



# UCloud安全屋：让数据流通起来

Peter司照凯 UCloud数据平台事业部总监

## ■ 申报主体单位

优刻得科技股份有限公司

## ■ 联合共建单位

复旦大学 数字和移动治理实验室

中国联合网络通信有限公司上海市分公司

上海大数据应用创新中心



提供安全、合规的数据流通和使用的在线服务  
实现数据所有权和使用权的分离

数据源方

数据需求方

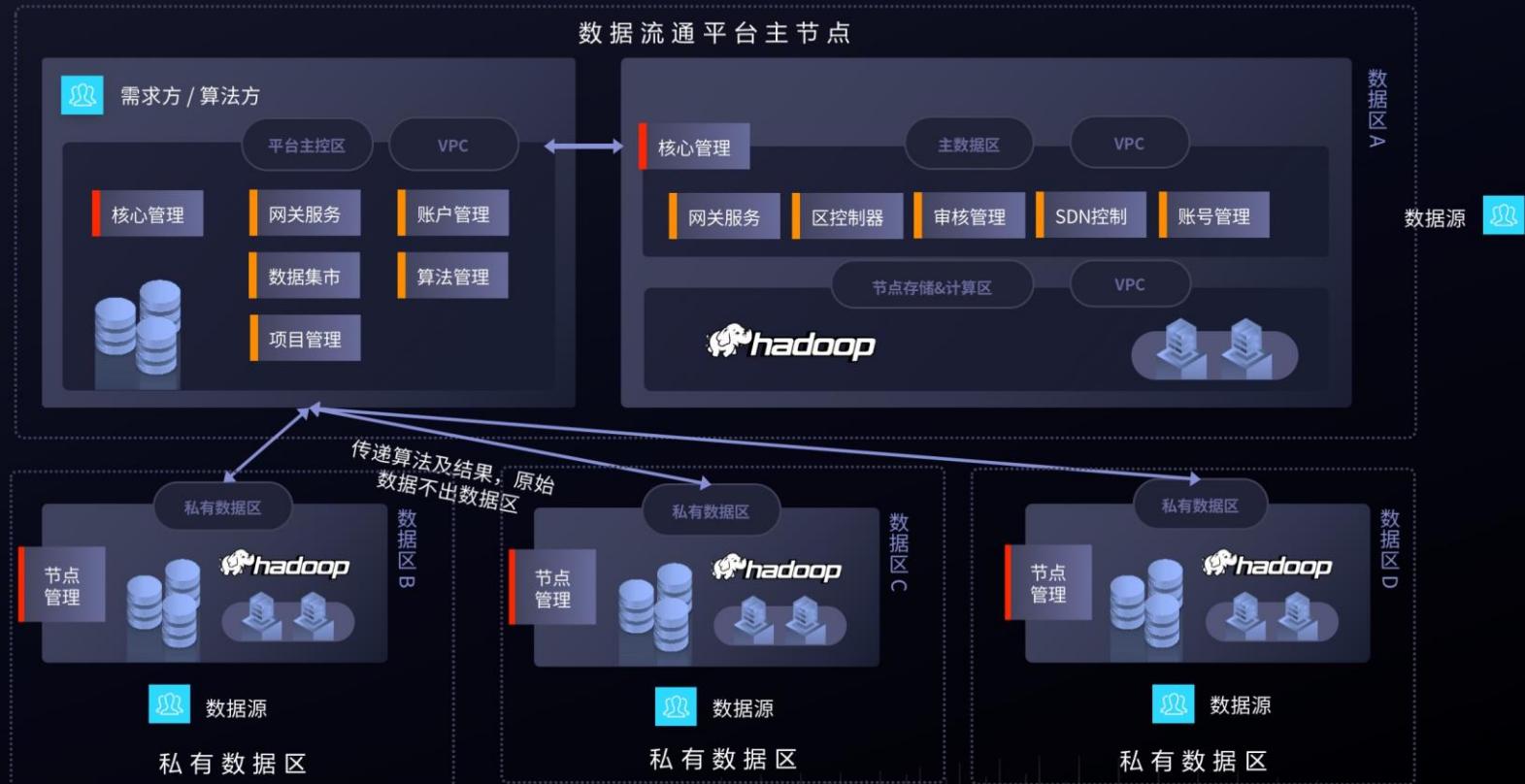


安全屋  
LET THE DATA FLOW

合作伙伴方

第三方监管方

# 如何实现数据所有权和使用权的分离



# 安全屋平台化产品的价值

## 需求痛点

- 对接繁琐
- 数据成本高
- 数据质量差
- 人力、精力投入大
- 数据无法融合
- 合法合规使用



## 安全屋使命愿景

让数据流通起来  
LET THE DATA FLOW

## 安全屋的定位

安全合规  
中立的流通平台



## 单数据源对外开放模式

- 数据积累越来越多
- 对数据可复用场景缺乏探索
- 需求方诉求多而无法消化

## 组织内的多部门间数据共享模式

- 各部门均业务导向，起初设计独立
- 部门间数据并非互补性特征，不愿意单方面提供
- 各部门重视自身权责，数据开放风险不愿承担

## 不同组织间共享数据的模式

- 业务互补相关，缺乏联合意识
- 数据安全风险过大，不敢开放

# 各种数据流通模式的比较

模式	优点	缺点
数据私下买卖	简单	隐私泄露 违法
API模式	数据安全性较高 方便快捷 使用门槛低	无法实现数据的深层次应用 (例如自定义算法、深度建模)
原始数据 / 线下模式	可以实现自定义算法和深度建模 通过物理手段实现数据安全	数据流通效率低 成本高昂 对政府来说可监管性差
原始数据 / 线上平台模式	可以实现自定义算法和深度建模 通过物理手段实现数据安全 效率高、成本低、政府可见惯性好	数据提供方的接受度需不断教育
区块链 / 去中心化模式	数据流通全流程采用区块链做记录，安全性高 接入方便，使用门槛低	算法执行无法在链上进行

- 安全屋采用的是“原始数据线上平台模式”，是最适合中国法规和安全体系的模式

The screenshot shows the UCloud Data Management Platform interface. At the top, there is a navigation bar with the UCloud logo, a '安全屋' (Safe House) icon, and menu items: 首页 (Home), 数据集市 (Data Marts), 数据管理 (Data Management), and 流程管理 (Process Management). On the right side of the header is a user profile with the email safehousedemoa@ucloud.

**我的项目** (My Projects): A donut chart showing 11 projects. The legend indicates 10 projects are '进行中' (In Progress) and 1 project is '已结束' (Completed). There is a '查看详情' (View Details) link.

**我的算法** (My Algorithms): Shows 77 items. There is a '查看详情' (View Details) link.

**我的数据** (My Data): Shows data management statistics. It includes '自有数据' (Own Data) with 10 items and '授权数据' (Granted Data) with 43 items. There is a '查看详情' (View Details) link.

**快速入门** (Quick Start): A flow diagram illustrating the data processing workflow:

- 使用数据 (Use Data) → 数据集市选择数据 (Data Mart Select Data)
- 数据集市选择数据 → 数据源授权 (Data Source Authorization)
- 数据源授权 → 编写算法 (Write Algorithm)

**使用帮助** (Usage Help): Includes links to '使用手册' (User Manual) and '帮助文档' (Help Documentation).

The screenshot shows the UCloud Data Marketplace interface. At the top, there is a navigation bar with the UCloud logo, a security icon labeled '安全屋', and menu items: 首页 (Home), 数据集市 (Data Marketplace), 数据管理 (Data Management), and 流程管理 (Process Management). On the right side of the header are icons for home, user profile, and settings.

## 数据集市

**sfds25740(SDK)**

DBA基础标签表 <a href="#">详情 &gt;&gt;</a> 覆盖范围：全国   数据量级：记录数量十... 更新周期：日更新   提供方式：原始数据 可使用ID：脱敏设备号 DBA基础标签表,包含性别,年龄段等设备使用者基础属性	DBAApp分类表 <a href="#">详情 &gt;&gt;</a> 覆盖范围：全国   数据量级：记录数量万级 更新周期：日更新   提供方式：标签数据 可使用ID：APP包名 DBA APP分类表,2000万+APP的一级二级分...类	DBA日志表 <a href="#">详情 &gt;&gt;</a> 覆盖范围：全国   数据量级：记录数量十... 更新周期：日更新   提供方式：原始数据 可使用ID：脱敏设备号 DBA日志表,单个设备号与单个APP的关联,使...用时长,使用日期	DBA设备位置信... 覆盖范围：全国 更新周期：日更 可使用ID：脱敏 DBA设备位置信...息,百度地图经纬...
DBA设备号映射表 <a href="#">详情 &gt;&gt;</a> 覆盖范围：全国   数据量级：记录数量十... 更新周期：日更新   提供方式：原始数据	DBA月活估计表 <a href="#">详情 &gt;&gt;</a> 覆盖范围：全国   数据量级：记录数量十... 更新周期：日更新   提供方式：原始数据	高级转码表 <a href="#">详情 &gt;&gt;</a> 覆盖范围：   数据量级： 更新周期：日更新   提供方式：原始数据	旅游数据-出境游 覆盖范围：全国 更新周期：日更

## sfds97880(运营商)

sfds97880用户基本属性表 [详情 >>](#)

覆盖范围：全国 | 数据量级：记录数量十...

更新周期：日更新 | 提供方式：原始数据

可使用ID：脱敏设备号

包含全国电信及江苏联通数据，通过运营商数据解析，对用户每日访问内容进行标签化及统计

sfds97880分级标签 [详情 >>](#)

覆盖范围：全国 | 数据量级：记录数量十...

更新周期：日更新 | 提供方式：标签数据

可使用ID：标签ID号

通讯访问内容对应整理的DMP库，访问内容的标签详情

sfds97880用户月表 [详情 >>](#)

覆盖范围：全国 | 数据量级：记录数量十...

更新周期：月更新 | 提供方式：原始数据

可使用ID：脱敏设备号

包含全国电信及江苏联通数据，通过运营商数据解析，按月统计的标签化用户访问内容

## sfds50284(垂直-POI)

sfds50284poi数据 [详情 >>](#)

覆盖范围：全国 | 数据量级：记录数量千...

更新周期：不指定 | 提供方式：原始数据

可使用ID：商铺名称及ID号

垂直领域数据服务商自有店铺、景区等详情及poi数据

UCLOUD 安全屋

首页 数据集市 数据管理 流程管理

safehousedemo@uclo

/ 数据市场 / 数据名称

数据详情

申请试用 申请授权

基本信息

数据名称	sfds97880用户基本属性表
描述	包含全国电信及江苏联通数据，通过运营商数据解析，对用户每日访问内容进行标签化及统计
数据条数	
覆盖范围	全国
更新周期	日更新
数据量级	记录数量十亿级

字段	类型	描述	值
type	STRING	UCloud ( 唯一 )	UCloud
user_id	STRING	用户基本属性-id号	561630522bc7cc6596ba4ef1544820f6
label_id	STRING	用户所属标签id	856723
log_count	INT	用户访问次数	2
log_day	FLOAT	运营商记录日期	2018-07-25
isp	STRING	用户所属运营商编码	ct
collect_env	STRING	用户使用网络环境编码	ps
province	STRING	用户所属省份编码	811
part_day	STRING	数据传输日期	2017-10-02

The screenshot shows the UCloud Data Management interface. At the top, there are tabs for '首页' (Home), '数据集市' (Data Mart), '数据管理' (Data Management, which is selected), and '流程管理' (Process Management). On the right side, there are user profile icons and the text 'safehousedem'. Below the tabs, there are three buttons: '文件上传管理' (File Upload Management), '自有数据表' (Own Data Table), and '第三方授权数据' (Third-party Authorization Data, which is selected). A search bar with the placeholder '请输入' (Please input) is located above the table. The main area displays a table with the following columns: '数据源' (Data Source), '授权表编号' (Authorization Table ID), '授权表名称' (Authorization Table Name), '授权语句' (Authorization Statement), '数据条数' (Number of Data Rows), '状态' (Status), '授权开始时间' (Authorization Start Time), '授权结束时间' (Authorization End Time), and '授权时长' (Authorization Duration). The table contains ten rows of data, each representing a different authorization entry.

数据源	授权表编号	授权表名称	授权语句	数据条数	状态	授权开始时间	授权结束时间	授权时长
sfds25740	acr5h4wvr98u2t2t	高级	<a href="#">查看</a>	--	<span>正常</span>	2018-12-15 00:00:00	2019-03-15 11:12:49	90天11小时12分钟49秒
sfds25740	acr5h4wvus7e2t2u	旅游数据-出境游	<a href="#">查看</a>	--	<span>正常</span>	2018-12-15 00:00:00	2019-03-15 11:13:05	90天11小时13分钟5秒
sfds25740	acr5h4wwwxf2t2v	旅游数据-国内游	<a href="#">查看</a>	--	<span>正常</span>	2018-12-15 00:00:00	2019-03-15 11:13:15	90天11小时13分钟15秒
sfds25740	acr5h4wyij4s2t2w	DBA基础标签表	<a href="#">查看</a>	--	<span>正常</span>	2018-12-15 00:00:00	2019-03-15 11:13:23	90天11小时13分钟23秒
sfds25740	acr5h4l1p4cx2t2o	高级	<a href="#">查看</a>	--	<span>正常</span>	2018-12-14 00:00:00	2019-03-14 15:20:03	90天15小时20分钟3秒
sfds25740	acr5h4l1qv072t2p	旅游数据-出境游	<a href="#">查看</a>	--	<span>正常</span>	2018-12-14 00:00:00	2019-03-14 15:20:11	90天15小时20分钟11秒
sfds25740	acr5h4l1s8db2t2q	旅游数据-国内游	<a href="#">查看</a>	--	<span>正常</span>	2018-12-14 00:00:00	2019-03-14 15:20:18	90天15小时20分钟18秒
sfds25740	acr5h4m8f7rs2t2r	DBA基础标签表	<a href="#">查看</a>	--	<span>正常</span>	2018-12-14 00:00:00	2019-03-14 17:19:39	90天17小时19分钟39秒
测试区数据	acr5h3rwxlyf2t2c	旅游数据-出境游	<a href="#">查看</a>	--	<span>正常</span>	2018-12-12 14:24:18	2028-12-12 14:24:18	3653天0小时0分钟0秒
测试区数据	acr5h3rwdxlo02t2h	旅游数据-出境游	<a href="#">查看</a>	--	<span>正常</span>	2018-12-12 14:24:18	2028-12-12 14:24:18	3653天0小时0分钟0秒

UCLOUD 安全屋 首页 数据集市 数据管理 流程管理

safehousedemoa@ucloud.

/ 算法管理 / 新建算法 取消 确定

数据表

- 目有表(10)
- 授权表(31)
  - acr5gv1ta9ce2t1y(sfds97880用户基本属性表)
  - acr5gy30ym4r2t20(sfds97880用户基本属性表)
  - acr5gyppyl702t10(测试-DBAApp分类表)
  - acr5gz3qch0o2t05(测试-DBAApp分类表)
  - acr5gz3sduat2t06(DBA月活估计表)
  - acr5h0ndv65i2t0w(test)
  - acr5h0nhfhmc2t0x(DBA基础标签表)
  - acr5h0nhhb0h2t0y(DBAApp分类表)
  - acr5h0nhjfmk2t0z(DBA日志表)
  - acr5h3gsunwy2t16(测试-DBA日志表)
  - acr5h3gsvm8i2t19(测试-DBA日志表)
  - acr5h3j8czfn2t1y(旅游数据-国内游)
  - acr5h3jh5mpu2t1z(测试-DBA基础标签表)
  - acr5h3rwt56n2t24(旅游数据-出境游)
  - acr5h3rwwy4d2t25(旅游数据-出境游)

算法代码

\* 算法名称: t1

算法描述: 请填写算法描述!

算法代码: 安全屋自定义函数

```
1 select device,imei_md5 from acr5h411p4cx2t2o limit 1200
```

UCLOUD 安全屋

首页 数据集市 数据管理 流程管理

safehousedev

/ 项目管理 / 花旗

### 作业管理

**新建作业**

作业提交时间 :  请输入

作业编号	作业类型	作业描述	操作对象	提交时间	更新时间	作业状态	操作
job5h5ovt64p0a1v	下载	--	--	2018-12-17 10:14:45	2018-12-17 10:22:38	● 作业已结束	下载 删除
job5h5ovroem0a1u	下载	--	--	2018-12-17 10:14:38	2018-12-17 10:22:44	● 作业已结束	下载 删除
job5h5huetc70a1r	计算	--	--	2018-12-16 22:25:24	2018-12-16 23:19:56	● 待保存	保存为工作表
job5h5cwggu0a1p	下载	--	--	2018-12-16 14:07:14	2018-12-17 10:23:28	● 结果审核驳回	删除
job5h5cumifh0a1o	计算	--	--	2018-12-16 14:02:07	2018-12-16 14:08:05	● 作业已结束	删除
job5h4wy5q710a0t	计算	--	--	2018-12-15 11:19:35	2018-12-15 14:13:52	● 待保存	保存为工作表
job5h4lple8b0a0j	计算	--	--	2018-12-14 16:27:03	2018-12-17 10:15:10	● 作业已结束	删除

14

# 安全屋的平台优势

## 平台化

包括数据提供方、数据需求方、合作伙伴方、第三方监管方

数据的采集、清洗、计算、使用全在云端打通

各参与方各司其职，形成数据流通生态闭环

## 线上化

客户线上自主化使用，无需太多线下流程

线上化的同时保证数据安全，通过全套云安全技术作为安全支撑

## 合规化

原始数据不泄露

数据的脱敏、加密、和匿名化

用户隐私授权许可

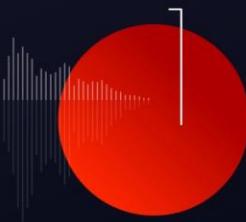
全流程符合《信息安全技术个人信息安全规范》

## 定制化

根据需求方自身业务属性定义标签

安全屋提供“加工原材料”

## 安全屋的技术优势



底层云计算技术



大数据技术



数据融合技术



安全技术

# 安全屋的安全机制



## 数据脱敏

原始数据不泄露；  
原始数据开放使用，需要脱敏、加密、匿名化处理；支持MD5，保序加密，同态加密等常用场景的加密标准机制



## 完善数据授权机制

数据使用授权，保障数据安全开放使用；  
数据结果/算法输出授权，保障据安全流通



## 数据使用审计

数据使用可追溯，确保数据使用场景可靠及稳定；  
数据源方可通过开放平台轻松查看谁在使用数据；  
模型运算脚本自动化检测，自动识别模型隐藏风险



## 多租户隔离

数据源方和数据需求方的数据都隔离存储在私有数据区；  
需求方在使用数据进行计算或者建模时完全独立，不受其他租户计算资源影响



## 模型结果标准化输出

平台化产品支持标准的行业报告，标签补充，营销投放等通用场景模型输出

私有化产品支持本地场景的模型输出能力组件集成

## 案例1：易居-用户画像



用户画像 >



营销策略优化 >



## 案例2：智子云-DMP



DMP >



< 广告曝光



## 案例3：百观-投资预测



企业数据 >



投资策略 >



金融市场

## 案例4：脉策-零售选址



人流数据 >  
用户画像



选址服务 >

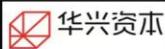




# 更多客户和合作伙伴

## 战略合作

## 典型客户



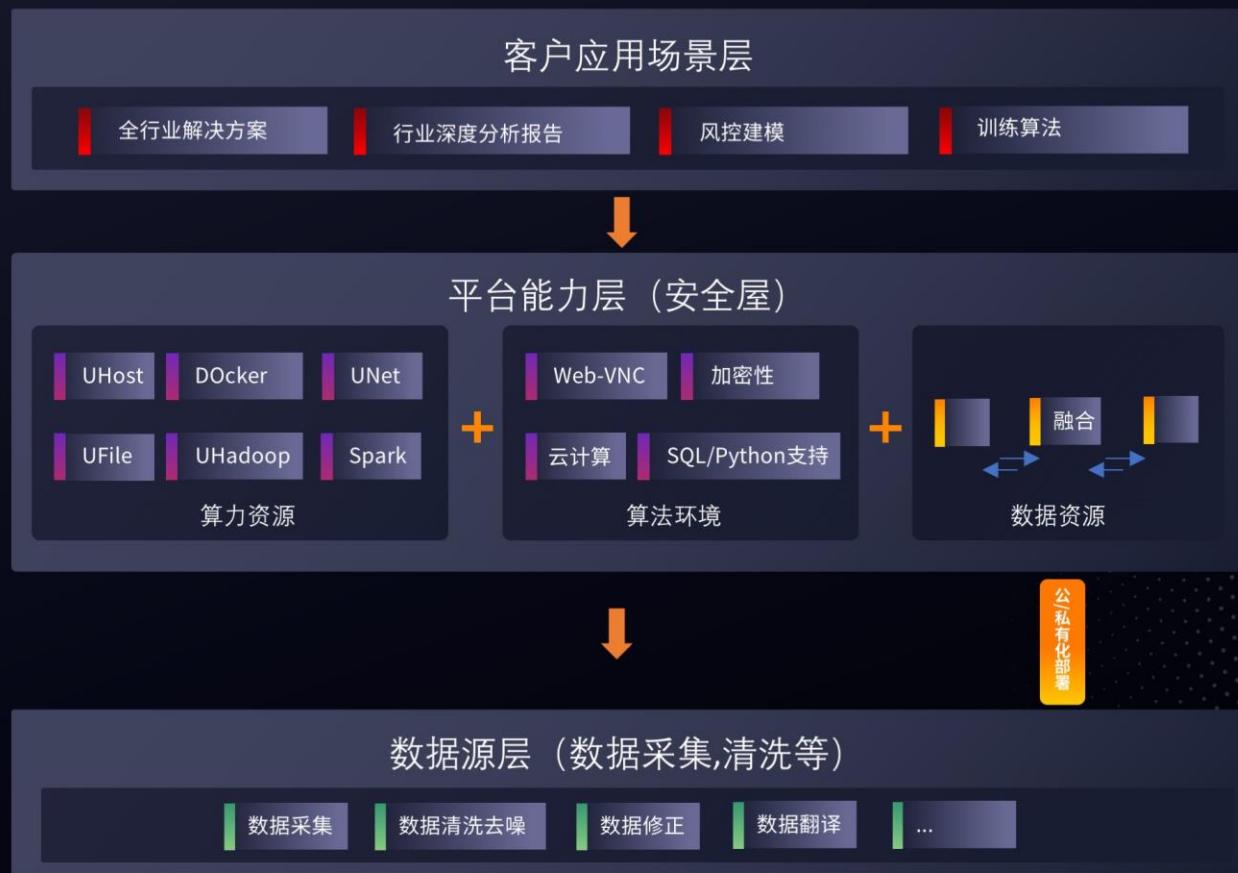
### 政府



### 企业



# 安全屋平台定位



# 让数据流通起来

LET THE DATA FLOW



[sf.ucloud.cn](http://sf.ucloud.cn)



极简 极真

T H A N K S