

## ADD 数据生态体系白皮书

版本 1.0

本白皮书仅供参考，不构成要约或任何形式的投资建议

随着项目的进一步发展，白皮书中的任何元素都可能发生重大的变化



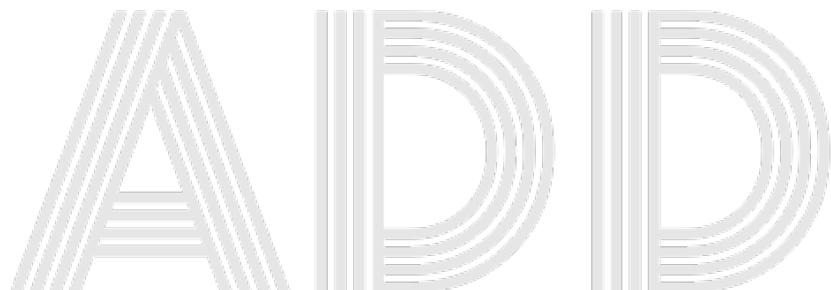
当大量杂乱无章的数据盘踞在我们的生活中

多数人感到的却是更多的恐慌和无助

你的想法和思路，才是最重要的资产

它们才能帮助你在繁杂的数据中找到自己想要的

进而让数据在你的手中发出无尽的光芒



Professor Viktor Mayer-Schönberger

世界大数据之父

ADD 数据生态基石理念奠基人



## 目录

序言

### 1、背景研究：数据产业 6

1.1	数据产业的定义及其涉及领域 .....	6
1.2	数据产业发展轨迹 .....	7

### 2、产业发展状况 10

2.1	数据时代已经到来 .....	10
2.2	数据获取途径多元化 .....	11
2.3	从流量价值论到数据价值论 .....	12
2.4	新时代的数据再定义 .....	13
2.5	技术进步促进生产关系变革 .....	14

### 3、数据产业需解决的问题 15

3.1	系统性数据安全 .....	15
3.2	个人隐私安全 .....	15
3.3	数据确权 .....	16
3.4	数据流转 .....	17
3.5	基础网络运营商 .....	18

### 4、ADD 数据生态区块链解决逻辑 19

4.1	区块链赋能数据产业 .....	19
4.2	数据确权与保护 .....	20
4.3	数据流转与交易 .....	20
4.4	去中心化云存储 .....	21
4.5	区块链+数据应用的发展 .....	22

### 5、ADD 数据生态架构与功能实现 24

5.1	ADD 基于 EOS 的数据侧链 .....	27
5.2	ADD 主要功能模块规划 .....	27

### 6、ADD 项目关键性内容披露 33

6.1	项目团队 .....	33
6.2	项目顾问 .....	34
6.3	代币分配与流通项目顾问 .....	35
6.4	ADD 项目所有人及 EOS 超级节点 (eosaddaddadd) .....	36

### 7、名词定义、免责声明与风险提示 36

7.1	免责声明 .....	36
7.2	风险提示 .....	38

结语

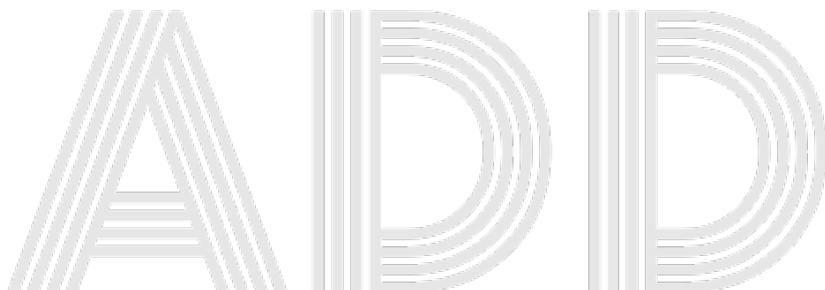


## 附录

附录 1：EOS 生态全景图 ..... 45

附录 2：区块链领域数据相关项目分析

(一)	Storj——去中心化的云存储应用 .....	46
(二)	IPFS——分布式存储及共享传输协议 .....	49
(三)	Lambda——去中心化存储底层公链 .....	50
(四)	Bluzelle——去中心化数据库服务商 .....	52
(五)	CPChain——面向物联网系统的基础数据平台 .....	53
(六)	Scry.info——去中心化的数据交易所 .....	54
(七)	Bottos——去中心化人工智能数据合约.....	55





## 序言

### ADD 愿景

ADD 将致力于建设基于 EOS 数据与应用生态，依托 EOS 搭建 ADD 数据侧链，提供互联网与大数据产业从“互联网+”到“区块链+”的转型的数据中枢与数据应用平台。

数据流动、数据分析与数据存储将不可阻挡的形成跨域趋势，ADD 试图打造的 ADD 数据生态是基于 EOS 体系的一条数据侧链，作为 ADD 的实时数据中枢与数据应用平台，这将是一个可扩展、低延迟、高交互、并拥有可扩安全链域护栏的分布式网络平台。

ADD 将自建数据模型、数据引擎与引索机制为生态用户提供可视化、便捷的自设更细颗粒度的分析、提供专项维度数据流转服务与数据分析应用体系，同时提供为用户提供一站式的去中心化数据应用与区块链服务的平台级解决方案。

ADD 数据生态愿景的 3 个关键词：

区块链的数据中枢（数据银行）、ADD 数据模型与引擎、数据可视化与数据应用平台



## 1、背景研究：数据产业

## 1.1 数据产业的定义及其涉及领域

所谓数据产业,是指在互联网时代,以数据为基础而产生的行业及相关产品,包含数据收集、数据存储、数据筛选、数据分析、数据服务、大数据等多段产业链条,虽然本白皮书中的“数据产业”与“大数据”并不能相互替换,但有较高重合度,可以将其理解为一种泛指。数据产业是一个通过数据收集、加工、应用与管理等业务,从众多数据中筛选有效信息,便于B端用户对市场以及现有受众的了解,有助于后续提出方案并展开商业运作的数据产业链。

眼下，随着PC端、移动端、IOT的发展，数据的收集、存储、流转正迅速由中心化向分布式发展；从线上线下数据化到数据流动时代前进。数据产生的惊人速度与数量，使人们开始思考数据的应用价值，如今人们对于数据产品的使用也正经历着本质的变化。各类数据产品正在改变着企业、个人甚至于物流运输、政府治理等诸多领域。

数据细分与行业应用分布如下：





图表 大数据产业应用全景图 (FIRSTMARK 2017)

AID

The letters A, I, D are formed by multiple parallel grey lines, creating a striped texture. The letter A has three vertical stripes. The letter I has two vertical stripes. The letter D has four vertical stripes.



## 1.2 数据产业发展轨迹

### 1.2.1 互联网推动数据产业演进

互联网时代，数据已成为一种新型的资产而逐渐被各行各业认可。数据产业的成长依赖于人们对互联网发展史的理解，从互联网诞生至今，其发展大致经历了三个阶段分别为：单向传输、双向互动、多方参与。

互联网发展的各个阶段			
	1970~2000 年	2000 年~2010 年	2010~至今
数据	数据单一	数据极大丰富	数据资产化
用户功能	信息单向传播	信息双向交互	多方参与, 信息自由流转
功能价值	网络化	交互化	终端化
商业应用需求	简单呈现, 可有可无	综合呈现, 成体系, 可统计	产品化呈现, 全民化, 资产化
典型类别	搜索引擎、门户网站、电子邮件	社交网络、电子商务、网络游戏、即时通信、门户网站	移动应用、应用商店、网络视频、社交网络、位置服务
工具演化	大型计算机	PC、手机	智能手机、智能可穿戴设备

图表 互联网发展阶段

第一个阶段是互联网发展的极早期，该阶段主要是以内容编辑、信息发布为主要功能，互联网所提供的数据是由网站运营者人工的录入并呈现给用户的，用户通过单向的数据传输进行交流，此时的互联网并不开放，是众多局域网的集合，彼此之间交流并不顺畅，虽初步实现网络的基本功能，但对于大多数民众而言互联网服务还浮于表面，可用性极低。

第二个阶段是互联网迅速在各领域渗透，PC 逐渐成为普通用户的必备品时，互联网的用户数量随之增长，互联网运营者开始重视与用户的互动交流，用户既是互联网数据的浏览者，也是互联网数据的创造者，互联网的数据量由于用户的双重角色而加速增长，此时为双向沟通阶段。

第三阶段以第二阶段的数据积累为基础，来自互联网的数据不再受限于单一网络，实现了互联网的多维数据整合。由“平台”向“生态”转型的互联网产业新格局形成，大型互联网公司开



始建立起数据库，并逐渐完善生态体系形成“用户闭环”，通过业务集中化实现源于生态内的数据整合应用。

以上是传统互联网即有的发展轨迹，其商业逻辑是通过将成本转嫁 B 端，从而向 C 端的提供免费服务，同时获取 C 端用户的个人数据。收集数据的真正意义并不完全在于掌握庞大的用户数据，而在于将用户数据进行分析整合，从而转换为具有战略价值的数据模型，数据的增值能力，由相关领域技术是否支持最终落地于业务场景而决定。

互联网的多维化造就了数据产业及大数据产业链，海量数据的分析整合可得出具有商业价值的结论：如数字营销的精准广告技术、物流系统管理的最优化路线计算、电商及内容分发平台的推荐算法、医疗产业的罕见疾病研究等，数据在实际应用中的角色正在从景上添花到不可或缺的角色转变，这些数据驱动的产业将成为下一个时代变革和发展的核心驱动力。

### 1.2.2 聚合及分析能力是数据产业的关键

数据是数据产业的基础资料，对于数据产业而言，数据源及已有数据是特殊资产，位于数据产业链的最前端。世界经济论坛报告认为，未来的大数据将成为新的财富高地，数据是未来科技时代的核心资产。

伴随各式终端设备及嵌入式计算的发展，数据源成为愈发重要的产业环节，数据收集是数据产业的源头，除了移动终端、可穿戴设备、IOT 传感器外，对于个人数据获取的“流量入口”也是重要的数据获取渠道之一，所谓的流量入口通常指搜索引擎、门户网站、网址导航、社交媒体、支付平台等，另外也包括共享行业、短途配送、内容分发平台等次级流量入口。

依据互联网的发展路线及趋势可大致分为四个阶段：“人口红利、流量红利、数据红利、信息红利”。具体来说，PC 及互联网普及带来的人口红利；移动互联网产生了流量红利；海量数据的数字化服务产生的数据红利以及大数据的互联互通的信息红利等，目前的互联网正从移动互联网的流量红利，过渡到基于海量数据的数据红利发展阶段。

在数据红利时代，数据的大量出现改变了原先的经验法则和指导发展的模式，而通过对于移动终端、传感器、流量入口的数据进行数据处理、分析及关联也是产业链的关键环节，数据产业的商业模式变革了传统产业的经验法则。数据应用的本质在于“提升运作效率”，而对于海量数据的搜集、处理和展示，使得数据具备的产业价值得以变现并使得人们可以无限接近、了解市场的真实情况、资源的配置状况、优化企业的运作方式，并且能够为企业的进一步决策奠定充分的基础，最终通过数据的积累实现在行业内建立起核心竞争优势的目的。

### 1.2.3 数据价值的流通将产生巨大的社会价值

数据应用的终极目标在于最终实现 B 端运作效率的提升，而数据价值的流通则意味着在全社会范围内的数据赋能，将为社会带来巨大的价值。



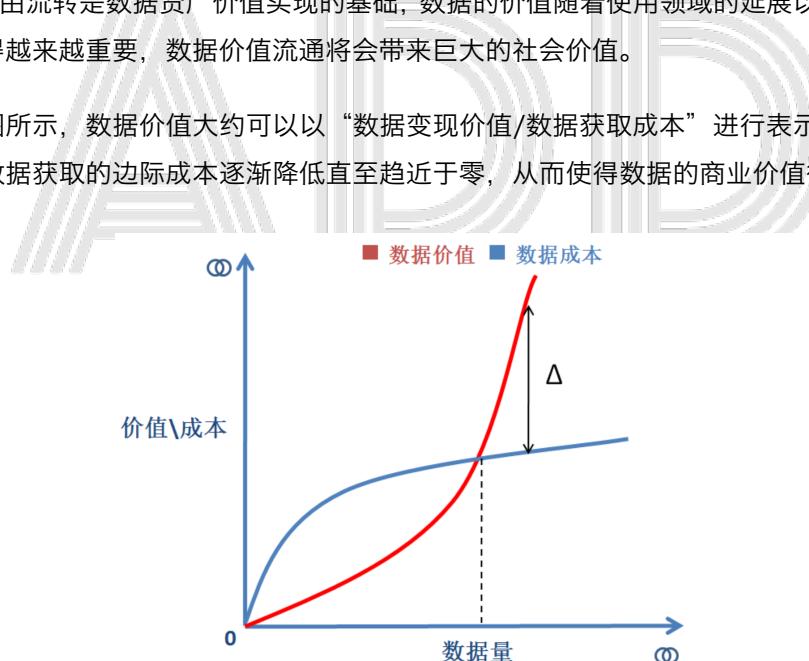
立足于数据产业，数据提供方可分为第一方数据、第二方数据以及第三方数据：第一方数据是指用户自有数据；第二方数据指用户所在行业的用户数据；第三方数据则是指脱离开业务背景和利益关系的机构提供的数据。

对于行业内部的组织数据，因其数据获取成本极高而数据备份成本极低，数据的归属问题使得数据流通受阻。且组织内部数据的物理及逻辑隔阂易形成数据竖井，因此在当前的环境下难以形成数据的流通价值。

在组织内部、行业内部的数据无法解决信任、共享、确权等问题的情况下，人类社会的数据价值难以充分利用。数据的有效交易是实现数据流通价值的关键，第三方数据服务商的产生成为必然——第三方数据服务商可利用自身的数据资源及收集数据的能力，通过售出经过加工的数据产品或是直接出售数据资源的模式获利，第三方数据交易所的应运而生。这种现象使得各类组织机构可以根据自身情况选择性购买即时数据，将外部数据通过交易机制引入组织内部，实现组织内外部数据的自由流转进而实现资源分配及运营策略的优化，为组织高效运转提供帮助。

数据资产的价值体现需要海量数据的流转聚合，孤立分散的数据资产并不具备完整的商业价值，数据自由流转是数据资产价值实现的基础，数据的价值随着使用领域的延展以及使用频次的增加而变得越来越重要，数据价值流通将会带来巨大的社会价值。

如下图所示，数据价值大约可以以“数据变现价值/数据获取成本”进行表示，随着数据量的增加，数据获取的边际成本逐渐降低直至趋近于零，从而使得数据的商业价值得以大幅上升。



图表 数据价值与成本示意图



## 2、产业发展状况

据相关报告显示，自 2009 年，全球互联网上的数据正以每年 50% 的速率增长，截至 2017 年全球的数据总量为 21.6ZB ( $1ZB=2^{12}TB$ )。未来数据增长率仍将维持在 40% 左右，预计到 2020 年全球的数据总量将达到 40ZB，与此对比的是 2025 年全球的数据总量将达到 163ZB。

### 2.1 数据时代已经到来

#### 数据获取的途径更加多元

数据量的快速增加离不开互联网化、移动互联网以及 IOT 的发展，在互联网+及云时代，线下数据的逐渐引向线上，数据大量产生，我们使用标签、流量来描述数据。

随着数据获取的途径逐渐扩展以及应用场景的渗透，数据已经开始影响到各行各业，成为重要的生产要素。海量数据的挖掘及运用，在当代商业社会中也成为重要的核心竞争力。

#### 从流量价值论到数据价值论

当数字设备逐渐取代了独立的模拟设备，算力为终端设备赋能，越来越多的数据可被自动生成，数据处理技术的快速迭代，数据存储成本显著降低，这一系列变化使得数据能够被标注、分析与利用，数据产业的崛起依赖于技术、设备以及资本的支持，其本质在通过技术手段对数据清洗、整理及分析，从而服务于越来越多的线上或线下业务场景。

在当前移动智能终端的用户逐渐趋于饱和，移动互联网的发展进入了变现以及转型阶段，互联网企业势必寻求新的增长点。由移动端生成的海量数据进行分析，通过大数据从而提高企业运作效率的技术手段正在逐渐被挖掘，互联网已由移动互联网的流量价值论，逐渐步入后移动互联网的数据价值论。

#### 新时代的数据再定义

当代社会对于数据的使用方式正在经历根本性的变革，随着越来越多的数据被生成以及信息技术的快速发展，人们越来越多地依赖全局的数据分析而不再依赖于随机抽样，数据研究的加深也使得人们不再需要追求精确度，更多的是通过数据之间的对比寻找其相关性，全面的数据分析不仅使得企业运营更加有效，而且正在改变个人生活的方方面面。

从“为什么”到“是什么”，我们不再需要知道事务背后的根本原因，我们只需要让数据呈现出结果。数据的本质正在被逐步地重新定义：从优化资源配置到企业的核心生产要素。

#### 技术迭代促进生产关系变革



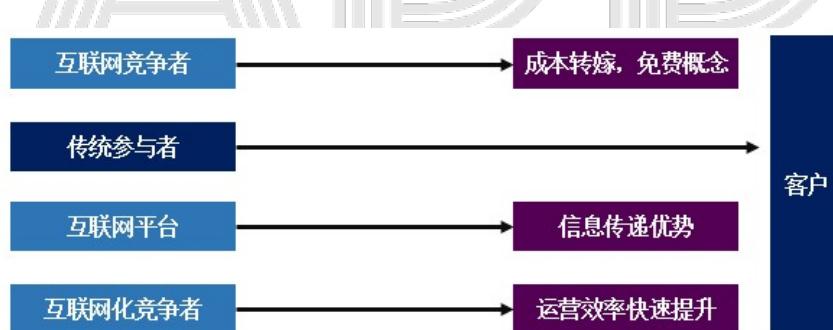
而随着互联网移动化以及端对端的信息交流快速增加，互联网的发展从中心化、平台化、生态化逐渐朝向分布式化发展，从互联网思维过渡到区块链思维：数据的生产、流通及存储正在逐渐下沉，带来的转变是云计算正逐渐向更加分散的雾计算以及边缘计算发展。

其趋势是由中心化的服务提供，如广播台、电视台、门户网站、数据库，转移至分布式的端对端分享，如微信、Line、Whatsup、滴滴打车。数据的下沉导致数量庞大的“小数据”正逐渐形成规模，并伴随着分布式网络的崛起，最终带来的是生产关系的变革，对于用户及商户的角色定位正经历转变，由消费者、生产者转型为生产消费者。

## 2.2 数据获取途径多元化

技术的快速发展带来了数据的大量涌现，互联网+、移动互联网以及物联网时代的到来，对于数据产业产生了深远的影响，互联网是个人与组织、社会互相关联、互联互通的时代；互联网+则是以互联网服务大众、以互联网技术对于产业环节进行优化，从而提高生产力的时代；移动互联网与物联网代表着联网终端设备的进一步扩展。

互联网+是通过互联网技术赋能传统产业，互联网企业凭借成本优势、渠道优势、运营优势三方面对于传统企业发起挑战，从而使得传统企业面临巨大的竞争压力，其最终的结果创新停滞、转型失败的企业退出市场，并在行业内重新形成一个以互联网为基础的竞争平衡。互联网+代表着数据获取的纵向扩展，并总是伴随着大量传统产业数据的互联网化，大量而庞杂的传统行业数据通过标签化、原子化逐渐由线下转移至线上。



图表 互联网+以及产业互联网化

相比于互联网+数据纵向扩展，移动互联网和物联网则是数据获取的横向扩张，受益于移动终端、可穿戴设备、物联网传感器等微型计算机的普及，并伴随着云计算、云储存、大数据的技术发展。智能手机、平板电脑等计算终端的快速出现提升了数据搜集的广度，催生出大量的联网用户并创造了巨大的商业利益。根据报告显示，移动互联网用户数量已在 2017 年 12 月达到 7.53 亿人，占到中国互联网网民的 97.5%。

尽管目前智能手机的高速增长阶段已成为过去式，但在智能手环/手表、VR 眼镜、服务机器人、游戏机、智能音箱、家用智能医疗器械、智慧家居、无人机等各类传感设备和智能设备的



快速发展下，数据将越来越广泛存在于现实生活中，互联网的接入方式、体验方式、使用习惯、操作模式发生巨大的变革，万物互联互通的物联网时代正在逐步到来，伴随着的数据量爆炸式增长为数据产业带来了发展机遇。

## 2.3 从流量价值论到数据价值论

移动互联网的普及激发了人们对于流量的需求，移动互联网渗透率以及用户数量的增长趋于平缓，运营商经历了从人口红利向流量红利的过渡，并开始朝向数据红利转型，从3G时代的语言、文字到4G时代的图像、视频，在宽带增加和移动互联网发展之间，创造了一个用户可以获得几乎任何所需要的互联网世界后，移动互联网用户的额外网络使用带来了越来越低的边际收益，智能手机普及率达到顶峰后，移动互联网和通讯速度发展带来的流量红利进入瓶颈，接入用户和用户使用互联网总时长已经到达天花板。

伴随着流量红利进入尾声，在移动端的用户使用习惯逐渐形成，互联网巨头对于现有流量的控制、引导以及对于用户的圈地运动日趋白热化，腾讯与今日头条的鹅头大战，反映的即是对于以新闻、短视频、直播、流媒体等移动互联网流量的最后窗口的争夺。以微信、淘宝、百度、QQ、微博、LOL所代表的超级流量入口，以及新闻、短视频、出行、短程配送、流媒体为代表的次级流量入口，移动互联网的市场版图构建已然接近完成，最终实现的是对于存量用户的管理以及运营，而新的市场入局者则难以在日趋完整的市场版图中杀出重围。

一方面是流量红利陷入瓶颈则代表着C端移动互联网用户的增长逐渐接近平缓；另一方面是经历十年发展的移动互联网带来的海量数据仍未得到有效的利用，伴随着数据在生产要素的地位逐渐上升，数据密集型产业逐渐取代原有的发展模式，以To B端以及To G端的企业数据服务正经历着高速发展的阶段，如亚马逊AWS、阿里云、腾讯云、政务云以及基于海量数据的人工智能、市场预测、大数据、云存储的崛起则成为必然趋势。

阿里研究院在一篇名为《新经济：从行业分工到平台共享》的报告中提出“信息革命”的概念：“第一次信息革命是计算革命，计算机、互联网的普及让专业化计算成为可能；第二次信息革命则是数据革命，“云计算+大数据”替代“计算机+软件”改变了技术发展轨道。”，因此在云计算、大数据、物联网、移动互联网、智能硬件的技术快速推进的新局面下，数据使用的效率将会得到快速的发展。

在数据红利时代，通过对于数据的合理应用，能够为企业带来成本降低以及收益提升，可以通俗地理解为“大数据对于生产要素以及企业增长的驱动力”，线上以及线下所产生的海量大数据，已然成为企业的重要生产要素之一，通过对于这些数据的开发和利用，一方面带来企业的数字化转型，另一方面实现互联网对生产和增长的深远影响。



## 2.4 新时代的数据再定义

互联网已经经历了移动互联网时代，传统产业的“互联网+”也在如火如荼的进行中。在这种转变发生之际，当前的互联网巨头早已完成原始资本积累，并逐渐向外延领域进行扩展，而数据产业正是其发展的核心领域。

在移动互联网时代，随着计算能力嵌入终端设备，数据生成的边界逐渐由线上数据下沉至线下，线下数据是指如贸易物流、供应链金融、农产品、制造业等传统产业互联网化，以及如短途配送、出行旅游、线下零售等生活关键型数据，借助互联网实现线下数据资源的整合及联网，从而使得庞杂的线下数据能够通过互联网实现数字化，数据的获取也逐渐与中心化的数据库脱钩，云计算的普及带来的万物互联，计算能力以及存储能力的分布化发展趋势，使得设备得以在任何时间、任何地点获取任何数据。

我们正生活在一个从任何意义而言数据量都极其充沛的时代，当代社会对于数据的使用方式正在经历根本性的变革，随着越来越多的数据被生成以及信息技术的快速发展，对于传统行业产生了极大的冲击，正如 Viktor Mayer-Schönberger 在其著作《大数据时代：生活、工作与思维的大变革》中写道，数据的快速增长并指出了数据的发展导致三个重大的思维转变：人们越来越多地仰赖全局的数据分析而不再依赖于随机抽样、数据研究的加深也使得人们不再需要追求精确度、更多的是通过数据之间的对比寻找其相关性，数据产业的发展使得人们可以从更全面的数据、更多样的维度去理解事务的全貌，并进一步窥探事务的本质。

全面的数据分析不仅使得企业运营更加有效，而且正在改变个人生活的方方面面。新产生的数据的使用情况越来越多地按照其重要程度进行分析，而数据的重要程度体现在诸多方面，包括实时处理和低时延的需求、使用的特殊性，以及数据不可用所造成后果的严重性等核心问题。

根据 IDC 的分析，在 2025 年全球数据中接近 20% 的数据将成为影响日常生活的关键数据，其中 10% 将成为超关键数据，这些超关键数据能够对用户的健康和幸福造成直接和即时的影响，例如商业航空旅行、医疗应用、控制系统和遥测技术数据等。

这些被大量进行标注或是分析的数据，重新定义了数据的社会意义，数据在企业及日常生活中越来越被频繁地整合并且正在快速地实现实时化，人们对于数据的使用需求以及依赖程度日益加深，数据也逐渐成为日常生活中至关重要的存在，从物流、通信、防伪溯源、船舶航运、航空航天等。从“为什么”到“是什么”，我们不再需要知道事务背后的根本原因，我们只需要让数据呈现出结果并遵从数据带给我们的启示，数据从优化资源配置的定位重新被定义为企业的核心生产要素，数据的交互不仅可以起到通知的作用，而且还可以帮助企业确定采取何种行动以及决策，并为企业的工作流程以及大众的食衣住行带来实质性的改善。



## 2.5 技术进步促进生产关系变革

早在 2001 年，麦塔集团 (META Group, 即现在的 Gartner) 的分析员道格 · 莱尼 (Doug Laney) 从三个维度定义了数据的增长特征，即 3V 模型：数据体量 (Volume)、速度 (Velocity) 与多样性 (Variety)，指出“大数据”是大容量、速度和各种信息资产，具有成本效益的需求，创新形式的信息处理增强洞察力和决策。

2012 年，IBM 在 3V 模型之上，加入了对于数据真实性 (Veracity) 的考虑，并指出“大数据时代较之以前具有两项显著区别，大量产生的新型数据不再适用于传统数据库，与此同时，分析能力对企业实施大数据具有至关重要的作用”，IBM 认为，尽管前 3 个 V 涵盖了大数据本身的关键属性，但真实性是当前企业亟需考虑的重要维度，将促使他们利用数据融合和先进的数学方法进一步提升数据的质量，从而创造更高价值 (Value)。

以“云计算+大数据”为主的息革命将带来数据传递、数据使用效率的快速提升并伴随着成本的降低以及安全性的提高，从而在互联网下实现更快速、更安全、更便宜的信息传递，数据产业也正从量变向质变过度，数据量的增长远大于技术进步的脚步，从各种意义而言也越来越难以被中心化计算存储设备所掌握，因此分布式系统也成为越加重要的存在。

在技术进步的基础上，互联网经济的商业模式也在不断革新，而这种革新始终围绕着去中心化的发展方向，从过去的 Web1.0、Web2.0 到 Web3.0，分别代表在功能上实现单向“网络的结合”、双向“内容的整合”、多方“功能的聚合”三个阶段，从门户网站时代、搜索引擎时代到社交网络及自媒体时代，嵌入式计算的发展、数据逐渐由线上到线下、中心节点到分布式网络节点下沉，商业模式演进的过程中，中心节点被不断降权。

随着互联网移动化以及端对端的信息交流快速增加，互联网的发展从中心化、平台化、生态化逐渐朝向分布式化发展，从互联网思维过渡到区块链思维：数据的生产、流通及存储正在逐渐下沉，即时通信、共享出行、二手交易电商、P2P 短租等，这样的数据下沉来自于多方面的因素，但终归离不开平台的不断降权、节点及个体意识的抬头。

数据的下沉导致数量庞大的“小数据”正逐渐形成规模，并伴随着分布式网络的崛起，安全可行且具备个人识别的小数据正在成为“大数据”之外的新领域，并最终带来生产关系的变革，对于用户及商户的角色定位正经历转变，由消费者、生产者转型为生产消费者，而平台也由交易的中介者逐渐转向规则的制定者、生态的参与者，平台的职能也从传统意义上的管理，向信任机制以及必要的激励措施制定的核心职能进过渡。



### 3、数据产业需解决的问题

在万物互联时代的全面快速发展下，越来越多的设备被接入互联网并且数据正在不经意间被动地被互联网公司所获取并使用，数据的快速增长使得数据的安全边界正在逐渐降低，数据相关的问题一再被凸显出来，诸如数据安全性、隐私安全性、数据确权、数据流通、运营商问题等。

#### 3.1 系统性数据安全

##### 3.1.1 数据泄漏

数据的大量汇集不可避免地增加了数据泄露的风险：一方面数据的集中存储无法避免数据丢失、被窃取的问题；另一方面，隐私及敏感数据的集中存储，无法将使用权以及所有权分开界定，对于大多数互联网公司而言，用户数据而非市场渠道（如手机 App）才是本质上的互联网公司的核心竞争力，即便在欧美等隐私数据保护法案（如 GDPR、DATA Act）相对完善的国家，仍然会出现如 Facebook Cambridge Analytic 事件所产生的高达 8700 万用户数据泄露的事件，基于数据的使用以及大数据分析都并未考虑到其中涉及的用户隐私问题。

##### 3.1.2 单点故障

数据存储于中心化的节点，并通过互联网以及云环境中进行传输和存储，用户对于自身数据在云中的安全风险并没有实际控制的能力，数据安全完全仰赖于服务商，如果服务商自身对于数据安全的风险控制存在不足，则会产生单点故障的问题，数据的可用性将会下降并导致数据的丢失，最终受到波及的仍然是使用产品的消费者。

##### 3.1.3 黑客攻击

对于中心化的存储节点而言，除了对于硬盘、带宽、服务器的高技术要求、高资本投入外，数据存储的安全性也将面临越来越大的考验，数据存储在网络空间中是更容易被发现及攻击的目标，随着数据的开放程度加深，有价值的、敏感性高的数据将会成倍提高并带来愈发频繁的黑客攻击事件，一旦出现数据安全技术问题，将会给用户和企业造成巨大的损失。

#### 3.2 个人隐私安全

隐私安全性通常是指个人的数据隐私及其相关衍生问题。

##### 3.2.1 个人数据的隐私安全

个人数据是指与“活着的自然人”（Living natural person）直接或间接相关的任何信息，无论是实际识别它们还是使其可识别。在云存储的使用中，用户可以依据自己的喜好将设备中的音频、视频、文件上传至云盘中。然而，目前的主流云存储平台对于数据的保密工作仍然存在诸



多缺陷，对于数据并未提供充足的保密技术，网盘的管理员能够从服务器中直接查看和删除用户上传的文件。

### 3.2.2 个人数据的被遗忘权

被遗忘权是一种已经在欧盟付诸实践的人权概念，即人们有权利删除或移除自己负面或过时的信息，但这项概念和言论自由间的冲突以及产生互联网审查的疑虑因此引发争议，根据 GDPR 所认可的广泛原则是：“在没有令人信服的理由继续处理个人数据的情况下，个人可以要求删除个人数据”。

根据 GDPR 规定，个人在以下情况有权要求删除或移除个人数据：个人资料对于搜集或处理目的不再需要者；数据主体撤回其同意，且该处理已无其他法律依据者；数据主体依对处理提出异议，且该处理无其他优先适用之法律依据者；该个人资料遭违法处理者；控管者依其受拘束之法律有义务应删除个人资料者。然而，在实际应用中，不考虑云存储服务提供商在个人要求删除的情况下是否会执行删除，仅从技术角度而言，由于当前互联网并无法得知数据的流向，因此数据是无法被（完全）删除的。

### 3.2.3 搜集与使用范围的界定

随着大数据产业的发展，对于数据的分析、处理、标注使得数据的价值密度正在逐渐增加，数据应用的范围也在逐步扩大。然而，人们往往忽视了**对于数据的搜集及使用范围界定**，最为切身的例子是手机 App，即便在《用户使用协议》或《法律声明及隐私权政策》中，以文字形式明确对于信息搜集的相关需求，该项内容却也往往流于形式，手机 App 被滥用于搜集个人数据、读取联系人、获取位置信息、运动数据、已安装应用列表，即便这些信息往往与该应用所提供的功能毫不相关。

对于数据的搜集与使用，以 2018 年 1 月 3 日所发生的支付宝事件最为著名：支付宝推出了“年度账单”产品，统计用户年度消费情况，给出“颜值正义”、“小确幸”等年度关键词，并通过字体大小和颜色削弱明显程度捆绑了《芝麻服务协议》，如果用户没注意到，就会直接同意这个协议，允许支付宝收集用户的信息，包括在第三方保存的信息。

## 3.3 数据确权问题

互联网企业首先看到了互联网数据所产生的价值，并通过服务协议要求用户对其获取用户数据进行授权以换取产品的使用权，从而获得数据的所有权以及数据的衍生价值。然而，数据的价值应当回归于用户本身，并且用户应当具备是否提供数据的选择权。

### 3.3.1 数据归属权

以智能手机为例，用户通过购买智能手机，即获取了该手机的所有权，因此同样也有权拥有对于该智能手机产生的相关数据所有权、使用权、收益权以及处置权，用户通过购买流量并提供



用户数据从而获得流量数据的使用权，这种使用权并不应以捆绑服务作为使用条件而向服务商提供用户数据的所有权。

在当代互联网社会，用户并没有对于使用这些产品所产生数据的所有权，互联网所建立的成本向 B 端转嫁的商业模式，互联网企业通过廉价的服务，在成功吸引 C 端用户的同时获得了用户数据的所有权，并将其应用于广泛的商业领域换取收益，而这种收益就是以牺牲用户权益为代价为前提的，起码对于用户而言，这是隐性成本。

### 3.3.2 数据使用权

用户在合法的情况下，有权控制自己数据的使用，并且对于自己的数据使用应当具有知情权以及选择权：知情权是指用户在使用服务时，企业应当提供用户数据的使用方式，并且对于用户数据的使用应当与其服务的目的相关；选择权是指用户在购买服务或是产品时，数据的使用条款必须是具有可选性的，且必须以明确的语言加以解释，如果用户对企业使用数据的方式不满意，用户有权要求该企业将其个人数据删除。

### 3.3.3 数据分润权

企业在使用用户数据时，最基本的原则是不能私自贩售用户的 data，而对于使用用户 data 所产生的相关收益，用户有权且应当获取。个人 data 经脱敏、脱密后，其所有权属于个人，使用权由 data 拥有者授权使用，收益权由个人或其授权的主体拥有。用户个人 data 所产生的价值应当归于用户本身。

## 3.4 数据流转

随着人工智能、数据产业的发展，企业以及个人用户对于 data 的使用需求正在逐步上升，而当前 data 的流通则受到行业内部、行业间广泛存在的 data 孤岛，以及 data 交易环节中，所面临的信息不对称、价格不透明、data 数量不足的问题阻挠，使得 data 的价值无法充分发挥。

### 3.4.1 数据孤岛

data 孤岛是 data 在互联网时代的碎片化现象，源自于公司内部各个部门、行业内各公司以及跨行业的各项 data 的物理性、逻辑性隔阂，据统计中小企业掌握着全球近半的数据信息，而非结构化、非标准化 data 的大幅度增加，也导致难以对隐藏在大数据中的有效信息进行统一的管理、高效的分析，并将其转化为指导决策和行动的有效 data 进而实现其市场价值。

在移动互联网、物联网时代，data 孤岛以及 data 碎片化的程度加剧，预计到 2020 年各种传感器、新兴物联网设备以及传统台式机、智能手机、服务机器人、平板电脑以及可穿戴设备，将到达 300 亿至 500 亿以上，data 分散，data 应用和数据分析受到严重阻碍，不同 data 源之间存在互不信任，从而导致数据分析结果偏差严重，data 的有效获取成本随着 data 断层的增加而递增，从而产生大量的 data 孤岛问题，因此区块链技术在这个领域具有十分巨大的商业潜力。



### 3.4.2 数据交易

在数据交易领域，数据交易所是一个不可获取的第三方中介平台，由于商业利益因素，企业内部数据难以在行业间进行流转，而第三方数据资产交易所作为平台作用，承担起数据资产流通的职能，数据资产交易是探索第三方数据变现的关键。

然而，对于传统的数据资产交易所而言，面临的最大问题是信息不对称、价格不透明、数据数量不足、数据安全以及隐私保护问题：由于价格以及数据的不透明性，数据的买卖双方对数据的质量经常会产生分歧；数据种类不全、数量不足导致交易效率低下；由于无法对数据资产进行确权，数据的易存储性以及易复制性导致非法使用与交易数据盛行；隐私数据在传统方式的数据交易中得不到保护等。

## 3.5 基础网络运营商

### 3.5.1 高昂使用成本

在当前的互联网产业，云计算、云存储、大数据等资本密集的服务产业，其市场仍处于快速发展阶段，对于大多数互联网企业而言，重资本的投入需要有在可见未来实现盈利的期望，对于每年十亿美元计的云存储行业而言，高昂的费用投入将最终转移至用户身上。当前云服务市场的高昂使用费用是行业难以解决的通病。

### 3.5.2 运营风险问题

云服务提供商方面，在 C 端用户大幅增加的情况下，中心化数据库的复杂性以及维护成本呈现指数上升的趋势，以 360 云盘为例，在 2016 年年中，其月活用户在 600 万以上，360 云盘为个人提供 1TB 以上的存储空间，若以每位用户使用 50GB 进行计算，在考虑到冗余备份等必要措施后，其数据的存储规模至少在数百 PB 量级。

数据存储的高昂成本以及对于带宽、服务器的极高要求，在对私提供云存储的盈利模式尚未清晰的情况下，其资本投入以及运营风险是远高于投资回报的，因此容易催生出云服务提供商的运营风险问题，并导致用户权益受损以及用户数据安全性问题。如金山快盘、新浪微盘退出云存储市场的决定对用户产生的负面影响。



## 4、ADD 数据生态的区块链解决逻辑

### 4.1 区块链赋能数据行业

在过去的 5 年里，大数据应用与区块链技术持续发展，借助分布式账本技术、加密算法、共识机制的结合，区块链技术为数据行业带来了大规模协作的基础，使得在去中心化的网络中，点对点的信息与数据交互、数据价值流通、更大量级的数据分析模型搭建成为了可能。

ADD 计划应用的被称作是区块链 3.0 的 EOS 底层平台，EOS 是计划支持百万级别高 TPS 的新一代主链，但其本质是通过去信任实现在非信环境下的有效交易或是价值流通，所谓价值流通即是通过 ADD 数字代币、数据市场、脱敏信息的有效数据互换，因此在数据行业的区块链转型中有着相对的应用优势，通过区块链技术可以在数据行业的多个环节带来升级与变革：

- 区块链的可追溯性使得数据从采集、清洗、分析到交易的每一步骤，皆可通过哈希值保留在 ADD 数据侧链上，从而确保数据的真实性及不可篡改性；
- 区块链的加密技术以及零知识证明可有效应用于数据交易上，脱敏数据购买者获得的仅是数据的使用权而非所有权，并可通过账本记录追踪数据在 ADD 数据生态及区块链网络中的流向；
- 区块链的加密技术同样可以运用在数据确权方面，用户可以通过 ADD 数据应用平台自主选择数据的开放程度以及授权范围，并通过加密 ADD 数字代币量化其价值；
- ADD 智能合约的使用可进一步规范数据的使用以及授权范围，在符合 ADD/EOS 智能合约规定的情况下自动触发 ADD 数据交易合约，在 ADD 数据市场内外流通得到更加有序的管理；
- ADD 基于 EOS 的数据侧链去中心化存储的功能使得原有孤立的互联网中数据存储空间得到提升与联通及价值体现，降低数据存储及计算成本，降低商业与数据行业的使用成本。



## 4.2 数据确权与保护

ADD 区块链的数据确权方案运用在多种领域，如：个体及商业数据确权、IP 确权、文字与著作、内容分发网络、数字营销等，其共通的特点是通过建立在区块链上的 ADD 数字凭证，通过多节点确认、验证及记录，从而赋予了该数据权利不可篡改与时间戳的特性。

基于区块链的数字凭证最早是由存在性证明（Proof of Existence）所提出，是指把数据白皮书的 SHA256 信息摘要嵌入比特币区块链来证明其存在性，存储的是文件的加密摘要（即哈希值）并连接提交文件的时间。通过这种数字凭证的方式，用户可以在不泄露文件信息或是个人数据隐私的情况下，公开证明自己拥有这些文件。

在 ADD 数据生态的确权与保护中，核心应用场景是在商业数据管理中，通过区块链技术实现商业数据的数字化管理、简化确权与时间认证过程，有助于更为清晰地披露时间戳情况。数据确权应用与保护是数据行业管理的一个重要基础也是区块链的重要应用领域之一，通过对所有人的数字身份、识别创建唯一 ADD 生态身份 ID，从而使得数据用户可以有效在多种应用、设备、平台甚至线下服务中安全、高效的接入及分析。其衍生意义在于可通过区块链技术实现多种应用、设备、平台系统中的快速切换，将有可能形成一种协作的、透明的身份管理方案，并通过平台实现对于数据的有效确权。

## 4.3 数据流转与交易

数据交易、流转与脱敏共享是数据行业在互联网与区块链相交的重要发展方向，而 ADD 数据市场链上的数据交易将是数据源端重要的变现方式，而数据价值不仅体现在数据量上，更体现在流通融合时在 ADD 生态上产生的更大价值上；数据同时也是 ADD 后期进行人工智能、机器学习快速发展的重要基础，人工智能算法通过对于海量大数据的学习，才得以不断优化其工作效率。

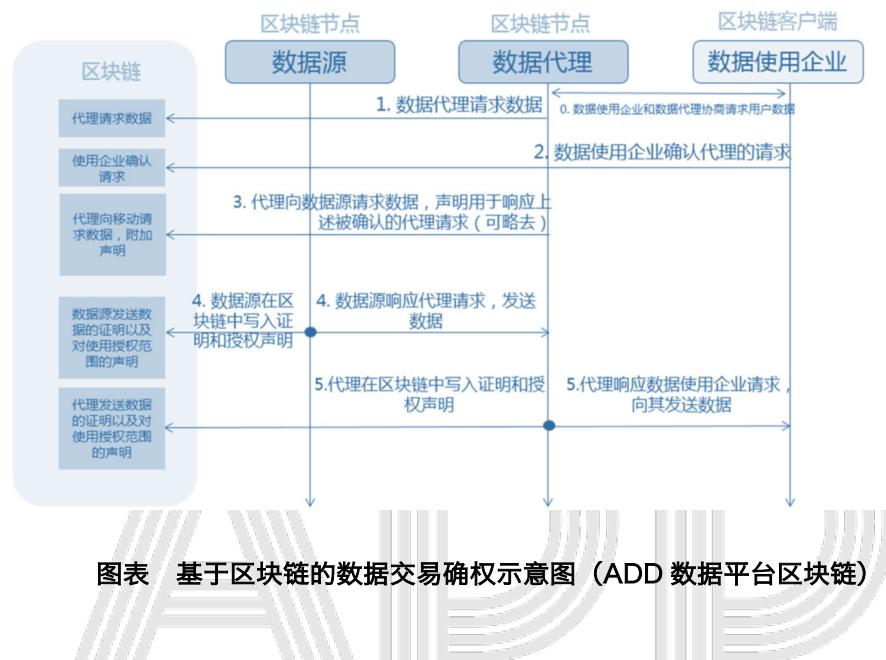
而对于数据搜集能力有限的企业而言，在 ADD 进行数据交易是一个快速而有效的方式，可以促进互联网公司创新并创造以数据作为新的收入来源之一。数据的交易有赖于 ADD 对于原始数据的有效确权，在确权完成后需要一个价值载体，实现数据的注册、交易和溯源，目前由于数据交易市场中普遍存在的数据非法倒卖，信息透明度低、数据篡改、数据质量参差不齐等情况，ADD 的数据市场将使得以上问题得到解决。

基于区块链技术的去中心化、安全性和不可篡改可追溯性，ADD 平台能够在参与主体之间建立信任，推进数据交易的可持续大幅增长，通过搭建基于区块链的去信任数据交易所，能够将数据的所有权、交易事项和授权范围记录在区块链上，数据所有权可以在交易时得到进一步确认，



精细化的授权范围可以规范数据的使用。同时，数据从采集到分发的每一步都可以记录在 ADD 数据侧链上，使得数据源可追溯，进而对数据源进行约束、加强数据质量。

基于 ADD 区块链的去中心化数据中枢，可以通过对于买卖双方的数据指纹进行多方确认，在世界范围内的不同平台上得到有效的数据确权，从而形成更大规模的全球化数据交易市场。



#### 4.4 去中心化云存储

在万物互联时代，移动设备、传感器的大量出现，使得硬件资源相比于数量庞大的数据逐渐稀缺，对于如何更有效地利用余裕的计算资源以及存储资源成为越来越迫切面对的问题，并催生出庞大的市场价值，中心化的云存储市场已在商业上证明具备十分广阔的市场空间，以 IaaS、PaaS 为主的云存储细分市场快速增长，并逐步在该领域内建立起高资金、高技术、高用户数量的竞争壁垒，逐渐呈现中心化的趋势。

在云存储、云服务领域内，是一个天然具备中心化倾向的市场，以亚马逊、阿里云为首的国际及国内云服务巨头分别在全球及中国市场占据近半壁江山，以其领先的技术优势、资金优势、用户优势逐渐拉开与其他竞争对手的差距，随着企业以及用户对于云存储的依赖性逐步加深，这种趋势将导致云服务的提供商具有愈来愈强的议价权。而中心化云存储的高昂的使用费、开发成本、隐私泄露、数据所有权及控制权、运营风险等问题，皆为中心化云存储的进一步发展带来了负面影响。



去中心化云存储的本质是一种资源层面的 IaaS 项目，在传统云存储的基础上免除了 IaaS 的巨大硬件成本投入，通过对现有用户余裕的硬盘空间的有效利用，从而降低了用户的使用成本；通过将数据分片并冗余存储于分布式网络中的不同节点，从而提高了文件的传输速度；通过加密算法对于文件进行非对称加密，提高了数据以及网络的安全性。

去中心化云存储是一种通过区块链技术，提供点对点加密的文件存储功能，其原理是通过将数据切割成固定大小的数据分片并进行加密，通过在区块链的分布式网络上租赁存储空间，实现数据在分布式网络的存储，并通过类似存储证明（Proof of Storage）的机制，在非固定的时间点对于存储主机的文件进行存储验证，提供存储空间的主机从而能够获得收益。

而对于去中心化存储项目而言，如何保证数据的可用性、可靠性以及性能是核心问题：

- 首先，基于分布式网络的存储需要保证数据的冗余备份足够充足，以避免在多个节点同时失效、存储硬盘损坏、存储上不可读取的情况下产生的存储数据丢失问题，并且在丢失的情况下，仍然能够保证数据是可恢复的。
- 其次，在存储的过程中，如何确保存储主机不会为了获得不正当利益，伪造数据的存储证明也是必须要衡量的重点，由于分布式网络中普遍存在的女巫攻击、数据生成攻击等问题，对于存储数据的有效验证，需要反映在制度的设计逻辑上。
- 其三，由于去中心化的数据存储项目是无需许可的，因此对于多数并不符合数据中心要求的用户的接入，必然导致网络的整体效率降低并主要反映在带宽、存取速度等指标上。

目前而言对这些问题，ADD 生态将逐步产出的解决方案。

## 4.5 区块链+数据应用的发展

李彦宏曾提出人工智能（AI）、大数据（Big Data）和云（Cloud）是“三位一体”的概念，随着计算终端的下沉、海量数据的广泛分布以及人工智能的发展，ABC 时代正在到来。而 ADD 认为数据的去中心化（Decentralized）将成为未来数据行业第四个决定因素，“ABCD”时代必将到来。数据行业将逐步完成区块链升级，从上链、存储、分析关键环节到数据的交易确权、交易、流转，用户与数据之间、数据节点与节点之间的交互呈现越来越紧密的趋势。

随着互联网的发展从中心化、平台化、生态化逐渐朝向分布式化发展，海量大数据越来越多的通过移动终端、传感器进行处理，数据生产、流通及存储的分散化导致了分布式网络的地位迅速上升，CDN、雾计算、边缘计算的概念被提出并逐步运用于商业领域，ADD 认为数据中心节点降权的趋势将是不可逆的。



区块链网络正如互联网离不开数据的存储及计算，互联网的底层架构决定了其上层建筑，上层建筑依托于底层技术的构建，随着当前互联网的分布式信息网络的轮廓逐渐明晰，我们对于分布式网络的要求也越来越多，区块链的特性正好能够因应越来越分散化的互联网需求，而区块链上层的通信协议以及 DApps 的数据计算、存储、流通，需要新的、更加符合分布式网络优势的底层架构，对于数据的处理、分析、交易也同样能够以区块链网络的形式解决。

尽管当前互联网中存在着许多中心化的云服务、数据服务架构，然而，通过它们来构建 DApps 或是分布式协议将会破坏区块链技术的固有优势。以分布式存储为例，我们可以采用不同的分布式方法存储数据。例如，完全可以将数据存储在区块链上，然而这是十分昂贵的：每个数据节点必须存储所有数据、每个节点必须广播给所有节点，高延迟以及低扩展性对于这样的底层架构将是不可接受的。而对于最近提出的许多解决方案而言，Plasma 以及 Lightening Network 允许离线存储事务，并且周期性地将它们与主链进行同步，然而在某些情况下，将数据传输到区块链以进行查询可能需要一些非区块链技术出身的应用者很长时间。

因此，简单地应用区块链技术并不能解决所有问题，这同样也发生在数据行业中，区块链+数据行业目前仍然处于萌芽期，ADD 认为当前互联网的区块链转型需要一个新的底层架构、新的通信协议，而由 ADD 打造的 EOS 侧链将在未来能够提升的数据行业运作效率，同时 ADD 区块链 3.0 EOS 技术的应用与 ADD 代币也将带来数据协作的信息红利，基于区块链技术的数据确权、隐私保护、安全性是其他技术无法替代的，具备十分庞大的潜在应用价值。

处于数据行业前沿的项目数量，自 2016 年以来仍在快速的增加中，大量的应用项目覆盖了数据行业的各个环节：数据存储、通信协议、数据库、数据+物联网、数据交易、数据搜集等。ADD 认为 ADD 的 EOS 数据侧链作为区块链世界的重要底层技术，将赋能以上数据行业的各个环节。



## 5、 ADD 数据生态架构与功能计划

在未来几年，有效数据将日益成为一种流动商品或数字资产。数据大量的存在互联网业、传统产业、服务行业和整个供应链中，传感器和连接设备产生了时间戳数据，这是现代经济的基础，其中大部分数据都是以流媒体方式生成的。

随着区块链的发展和无处不在的应用，数据量级将呈指数增长。而数据应用从互联网业、传统行业转移或在区块链上进行同步将是一个重要趋势。

IHS Markit 预测，在全球物联网市场，已安装的基础将为从 2015 年的 154 亿部增长到 2020 年的 307 亿部，到 2025 年的 754 亿部。而 ADD 预计到 2023 年，其中 25%-30% 甚至更多设备将由区块链底层技术提供数据与交互服务。许多新生成的数据是有价值的：它可以用来优化市场操作，以更高的精确度跟踪资产，以高粒度的现有主数据服务为目标，并创建全新的服务和业务模型。

ADD 团队在项目启动前，曾多次和世界顶级的数据服务商、国家级数据源、大规模数据应用使用者、区块链投资人及来自牛津大学、剑桥大学的大数据学者进行过已有数据上链必要性访谈与“商业与公共数据上链是否是伪需求？”的探讨。

ADD 在本白皮书分享的结论是：已有商业与公共数据上链是各领域区块链转型中，保证已有数据未来在区块链上流转、确权、比对、应用、更新的基石和必要步骤。

在 ADD 的访谈中，来自中国三大通信服务商之一的数据中心负责人就明确表示：“数据与区块链应用方向的结合，运营服务与区块链应用的结合已经写到了中国 XX 总部领导(根据被访谈人需求隐去企业和领导名)未来 3 年规划和 5 年战略中去了。营业所和坐席现在都是用已有数据库比对和更新，这个数据库不迁移到链上，领导的未来实现业务的区块链转型等等就是空谈，因为一线的同事用的还是原有数据库。现在区块链方向定了和预算也批了，我们在做的就是规划一个区块链运营商解决方案，但重中之重的第一步就是完成现有数据上链。”



## ADD——基于区块链 3.0 EOS 的大数据中枢与数据分析、数据应用平台

通过 ADD 区块链应用平台,企业积累的各种来自内部和外部的数据, 比如网站数据、销售数据、ERP 数据、财务数据、大数据、社会化数据、各种数据库等,都可进行整合、探索、挖掘、分享和控制, 实现数据价值到商业与数据个体价值的完美蜕变。ADD 区块链数据应用平台简单易用, 无技术壁垒。

用户可迅速使用 ADD 平台进行数据上链, 或定制化进行大规模互联网数据的区块链上链迁移, 轻松构建企业基于区块链数据的价值挖掘体系。ADD 区块链应用平台拥有全面、丰富的数据可视化效果库和数据源连接库, 将开发多种数据挖掘算法、行业模型和海量数据的实时计算。ADD 区块链应用平台为企业提供从数据清洗处理、数据银行、数据分析挖掘到数据可视化展示的全套解决方案, 同时针对企业的特定需求, 提供定制化的大数据解决方案, 从而推动企业实现数据区块链化与智能化管理。

ADD 区块链应用平台行业覆盖广泛, 将通过数据合作源和大数据合作伙伴, 应用于电商、制造业、政府、金融、医疗、银行、保险、电信、教育、互联网企业等多个行业。

ADD 计划打造基础平台的是一个基于 EOS 的数据侧链, 在安全性的前提下, ADD 打造的 EOS 数据侧链提供与主网无缝衔接, 使用智能合约和 ADD 数字代币, 使侧链数据完全存储于主链, 同时增加主链与侧链的数据可信性, 同时使用智能合约, 实现侧链与主链的交互和数据共享, 使 DAPP 开发者可以同时获取侧链与主链的双用户体系, 减少 DAPP 的开发成本, 让整个 EOS 社区更加繁荣。

我们相信, 区块链的持续增长将通过适当数据上链存储和未来链上数据操作得到促进, 好比经历“互联网转型”的历史轨迹加速后, 区块链转型的轨迹亦应得到加速。比如通过 ADD 数据生态简易工具, 非专家或开发者可以便捷的将数据进行上链与存储的, 同时 ADD 数据银行和 DApps 也连接到可靠的数据来源, 提供应用型的数据分析解读。



在 ADD 数据生态层面，ADD 将在 ADD 数据侧链建立后搭建如下模块：

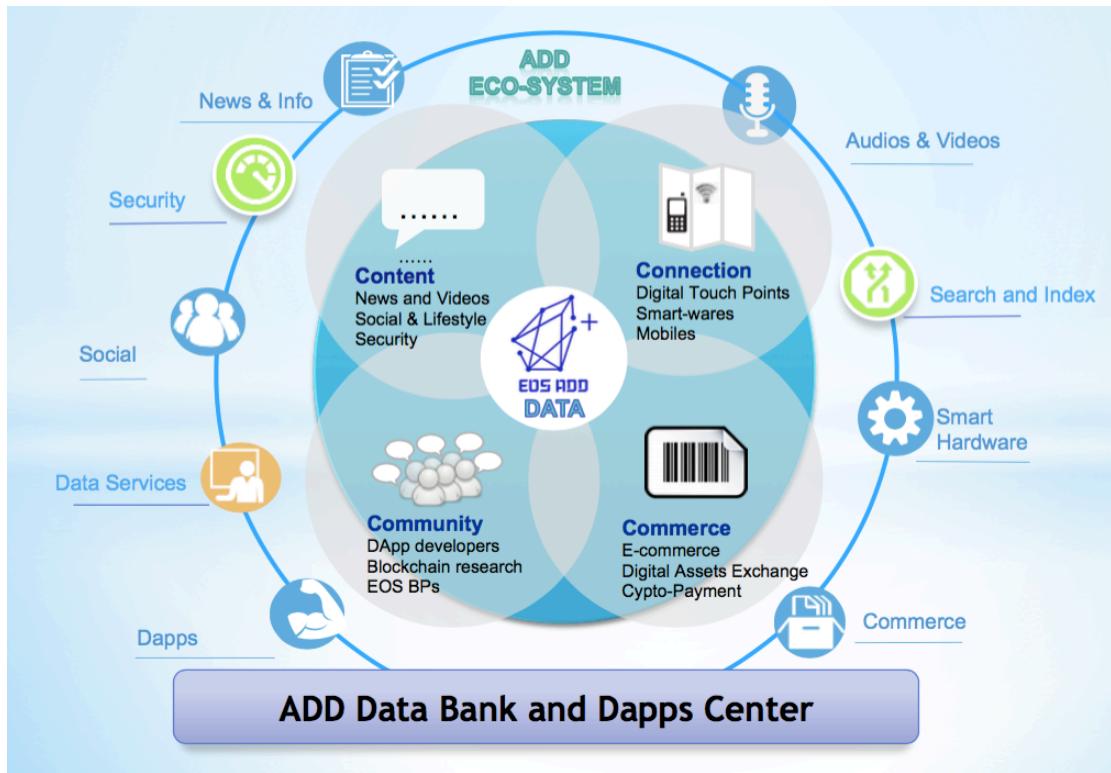
ADD 数据中心：一个全域、跨链的共享数据流，每个人与数据源的 4C（内容、链接、社群、商业交易）相关数据都可以贡献和脱敏流转并以更小颗粒度获取及筛选信息

ADD 数据模型、引索与数据引擎是一个高性能的数据处理和分析引擎以分布式的方式执行存储，它将提供可视化、便捷的自设更细颗粒度的分析、提供专项维度与行业报告与评级体系。

ADD 数据应用平台构成一个可用性层和工具箱使数据上链与大数据区块链转型快速发展，数据驱动的应用程序平台

ADD 生态架构以链接不同来源的元数据 (Raw Data) 为核心，搭建数据银行和数据应用中心，围绕 4C 机构，打造数据模型服务多种业态，同时促进数据安全传输与交易。ADD 数字代币将在数据市场与 ADD 生态进行流转。

以上模块均基于 EOS 的 ADD 数据侧链



## 5.1 ADD 基于 EOS 的数据侧链

ADD 数据侧链按照不同需求以实现以下方面功能：

- 去除重复账户注册，使用公钥与用户名绑定，与 EOS 主网账户名一致，未来打通与主网信息交互
- 删除或调整 RAM, CPU, NET 等 EOS 主网需求功能与设置，提供更纯粹的数据存储、数据应用场景和交互体验
- 在 EOS 主链的安全性之上增加 ADD 独有的数据安全模块，为每一位用户提供数字及账户资产安全性提醒和预防

部署智能合约，完成和打造 EOS 主网和 ADD 侧链通信交互，ADD 侧链的所有数据都在主网智能合约中展现，合约对侧链每一笔交易都在主网同时进行，同时预留接口，回调主网 ADD 数据到侧链。

同时使用智能合约，实现侧链与主链的交互和数据共享，使 DAPP 开发者可以同时获取侧链与主链的双用户体系及并将数据低成本的在 ADD 侧链上实现读取，减少 DAPP 的开发与数据存储、交互成本，让整个 EOS 生态更加繁荣。

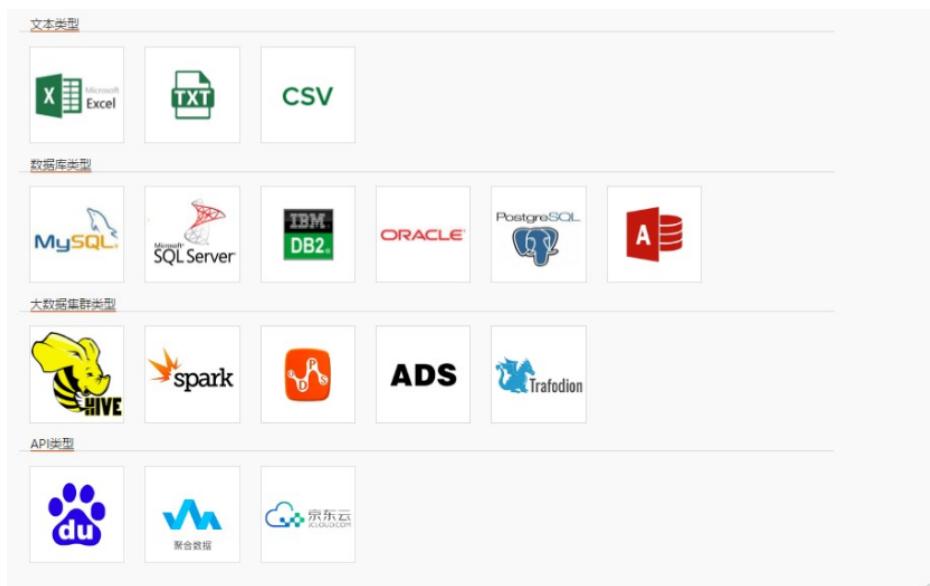


## 5.2 ADD 主要功能模块规划：

ADD 数据中枢应用平台拥有包括数据中枢 Dashboard、多图表类型支持、数据源上链一键式接入支持、数据安全、权限管理，数据共享、数据源接入库、数据可视化解决方案、数据机器学习与挖掘模型、智能分析、动态图表、跨表分析、海量大数据处理、数据银行、增值和定制化服务等计划功能，充分满足 B 端及 C 端数据相关需求。

### 5.2.1 ADD 数据上链功能与迁移支持

ADD 数据中枢与数据应用平台将支持多种常见数据库的连接与上链服务，包括 mysql、SqlServer、Oracle、excel 等数据源，用户只需接入数据源进行数据配置和管理，就可以按照需求进行加密上链，同时还将提供对常用文本文件能提供一键上传实时上链服务。



ADD 团队也将提供定制化企业级大规模数据上链解决方案，协助已有数据的上链与加密。

### 5.2.2 数据银行功能

数据上链完成后，ADD 区块链数据应用平台通过数据获取，数据清洗，数据整合的技术，针对企业不同需求，为企业建立数据银行，包括传统数据库，ADD 侧链数据库，新一代动态数据库，和打造基于区块链的数据银行。



### 5.2.3 数据模型、分析引擎与数据可视化分析功能

ADD 把对业务数据的分析转换成从数据关系图中查找符合条件最优路径的过程，系统根据权重自动排序，使得每个查询都使用最少的资源却以最高的效率计算。

ADD 生态将内置聚类分析、挖掘预测等高端数据挖掘功能模型，同时整合了多种数据挖掘功能并根据行业和用户需求持续更新。

ADD 数据模型支持对海量数据的分析和挖掘，处理数据目标达到 PB 级别。ADD 数据中枢与生态在 EOS 侧链上实现高 TPS 存贮与读取，通过 ADD 数据模型实现高效海量数据挖掘，基于 ADD 生态搭建高效数据挖掘框架，利用数据库来模拟链表结构，提供树形结构、图模型的分布式计算方法。

同时用户使用 ADD 数据中枢与生态，直接通过简单的拖拽就一步生成分析模型，比如精准营销、用户分析、用户画像等。

### 5.2.4 ADD 机器学习功能

ADD 机器学习系统将根据给定的训练样本对分析系统输入输出之间的依赖关系进行估计，能够对未知输出做出尽可能准确地预测。通过使用 ADD，用户则根据现有的数据状况进行自动建中间表，学习和优化用户的分析性能。

### 5.2.5 数据可视化功能

ADD 平台将研发并打造行业领先的可视化效果库，拥有多种分析图表类型，给使用者带来全新的视觉享受。其中包含多种常规图表和多种大数据可视化图表效果及定制化图表这些非传统交互式可视化效果。常规分析图表、大数据可视化效果、动态高级效果、3D 效果以及定制化具象效果，同时附有现实意义和分析内涵。让分析从困难枯燥走向简易。用户可以在 ADD 数据中枢与平台的多种可视化效果库里获取任意需要的可视化效果。另外，ADD 也将打造领先的大屏幕可视化解决方案，完美支持三屏至多屏合一。



### 5.2.6 数据安全与权限功能设计

ADD 数据中枢与生态在对用户的数据传输上链进行安全支持，用户可同步通过 ADD 选择进行个性化加密、解密，私有链上数据在缺失用户自己持有的 ADD 数据密钥时可视为加密后代码，可视但无法提取有效数据。

用户可以使用 ADD 数据中枢与生态一键完成安全授权，保证数据使用安全。

通过 ADD 数据中枢与生态的权限管理，用户可以瞬间搭建企业数据价值挖掘体系，增强团队协作能力。用户管理员可分配不同二级用户角色与权限，通过权限管理功能，所进行的分析权限和数据权限都可自行进行配置，实现数据化管理体系。

### 5.2.7 数据分享功能

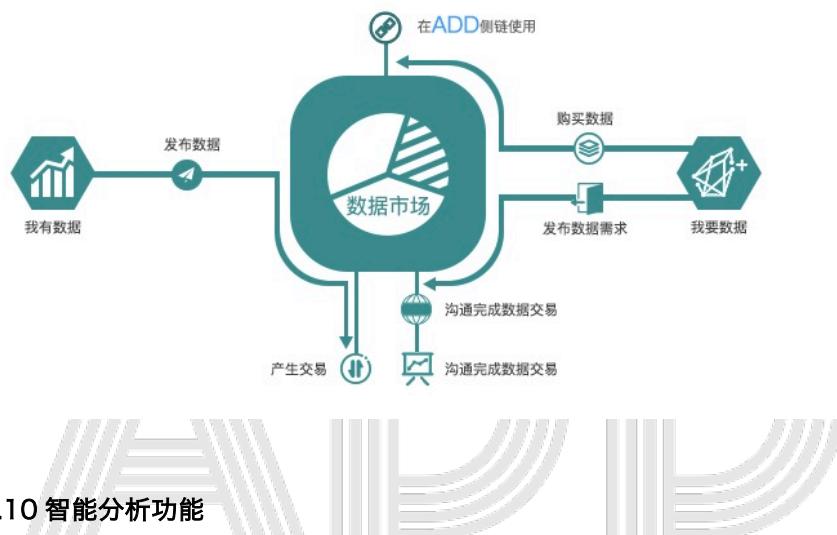
用户使用 ADD 数据中枢与生态，将生成的数据分享链接实现分享，分享至用户分享对象，配合数据权限功能，轻松实现实时化的数据保密、公开和便捷查看。



### 5.2.8 增值和定制化模块

针对不同企业和用户的需求, ADD 区块链数据应用平台将提供多个增值和定制化模块, 包括机器学习, 自动建模, 可定制化图表支持, 跨数据库、数据源支持, 行业数据分析(项目), 可定制化分析组合, 定制化分析挖掘模型和解决方案等。

### 5.2.9 ADD 数据市场架构



### 5.2.10 智能分析功能

ADD 区块链生态平台将内置聚类分析、数据清洗、相关性分析、行业与数据解决方案等高端数据挖掘功能模型, 同时整合了多种数据挖掘功能并根据行业和用户需求持续更新





### 5.2.11 海量大数据批量处理功能

ADD 数据中枢与生态支持对海量数据的分析和挖掘, 处理数据达到 PB 级别。用户在 ADD 生态里实现高效海量数据挖掘, 基于 EOS 搭建高效数据挖掘框架, 利用数据库来模拟链表结构, 管理挖掘出来的知识, 提供树形结构、图模型的分布式计算方法。



## 5.3 ADD 动态数据工作模型架构图





## 6、 ADD 项目关键性内容披露

### 6.1 项目团队

Lucas Shang - 联合创始人

香港公开大学毕业、商科学士，资深数据及区块链专家，他有着丰富的数据与区块链经验。曾任职于香港新鸿基金融、三星电子通讯、韩讯、星云比特等负责数据模型、量化交易、区块链技术、矿池技术，并先后负责了 Zcash、EOSBeijing 超级节点和 EOS ADD 超级节点的技术搭建与 EOS 启动前主网安全扫描。现全面负责 ADD 项目推进与实施。

他是 EOS 主网中国区创世节点群成员、EOS 主网中国安全启动团队成员。

Beverly Han - 联合创始人

华中科技大学毕业，广告学硕士，资深媒体人、连续创业者，曾在国内顶级经济类媒体、海外华文留学生在线社区、网络营销等领域就职和连续创业。2017 年作为联合创始人创立区块链服务平台，为区块链项目提供咨询与顾问服务，同时提供资源支持，后期负责区间资本的管理工作。现主要负责 ADD 的宣发、媒体及商务合作。

Cici Liao ADD-创始合伙人

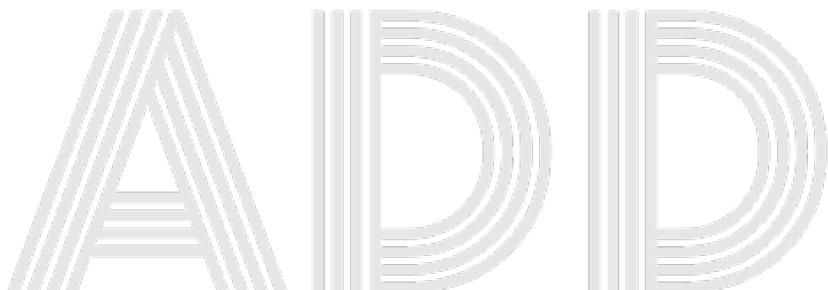
英国帝国理工大学博士，曾先后任职于英国 HSBC 银行、Barclays 欧洲投行，负责科技板块投资与并购重组业务；有超过 15 年的科技资本运作经验，并长期志愿担任英国科技协会高级顾问。现负责基金会管理与海外投融资操作，ADD 项目资金与财务管理工作，兼 CFO。

### 6.2 项目顾问



### Charley Dong - 天使投资人、顾问

英国皇家市场学院院士 (FCIM), 天使投资人、互联网营销与数字化转型专家, 曾任职于 Yum!、Coca-Cola 等国际顶级品牌, 在数字商业模型与互联网营销领域有着超过 15 年的实战及顾问经验。他曾见证并参与过多家海内外主流数字媒体与数据平台的第一次商业化转型进程, 从 2017 年 7 月开始涉足 EOS 领域投资。他同时也是 EOS 生态孵化基金的顾问, 多个 EOS 超级节点与生态项目顾问, 他深度参与 EOS 生态建设与孵化、投资, 曾出席 2018 年 6 月 9 日 EOS 主网是否启动的全球节点关键性投票, 他也是 EOS 中国创世纪群的成员、ECAF 创始成员。





## 6.3 代币分配与流通

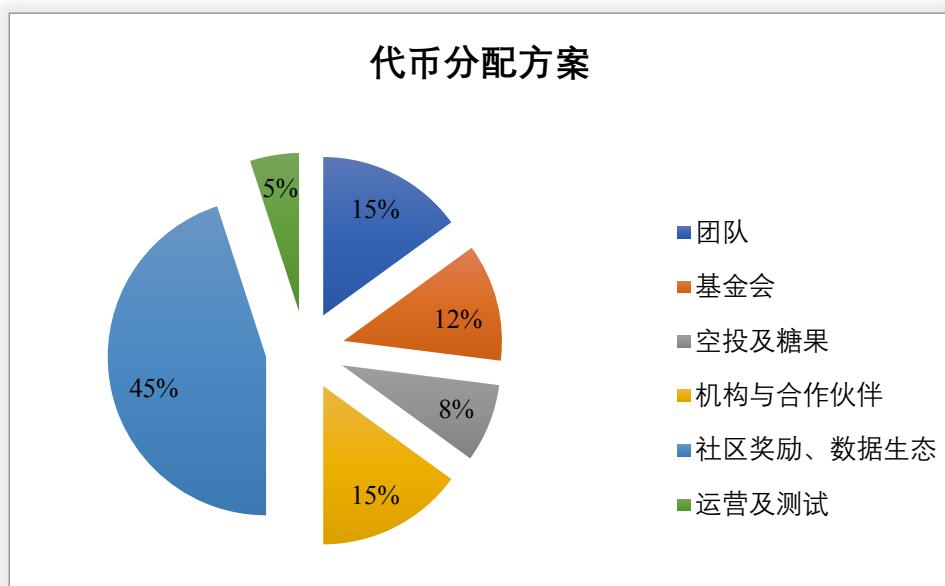
### 6.3.1 代币发行与分配、锁仓计划

ADD 代币总发行量为 100 亿，永不增发，

代币分配与锁仓计划如下：

团队 15% (锁仓 5 年，逐年定期线性释放，锁仓公开地址：eosaddlockup)  
基金会 12% (锁仓 4 年，逐年定期线性释放，锁仓公开地址：eosaddlockup)  
全网空投及糖果 8%  
机构与合作伙伴 15% (首次发放 7.5%，剩余锁仓 3-6 个月)  
生态社群奖励、数据生态共建、数据源 45% (按照数据挖矿与相关奖励机制)  
运营与技术测试 5%

ADD 项目首次空投完成后 45 天内（首次空投时间定于 2018 年 6 月 22 日），实际市场流通量为总发行额的 15.6%；项目运营及 ADD 首次空投满 1 年时（即 2019 年 6 月 22 日），预计总流通额为代币总额的 20-30%





### 6.3.2 ADD 数字合约、数字代币 [ADD]与空投信息

#### a. 合约

合约应用 EOS 主网合约

合约地址: eosaddaddadd

合约精度: 0.0000

#### b. 数字代币及空投计划:

数字代币名称: ADD

空投时间: 英国伦敦当地时间 6 月 22 日零时起 (夏令时 UTC+1)

空投比例: 2 EOS : 1 ADD

空投条件: 对创世快照时账户 EOS 数量大于等于 100 个的地址

ADD 空投说明:

c. ADD 使用了 2018 年 6 月 1 日 7 时的 EOS 主网启动创世快照进行空投。

d. 在 6 月 1 日 7 时前完成映射的账户且余额大于 100 EOS 的用户都可以收到空投，包括交易所账户。

e. 未映射的用户无法收到空投。

f. 创始快照地址 :

<https://raw.githubusercontent.com/meet-one/genesis-snapshot/master/snapshot.csv>

### 6.3.3 代币权利

在本白皮书书写时，ADD 数字代币的权利被设定为使用 ADD 数据中枢系统或数据应用服务、数据分析或 ADD 数据市场流通的数字代币。项目方保留更改或增加本项权利的权利，如有相关更改，项目方将提前通过公开渠道公布。



## 6.4 ADD 项目所有人及 EOS 超级节点(eosaddaddadd)

ADD 项目概念、商业模式、品牌权利、数据归属于注册在英格兰和威尔士的 EOS 基金会 (EOS FOUNDATION) 基金会系独立运营的非盈利组织, 请注意本基金会同 BLOCK.ONE 公司无任何直接关联。

EOS FOUNDATION 将以上 ADD 项目权利授予给位于开曼群岛的豁免公司 ADD BLOCKCHAIN 以 ADD 名义进行项目的技术开发与相关项目运营, ADD BLOCKCHAIN 承担项目全权责任与义务。

超级节点候选人 eosADD (eosaddaddadd) 系 ADD 项目的基石与项目组成部分之一, 亦为实现 ADD 生态愿景专门设立的 EOS 备选超级节点。

## 7. 名词定义、免责声明与风险提示

### ADD 定义 (本白皮书、本条款中称 ADD 或本项目)

ADD 定义的包含但不限于以下: ADD 项目、ADD 项目方主体及为其直接或间接服务的自然人、ADD 项目计算能力与数据中心、ADD 生态的开发者与贡献者、受项目方委派及认可的兼职社群运营参与者、ADD 投资人与投资机构、ADD 员工及受认可的志愿者、EOSADD 超级节点、英国 EOS Foundation 及有利益或股份关联的公司或生意伙伴。

**[ADD] 定义 (在本白皮书、本条款中称[ADD]或 ADD 数字代币、ADD 权证)** 本条中的[ADD]代表由 ADD 发行的数字代币符号为 ADD 的 EOS 主网代币, 代币由合约地址: eosaddaddadd (合约精度: 0.0000) 发放, 并于 2018 年 6 月 22 日通过 EOS 主网按比例空投给到创世快照中持仓超过 100EOS 的 EOS 账户的 EOS 主网 ADD 数字代币。

### 7.1 免责声明

7.1.1 本项目生态参与者、代币持有者、数据源提供者、数据市场与数据银行使用者 (以下简称参与者) 采取自愿参加、风险自担责任、接受和遵守活动发布内容中的规则事项。一切因参与者直接或间接引起的法律责任由参与者自行承担。

7.1.2 本项目参与者已确认自己有充分的身体、心理和物质上的准备而参加, 对项目中的一切风险及导致的各种后果均可自我承担, 并承诺在项目中发生的一切有关自己人身、财



产和精神的损失均不会向 ADD 或 ADD 定义中未涵盖的本项目关联自然人、机构或商业组织追究法律上及任何形式上的责任。

7.1.3 本项目组织者在白皮书及其相关内容宣发中的一切说明和安排均为构想，不排除在项目开始后，因受人为或自然不可抗力因素影响导致对本项目原计划的变更和取消。ADD 仅保证在取消项目活动或变更相关内容前，尽可能通知到报名参与者（通过自有社交媒体平台如 twitter\github\telegram, 发布将视为有效通知），并说明原因，但不负责承担由此给项目参与者、[ADD]持有人造成的人身、财产和精神上的损失。

7.1.4 任何由于黑客攻击、计算机病毒侵入或发作、因政府管制等造成的暂时性关闭或影响网络正常经营的不可抗力而造成的个人资料泄露、丢失、被盗用或被篡改等，ADD 均不负任何责任。

7.1.5 ADD 如因系统维护或升级而需暂停服务时，将会事先公告。若因线路及非 ADD 可控制范围外的硬件故障或其他不可抗力而导致暂停服务，与暂停服务期间造成的一切不便与损失，ADD 不承担任何责任。

7.1.6 凡以任何方式直接、间接使用 ADD 项目或持有[ADD] 权证的个人、机构及无法用个人或机构的定义可以涵盖的 Subjects，均视为自愿认可、接受并遵守本声明的约束，同时自愿永久放弃对 ADD 发起法律诉讼、提起任何形式的仲裁之权利。

7.1.7 此章所包含的信息为风险提示，请相关意向爱好者仔细阅读。该白皮书 只用于传达信息之用途，并不构成买卖 ADD 项目方、所有者、投资人及投资机构、EOS 基金会及关联公司、商业伙伴股份或证券的相关意见。任何类似的提议或征价将在一个可信任的条款下并在可应用的证券法和其它相关法律允许下进行，以上信息或分析不构成投资决策，或具体建议。

7.1.8 本白皮书不构成任何关于证券形式的投资建议，投资意向或教唆投资。

7.1.9 本白皮书不组成也不理解为提供任何买卖行为，或任何邀请买卖任何形式证券的行为，也不是任何形式上的合约或者承诺。

7.1.10 ADD 明确表示相关意向用户明确了解 ADD 项目的风 风险，投资者一旦参与投资即表示了解并接受该项目风险，并愿意个人为此承担 一切相应结果或后果。

7.1.11 ADD 明确表示不承担任何参与 ADD 项目造成的直接 或间接的损失，包括：本白皮书提供所有信息的可靠性；由此产生的任何错误， 疏忽或者不准确信息；或由此导致的任何行为。



7.1.12 在写这段文字时，「ADD」不能用来购买相关物品或者服务；ADD 项目尚未有实质上商业应用。

7.1.13 ADD 项目不是一种投资。我们无法保证本项目及「ADD」将会增值，但其也有在某种情况下出现价值下降的可能，那些没有正确地使用「ADD」的人有可能失去使用 ADD 的权利，甚至会有可能失去他们的「ADD」。

7.1.14 「ADD」不是一种所有权或控制权。控制「ADD」并不代表对 ADD 应用的所有权，ADD 并不授予任何个人或机构任何参与、控制、或任何关于 ADD 应用决策与商业方向方向及本项目其他事项决策的任何权利。

本声明以及其修改权、更新权及最终解释权均属 ADD 所有。

## 7.2 风险提示

### 7.2.1 ADD 项目投资均属于风险投资

投资过程中可能存在市场风险、经营风险、信用风险、管理风险、政策风险及其他相关风险。

### 7.2.2 风险投资收益来自于项目成长和经营

ADD 不承诺任何固定回报，不承诺保本和最低收益，不做任何形式的担保。

### 7.2.3 私钥或账号丢失泄漏导致的丢失「ADD」的风险

「ADD」在分配给所有者时很可能关联至一个 EOS 体系的 EOS 账号，进入 EOS 账号的唯一方式就是所有者选择的相关登录凭证，遗失私钥或这些凭证将导致「ADD」的遗失。

最好的安全储存登录凭证的方式是所有者将凭证分开到一个或数个地方安全储存，且最好不要储存、暴露在工作的地方。

### 7.2.4 EOS 核心协议相关的风险

ADD 和「ADD」基于 EOS 协议开发，因此任何 EOS 核心协议发生的故障，不可预期的功能问题或遭受攻击都有可能导致 ADD 和 [ADD] 以难以意料的方式停止工作或功能缺失。此外，EOS 协议中账号的价值也有可能以跟「ADD」相同方式或其它方式出现价值上下降。



### 7.2.5 购买或持有者凭证相关的风险

任何第三方获得购买或持有者的登录凭证（如数币交易所登录密码）或私钥，即有可能直接控制购买者的「ADD」，为了最小化该项风险，购买或持有者必须保护其电子设备以防未认证的访问请求通过并访问设备内容。

### 7.2.6 司法监管相关的风险

区块链技术已经成为世界上各个主要国家的监管主要对象，如果监管主体插手或施加影响则 ADD 应用或 [ADD] 可能受到其影响，例如法令限制使用、销售、交易电子代币诸如「ADD」有可能受到限制，阻碍甚至直接终止 ADD 应用的发展。

### 7.2.7 ADD 应用缺少关注度的风险

ADD 应用存在没有被大量个人或组织使用的可能性，这意味着公众没有足够的兴趣去开发和发展这些相关分布式应用，这样一种缺少兴趣的现象可能对 ADD 或其相关应用造成负面影响。

### 7.2.8 ADD 相关应用或产品达不到 ADD 自身或购买或持有者的预期的风险

ADD 应用当前正处于开发阶段，在发布正式版之前包含白皮书均有可能会进行比较大的改动，任何 ADD 自身或购买者对 ADD 应用或 ADD 的功能或形式(包括参与者的行为)的期望或想象均有可能达不到预期，任何错误地分析或一个设计的改变等均有可能导致这种情况的发生。

### 7.2.9 黑客或盗窃的风险

黑客或其它组织或国家均有以任何方法试图打断 ADD 应用或 ADD 功能的可能性，包括服务攻击，恶意软件攻击或一致性攻击等。

### 7.2.10 漏洞风险或密码学科突飞猛进发展的风险

密码学的飞速发展或者科技的发展诸如量子计算机的发展，或将破解的风险带 给加密货币和 ADD 平台，这可能导致「ADD」的丢失。

### 7.2.11 EOS 节点攻击的风险



如其它去中心化密码学代币和加密代币一样，用于 ADD 合约应用的 EOS 区块链的节点或节点们也容易受到攻击，例如双花攻击、高算力比例攻击、过度竞争攻击，任何成功的攻击对 EOS 主链、ADD 应用，「ADD」来说一种风险，尽管 ADD 非常努力地提升系统的安全性，但以上所述的攻击风险是真实存在的。

#### 7.2.12 缺少维护或使用的风险

首先「ADD」不应该被当做一种投资，虽然「ADD」在一定的时间后可能会有一定的价值，但如果「ADD」缺少维护或使用的话，这种价值可能非常小。如果这种情况发生，那「ADD」后续的跟进持有者或少有跟进持有者，这显然会对「ADD」不利。

#### 7.2.13 ADD 存在的解散风险

存在这样的可能，出于各种原因，包括「ADD」自身价格的波动，ADD 应用发展遭遇问题，资金链关系、生意关系的破裂或知识产权等可能性原因，ADD 项目随时都有可能遭遇重大打击或直接解散。

#### 7.2.14 应用存在的故障风险

EOS 主链或 ADD 项目可能因各方面的原因故障，无法正常提供服务，严重时可能导致用户「ADD」的丢失。

#### 7.2.15 无法预料的其它风险。

EOS 主网代币或密码学代币是一种全新且未经全面测试的技术，除了本白皮书内提及的风险外，此外还存在着一些 ADD 团队尚未提及或尚未预料到的风险，其它风险也有可能突然出现，或者以多种已经提及的风险的组合的方式出现。



## 结语

白皮书概述了我们对 EOS 主链与打造区块链数据生态的信心，搭建 EOS 侧链及 ADD 本地数据银行与安全中心、ADD 生态体系的设想。我们相信横跨传统互联网的实时数据和区块链数据的结合管道将变革大数据行业和数字生态系统、数据引索及可视化系统，甚至整个互联网及区块链生态。我们的目标是建立一个经过深思熟虑和专业实施的数据生态与应用平台。

ADD 将是区块链数据转型革命的一部分，在这场区块链数据革命中，ADD 相信原有的中心化的云数据服务将被取代。在分布式计算中，Golem 取代了 Azure 虚拟机，而 IPFS 取代了 Azure Blob 存储。ADD 将率先应用 EOS 提供了数据消息传递和事件处理的分布解决方案，ADD 为加入这场革命而自豪。

ADD 生态架构服务于区块链及互联网的未来需求，并为 EOS 及区块链的应用程序提供有效的数据及系统化支持。ADD 的技术生态架构将是分层的、模块化的，并且构建在高效能的 EOS 之上（未来也许有更优的性能主链，ADD 生态中的数据亦可跨链集合）。由一个由可被激励数据节点逻辑组成的点对点网络矩阵，网络托管发布/订阅机制，并支持加密事件的分布存储。吞吐量与参与节点的数量成线性关系（理想状态下 EOS 网络每秒可以处理数百万个事件）。

### ADD 团队对 ADD 生态远景的目标：

ADD 生态未来将拥有自建的 ADD 数据侧链（基于 EOS 体系），计划通过数据节点与数据市场贡献并整合的全球最大数据交互的能力，如数据存储、处理与分析算力、跨链域通信及数据资源交互等能力，并用自身的 ADD 数据模型、数据分析引索及可视化引擎、数据应用中心产生结合现有大数据行业，进而产生应用与商业价值。

ADD 将用此数据资源能力自行研发或孵化出更便捷的助力互联网转型区块链的数据应用平台，具有超高算力与颗粒度超细的脱敏跨链域数据银行与商业应用服务、智能与动态的数据安全防护与审计体系、数据相关应用与分析体系等等。

ADD 数字代币将是此数据生态体系中流通的主要介质，为数据提供者支付对价、由数据与平台生态使用者支付对价、奖励生态开发者与贡献者，达成更低数据交易与应用成本。

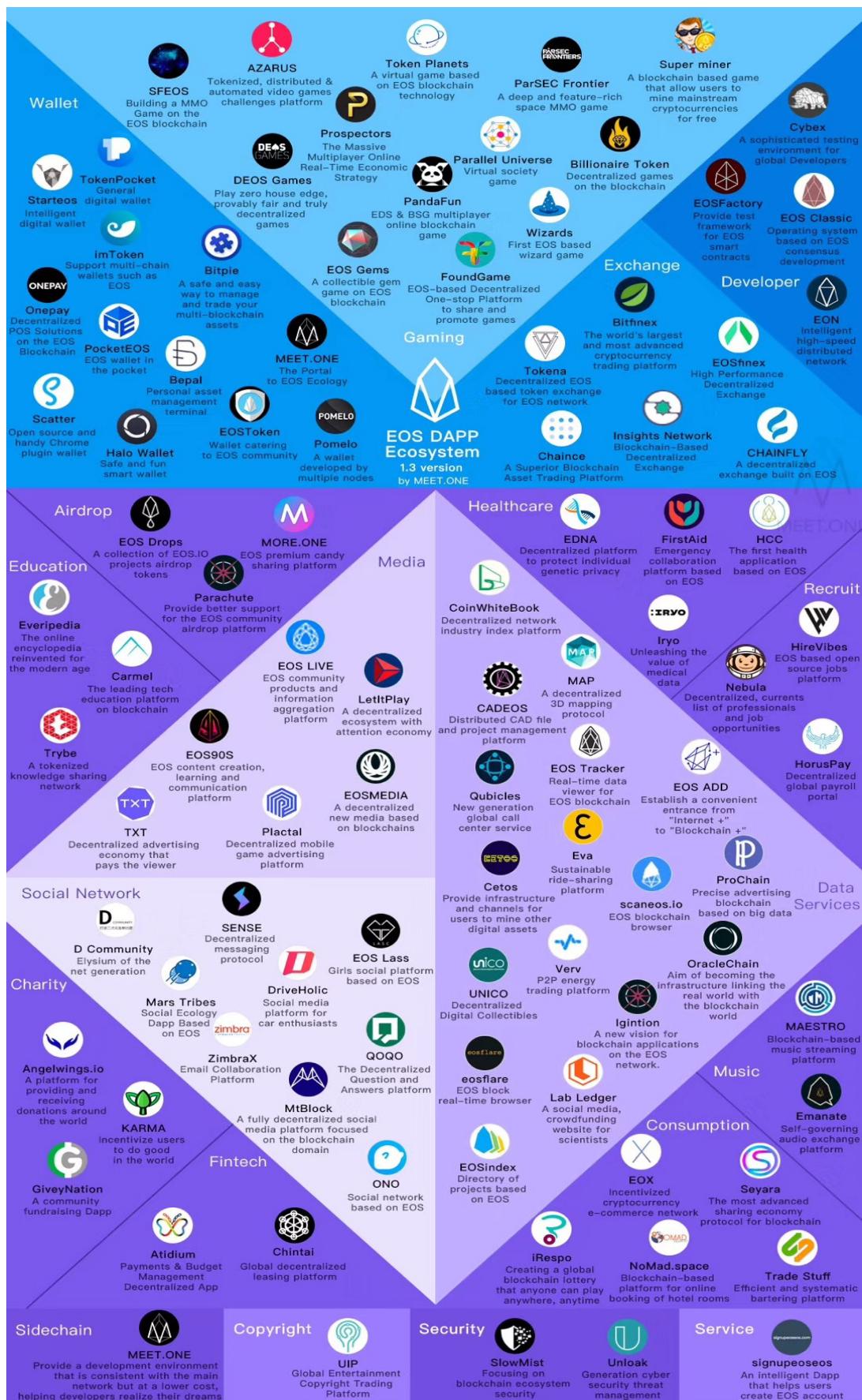


(以下无正文，附录及参考内容来源于第三方)

AIDP



## 附录 1：EOS 生态全景图（特别致谢：MEETONE 团队）



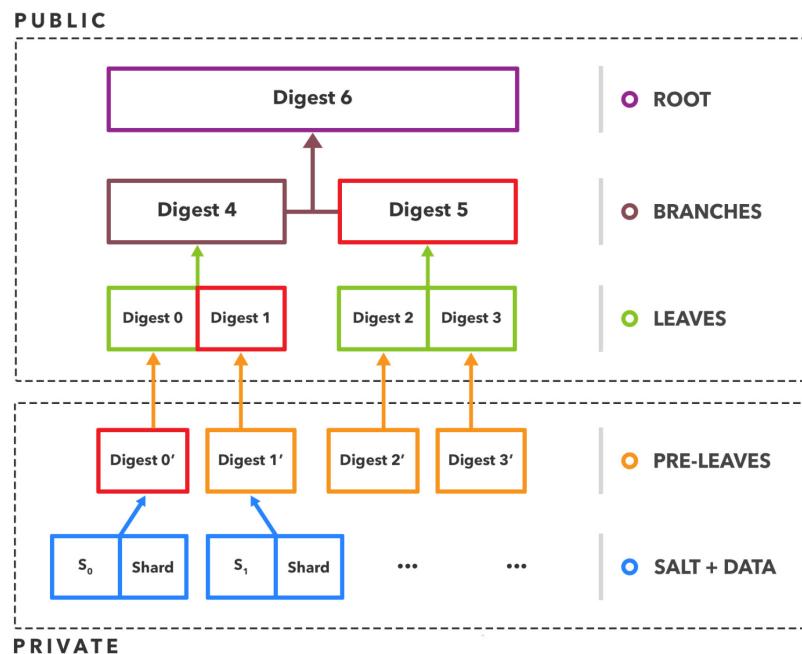


## 附录二：数据相关区块链项目竞品分析

### (一) Storj——去中心化的云存储应用

Storj 于 2014 年推出，是一个基于以太坊区块链的开源的、去中心化的、分布式存储系统，功能与中心化的 Dropbox, Onedrive 类似，Storj 创造了一个用于建立和执行点对点的存储合约的分布式网络，允许网络中的节点进行合约协商、数据传输、数据检索、存储、支付并验证远程数据的完整性和可用性。

Storj 的核心目标是改善传统云存储服务，致力于成为免审查、免监控、不会停机的云存储平台，让用户重新掌握设备和数据的“控制权”，激励用户（Farmers）向数据所有者（Renters）分享自己的剩余空间和流量以获得区块奖励（STORJ），通过 AWS S3 兼容接口，每个文件都经过加密、分解为多个部分并通过网络进行分发，只有数据所有者才能拥有访问其加密文件的密钥，从而提供开箱即用的无缝安全性和隐私。



附录图表 1 Storj 文件存储工作原理

其原理是通过对于待存储的文件进行加密得到对应密，并将密文切分为等长的分片（Shards）并同时生成一段随机数 Salts（挑战值）。将这些 Salts 添加到 Shards 前面并进行哈希计算得到 Pre-Leaves，对于这些 Pre-Leaves 进行再次哈希计算得到 Leaves，基于 Leaves 可以得到整个 Merkle Tree。

对于分布式云存储项目而言，在数据的可用性方面是巨大的挑战，因此 Storj 提出了可恢复性证明（Proof of Retrievability），避免 Farmers 仅仅存储了整个 Merkle Tree 而实际上却



没有存储 Shards 数据的情形发生，由于 Merkle Tree 下的 Salts 对于 Farmer 是不可见的，对于文件进行审计时，数据所有者（Renter）将向 Farmer 提供已存储的 Salts 值以及文件 Shards 的哈希值，而 Farmer 则根据其哈希值匹配出具体的 Shards 信息并结合之前的 Merkle Tree 返还一个 Merkle 证明，从而确保了数据的可用性验证。

由于 Storj 引入了 Bridge 机制，从而使得交易、存储、协商双方皆不可见并保障了隐私性。而附带一提的是，当一个 Salt 值被使用后，Farmer 可以保留 Salt 值并删除源文件，依然可以提供数据存储的审计证明，导致 Salt 值实际上是无法复用，有效验证只能进行一次，这种可用性问题同样也发生在 Siacoin 以及多数分布式存储项目中上。

Storj 与 Siacoin 类似，其主要区别在于 Siacoin 支持的 On-blockchain 区块链存储智能合约，创建智能合约后，即使数据所有者从不访问该文件也可保证 Farmer 获得付款，智能合约还允许对脱机或丢失数据的 Farmer 实施处罚。相比于 Siacoin 的存储合约，Storj 则提供按需付费的模式，主机由用户以 0.015USD/GB 进行付费，略高于 Siacoin 的收费，但两者皆显著低于中心化存储的费用。

下图是由 Davit K 所整理的关于 Siacoin 与 Storj 的比较：





	<b>SIACOIN ( SC )</b>	<b>STORJ ( STORJ )</b>
<b>Date Created</b>	June 2015	July 2014
<b>Purpose</b>	Decentralized cloud storage platform that uses a blockchain to facilitate payments. You can rent storage, get paid to host files, mine Siacoin, or contribute to the project	Decentralized, end-to-end encrypted cloud storage that uses blockchain technology and cryptography to secure your files. Users act as decentralized cloud nodes essentially.
<b>Feature</b>	Hosts receive money for storing files even if user doesn't access the files	Pay-as-you-go model, hosts are paid by the frequent use of the users
<b>Fully Decentralized?</b>	Yes	No
<b>Requires Deposit to Host?</b>	Yes	No
<b>Dynamic Pricing</b>	Yes	No ( Fixed )
<b>Cheaper Pricing</b>	1	2
<b>Scalability</b>	1	2
<b>Security</b>	1	2
<b>Trustworthiness of the Team</b>	1	2
<b>Coin Inflation Rate</b>	2	1
<b>Graphical User Interface</b>	2	1
<b>Speed of Data Access</b>	2	1
<b>Marketing</b>	2	1
<b>Exchanges</b>	<ul style="list-style-type: none"> <li>• Bittrex</li> <li>• Cryptopia</li> <li>• HitBTC</li> <li>• Poloniex</li> <li>• alcurEX</li> <li>• Bitsquare</li> <li>• ShapeShift</li> </ul>	<ul style="list-style-type: none"> <li>• Liqui</li> <li>• EtherDelta</li> <li>• Tidex</li> </ul>
<b>Price in CAD</b>	\$0.017 ( Bittrex July 4, 2017 )	\$0.97 ( Liqui July 4, 2017 )
<b>Circulating Supply</b>	27,474,953,847 SC	43,218,141 STORJ
<b>Market Cap</b>	\$473,540,728 ( CoinMarketCap July 4, 2017 )	\$42,075,197 ( CoinMarketCap July 4, 2017 )

附录图表 2 Siacoin 与 Storj 比较



## (二) IPFS——分布式存储及共享传输协议

星际文件系统 IPFS (Inter-Planetary File System) 是由 Protocol Labs 发起的，一种用于在互联网上创建服务器信息的协议，同时是一种全球的、点对点的分布式文件传输系统，旨在补充或取代当前的超文本传输协议 (HTTP)，可以让存储在 IPFS 系统上的文件，在全世界任何一个地方快速获取。

目前，互联网采用的超文本传输协议是基于一个位置进行处理，用户可以访问具有 IP 形式为 X.X.X.X 的网站并获得网站上的信息，这些 URL 指向世界各地的服务器，如美国、英国、新加坡等，这样的协议显然是鼓励中心化的，并且由于中心化的机制同时带来了安全性、低效性、错误性、集中化以及过度依赖主干网等问题。

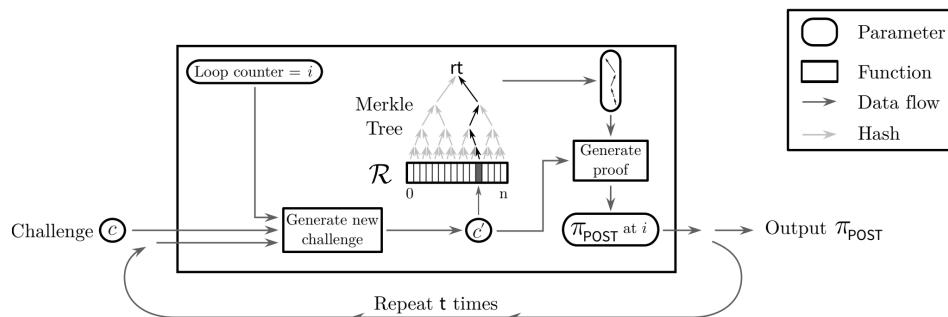
IPFS 综合了过去的对等系统的成功想法，包括分布式哈希表 (DHT)、BitTorrent、Git 和 SFS 的，IPFS 所做的是它根据信息是什么而不是信息在哪里来提供信息，通过 IPFS 的路由算法，用户可以选择从哪里获取内容，并且可以设置信任的 Peers 以及 Nodes 隐私以接收文件。

IPFS 的贡献是简化、发展和将成熟的技术连接成单一的内聚系统，IPFS 提供了高吞吐量的内容寻址块存储模型以及具有内容寻址的超链接，用以替代 HTTP 的基于域名的地址，并形成了一个广义的 Merkle DAG 数据结构，可用于构建版本文件系统、区块链或是永久性网站。而这样的系统也同样带来了许多特点：哈希寻址使内容不可变、从多个节点收集内容来节省带宽、访问内容离线或在低连接性网站与 Git 离线工作的意义相同、审查抵制。

为了进一步利用 IPFS 技术，FileCoin 被提议作为一种创建分布式存储网络的方法，通过使用 IPFS 协议以及 FileCoins (FIL)，使用互联网中未被使用的存储空间并激励用户成为共享经济的一部分，作为 2017 年最为成功的首次代币发行项目之一，Filecoin 的众筹吸引了超过 2.5 亿美元的资金。

与其他分布式存储解决方案类似，FileCoin 围绕激励和参与建构了基于区块链的经济和思维方式，Filecoin 采用复制证明 (Proof of Replication) 和时空证明 (Proof-of-Spacetime) 建立基于分布式存储的激励系统，通过复制证明解决女巫攻击、外部数据源攻击以及生成攻击，通过存储证明，矿工通过提供存储空间、写入和检索数据来实现挖矿。

目前，Filecoin 已推迟了其主网上线日期。



附录图表 3 PoSt 共识机制示例



### (三) Lambda——去中心化存储底层公链

Lambda 项目是一个高速、安全、可扩展的区块链基础设施项目，主要解决区块链领域数据存储问题，通过对 Lambda Chain 和 Lambda DB 的逻辑解耦和分别实现，向去中心化应用提供可无限扩展的数据存储能力，并实现多链数据协同存储、跨链数据管理、数据隐私保护、数据持有性证明、分布式智能计算等服务。

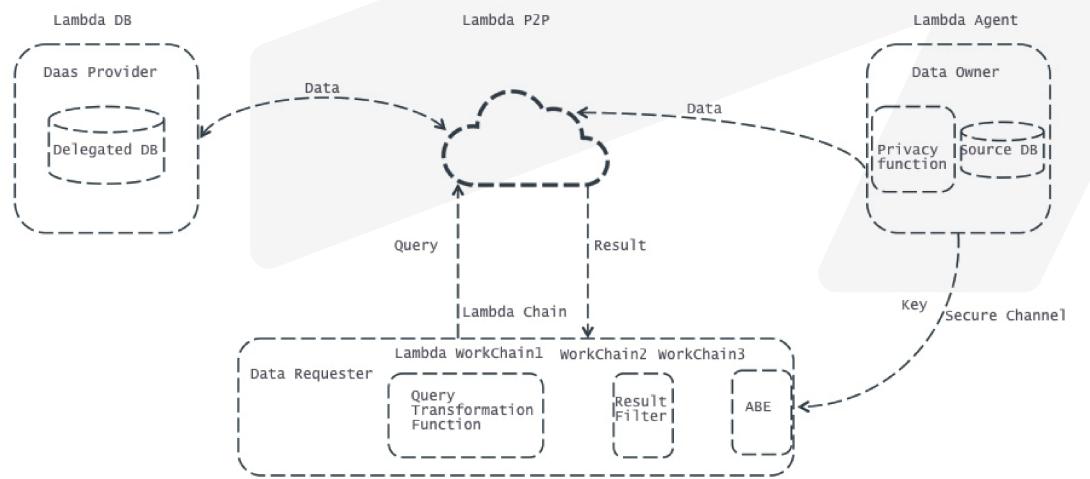
Lambda 项目本质是一个去中心化的数据库系统，计算能力依赖于矿工所共享的计算、存储和带宽能力，通过验证节点的共识完成数据的持有性验证 (PDP, Provable Data Possession) 和可恢复性证明 (POR, Proofs of Retrievability)，从而保证了存储在「非可信存储节点」上数据的完整性和可恢复性，这是该项目的最大创新点，从而实现成不可信机构之间的授权、加密与通信，向去中心化应用提供数据存储和管理服务。

Lambda 项目致力于为区块链和去中心化应用提供一个数据存储基础设施，通过去中心化的云数据库为 DApp 提供无限扩展的存储和访问能力；并且 Lambda 基于水平可扩展性和分片技术，提供了高速的交易能力；通过中继链技术和跨链交易验证，提供了系统内的跨链交易和数据访问及验证能力，通过对 BCP 协议的支持，提供和其他链系统的跨链通信能力。

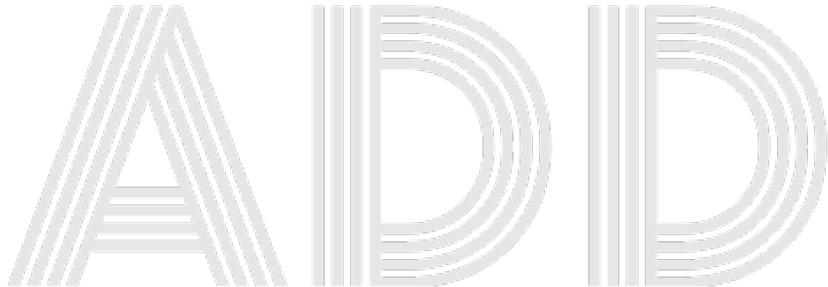
其整体核心的设计思想是链库分离机制和功能子链设计。去中心化应用可以按照数据不同的信任和公开验证等级，将数据分别存放在链上和数据库系统中，Lambda 项目提供了不同类型不同层级的数据协同管理。并且由于整个 LambdaDB 是一个无需许可的环境，Lambda 还完成了基于多授权机构属性基 (Attribute-Based Encryption) 加密的访问控制机制，以及完整的对存储数据的持有性证明。



### LAMBDA ABSTRACT ARCHITECTURE



附录图表 4 Lambda 抽象架构





#### (四) Bluzelle——去中心化数据库服务商

Bluzelle 是一个去中心化的、基于客户需求的、可扩展且可靠的数据库服务，与其他去中心化云存储相比，Bluzelle 所提供的是类似甲骨文，微软的 SQL Server 或 Mongo 的数据库访问并且主要针对 DApp 的开发者设计，预计该服务的潜在市场将包括基于区块链技术的预测市场、交易协议和数据流网络。

Bluzelle 致力于创建一种全新的代币共享经济和存储共享生态，Bluzelle 允许用户通过共享自己的计算机存储空间获取 BNT 代币，与此同时，允许分布式应用 DApp 开发者支付代币来存储他们的数据并尽可能有效的利用这些存储空间。Bluzelle 希望通过各种分发点连接消费者，例如 AWS, Android Studio, Quorum 和 Microsoft Azure，该项目侧重于亚太地区的软件开发人员和企业软件架构师，为确保开发人员在性能，可靠性和可扩展性方面获得最高吞吐量，Bluzelle 实施了集群技术。

**Bluzelle 的数据库就像 Torrent 或内容交付网络一样，可以实时从多个节点检索数据。**当数据上传到网络时，它会被分割成更小、更快、更容易管理的部分，称为数据分片（Shards），这些数据分片会完全复制到群集中的每个节点。当 DApp 请求数据时，从最近的本地节点或与网络上较快的节点并行地提取分片，这可以减少延迟并提高性能。

群集中的节点分布在多个地理位置，以减少当地事件（如停电或自然灾害）的影响并有助于降低由于安全漏洞，人为干扰而导致单点故障的风险，如果某个节点在群集中脱机，新节点可以在对网络造成最小影响的情况下出现，而数据仍由该组中的其他节点进行维护。在群体层面创建共识，其中节点组就当地共识达成一致，这与大多数区块链应用程序不同，后者依赖于全网络范围内的共识状态。

图表 1 过去、现在与未来的数据库比较

过去	现在	未来
单一系统数据库	云数据库	分布式数据库
多个源的失效	单一源的失效	无失效
规模化昂贵	规模化昂贵	高效规模化
无隐私性	无隐私性	高度隐私性
无不可变	无不可变	不可篡改
存在性能瓶颈	部分性能瓶颈	无性能瓶颈

(数据来源：Bluzelle 白皮书)



## (五) CPChain——面向物联网系统的基础数据平台

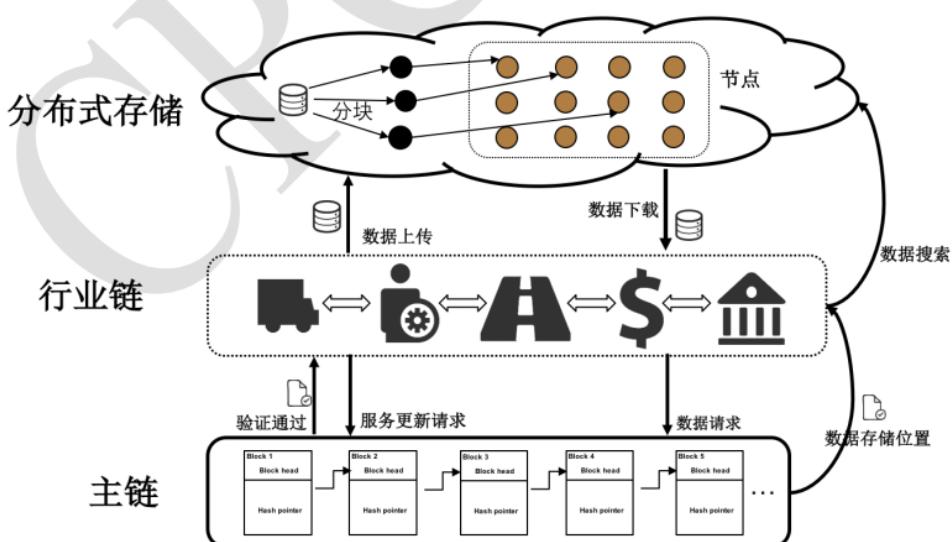
CPChain 平台致力于构建一个面向物联网系统的基础数据平台, CPChain 重点围绕区块链技术应用于物联网行业所面临的扩展性、安全性、和实时性问题, 结合区块链、物联网、分布式存储, 构建物联网体系架构, 提供从数据获取、存储、分享到应用的全流程解决方案。

CPChain 专注于多方参与的数据交易和基于物联网大数据的人工智能决策应用场景, 建立多方信任和实现异构数据的互联互通。当前, 去中心化的区块链系统要求全网节点对同一交易(数据)进行运算, 从计算和存储角度来说具有很大的弊端, 无法充分发挥分布式网络系统的协同能力, 其只能遵循“木桶原理”, 因此不具有扩展性。

CPChain 应用区块链解决方案, 在解决数据存储、分享、交易的可扩展性问题方面, 提出数据层与控制层分离思想, 构建平行架构来增强系统的可扩展性。如下图所示, 将去中心化的主链、行业链和分布式存储系统有机结合, 区块链作为 CPC 的控制层, 不再负责存储区块链系统的全部信息, 而是仅上传数据的标志与凭证。

CPChain 平行分布式架构中, 分布式云存储层与区块链层作为两层平行的分布式网络, 分别承担数据存储和计算任务的功能。在云存储层, 用户数据被加密上传到云端; 在区块链层, 用户数据在客户端加密后分块, 各个部分进入不同的存储节点, 同时哈希凭证上传到区块链网络的所有节点中, 保证数据完整和准确。

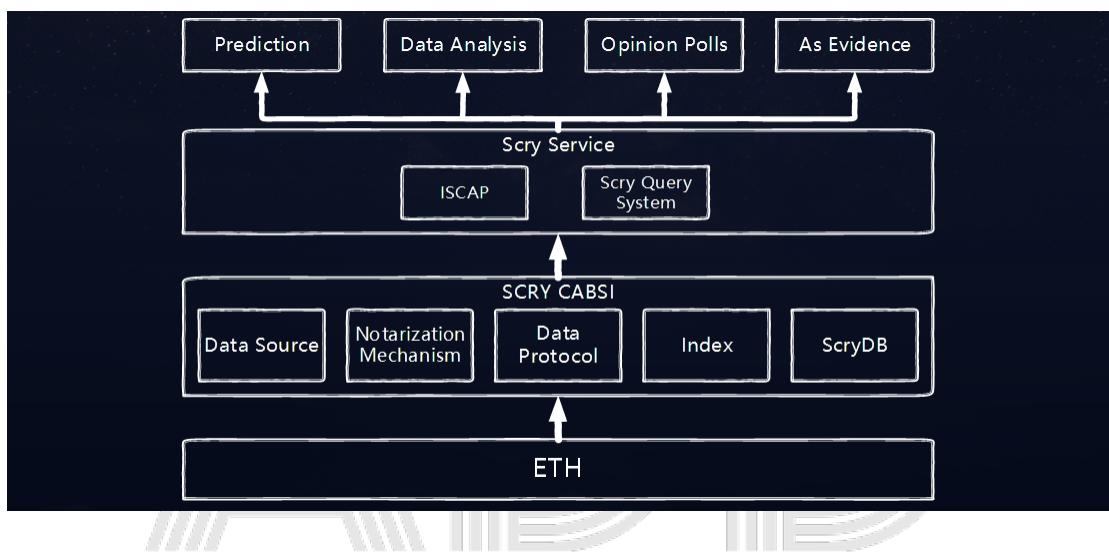
这种平行分布式架构, 将数据层从区块链中剥离, 既保留了区块链系统的安全性、去中心化特性, 又降低区块链存储负担并提高可扩展性; 而在区块大小不变的情况下, 单个区块链可以打包的交易数量大大增加, 提升平台的交易处理速度





## (六) Scry.info——去中心化的数据交易所

Scry.info 是全球首个区块链可量化数据交易平台，实现真实数据的存储，验证，共享，分析和交易。利用 Scry.info 内置智能合约，对跨行业资源的分布式用户提供集群智能数据验证，并以 Scry 智能合约的形式在 Scry 平台上提供调用、工具、分享和交易。Scry.info 还提供对市场端客户与开发者提供对特定行业/事件发起智能合约协议，应用于各种商业数据场景，例如体育，娱乐，农业，金融市场等的民调或交易市场。。



图表 2 Scry.info 技术架构

Scry.info 的底层是双链结构的公有链以及协议层，双链中一条链侧重于做数据，另外一条链专门做合约的结算，这条双链交织并不断利用算法进行结算和通信，将交易、存储于结算分离，从而保证了高 TPS、高可扩展性的情况下，保证区块链网络的安全性。

在双链上是一个灵活而具备可扩展性的 Scry.info CABS，Scry.info CABS 是创造性定量数据协议机制的核心技术，Scry.info 为跨行业的分布式用户，提供基于集群的智能数据验证并在 Scry 平台上实现访问、工具、交易和共享，基于 Scry Protocol 的开发人员可以调用 Scry.info 中不同行业的标记数据源以形成智能合约以构建所需的 DApp。

Scry.info 具备五大核心技术板块：

- Data Source：对 SCRY 底层数据进行索引、查询等；
- Scry Service：提供对 Scry CABS 的部分封装，为上层平台及应用提供方便的调用 API；
- Data Protocol：为数据写入者、数据需求者、数据公投人创建智能合约提供条件；
- 数据公证：对数据提供者的数据进行盲投，公证人参与数据盲投，核实数据的真实性；



- Scry DB: 将数据放置于外部去中心化数据存储链中，把外部存储的索引保存在数据合约、注册合约内部；

## (七) Bottos——去中心化人工智能数据合约

铂链（BOTTOS）是基于区块链技术，实现将人工智能的数据和模型通过点对点网络进行登记发行、转让交易的去中心化网络协议，是一个基于区块链的去中心化的数据交易平台、数据集合池，通过智能合约帮助人工智能项目简单、快速获取训练数据，通过数据挖矿使得个人数据产生财富价值的基于共识的、可扩展的、标准化的、特性完备、易于开发的和协同的一站式应用平台。

铂链主要面向机器人、物联网、VR/AR 等领域，连接有数据需求的 AI 初创公司和能提供特定数据的个人，解决人工智能优质数据获取难的痛点，通过智能数据入股合约帮助人工智能模型快速进化，用数据挖矿方式实现个人数据共享并实现原生数据/模型资产的经济激励，在个人数据和 AI 模型之间建立价值联系。因此可以将项目理解为“区块链+数据×人工智能”。

- 从数据共享角度：铂链将基于区块链技术构建全球最大的数据集合池，将极大推进现有价值属性的优质数据/模型资产 IP 的注册登记、交换、交易和流转。
- 从人工智能角度：铂链借鉴以太坊生态的成功经验，帮助人工智能技术实现场景化迭代；通过高扩展性的智能合约实现“数据入股”，帮助数据需求方获得训练数据，为数据贡献方实现贡献价值，解决人工智能（AI）行业优质数据获取难的痛点。

DataMarket 是铂链（BOTTOS）连接现实世界和数据世界的桥梁，包含 Oracle 和 DataFeed 机制。DataMarket 的最终目标是通过大量的数据交易，最终可以实现任何数据的市场估值，所以当业务达到一定规模以后，DataMarket 的进阶版本是提供数据/模型的共享平台，是实现预估价机制的组件（类似 Gnosis 的预测机制），通过设置多重权重（例如前期会引入历史同类数据对标和节点投票双权重），来给数据/模型的交易估价提供参考，帮助数据/模型需求方和数据/模型供应方完成数据的交易流转，为特定数据预估价市场提供资源和实现数据细分。



附录图表 6 BOTTOS 去中心化数据流转平台

特别致谢：维京资本为附录与白皮书提供的竞品项目信息，与重要咨询意见

