

Toward a Decentralized Data & AI Marketplace

@Qibing Li, BitHacks

What People Say for Data?

CURRENT SITUATION

Precision marketing. Big companies collect user information for data-driven advertising, while users get nothing.

Heavy processing on **centrally controlled private data.**

Manipulation, surveillance, and frequent data leaks.

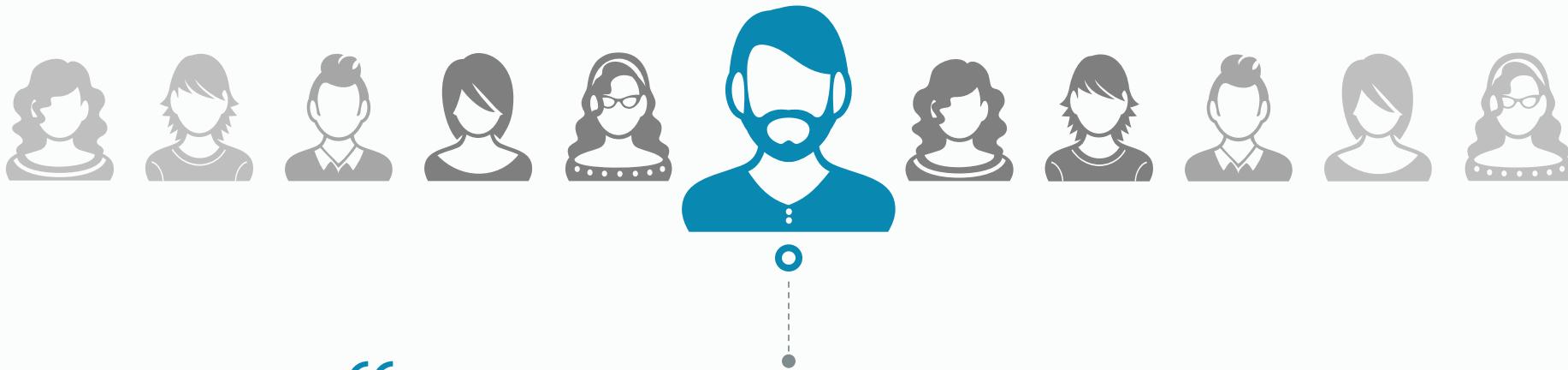
KEY POINTS

Decentralized identity and private storage are required for preserving user-controlled personal data.

Accurate data modeling is essential in the commercial world!

Users monetize their data and get incentives from data & AI marketplace.

OpenHealth



*More **private**, more **accurate**, much **cheaper***

Present Healthcare Ecosystem

What they want ...



Government

Less medical expenses,
regulation (e.g., drug traceability)



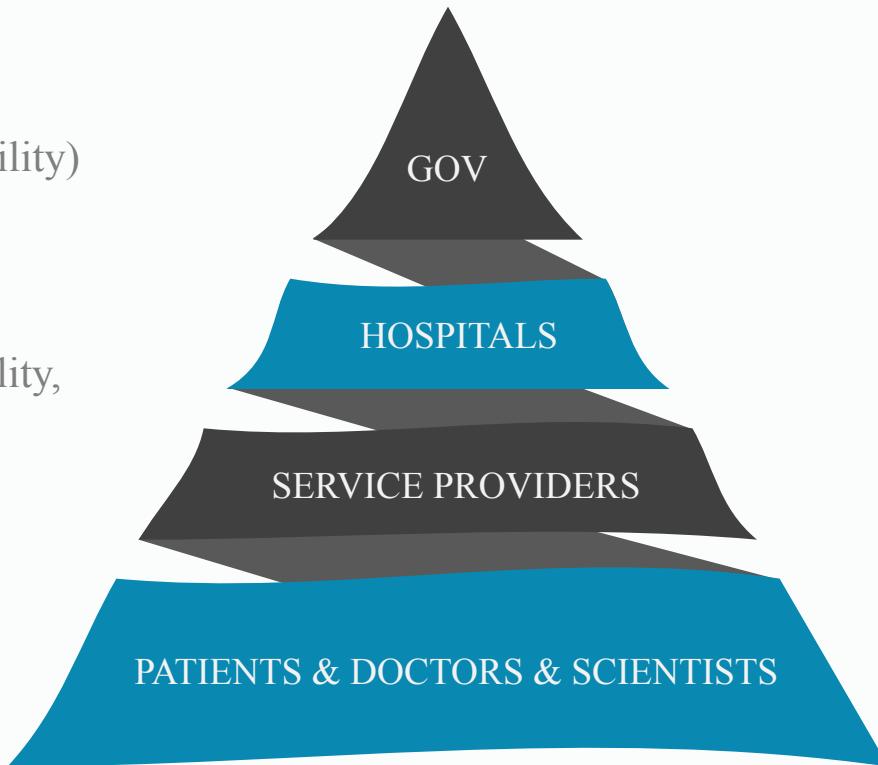
Hospitals

Unified HER & interoperability,
clinical trial recording



Service Providers

Guaranteed private storage,
more IOT data for AIs,
Interoperability



Patients

Privacy, data ownerships,
cheaper and more accurate



Doctors

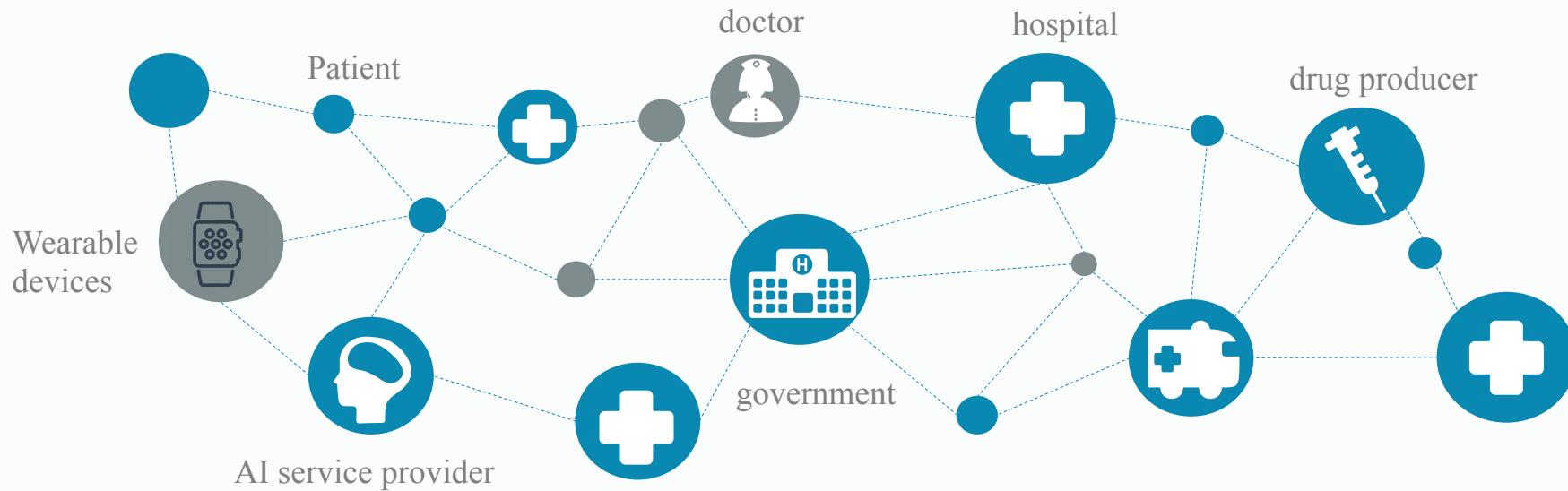
More accurate medicine,
doctor-patient relationships



Scientists

Open data marketplace,
federated AIs for good

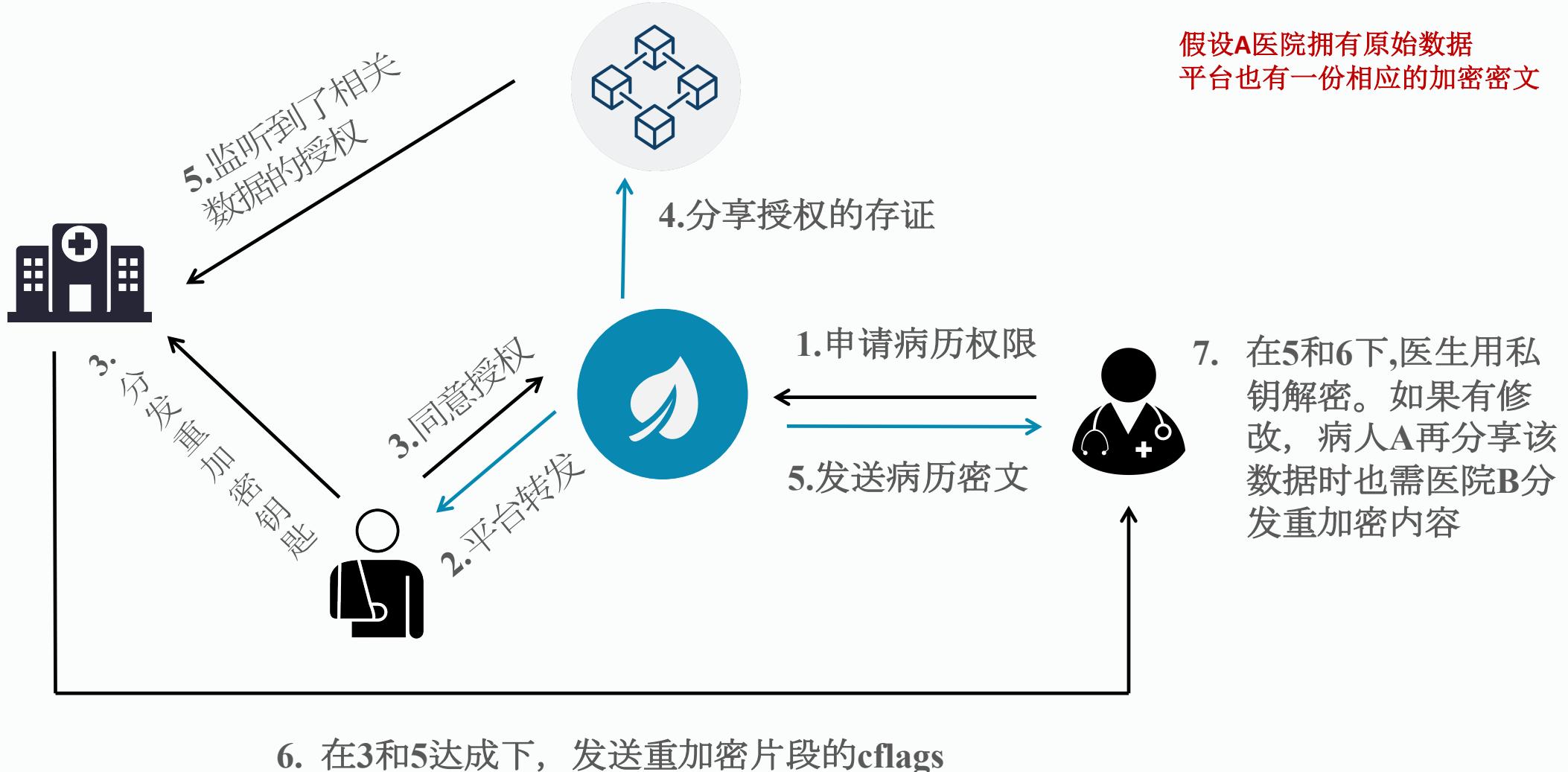
Federated Healthcare Ecosystem?



OpenHealth:

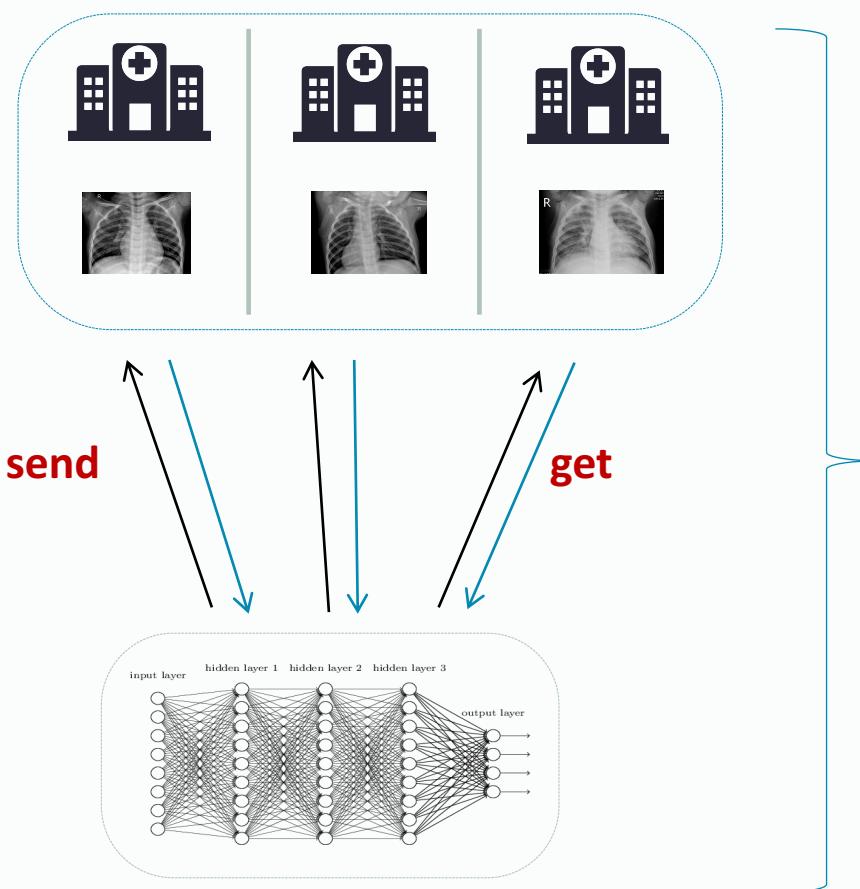
1. manage **data authorities** of private patient-controlled HERs
2. import existing isolated sources and maintain **shared, global and encrypted** records
3. obtain more accurate AIs through **private learning** from multiple patient data

Encrypted Storage Sharing



Health Data & AI Marketplace

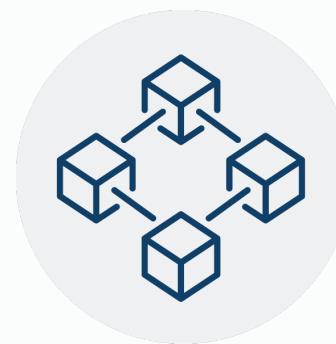
Federated workers



Model client

Federated Learning + Blockchain

distribute model and reward to data places!

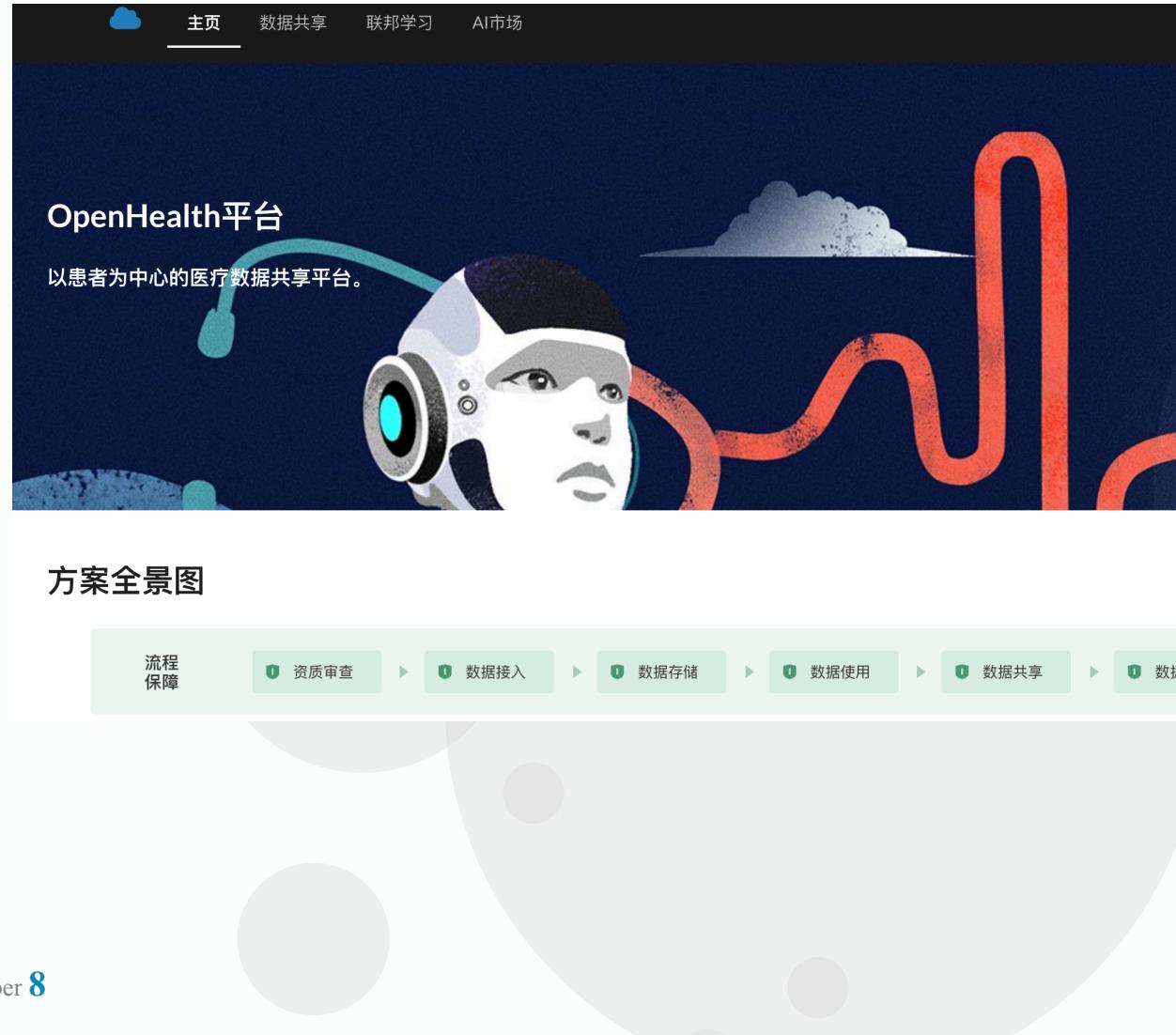


- On-chain market matching for joint AI modeling
- Off-chain routing & Training auditability
- ↳ Reward distribution

Demo Project

OpenHealth

1. Patient-controlled HER data platform, maintaining **shared, global and encrypted records**;
2. Federated data marketplace for private learning, jointly modeling three remote X-ray datasets of three hospitals;
3. Decentralized AI service marketplace, including three AIs now, pneumonia detection AI, reading comprehensive AI from EHR



基于DID的借贷联合风控平台

网络借贷爆发

近年来网络借贷业务的快速增长，一方面推动了普惠金融的发展，另一方面也滋生了诸多新型的欺诈手段，金融欺诈风险不断升级。



行业痛点

- ❖ 征信体系和数字身份体系不健全，用户逾期、多头借贷、身份冒用问题频出；
- 不同借贷平台各自为营，不愿共享彼此的客户记录，产生了一个个数据孤岛；
- ✓ 部分共享数据的质量和来源难以得到有效的验证，限制了AI风险定价能力；
- 用户隐私保护和身份主权意识觉醒，揭示了未来机构间共享客户数据的难度。

信贷体系中的多方诉求



借贷 平台



身份核实、信息可溯

核实借贷申请人的身份数据，防止虚假信息申请、冒用身份申请，保证数据可验证可追溯



信贷反欺诈、风险定价

协同其他平台进行联合反欺诈，如识别多头借贷等行为，这需要建立统一的客户数据标识符，以及让AI联合多平台特征对用户信用建模。

用户

数据隐私、身份主权

可多方验证的网络自主身份，并将身份数据归还于个人，保证数据的隐私存储和授权共享；

政府

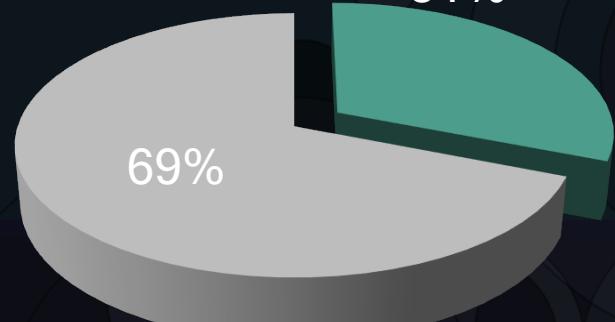
金融监管

确保用户利益，防范系统性风险，规范行业标准

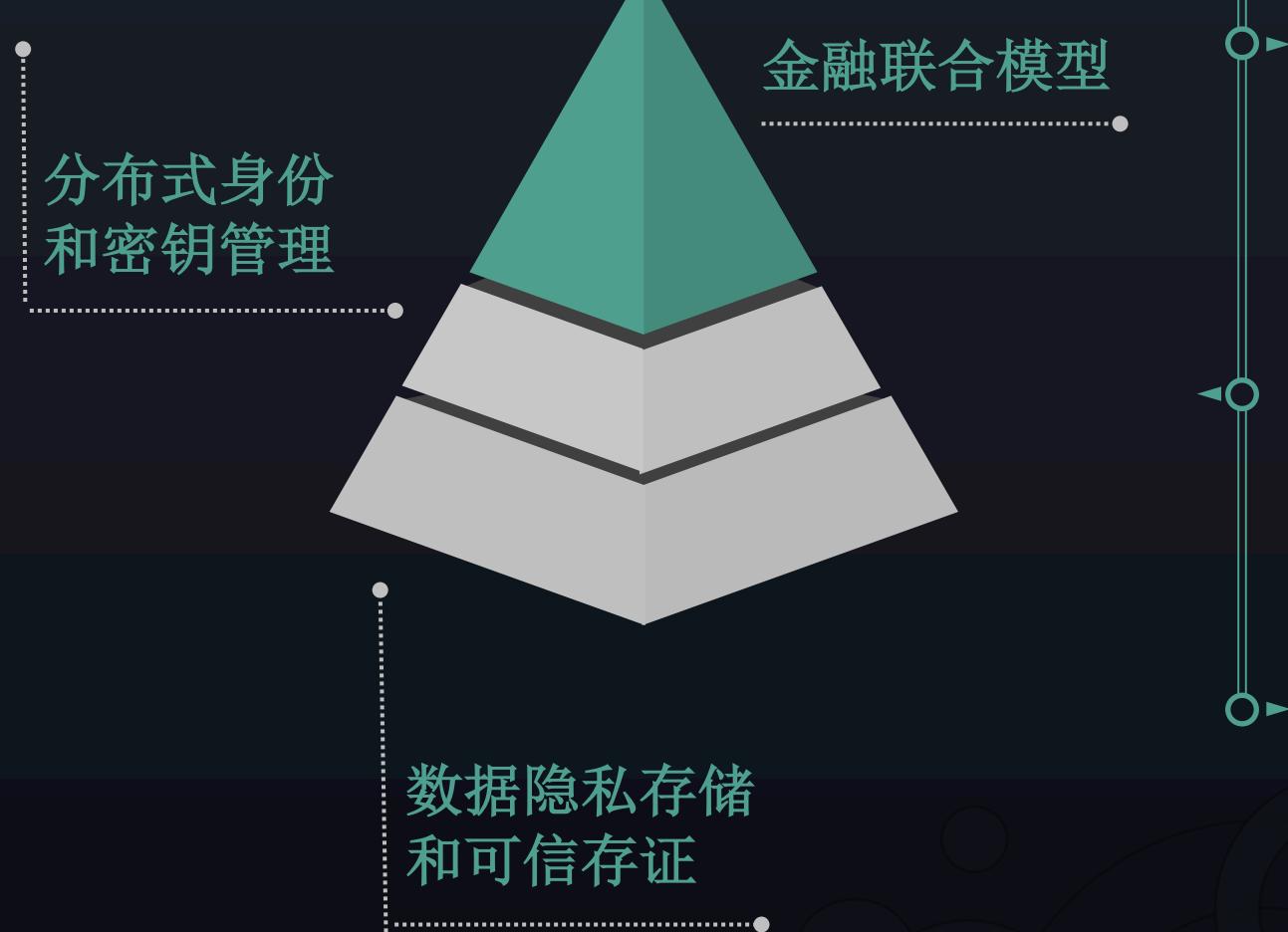
“据OTA统计，2018年有近6万起数据泄漏事件！”



欺诈受骗分布



多方诉求的内涵



Web3 + FinTech?

Start to be great

多方诉求的本质是对网络金融基础设施的需求。我们的方案可划分为三层架构：底层把数据主权归还用户，中层形成统一的用户网络身份标识符和凭证验证标准，上层进行金融的联合风控。



Authentication

Users

Sign-in method

Templates

Usage

Search by email address, phone number, or user UID

No users for this project yet

小额贷款行业联盟？

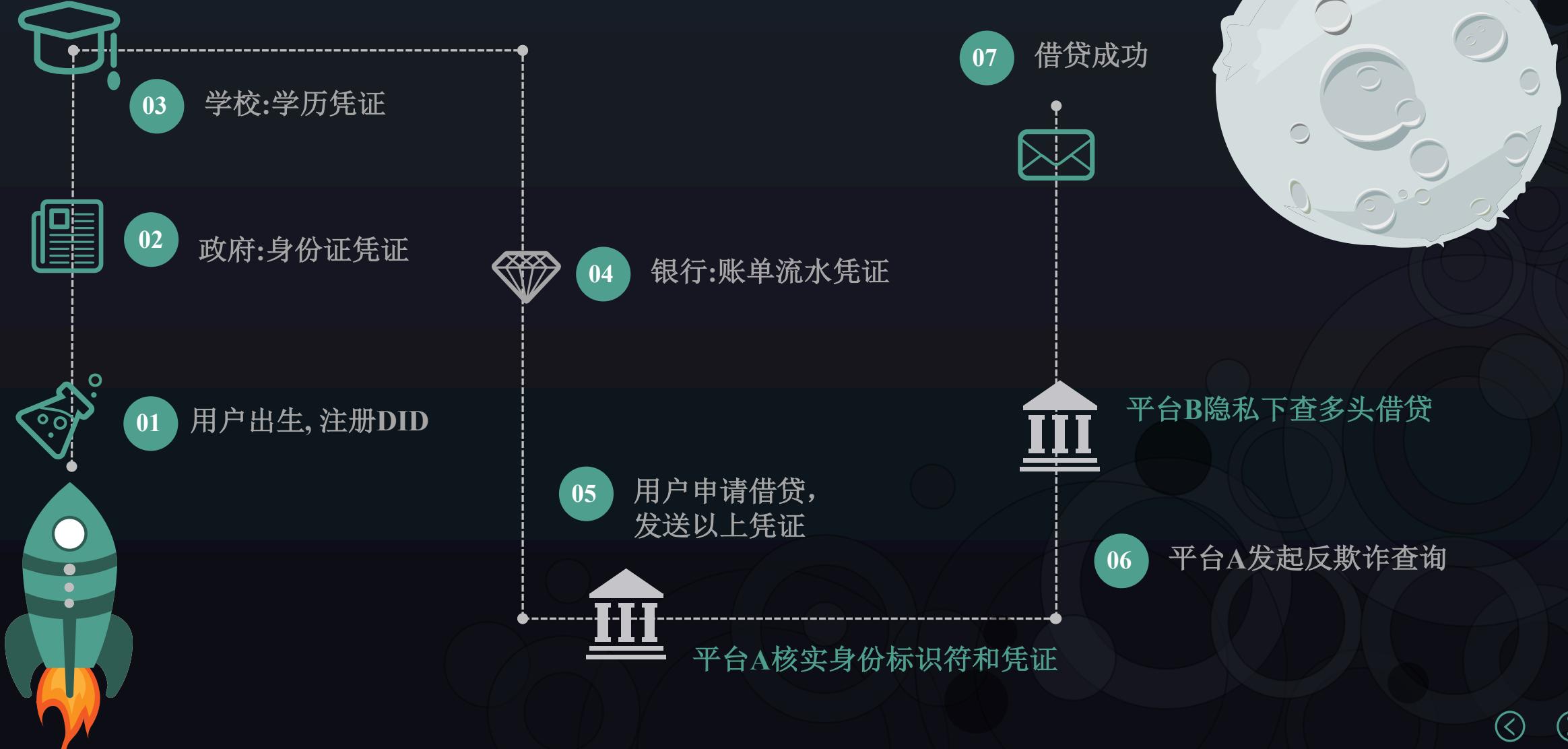
以分布式身份作为多方桥梁
在数据验证/计算层面互联互通！



分布式自主身份

To the Moon

凭证跟DID关联、DID跟公私钥对关联！



To the Moon

分布式自主身份

凭证跟DID关联、DID跟公私钥对关联！



03 学校:学历凭证



02 政府:身份证件凭证

凭证管理插件



04

银行:账单流水



01

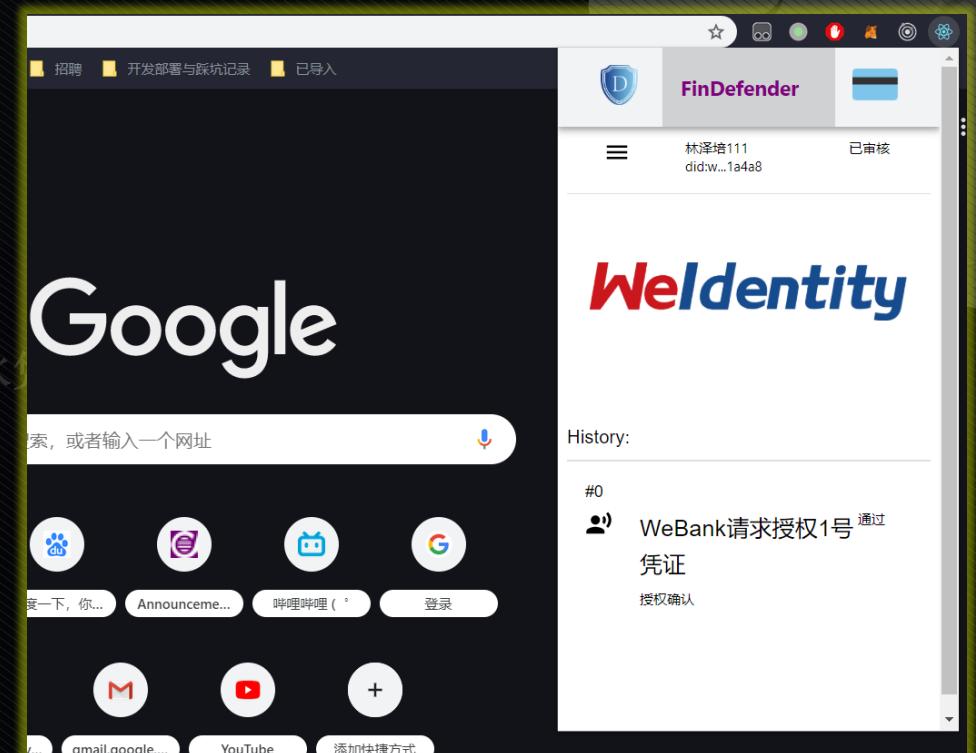
用户出生,注册DID
将数据归还给用户
由用户控制数据的分发

05

用户申请借贷,
发送以上凭证



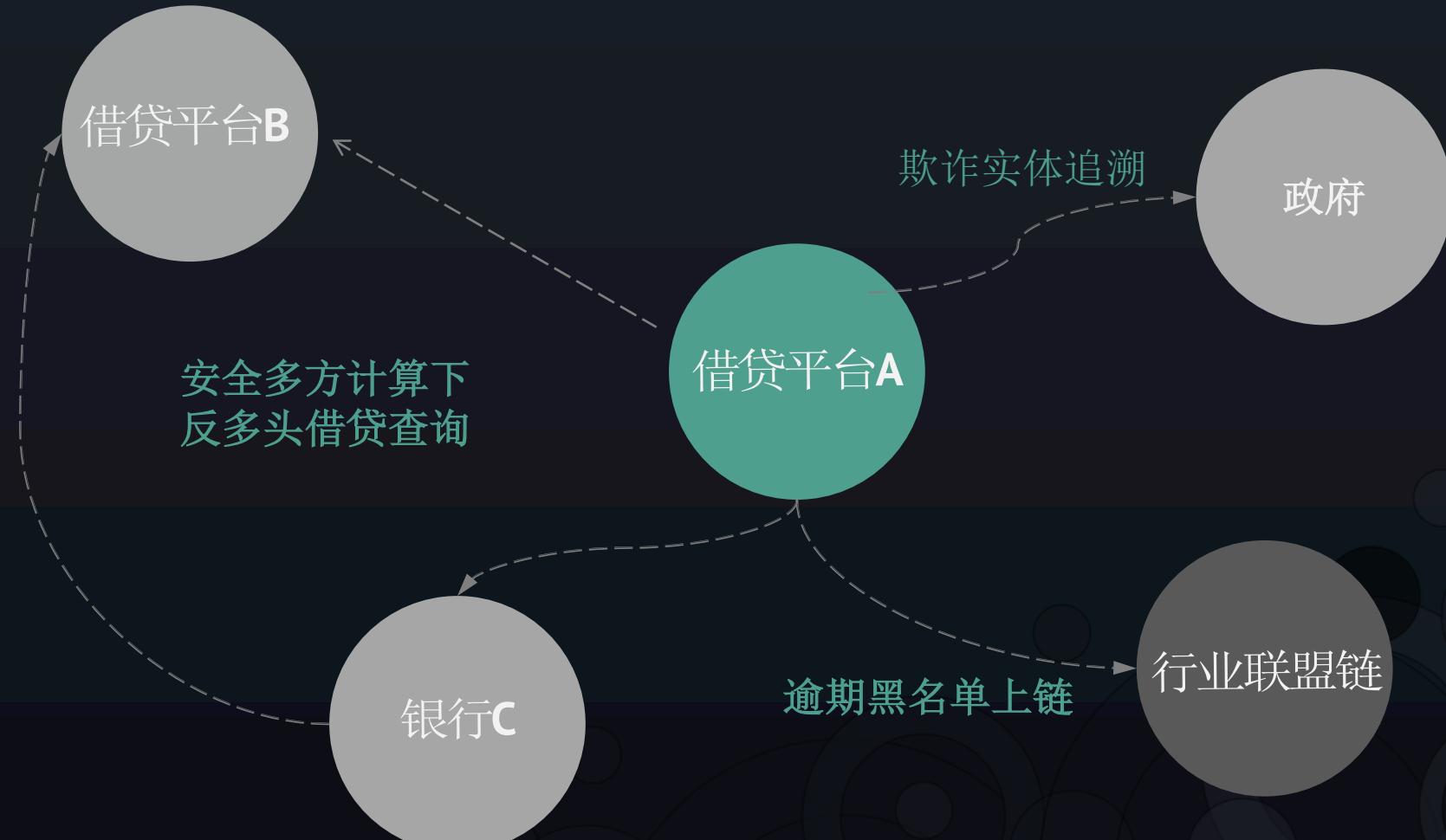
平台A核实身份标识符和凭证



06 平台A发起反欺诈查询

借贷联合风控

基于统一的**DID**标识符及可信凭证



技术方案

FISCO-BCOS联盟链
数据存证和可信溯源、
维护行业联合反欺诈库

01

代理重加密和IPFS
个人凭证数据的自主
管理、隐私云存储共享

03

02

WeIdentity分布式身份
DID和凭证生成、选择性披露、
微贷身份标识符、信息核实时

TEE、MPC隐私计算
反多头借贷查询、
部署隐私借贷合约

05 Federated AIs
联邦学习、金融大脑

Thank you