



## 指标体系

### 1. 痛点分析

主要从**业务、技术、产品**三个视角来看：

- **业务视角**

业务分析场景指标、维度不明确；

频繁的需求变更和反复迭代，数据报表臃肿，数据参差不齐；

用户分析具体业务问题找数据、核对确认数据成本较高。

- **技术视角**

指标定义，指标命名混乱，指标不唯一，指标维护口径不一致；

指标生产，重复建设；数据汇算成本较高；

指标消费，数据出口不统一，重复输出，输出口径不一致；

- **产品视角**

缺乏系统产品化支持从生产到消费数据流没有系统产品层面打通；

### 2. 管理目标

- **技术目标**

统一指标和维度管理，指标命名、计算口径、统计来源唯一， 维度定义规范、维度值一致

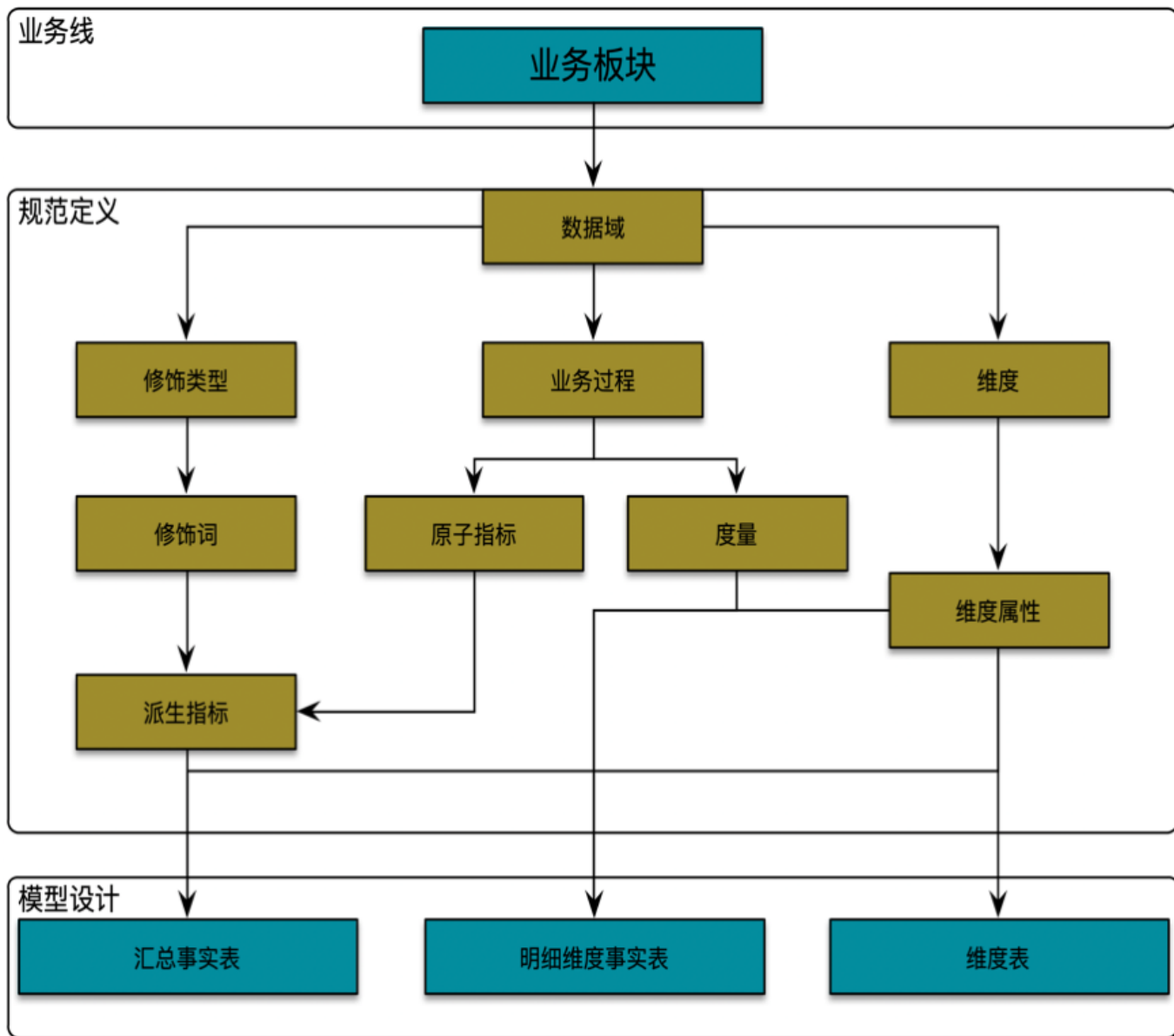
- **业务目标**

统一数据出口、场景化覆盖

- **产品目标**

指标体系管理工具产品化落地；指标体系内容产品化落地支持决策、分析、运营例如决策北极星、智能运营分析产品等

### 3. 模型架构



## 业务线

**业务板块定义原则：**业务逻辑层面进行抽象、物理组织架构层面进行细分，可根据实际业务情况进行层级分拆细化，层级分级建议进行最多进行三级分拆，一级细分可公司层面统一规范确定，二级及后续拆分可根据业务线实际业务进行拆分。

例如滴滴出行领域业务逻辑层面两轮车和四轮车都属于出行领域可抽象出行业务板块(level一级)，根据物理组织架构层面在进行细分普惠、网约车、出租车、顺风车（level二级），后续根据实际业务需求可在细分,网约车可细分独乘、合乘，普惠可细分单车、企业级。

## 规范定义

### ● 数据域

**指面向业务分析，将业务过程或者维度进行抽象的集合。**其中，业务过程可以概括为一个个不拆分的行

为事件，在业务过程之下，可以定义指标；维度，是度量的环境，如乘客呼单事件，呼单类型是维度。为了保障整个体系的生命力，数据域是需要抽象提炼，并且长期维护更新的，变动需执行变更流程。

- **业务过程**

**指公司的业务活动事件**，如呼单、支付都是业务过程。其中，业务过程不可拆分。

- **时间周期**

**用来明确统计的时间范围或者时间点**，如最近30天、自然周、截止当日等。

- **修饰类型**

**是对修饰词的一种抽象划分**。修饰类型从属于某个业务域，如日志域的访问终端类型涵盖APP端、PC端等修饰词。

- **修饰词**

**指的是统计维度以外指标的业务场景限定抽象**，修饰词属于一种修饰类型，如在日志域的访问终端类型下，有修饰词APP、PC端等。

- **度量/原子指标**

原子指标和度量含义相同，**基于某一业务事件行为下的度量，是业务定义中不可再拆分的指标**，具有明确业务含义的名称，如支付金额。

- **维度**

**维度是度量的环境，用来反映业务的一类属性**，这类属性的集合构成一个维度，也可以称为实体对象。维度属于一个数据域，如地理维度（其中包括国家、地区、省市等）、时间维度（其中包括年、季、月、周、日等级别内容）。

- **维度属性**

维度属性隶属于一个维度，如地理维度里面的国家名称、国家ID、省份名称等都属于维度属性。

- 指标分类主要分为**原子指标、派生指标、衍生指标**

- a. **原子指标**

基于某一业务事件行为下的度量，是业务定义中不可再拆分的指标，具有明确业务含义的名称，如呼单量、交易金额

- b. **派生指标**

是1个原子指标+多个修饰词（可选）+时间周期，是原子指标业务统计范围的圈定。派生指标又分

以下二种类型：

i. 事务型指标：

是指对业务过程进行衡量的指标。例如，呼单量、订单支付金额，这类指标需要维护原子指标以及修饰词，在此基础上创建派生指标。

ii. 存量型指标：

是指对实体对象（如司机、乘客）某些状态的统计，例如注册司机总数、注册乘客总数，这类指标需要维护原子指标以及修饰词，在此基础上创建派生指标，对应的时间周期一般为“历史截止当前某个时间”。

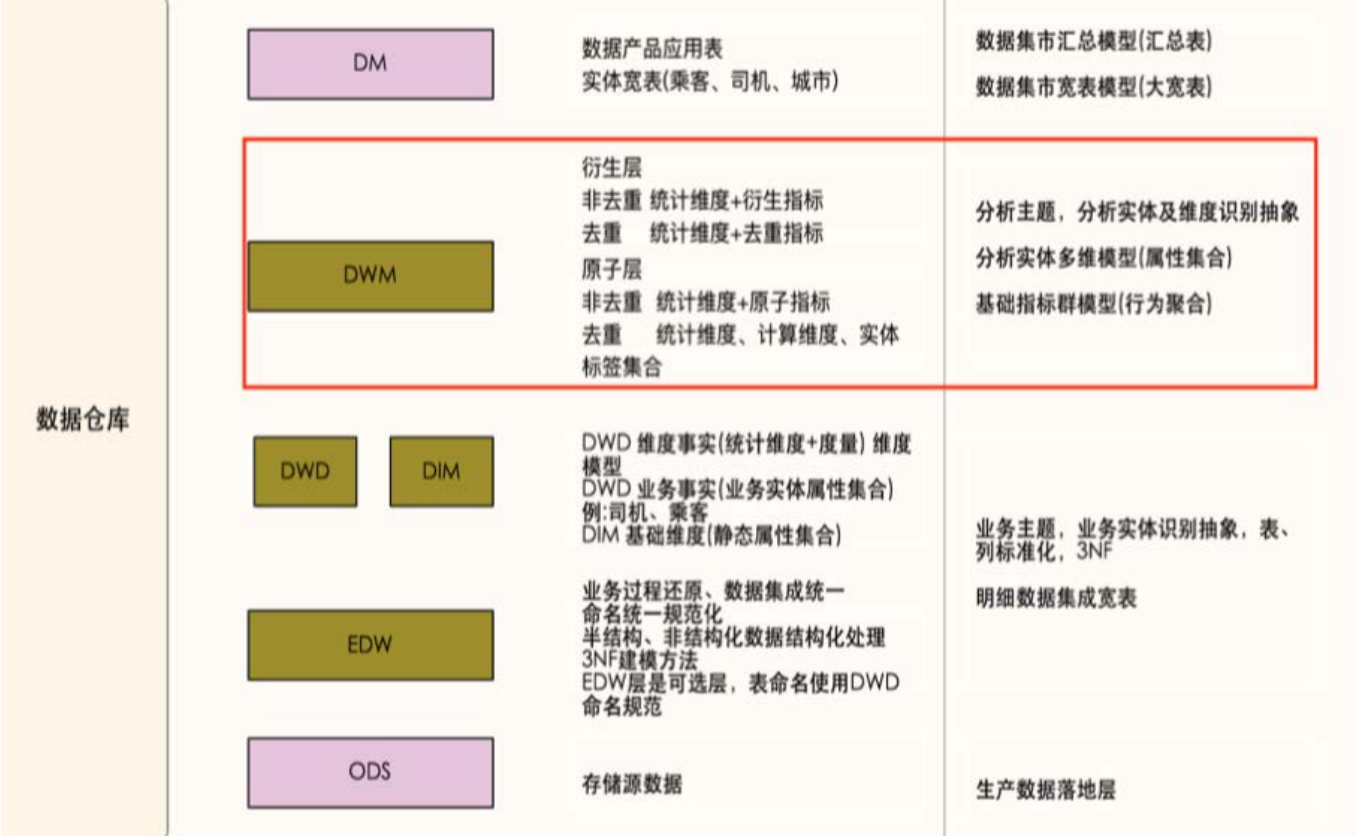
C. 衍生指标

是在事务性指标和存量型指标的基础上复合成的。主要有比率型、比例型、统计型均值

模型设计

主要采用**维度建模方法**进行构建，基础业务明细事实表主要存储维度属性集合和度量/原子指标；分析业务汇总事实表按照指标类别(去重指标、非去重指标)分类存储，非去重指标汇总事实表存储统计维度集合、原子指标或派生指标，去重指标汇总事实表只存储分析实体统计标签集合。

**指标体系在数仓物理实现层面主要是结合数仓模型分层架构进行指导建设**，滴滴的指标数据主要存储在DWM层，作为指标的核心管理层。



## 4. 指标体系元数据管理

### Apache Atlas | 元数据管理框架

#### 维度管理

包括基础信息和技术信息，由不同角色进行维护管理。

- 基础信息对应维度的业务信息，由业务管理人员、数据产品或BI分析师维护，主要包括维度名称、业务定义、业务分类。
- 技术信息对应维度的数据信息，由数据研发维护，主要包括是否有维表（是枚举维度还是有独立的物理维表）、是否是日期维、对应code英文名称和中文名称、对应name英文名称和中文名称。如果维度有维度物理表，则需要和对应的维度物理表绑定，设置code和name对应的字段。如果维度是枚举维，则需要填写对应的code和name。维度的统一管理，有利于以后数据表的标准化，也便于用户的查询使用。

#### 指标管理

包括基础信息、技术信息和衍生信息，由不同角色进行维护管理。

- 基础信息对应指标的业务信息，由业务管理人员、数据产品或BI分析师维护，主要包括归属信息(业务板块、数据域、业务过程)，基本信息(指标名称、指标英文名称、指标定义、统计算法说明、指标类型(去重、非去重))，业务场景信息(分析维度，场景描述)；
- 技术信息对应指标的物理模型信息，由数据研发进行维护，主要包括对应物理表及字段信息；
- 衍生信息对应关联派生或衍生指标信息、关联数据应用和业务场景信息，便于用户查询指标被哪些其它指标和数据应用使用，提供指标血缘分析追查数据来源的能力。

原子指标定义归属信息 + 基本信息 + 业务场景信息派生指标定义时间周期 + 修饰词集合 + 原子指标修饰类型主要包含类型说明、统计算法说明、数据源(可选)

## 5. 指标体系建设流程

#### 建模流程

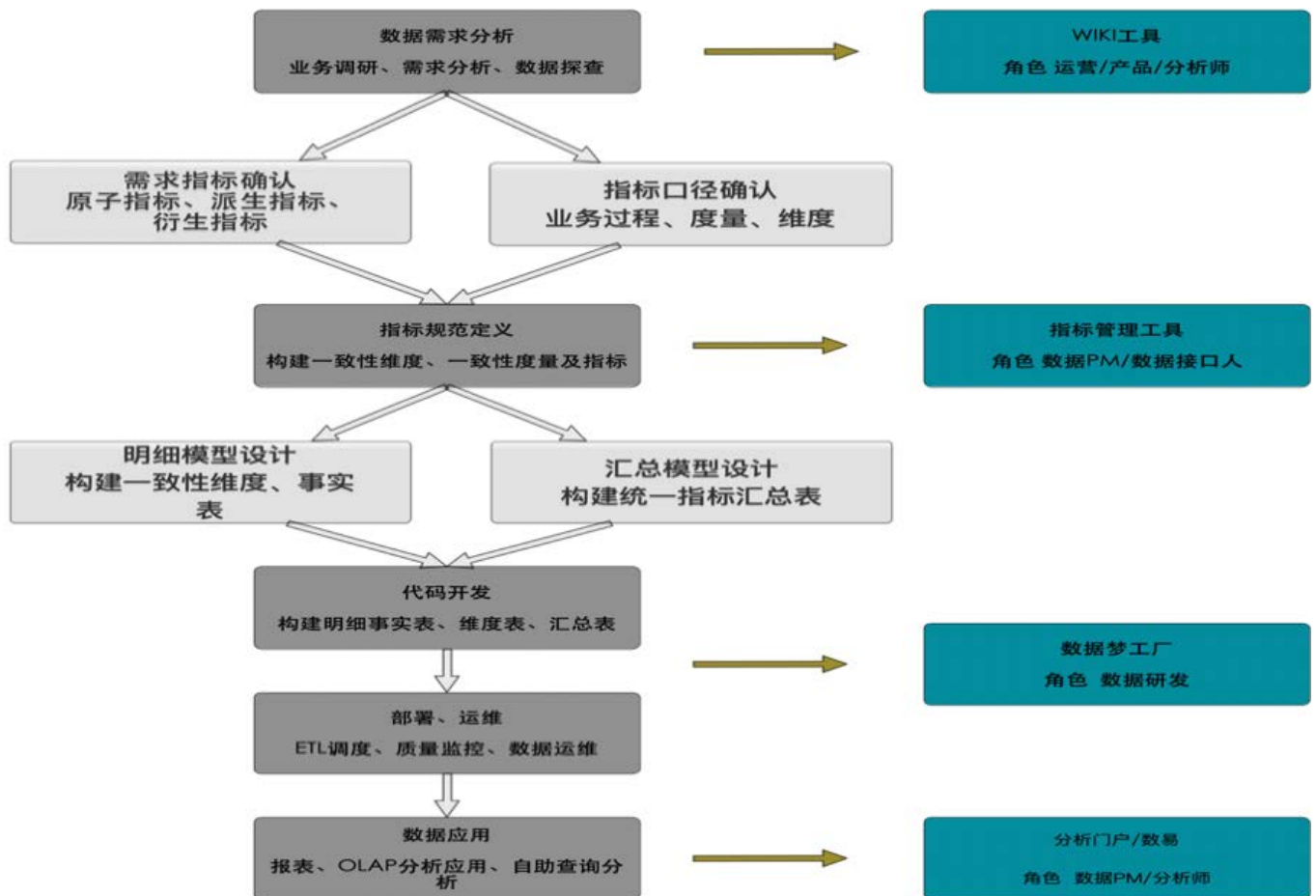
建模流程主要是从业务视角指导工程师对需求场景涉及的指标进行主题抽象，归类，统一业务术语，减少沟通成本，同时避免后续的指标重复建设。



分析数据体系是模型架构中汇总事实表的物理集合，业务逻辑层面根据业务分析对象或场景进行指标体系抽象沉淀。滴滴出行主要是根据分析对象进行主题抽象的，例如司机主题、安全主题、体验主题、城市主题等。指标分类主要是根据实际业务过程进行抽象分类，例如司机交易类指标、司机注册类指标、司机增长类指标等。基础数据体系是模型架构中明细事实表和基础维度表的物理集合，业务逻辑层面根据实际业务场景进行抽象例如司机合规、乘客注册等，还原业务核心业务过程。

## 开发流程

开发流程是从技术视角指导工程师进行指标体系生产、运维及质量管控，也是数据产品或数据分析师和数仓研发沟通协调的桥梁。



6. 指标体系图谱建设

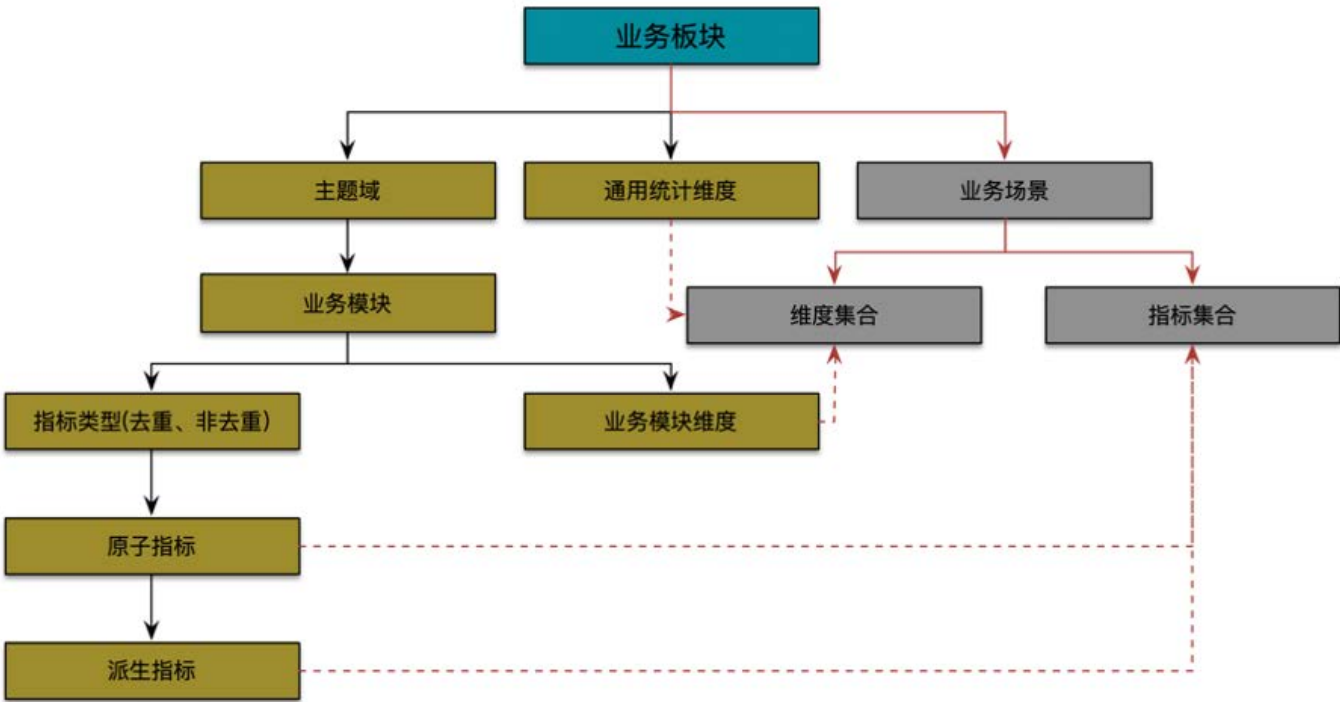
指标体系图谱概述

指标体系图谱也可称为数据分析图谱主要是依据实际业务场景抽象业务分析实体，整合梳理实体涉及的**业务分类、分析指标和维度的集合**。建设方法：主要是通过业务思维、用户视角去构建，把业务和数据紧密关联起来，把指标结构化分类组织。

建设目的：

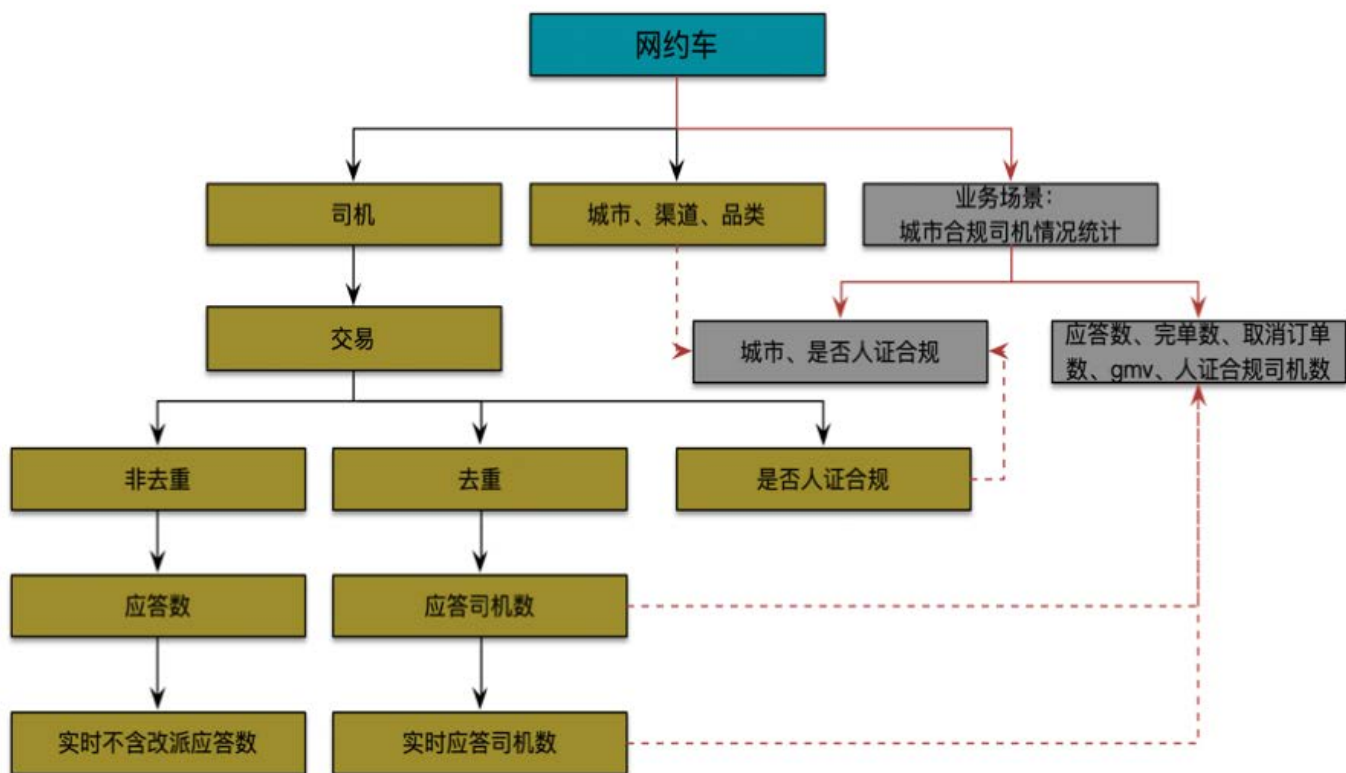
- 对于用户：  
便于用户能够快速定位所需指标和维度，同时通过业务场景化沉淀指标体系，能够快速触达用户数据诉求。
- 对于研发：  
利于后续指标生产模型设计、数据内容边界化、数据体系建设迭代量化和数据资产的落地。

指标体系图谱模型

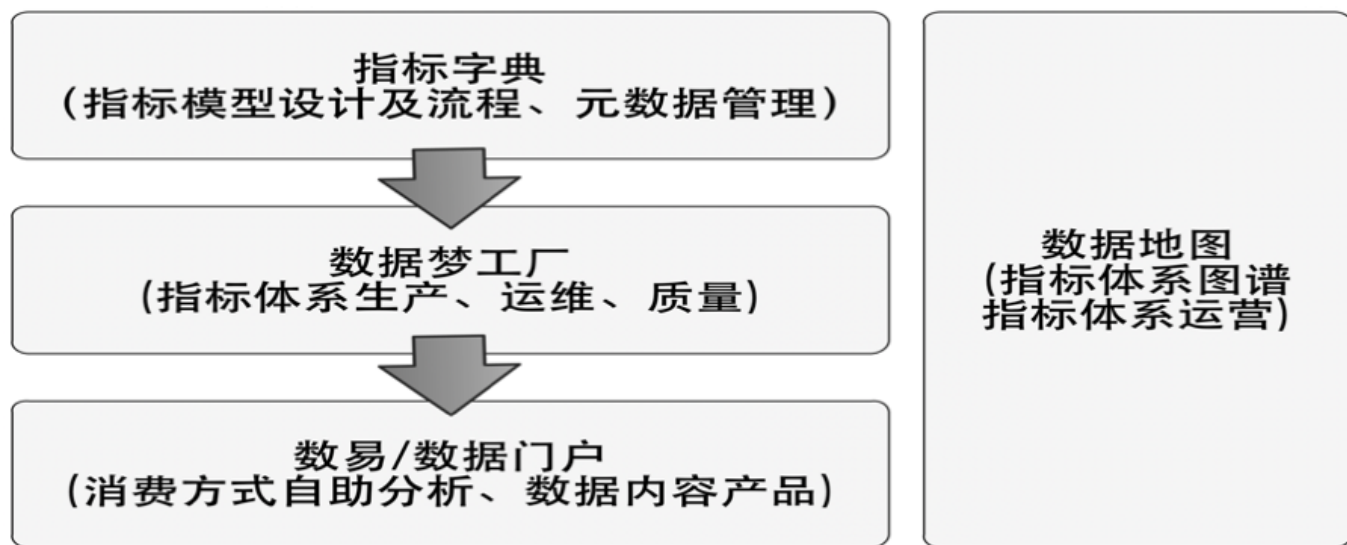


指标体系图谱实例





## 指标体系产品化



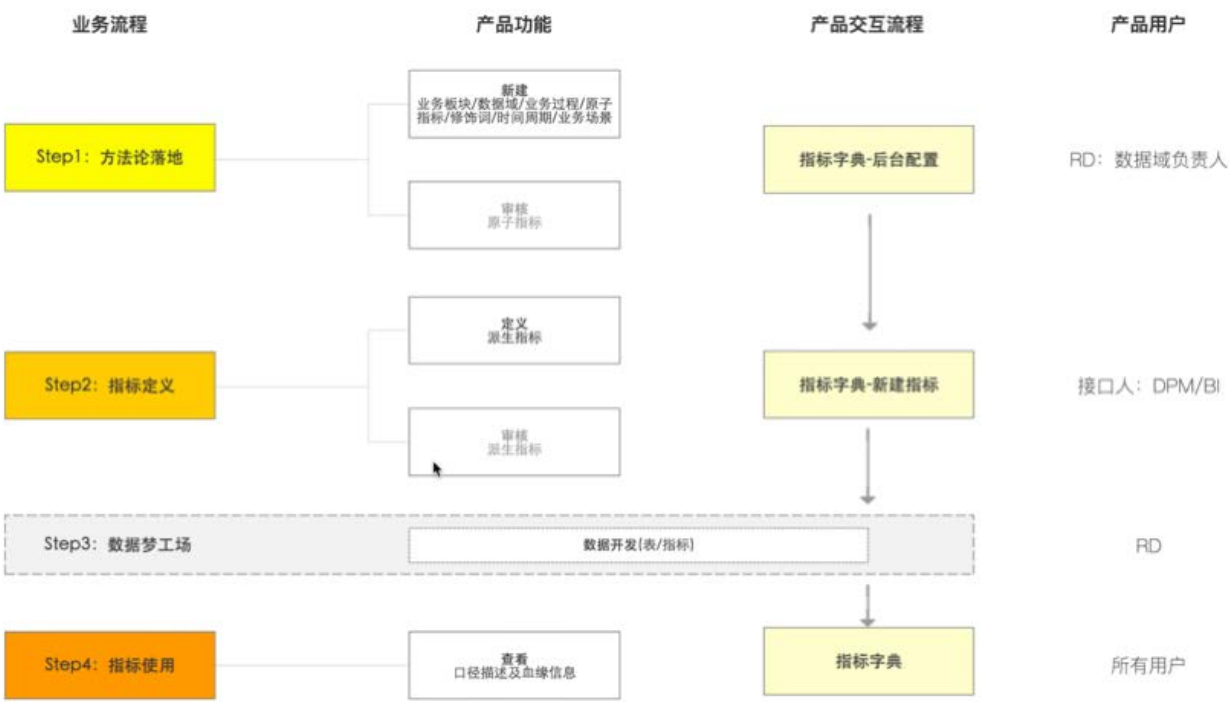
指标体系涉及的产品集主要是依据其生命周期进行相应建设，通过产品工具打通数据流，实现指标体系统一化、自动化、规范化、流程化管理。因为指标体系建设本质目标是服务业务，实现数据驱动业务价值，所以**建设的核心原则是“轻标准、重场景，从管控式到服务式”**。通过工具、产品、技术和组织的融合提高用户使用数据效率，加速业务创新迭代。

其中**和指标体系方法论强相关产品就是指标字典工具的落地**，其产品的定位及价值：

- 支撑指标管理规范从方法到落地的工具，自动生成规范指标，解决指标名称混乱、指标不唯一的问题，消除数据的二义性



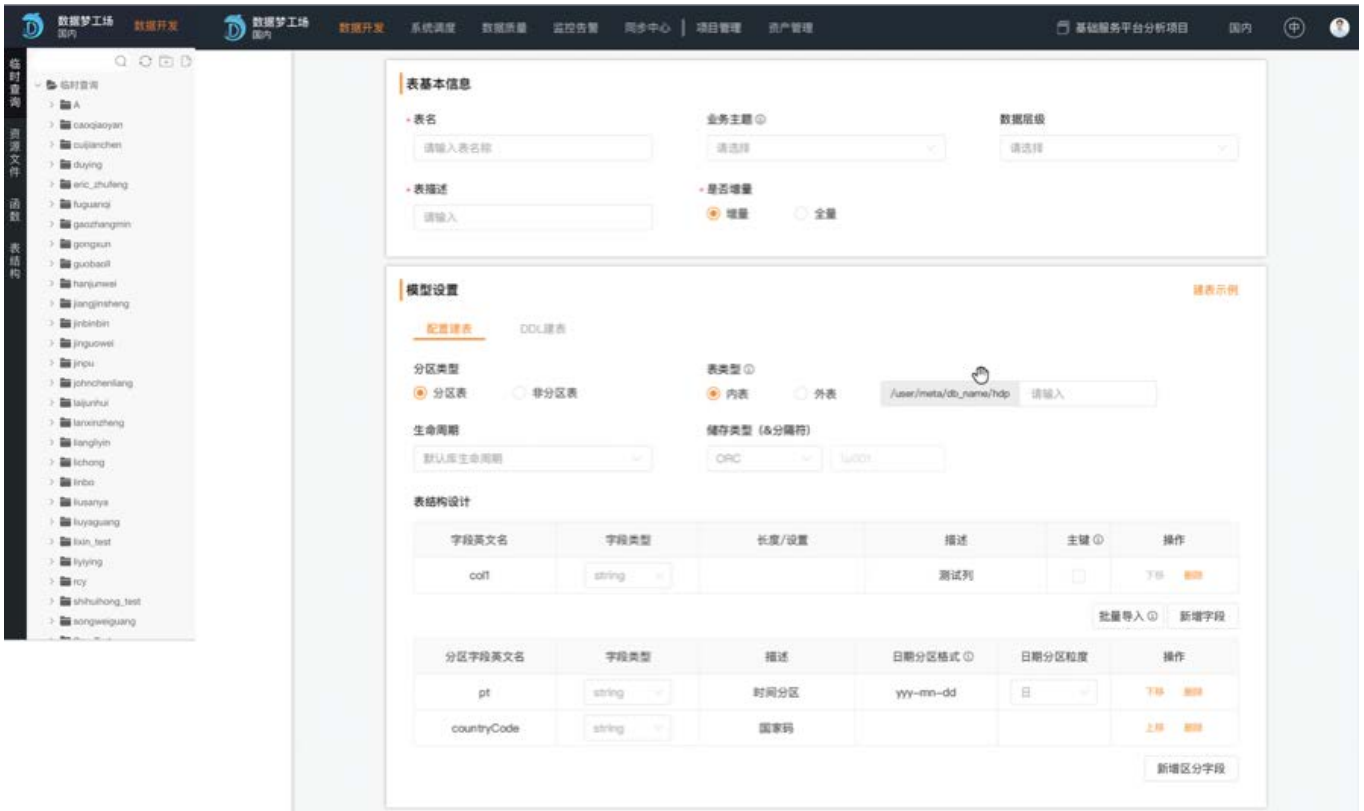
- 统一对外提供标准的指标口径和元数据信息



工具设计流程 (方法论->定义->生产->消费)



指标定义



指标生产

## 结束语

文章整体介绍了指标体系建设方法论&实践和工具产品的建设情况，指标字典和开发工具已实现流程打通，与数据消费产品的打通后续会通过DataAPI方式提供数据服务。