

# 团 体 标 准

T/CCSA 555—2024

T/CAAAD 008—2024

## 数字营销技术 数据管理平台技术要求

Digital marketing technology - Technical requirements on data management platform

2024 - 07 - 03 发布

2024 - 10 - 01 实施

中国广告协会

中国通信标准化协会

发 布



## 目 次

前言 .....	II
引言 .....	III
1 范围 .....	1
2 规范性引用文件 .....	1
3 术语和定义 .....	1
4 缩略语 .....	2
5 概述 .....	3
6 平台功能要求 .....	3
6.1 数据收集层要求 .....	3
6.2 数据治理层要求 .....	4
6.3 数据管理层要求 .....	5
6.4 数据应用层要求 .....	6
6.5 数据输出层要求 .....	7
6.6 其他技术要求 .....	8
7 平台接口要求 .....	9
7.1 数据接入接口要求 .....	9
7.2 标签查询接口要求 .....	9
7.3 人群查询接口要求 .....	10
7.4 接口其他技术要求 .....	10
8 安全要求 .....	10
8.1 系统安全 .....	10
8.2 通信安全 .....	11
8.3 数据安全 .....	11
附录 A (资料性) 数据管理平台概述 .....	12
A.1 关于数据管理平台 .....	12
A.2 数据管理平台类型 .....	12
参考文献 .....	14

## 前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国广告协会和中国通信标准化协会共同提出，并分别归口。

本文件起草单位：北京腾云天下科技有限公司、上海亦拓广告有限公司、中国信息通信研究院、上海腾徽软件科技有限公司、加和（北京）信息科技有限公司、北京泛为信息科技有限公司、北京深演智能科技股份有限公司、北京巨量引擎网络技术有限公司、北京勾正数据科技有限公司、北京国双科技有限公司、秒针信息技术有限公司、北京京东世纪贸易有限公司、阳狮广告有限公司、华为终端有限公司、中国传媒大学、上海外国语大学、华扬联众数字技术股份有限公司、陕西省信息化工程研究院、上海优比客思科技有限公司、北京快手科技有限公司。

本文件主要起草人：闫辉、杨正军、朱岩、杨阳、张家绮、马良骏、张旭、顾明毅、范秋华、彭喜喜、曹玉骥、赵乃萱、马磊、吴充、于晓蕾、刘蒙恩、杨燕、欧阳书馨、张勇、蒲杰、落红卫。

## 引 言

为适应信息通信发展对标准文件的需求，由中国通信标准化协会和中国广告协会共同组织制定本文件，推荐有关方面采用。有关对本文件的建议和意见，向中国通信标准化协会和中国广告协会反映。

程序化广告类技术（DSP/SSP/ADX）诞生后，诞生了数据管理平台（DMP），主要是将多方分散的、可用于程序化广告业务的数据整合集中起来，整理加工成可以直接利用的数字资产。DMP可以通过对用户数字广告与营销相关行为轨迹的记录和分析，形成数字用户ID的标签数据库，提供给广告主与广告投放服务的相关平台，为其提供“准确识别”广告活动目标用户的数据支持。

当前，数字营销的数据管理领域已面临如下主要问题：首先，DMP的应用范畴在扩大，其分类和定位在多元化，业内缺乏统一认知的专业知识，不利于行业的交流和发展。其次，由于数字商业增长和竞争压力增大，DMP的功能不清晰和边界不明确，不利于营销技术行业有序竞争和核心技术发展。

另一方面，在移动互联网极速发展大背景下，无论从终端类型，还是数据维度，以及数据量级都攀升到很高水平，DMP的技术实现参差不齐，跨平台的技术对接协议和技术参数也缺乏标准，极大影响行业协作效率。

最后，随着我国颁布实施《中华人民共和国数据安全法》、《中华人民共和国个人信息保护法》等法律法规，数字营销行业对数据收集、使用、加工、提供和委托处理都提出了更高的安全合规要求。

所以，DMP在技术实现、安全合规处理、接口定义上都需要更清晰的行业技术规范来遵循。本标准的制定，一方面是期望向全行业普及DMP的基础概念并说明分类，来解决概念混淆和定义不清问题；另外更重要的是，在广告场景上建立标准规范化的数据接入、数据管理和数据使用等各项技术要求，达成行业共识，促进DMP及其上下游数字营销技术系统的健康、稳定合作与发展。



# 数字营销技术 数据管理平台技术要求

## 1 范围

本文件规定了数字营销领域中数据管理平台（DMP）的整体功能要求、技术实现要求、系统对接要求、安全要求等内容。

本文件适用于指导数字营销领域互联网广告场景下，相关的广告主、营销科技企业、互联网广告相关企业、监测服务企业、数据服务企业等，对数据管理平台（DMP）的设计、研发、合作及应用。

## 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 35273 信息安全技术 个人信息安全规范

T/CAAAD 003—2020 移动互联网广告标识技术规范

T/CAAAD 004-2022 T/CCSA 424-2022 互联网广告 匿名化实施指南

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

**数据管理平台 data management platform**

是指用于收集、组织和激活数字营销过程中来源数据，以支持受众挖掘、洞察、定向等程序化广告场景对数据应用的技术平台。

### 3.2

**数据分级 data classification**

数据分级按数据的重要程度和敏感程度，通过数据遭到破坏和泄露的影响、后果、损失与责任。

### 3.3

**数据资产 data asset**

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效益的数字化资源。在数字营销领域，也包括元数据、标签、人群数据。

[来源：GB/T 37550—2019, 2.4]

### 3.4

**源数据 raw data**

指在平台中收集到的，未经过处理或简化的原始数据。

### 3.5

**元数据 metadata**

描述其他数据或数据元素的数据，以及关于数据拥有权、存取路径、访问权和数据易变性的数据。

[来源：GB/T 5271.17-2010, 17, 06.05]

### 3.6

**标签 tag**

数字营销领域中的标签，是指通过一定的数据加工逻辑产出，能够为业务所直接使用的可阅读、易理解、有业务价值的信息。

### 3.7

**人群 audience segmentation**

也称受众群体，指具有共同行为特征的群体。

### 3.8

#### 人群画像 audience insights

也称受众洞察，指通过收集、汇聚、分析人群信息，对某特定人群的个人特征，如人口属性、兴趣爱好、行为偏好等方面作出分析或预测，形成画像统计分析结果。

### 3.9

#### 广告主 advertiser

为推销商品、提供服务或推广概念而发布广告信息的市场主体。

### 3.10

#### 媒体 media

发布、展示广告的载体。

### 3.11

#### 匿名化 anonymization

通过对信息的技术处理，使信息主体无法被识别或关联，且经处理的信息不能复原的过程。

[来源：GB/T 25069—2022，3.438]

### 3.12

#### 标识符 identifier

在数据组织中，一种用于标识某一数据元素或为其命名，并可能指明其某些性质的一个或多个字符。

[来源：GB/T 25069—2022，3.48]

### 3.13

#### 广告标识符 advertising Identifier

应用于互联网广告业务的一类非永久性标识符，能够标记、识别互联网广告服务对象，可以在保障用户隐私的前提下，通过此标识在互联网广告业务中建立广告数据基础级服务，促进数据流通。

### 3.14

#### 去标识化 de-identification

通过对信息的技术处理，使其在不借助额外信息的情况下，无法识别或者关联信息主体的过程。

[来源：GB/T 25069—2022，3.475]

### 3.15

#### 需求方平台 demand side platform

帮助广告主执行广告投放策略的平台，可以设定投放金额、单价、数量、物料等执行策略。

### 3.16

#### 供应方平台 supply side platform

帮助互联网广告媒体，经营广告资源销售的平台，记录了媒体销售的广告位、物料尺寸、售卖金额、库存等信息，主要包括流量管理、采买管理、售卖管理、广告源管理等核心功能。

### 3.17

#### 广告监测 ad monitoring

为监测和衡量广告效果、确保广告数据真实性、归因分析等目的而进行的广告相关数据监测与分析的活动。

### 3.18

#### 数据服务供应商 data service provider

数据服务提供者是一种拥有或可获得大数据服务所需数据资产的网络运营者，包括数据平台提供者、数据应用提供者和数据服务协调者三种角色的数据处理者。

[参考：GB/T 35274 信息安全技术 大数据服务安全能力要求 术语]

## 4 缩略语

下列缩略语适用于本文件。

ADX：广告交易平台（Advertising Deal Exchange）

API：应用程序接口（Application Programming Interface）

CDP：客户数据平台（Customer Data Platform）



DMP: 数据管理平台 (Data management platform)  
 DSP: 需求方平台 (Demand Side Platform)  
 FTP: 文件传输协议 (File Transfer Protocol)  
 RTB: 实时竞价广告 (Real Time Bidding)  
 SDK: 软件开发工具包 (Software Development Kit)  
 SSP: 供给方平台 (supply side platform)

## 5 概述

DMP同DSP、ADX、SSP等系统完成系统对接,进行数据交互,其在数字营销行业中的位置及与各平台间的业务交互逻辑可参考附录A。DMP总体功能架构如图1所示:



图1 DMP总体功能架构示意

DMP的功能架构应包含以下部分:

- 数据收集层负责DMP的数据接入,主要技术分为:批量数据接入和流式数据计入,方式分为主动采集和接入其他系统数据;
- 数据治理层负责针对进入DMP的数据,进行抽取、转换、加载等操作,功能上主要分为:源数据管理、数据关联融通、数据清洗;
- 数据管理层负责数据的区分、存储、加工、销毁等全生命周期的管理。功能上主要分为:标签管理、人群管理两个数据实体的全面管理,以及数据销毁、数据分级等;
- 应用层负责将加工后的数据应用于业务场景,具体到本标准就是应用于数字营销领域的广告投放场景。为满足此类场景,功能上主要分为:数据分析、人群画像、人群放大;
- 数据输出层负责DMP的服务发布、服务接口、服务规范等的制定和维护,还包含导出形式和前置机形式的输出功能。

## 6 平台功能要求

### 6.1 数据收集层要求

#### 6.1.1 数据采集要求

DMP应具备数据采集能力,通过自动化技术方法主动获取用户已授权数据。

数据采集功能相关要求如下:

- 数据采集宜支持以下SDK类型:Web JavaScript SDK、小程序SDK、APP SDK;
- 数据传输上,应支持HTTPS传输通信协议,提供IPv4、IPv6协议的支持能力,应支持基于消息和文件的传输方式,提供可靠的数据传输机制;
- 数据采集宜具备监控能力,提供网络运行状态、采集对象、任务执行情况监控,能提供异常情况预警,监控不影响任务正常执行;
- 数据采集宜通过对采集过程和结果的统计,验证是否正常采集数据,宜提供采集任务运行情况、任务资源使用、任务异常事件统计报表;
- 数据采集应主动披露使用的权限列表、收集使用的信息及目的等相关内容;

f) 应支持对采集对象的采集频度、数量、类型等进行管理，不过分采集。

### 6.1.2 数据接入要求

DMP应具备接入外部系统数据的能力，这些外部系统为广告流量系统、广告监测系统、客户数据平台、其他第三方DMP系统等，要求如下：

- 数据接入范围可涵盖广告行为日志、广告流量问询日志、数据标签；
- 数据接入方式应支持流式处理、批量处理多种可选方式。批量接入方式应支持数据的压缩和加密；
- 数据接入应具备对接入数据端的令牌管理和权限控制功能，应能够识别和校验数据端身份；
- 数据接入功能宜具备可视化界面来清晰查阅接入的数据源类型、数据类型、数据量级等信息，并可在界面端管理接入数据端的令牌和权限；
- 平台宜具备对接入数据质量监控的功能，并能够以可视化报表形式提供接入监控报告；
- 接口技术要求详见：7.1 数据接入接口要求。

## 6.2 数据治理层要求

### 6.2.1 数据模型和对象要求

DMP的整体数据模型和数据对象间关系可描绘为图2：

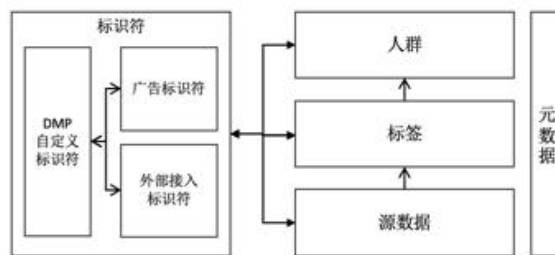


图2 DMP主要数据对象关系示意

整体数据模型应由两大类数据对象构成，数据的ID标识和主要数据实体源数据、元数据、标签、人群，DMP整体管理了ID标识与数据实体间的关系：

- 标识符方面宜具备DMP内部的匿名化标识符，用来与广告标识符以及外部接入标识符形成整体的统一ID视图；
- 数据实体中应以元数据来管理底层的源数据数据资产，并为标签和人群加工提供数据原料，标签一方面是源数据加工后的业务可用资产，也同时为人群构建、人群圈选提供数据原料。

### 6.2.2 标识符与关联要求

DMP中使用的标识符要求如下：

- 若包含移动互联网广告标识，应符合T/CAAAD 003-2020中对广告标识符的要求；
- DMP应对收集到的广告标识符做匿名化处理，匿名化标识符应符合T/CAAAD 004-2022 T/CCSA 424-2022中相关要求。

DMP应具备对相同ID标识符下一方数据、二方数据、三方数据的ID关联匹配能力，可通过关联匹配、知识图谱聚合等技术以匿名化ID和广告标识符来创建多ID间关系和关联。

可支持对一方、二方、三方数据的以下几类关联匹配：

- 可支持对不同数据源的、同类标识符的匹配识别ID进行关联匹配；
- 可支持对相同操作系统平台下不同广告标识符格式的标识ID进行关联匹配，例如对不同时期和版本的Android ID与OAID的关联，不同时期和版本的IDFA与CAID的匹配关联等；
- 可支持对异类标识符数据的关联打通，如：OTT广告标识符和移动设备广告标识符的关联匹配等；
- 可支持匿名化ID创建，是在关联匹配过程中对不同格式的广告标识符归一至新生成的匿名化ID，并将广告标识符作为属性挂载在匿名化ID中的过程；

- e) 可支持 ID 关联匹配的统计分析功能，是通过指标监控和追踪匹配关联的质量、比例情况的功能。常用监控指标为：
  - 1) 识别率：DMP 可以匹配识别其他数据源的同类标识符的比例。
  - 2) 打通率：DMP 可以映射打通异类标识符的比例。

### 6.2.3 源数据管理要求

DMP的源数据管理功能要求如下：

- a) 应支持数据目录的管理，提供不同用户逻辑独立的目录服务，支持建立数据采集对象、数据存储、数据服务目录；
- b) 应具备源数据存储功能，提供源数据库，用于源数据的持久化存储及维护；源数据库应支持数据库管理系统、文件系统等不同物理存储方式；
- c) 宜具备源数据查询功能，支持类 SQL 查询语言，支持对文本类数据类型的全文检索功能，提供查询结果的对象元组集的过滤、排序等功能；
- d) 宜支持源数据变更追溯功能，支持变更通知功能，支持版本管理功能，包括最新、历史版本的增删改查、恢复、分支与合并等，以便管理和跟踪源数据的整个生命周期状态；
- e) 源数据写入操作宜提供事务与回滚功能。

### 6.2.4 数据清洗要求

DMP的数据清洗功能要求如下：

- a) 数据抽取过程应支持全量、增量、和基于日志等数据抽取模式，应支持不同数据格式的抽取，例如：MySQL binlog、CSV、Sequence、Parquet 等文件格式；应支持主流安全可控的数据库系统、数据仓库系统或文件系统，例如：MySQL、Oracle、DB2、HIVE、HDFS 等；可支持常见关系型数据库、数据仓库、文件系统或队列的数据加载，比如：MySQL、Oracle、DB2、HIVE、HDFS、Apache Kafka 等；
- b) 应提供数据校验功能，如空值检测、数值范围检查等；
- c) 应提供数据表转换功能，如字段值替换、值映射、记录合并、去除重复记录等；
- d) 应提供数据运算处理功能，如求和、平均值等，提供基于代码映射的自动转换功能。支持数据的计算、清洗、比对的服务，排除错误、重复、不完整的数据；
- e) 数据加载应支持容错处理，如：重新加载、根据加载错误率、错误数进行熔断；
- f) 应提供监控管理功能能够对数据处理状态进行实时监控和预警，例如：抽取条数、转换出错数、空值数等指标，同时能够根据指标进行预警，对于达到预警的指标能够及时发送预警信息；
- g) 应具有完整的日志管理功能，提供对监测数据的查询及汇总功能和对清洗过程的日志进行查询。

## 6.3 数据管理层要求

### 6.3.1 标签管理要求

DMP应具备标签管理功能，标签管理功能的要求如下：

- a) 应支持新建/定义标签的功能：对于定义标签，需保证用户使用清晰的数据源、数据操作方式，完成组合加工，并了解其对应的设备、人群数量；
- b) 应支持规则标签的加工方法，通过人工经验定制标签运算规则，比如交集、并集、差集等；
- c) 宜支持算法标签的加工方法，通过一些真值数据，作为正负样本，再通过模型预测出的标签；
- d) 应支持删除标签功能：DMP 平台需在用户账户下提供相关功能，删减在平台上的标签。通过逻辑、物理删除方法，确保此标签不再被使用；
- e) 应支持修改标签功能：DMP 平台需在用户账户下提供相关功能，对“标签名称、标签定义”等基础信息做修改。如果希望修改标签的加工逻辑，对于自定义标签，需实现自定义修改功能。对于逻辑较复杂的标签，可支持后台加工的自动应用；
- f) 应支持查询和检索标签功能：由于标签的树状层级关系，在查看过程中，需要保证以下几个信息的清晰展示。包括”标签的层级、最近更新日期/时间，此标签对应的最新的人群数量、

设备数量，标签变化趋势，标签被哪些人群包所使用，标签的加工逻辑，标签所使用的数据来源，标签当前的加工状态”；

- g) 标签管理过程宜支持以下信息字段：标签版本号，标签根目录，标签类目，标签名称，标签加工类型，标签逻辑，值字典，取值类型，示例，标签描述，对应的设备/人群数量，创建、更新时间，加工计算状态，更新频次。

### 6.3.2 人群管理要求

DMP应具备人群管理功能，人群管理功能的要求如下：

- a) 人群创建/圈选规则应支持通过多个标签的选择、排除等组合，或者多个已知人群包的选择、排除组合，构建人群形成人群包。功能上应支持标签的组合，支持交集、并集、差集、非集的操作，支持多层级的组合操作；
- b) 人群信息中宜包括以下字段：人群名称，人群 ID 类型，构建状态，人群量级，创建规则，人群数据源，最近更新时间，创建时间，创建规则；
- c) 人群计算应支持多种更新方式，如定期更新、触发式更新。定期更新时间粒度应支持天级更新；触发式更新应至少支持人工触发；
- d) 应支持删除人群功能：支持人群的下线处理，不再被使用，清除人群相关的数据；
- e) 应支持修改人群功能：支持对人群的名称、创建规则的修改；
- f) 应支持查询和检索人群功能：支持按照人群名称搜索，展示最近更新日期/时间、人群对应的最新的广告标识符数量、人群变化趋势、人群的加工逻辑、人群所使用的数据来源、人群当前的加工状态、和其他人群的关系；
- g) 应支持复制人群功能：支持复制已知人群包的基本信息、规则、加工逻辑；
- h) 宜支持推送人群数据功能：通过安全对接的方式，推送人群数据至投放系统、平台，对接其他来源的人群，并在 DMP 平台展示、查询、使用。支持对接以下几类平台，包括媒体，其他类型的 DMP，DSP，广告主内部数据平台。可通过 API、前置机、FTP 等方式对接；
- i) 人群圈选针对千万级的所选人群，DMP 宜保障在 30 秒内完成人群数量的统计计算，并呈现给平台操作人员。

### 6.3.3 数据分级要求

DMP宜具备数据分级功能，对不同源数据、标签和人群按重要程度和敏感程度设定不同分级。

具体功能要求如下：

- a) 应为 DMP 的使用者提供清晰的操作界面，实现对数据的分级标注。标注的维度应包括数据来源、数据内容；
- b) 宜建立数据分级清单：根据数据分级情况对业务运营者数据资源进行分类分级后，输出业务运营者的数据分级清单；
- c) 可在交互界面，针对操作人员，针对不同数据分级提供及时和清晰的安全提示；
- d) 可定期提供不同分级数据使用情况的分析报告，及时发现问题。

### 6.3.4 数据销毁要求

DMP应具备数据销毁功能，在确保数据安全、确认可被删除的前提下，实现物理删除、逻辑删除功能。应符合以下要求：

- a) 应支持数据有效期过期的数据删除功能，包括授权过期数据、备份过期数据、临时数据、中间数据、日志数据等；
- b) 可根据不同的类型和级别的数据制定不同的删除策略，如自动化删除、人工确认删除。

## 6.4 数据应用层要求

### 6.4.1 数据分析要求

DMP应具备对数据资产的分析功能，要求如下：

- a) 宜支持标签相关资产的业务指标分析，包括：标签饱和度，说明某个标签在人群中的分布比例；标签覆盖度，说明全部标签可以覆盖人群的比例；标签使用率，说明此标签被多少个人群包使用；
- b) 宜支持人群相关资产的业务指标分析，包括：人群覆盖度，反映目标人群包覆盖 ID 占 DMP 总体 ID 的比例；人群查得率，反映人群包在投放流量查询时的查中比率；人群活跃度，反映人群包在近来一段时间周期中活跃 ID 的占比，常用周期 7 天、14 天、30 天、90 天等；
- c) 宜支持源数据相关资产的业务指标分析来考核源数据资产的可用性：在源数据的质量方面的分析，可包括：记录数、非空数、空值数、有效数、重复数、非重复数，质量分析指标上可支持饱和度、填充率、准确率、缺失率。分析中采用的数据计算方法可包括：最大值、最小值、平均值、众数、标准差、方差、变异系数、特殊值、精度、频率分布等；
- d) 应支持数据可视化方式构建和查阅数据分析结果。宜以图形化方式呈现数据总量、条数、去重条数、数据增量、数据质量、数据分级情况、分布情况，图表类型可选用线图、面积图、柱形图、饼图、表格图等，宜支持不同图形形式的切换；
- e) 应支持主要的数据筛选维度，宜包括：时间维度（年月日的周期）、标签过滤筛选维度、人群过滤筛选维度等；
- f) 对标签和人群宜支持数据目录的展示形式，以多层级表格列表、树状列表展示全量的标签和人群资产，并可查阅其层级结构和位置。宜提供标签、人群数据间关联关系的查阅，例如人群构建使用了哪些标签，便于分析其相关性；
- g) 宜支持数据检索分析能力，通过对关键字的匹配和模糊匹配，检索出相关数据集结果，并提供数据列表展示和统计关联检索结果总量等功能。

#### 6.4.2 人群画像要求

DMP 应具备人群画像功能，分析目标受众人群宏观特征。要求如下：

- a) 应支持人群画像任务管理功能，宜包括以下信息：人群名称、人群 ID 类型、画像构建状态、画像人群量级、最近更新时间和创建时间；
- b) 人群画像结果应支持以下分析指标：识别率：能够被识别的，且可以被描述的实体标识数量；TGI 指数： $[\text{目标群体中具有某一特征的群体所占比例} / \text{总体中具有相同特征的群体所占比例}] * \text{标准数 } 100$ ；特征浓度：符合某特征的人群数量/画像人群总数量；分布比例：在某个区域的人群数量/画像人群总数量；
- c) 人群画像宜支持以下维度：时间维度（年季度月日时分秒的周期）、地域、标签、设备类型、设备信息、行为事件维度等；
- d) 人群画像的可视化宜支持以下图表类型：趋势图、分布图、漏斗图、热力图、路径图、关系图、仪表图、比例图，宜支持以下图表图形的切换展示：表格、柱状图、线形图、饼图、雷达图。

#### 6.4.3 人群放大要求

DMP 宜支持人群放大功能，人群放大功能的要求如下：

- a) 人群放大功能应具备便捷的种子人群选择功能，可提供从系统已有人群全部抽取和随机抽取指定数量来构建种子人群的功能。系统应对种子人群具备最小 ID 量和最大 ID 量的限制功能，以确保放大执行的有效性；
- b) 人群放大功能应提供放大后目标量级和目标倍率的设定功能，功能中应具备放大后的最大人群包量级控制功能，避免无限制不计效果的无限放大；
- c) 人群放大功能应支持使用者来选择放大后人群包是否要包含原有种子人群的功能；
- d) 经由放大生成的人群应支持用于人群画像和定向投放，放大后的人群宜根据数据授权和安全性考虑来增加对人群包数据导出的限制；
- e) 人群放大功能应具备足够的计算效率，具体耗时应根据种子量级和目标人群量级来控制，推荐所有的放大任务宜不超过 6 小时来提供结果。

### 6.5 数据输出层要求

#### 6.5.1 数据导出功能要求

DMP对于使用者享有合法权属的数据应提供数据导出功能，应支持以下能力：

- a) DMP 可支持多种形式的导出数据，至少支持以下任意一种：
  - 1) API 对接导出：双方通过 open-api 对接的方式进行数据传输；
  - 2) SFTP 导出：由使用方提供 FTP/SFTP 服务器，DMP 根据使用方给出的连接方式将数据传输至 FTP/SFTP 服务器；
  - 3) UI 界面手工导出：由使用方在 DMP 页面构建下载任务进行下载。
- b) 导出应支持多种数据格式，如数据文件：例如纯文本、带格式的文件；报表文件：例如图表、表格等。

### 6.5.2 接口服务功能要求

DMP应具备接口服务功能，为DSP、ADX以及其他的数据管理型系统提供数据应用能力，要求如下：

- a) 接口服务功能可涵盖标签查询、人群命中查询、画像统计结果等主要投放查询场景；
- b) 标签查询功能，应支持查询标签列表，应支持查询以广告标识符入参的标签明细值。接口技术要求详见：7.2 标签查询接口要求；
- c) 人群查询功能，应支持查询人群概要信息，应支持以广告标识符入参查询人群命中状态结果，应至少涵盖命中、未命中、不识别等不同状态结果。查询接口宜同时支持 RTA 场景的人群命中查询。接口技术要求详见：7.3 人群查询接口要求；
- d) 接口服务应具备令牌管理和权限控制功能，应能够识别和校验调用方身份，针对不同的访问方分配令牌和密钥。

### 6.5.3 前置机服务要求

为解决跨域访问时效问题，DMP宜支持前置机功能，要求如下：

- a) 前置机服务应支持查阅查询量、查中量、查中率等统计指标的功能。应支持 T+1 日查阅依据前置机、活动、媒体、按设备标识符类型等维度的指标分析和洞察；
- b) 前置机服务应支持节点自动部署功能，允许跨域使用方自助快速部署及扩展服务环境；
- c) 前置机服务应支持人群命中查询功能，宜根据流量类型（PC/MO/OTT）、TA(性别年龄)、投放项目信息封装订单模式人群包以供查询，查询结果应能区分查中、未查中、未识别等多种状态。接口技术要求详见 7.3 人群查询接口要求。

### 6.5.4 数据输出的控制功能

DMP应具备按不同数据分级控制数据输出的功能，要求如下：

- a) 在交互界面中应支持查阅到目标数据的分级类型，针对输出操作提供及时和清晰的安全提示，经单独的确认交互操作后方可执行导出；
- b) 在权限设定中应针对分级数据导出权限，提供单独的功能设定，支持对单操作人员级进行权限的增设、编辑、取消；
- c) 在无交互界面，以程序自动化形式对接输出数据时，应具备明确的白名单和访问令牌机制，令牌需定时更新刷新；
- d) 系统中应保存所有数据导出的操作日志，至少包含操作账户、操作时间、导出数据概况、操作者 IP 等信息；
- e) 系统宜定期提供不同分级数据输出情况的统计分析报告，以便管理人员及时发现问题。

## 6.6 其他技术要求

### 6.6.1 可用性要求

系统应按照最大容量的80%或标准压力（系统的预期日常压力）情况下运行，能够稳定运行7×24小时，并且系统的各项资源指标没有泄漏或异常。

### 6.6.2 可扩展性要求

系统应支持通过增加关键通信线路、关键网络设备和关键计算设备的硬件来提高系统的可扩展性。

### 6.6.3 可靠性要求

系统应提供通信线路、关键网络设备和关键计算设备的硬件冗余部署来保证系统的可靠性；还应提供数据的备份和恢复机制或通过数据多副本存储来提高系统的可靠性。

#### 6.6.4 易用性要求

系统的表现和表述应贴近用户所在的环境和专业领域。对不同组成部分之间的设计目标、元素外观、交互行为模式应保持一致。针对容易引起错误发生的情况，应在用户完成提交之前系统能自行进行检查，并让用户确认。

#### 6.6.5 可观测性要求

系统应记录详细的系统运行日志，在业务的关键点应记录业务日志、同时针对关键数据的变更也应有日志。日志的内容应包含：时间、日志级别、文件、行号、说明信息。

日志记录应包括事件的日期、用户、事件类型、事件是否成功及其它相关的信息；应对日志记录进行保护，定期备份，避免受到未预期的删除、修改或覆盖等。备份数应压缩存储并异地备份。

### 7 平台接口要求

#### 7.1 数据接入接口要求

DMP数据接入接口要求如下：

- 对于广告行为日志、数据标签等接入场景应满足 T+1 的数据处理能力；
- 传输字段宜包含下述表格中参数见表 1。

表 1 数据接入接口传输请求参数

必要参数	描述	样例
广告标识符类型	广告标识符类型即广告标识符的类型，其包括但不限于 IDFA, OAID及其MD5	idfa_md5
广告标识符	广告标识符	D4C00E03AC3BFAFE392A310E5D35807F
时间戳	事件发生时间，为long型毫秒级UTC时间（即1970年1月1日到当前时间的毫秒数）	1437118969422
ip地址	设备的IP地址	36.110.87.173
广告点位ID	投放媒体广告位ID（通常广告流量系统线下提供媒体名称和媒体标识的对应关系）	12345
媒体ID	投放媒体标识ID（通常广告流量系统线下提供广告位ID和广告位名称的对应关系）	123
广告行为类型	如：曝光、点击	impression
标签列表	客户数据平台中向匿名ID标记的标签code列表，通常每个标签需附带权重信息和标签版本信息，数据格式需经双方系统共同约定	030101:0.76:1,030102:0.88:1

#### 7.2 标签查询接口要求

DMP标签查询接口要求如下：

- 接口应使用 https 通信协议；
- 标签查询实时接口的响应时间应小于 5 秒，批量异步接口的结果返回时间应小于 24 小时；
- 传输字段宜包含下述表格中参数见表 2、表 3。

表 2 标签查询接口请求字段

参数名	类型	描述	是否必选
id	string	广告标识符，应符合T/CAAAD003-2020对广告标识符的要求	是
idType	string	广告标识符类型	是
otype	string	输出的标签列表	是
filter	string	过滤条件	是

表 3 标签查询接口返回字段

字段	类型	描述
code	int	业务状态码
msg	string	状态值描述
seq	string	流水号，全局唯一
data	string	数据结果

### 7.3 人群查询接口要求

人群归属查询接口要求如下：

- 如果接口在公网环境下，则应使用 https 通信协议；
- 接口响应时间应在 20ms 以内，应具备 20 万 QPS 的数据响应能力；
- 传输字段宜包含下述表 4、表 5 信息。

表 4 人群查询接口请求字段

参数名	类型	描述	是否必选
mediaId	string	媒体标识，由数据平台方提供	是
campaignId	string[]	活动人群标识列表	是
id	string	通常为广告标识符，广告标识符应符合 T/CAAAD003-2020 对广告标识符的要求	是
idType	int	广告标识符类型，通过约定数值表达对标识符的加密类型	是

表 5 人群查询接口返回字段

字段	类型	描述
code	int	业务状态码
result	string	查询结果，数据结构由数据平台提供
campaignId	string	人群标识
matched	string	匹配状态，返回结果如下： true: 查中 false: 未查中, DMP 未识别 N: 未查中, DMP 可识别

### 7.4 接口其他技术要求

DMP 平台各类接口的通用技术要求如下：

- 应具备接口节点的扩缩容能力，宜具备节点自动化部署能力；
- 接口的 SLA 应不低于 99.9%；
- 接口宜具备降级和熔断机制，宜具备高可用的多区部署特性；
- 接口应能够按版本实现向下兼容，也应支持不同版本分别部署或并存部署；
- 接口的流量负载节点应具备应对常见网络攻击的防护能力；
- 接口的安全控制应在系统接口接入点的网络边界实施。应具备对接口的访问控制和安全审计能力。

## 8 安全要求

### 8.1 系统安全

系统安全需符合下列要求：

- 平台应采用安全网络设备保证平台系统的硬件层安全，如路由器、交换机、防火墙、入侵检测系统等；
- 应关闭非必要的业务数据（外部）接口以保证接口安全，应支持接口访问的授权验证和来源验证；



- c) 应支持不同账户的分级和最小必要权限授权管理，如管理员、运维人员、访客等，根据不同等级账户需要应具有不同的管理权限；
- d) 应具备系统漏洞修复能力，不应含有 6 个月上的 CNVD, CNNVD 高危及以上漏洞；
- e) 应具有安全报警、安全审计、日志记录功能，以保证各操作行为的可追溯性；
- f) 应具有更新功能，更新时应能够对更新来源进行鉴别，并对更新文件进行完整性校验；
- g) 应采取白名单控制策略，并使用隐藏或掩码方式防止各类攻击；
- h) 平台宜符合 GB/T 22239 三级要求。

## 8.2 通信安全

接口安全需符合下列要求：

- a) 应采用双向身份认证技术保证平台各接入方的身份合法性；
- b) 应采用安全通信协议保证平台各接入方的通信安全，如 1.2 及以上版本 TLS 等；
- c) 应采用密码技术保证通信中关键数据的保密性和完整性等。密码算法应符合相关国家标准和行业标准的的要求，宜使用商用密码算法。

## 8.3 数据安全

数据安全应符合以下要求：

- a) 平台数据处理应符合 GB/T 41479-2022 相关要求；
- b) 对平台中涉及到个人信息处理的应符合 GB/T 35273 相关要求；
- c) 平台应遵循最小化数据原则，数据采集与接入、加工治理、管理与应用、输出均应控制最小必要信息，并降低量级、时限、频次；
- d) 对数据和个人信息处理中采用的密码技术应符合相关国家标准和行业标准的的要求，宜使用商用密码算法；
- e) 宜根据数据分级要求，实施严格的数据访问控制措施，包括强密码策略、多因素身份验证等，以限制对敏感数据的访问权限。在与不同平台进行数据的接入和输出时，需确保数据安全标准和合规要求；
- f) 平台宜采用多租户的设计，各租户之间的数据、功能、用户权限、角色和性能应相互隔离互不影响，以确保不同组织和租户的数据安全级别。

## 附录 A (资料性) 数据管理平台概述

### A.1 关于数据管理平台

数据管理平台DMP同DSP、ADX、SSP等系统完成系统对接，进行数据交互，常附属于广告投放系统而存在。在更成熟的数据营销全场景下，DMP系统与更多系统产生对接，形成多方协助的DMP，以及独立技术服务商搭建的DMP。

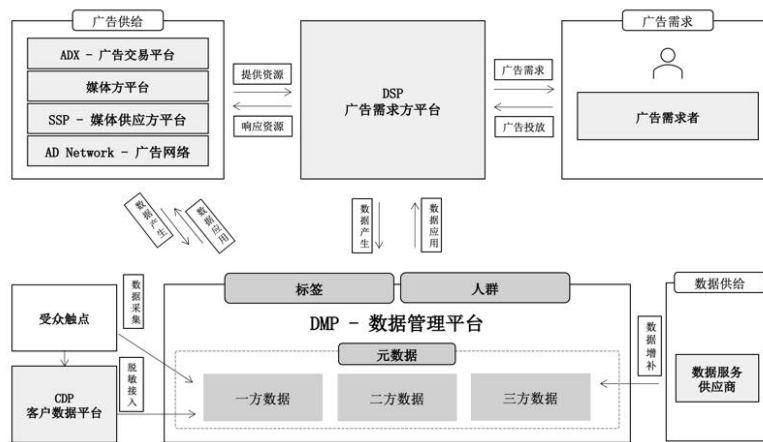


图 A.1 DMP 在数字营销中的业务逻辑示意图

如数字营销中的业务逻辑示意如图A.1所示：

- a) DSP、ADX、SSP、广告网络、媒体方平台等系统平台中所产生的实时竞价广告过程数据成为DMP的数据源，接入DMP中。
  - b) DMP系统提供统一的数据聚合、加工能力，以加工后的标签、人群形式为DSP、媒体方平台提供目标受众定向能力，实现数据的应用。
  - c) DMP系统中可对具备授权的匿名化受众触点数据进行直接数据采集。
  - d) 实名化受众触点数据，可经由CDP系统平台接入后，经过脱敏处理后再行接入DMP。
- DMP可由数据服务供应商进行数据增补，补充三方数据标签和人群。

### A.2 数据管理平台类型

以广告主视角来看待，数据管理平台可分为第一方DMP、第二方DMP、第三方DMP三种类型，这样的分类方法是以平台的建设和管理方为判定依据。

第一方DMP是以广告主第一方数据为主要数据源和数据管理对象的数据管理平台，用于聚集、分析和管理的数字营销相关数据，为营销环节提供决策支撑和数据支撑。

DMP可以收集和分析一方数据，这样做通常为了采集营销着陆后的触点行为信息和访客的内容兴趣信息，采集基于广告标识符的信息。基于用户授权的一方数据，通常使用客户数据平台（CDP）进行收集和管理，然后经脱敏提供给DMP，用于形成第一方DMP的标签和画像。

第一方DMP与CDP的比较，见下表A.1。

表 A.1 第一方 DMP 与 CDP 比较

主要差异	第一方DMP	CDP
数据类型	脱敏数据为主	用户数据为主
数据存储	临时数据+长期数据	长期数据为主
处理量级	极大	大
多触点数据打通	是	是
数据输出	标签、人群	标签、人群

第二方DMP是以互联网广告媒体产生的数据为主，用于互联网广告程序化投放的数据管理平台。第二方DMP一般是由媒体、广告投放方、广告监测服务方等提供。

第三方DMP是以数据增补、数据交易为主要形式的DMP，一般是其他行业领域的数据供应商提供的数据管理产品，为需求方提供数据交换、数据分析等服务，但是有的应用场景下，会将第三方数据监测机构提供的数据工具也称为“第三方DMP”。第三方数据同广告主无直接关系，第三方数据供应商提供的数据对广告主而言就是第三方数据，这些数据是指从多种来源收集、打包到一起，并可供获取使用权的数据。

无论哪一方为主导来建立的数据管理平台，都会涉及管理多种来源的数据，从来源方来评判，可以总体将数据分为一方数据、二方数据、三方数据，要求的数据来源授权路径有所不同。

一方数据 first party data，是指企业经用户授权后，直接从受众（包括客户、网站访问者、移动应用程序或小程序的使用者和社交媒体账户的关注者）收集的数据。这里的“一方”是指收集第一手数据用于重新处理、分析和使用的一方。简单来说就是广告主自己采集、处理、加工并使用的数据，并且是这些数据的所有者。这类数据中包含人口统计信息，个人信息或个人敏感信息，与广告主方的通信、互动和消费行为信息等。

二方数据 second party data，是指数字营销过程中的合作企业（通常包含媒体方、广告投放技术提供方、广告监测技术提供方等）所产生和掌管的合规授权用于数字营销投放的过程数据。广告主并不直接采集、处理、加工二方数据，也没有这些数据的所有权，仅有这些数据的有限使用权。这些数据同样可能包含人口统计信息，个人信息（通常不再包含个人敏感信息），并且数据规模和数据字段要比广告主自己的第一方数据要丰富。

三方数据 third party data，是指由广告业务中没有直接合作关系的第三方主体所享有，并用于为需求方提供数据服务的数据。

### 参 考 文 献

- [1]GB/T34090.2-2017 互动广告 第1部分：术语概述
  - [2]GB/T 34090.2-2017 互动广告 第2互动广告部分：投放验证要求
  - [3]GB/T 34090.3-2017 互动广告 第3部分：效果测量要求
  - [4]GB/T 25069-2010 信息安全技术
  - [5]GB/Z 28828-2012 信息安全技术 公共及商用服务信息系统个人信息保护指南
  - [6]GB/T 35273-2020 信息安全技术 个人信息安全规范
  - [7]GB/T 37964-2019 信息安全技术 个人信息去标识化指南
  - [8]GB/T 37932-2019 信息安全技术 数据交易服务安全要求
  - [9]GB/T 41479-2022 信息安全技术 网络数据处理安全要求
  - [10]GB/T 22239-2019 信息安全技术 网络安全等级保护基本要求
  - [11]T/CAAAD003-2020 移动互联网广告标识技术规范
  - [12]T/CAAAD002-2020 互联网广告 投放监测及验证要求
  - [13]T/CAAAD001-2021 互联网广告 数据应用和安全技术要求
  - [14]T/CCSA 424-2022 互联网广告 匿名化实施指南
-