

数据资产运营人员能力要求

1 范围

本文件规定了数据资产全生命周期中数据资产运营人员能力的的能力要求，以支持实现组织目标。
本文件适用于数据资产运营人员开展数据资产运营工作。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 33173 资产管理 管理体系 要求

GB/T 25000.2 《系统与软件工程系统与软件质量要求和评价（SQuaRE）第 24 部分：数据质量测量》

GB/T 37550 《电子商务数据资产评价指标体系》

3 术语和定义

下列术语与定义适用于本文件。

3.1

数据 data

信息的可再解释的信息化表示，以适用于通信、解释和处理。

注：可以通过人工或自动手段处理数据。

[来源：GB/T 37550-2019, 2.1]

3.2

数据资产 data assets

以数据为载体和表现形式，能够持续发挥作用并且带来经济利益的数字化资源。

注 1：数据资产能够为组织带来有潜在价值或实际价值。

注 2：数据资产能够估值、交易，并以货币计量。

注 3：数据资产包含结构化数据、非结构化数据和半结构化数据。

[来源：GB/T 37550-2019, 2.4]

3.3

抽取转换加载技术 Extract Transform Load

指用户从数据源抽取出所需的数据，经过数据清洗，最终按照预先定义好的数据仓库模型，将数据加载到数据仓库中去，简称ETL。

3.4

数据资产指标体系 data asset indicator system

指相对独立又相互联系的统计指标所组成的数据资产。

4 能力要求概述

4.1 数据资产运营人员能力从制订目标到最后基于目标进行产品评估与运营优化，形成数据运营闭环的能力支撑，应与GB/T 33173 的要求相一致。见图 1。

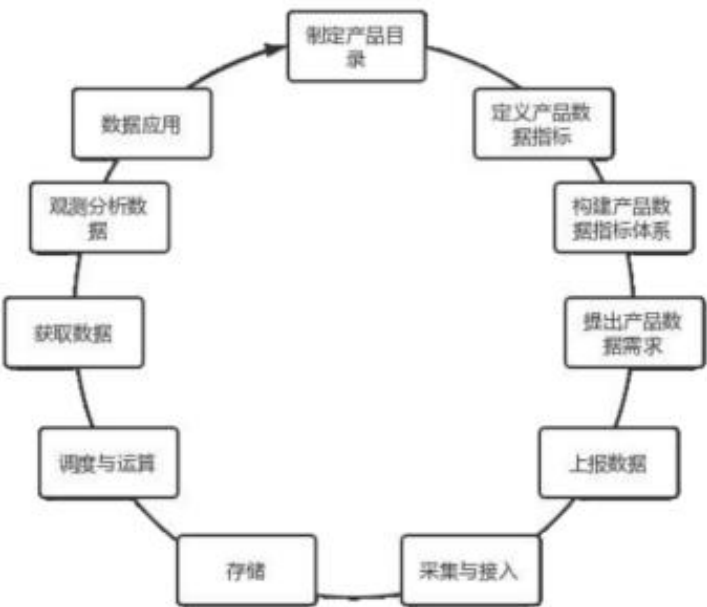


图 1 数据运营流程图

4.2 流程和规范应主要包括以下 7 个层次（见表 1），数据资产全生命周期应建立数据统一流程规范，建立与组织目标一致的运营方式。构建统一的数据中心，开展数据仓库建设，通过运营产生数据价值，通过数据驱动组织各项工作开展。

表1 流程和规范层次图

层次名称	具体内容
数据采集	数据需求、数据上报、数据传输、数据存储、调度运算
业务理解	产品场景、业务逻辑、数据定义、指标体系
数据模型	业务模型、计费模型、用户模型
分析模型	漏斗模型、增长模型、流失模型、催付模型、健康模型
数据指标	活跃数据、留存数据、渠道数据、收入数据、环境数据
数据挖掘	相关分析、关联分析、聚类分析、决策树、回归分析
数据意识	分析工具、分析案例、数据应用、行业数据、数据决策

5 产品目标制定能力

数据资产运营人员应根据要求设定是数据运营的起点，设定数据资产运营后进行评估的指标，建立闭环。目标可以根据组织发展、行业发展、市场分析、历年发展走势、转化规律等分析得出。目标的制定可用SMART原则来衡量：

- a) S(Specific, 具体) 指工作指标要具体可评，不能笼统
- b) M(Measurable, 可度量) 指绩效指标是数量化或者行为化的，验证这些绩效指标的数据或者信息是可以获得的
- c) A(Attainable, 可实现) 指绩效指标在付出努力的情况下可以实现，避免设立过高或过低的目标
- d) R(Relevant, 相关性) 与工作的其它目标具有相关性
- e) T(Time-bound, 时限) 注重完成目标的特定期限

6 数据资产指标定义能力

数据资产指标是反应数据运营能力的具体的数值，需要对数据指标给出明确定义，还包括数据上报方法、计算公式等。数据资产运营人员应能：

- a) 根据产品目标来选择数据指标，定义产品指标体系，协调各相关方；
- b) 设定清晰的数据指标，并且有据可查，不会引起数据解读的理解差异；
- c) 根据产品在不同生命周期调整关注的数据指标侧重点。

7 数据资产指标体系构建能力

数据资产运营人员应能在数据指标提出的基础上，构建逻辑进行指标的归纳整理并体系化，指标体系建设可参考GB/T 37550 开展。如图 3。



图 3 数据资产指标体系图

8 确认数据资产需求能力

8.1 数据资产运营人员应能推动组织业务开展的指标体系的建立，并组织各职能部门根据不同时期的特点和要求，提出对应的数据需求。

8.2 数据资产运营人员应能制定需求文档的模板，方便相关数据上报开发、协调数据平台相关人员进行沟通，开展数据建设。

9 数据上报能力

数据资产运营人员应根据数据需求和数据上报规范，组织完成上报功能开发，将数据及时上传到数据服务器，建立数据上报通道。

10 数据采集与接入能力

10.1 数据统一

10.1.1 多个数据源的统一。一般实际的应用过程中，存在不同的数据格式来源，需要把数据源进行统一转化

10.1.2 采集的实时高效。由于大部分系统都是在线系统，对于数据采集的时效性要求会比较高

10.1.3 脏数据处理。对于一些会影响整个分析统计的脏数据，需要在接入层的时候进行逻辑屏蔽，避免影响后续的统计分析和应用

10.2 数据采集

数据资产运营人员应能通过接口创建开展采集数据，应能考虑数据字段的拓展性、数据采集过程中的ETL数据清洗流程、客户端数据上报的正确性校验等。数据采集分为两步：

- a) 数据从业务系统上报到服务器并存储到文件系统；
- b) ETL 通过抽取、转换、加载把日志从文本中，基于分析的需求和数据维度进行清洗，然后存储在数据仓库中。

11 存储与计算能力

数据资产运营人员应能在完成数据上报和采集和接入之后，实现数据的有效存储。包括：

- a) 数据安全性。很多数据是不可恢复的，数据存储的安全可靠永远最重要，要投入最多的精力来关注。
- b) 数据计算和提取的效率。做为存储源，会面临很多数据查询和提取分析的工作，需要确保实现效率。
- c) 数据一致性。存储的数据主备要保证一致性。

12 获取数据能力

数据资产运营人员应能协助相关方从数据系统获得数据，常见方式包括数据报表和数据提取。要求包括：

- a) 提供连续周期的查询功能，报表要提供查询的起始时间，可以查看指定时间范围内的数据；

- b) 对一段时间范围内的数据能够分段或汇总，能够对不同阶段进行比较；
- c) 查询条件与维度相匹配，提供与维度对应数目的查询条件，尽量满足每个维度都能分析；查询条件要提供开、合，以及具体值的过滤功能；查询条件的顺序，尽量与维度的顺序对应，最好按从大到小的层次；
- d) 图表与数据要一致，图表显示的趋势，要与相应的数据一致，避免数据有异议；有图就必须有数据，但有数据可以没有图；图表内的指标不要太多，并且指标间的差距不要太大；
- e) 报表要单一，一张报表，只做一份分析功能，多个功能尽量拆到不同的表报中；在报表中尽量不要有跳转；报表只提供查询功能；
- f) 数据提取，在做产品运营中，是很常见的需求。功能比较完备的数据平台，会有数据自助提取系统，不能满足自助需求，则需要数据开发写脚本进行数据提取。

13 数据观测和分析能力

数据资产运营人员应能开展数据变化的监控和统计分析(见图4)，对数据进行自动化的日报表输出，并标识异动数据，形成数据的可视化输出。包括：

- a) 在进行数据分析之前，先进行数据准确性的校验，判断这些数据是否是你想要的，例如从数据定义到上报逻辑，是否严格按照需求文档进行，数据的上报通道是否会有数据丢包的可能，建议进行原始数据的提取抽样分析判断数据准确性
- b) 数据分析结果的输出，通常采用直观的可视化展现方式，选择一种合理的图表，使得分析结果更直观。

数据现状	• 日常监控、用户规模、用户来源、PV、UV、人数、次数 • 收入、arup、用户属性、活跃度；渠道分析；
了解趋势	• 环比、同比、流动模型 • 增长率、留存率、流失率
发现问题	• 漏斗模型、突发事件分析 • Ab test、调查问卷
认清用户	• 功能使用情况、用户留存、用户路径 • 热度分析、用户画像
应用与推广	• 精准化投放 • 用户生命周期管理：挽留、拉新等

图4 数据观测和分析层次图

14 产品评估与数据应用能力

数据资产运营人员应能持续进行数据资产的观测分析，评估数据资产健康度，同时将积累的数据应用到业务运营环节，数据质量参照GB/T 25000.24-2017。数据资产应用可包括：

- a) 以效果算法进行精准投放
 - 推荐周期短，实时性要求高；
 - 用户短期兴趣和即时行为影响力大；

- 投放场景上下文和访问人群特性
 - b) 基于音频、视频算法推送内容
 - 长期兴趣的累积影响力大；
 - 时段和热点事件；
 - 多维度内容相关性很重要；
 - c) 基于电商算法推送推送建议
 - 长期+短期兴趣+即时行为综合；
 - 最贴近现实，季节与用户生活信息很关键；
 - 追求下单与成交，支付相关；
 - d) 其他对运营有正面作用的行为。
-