阿里云数据安全治理思路

扫尘

2022.11

数据安全治理的误区

堆功能

采集	传输	存储	处理	使用	交换	销毁
数据识别	数据识别	数据识别	数据识别	数据识别	数据识别	数据识别
身份鉴别	身份鉴别	身份鉴别	身份鉴别	身份鉴别	身份鉴别	身份鉴别
访问控制	访问控制	访问控制	访问控制	访问控制	访问控制	访问控制
行为审计	行为审计	行为审计	行为审计	行为审计	行为审计	行为审计
显式授权/声明	传输加密	存储加密	数据脱敏	数据脱敏	隐私计算	安全销毁
端采集行为识别	源/目的端管控	备份恢复	生产/开发态隔离		数据溯源	
端权限检测						

数据安全治理的本质

处理好身份、行为、资产之间的关系

【Example】

电商行业分析师 -> 订单事实表 -> 统计月收入环比。 ✅

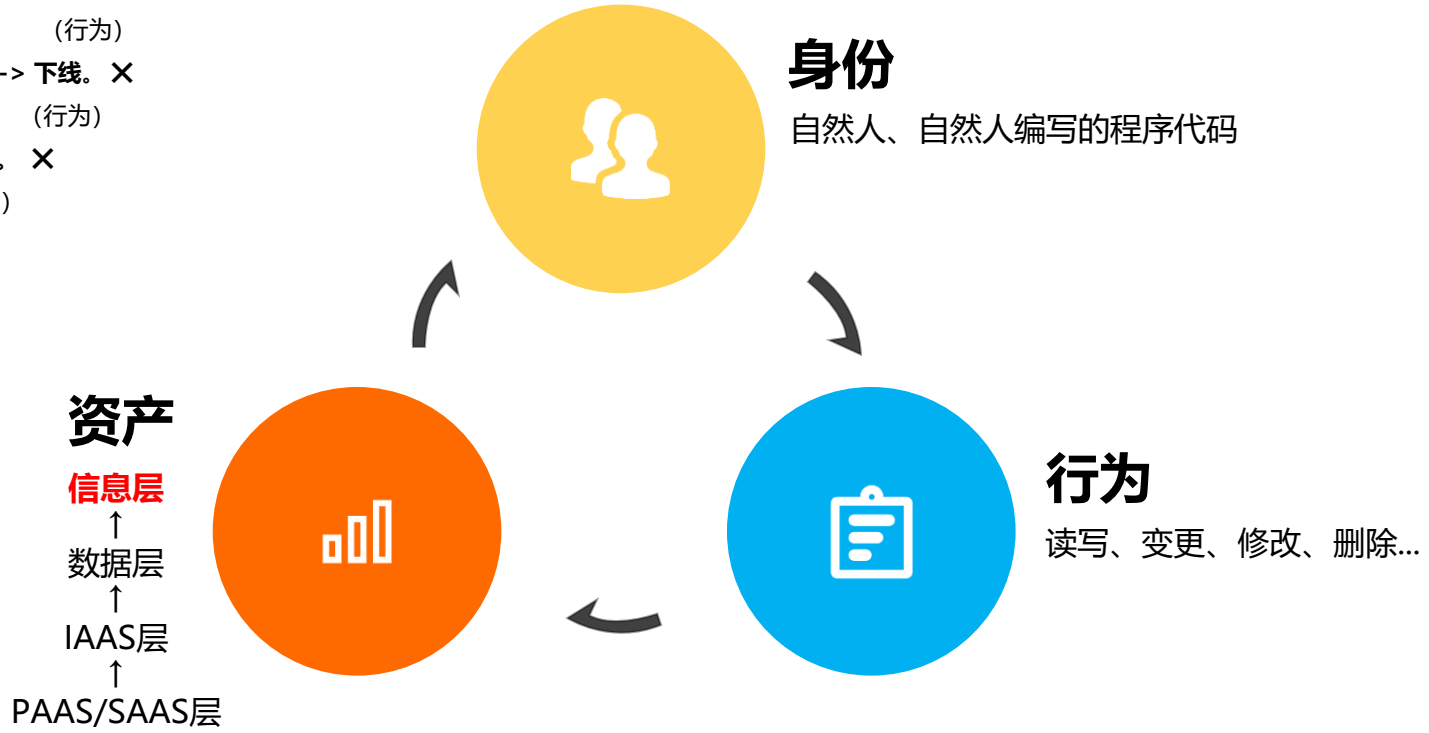
（身份） （资产） （行为）

分析师 -> Hadoop Master -> 下线。 ❌

（身份） （资产） （行为）

分析师 -> Hive组件 -> 重启。 ❌

（身份） （资产） （行为）



数据安全治理思路变化

从粗暴式隔离与防御，到精细化的管理与融合

数据安全治理思路

	从前	现在
目标	南北向安全（外部威胁）	东西向安全（内部威胁）
对象	外部黑客入侵	内部人员行为
理念	隔离	分级分类+合理流动
手段	边界安全技术	过程安全技术
实施	技术、管理分离	技术、管理融合

零信任为什么火起来了？

数据安全治理体系

愿景是什么？ 目标是什么？ 谁来做？

组织建设

- 决策层（高管，CIO）
- 管理层（数据安全管理团队，CISO）
- 执行层（数据安全运行团队、技术团队）

审计监督层

怎么做？

制度流程

- 方针 总纲
- 管理制度 管理办法
- 流程 规范 指南 模板
- 计划 报告 记录 日志

用什么做？

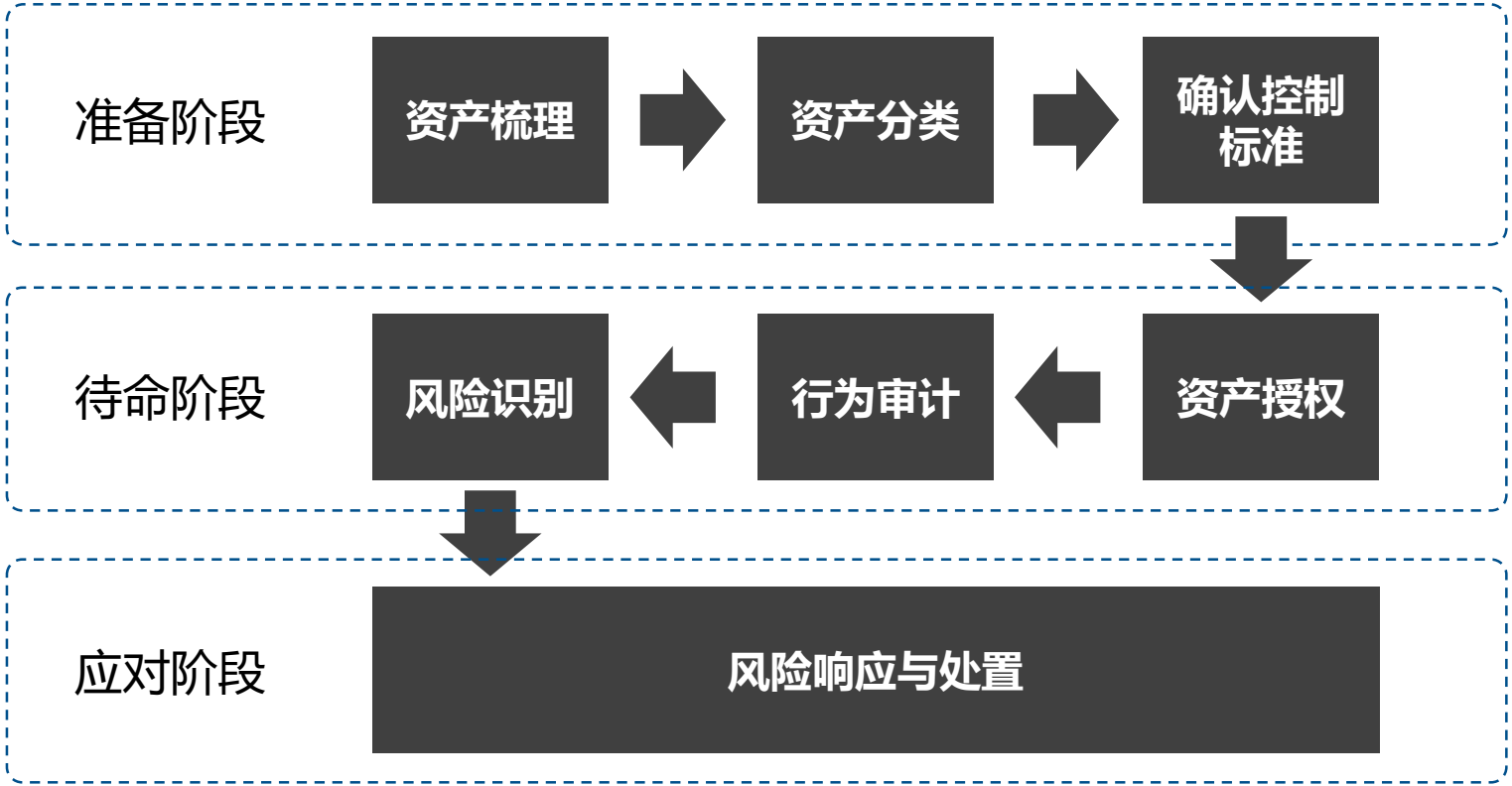
技术工具

- 分级分类
- 权限管理
- 审计
- 脱敏加密
- 开发管理
- ……

数据全生命周期安全治理

数据采集（传输） | 数据存储 | 数据处理 | 数据交换 | 数据销毁

数据安全治理核心链路



资产梳理

从有形资产到无形资产

什么是应被保护的资产?

判断标准

（1）如遭到滥用、泄露则会对企业信息的**机密性（Confidentiality）、完整性（Integrity）、可用性（Availability）造成损害。**

（2）如遭到滥用、泄露则会对**企业业务连续性造成负面影响。**

有形资产	无形资产
设备	IAAS：ECS/VPC/高速通道
文件	PAAS：RDS/容器
人	SAAS：算法/DataWorks
	DAAS：API

资产梳理

从有形资产到无形资产

工作空间类	业务空间	数仓分层空间		
资源类	引擎实例	资源组实例		
数据类	表	UDF	数据源	数据服务API
任务类	数据计算任务	数据传输任务		
规则类	数据质量规则	任务报警规则		
策略类	审批策略	查询结果策略		
	数据识别策略	风险识别策略		
建模类	数据标准	维度/事实模型	指标	

数据源

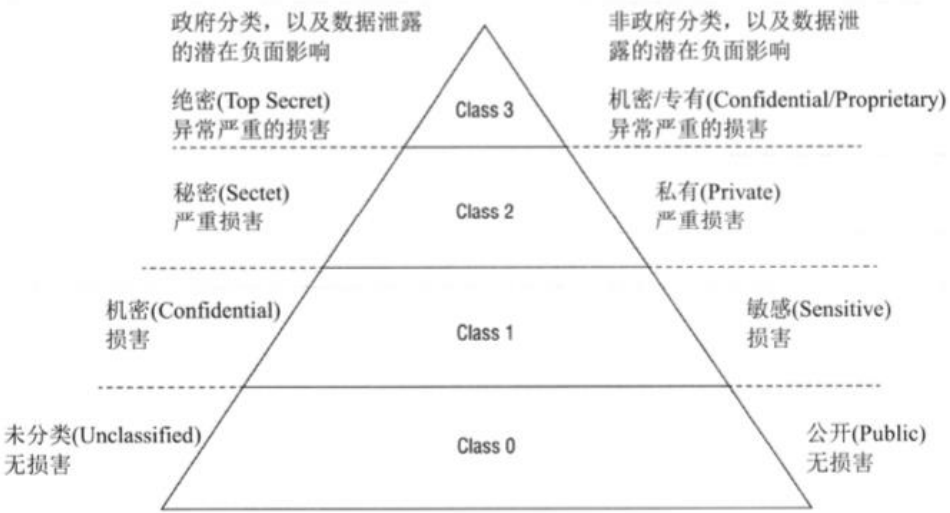
表/函数/任务/数据API

数据质量规则

模板名称
表行数,1,7,30天波动率
表行数，7天平均值波动率
表行数，30天平均值波动率
表行数，1天波动率
表行数，7天波动率

资产分类

资产分类标准：来自CISSP



1.个人身份信息（PII）

任何可以识别个人的信息，如姓名、社会保险账号、出生日期、出生地点、母亲的娘家姓或生物识别记录，也包括关联信息，如医疗、教育、金融和就业信息等。

2.受保护的健康信息（PHI）

与个人有关的任何健康信息，如身体健康状况、病例信息、医疗费用等。从另一方面来说，PHI也属于PII。

3.专有数据（Proprietary Data）

影响组织核心竞争力、一旦泄露会对组织造成损害的数据，典型例子有设计图纸、药物配方、客户信息等。

资产分类

数据分类标准

阿里云

国家标准：

《信息技术 大数据 数据分类指南 GB/T 38667-2020 》

《信息安全技术 健康医疗数据安全指南 GB/T 39725-2020》

行业标准：

《个人金融信息保护技术规范 JR/T 0171-2020》

《金融数据安全 数据安全分级指南 JR-T 0197-2020》

《基础电信企业重要数据识别指南 YD/T 3867-2021》

《基础电信企业数据分类分级方法 YDT 3813-2020》

《证券期货业数据分类分级指引 JR/T 0158-2018》

地方标准：

《贵州省政府数据分类分级指南 DB 52/T 1123-2016 》

《重庆市公共数据分类分级指南》

确认控制标准

不同分类不同标准

	1 采集	2 传输	3 存储	4 处理	5 使用	6 交换	7 销毁
机密	不采集	不传输&阻断	不存储	-	-	不交换共享	-
私有	脱敏采集 声明&授权采集 183日审计.日审	高层审批 传输加密 双因素身份鉴别 183日审计.日审	180日内快照 LifeCycly冷存归档 存储加密 不删除	细粒度授权 动/静态脱敏 生产/开发隔离 183日审计.日审	阻断下载 服务化调用 无密登录 183日审计.日审	求交&联邦 双因素身份鉴别 脱敏共享 183日审计.日审	安全销毁
敏感	声明&授权采集 183日审计.月审	一线主管审批 传输加密 单因素身份鉴别 183日审计.月审	90日内快照 LifeCycly冷存归档 存储加密 不删除	表级别授权 生产/开发隔离 183日审计.月审	申请下载&下载告警 服务化调用 授权查询 183日审计.月审	鉴权共享 明文共享 单因素身份鉴别 183日审计.月审	安全销毁
公开	直接采集 不审计	直接传输 不审计	30日内快照 默认LifeCycle删除	直接处理 不审计	直接下载 服务化调用 直接查询 不审计	公开共享 不审计	不审计

风险级别低：授权　　风险级别高：授权 + 响应处置

资产授权

登陆地&登录方式控制

资产授权

事前授权、事中授权

自然人/程序

开发人员

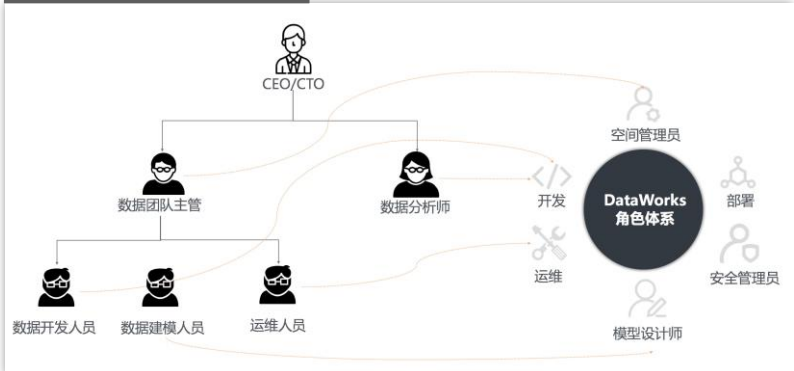
运维人员

分析师

CISO

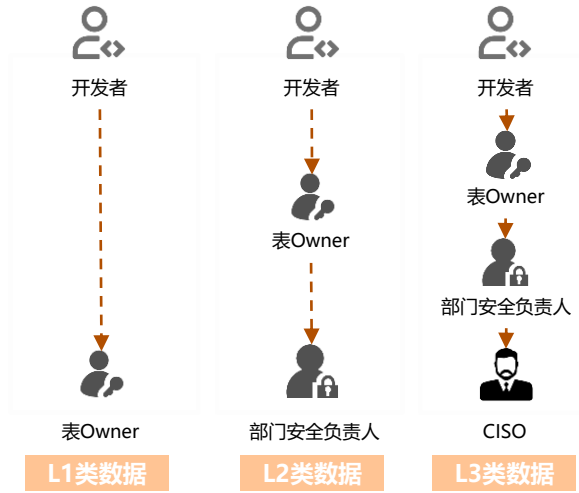
生产账号身份

RBAC&ABAC&Policy




```
"Statement":[
  {
    "Effect":"******",
    "Principal":"****",
    "Action":[
      "odps:CreateTable",
    ],
    "Resource":"acs:odps:*:****
    "Condition":{
      "DateLessThan":{
        "acs:CurrentTime":"2013-11-11T23:59:59Z"
      }
    }
  }
]
```

按需申请

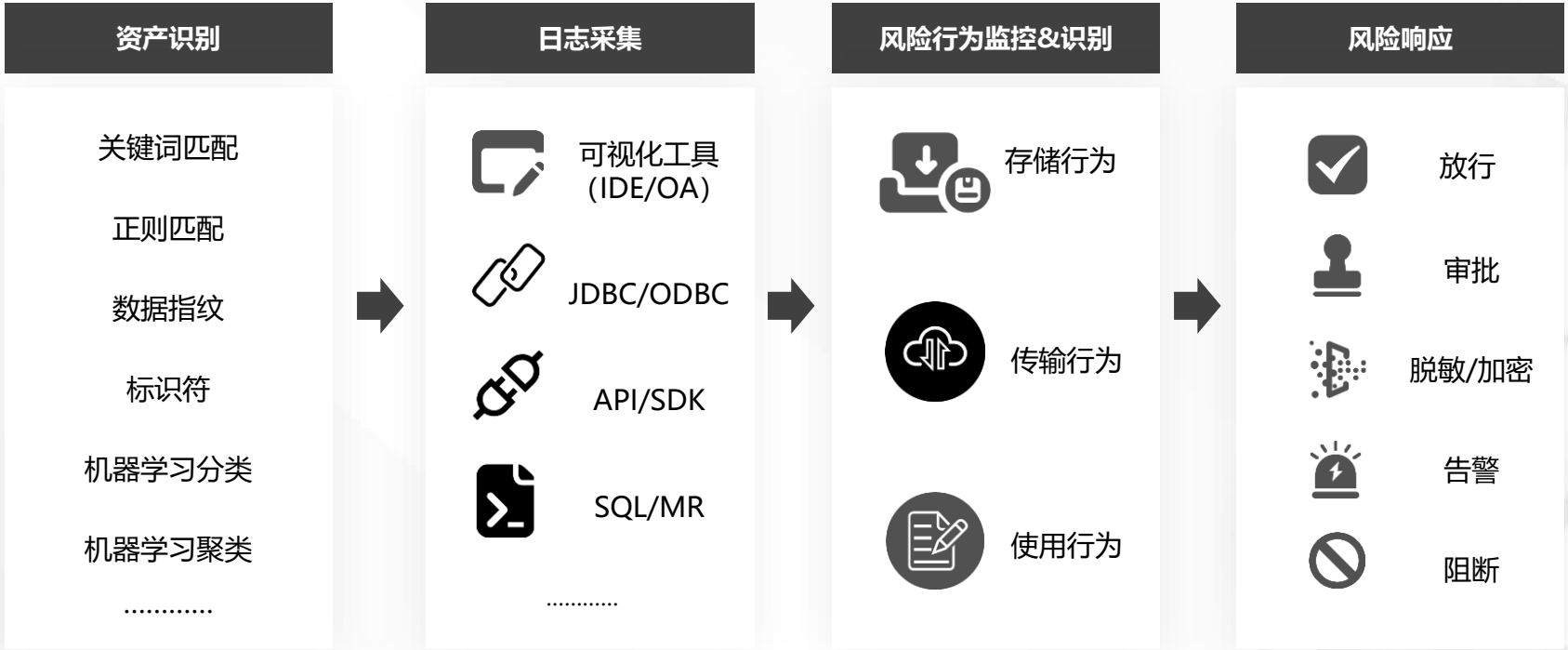


操作审计

谁?	在哪里?	什么时候?	什么方式?	什么资产?	什么操作?	结果如何?	风险如何?	怎么处理?
Who?	*Where?*	*When?*	*Method?*	*Asset?*	*R?W?X?*	*Result?*	*Risk?*	*Action?*
开发人员	外网VPN (47.xx.xx.xx)	2022-06-01 23:59:59	Online IDE	db.user_info	r:select .. form .. where...	成功	**4** **(top secret&Non Office Hour)**	**Alert**
运维人员	外网VPN (121.xx.xx.xx)	2022-05-29 10:00:00	OPEN API	host_1	x:Restart	失败	2	Notice
运维人员	外网VPN (121.xx.xx.xx)	2022-05-29 01:00:00	JDBC	db.user_info	x:Delete db	成功	**5** **(top secret&Non Office Hour&Delete)**	**Block**
分析师	外网VPN (59.xx.xx.xx)	2022-06-12 10:00:00	Online IDE	db.table_2	r:select .. form .. where...	成功	1	Null
某服务身份	办公网	2022-06-10 13:00:00	Inner API	host_2	x:CallBack	成功	1	Null
分析师	办公网	2022-06-12 10:00:00	Online IDE	db.table_2	x:Download	失败	**3** **(top secret&Download)**	**Approval**

风险识别&响应

常见风控范式



风险识别&响应

常见风控范式

风险类型

内部人员 | 外包人员 | 合作伙伴

更多数据来源 | 更多基础应用

更多外部合作 | 更多跨域流动

数据流动保护风险

数据业务过程风险

数据权限风险 | 数据导出风险

数据访问风险 | 数据操作风险

个人隐私 | 商业机密 | 国家机密

风险发现

利用统计模型、机器学习、规则策略、基线分析等方法构建安全预置并检测异常行为。

以风险场景为核心制定不同分析规则和策略组合

从多种数据源汇总数据

风险处置与响应

数据操作行为审计

数据访问行为

数据导出行为

数据操作行为

风险总览

风险检测趋势

风险类型分布

风险分布排名

风险明细

风险明细查询

风险触发条件

风险处置标记

放行

审批

去标识化

告警

阻断

DataWorks安全能力概览