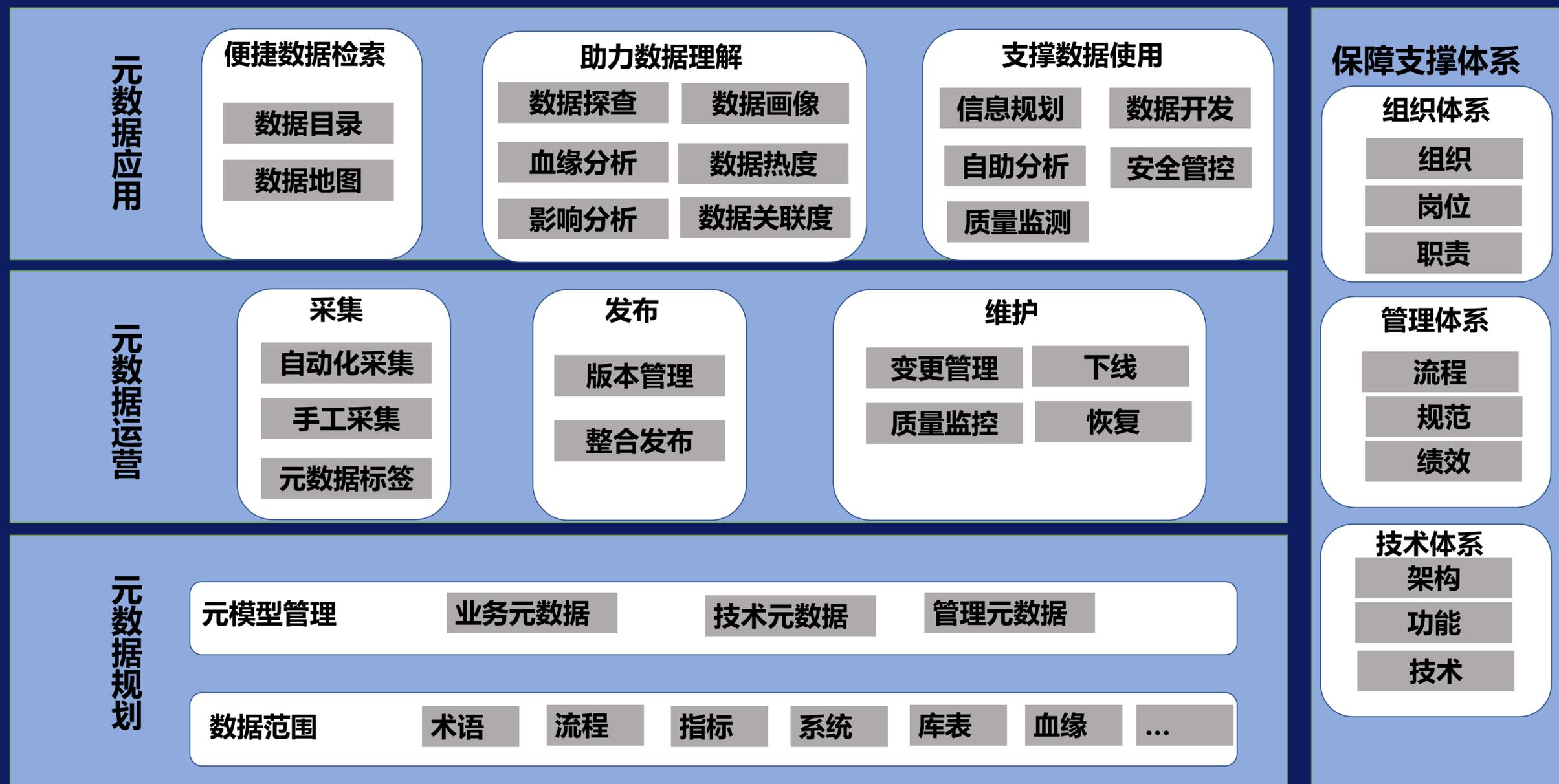
指标合理取值范围:

编写说明:

1、数值范围: 如大于1且小于100, 填写 1-100; 如大于等于1且小于100, 填写 [1-100]; 如大于1, 填写 1-5;

2、枚举类型: 如取值为a,b,c, 填写 a,b,c; 如取值为1,2,3, 填写 1,2,3;

元数据管理工作框架



源系统



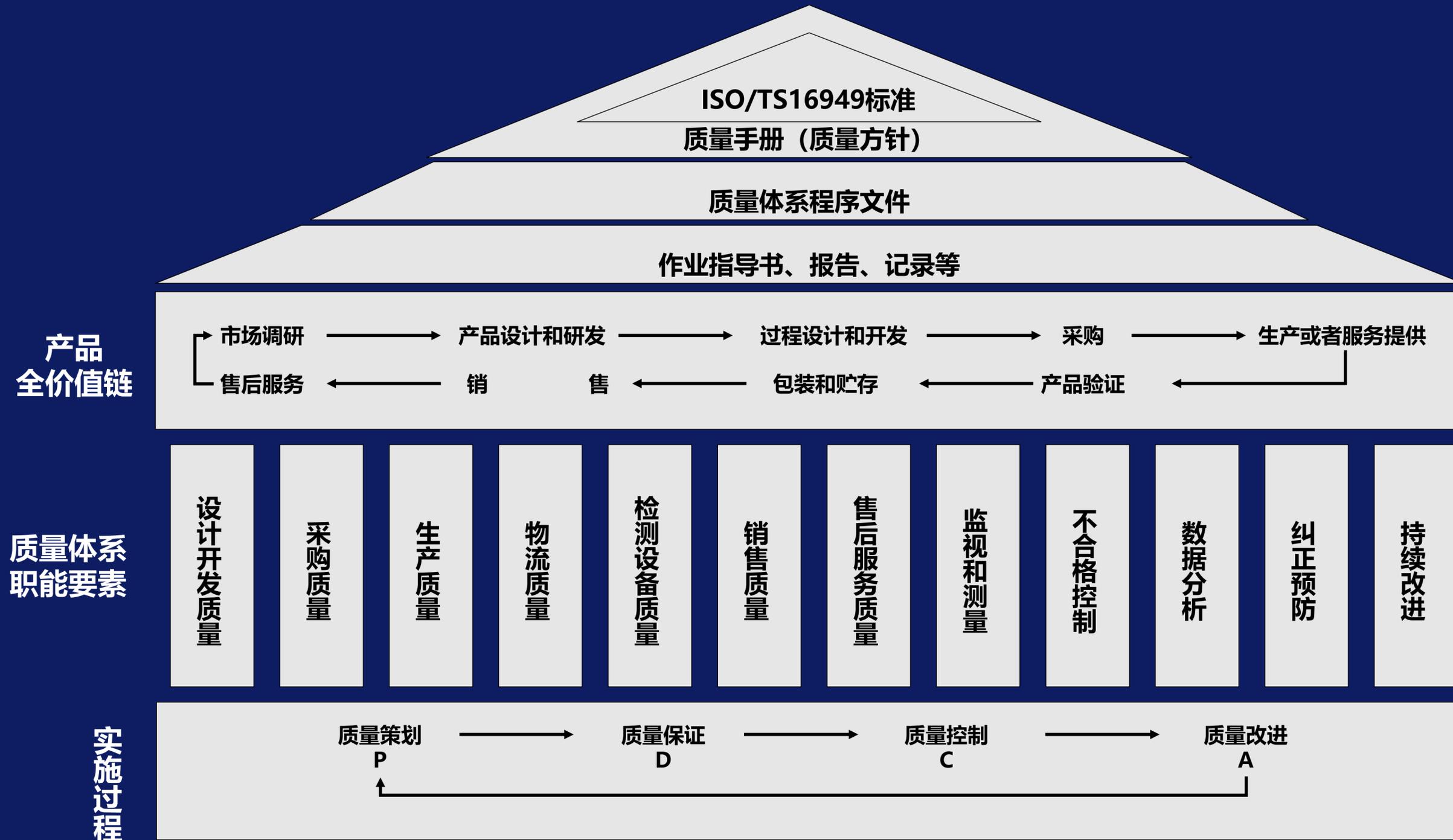
数据质量管理框架



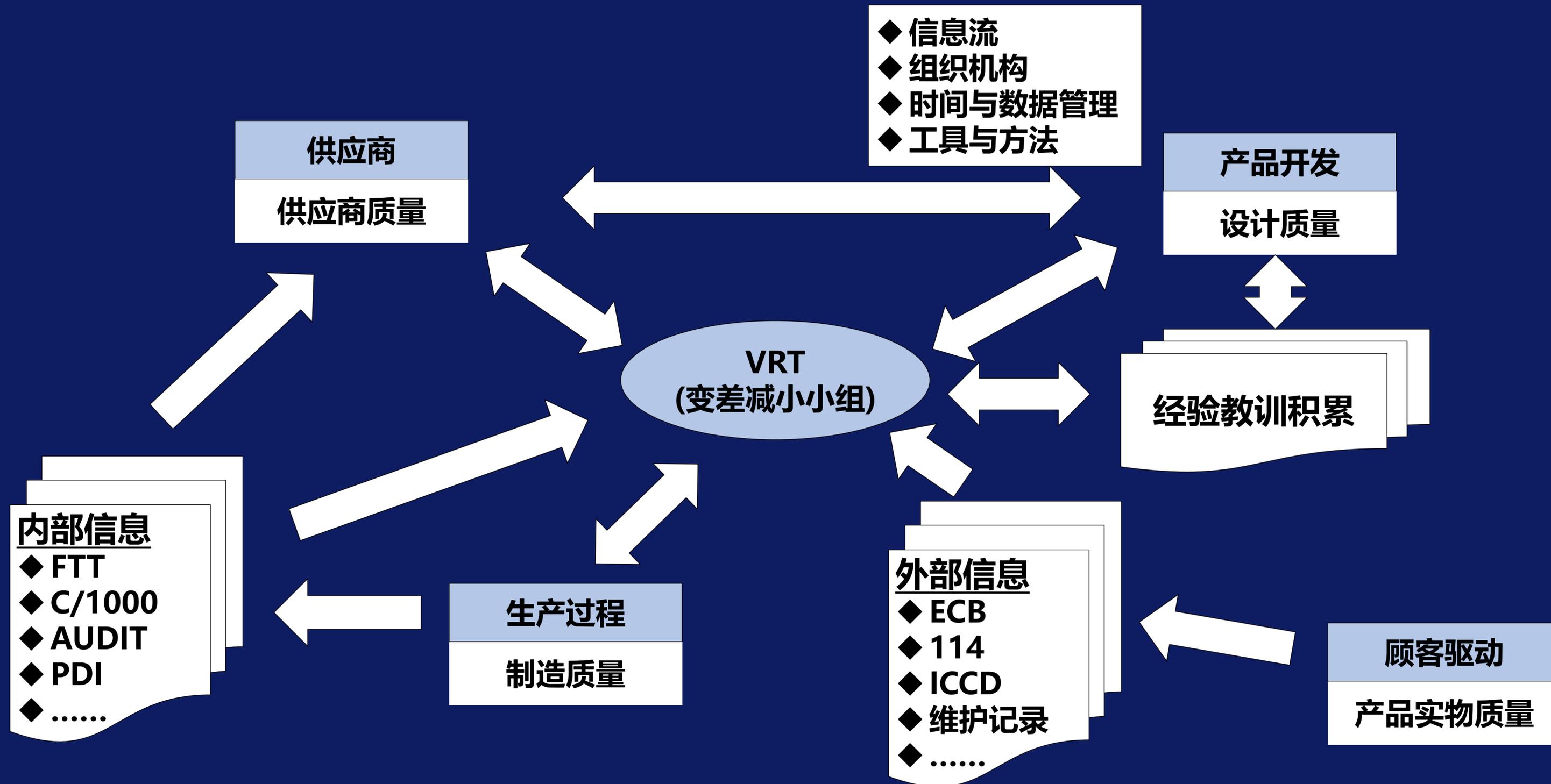
数智人力管理——构建方法



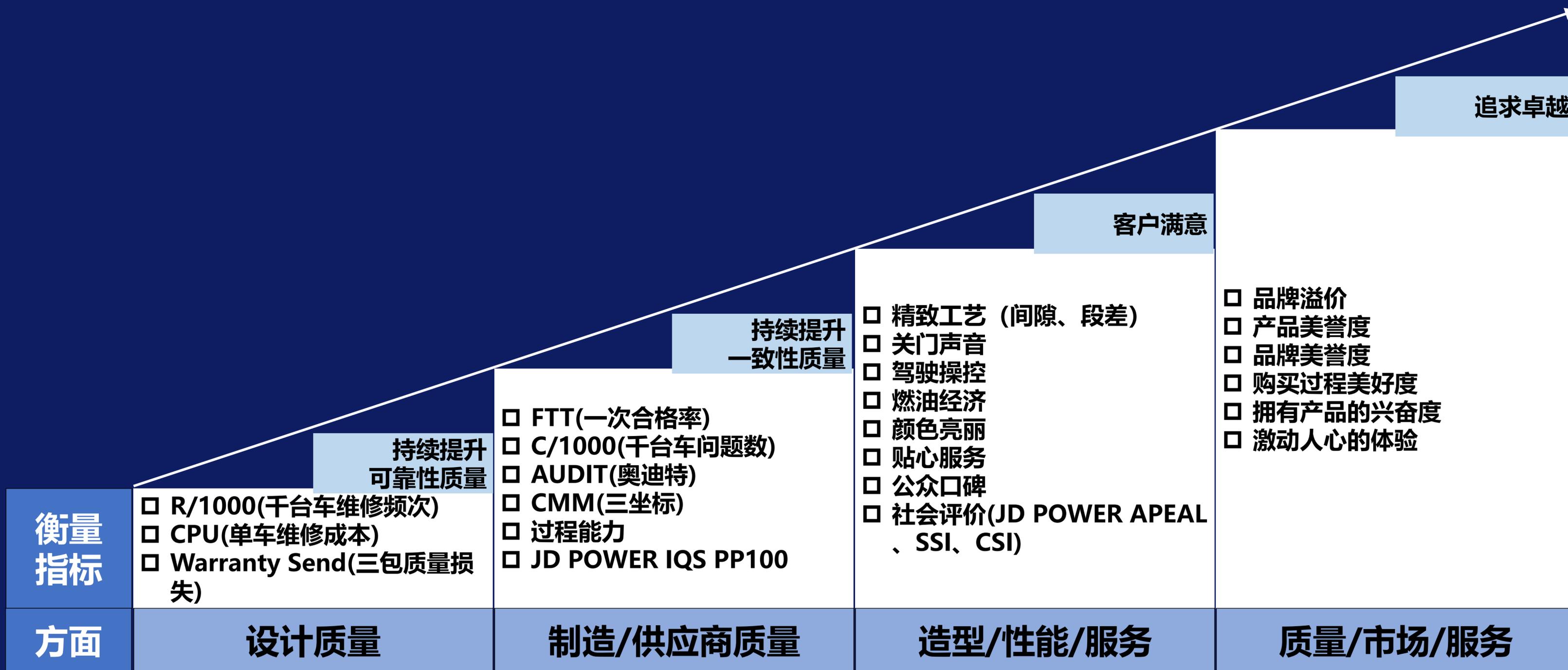
全面质量管理



全面质量持续改进



质量管理指标体系



数智车联网——应用场景

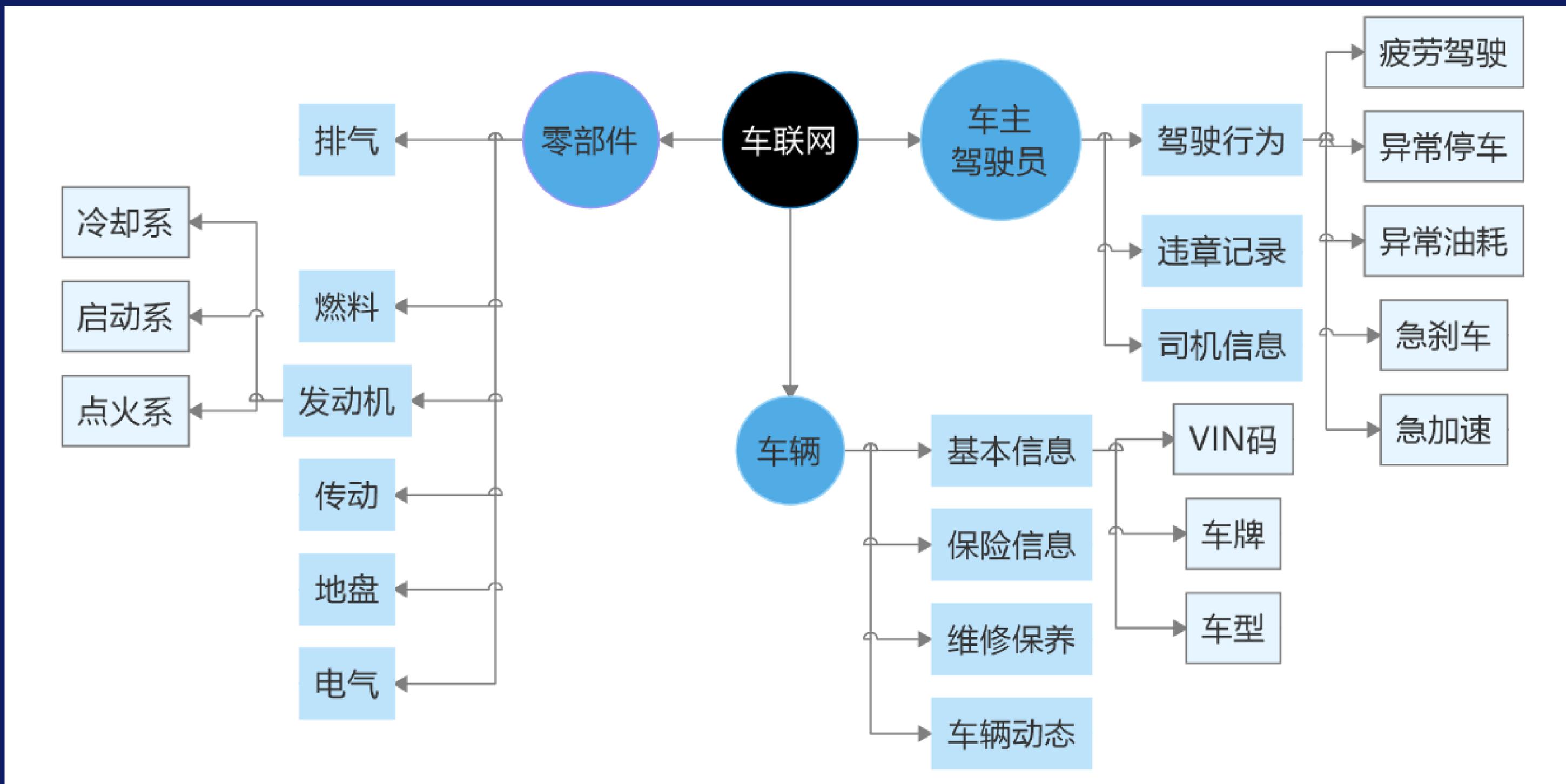


驾驶行为分析

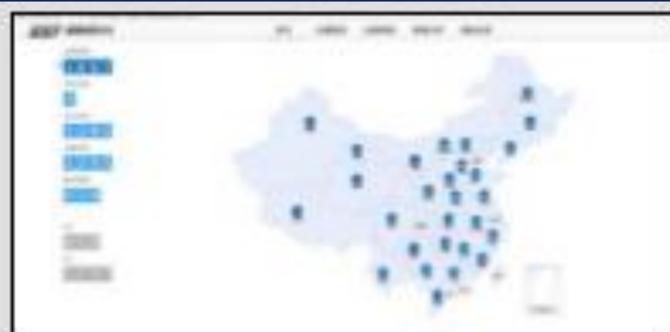


车辆故障诊断与预测

数智车联网——指标体系



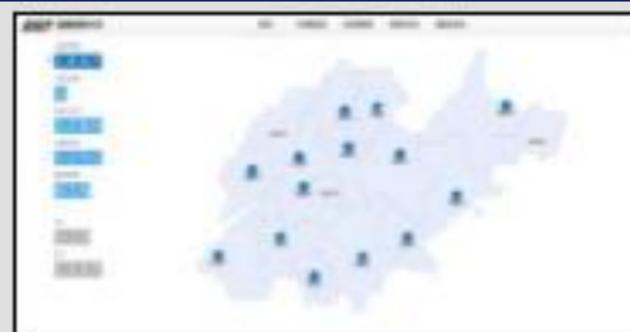
数智车联网——数据可视化分析



整体运营分布



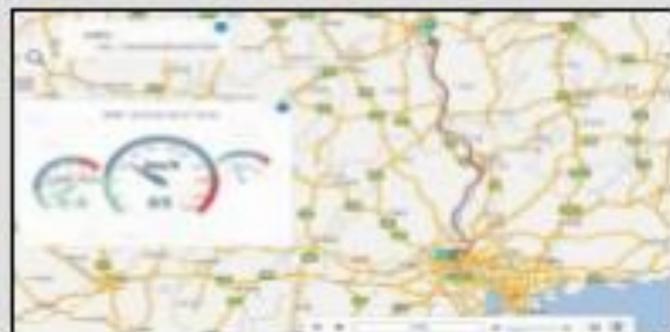
车辆运营热点



区域运营分布



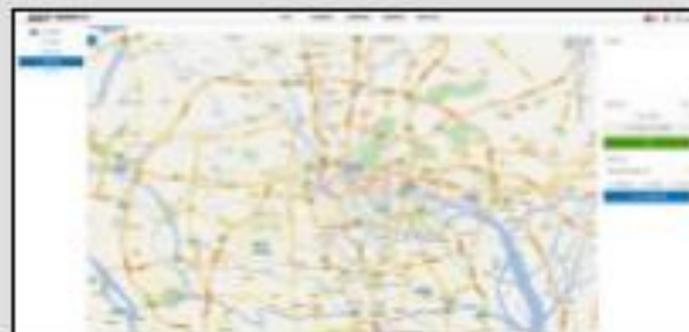
运营里程统计



位置定位及轨迹回放



车辆运行数据采集



金融服务



客户服务



驾驶行为分析



车辆运行故障监控



车辆运行状况分析



车辆报警统计

GRIDSUM 国双

www.gridsun.com



扫一扫 了解国双

国双公司为“国双”商标及产品、服务（包括但不限于在本演示文稿中提及的产品、服务）的合法所有权人和知识产权权利人。

国双公司未授予他人单独享有或与国双公司共同享有上述所有权及（或）知识产权的权利。

任何人不得擅自使用或侵犯“国双”商标及（或）国双公司产品、服务的合法权利。