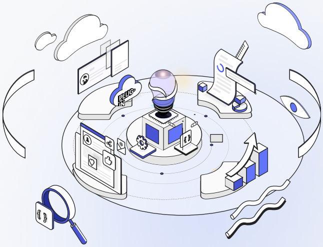
by leeC

Convertlab

Data Hub

一款基于多年业务最佳实践沉淀出的低代码 CDP。跨平台整合全渠道客户数据，输出精准客户洞察，指导营销决策并高实时性地支撑用户运营，实现企业数据资产的高效变现与增值。



↓

Background

项目 — 背景 & 功能 — 介绍

Convertlab 全新 CDP 产品。要打造世界一流的客户枢纽平台

0 - 1

产品 0 -1，团队也大多为新员工

数据整合

多端数据进行整合管理

数据清洗

丰富的标签体系和精准的数据分析

安全与合规

去标识方案、溯源链路

User Properties

用 — 户 — 属 — 性

因为产品专业性很强，决策者用产品的可能性微乎其微，购买流程一向是由使用部门试用，然后向决策者推荐并付费（如 figma），既然试用的决策者不多，那么目前基于工作流的功能体验，尽量减少使用者的学习门槛，用户属性的研究也着重在使用者。Datahub 使用门槛极高，对用户的专业性要求很高，通常为客户公司的数据运营、数据分析师。

调研途径

因为时间紧迫，采取下列性价比最高的方式

- ✓ 竞品用户：包括神策、明略、Segment...
- ✓ 以及预购的内部测试用户：如吉利汽车

—— Convertlab

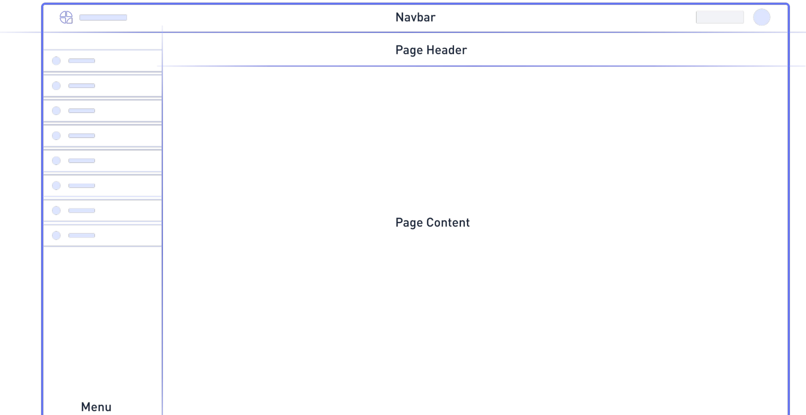
初步结果

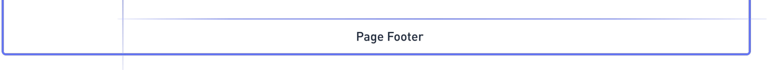
通过竞品的用户报告和对自己实际用户的定性分析，总结出以下标签：

- ✓ 数据运营
- ✓ 专业性强
- ✓ 数据分析
- ✓ 性别男
- ✓ SQL
- ✓ 理性

—— Convertlab

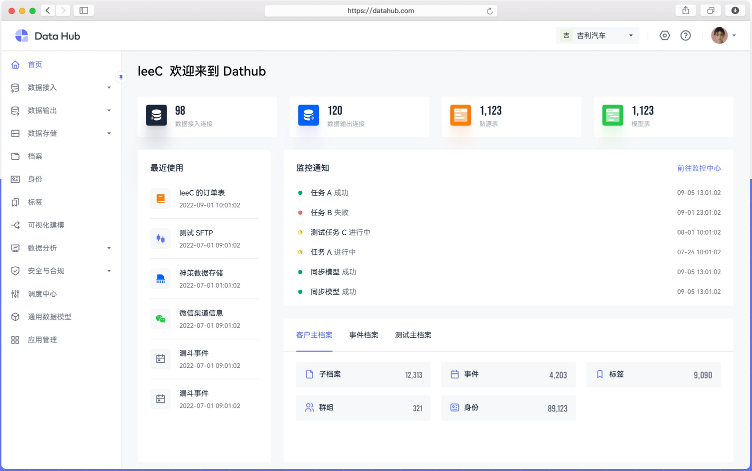
基础页面布局


Navbar
Page Header
Page Content
Menu


Page Footer

首页

除了预览重要数据意外，用户刚进 Datahub 还比较关心进行中的任务状态。最近使用的实体在首页也方便用户继续进行上次的任务。


https://datahub.com
Data Hub
吉 吉利汽车
首页
数据接入
数据输出
数据存储
档案
身份
标签
可视化建模
数据分析
安全与合规
调度中心
通用数据模型
应用管理

leeC 欢迎来到 Dathub

98	数据接入连接
120	数据输出连接
1,123	贴源表
1,123	模型表

最近使用

leeC 的订单表	2022-09-01 10:01:02
测试 SFTP	2022-07-01 09:01:02
神策数据存储	2022-07-01 01:01:02
微信渠道信息	2022-07-01 09:01:02
漏斗事件	2022-07-01 09:01:02
漏斗事件	2022-07-01 09:01:02

监控通知 前往监控中心

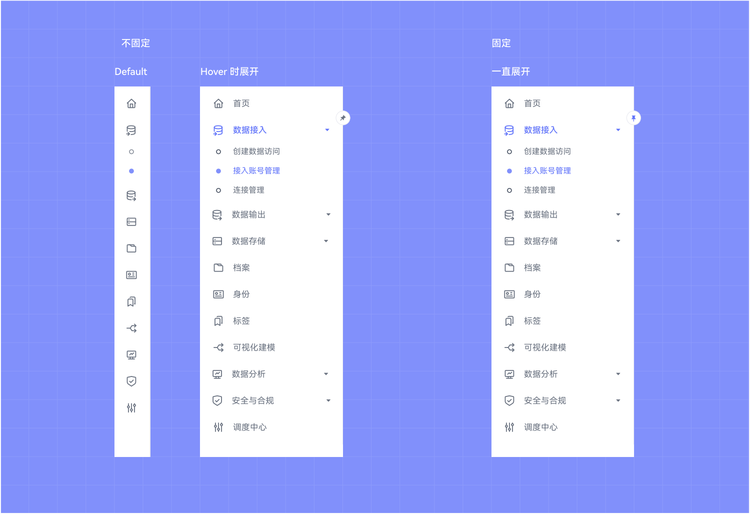
● 任务 A 成功	09-05 13:01:02
● 任务 B 失败	09-01 23:01:02
◔ 测试任务 C 进行中	08-01 10:01:02
◔ 任务 A 进行中	07-24 10:01:02
● 同步模型 成功	09-05 13:01:02
● 同步模型 成功	09-05 13:01:02

客户主档案 事件档案 测试主档案

子档案	12,313
事件	4,203
标签	9,090
群组	321
身份	89,123

菜单

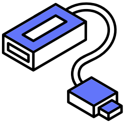
B 端项目的页面空间十分宝贵，当用户专注于页面内容时，占据不少横向空间的菜单收起来更好，但通过频繁的单击来让菜单展开/收起是比较麻烦的。采用 pin 的交互方式，更有一种用完即走的感觉。

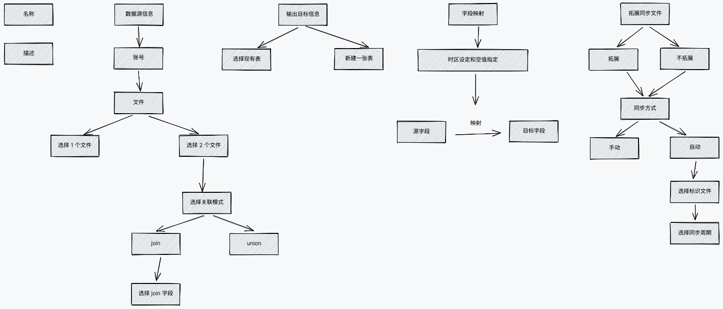

不固定
Default
Hover 时展开
首页
数据接入
创建数据访问
接入账号管理
连接管理
数据输出
数据存储
档案
身份
标签
可视化建模
数据分析
安全与合规
调度中心
固定
一直展开
首页
数据接入
创建数据访问
接入账号管理
连接管理
数据输出
数据存储
档案
身份
标签
可视化建模
数据分析
安全与合规
调度中心

Data Connection

数 — 据 — 连 — 接

创建一个连接器需要的信息多且杂，相互之间的联动也多，拿到PM的文档后进行大致的分析，整理出用户需要填写的内容并联动，从而理出更合适的交互。

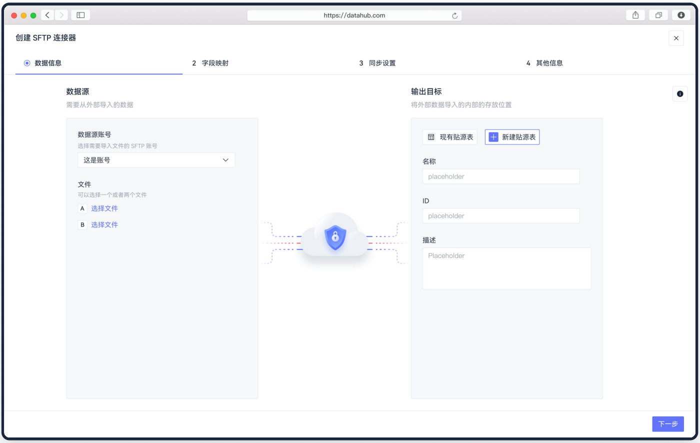



名称
描述
数据源信息
账号
文件
选择 1 个文件
选择 2 个文件
选择关联模式
join
union
选择 join 字段
输出目标信息
选择现有表
新建一张表
字段映射
时区设定和空值指定
源字段
映射
目标字段
拓展同步文件
拓展
不拓展
同步方式
手动
自动
选择标识文件
选择同步周期

1

建立连接模型

根据尼尔森可用性原则， Match between system and the real world。强调连接器是将两个东西连接，东西从一个地方导入另一个地方，所以在用分布表单完成创建的路径中，第一步就是建立该模型。


https://datahub.com

创建 SFTP 连接器

×
● 数据信息
2 字段映射
3 同步设置
4 其他信息

数据源

需要从外部导入的数据

数据源账号

选择需要导入文件的 SFTP 账号

这是账号

文件

可以选择一个或者两个文件

A 选择文件

B 选择文件

输出目标

将外部数据导入的内部的存放位置

现有贴源表
新建贴源表

名称

placeholder

ID

placeholder

描述

Placeholder

下一步

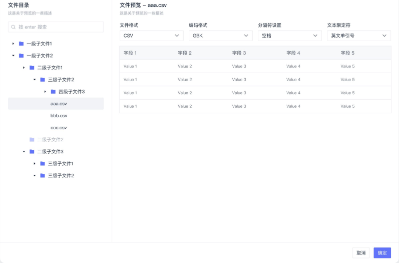
选择文件

THE MOST IMPORTANT

选择文件

选择文件的 Drawer 承担下述两种功能，而一个 Drawer 内完成两个文件的选择和预览，会给用户造成迷惑性。所以每个文件都单独选择。

- ✓ 选中需要选择的文件
- ✓ 预览并设置文件属性


文件目录
这是关于预览的一些描述
按 enter 搜索
一级子文件1
一级子文件2
二级子文件1
三级子文件2
四级子文件3
aaa.csv
bbb.csv
ccc.csv
二级子文件2
二级子文件3
三级子文件1
三级子文件2
文件预览 – aaa.csv
这是关于预览的一些描述
取消
确定

文件格式：CSV
编码格式：GBK
分隔符设置：空格
文本限定符：英文单引号

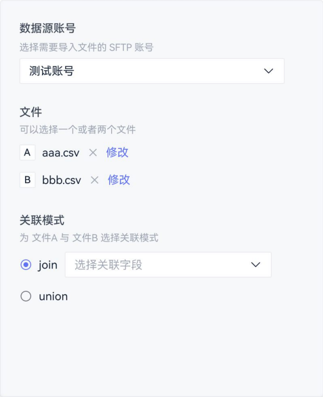
字段 1	字段 2	字段 3	字段 4	字段 5
Value 1	Value 2	Value 3	Value 4	Value 5
Value 1	Value 2	Value 3	Value 4	Value 5
Value 1	Value 2	Value 3	Value 4	Value 5
Value 1	Value 2	Value 3	Value 4	Value 5

THE MOST IMPORTANT

文件关联模式

关联模式是选择两个文件之后才有的设置，以及关联模式内的关联字段也是模式为 join 的时候才展示。

这种显示 / 隐藏的方法允许用户专注于相关信息，同时保持工作流程简短。在 datahub 项目被大量使用。
相关链接： 🔗 Progressive disclosure



数据源账号
选择需要导入文件的 SFTP 账号
测试账号

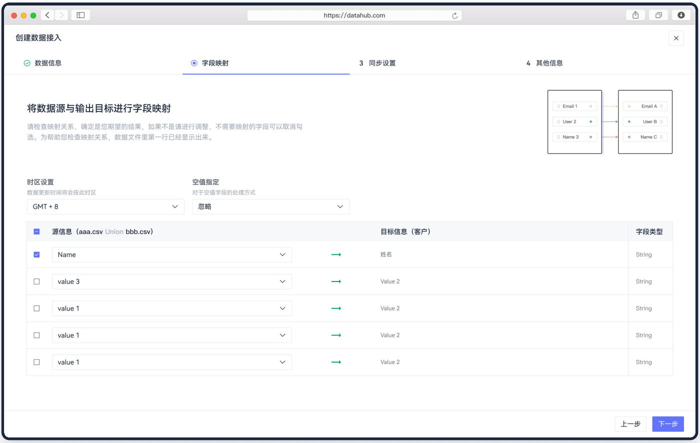
文件
可以选择一个或者两个文件
A aaa.csv × 修改
B bbb.csv × 修改

关联模式
为 文件A 与 文件B 选择关联模式
● join 选择关联字段
○ union

2

字段映射

根据尼尔森可用性原则， Recognition rather than recall。这一步将上一步（数据信息）中用户所填写的关键信息展示在映射表里。
同时为了帮助用户理解映射，增加了一个小插画。


https://datahub.com
数据信息
字段映射
3 同步设置
4 其他信息
Email 1
User 2
Name 3
Email A
User B
Name C
上一步
下一步

创建数据接入

将数据源与输出目标进行字段映射

请检查映射关系，确定是您期望的结果，如果不是请进行调整，不需要映射的字段可以取消勾选。为帮助您检查映射关系，数据文件里第一行已经显示出来。

时区设置
数据更新时间将会按此时区
GMT + 8

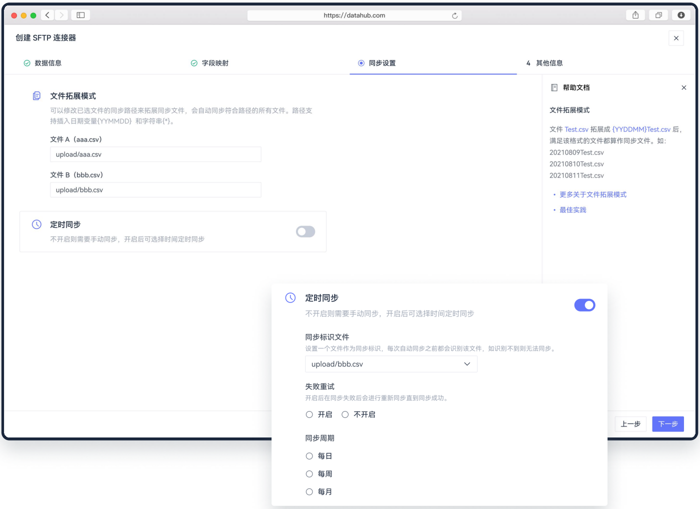
空值指定
对于空值字段的处理方式
忽略

☐	源信息（aaa.csv Union bbb.csv）		目标信息（客户）	字段类型
☑	Name	→	姓名	String
☐	value 3	→	Value 2	String
☐	value 1	→	Value 2	String
☐	value 1	→	Value 2	String
☐	value 1	→	Value 2	String

3

同步设置

较难理解的信息，会有文档帮助用户理解，降低使用门槛


https://datahub.com
数据信息
字段映射
同步设置
4 其他信息
上一步
下一步

创建 SFTP 连接器

文件拓展模式

可以修改已选文件的同步路径来拓展同步文件，会自动同步符合路径的所有文件。路径支持插入日期变量{YYMMDD} 和字符串{*}。

文件 A（aaa.csv）
upload/aaa.csv

文件 B（bbb.csv）
upload/bbb.csv

定时同步
不开启则需要手动同步，开启后可选择时间定时同步

帮助文档

文件拓展模式

文件 Test.csv 拓展成 {YYDDMM}Test.csv 后，满足该格式的文件都算作同步文件。如：
20210809Test.csv
20210810Test.csv
20210811Test.csv

- 更多关于文件拓展模式
- 最佳实践

定时同步
不开启则需要手动同步，开启后可选择时间定时同步

同步标识文件
设置一个文件作为同步标识，每次自动同步之前都会识别该文件，如识别不到则无法同步。
upload/bbb.csv

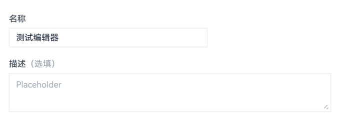
失败重试
开启后在同步失败后会进行重新同步直到同步成功。
○ 开启　○ 不开启

同步周期
○ 每日
○ 每周
○ 每月

4

基本信息

基本信息放在最后填写是因为：连接器的名称不允许重复，同时创建流程很长，如果放在前面的步骤，填写名称的时间和提交表单的时间会间隔较久，导致填写的时候名称没有重复，提交的时候名称可能重复。



名称

测试编辑器

描述（选填）

Placeholder

Convertlab

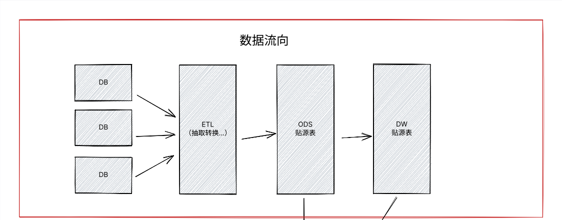
Data Storage : ODS & DW .

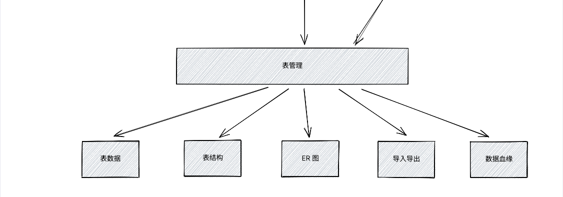
数 — 据 — 存 — 储

进入 Datahub 的数据会存在两种表内：

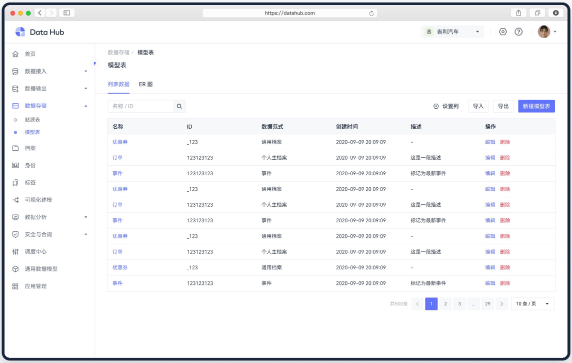
ODS 表（贴源表）：直接从客户的数据系统抽取而来；

DW 表（模型表）：数据仓库，所存数据为经过数据清洗后的规范数据。


数据流向
DB
DB
DB
ETL
（抽取转换...）
ODS
贴源表
DW
贴源表


表管理
表数据
表结构
ER 图
导入导出
数据血缘

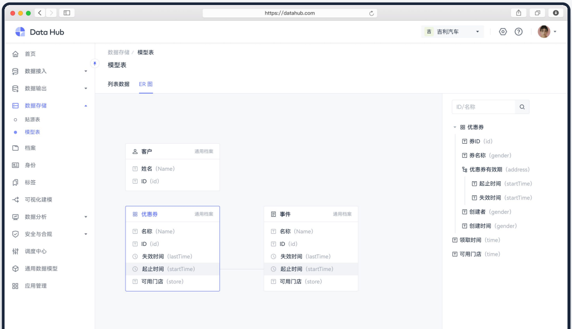
模型表列表


https://datahub.com
Data Hub
吉 吉利汽车
首页
数据接入
数据输出
数据存储
贴源表
模型表
档案
身份
标签
可视化建模
数据分析
安全与合规
调度中心
通用数据模型
应用管理
数据存储 / 模型表
模型表
列表数据
ER 图
名称 / ID
设置列
导入
导出
新建模型表

名称	ID	数据范式	创建时间	描述	操作
优惠券	_123	通用档案	2020-09-09 20:09:09	-	编辑 删除
订单	123123123	个人主档案	2020-09-09 20:09:09	这是一段描述	编辑 删除
事件	123123123	事件	2020-09-09 20:09:09	标记为最新事件	编辑 删除
优惠券	_123	通用档案	2020-09-09 20:09:09	-	编辑 删除
订单	123123123	个人主档案	2020-09-09 20:09:09	这是一段描述	编辑 删除
事件	123123123	事件	2020-09-09 20:09:09	标记为最新事件	编辑 删除
优惠券	_123	通用档案	2020-09-09 20:09:09	-	编辑 删除
订单	123123123	个人主档案	2020-09-09 20:09:09	这是一段描述	编辑 删除
优惠券	_123	通用档案	2020-09-09 20:09:09	-	编辑 删除
事件	123123123	事件	2020-09-09 20:09:09	标记为最新事件	编辑 删除

共500条
1
2
3
...
29
10 条 / 页

模型表 ER 图


https://datahub.com
Data Hub
吉 吉利汽车
首页
数据接入
数据输出
数据存储
贴源表
模型表
档案
身份
标签
可视化建模
数据分析
安全与合规
调度中心
通用数据模型
应用管理
数据存储 / 模型表
模型表
列表数据
ER 图
ID/名称

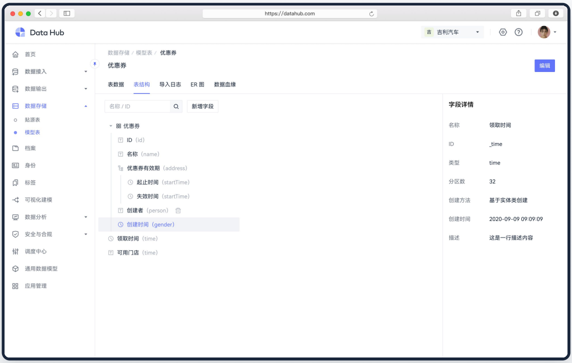
客户 通用档案
姓名 (Name)
ID (id)

优惠券 通用档案
名称 (Name)
ID (id)
失效时间 (lastTime)
起止时间 (startTime)
可用门店 (store)

事件 通用档案
名称 (Name)
ID (id)
失效时间 (lastTime)
起止时间 (startTime)
可用门店 (store)

- 优惠券
  - 券ID (id)
  - 券名称 (gender)
  - 优惠券有效期 (address)
    - 起止时间 (startTime)
    - 失效时间 (startTime)
  - 创建者 (gender)
  - 创建时间 (gender)
- 领取时间 (time)
- 可用门店 (time)

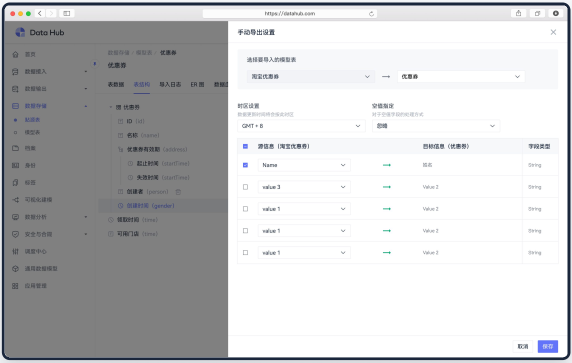
模型表详情 - 表结构


https://datahub.com
Data Hub
吉利汽车
首页
数据接入
数据输出
数据存储
贴源表
模型表
档案
身份
标签
可视化建模
数据分析
安全与合规
调度中心
通用数据模型
应用管理
数据存储 / 模型表 / 优惠券
优惠券
编辑
表数据
表结构
导入日志
ER 图
数据血缘
名称 / ID
新增字段
优惠券
ID (id)
名称 (name)
优惠券有效期 (address)
起止时间 (startTime)
失效时间 (startTime)
创建者 (person)
创建时间 (gender)
领取时间 (time)
可用门店 (time)
字段详情

名称	领取时间
ID	_time
类型	time
分区数	32
创建方法	基于实体类创建
创建时间	2020-09-09 09:09:09
描述	这是一行描述内容

表结构决定数据存储的内容。从模型表结构而言，表结构在表数据后面，但是查看和更改表结构是使用频率最高的操作，所以从列表页进入详情页，自动选中第二个 tab。

贴源表导入模型表及其记录


https://datahub.com
Data Hub
首页
数据接入
数据输出
数据存储
贴源表
模型表
档案
身份
标签
可视化建模
数据分析
安全与合规
调度中心
通用数据模型
应用管理
数据存储 / 模型表 / 优惠券
优惠券
表数据
表结构
导入日志
ER 图
优惠券
ID (id)
名称 (name)
优惠券有效期 (address)
起止时间 (startTime)
失效时间 (startTime)
创建者 (person)
创建时间 (gender)
领取时间 (time)
可用门店 (time)
取消
保存

手动导出设置

选择要导入的模型表

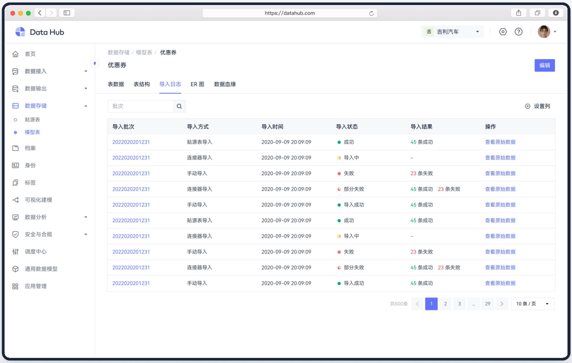
淘宝优惠券 → 优惠券

时区设置
数据更新时间将会按此时区
GMT + 8

空值指定
对于空值字段的处理方式
忽略

▬	源信息（淘宝优惠券）		目标信息（优惠券）	字段类型
☑	Name	→	姓名	String
☐	value 3	→	Value 2	String
☐	value 1	→	Value 2	String
☐	value 1	→	Value 2	String
☐	value 1	→	Value 2	String

模型表的数据来源多样化，贴源表导出就是其中之一。出于一致性考虑，导出设置在连接器的字段映射阶段稍作调整。


https://datahub.com
Data Hub
吉 吉利汽车
首页
数据接入
数据输出
数据存储
贴源表
模型表
档案
身份
标签
可视化建模
数据分析
安全与合规
调度中心
通用数据模型
应用管理
数据存储 / 模型表 / 优惠券
编辑
表数据
表结构
导入日志
ER 图
数据血缘
批次
设置列

优惠券

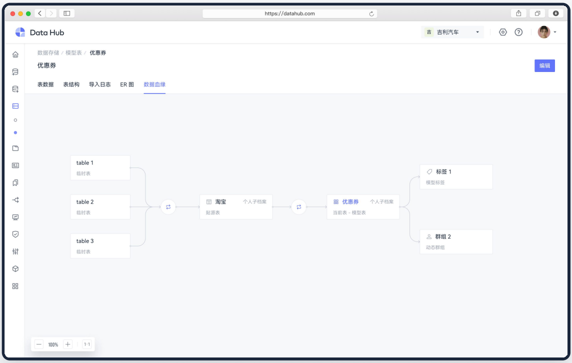
导入批次	导入方式	导入时间	导入状态	导入结果	操作
2022020201231	贴源表导入	2020-09-09 20:09:09	● 成功	45 条成功	查看原始数据
2022020201231	连接器导入	2020-09-09 20:09:09	◔ 导入中	-	查看原始数据
2022020201231	手动导入	2020-09-09 20:09:09	● 失败	23 条失败	查看原始数据
2022020201231	连接器导入	2020-09-09 20:09:09	◔ 部分失败	45 条成功 23 条失败	查看原始数据
2022020201231	手动导入	2020-09-09 20:09:09	● 导入成功	45 条成功	查看原始数据
2022020201231	贴源表导入	2020-09-09 20:09:09	● 成功	45 条成功	查看原始数据
2022020201231	连接器导入	2020-09-09 20:09:09	◔ 导入中	-	查看原始数据
2022020201231	手动导入	2020-09-09 20:09:09	● 失败	23 条失败	查看原始数据
2022020201231	连接器导入	2020-09-09 20:09:09	◔ 部分失败	45 条成功 23 条失败	查看原始数据
2022020201231	手动导入	2020-09-09 20:09:09	● 导入成功	45 条成功	查看原始数据

共500条 < 1 2 3 ... 29 > 10 条 / 页

导入操作会在贴源表生成导出记录，在模型表生成导入记录，此处展示导入记录。

DW Data lineage

数据血缘


https://datahub.com
Data Hub
吉 吉利汽车
数据存储 / 模型表 / 优惠券
优惠券
编辑
表数据
表结构
导入日志
ER 图
数据血缘
table 1
临时表
table 2
临时表
table 3
临时表
淘宝
个人子档案
贴源表
优惠券
个人子档案
当前表 - 模型表
标签 1
模型标签
群组 2
动态群组
100%
1:1

通过数据血缘分析实现数据融合处理的可追溯。数据血缘是指数据产生的链路，直白点说，
就是我们这个数据是怎么来的，经过了哪些过程和阶段。当数据出现问题，能更快溯源。

Data Analysis

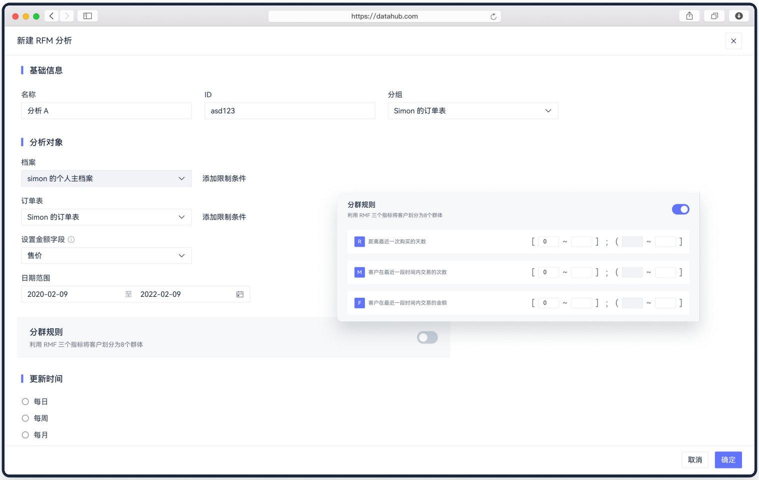
数 — 据 — 分 — 析

THE MOST IMPORTANT

RFM 分析

需要根据用户的属性进行归类，也就是通过分析数据，对用户进行归类，以便于在推送及转化过程中获得更大的收益。

- ✓ 确定分析对象
- ✓ 设置分群规则
- ✓ 基础信息
- ✓ 更新时间



https://datahub.com

新建 RFM 分析

基础信息

名称: 分析 A

ID: asd123

分组: Simon 的订单表

分析对象

档案: simon 的个人主档案 添加限制条件

订单表: Simon 的订单表 添加限制条件

设置金额字段: 售价

日期范围: 2020-02-09 至 2022-02-09

分群规则

利用 RMF 三个指标将客户划分为8个群体

R	距离最近一次购买的天数	[0 ~] ; (~]
M	客户在最近一段时间内交易的次数	[0 ~] ; (~]
F	客户在最近一段时间内交易的金额	[0 ~] ; (~]

更新时间

○ 每日

○ 每周

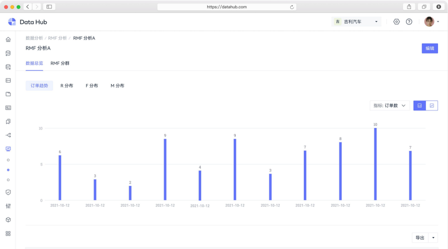
○ 每月

取消
确定

THE MOST IMPORTANT

详情页 – 数据总览

创建成功后自动跳转到详情页。这样的交互在 Datahub 也比较普遍，这样做的目的是用户创建完是需要分析结果来确定创建是否符合预期。



https://datahub.com
Data Hub
吉 吉利汽车
数据分析 / RMF 分析 / RMF 分析A
RMF 分析A
编辑
数据总览
RMF 分群
订单趋势
R 分布
F 分布
M 分布
指标: 订单数
10
5
0
6
3
2
9
4
9
3
7
8
10
7
2021-10-12
2021-10-12
2021-10-12
2021-10-12
2021-10-12
2021-10-12
2021-10-12
2021-10-12
2021-10-12
2021-10-12
2021-10-12
导出

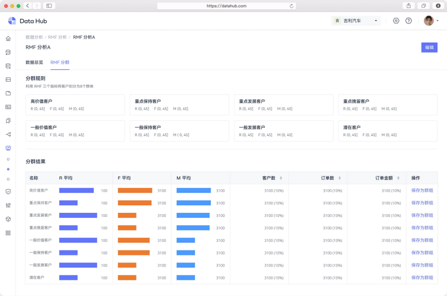
日期	订单数	对象数	订单金额(元)
2022-03-06	123123123	12323	123123123123
2022-03-06	123123123	12323	123123123123
2022-03-06	123123123	12323	123123123123

THE MOST IMPORTANT

详情页 - RMF 分群

如果设置了 RMF 分群，会自动将分析的客户划分为 8 部分，识别优质客户。可以指定个性化的沟通和营销服务，为更多的营销决策提供有力支持。

- ✓ 获取分群结果
- ✓ 将结果保存为群组
- ✓ 将分群逻辑保存为标签（后续拓展功能）


https://datahub.com
Data Hub
吉 吉利汽车
编辑

数据分析 / RMF 分析 / RMF 分析A

RMF 分析A

数据总览　RMF 分群

分群规则

利用 RMF 三个指标将客户划分为8个群体

高价值客户 R (0, 45]　F (0, 45]　M (0, 45]	重点保持客户 R (0, 45]　F (0, 45]　M (0, 45]	重点发展客户 R (0, 45]　F (0, 45]　M (0, 45]	重点挽留客户 R (0, 45]　F (0, 45]　M (0, 45]
一般价值客户 R (0, 45]　F (0, 45]　M (0, 45]	一般保持客户 R (0, 45]　F (0, 45]　M (0, 45]	一般发展客户 R (0, 45]　F (0, 45]　M (0, 45]	潜在客户 R (0, 45]　F (0, 45]　M (0, 45]

分群结果

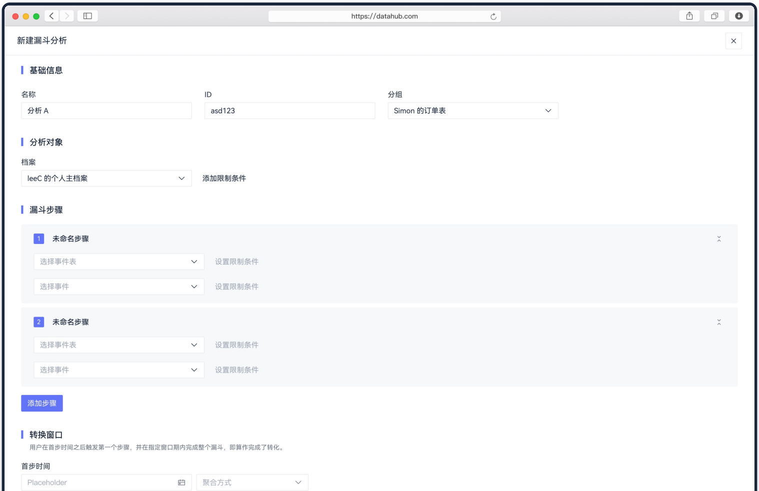
名称	R 平均	F 平均	M 平均	客户数	订单数	订单金额	操作
高价值客户	100	3100	3100	3100 (10%)	3100 (10%)	3100 (10%)	保存为群组
重点保持客户	100	3100	3100	3100 (10%)	3100 (10%)	3100 (10%)	保存为群组
重点发展客户	100	3100	3100	3100 (10%)	3100 (10%)	3100 (10%)	保存为群组
重点挽留客户	100	3100	3100	3100 (10%)	3100 (10%)	3100 (10%)	保存为群组
一般价值客户	100	3100	3100	3100 (10%)	3100 (10%)	3100 (10%)	保存为群组
一般保持客户	100	3100	3100	3100 (10%)	3100 (10%)	3100 (10%)	保存为群组
一般发展客户	100	3100	3100	3100 (10%)	3100 (10%)	3100 (10%)	保存为群组
潜在客户	100	3100	3100	3100 (10%)	3100 (10%)	3100 (10%)	保存为群组

THE MOST IMPORTANT

漏斗分析

客户关注他们的用户在重要环节上的转化率，可以通过漏斗分析来实现。

- ✓ 确定分析对象
- ✓ 设置漏斗步骤
- ✓ 基础信息
- ✓ 转换窗口


https://datahub.com
添加步骤

新建漏斗分析

基础信息

名称：分析 A

ID：asd123

分组：Simon 的订单表

分析对象

档案：leeC 的个人主档案　添加限制条件

漏斗步骤

1 未命名步骤

选择事件表　设置限制条件

选择事件　设置限制条件

2 未命名步骤

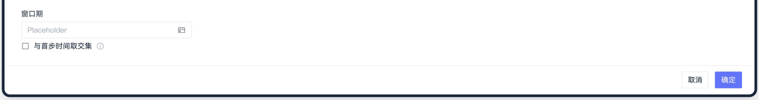
选择事件表　设置限制条件

选择事件　设置限制条件

转换窗口

用户在首步时间之后触发第一个步骤，并在指定窗口期内完成整个漏斗，即算作完成了转化。

首步时间：Placeholder　聚合方式



窗口期

Placeholder

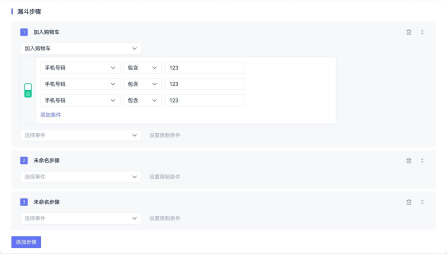
☐ 与首步时间取交集

THE MOST IMPORTANT

设置漏斗步骤

因为可以对事件和事件所在的表进行筛选，所以把选择事件拆成两步并独立设置限制条件。
步骤设置需要很多空间，但页面空间寸土寸金，所以设置区域支持展开收起。

一个 fantasy 的点：用户是否需要在一个步骤里添加多个或关系的事件。因为资源紧张且竞品都没有所以搁置中...



漏斗步骤

1 加入购物车

加入购物车

且

手机号码	包含	123
手机号码	包含	123
手机号码	包含	123

添加条件
选择事件
设置限制条件

2 未命名步骤

选择事件
设置限制条件

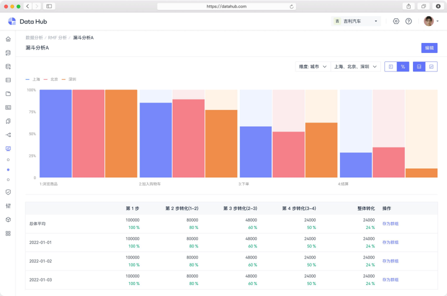
3 未命名步骤

选择事件
设置限制条件
添加步骤

THE MOST IMPORTANT

漏斗分析图

漏斗图展示漏斗分析原本是最好的方案，但是 Datahub 的漏斗支持多维度查看，所以这种情况用柱状图展示漏斗有更好的效果。



https://datahub.com

Data Hub

吉利汽车

数据分析 / RMF 分析 / 漏斗分析A

漏斗分析A

维度：城市　上海、北京、深圳

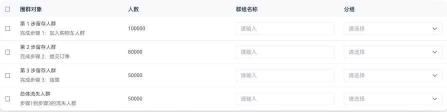
编辑
%
上海
北京
深圳
100%
75%
50%
25%
0
1:浏览商品
2:加入购物车
3:下单
4:结算

	第 1 步	第 2 步转化(1–2)	第 3 步转化(2–3)	第 4 步转化(3–4)	整体转化	操作
总体平均	100000 100 %	80000 80 %	48000 60 %	24000 50 %	24000 24 %	存为群组
2022-01-01	100000 100 %	80000 80 %	48000 60 %	24000 50 %	24000 24 %	存为群组
2022-01-02	100000 100 %	80000 80 %	48000 60 %	24000 50 %	24000 24 %	存为群组
2022-01-03	100000 100 %	80000 80 %	48000 60 %	24000 50 %	24000 24 %	存为群组

THE MOST IMPORTANT

将漏斗保存为群组

同样是保存为群组，漏斗的圈群对象要多一点，所以针对保存做了一个集成，以 Table 的形式让用户快速保存所需要的群组。



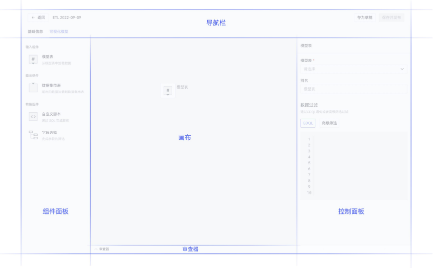
□	圈群对象	人数	群组名称	分组
□	第 1 步留存人群 完成步骤 1：加入购物车人群	100000	请输入	请选择
□	第 2 步留存人群 完成步骤 2：提交订单	80000	请输入	请选择
□	第 3 步留存人群 完成步骤 3：结算	50000	请输入	请选择
□	总体流失人群 步骤1到步骤3的流失人群	50000	请输入	请选择

Visual Modeling

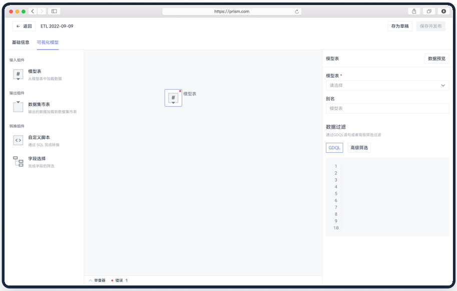
可 — 视 — 化 — 建 — 模

利用可视化建模来完成数据的清洗、转换、存储。

- ✓ 将页面操作区域划分：组件、导航、画布、控制面板、审查器
- ✓ 用 tab 将模型和基础信息分隔，保证专注



← 返回
ETL 2022-09-09
导航栏
存为草稿
保存并发布
基础信息
可视化模型
输入组件
模型表
从模型表中加载数据
输出组件
数据集市表
输出的数据加载到数据集市表
转换组件
自定义脚本
通过 SQL 完成转换
字段选择
完成字段的筛选
组件面板
模型表
画布
模型表
模型表 *
请选择
别名
模型表
数据过滤
通过GDQL语句或者高级筛选过滤
GDQL
高级筛选
控制面板
审查器
审查器



https://prism.com
← 返回
ETL 2022-09-09
存为草稿
保存并发布
基础信息
可视化模型
输入组件
模型表
从模型表中加载数据
输出组件
数据集市表
输出的数据加载到数据集市表
转换组件
自定义脚本
通过 SQL 完成转换
字段选择
完成字段的筛选
模型表
模型表
数据预览
模型表 *
请选择
别名
模型表
数据过滤
通过GDQL语句或者高级筛选过滤
GDQL
高级筛选
审查器
● 错误 1

THE MOST IMPORTANT

输入组件 - 模型表

输入组件承载数据输入，当前版本只做了模型表，后续会持续增加

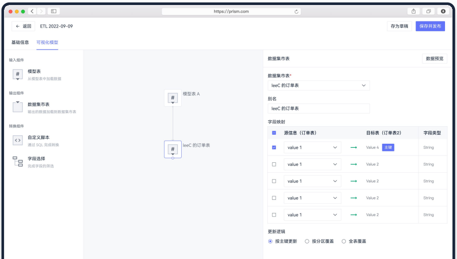
- ✓ 可以做数据过滤
- ✓ 给组件起别名（别名默认为组件名，若选择模型表则默认为表名）
- ✓ 数据预览：预览从当前组件输出的数据

THE MOST IMPORTANT

输出组件 - 数据集市表

输入组件承载数据输入，当前版本只做了数据集市表，后续会持续增加

- ✓ 从输入组件接受数据
- ✓ 设置字段映射



https://prism.com
← 返回
ETL 2022-09-09
存为草稿
保存并发布
基础信息
可视化模型
输入组件
模型表
从模型表中加载数据
输出组件
数据集市表
输出的数据加载到数据集市表
转换组件
自定义脚本
通过 SQL 完成转换
字段选择
完成字段的筛选
模型表 A
leeC 的订单表
数据集市表
数据预览

数据集市表*

leeC 的订单表

别名

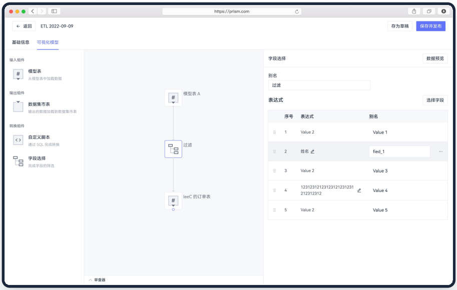
leeC 的订单表

字段映射

☐	源信息（订单表）		目标表（订单表2）	字段类型
☑	value 1	→	Value 4 主键	String
☐	value 1	→	Value 2	String
☐	value 1	→	Value 2	String
☐	value 1	→	Value 2	String
☐	value 1	→	Value 2	String

更新逻辑

● 按主键更新 ○ 按分区覆盖 ○ 全表覆盖


https://prism.com
返回
ETL 2022-09-09
存为草稿
保存并发布
基础信息
可视化模型
输入组件
模型表
从模型表中加载数据
输出组件
数据集市表
输出的数据加载到数据集市表
转换组件
自定义脚本
通过 SQL 完成转换
字段选择
完成字段的筛选
模型表 A
过滤
leeC 的订单表
字段选择
数据预览
选择字段
审查器

别名：过滤

表达式

序号	表达式	别名
1	Value 2	Value 1
2	姓名	fied_1
3	Value 2	Value 3
4	123123121231231212312312 12312312	Value 4
5	Value 2	Value 5

THE MOST IMPORTANT

转换组件 - 字段选择

转换组件为输入和输出组件之间的组件，用于加工从输入组件处理的数据。已字段选择为例，为了节约篇幅，自定义脚本组件不展示。

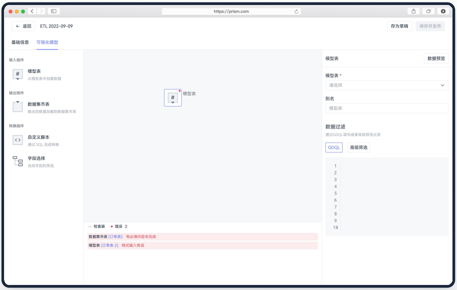
- ✓ 过滤字段
- ✓ 给字段设置表达式，即通过函数加工字段

THE MOST IMPORTANT

审查器

可以随时检查当前页面的一些问题，并能通过点击去定位问题所在的组件并进行修改

- ✓ 实时校验
- ✓ 快速定位修改


https://prism.com
返回
ETL 2022-09-09
存为草稿
保存并发布
基础信息
可视化模型
输入组件
模型表
从模型表中加载数据
输出组件
数据集市表
输出的数据加载到数据集市表
转换组件
自定义脚本
通过 SQL 完成转换
字段选择
完成字段的筛选
模型表
模型表
数据预览
模型表 *
请选择
别名
模型表
数据过滤
通过GDQL语句或者高级筛选过滤
GDQL
高级筛选
检查器
错误 2

数据集市表 [订单表]：有必填内容未完成

模型表 [订单表 2]：格式输入有误

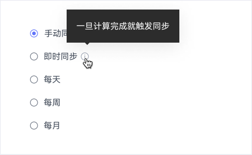
Something Else

THE MOST IMPORTANT

引导 & 反馈

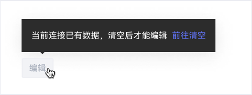
面对一个庞大复杂的系统，引导和反馈都是降低使用门槛的重要手段。

- ✓ 完善的帮助文档
- ✓ 见缝插针的描述
- ✓ 使用 disable 状态而不是直接隐藏
- ✓ 给 disable 加上 tooltip 告知原因或引导解决方式
- ✓ 善用 Inline - message、Toast - message 和 Notification
- ✓ 合理使用结果页

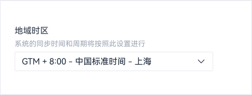


一旦计算完成就触发同步

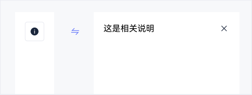
● 手动同
○ 即时同步
○ 每天
○ 每周
○ 每月



当前连接已有数据，清空后才能编辑 前往清空



地域时区
系统的同步时间和周期将按照此设置进行
GTM + 8:00 - 中国标准时间 - 上海



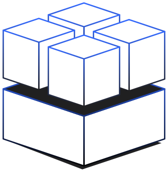
这是相关说明

THE MOST IMPORTANT

规范和设计系统

Datahub 的规范由最开始的 Ant Design， 通过样式覆盖变成一个属于自己的组件库，最后演变成 Prism Design System。**因为复杂且篇幅较长所以放在后续单独介绍。**

B 端注重效率，避免重复造轮子，应尽量将所有页面的组成元素规范化。和前端密切交流产出最适合当前业务的规范。

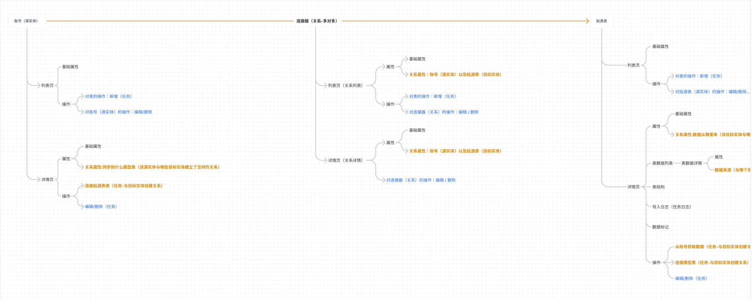


THE MOST IMPORTANT

具象的业务和抽象的信息架构

B端设计师本身大多数情况下并不是产品的目标用户，要做出合理的设计，第一时间应该投身业务，尽可能了解需求。但也要实时回头，将业务抽象化才能将产品完善和系统化。

如：把具体的业务抽象成 实体、属性、关系、任务...


账号（源实体）
连接器（关系-多对多）
贴源表
列表页
基础属性
操作
对表的操作：新增（任务）
对账号（源实体）的操作：编辑/删除
详情页
属性
基础属性
关系属性:同步到什么模型表（该源实体与哪些目标实体建立了怎样的关系）
操作
连接贴源表表（任务-与目标实体创建关系）
编辑/删除（任务）
列表页（关系列表）
属性
基础属性
关系属性：账号（源实体）以及贴源表（目标实体）
操作
对表的操作：新增（任务）
对连接器（关系）的操作：编辑 / 删除
详情页（关系详情）
属性
基础属性
关系属性：账号（源实体）以及贴源表（目标实体）
对连接器（关系）的操作：编辑 / 删除
列表页
基础属性
操作
对表的操作：新增（任务）
对贴源表（源实体）的操作：编辑/删除...
详情页
属性
基础属性
关系属性:数据从哪里来（该目标实体与哪
表数据列表
表数据详情
属性
数据来源（与哪个实
表结构
导入日志（任务日志）
数据标记
操作
从账号获取数据（任务-与目标实体创建关
连接模型表（任务-与目标实体创建关系）
编辑/删除（任务）

写在本篇最后

Convertlab Datahub

B 端项目页面众多，迫于篇幅 [安全与合规]、[配置迁移]、[admin 管理系统] 等交互无法尽述。

另外本篇作品集涉及到不少前司核心业务，仅用于个人面试展示使用，希望您不要传播，谢谢。