



探索阿里数据藏宝图——数据地图

阿外(刘应耀)

大数据该关注什么？

阿里有多少数据？如何存储的？

数据与数据间什么关系？

如何找？如何用？

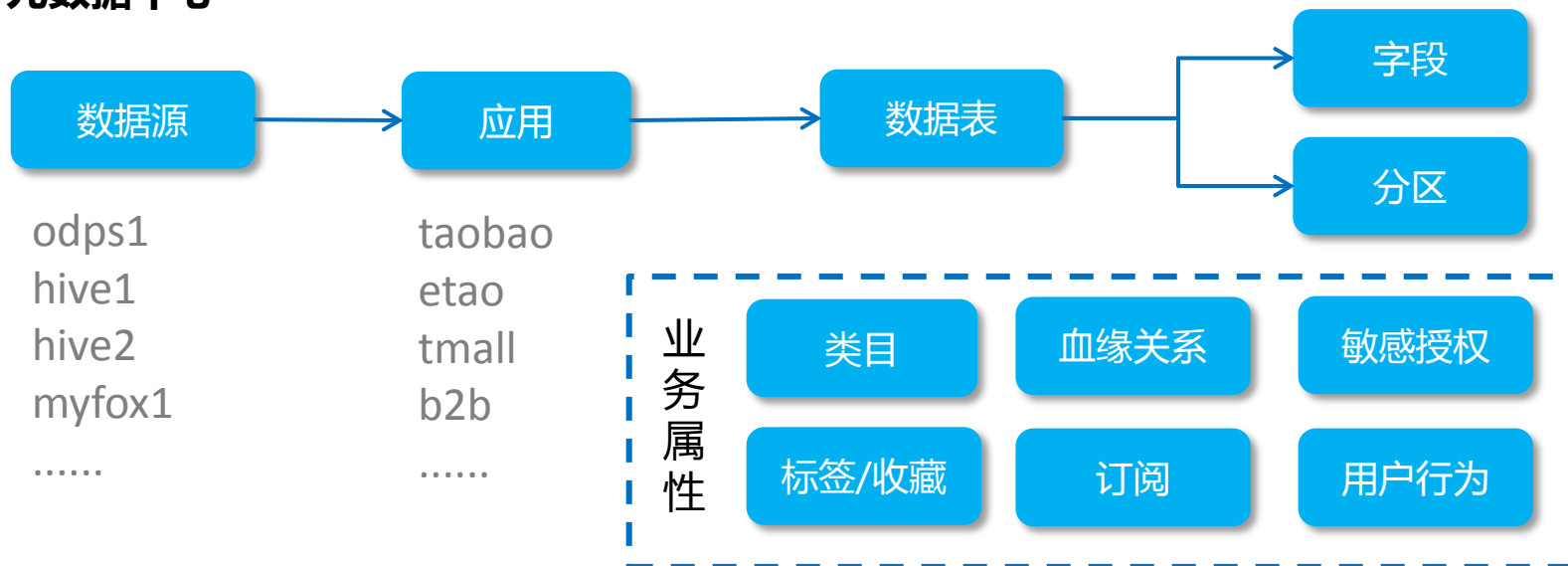
数据的未来？

数据管理模块结构



元数据模块结构

元数据中心



数据源：数据系统集群实例命名，系统种类包括iDB/ODPS/OTS/hive/hBase/myfox/garuda/galaxy/itier；

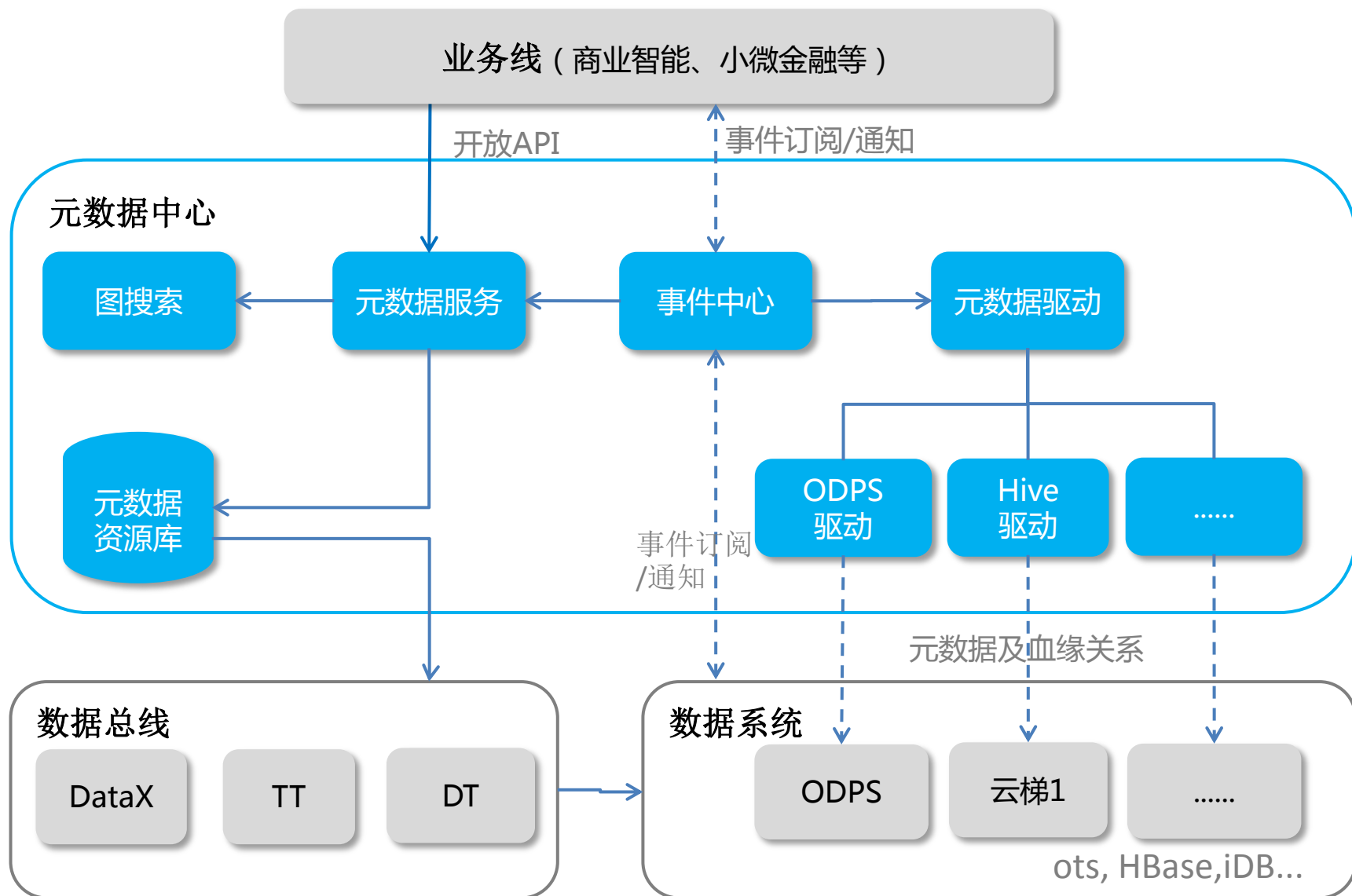
应用：数据源物理上的统一命名空间，如oracle的schema、odps的project、mysql的database等；

数据表：统一的 GUID命名规范，datasource1.app1.table1，直观易用，简化API接口；

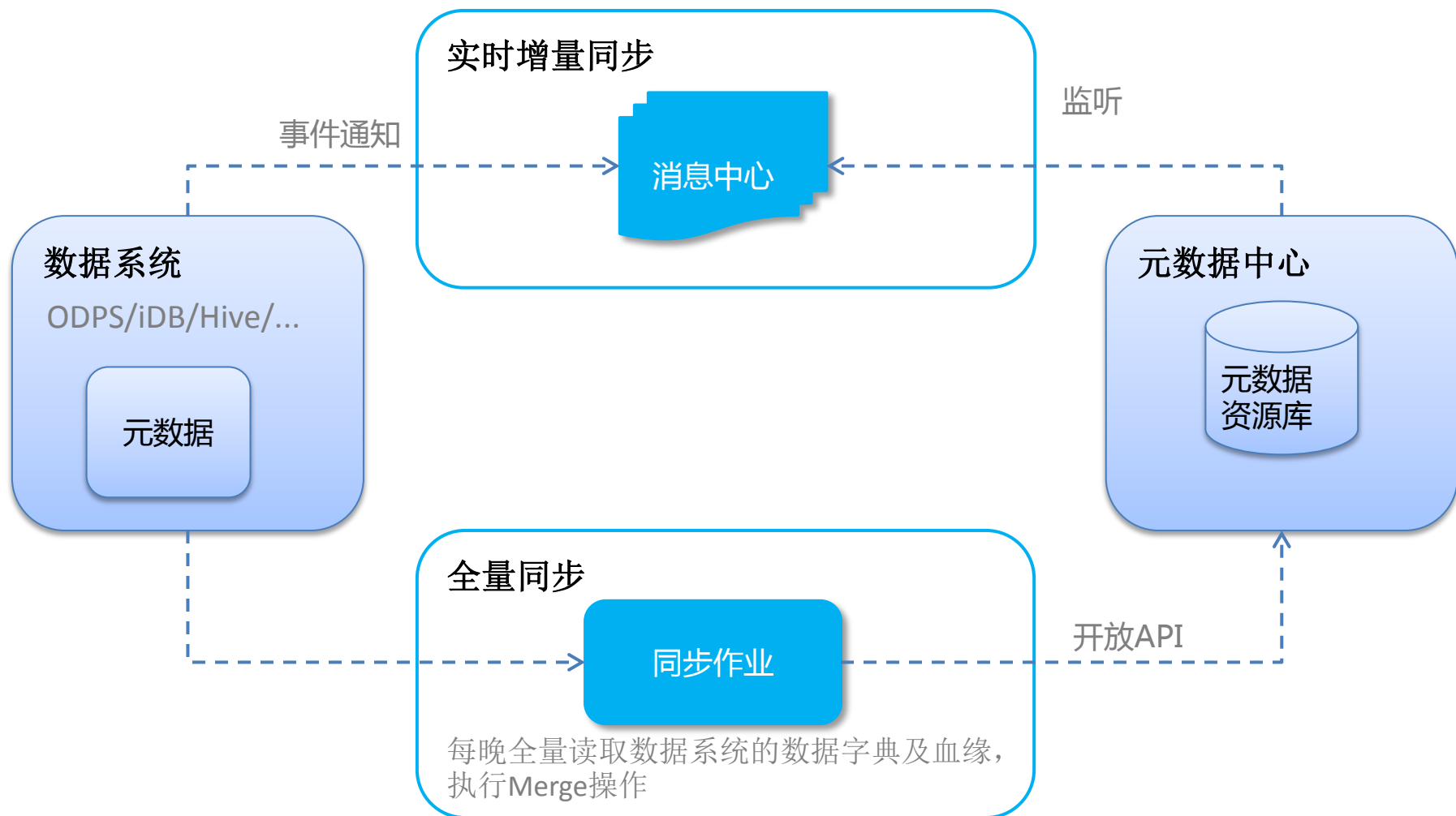
血缘关系：数据流向（表→表，字段→字段），数据关联（外键）；

敏感授权：应用或表的owner可对全表或字段进行敏感级别设置；

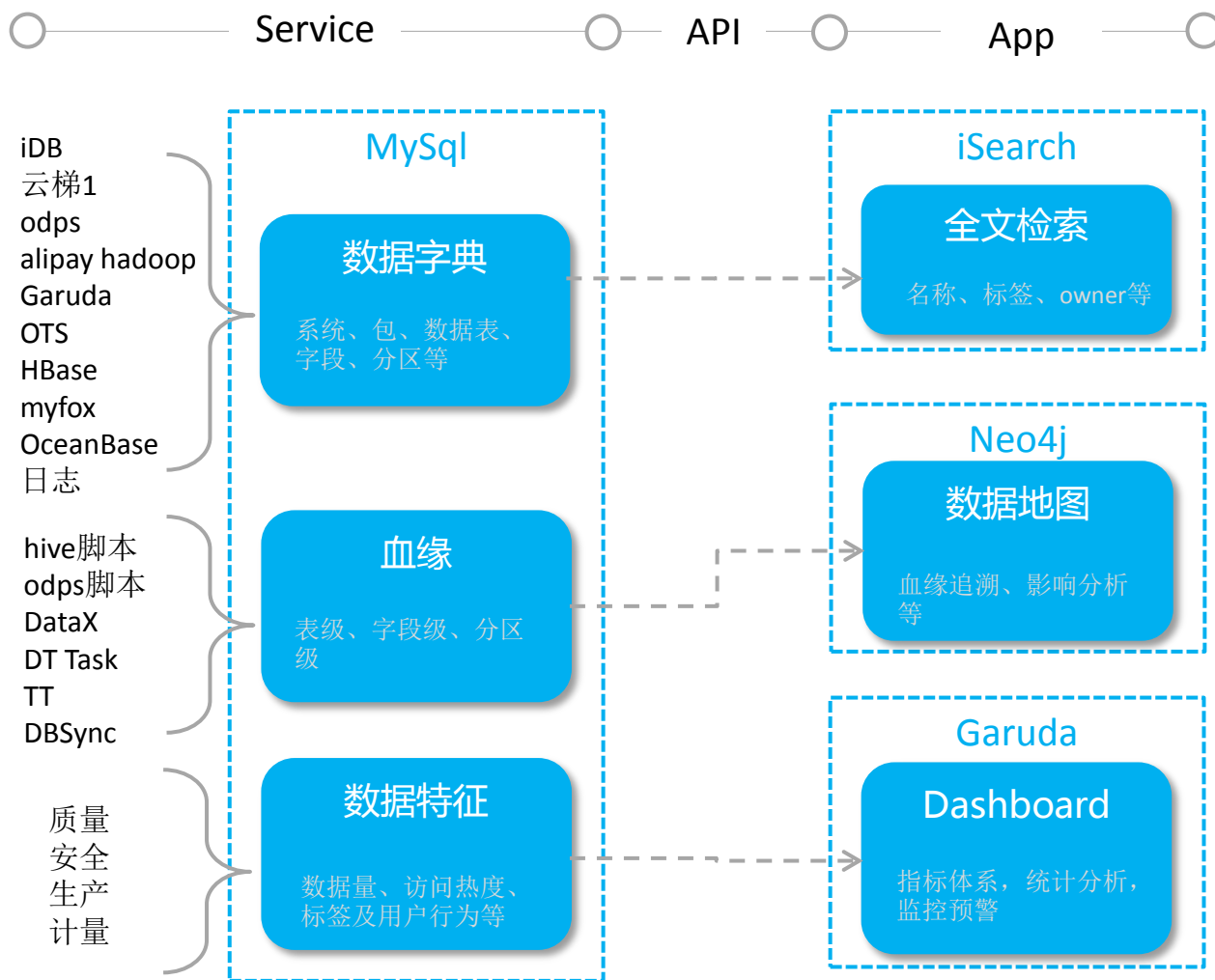
元数据中心



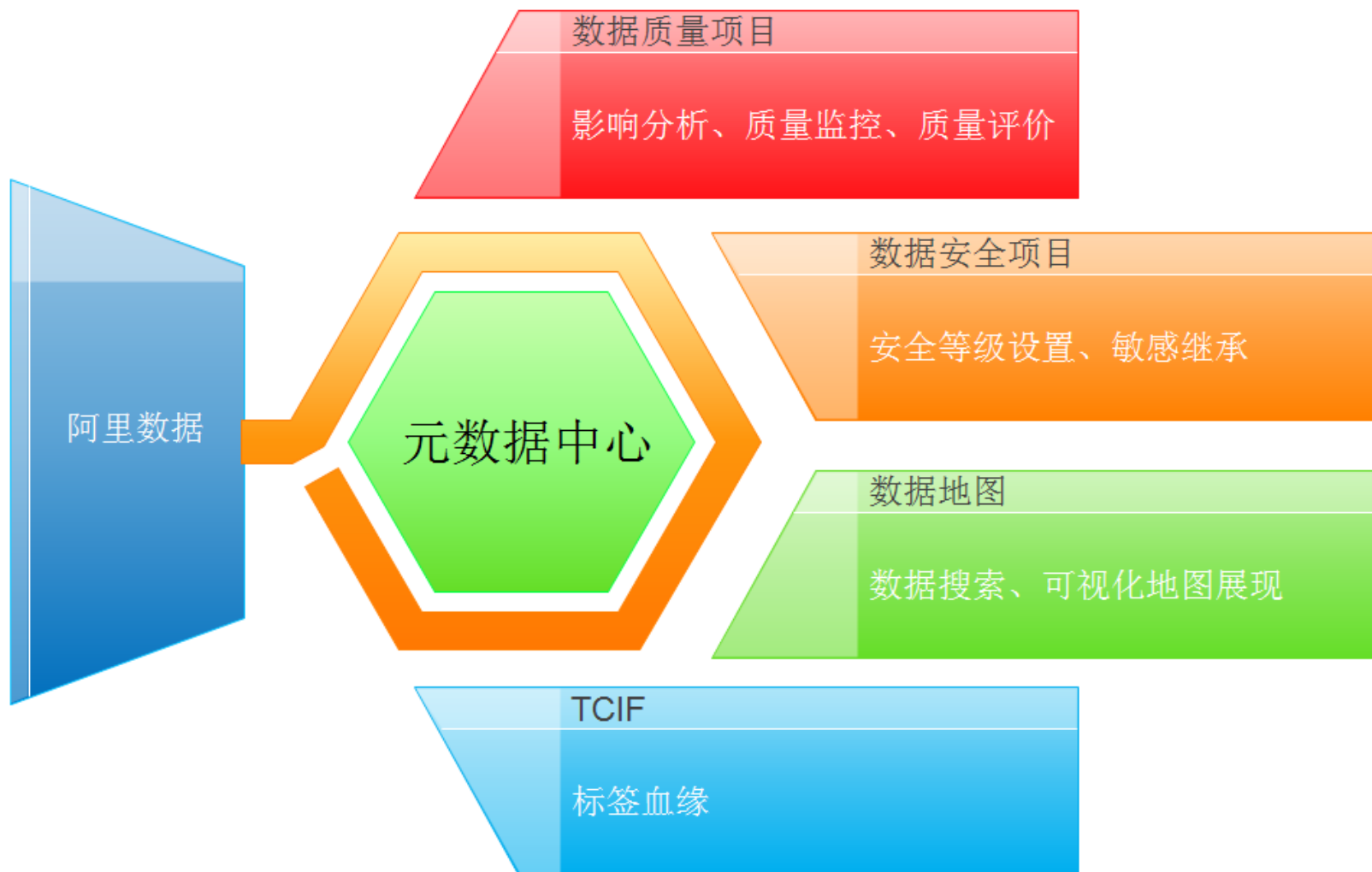
同步方案



系统建设方案



元数据中心应用



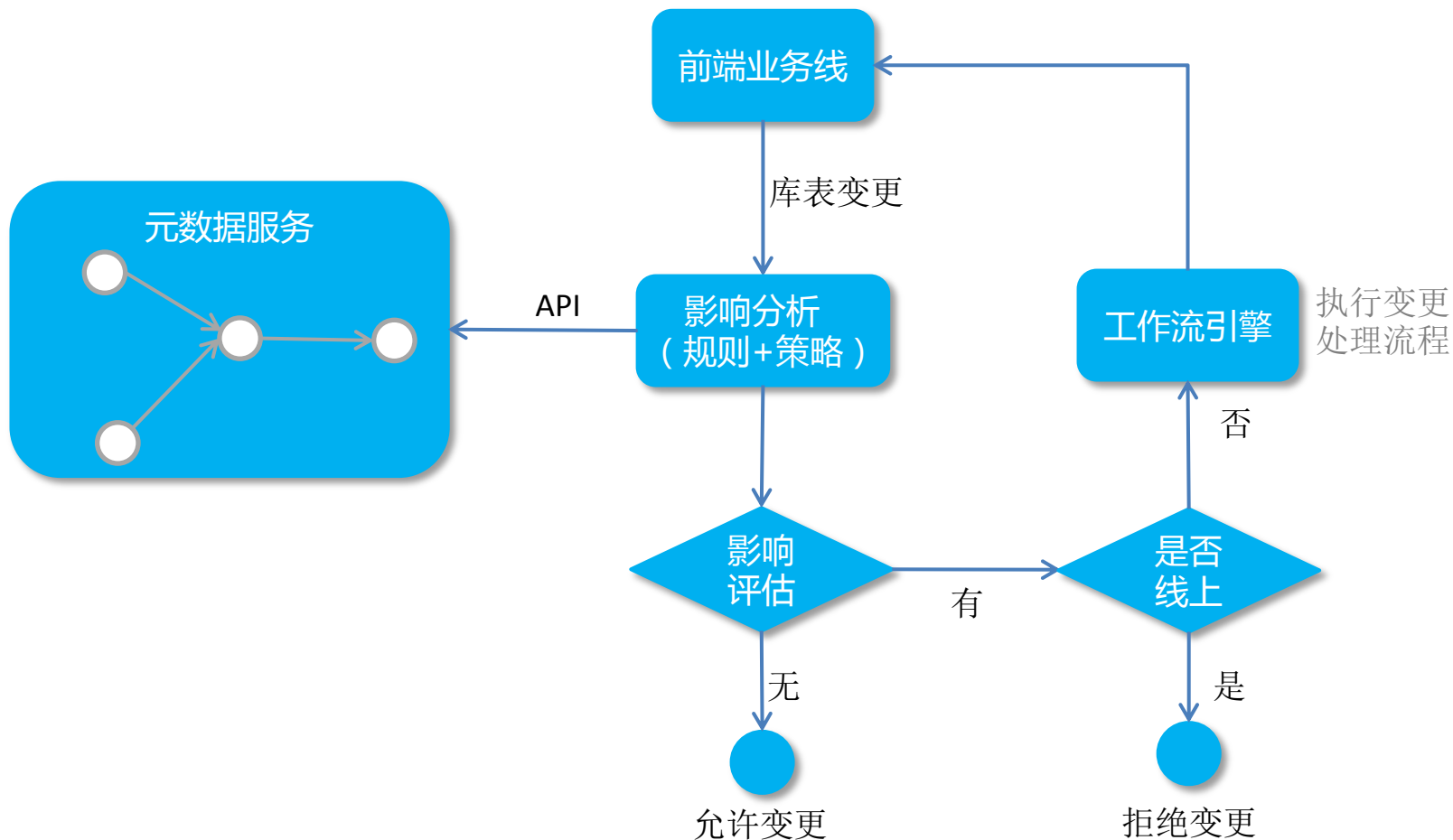
- 信息可互通（-Information）
- 规范可遵循（-Rule）
- 组织可信任（-Organisation）
- 系统可依赖（-System）
- 环境可发挥（-Environment）



血缘应用：影响分析

商业智能、小微金融等要求源数据表变更时有影响评估和变更流程协同。

- ✓ **业务规范约束**：前端业务线发起变更申请；
- ✓ **技术体系保障**：数据系统DDL执行前事件触发影响分析流程；

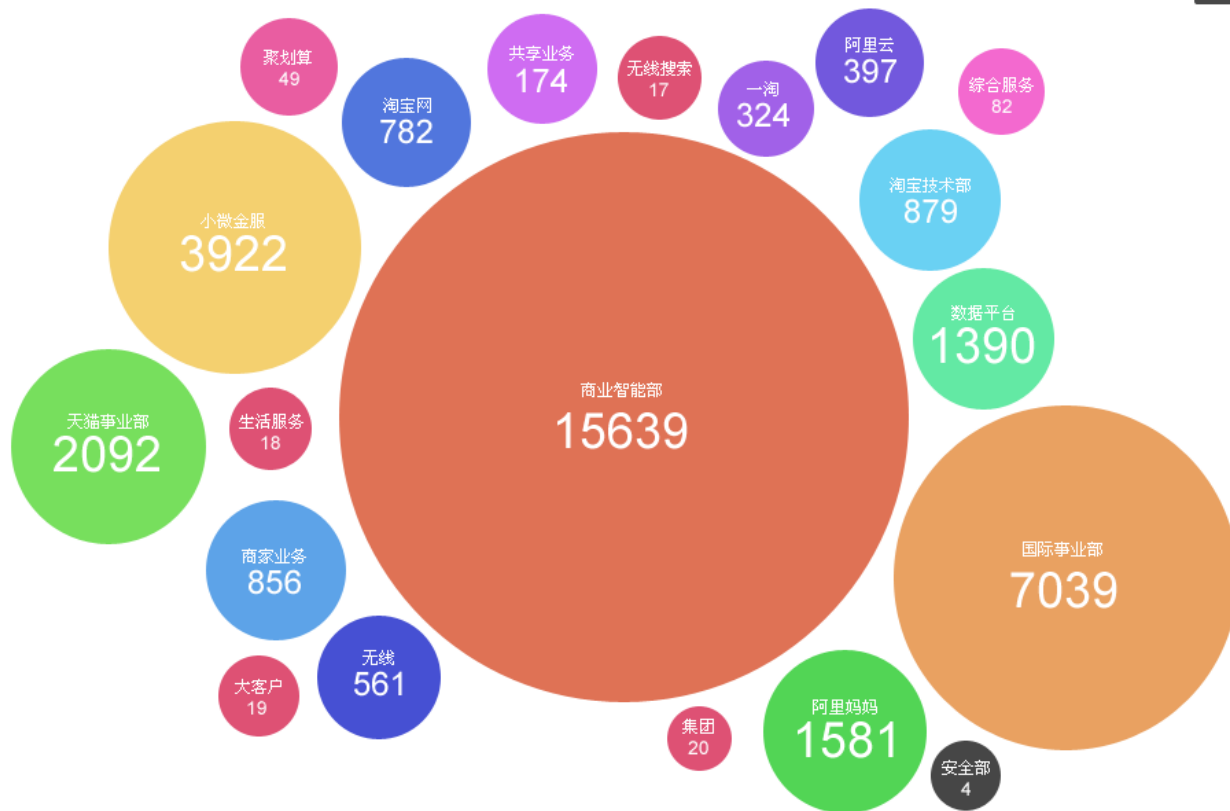


数据地图 (1)

阿里数据地图



高级筛选



阿里数据地图

rkadm



rkadm (100 条结果)

高级筛选

rkadm_agds_rate_discount_ds

agds_差异定价——利率折扣 (所有获贷客户, 按照 cust_i...
GUID: odps.alifin_dw_dev.rkadm_agds_rate_discount_ds...

rkadm_pub_agent_status_log

集团通话记录
GUID: odps.alifin_dw.rkadm_pub_agent_status_log

rkadm_agds_artificial_audit_dd

人工审批信息
GUID: odps.alifin_dw_dev.rkadm_agds_artificial_audit...

rkadm_mntr_datacube_p4_q591

风险应用_监控预警需求_迭代四_数据魔方_电话信息表
GUID: odps.alifin_odps_dw.rkadm_mntr_datacube_p4_q59

rkadm_cc_ledger_bill_d

后台帐账单表
GUID: odps.alifin_dw_dev.rkadm_cc_ledger_bill_d

t_rkadm_masterdata_name_dd_1

客户主题_email属性信息表--临时表1
GUID: odps.alifin_dw_dev.t_rkadm_masterdata_name_dd_

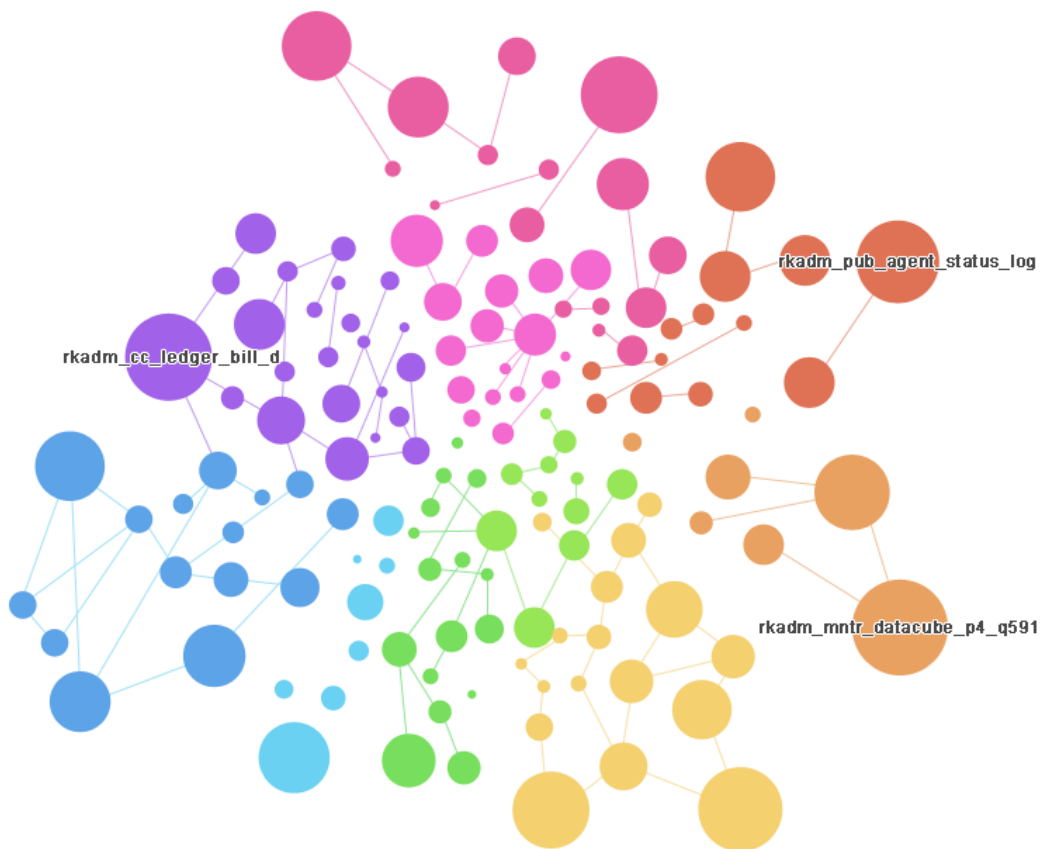
rkadm_agds_guarantee_dd

个人担保信息
GUID: odps.alifin_dw_dev.rkadm_agds_guarantee_dd

v_d_rkadm_agds_rate_discount_ds

差异定价——利率折扣 (所有获贷客户, 按照 cust_id汇总, 再...
GUID: idb2_1583.fdx_agds.v_d_rkadm_agds_rate_discount_ds

rkadm_agds_alifi_alarm_info_temp_2013_05_21



阿里数据地图

rkadm



rkadm (100 条结果)

高级筛选

[rkadm_agds_rate_discount_ds](#)

agds_差异定价——利率折扣 (所有获贷客户, 按照cust_id...
GUID: odps.alifin_dw_dev.rkadm_agds_rate_discount_ds...

[rkadm_pub_agent_status_log](#)

集团通话记录
GUID: odps.alifin_dw.rkadm_pub_agent_status_log

[rkadm_agds_artificial_audit_dd](#)

人工审批信息
GUID: odps.alifin_dw_dev.rkadm_agds_artificial_audit...

[rkadm_mntr_datacube_p4_q591](#)

风险应用_监控预警需求_迭代四_数据魔方_电话信息表
GUID: odps.alifin_odps_dw.rkadm_mntr_datacube_p4_q59

[rkadm_cc_ledger_bill_d](#)

后台帐单表
GUID: odps.alifin_dw_dev.rkadm_cc_ledger_bill_d

[t_rkadm_masterdata_name_dd_1](#)

客户主题-email属性信息表-临时表1
GUID: odps.alifin_dw_dev.t_rkadm_masterdata_name_dd_

[rkadm_agds_guarantee_dd](#)

个人担保信息
GUID: odps.alifin_dw_dev.rkadm_agds_guarantee_dd

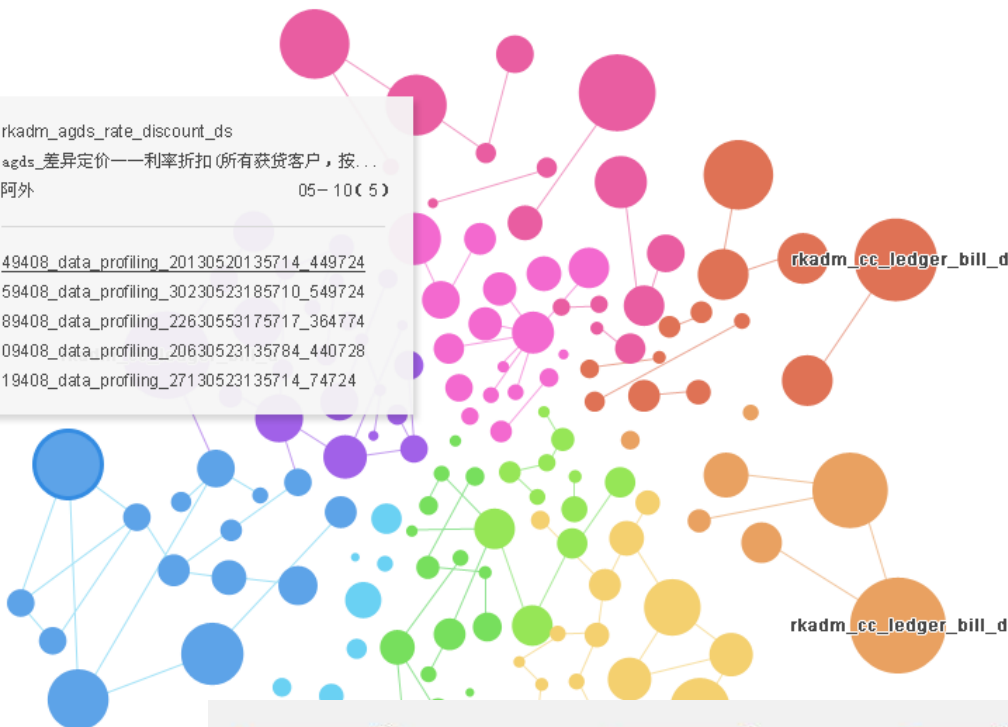
[v_d_rkadm_agds_rate_discount_ds](#)

差异定价——利率折扣 (所有获贷客户, 按照cust_id汇总, 再...
GUID: idb2_1583.fdx_agds.v_d_rkadm_agds_rate_discount_ds

[rkadm_agds_alifi_alarm_info_temp_2013_05_21](#)

rkadm_agds_rate_discount_ds
agds_差异定价——利率折扣 (所有获贷客户, 按...
阿外 05-10(5)

49408_data_profiling_20130520135714_449724
59408_data_profiling_30230523185710_549724
89408_data_profiling_22630553175717_364774
09408_data_profiling_20630523135784_440728
19408_data_profiling_27130523135714_74724





首页

应用开发

数据管理

社区

数据中心

数据字典

监控中心

我的空间

配置管理

数据地图

数据质量

生命周期

rkadm_cc_ledger_bill_d (如有疑问请联系该表顾问: 2161111111111111)

收藏 订阅

中文名:	GUID:		odps.alifin_dw_dev.rkadm_cc_ledger_bill_d		编辑
责任人:	数据源:	odps	数据应用:	odps.alifin_dw_dev	编辑
存储方式:	表容量:		表行数:		
数据粒度:	是否分区表:	Y	分区规则:		
DDL最后变更时间: 2013-05-09 16:10:37		最后数据变更时间:	最后数据查看时间: 0		
访问热度:	未统计	标签:			
描述:	后台帐 账单表				

字段信息

产出信息

已知问题

上游血缘

下游血缘

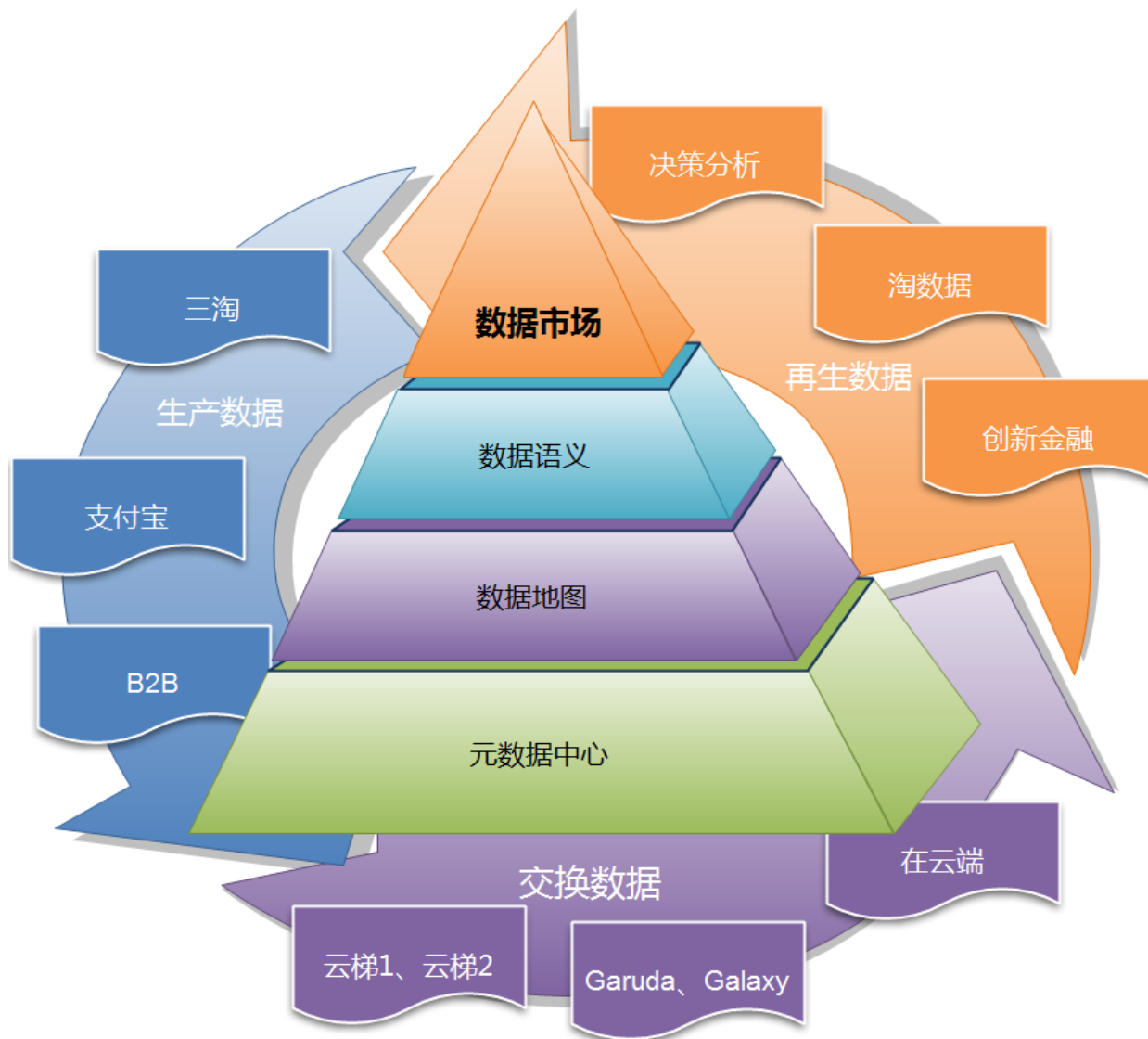
操作纪录

序号	字段名	中文名	类型	描述信息	编辑
0	bill_no		string	账单编号	
1	bill_period		string	如201301, 13年1月的账单	
2	account_no		string	贷记账户ID	
3	prod_id		string	任我付产品码	
4	begin_date		string	账单起始日	

数据地图价值



数据市场



谢谢！

<http://www.alidata.org>

yingyao.lyy@alibaba-inc.com

@阿里之外