**会员标签方案**

**一、目的：**通过对会员打标更了解每个会员，对各大业务线产生数据价值。

1. **当前会员标签：**
2. 当前会员标签在HANA上进行计算，主要用途有以下几点：

A 精准营销计算需要

B 外部业务运营依赖数据需要

C 分析需要

对于A类需求，当前HANA内部自产自销；

对于B类需求，当前的方式是在HANA上计算落表，然后通过视图权限开放给外部，外部通过视图访问HANA表，然后将数据在他们自己的ES中落表，再封装前端提供服务；

对于C类需求，主要贡献在于简化了分析过程，当前也有一定应用。

总结来说，目前会员标签的情况为：**HANA作为会员标签主要计算系统，也承担会员标签的历史存储，同时，HANA会将最新的会员标签通过视图推送到电商系统建立ES查询库。**

截至20191025，具体会员标签详细情况见EXCEL。

当前这种模式下存在问题如下：

1. 从应用的角度说，数据由大数据管理部产生，由业务通过业务系统使用，但是整个过程大数据管理部无法监控标签，业务方需要增加标签只跟标签前端产品经理提需求，然后由标签前端产品经理跟大数据管理部提标签需求，导致在标签这个项目上大数据管理部远离了业务一步，无法参与标签内容及口径确定；
2. 从数据安全的角度说，标签数据整个放入外系统库导致无法对数据内容安全负责；
3. 从数据准确性的角度说，同步数据导致数据延时，两边至少存在半天的数据延时，导致标签不及时，如果数据产生系统报错，标签延时更严重；
4. 从标签内容本身来说，当前有太多统计定量数据，由于计算口径及延时问题，加上定量数据的非标准化问题，导致运营方式混乱并且依赖数据错误。
5. **目标：**

会员标签是标签体系的一部分，从数据安全性、唯一性、准确及时性的角度考虑，整个标签体系需要一个统一的数据服务中心，由该服务中心对标签进行发布、管理、下架、权限管控。

1. **会员标签目标：**

结合总体标签体系目标和目前实际情况，建议当前会员标签做如下改进：

可以关联的数据不落表，通过关联的方式以视图提供给外系统；

新上线会员标签除了特殊场景信息（如特殊时间或商品等）外只提供定性标签，即可选项有限的分类标签；

规划会员标签历史拉链表，提升标签数据查询服务能力；

下线当前会员标签时间切片表中不合理字段，减少HANA存储。

1. **资源预估：**

根据当前会员标签目标情况，预计达成目标具体情况及所需资源如下表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 目标描述 | 具体内容 | 预计所需人天 | 预计效果 | 难度 |
| 删除原标签中非刚需落表字段 | 1、在标签视图中修改会员标签表中依赖，能关联字段通过关联产生；2、删除这些字段 | 10天 | 释放原标签将近一半存储资源、至少缩短近1/3当前会员标签调度时间 | 中等 |
| 新会员标签视图开发 | 1、确定新会员标签内容范围；2、整理并上线相关必要数据；3、定义相关标签具体口径并与业务方确认；4、标签逻辑开发并上线 | 30天 | 业务系统可以使用大数据部门科学自定义定性标签，为业务人群营销及会员认知赋能 | 困难 |
| 上线会员标签历史拉链表 | 1、确定会员标签历史拉链表使用场景；2、确定会员标签历史拉链表标签内容、更新方式；3、会员标签历史拉链表逻辑开发并上线使用 | 10天 | 能很好地平衡查询效率与存储问题，是大数据管理部具备数据服务能力的重要基础工作 | 中等 |
|  |  |  |  |  |