数据流通 -> 数据外循环 -> 技术信任 -> 隐私计算内涵 -> 开源secretflow

# 数据：要素流转和内外循环

数据生命周期：采集、加工、存储、使用、提供服务、销毁

可以划分为两个阶段：数据采集加工、数据价值释放

数据价值释放有内外循环两个模式。

内循环只在持有方的运维域使用

外循环会离开运维域，所以持有方有数据管控的需求和责任

外循环是构建数据要素市场的核心。数据提供方、使用方都可以获取收益

# 数据外循环和信任焦虑

## 信任焦虑

数据市场需要足够多的数据提供方的参与，但是这些人目前因为各种原因不敢加入。这些原因主要是对流通主体的担心，包括：

1、平台的运维人员会不会泄露

2、数据使用方会不会按照约定使用数据

3、数据使用方对数据源的合规性也有质疑

（信任焦虑）

过去数据流通的主要模式是数据打包交付，也就是数据所有权的完全交付

数据二十条提出三权分置：

|  |  |  |
| --- | --- | --- |
| 数据生产 | 数据资源持有权 | 来源合规、安全保障 |
| 数据流通 | 数据加工使用权 | 流转合规、安全 |
| 数据消费 | 数据产品经营权 | 定价交易、确权 |

分置的目的是期望使用权和持有权的分离

## 技术信任

信任焦虑的解决方案：主体信任到技术信任

主体：企业和工作人员

技术：技术性能、安全性

技术信任体系需要三个基础：

1、匿名化技术

2、数据互联

可以通过隐私计算解决

3、数据使用权跨域管控

需要区块链、可信计算技术

技术信任需要信任链，限制运维权限，同时还要提供远程认证、可信安全模块

核心：通过技术可信代替人员可信

# 数据要素流通和隐私计算

隐私计算内容，扩展到了这三个方面

1、数据可用不可见（传统方案侧重）

2、数据可算不可识（现代方案要求），也就是不能反推出用户的隐私信息

3、可控，可以计量

隐私计算产品还需要：

1、通用的安全分级、评测方式

2、降低使用门槛，通过开源得到大规模的普及应用，从而促进数据流通

# 隐私计算开源助力数据要素流通

隐语做了一些工作助力数据流通。

SF的优势：

1、统一架构（明密文混合）

2、原生应用（密态编译器支持原生SQL、AI训练和预测）

3、开放扩展（模块化设计，存储可以扩展，硬件/计算/算法层都可以扩展）

4、性能卓越（支持poc认证，也支持十亿数据psi、千万级建模）

隐私性得到过认证