

字节跳动企业级埋点设计方法论及实践分享

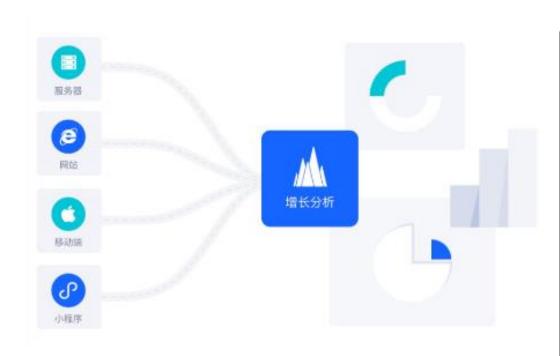
火山引擎数据分析师:华雷

2022/09

关于火山引擎增长分析DataFinder



支持多场景业务决策的数据分析平台



点击进入

火山引擎增长分析DataFinder官网

跨端跨触点,聚合多种数据源

多种埋点形式、采集APP、网站,小程序等各平台数据,并通过API,聚合跨端、跨触点、用户全生命周期的数据,实现对用户的全面洞察

● 搭建指标体系,业务效果可衡量

以统一的数据指标系体系链接不同业务,对关键节点的数据进行深度下钻分析,更好地用数据衡量业务效果,洞察业务问题

● 从数据中挖掘增长点

站式用户分析和运营平台

聚合多种高级分析模型,灵活组合各项指标,实时数据灵活洞察用户全生命周期的行为表现,从而发现指标背后的增长可能

● 为业务增长进行归因

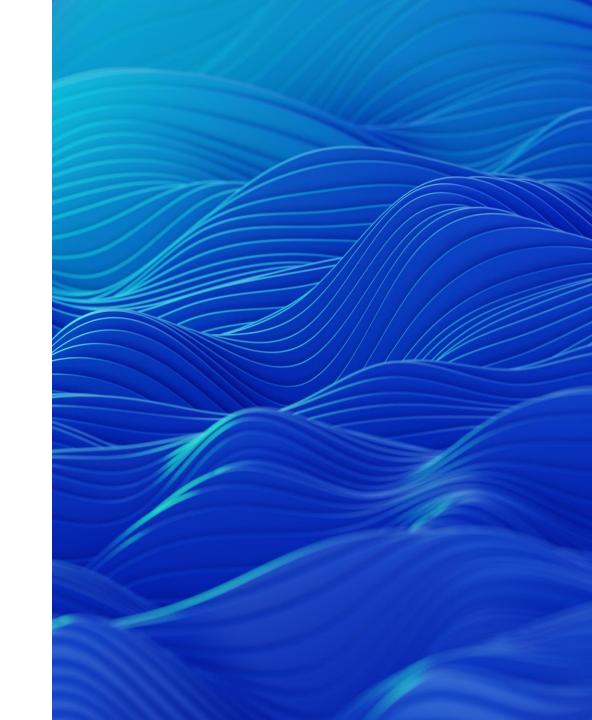
什么渠道的投放效果最好?哪个资源位为运营指标的上升带来了最大贡献?用归因模型找到增长的真正原因,并固化为行之有效业务范式

● 细化分群并精确运营

根据各种特征数据,精确洞察与运营,通过透弹窗,短信,邮件等渠道 推进个性化运营活动,实现分层运营

目录

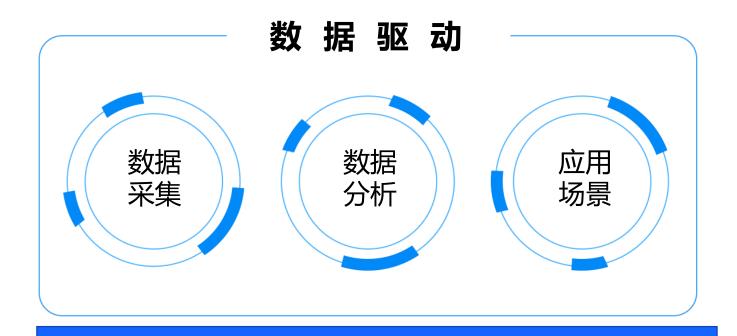
- 1.企业埋点过程中的痛点以及解决方案
- 2.埋点设计整体的规划思路
- 3.埋点的团队协作流程
- 4.行业埋点实践



增长分析是什么?



一站式用户分析与运营平台,为企业提供数字化消费者行为分析洞见,优化数字化触点、用户体验,支撑精细化用户运营,发现业务的关键增长点,提升企业效益。



支持多场景业务决策的数据分析平台

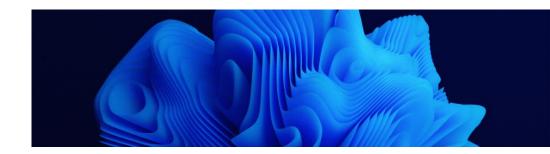
- 采集并整合全域数据: 采集并整合包括用户行为、经营数据、平台数据等多方数据来源;
- **覆盖客户业务场景分析需求**:在产品、运营、用增等业务场景提供所需要的业务数据;
- **智能决策并驱动业务**:引导实际的 业务动作



01

埋点的痛点以及解决方案

企业数据治理的整体目标是持续运营数据价值,推动企业数字化转型。企业在转型过程中会面临来自业务、组织架构、技术等等方面的问题..企业在转型过程中会面临来自业务、组织架构、技术等等方面的问题..



日常数据应用过程中常见埋点问题



数据采集方案设计作为数据建设的重要组成的部分,直接影响到后续的数据应用质量和数据回溯,而我们在日常中是不是经常会碰到如下问题:

01

02

03

04

埋点重要信息缺失

作为一个新入职员工,面对环境中的几百个事件,或者afsdfgfhtr无任何标注的属性名,茫然不知所措,而前任留下的文档也是似有若无,询问身边的老员工则众口不一……

埋点用时方知少

作为一个运营人员,活动策划写好了,活动页面做好了,哈哈,到了大展身手的时候啦,想详细分析不同人群的数据反馈。突然发现很多属性信息没有,不足以细分,好的,什么也不用干了……

埋点口径不统一

作为一个产品人员,新版本上线后,想详细分析新版本上线后的数据变化。内心os:还好之前提了埋点的需求,不像上面那个运营没有数据可以查。几分钟后……等等,为什么这个数据和后台业务数据很不一样……

埋点浪费

过了一段时间后,会有些历史埋 点无人查询,导致环境中事件人 越来越多。

大数据时代,企业流量数据的建设困境



生产链路长

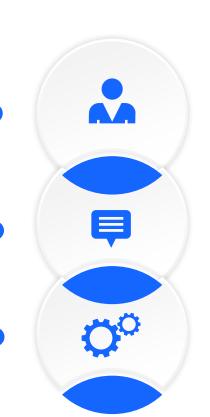
埋点采集生链路长角色多,管理难度高

设计不规范

设计不规范导致埋点离散混乱,不利于使用

业务利用率低

埋点使用情况无处追踪, 存在大量无用埋点



埋点质量差

错埋漏埋,质量参差不齐,使用数据有风险

数据管理难

各类数据缺少统一管理,难以发挥数据价值

消耗成本高

机器和运维人员消耗成本高

各角色在流量数据实践中的痛点





数据使用者

埋点无统一管理,逻辑 含义、属性内容、负责 人等信息靠打听。找不 到现有可用埋点、不敢 用无负责人埋点,只能 重新埋。



埋点开发者

埋点新增、变更、废弃、 下线无处记录,过去的 埋点代码不敢动,越埋 越多,原来越混乱,开 发维护成本高。



埋点需求方

埋点需求无处维护,无 法进行灵活的有管控的 版本管理。PM/运营都 来找我问埋点,无查询 工具辅助我进行管理。



数据管理者

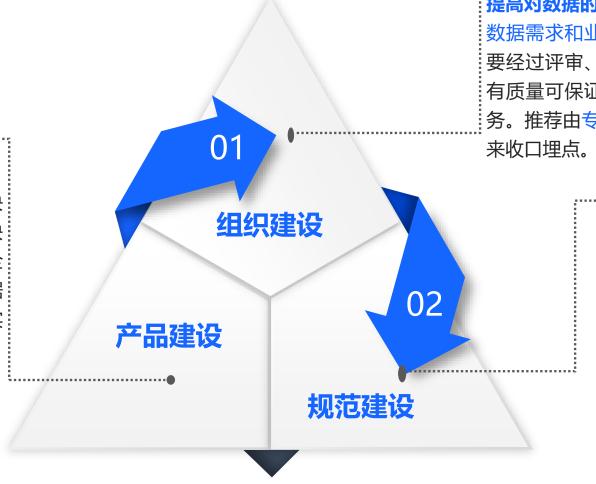
流量数据资产现状如何, 埋点个数、数据量、成 本消耗、有效埋点个数 是多少。是否有高效的 治理管理工具。

火山引擎埋点治理解决方案:产品+方法论



用工具来管理,提高效率

通过一站式埋点治理平台解决 埋点方案线上化、按照方案快 速验证埋点数据,管理集团全 部埋点资产等,提升整体数据 质量和可维护性,让规范扎实 落地。



提高对数据的重视

数据需求和业务需求一样重要。埋点需要经过评审、测试等环节才能上线。只有质量可保证的数据才能真正地指导业务。推荐由专门的数据分析师/数据产品来收口埋点。

建立一套规范

规范埋点工作流程,明确埋点流程中各角色的责任和权利。

确定埋点注册机制,将埋点资产线 上化管理,方便查询语义、负责人 和版本记录等。

规范埋点设计,需要具有可操作性和可验证性。



组织建设 提高组织对数据的重视。数据需求和业务需求一样重要!



| 火山引擎埋点治理解决方案: 组织建设



在埋点开始之前要思考的问题?

整体性思考

定位

- 系统在企业内部整体数据体系中的承担边界
- 数据源的导入和导出流向
- 埋点业务的范围

项目管理

- 清晰的项目目标
- 规范的业务流程
- 明确的优先级顺序
- 清晰的角色负责人



组织建设-定位



工具冲突

- 矛盾1

数据体系中众多数据产品,没有想清楚具体每个系统所 承担的作用分别是什么,如何避免各司其职而定位明确, 不造成重复浪费资源和人力。

埋点业务范围

- 矛盾2

承接的业务线和埋点的范围,例如业务数据到底是导入还是把行为数据导出? 多业务线数据分还是合? **一** 行为数据的采集和分析

最本质的行为数据采集上报和各种应用模型的分析。

为其他系统提供数据

为消费画像、AB、GMP、ABI等系统提供基础 行为数据,作为重要数据来源,与其他信息整合。

简易BI

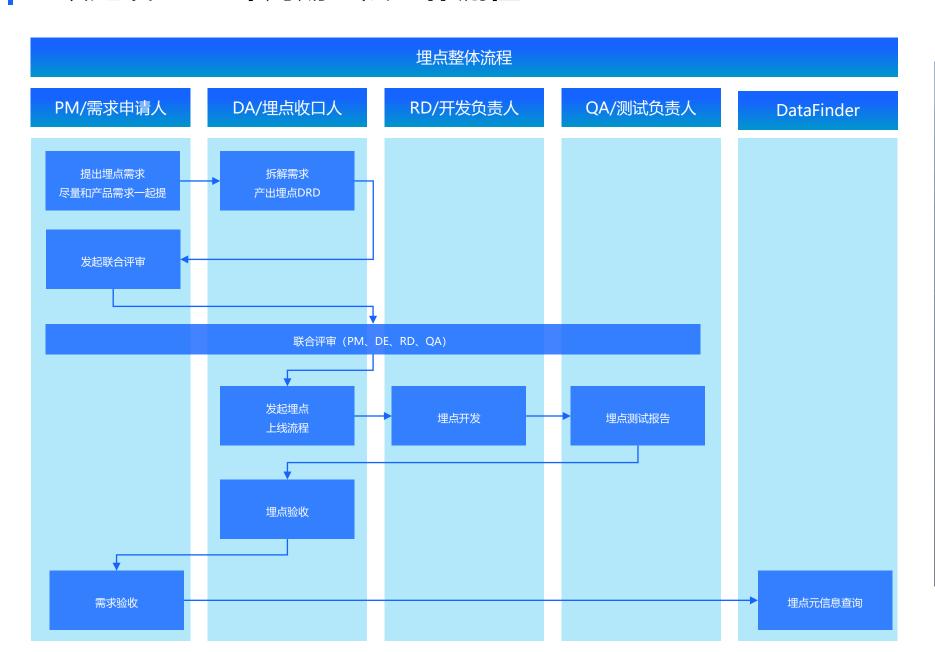
基础数据可视化展现。

4 行为数据与业务关联

作为行为数据源将行为数据和业务数据进行关联, 可进行导入或者导出。

组织建设-全生命周期埋点工作流程

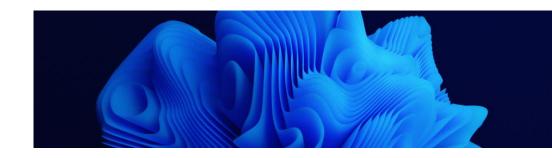




埋点设计手册1.0

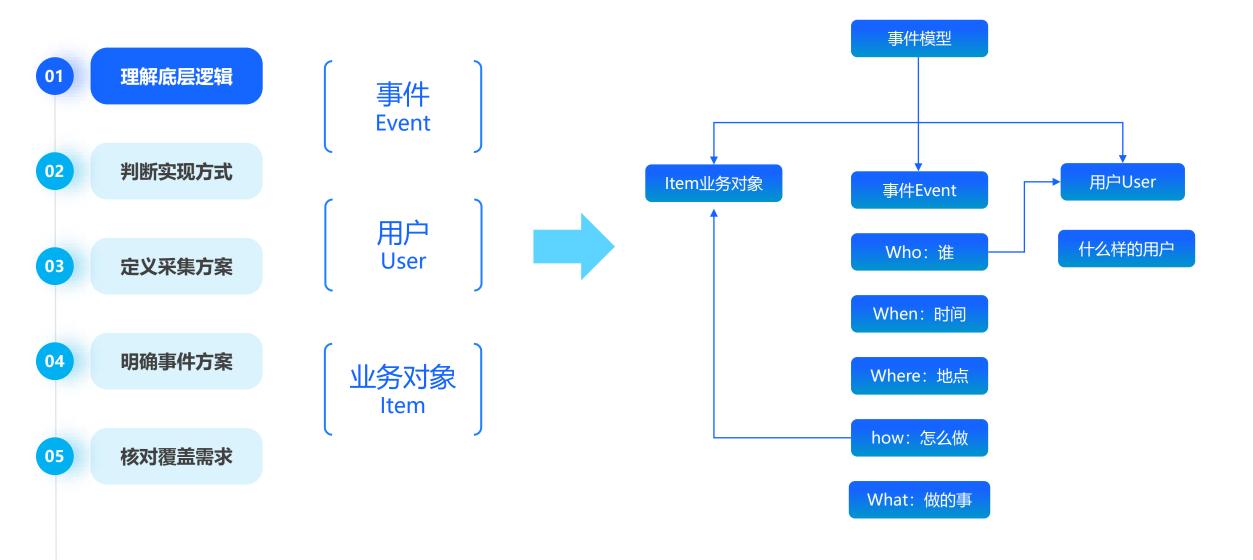
- 1. 埋点是什么
 - 1.1埋点构成:事件、属性
 - 1.2埋点类型
- 1.2.1launch(App启动)
 - 1.2.2terminate (App退出)
 - 1.2.3predefine_pageview (默认PV事...
 - 1.2.4停留时长
 - 1.2.5自定义事件
- 2. 埋点流程规范
 - 2.1埋点设计
 - 2.2埋点录入
 - 2.3埋点的评审和开发
 - 2.4埋点测试验收
 - 2.5埋点维护
- 3. 埋点设计规范
 - 3.1埋点设计规范
 - 3.1.1埋点方案设计原则
 - 3.1.2埋点方案设计思路
 - 3.2埋点命名规范
 - 3.2.1Event
 - 3.2.2Param





| 火山引擎埋点治理解决方案: 规范建设





数据采集方案设计-业务对象



什么是业务对象?

业务对象,又称item。item是对事件+用户模型的补充,用户在我们的 App中发生行为记录的时候,都会有一个业务对象,不同种类的App业务对 象不同。例如电商中业务对象为商品,在线教育中业务对象为课程,银行中 业务对象为理财产品、贷款产品等,短视频中业务对象为短视频等。

以内容类app举例:

事件名称	属性名称
浏览内容详情页	预置属性
	内容ID
	内容名称
	内容类型

ltem表
内容ID
内容名称
内容类型
内容关键词
是否付费
发布时间
观看方式
搜索标识
内容提供商标识
媒资分类
年份
产地
演员 (电竞为主播/解说)
导演
语言
描述
版权方
审批号
授权时间

数据采集方案设计-业务对象





规范建设-数据采集方案设计





| 规范建设-各种埋点优劣势



	全埋点	代码埋点 (前端)	代码埋点 (后端)	预置埋点
定义	• 通过标准SDK采集的页面浏 览和按钮点击数据,可选开 启or 不开启。	• 在app\web\h5\小程序等 需要前端研发代码埋点。	• 在服务器端进行代码埋点。	• 通过标准SDK采集的预置事件,例如App启动和App退出,详细见说明文档,默认开启。
优势	开发成本低采集范围广用于可视化埋点,操作门槛低。	可满足深入场景定制化分析需求。取值规范,上报数据准确度可控。	数据准确度高。不依赖前端发版。取值规范。	・ 采集基础数据,满足基本分析需求。・ 减少开发成本,对基础实践默认采集。
劣势	・ 依赖于代码规范程度,会有一定程度脏数据。・ 采集数据量大。・ 只满足基础的PV,UV分析。	需要一定程度开发投入。前端会有一定程度上的数据 丢失。依赖发版。	• 需要一定程度开发投入。	• 无

规范建设-数据采集方案设计



01 理解底层逻辑

02 判断实现方式

03 定义采集方案

04 明确事件方案

05 核对覆盖需求

01 用户标识

业务唯一ID确定

了解用户识别逻辑,查看复杂场景下用户是否能够关联。

02 文档维护规范

- 修改的changelog记录。
- · 记录修改人、修改时间、修改内容、确切表达、保证可溯期、可迭代。

03 设计规范

需包含序号、事件名称、事件属性名称、属性类型、属性值含义或示例、事件的触发时机、埋点形式、备注等

04 数据格式

Int, string, float, datetime, list

数据采集方案设计-用户标识



火山引擎增长分析使用 device_id、user_unique_id、ssid 三种 id 标识设备和用户。

device_id/web_id

设备的唯一标识,我们通过设备 注册服务根据获取到的设备信息 为每个设备生成唯一的标识,该 标识会通过客户端sdk在设备本 地进行存储。另,网页端、小程 序使用web_id,作用与 device_id 基本相同

user_unique_id

用户唯一标识,一般情况直接使用产品业务中使用的用户标识,比如登录账号。当user_unique_id未设定时,在SaaS版本中,系统会自动使用device_id/web_id替代,在私有化版本中,会显示为空

ssid

火山引擎增长分析为用户/设备 自动分配的唯一标识,并与 device_id/web_id、 user_unique_id 相关联,以实 现对匿名和实名状态用户的统一 标识。

数据采集方案设计-设计规范

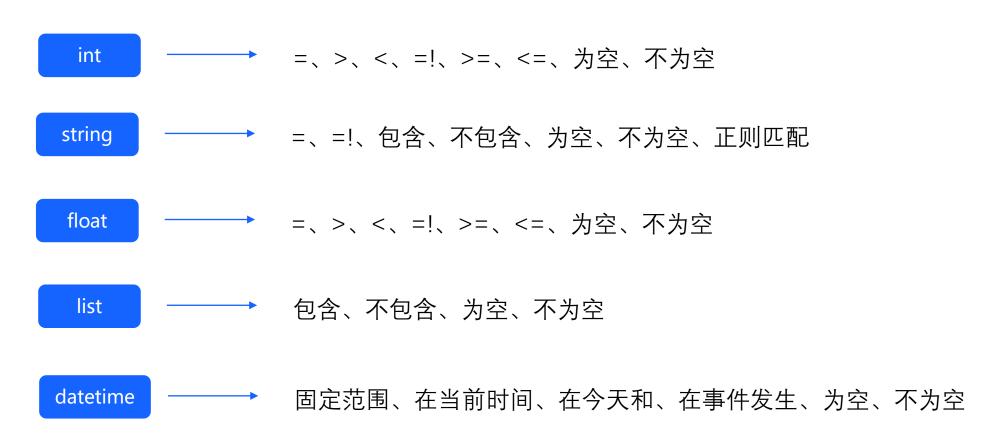


序号	event	埋点描述	属性名称	属性类型	属性描述	属性备注	是否必传	属性取值	其它描述性信息
- /1			1. 如果定义了属性,则名称 、类型、描述和是否必传字段 必填 2. 表格内复用已有属性不允	则必填) 属性类型: 限制integer、string 、float、list、		选填 属性取值相关备注, 长度要求200字符以内	选填(若属性名称非空,则必填) 是否必传(大小写均可): TRUE/FALSE 如果选择TRUE, 会在埋点验证的时候校验该埋点下的必传字段是否为空	选填 字段值和字段值说明,格式为 (冒号分隔): 取值:取值描述	选填 填写其他备注信息,系统不做解析
	homepage_hot_slide_up	推荐页视频上滑	enter_from	string	页面	根据上一级页面上报	TRUE	homepage_hot: 个人主页 homepage_follow: 关注页	属性取值为:字符串枚举类型
1			is_follow	string	是否关注		TRUE	0: 是1: 否	boolean类型用于过滤筛选的属性
			style_id	integer	样式号		FALSE		
			type	string	样式类型	取固定值: normal	TRUE		
2	video_play	视频播放	rating	integer	用户提交的评分星级,1星	取1到5整数	TRUE	[1, 5]	属性取值为:数值范围类型
3	video play	视频播放	type	string	样式类型		TRUE		
J		7元少央1田八人	rating	integer	用户提交的评分星级,15	星到5星	TRUE	[1, 5]	
4	test_event (注:以下两行可不必合	测试	test_param1	string	属性1		TRUE		
	# ,		test_param2	integer	属性2		FALSE		
			test_param3	integer	属性3		TRUE		

数据采集方案设计-数据格式

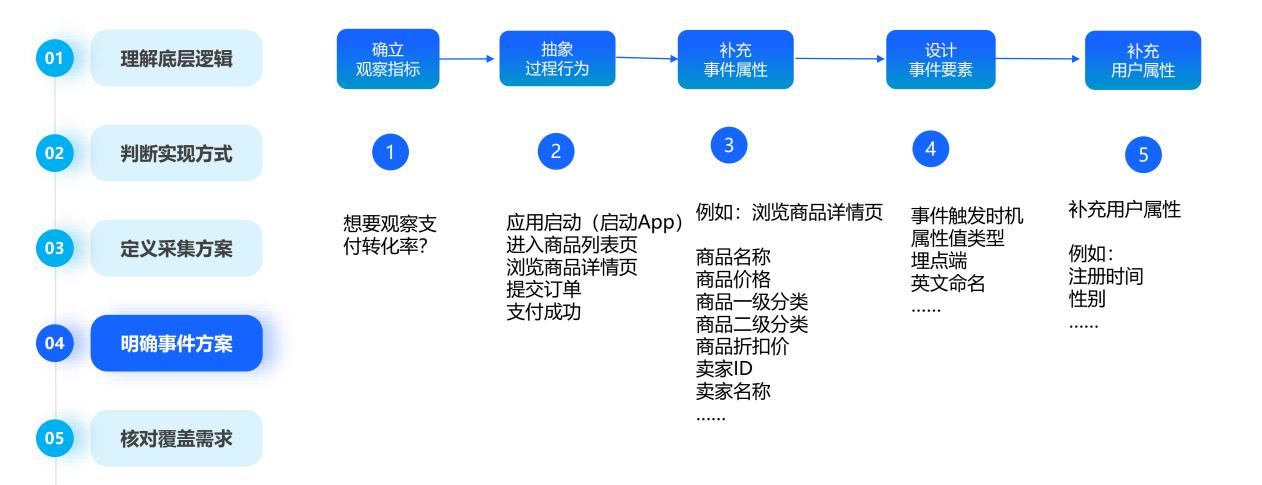


产品内筛选条件:



数据采集方案设计





数据采集方案设计



01 理解底层逻辑

02 判断实现方式

03 定义采集方案

04 明确事件方案

05 核对覆盖需求

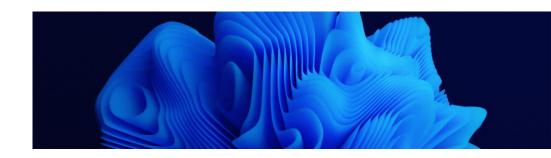
回溯需求是否满足

排开发的优先级

规划后续开发计划



产品建设 可通过DataFinder来管理埋点数据,提高效率



火山引擎埋点治理解决方案: 埋点治理





扫码进群*全* 领取**免费埋点采集方案模板**



火山引擎埋点治理解决方案:埋点需求



基于需求的埋点方案,在需求内列举了你想要追踪的埋点和属性、对应的业务场景,开发同学就可以根据 此需求进行开发、QA同学可以据此进行测试、最后进行上线使用。

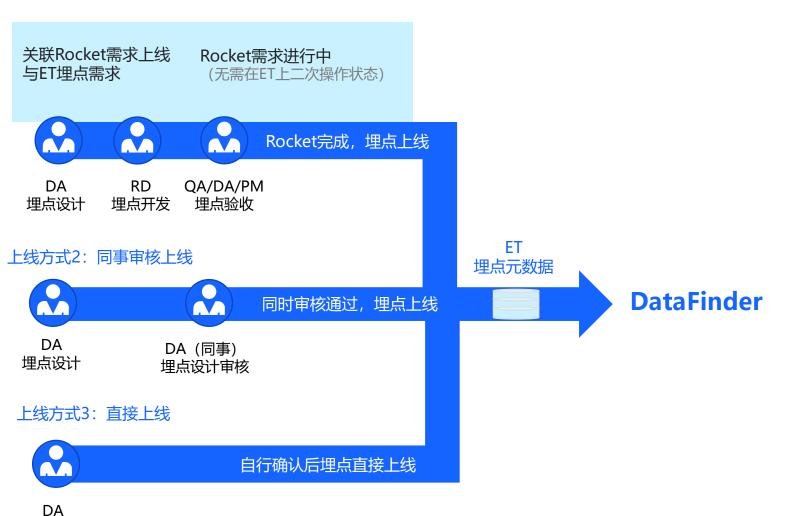


火山引擎埋点治理解决方案: 需求上线



上线方式1: 关联Rocket需求上线

埋点设计



基于需求的埋点上线,

提供直接上线和同事审核 上线流程。

| 火山引擎埋点治理解决方案: 埋点设计



基于场景的埋点设计,支持分场景来填写埋点的属性内容。

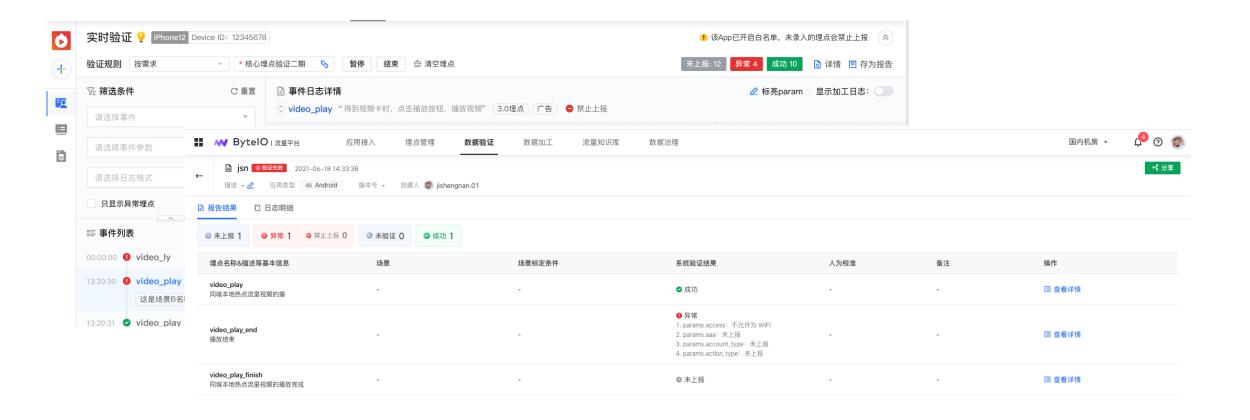
标签管理, 支持给埋点填写多个标签将埋点按照功能模块分类。



火山引擎埋点治理解决方案: 埋点验证



基于需求的埋点验证,针对指定的埋点需求,QA同学可以在强大的系统支持下(系统会根据需求中的验证规则自动校验),进行快速的验证。不仅能实时的了解验证的进展,还可以一键生成对应的验证报告。

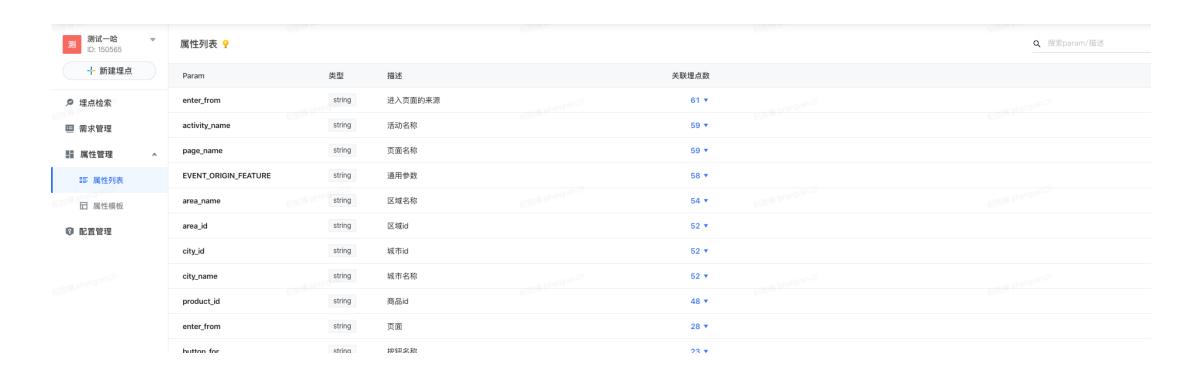


| 火山引擎埋点治理解决方案: 属性管理



属性管理规范,一个业务的属性是规范的,标准的。

同名属性不允许有不同的类型和描述。



火山引擎埋点治理解决方案: 埋点资产



埋点资产有多少,有用的有多少,禁用的有多少。

	~	埋点名称&描述等基本信息	操作系统	标签	属性	埋点负责人	昨天上报PV 💠	状态①	操作 ③
- - 新建埋点		event_list_click 埋点检索列表页点击	Web	埋点管理	_staging_flag、all_para m、anal 8 个	@ ?	795	■已上线	回 分析 ○ 禁用
▶ 埋点检索■ 需求管理		logverify_fold_filter_condition_panel 埋烏验证界固清选条件面板収起按钮统计	Web	- 宏胜權 (sterior or or	fold 1 ↑	shaozhirong zhaozhirong	590	■已上线	□ 分析 ○ 禁用
■ 属性管理	^	event_verify_debug_device 埋点验证开启连接-调试设备	Web	埋点验证	app_id、platform、user 3 ↑	₫3	560	■ 已上线	副 分析 ○ 禁用
器 属性列表 □ 属性模板		create_event_click 包含新建埋点需求,多需求批量录入,直接录入	Web	新建埋点	_staging_flag、event_ind ex、me 5 个	4 € +1	255	■已上线	□ 分析 ○ 禁用
② 配置管理		component version switch to 验证组件新旧版切换	Web	-	version 1 个	😰 lijiawei.0602	240	■ 已上线	副 分析 ○ 禁用
		event_verify_filter -	-	EIEM Identifier (1)	_staging_flag、app_id、 event_in 6 个	A admin-et	235	■未录入	□ 直接录入
		enter_transformation_detail 进入数据转换详情	Web	-	from 1 个		135	■已上线	■ 分析 ○ 禁用
		save_demand 保存需求通用信息按钮点击	Web, Android, iOS	- marking let engran (1)	_staging_flag、event_ind ex、save、user 4 个	ishengnan.01	120	■ 已上线	回 分析 ○禁用
		create_demand_click 新建埋点和新建埋点需求按钮点击	Web, Android, iOS	-	_staging_flag、event_ind ex、position、user 4 个	🧖 jishengnan.01	100	■已上线	回 分析 ○ 禁用
		kg_search -	-	-	_staging_flag、condition s、even6个	(8) admin-et	75	■未录入	□直接录入
		submit_online 規交上线按钮	Web, Android, iOS	宏胜權 阿里安斯	_staging_flag、online_m ethod、 4 个	🍘 jishengnan.01	75	■已上线	回 分析 ○ 禁用
		number_click 关联理点个数点击	Web, Android, iOS	属性管理页	_staging_flag、click_typ e、even 5 ↑	pishengnan.01	65	■已上线	回 分析 ○ 禁用
	≡	共84行 每页行数 20 ▼							< 1 2 3 4 5 >

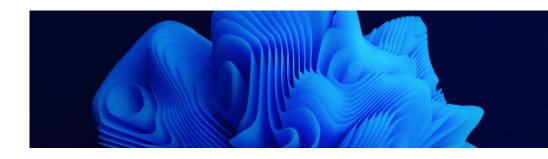
| 火山引擎埋点治理解决方案: 埋点下线







少人 埋点行业实践 x



常见行业埋点方案实践



设计为同一事件	例如相似场景下的按钮点击可合并,不必一 个点击一个事件,需合并为一个事件。	对于全局性的事件,我们建议设计为同一事件,通过特定的属性值对特定操作进行区分,而不是针对每一个操作设计一个事件。	事件名称	事件履 频道位 频道ID 频道名	字符串 字符串		大 例 支,生活、娱乐、财经
设计为同一事件	例如:点击banner、点击热门活动位,都是 点击首页的推荐位,可通过增加属性区分。	各事件所需属性相差不大,平时分析场景多 整体分析。	点击推荐位	事件属性 模块类型 点击位置 点击ID	属性值类型 STRING STRING STRING STRING	banner、 1, 2, 3· 如果是点 如果是点 如果是点	活动推荐位 活动推荐位的位置排序序号:
设计为不同事件	例如一个内容平台,有视频,有文章,因视频和文章所记录属性差异较大,浏览内容详情应区分为浏览视频详情和浏览文章详情	各事件所需属性相差很大,分析场景多分别 分析。	事件名称 浏览资讯文章 视频播放	内内内当发作新责记浏内内内视是是内内内当发作新责记浏内内内视是是	: 标题 : 频道 : 类别 : 讨问 : 时间 : 在称 : 来源 : 来源	属字字字字数时字字字字字字字字数BBB数性符符符符值间符符符符符符符符符符值 Le由串串串串串串串串串串串串串串串串串串串串串串串	属性值示例 热门,热评,置顶······ 娱乐、财经、社会等 娱乐、财经、社会等
设计为不同事件	例如:收藏商品、浏览商品详情,虽属性差异不大,但是收藏和浏览业务关系不大,且通常为单独分析。	各事件所需属性相差不大,但平时分析场景 单一分析,并且业务涵义区别较大。		品详情页 i-收藏	事商商商商商商商商商商商商商商品品品品品品品品品品品品品品品品品品品品品品品	属性值类型 string string string string string string string string	属性值示例 商品的唯一ID 商品对应的名称 商品对应的品类 商品对应的品牌 商品对应的名称 商品对应的名称 商品对应的品类

相似场景是合并一个 事件还是分不同的事 件



www.volcengine.com

常见行业埋点方案实践



曝光事件的处理

和其他事件一样,只是曝光事件的触发时机需要注意,例如某平台内容曝光事件触发时机为:

- 1、内容露出全部,且feed流静止状态超过2s算曝光
- 2、限制单一内容一次请求只会出现一次曝光。(比如上下滑动屏幕,只要不刷新发生新请求,算一次曝光)

当然具体的规则可根据需求和研发的实现成本灵活变动,以上仅为示例,可做调整。另外,需要注意的是,曝光触发事件量巨大,一般分析CTR,或者有推荐算法数据需求时需要曝光事件,其他场景请根据需求慎重埋点

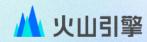
虚拟事件

虚拟事件是对元事件的合并和拆分,是一个特殊功能。所以在事件埋点设计时,如果虚拟事件可满足,不必增加新埋点。

合并事件:例如社交平台的互动行为包括:点赞、评论、喜欢、发送私信等很多行为,这时需求想分析当天发生互动行为的用户数去重。可对事件合并。

拆分事件:例如618电商活动期间,频繁看带满200减20标签的商品的加入购物车数量。不想重复性的配置指标,可设置虚拟事件。

新建虚拟事件			×
虚拟事件名称:	activelike		
展示名:	互动		
组合一般事件:	*一般事件之间是or的关系。即当任意一个事件触发时,视作该虚拟事件被触发		
	like_success (like_success)	~	
1	或 send_gift (send_gift)	~	
	touchpoint_click (touchpoint_click)	~	
ı	新增事件		
		取消	确定



THANKS FOR YOU

感谢观看

Design by DataFinder 2022/09

WWW.VOLCENGINE.COM