



数据要素研究系列

编号：DE-2021-01

# 中国数据要素市场发展报告 (2020~2021)

国家工业信息安全发展研究中心

2021年4月

## 报告编写组

### 参编单位：

国家工业信息安全发展研究中心

蚂蚁科技集团股份有限公司

天聚地合（苏州）数据股份有限公司

### 参编人员：

乔思渊 杨捷 王梦辰 周易江 孙璐

骆伊宁 杨玫 刘巍 纪润博 毛庆凯

袁桂安

## 版权声明

本报告版权属于国家工业信息安全发展研究中心，并受法律保护。转载、编撰或以其他方式使用本报告文字或观点，应注明“来源：国家工业信息安全发展研究中心”。违反上述声明者，将追究其相关法律责任。

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

## 前 言

“十三五”时期是我国大数据产业蓬勃发展的阶段，根据国家工业信息安全发展研究中心产值测算数据，截至 2020 年底，我国大数据产业规模已达万亿元。随着我国大数据产业与实体经济深度融合、产业发展不断壮大，数据作为生产要素的属性不断凸显。如何实现数据要素市场化配置，激活数据要素潜在价值，推动与实体经济继续深度融合，是“十四五”期间我国需要面临的重要课题之一。

本报告提出了数据要素及数据要素市场的边界，梳理了国内外数据要素市场发展现状，重点从市场概况、政策脉络、产业图谱及市场运行机制等角度，并结合相应的产值规模、技术水平、产品和服务、商业模式等情况，阐释了我国数据要素市场的发展现状，分析了现阶段我国数据要素市场存在的问题及未来发展趋势，提出了对策与建议。

在本报告撰写过程中，国家工业信息安全发展研究中心（以下简称“国家工信安全中心”）通过专家访谈、企业调研、案头研究等方式开展数据要素市场相关研究，并结合自有的逾 5000 家企业大数据案例库对报告界定的产值规模进行了测算和分析。

由于编者水平有限，本报告难免存在疏漏与不足，恳请各界同仁阅后批评指正，加强合作交流。

报告编写组

二〇二一年四月

# 目 录

一、数据要素市场相关概念界定.....	1
(一) 数据.....	1
(二) 数据资源和数据资产.....	2
(三) 数据要素和数据要素市场.....	2
二、国外数据要素市场发展概况.....	6
(一) 美国：数据交易模式多样，数据市场政策开放.....	6
(二) 欧盟：数据立法顶层设计，加强数据主权建设.....	7
(三) 德国：率先打造数据空间，建立可信流通体系.....	8
(四) 英国：金融行业先行先试，促进数据市场交易.....	8
(五) 日本：创新设立数据银行，释放个人数据价值.....	9
三、我国数据要素市场发展现状.....	10
(一) 发展概况.....	10
(二) 政策脉络.....	11
(三) 产业图谱.....	12
1. 数据采集.....	13
2. 数据存储.....	15
3. 数据加工.....	17
4. 数据流通.....	19
5. 数据分析.....	29
6. 生态保障.....	31
(四) 运行机制.....	34

1.定价机制.....	35
2.收益分配机制.....	38
3.政企合作机制.....	40
<b>四、我国数据要素市场发展面临的问题与挑战.....</b>	<b>43</b>
(一) 数据权属界定问题尚不明晰.....	43
(二) 数据要素市场存在安全风险.....	44
(三) 数据流通交易机制有待完善.....	44
(四) 数据监管法律法规存在空缺.....	45
<b>五、我国数据要素市场未来的发展趋势.....</b>	<b>47</b>
(一) 数据要素市场主体趋于多元化.....	47
(二) “数据可用不可见”成为技术趋势.....	48
(三) 数据要素市场行业应用不断拓展.....	50
(四) 数据要素市场生态体系日趋完善.....	51
<b>六、对策与建议.....</b>	<b>53</b>
(一) 重塑数据价值流动规则.....	53
(二) 扩大公共数据开放程度.....	54
(三) 促进数据要素交易流通.....	55
(四) 提升数据安全保障能力.....	56

# 图 目 录

图 1	数据要素理论图.....	3
图 2	数据要素市场构成关系图.....	4
图 3	数据流通金字塔模型.....	5
图 4	2016-2025 年中国数据要素市场规模.....	10
图 5	数据要素市场政策梳理.....	12
图 6	数据要素市场产业图谱.....	12
图 7	API 技术服务模式图.....	22
图 8	隐私计算原理.....	23
图 9	安全多方计算架构图.....	24
图 10	联邦学习医疗行业应用.....	25
图 11	数据沙箱技术图.....	26
图 12	个人数据出境安全监督.....	29
图 13	数据要素市场机制一览.....	35
图 14	数据要素市场收益分配机制.....	40

## 一、数据要素市场相关概念界定

数据要素是一个经济学概念，对数据要素市场相对准确、清晰的认识和界定，是探索和培育数据要素市场模式和方向的重要前提，也是值得各界商榷的难点所在。本节探索对数据、数据资源、数据资产、数据要素以及数据要素市场的含义进行界定和分析。

### （一）数据

数字经济的背后实际上是数字经济，甚至可以说“无数据，不经济”<sup>1</sup>，数据是整个数据要素市场最基本的构成元素。计算机科学将数据定义为“对所有输入计算机并被计算机程序处理的符号的总称”<sup>2</sup>。国际数据管理协会（DAMA）也给出了相似的定义：“数据是以文字、数字、图形、图像、声音和视频等格式对事实进行表现”<sup>3</sup>，而国际标准化组织（ISO）对以上两种定义进行了进一步概括，认为“数据是对事实、概念或指令的一种形式化表示”。

以上定义各有侧重，一方面，数据若想为人所用，必须能够被计算机以数字化、可视化的形式呈现出来，这是数据必备的外在形态；另一方面，数据之所以有价值，是因为其承载着某些客观事实，这是数据的内在实质。因此我们认为，数据是指所有能够输入计算机程序处理、反应一定事实、具有一定意义的符号介质的总称。

<sup>1</sup> 赵刚. 数据要素[M]. 人民邮电出版社, 2021年.

<sup>2</sup> 严蔚敏, 吴伟民. 数据结构[M]. 清华大学出版社, 2002年.

<sup>3</sup> DAMA International. DAMA 数据管理知识体系指南[M]. 清华大学出版社, 2016年.



## **(二) 数据资源和数据资产**

数据资源和数据资产都是数据汇聚产生的结果，数据资源是数据的自然维度，数据资产是数据的经济维度，两个概念相互融合，且不冲突。

数据资源，是载荷或记录信息的按一定规则排列组合的物理符号的集合。可以是数字、文字、图像，也可以是计算机代码的集合。数据是一种宝贵的资源，但是数据的法律权属界定仍然是一个世界性难题，传统的法学确权理论还无法移植到数据这种易复制的无形资源上。

数据资产，从本质上来讲是产权的概念，是指由个人或企业拥有或者控制的，能够为个人或企业带来经济利益的，以物理或电子的方式记录的数据资源。从会计学角度看，数据资产也不完全符合会计准则中对资产及无形资产的定义，很难将数据计入财务报表。因此，数据目前还不能被视为传统意义上的资产。但数据资产化是世界经济发展的必由之路，也是数据成为一种生产要素的必然要求。

## **(三) 数据要素和数据要素市场**

进入信息时代后，最重要的生产资料是用“比特”来描述的数字化信息，人类的生产活动正逐渐由物理世界深度转向比特世界，越来越多的生产环节需要在赛博空间中独立完成。数据对生产的贡献越来越突出，同时也显著提升了其他生产要素在生产中的利用效率，因此，数据已成为当今经济活动中不可或缺的生产资料。数据作为生产要素参与生产，需要进行市场化配置，形成生产要素价格及其体系。数据要

素价格体系的建立，又是建立在数据所有制基础上的。因此谁掌握数据资产，在一定程度上就可以影响体系建立。

生产要素主要包含土地、资本、技术、劳动力和数据。数据作为新型生产要素，具有劳动工具和劳动对象的双重属性。首先数据作为劳动对象，通过采集、加工、存储、流通、分析环节，具备了价值和使用价值；其次，数据作为劳动工具，通过融合应用能够提升生产效能，促进生产力发展。

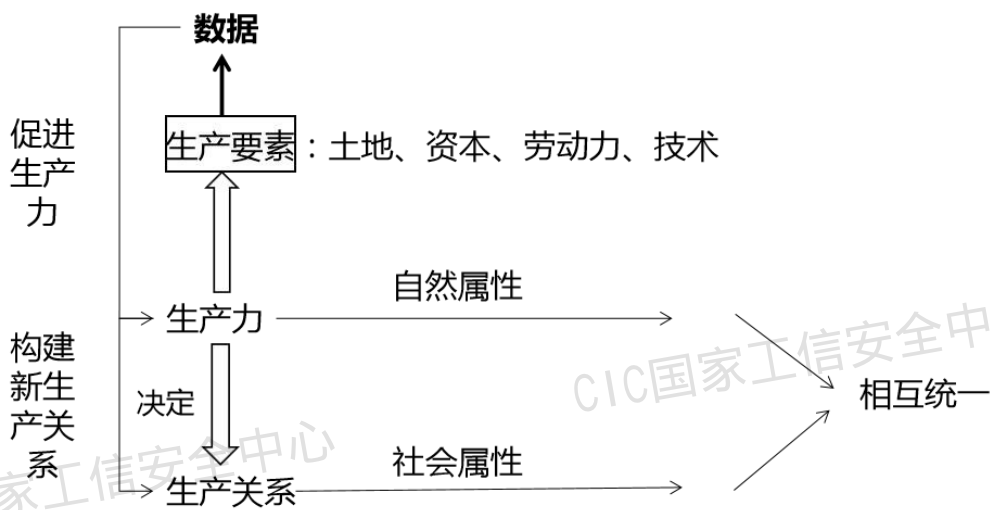


图1 数据要素理论图

来源：国家工业信息安全发展研究中心

数据要素市场就是将尚未完全由市场配置的数据要素转向由市场配置的动态过程，其目的是形成以市场为根本调配机制，实现数据流动的价值或者数据在流动中产生价值。数据要素市场化配置是一种结果，而不是手段。数据要素市场化配置是建立在明确的数据产权、交易机制、定价机制、分配机制、监管机制、法律范围等保障制度的基础上。未来数据要素市场的发展，需要不断动态调整以上保障制度，最终形成数据要素的市场化配置。

本报告从产业链的角度出发，将我国数据要素市场归结

为数据采集、数据存储、数据加工、数据流通、数据分析、数据应用、生态保障七大模块，覆盖数据要素从产生到发生要素作用的全过程。其中数据采集、数据存储、数据加工、数据流通、数据分析、生态保障六大模块，主要是数据作为劳动对象，被挖掘出价值和使用价值的阶段；而数据应用模块，主要是指数据作为劳动工具，发挥带动作用的阶段。本报告当前关注数据作为劳动对象，挖掘数据价值的阶段，聚焦于数据采集、数据存储、数据加工、数据流通、数据分析、生态保障六大模块。

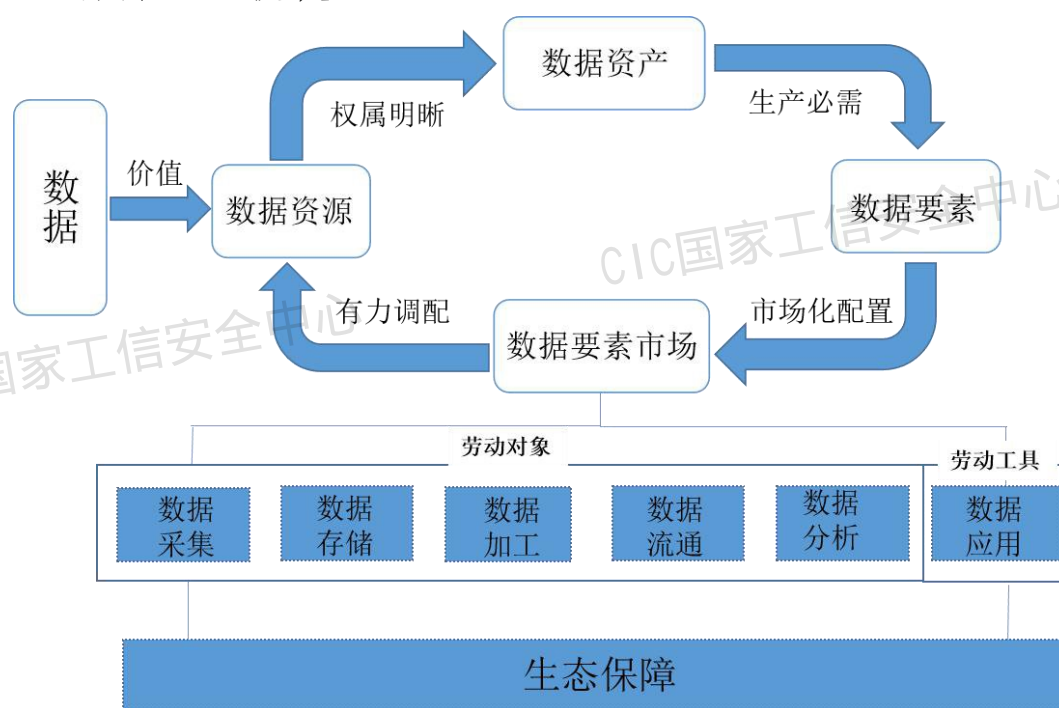


图2 数据要素市场构成关系图

来源：国家工业信息安全发展研究中心

保障数据要素市场化配置这一结果，不同产业链环节均被赋予了独特使命。数据采集环节，关注数据采集的准确度、全面性；数据储存环节，关注数据储存安全性，调用实时性；数据加工环节，关注数据加工精度；数据流通环节是数据要素市场的核心环节，关注在保障所有者权利的前提下，进行

合理合规流通；数据分析环节，关注数据深度分析挖掘；数据应用环节，关注数据作为要素在合理、充分应用中产生价值，降低生产要素获取成本及提升其赋能水平。其中，数据流通作为数据要素市场的核心环节，需要针对不同类型数据，提出不同的解决方案。国家工信安全中心认为，需要针对不同数据分级分类进行数据要素市场化配置，并提出了“数据流通金字塔模型”。该模型将数据分为公开数据、低敏感度数据、中敏感度数据、高度机密数据四种，提出针对不同数据类型，应用不同的数据流通技术和模式。



图3 数据流通金字塔模型

来源：国家工业信息安全发展研究中心

## 二、国外数据要素市场发展概况

### （一）美国：数据交易模式多样，数据市场政策开放

美国发达的信息产业提供了强大的数据供给和需求驱动力，促进其数据交易流通市场的形成和发展。美国在数据交易流通市场构建过程中，通过数据交易产业推动政策和法律制定，开放的政策和法律又进一步规范了数据交易产业的发展。

首先，建立了政务开放机制。美国联邦政府自 2009 年发布《开放政府指令》之后，便通过建立“一站式”的政府数据服务平台 Data.gov 加快开放数据进程。联邦政府、州政府、部门机构和民间组织将数据集统一上传到该平台，政府通过此平台将经济、医疗、教育、环境与地理等方面的数据以各种可访问的方式发布，并将分散的数据整合，开发商还可通过平台对数据进行加工和二次开发。

其次，发展多元数据交易模式。美国现阶段主要采用 C2B 分销、B2B 集中销售和 B2B2C 分销集销混合三种数据交易模式，其中 B2B2C 模式发展迅速，占据美国数据交易产业主流。所谓数据平台 C2B 分销模式，即个人用户将自己的数据贡献给数据平台以换取一定数额的商品、货币、服务、积分等对价利益，如 personal.com、Car and Driver 等；数据平台 B2B 集中销售模式，即以美国微软（Azure）为首的数据平台以中间代理人身份为数据的提供方和购买方提供数据交易撮合服务；数据平台 B2B2C 分销集销混合模式，即以数据平台安客诚（Acxiom）为首的数据经纪商（Databroker）

收集用户个人数据并将其转让、共享给他人。

**最后，平衡数据安全性与产业利益。**在涉及到数据保护等方面，目前美国尚没有联邦层面的数据保护统一立法，数据保护立法多按照行业领域分类。虽然脸书（Facebook）、雅虎（Yahoo）、优步（Uber）等公司近些年来均有信息失窃案件发生，但由于硅谷巨头的游说使得美国联邦在个人数据保护上进展较为缓慢。

## **（二）欧盟：数据立法顶层设计，加强数据主权建设**

欧盟委员会希望通过政策和法律手段促进数据流通，解决数据市场分裂问题，将 27 个成员国打造成统一的数字交易流通市场；同时，通过发挥数据的规模优势建立起单一数字市场，摆脱美国“数据霸权”，回收欧盟自身“数据主权”，以繁荣数字经济发展。

**首先，建立数据流通法律基础。**2018 年 5 月，《通用数据保护条例》（GDPR）在欧盟正式生效，特别注重“数据权利保护”与“数据自由流通”之间的平衡<sup>4</sup>，这种标杆性的立法理念对中国、美国等全球各国的后续数据立法产生了深远而重大的影响。但由于 GDPR 的条款较为苛刻，使得推出后，欧盟科技企业筹集到的风险投资大幅减少，每笔交易的平均融资规模比推行前的 12 个月减少了 33%。

**其次，积极推动数据开放共享。**2018 年，欧盟提出构建专有领域数字空间战略，涉及制造业、环保、交通、医疗、财政、能源、农业、公共服务和教育等多个行业和领域，以

<sup>4</sup> 何渊. 数据法学[M]. 北京大学出版社, 2020 年.

此推动公共部门数据开放共享、科研数据共享、私营企业数据分享。

最后，完善顶层设计。欧盟基于 GDPR 发布了《欧盟数据战略》，提出在保证个人和非个人数据（包括敏感的业务数据）安全的情况下，有“数据利他主义”（Data altruism）意愿的个人可以更方便地将产生的数据用于公共平台建设，打造欧洲公共数据空间。

### （三）德国：率先打造数据空间，建立可信流通体系

德国提供了一个“实践先行”的思路，通过打造数据空间构建行业内安全可信的数据交换途径，排除企业对数据交换不安全性的种种担忧，引领行业数字化转型，实现各行各业数据的互联互通，形成相对完整的数据流通共享生态。数据空间是一个基于标准化通信接口并用于确保数据共享安全的虚拟架构，其关键特征是数据权属。它允许用户决定谁拥有访问他们专有数据的权利并提供访问目的，从而实现对数据的监控和持续控制。目前，德国数据空间已经得到包括中国、日本、美国在内的 20 个多个国家及 118 家企业和机构的支持。

### （四）英国：金融行业先行先试，促进数据市场交易

作为高度重视数据价值的国家，英国采用开放银行战略对金融数据进行开发和利用，促进数据的交易和流通。该战略通过在金融市场开放安全的应用程序接口（API）将数据提供给授权的第三方使用，使金融市场中的中小企业与金融服务商更加安全、便捷地共享数据，从而激发市场活力，促

进金融创新。开放银行战略为具有合适能力和地位的市场参与者提供了六种可能的商业模式：前段提供商、生态系统/引用程序商店、特许经销商模型、流量巨头、产品专家以及行业专家。其中，金融科技公司、数字银行等前端提供商通过为中小企业提供降本增效服务来换取数据，而流量巨头作为开放银行产业链的最终支柱掌握着银行业参与者所有的资产和负债表，控制着行业内的资本流动性<sup>5</sup>。目前，英国已有100家金融服务商参与了开放银行计划并提供了创新服务，数据交易流通市场初具规模。

#### **（五）日本：创新设立数据银行，释放个人数据价值**

日本从自身国情出发，创新“数据银行”交易模式，最大化释放个人数据价值，提升数据交易流通市场活力。数据银行在与个人签订契约之后，通过个人数据商店（Personal Data Store, PDS）对个人数据进行管理，在获得个人明确授意的前提下，将数据作为资产提供给数据交易市场进行开发和利用。从数据分类来看，数据银行内所交易的数据大致分为行为数据、金融数据、医疗健康数据以及行为嗜好数据等；从业务内容来看，数据银行从事包括数据保管、贩卖、流通在内的基本业务以及个人信用评级业务。数据银行管理个人数据以日本《个人信息保护法》（APPI）为基础，对数据权属界定以自由流通为原则，但医疗健康数据等高度敏感信息除外。日本通过数据银行搭建起个人数据交易和流通的桥梁，促进了数据交易流通市场的发展。

<sup>5</sup> 毕马威.《从开放银行到业务开放——英国企业家对开放银行服务的展望》，2018年.



### 三、我国数据要素市场发展现状

#### (一) 发展概况

当前，我国数据要素市场处于高速发展阶段。“十三五”期间，我国各要素市场规模实现不同程度的增长，以数据采集、数据储存、数据加工、数据流通等环节为核心的数据要素市场增长尤为迅速。据国家工信安全中心测算数据，2020年我国数据要素市场规模达到545亿元<sup>6</sup>，“十三五”期间市场规模复合增速超过30%；“十四五”期间，这一数值将突破1749亿元，整体上进入高速发展阶段。

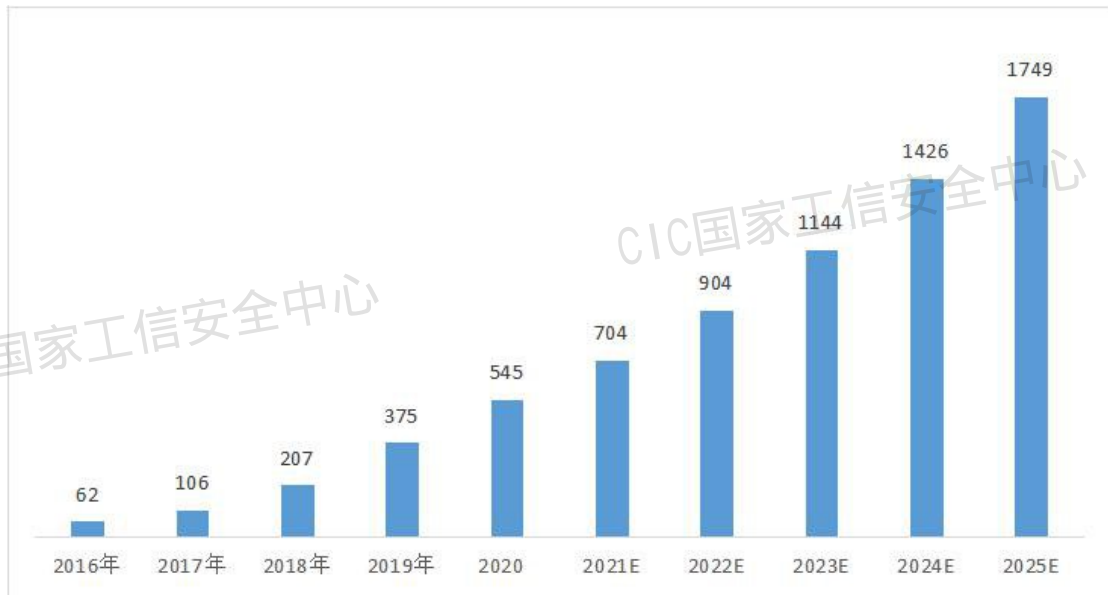


图4 2016-2025年中国数据要素市场规模

来源：国家工业信息安全发展研究中心

在技术融合层面，以联邦学习、安全多方计算、数据沙箱为主的新技术不断助力我国数据要素市场发展。在区域发展层面，数据要素市场区域分工协作格局逐渐形成，北、上、广、深等城市依托自身人才与技术优势大力发展数据流通交

<sup>6</sup> 如前文所述，本报告当前关注数据作为劳动对象，挖掘数据价值的阶段，聚焦于数据采集、数据存储、数据加工、数据流通、数据分析、生态保障六大模块，因此测算数据中未包含数据应用的部分。以下研究和阐述皆遵循这六个模块或环节。

易与数据技术研发等高精尖业务，而围绕中心经济带的欠发达地区则利用人力密集特点开展数据标注、清洗等传统数据服务。通过技术创新及统筹发展，我国数据要素市场相对完善的产业生态初见雏形。

## （二）政策脉络

充分发挥数据要素市场化配置是我国数字经济发展水平达到一定程度后的必然结果，也是数据供需双方在数据资源和需求积累到一定阶段后产生的必然现象。2014年，“大数据”第一次写入政府工作报告，标志着我国对大数据产业顶层设计的开始。在“十三五”期间，大数据相关的政策文件密集出台，为数据作为生产要素在市场中进行配置，提供了政策土壤，也推动了我国大数据产业不断发展，技术不断进步，基础设施不断完善，融合应用不断深入。各个地方积极先行先试，探索出了一条适合我国大数据产业发展的路径。

2020年是我国“十三五”规划收官之年，根据国家网信安全中心测算数据，我国大数据市场规模已经达到1万亿元。数据在国民经济中的地位不断突出，要素属性逐渐凸显。2020年4月，中共中央、国务院印发《关于构建更加完善的要素市场化配置体制机制的意见》，将数据列为生产要素，明确指出了市场化改革的内容和方向。数据要素市场的培育将消除信息鸿沟、信任鸿沟，促进数据资源要素化体现，推进各方对数据资源的合作开发和综合利用，实现数据价值最大化，以新动能、新方向、新特征开启数据生态体系培育新

征程。



图 5 数据要素市场政策梳理

来源：国家工业信息安全发展研究中心

### (三) 产业图谱

数据要素市场产业图谱，主要是指对我国数据要素市场产业链各环节市场主体的梳理情况，包含对政府、企业等不同主体提供从数据产生到数据产生价值服务的过程。



图 6 数据要素市场产业图谱

来源：国家工业信息安全发展研究中心

## 1. 数据采集

数据采集是数据要素市场的基石。企业在经营过程中，需要针对不同的数据类型进行采集，并进行分析处理训练。企业数据采集主要包括内部数据采集、外部数据采集和定制化数据采集等三种类型。

### (1) 企业内部数据

一是企业在日常经营活动中产生的数据。通过 Excel、数据仓库、数据库、ETL 等技术进行储存和调取；政府部门和企业经营数据的采集主要与主体的信息化水平、数据治理能力提升有直接联系，因此从上到下的整体设计，才能保证数据的一致性、可用性。该部分数据是现阶段大数据、人工智能分析与训练的主要数据来源。现阶段，提供该部分数据采集的公司主要是企业基础软件公司，如用友、SAP 等。

二是生产环节的数据采集。该环节主要通过物联网设备，如温度传感器、PLC 传感器、MES 等。但该环节数据采集主要受到两个方面的制约，一是部分机器没有数据接口，二是存在大量异构的通信规范。未来需要对现有机器设备进行信息化改造升级，以及完善定义通信标准。现阶段，该部分数据采集主要集成在相关领域解决方案中进行整体落地。在工业领域，相关企业有树根互联、航天云网等，互联网领域，相关企业有诸葛、神策等。

### (2) 外部数据采集

企业对于外部数据的采集主要是用于网络舆情分析、广告投放检测，以调整自身的产品和营销策略。在外部数据采

集方面，主要通过爬虫技术、埋点检测技术、用户调研等方法，综合给出外部数据的检测结果。在该方面，主要公司有 AdMaster 和秒针等。其中，AdMaster 主要定位在社交媒体舆情监控方向；秒针主要定位在曝光监测方向，其在国内市场占有率占有较大份额。企业在进行外部数据采集时，需要保证数据采集不侵犯个人隐私和国家利益，需要保证采集数据在数据需要的最小范围。另外，现阶段国家在数据采集方面，也缺少相关的法律法规以及行业标准。未来需要加强该方面的法规、标准研制，增强数据采集的复用性，更好的保护社会信息安全。

### （3）定制化数据采集

定制化数据采集，主要是用于非结构化数据采集。定制化数据采集广泛应用在各个领域。比如摄像头、麦克风、激光雷达都是数据采集工具，采集的数据类型也包含文本、语音、图像、云点图等。现阶段定制化数据采集主要采用人工采集的方式，但也出现了利用人工智能手段进行智能数据采集的技术。

**一是人工数据采集。**即根据方案设计，对目标领域、场景的特定数据进行采集，多以人像、车辆周遭环境、声音片段等内容为主，采集需要满足数据量大、数据精度高、数据丰富度高的要求。现阶段提供相关服务的公司有海天瑞声、倍赛数据等。

**二是智能化数据采集。**人工智能（AI）合成图像和视频，即利用生成对抗网络（GAN）合成图像、视频，所生成的图

片，可以为 AI 提供模拟的训练数据，但该技术现阶段仍处于探索阶段，生成数据的质量仍处于较低水平。

## 2. 数据存储

企业在完成数据采集环节后，需对有价值的数据进行有效存储，以便进一步对数据进行加工和处理。在进行数据存储过程中，企业主要会从数据敏感度、数据时效性、企业硬件水平等维度考虑，选择不同的数据存储服务。现阶段，我国数据存储主要包括公有云、私有云以及混合云等方式。

### （1）公有云数据存储

现阶段，公有云是我国数据存储最大的平台。根据工信部《全国数据中心应用发展指引（2020）》数据，截至 2019 年底，我国在用数据中心机架总规模达到 314.5 万架，同比增长 39%。从技术角度看，我国的数据存储和处理市场成熟度相对较高，已经成长出 Oceanbase 等分布式数据库，数据存储能力有了大规模上升。同时，公有云提供商已经从早期的阿里云、腾讯云等几家互联网公司，发展成为超过 60 家企业的供应商群体。

虽然我国公有云规模有了较大上升，但公有云建设仍存在一定的盲目性，需统筹规划公有云平台建设。例如，现阶段我国数据中心平均上架率仅为 53.2%，使用情况较差。

### （2）私有云数据存储

对数据安全性要求较高的企业可构建属于该企业的数据中心并设立防火墙，将数据存储在私有云。对于某些企业如科技研发企业，程序、数据等无形资产是最关键的核心资

产，为防止数据泄露会采取网络隔离的方式，使数据存储在本本地。同时，对于某些时序性要求较高的数据，云端访问在流量峰值期间容易出现性能问题（如网络拥堵），在局域网访问数据的本地化部署可以有效避免外部流量峰值的冲击。由于局域网带宽更大，数据传输也会更快，下载大文件时更有优势。根据 Gartner 预计，我国私有云市场将在 2023 年升至全球第一位，占据全球市场 25% 的份额。

由于企业私有云搭建和维护成本较高，且部署在私有云上的数据较难通过远程访问，因此私有云的广泛部署仍存在限制。未来，我国数据要素市场私有云存储仍主要集中在政府部门、大型企业及高新科技企业等主体。

### （3）混合云数据存储

混合云是公有云和私有云两种数据存储方式的结合。一方面，部分企业依数据分类分级选择混合云存储方式。企业可将自己核心业务及重要数据资源存放在私有云上，而将次要及需要外部访问的数据存放在公有云上，可提高数据资源的整体安全性能；另一方面，部分企业依时序需求选择混合云存储方式，例如，旅游服务企业或零售企业的数据存储需求会随业务淡季或旺季而产生周期性变化，当数据存储需求量较大时可由私有云切换为公有云存储方式。此外，公有云也可作为私有云的存储备份，当发生灾难性事件时将私有云数据转移至公有云，提高了数据存储的抗风险能力。

混合云数据存储方式兼具公有云和私有云的优势，提高了数据存储的安全性和灵活性，降低了存储成本。然而，混



合云存储也提高了数据维护的复杂性。未来，混合云数据存储将探索更加灵活合理的结合方式，规避可能出现的兼容性问题，从而使混合云数据存储被更多企业所采用。

### 3. 数据加工

数据加工是指对企业采集和存储的数据进行筛选和处理，提高数据可用性，为数据资源的挖掘和分析奠定基础，主要包括数据清洗、数据标注、数据审核以及数据融合处理等方式。

#### (1) 数据清洗

数据清洗是指对数据进行校验的过程，目的在于删除重复信息、纠正存在的错误，提升数据质量。不符合要求的数据主要是有不完整的数据、错误的数据、重复的数据三大类。录入后的数据一般是由数据分析师设置某些筛选条件通过程序而非人工完成。现阶段数据治理工具、大多数大数据平台都提供自动化的数据清洗功能，简化数据加工过程。例如，百度的 EasyData 智能数据服务平台可提供图片去模糊、图片去重、图片批量裁剪、图片旋转等功能，利用机器和人工双重检验，保证数据质量。

#### (2) 数据标注

数据标注是指借助特定软件标注工具以人工的方式将图片、语音、文本、视频等数据内容打上特征标签，使计算机通过大量学习这些带有特征标签的数据，最终具备自主识别特征的一种行为。数据标注质量主要关系到三方面：一是数据标注人员的素质问题，二是标注工具的效率，三是企



业的管理能力。数据标注行业主要有两种服务模式：一是数据外包形式，由数据标注公司完成人工智能项目的数据采集、数据标注；二是部署标注平台，由企业内部人员利用本地化的数据标注平台标注完成企业内部数据。例如，倍赛数据利用自身研制的标注工具，汇集标注团队，为人工智能企业提供数据标注服务。现阶段，我国数据标注产业快速发展，已经吸纳就业超过 10 万人，成为欠发达地区发展人工智能、大数据产业的选择。

### （3）数据审核

数据审核主要是指对非结构化数据进行涉及色情、反动、欺诈之类的非法内容审核。一般是机器和人工审核的双重过滤。根据不同业务需求、信息量及风险权衡，人机审核比例会有所不同。对于机器审核，主要利用自然语言处理（NLP）、计算机视觉等技术。现阶段，各企业在敏感性数据审核上主要通过购买成熟的审核产品来实现，例如图普科技、百度、腾讯、依图等公司可提供通过算法模型进行数据敏感性审查的技术。对于人工审核，数据公司主要通过自建团队或选择第三方审核团队，进行内容审核。例如，人民网就利用自身第三方身份，为今日头条、梨视频等平台提供第三方内容审核。

### （4）数据融合处理

数据融合处理是指将多源、多模态数据互相融合、形成可以被挖掘分析的数据集的技术过程。企业在信息化建设过程中由于受各业务系统建设和实施数据管理系统的阶段性、

技术性以及其它经济和人为因素的影响，导致企业在发展过程中积累了大量采用不同存储方式以及不同模态的业务数据。多源、多模态数据采用的数据管理系统有所不同，从简单的文件数据库到复杂的网络数据库，它们构成了企业的异构数据源。数据融合处理将多源异构数据进行融合，对多模态数据进行知识融合与挖掘。目前，不同数据管理平台均开始在平台中集成相关多源异构数据融合和多模态数据挖掘的功能，使企业数据可以被分析和利用，充分释放数据价值。

#### 4.数据流通

数据流通在提升公共决策效率、扩展商业应用场景等方面有着显著的作用。数据流通根据技术方式的不同，可以分为数据开放共享、数据交易、API 技术服务、“数据可用不可见”模式、“数据可算不可识”模式、数据跨境流动等形式。

##### (1) 数据开放共享

根据数据所有权的不同，可将流通数据分为政府数据和企业数据。政府数据开放共享主要有两层内涵，一是各级政府以及政府部门之间的数据共享，二是政府数据对外开放；企业数据开放共享是指企业之间的数据共享与开放。

##### 1) 政府数据开放共享

在政府数据开放共享方面，现阶段已经形成了以“国家电子政务网站”为平台，促进各部委、省市政务数据共享的内部共享体系和以部委、地方政府为主体，对社会进行数据开放的外部数据开放体系。

**内部共享体系方面。**国家电子政务网站接入中央部门和相关单位共计 162 家，接入全国政务部门共计约 25.2 万家，形成了国家共享平台。31 个国务院部门在国家共享平台注册发布实时数据共享接口 1153 个，约 1.1 万个数据项，涵盖个人身份、出生、教育、婚姻、社保等自然人相关信息，企业基本信息、信用信息、资质信息等法人相关信息。国家共享平台累计为生态环境部、商务部、税务总局等 27 个国务院部门、31 个省（区、市）和新疆兵团提供查询核验服务 9.12 亿次，有力支持网上身份核验、不动产登记、人才引进、企业开办等业务。

**对外数据开放方面。**随着国家政策的引导以及各地数据开放体制机制的完善，我国地方政府数据开放平台数量和开放的有效数据集数量呈现爆发式增长。其中，数据开放平台达到 142 个，在四年内增长近 7 倍；有效数据集达到 98558 个，在四年内增长超过 10 倍。各地也不断通过举办开放数据的活动，推动数据开放应用的发展，挖掘政府开放数据价值。2020 年，全国已经有 12 个省市及地级政府举办了开放数据利用互动<sup>7</sup>。

## 2) 企业数据开放共享

现阶段数据开放的主体是政府。与政府数据开放共享相比，企业数据开放共享仍处于较低水平。企业数据开放共享对提升供应链协同效应、提升产业竞争力有明显的带动作用。我国政府正在推出相关政策，鼓励支持企业数据开放共

<sup>7</sup> 复旦大学，国家信息中心数字中国研究院，《2020 中国地方政府数据开放报告》，2020 年。

享。例如，《工业和信息化部关于工业大数据发展的指导意见》提出支持优势产业上下游企业开放数据，加强合作，共建安全可信的工业数据空间，建立互利共赢的共享机制。对于企业数据的开放共享，一方面需要政府出台政策进行鼓励引导，另一方面需要形成良好的市场机制，使数据要素可以进行合理合规流通。

## （2）数据交易

数据交易是指数据买卖双方就数据所有权进行交易。其典型模式为数据交易平台模式，该模式通过构建数据交易平台，吸收第三方数据，撮合数据供给者和数据需求者发生数据所有权交易，并获取交易的服务费。数据交易模式是早期的数据流通模式，但由于我国数据确权相关法律法规不明晰，该模式有较高的数据安全风险，较难保护数据所有者利益，因此近些年来该模式增长开始逐渐放缓。

## （3）API 技术服务

API 技术服务主要利用 API 接口技术，通过程序对元数据进行隔离，在用户发出数据使用请求后，由程序从元数据中抽取、调用数据反馈给用户。目前，API 技术服务模式的供应商主要有聚合数据、京东万象等。例如，API 技术服务头部企业聚合数据已经沉淀了超过 500 个分类的 API 接口，日调用次数已经达到 3 亿次，目前已合作客户逾 120 万家，涵盖智能制造、人工智能、5G 应用等领域。

API 技术服务模式具有安全性、便利性等优势。一是保证数据流通安全性。API 技术服务公司提供 API 接口，数据

依然存储在客户的数据中心，用户发出的数据验证请求通过 API 接口传递给供应商，并将供应商的验证结果反馈通过 API 接口反馈给用户。二是 API 技术成熟、使用门槛低。API 技术基于互联网的应用正变得越来越普及，数据 API 具有很好的可控性，数据 API 是按需消费，无需购买全量数据，同时很大程度地减少数据质量的争议。三是在技术上易监管。因为用户在进行数据调用时，并不能获取全部元数据，只能调用业务相关数据，因此可以对相关数据调用进行记录，并形成一套国家自主可控的标准化、安全可信的技术来实现 API 的网络管控。

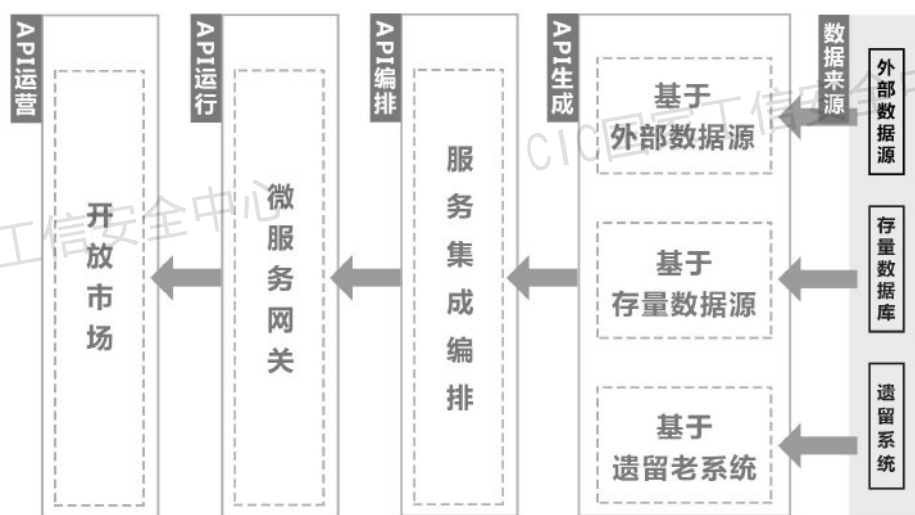


图 7 API 技术服务模式图

来源：聚合数据

#### (4) “数据可用不可见”模式

“数据可用不可见”模式是指通过隐私计算技术，实现数据在加密状态下被用户使用和分析。隐私计算，广义上是指带有隐私机密保护的计算系统与技术，能够在不泄露原始数据的前提下对数据进行采集、加工、分析、处理与验证，

涉及数据的生产、存储、计算、应用等数据处理过程。隐私计算能够在保证数据所有者权益、保护用户隐私和商业秘密的同时，充分挖掘发挥数据价值。

隐私计算重点强调同时处理隐私和商业机密以及处理复杂数据的算法能力，类似的概念有隐私计算、隐私安全计算、机密计算等。在具体的商业场景中，隐私计算具有在保护数据安全的情况下处理海量数据的基本能力。其技术方案主要有安全多方计算、联邦学习、可信执行环境等。

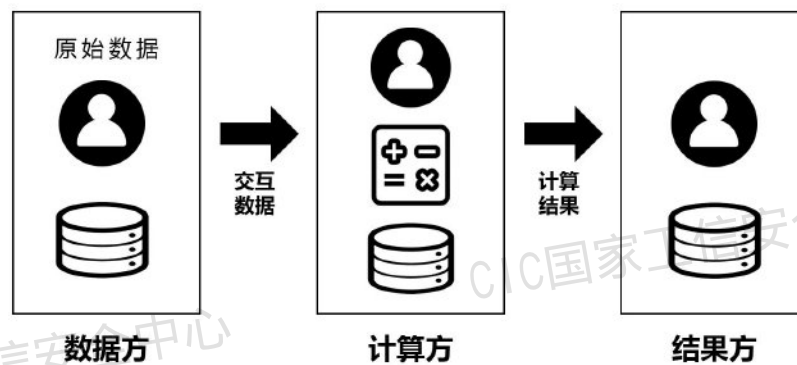


图 8 隐私计算原理

来源：国家工业信息安全发展研究中心

### 1) 安全多方计算

安全多方计算 (MPC) 是指在无可信第三方的情况下安全地进行多方协同计算。在一个分布式网络中，多个参与实体各自持有秘密输入，各方希望共同完成对某函数的计算，要求每个参与实体除计算结果外，均不能得到其他参与实体的任何输入信息。安全多方计算包含的基础技术有很多，比如同态加密、秘密分享、不经意传输、混淆电路等。

基于混淆电路的协议更适用于两方逻辑运算，通讯轮数固定，但是拓展性稍差。在另一类基于秘密分享的安全多方

计算中，数据输入和计算中间值都会以“密文分片”的方式存在。秘密分享技术可以把隐私数据切割为 2 份或更多份，并将随机分片分发给计算参与方，这个过程保护了数据隐私又允许多方联合对数据进行计算。之后，就可以利用分片间存在的同态计算性质来实现在分片上计算并重建得到隐私数据计算结果的目的。安全多方计算技术主要可应用于联合风控、联合营销、联合商业智能等业务应用场景。当前探索安全多方计算技术的厂商主要有蚂蚁集团、阿里云等。

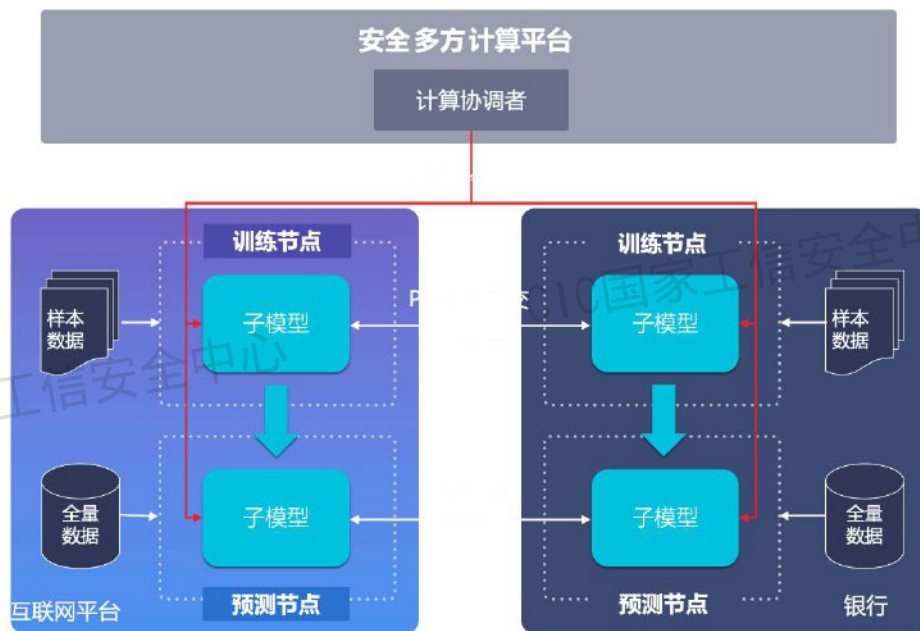


图 9 安全多方计算架构图

来源：蚂蚁科技集团

## 2) 联邦学习

联邦学习（Federated Learning）是指基于多方数据进行联合建模。各自原始数据不对外输出、由中心方进行协调的建模，都可成为联邦学习。联邦学习更多地用于解决联合建模的业务问题，比如信贷风控中的常用的逻辑回归建模评

分、精准营销中的常用的 XGBOOST 分类等建模。比较有代表性的企业和平台包括蚂蚁集团、镭崑科技、富数科技（Avatar）、平安科技（蜂巢）、数牍科技等。

联邦学习的技术原理在于参与学习的各方机器上部署的客户端从服务器端下载现有模型以及参数，而后根据各自所有的数据对模型的参数进行更新，并把结果传回客户端并更新模型。使用该方法不会与服务器交互原始数据，因此保证了原始数据不会出库。另外，原始数据或者梯度在使用前都可以使用某种方式进行加密或加噪，来增强安全性。目前，联邦学习已经开始用于医疗、基因分析等领域。

需要注意的是，无论是安全多方计算还是联邦学习，在涉及个人隐私信息的事情，都需要得到法律法规所规定的使用场景授权。

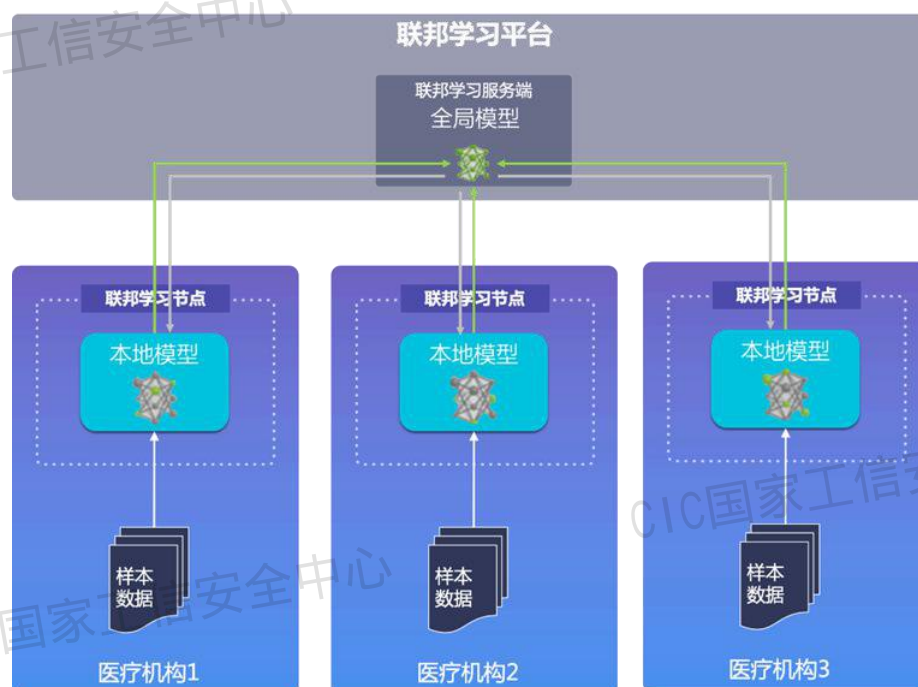


图 10 联邦学习医疗行业应用

来源：蚂蚁科技集团



### 3) 可信计算环境

即 TEE，是指主处理器的安全区域，它保证装载在内部的代码和数据在保密性和完整性方面受到保护。TEE 作为一个独立的执行环境，提供了完整的安全特性，如执行的独立性、使用 TEE 执行的应用程序的完整性以及其数据的机密性。为了满足对隐私的紧迫需求，越来越多的云提供商正在转向采用 TEE 技术方案。Gartner 在其报告中将 TEE 列为 30 多种“关键安全技术”之一。2019 年，百度、阿里巴巴及腾讯等公司联合国外大型科技企业启动了“机密计算联盟”项目，联盟成员共享 TEE 开源工具，实现 TEE 标准互认。

### 4) 数据沙箱技术

数据沙箱技术是将调试环境和运行环境分离，数据分析师在调试环境中使用样本数据调试代码，然后将代码发送到运行环境中运行全量数据，从始至终数据分析师无法接触全量数据，从而达到保护数据隐私的目的。数据沙箱技术实现了“数据不动程序动，数据可用不可见”，既确保原始数据不泄露，促进合法合规的数据开放，又充分发挥了数据的最大价值。

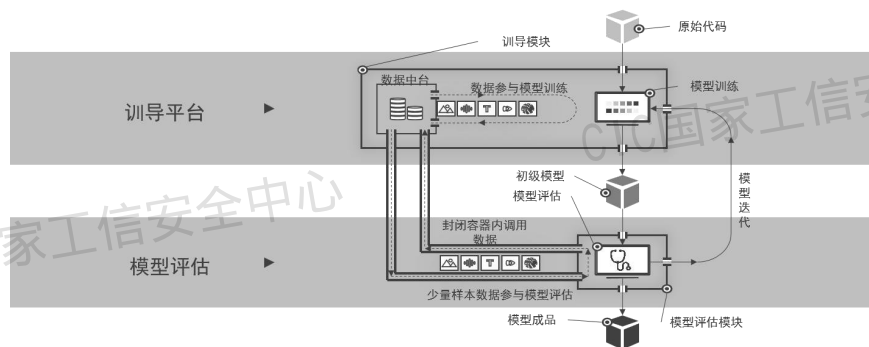


图 11 数据沙箱技术图

来源：倍赛科技

实际数据应用场景中，对安全性和计算性能有不同要求，以上的数据技术并不是独立的，所以在实际应用中需要将以上的数据安全技术进行有机的融合方能形成有效的技术解决方案。

#### (5) “数据可算不可识”模式

即数据可证去标识。在数据规模较大（比如大于百万条记录）或者对计算性能要求较高时，基于安全多方计算或联邦学习的技术可能难以满足性能或者实时性需求。在大规模或实时性要求较强的数据分析场景下，可证去标识是目前唯一能同时满足隐私合规要求和计算性能要求的新技术。这种技术确保数据去标识后，数据接收方无法重新识别或者关联个人信息主体。可证去标识首先对参与计算的多方数据可信去标识管控，确保所有计算基于去标识化数据展开；其次构建集中式的可信计算环境，通过对试图关联或还原个体身份的高危行为进行拦截，实现挖掘过程中个人数据“可算不可识”；最后在结果输出阶段对输出数据进行原始数据拥有主体及用户的双重确权，实现了价值输出时各方权益可保障。该方案可与现有大数据技术栈无缝集成，且采用集中式计算规避了跨网延时成本，可支持大规模数据的高性能分析和计算，且计算场景受限较小，支持几乎所有类型的数据分析和建模，较好地平衡了个人隐私权保障、数据处理规模和业务实时性，适用于对计算环境存在信任基础的多方大规模数据挖掘场景。

## （6）数据跨境流动

近年来，数据跨境流通已经成为全球经济的重要驱动力。根据美国著名智库布鲁金斯学会的相关研究，过去十年间，全球数据跨境流动对全球经济增长贡献度超过10%。我国数据要素市场需建立健全法律法规、完善监管制度以维护数据主权。

**数据跨境立法方面。**2016年11月7日，全国人大常委会表决通过《中华人民共和国网络安全法》（简称“《网络安全法》”），并于2017年6月1日起正式实施。该法是我国网络安全管理领域的基础性法律，对关键信息基础设施的运营者数据跨境传输的义务进行原则性规定。然而，该法规中涉及数据跨境流通的条款和规定相对比较笼统，而其他相关的配套法规目前尚未出台（仍处于征求意见阶段）。这一立法现状导致了我国目前在数据跨境流通监管方面的执法力度仍然相对宽松。

**数据跨境政策制度方面。**2019年6月13日，国家互联网信息办公室发布《个人信息出境安全评估办法（征求意见稿）》（简称“《办法》”）。《办法》指出国家网信部门负责统筹协调数据出境安全评估工作，具体数据出境安全评估工作由各行业主管或监管部门负责，而将组织开展个人信息出境安全评估工作的职责统一归于省级网信部门。

**地方试点方面。**现阶段关于数据跨境流通的政策性文件主要为商务部印发的《全面深化服务贸易创新发展试点总体方案》（以下简称“《方案》”）。《方案》指出要在条件

相对较好的试点地区开展数据跨境传输安全管理试点。其中，“探索跨境数据流动分类监管模式，开展数据跨境传输安全管理试点”这一试点任务，由中央网信办指导并制订政策保障措施；北京、上海、海南、雄安新区等试点地区负责推进。

个人信息出境安全评估流程图

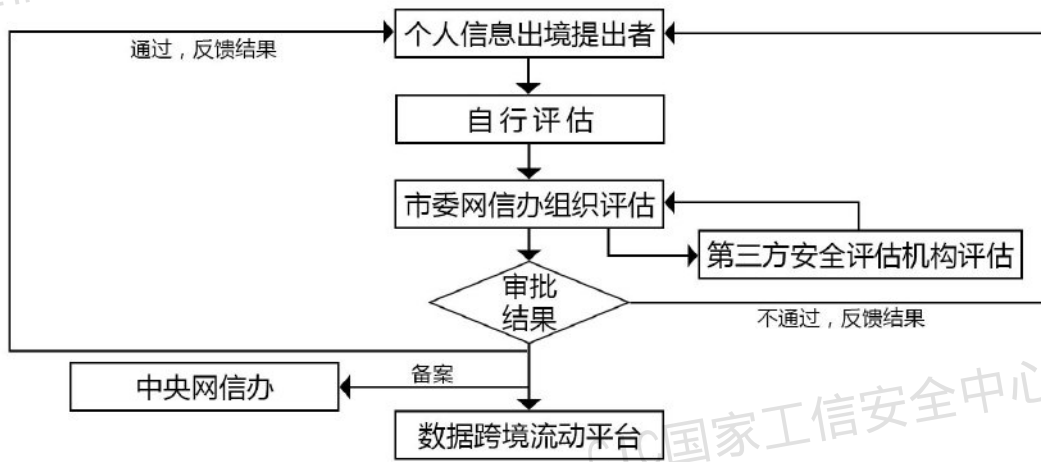


图 12 个人数据出境安全监督

来源：商务部《全面深化服务贸易创新发展试点总体方案》

## 5. 数据分析

数据分析主要是指最大化地开发数据的功能，发挥数据的作用。从数据来源来讲，主要包含内部数据分析和外部数据分析。

### (1) 内部数据分析

企业内部数据分析主要是针对企业内部经营分析，进行检测预警。内部数据分析可以分为离线数据分析和在线数据分析两种。

一是**离线数据分析**。用于较复杂和耗时的数据分析和处理，一般构建在云计算平台之上，如开源的 HDFS 文件系统和 MapReduce 运算框架。Hadoop 机群包含数百台乃至数千台服务器，存储了数 PB 乃至数十 PB 的数据，每天运行着成千上万的离线数据分析作业，每个作业处理几百 MB 到几百 TB 甚至更多的数据，运行时间为几分钟、几小时、几天甚至更长。日常的离线数据分析可以通过 EXCEL 等软件完成，较复杂任务则可通过商业智能领域的 Oracle 以及 Yonghong Z-Suite BI 套件等国产专业软件完成。

二是**在线数据分析**。在线数据分析也称为联机分析处理，用来处理用户的在线请求，它对响应时间的要求比较高（通常不超过若干秒）。与离线数据分析相比，在线数据分析能够实时处理用户的请求，允许用户随时更改分析的约束和限制条件。与离线数据分析相比，在线数据分析能够处理的数据量要小得多，但随着技术的发展，当前的在线分析系统已经能够实时地处理数千万条甚至数亿条记录。在线数据分析系统一般构建在云计算平台的 NoSQL 系统上。在线数据分析广泛应用于移动互联网应用中，如在微博、网约车软件、短视频平台、电商平台、搜索引擎中均有广泛应用。在线数据分析提供厂商主要包括华为、星环、阿里等。

## （2）外部数据分析

外部数据分析主要是针对企业外部数据，进行同业竞对分析、营销投放检测等。外部数据分析主要通过抓取各个企业数据，形成不同领域的行业数据库，并通过对数据库产品

进行分析，为用户提供咨询服务。企业通过脱敏数据分析研究服务既可以帮助客户了解市场发展趋势和行业竞争格局，通过理解用户特征和全景画像优化自身运营效率，也可以帮助客户前瞻性地发现市场机会。伴随着产业竞争力不断加大，外部数据分析企业会有较大规模的增长。现阶段外部数据分析供应商主要包括 Questmobile、艾瑞数据、易观数据等。

## 6.生态保障

数据要素市场除包含数据采集、数据存储、数据加工、数据流通等直接面向数据要素进行处理的环节外，还需数据要素市场各方主体为数据交易流通提供有效保障，构建良好的市场生态。数据要素市场生态保障主要包括数据资产评估、登记结算、交易撮合、争议仲裁及跨境流动监管等环节。

### (1) 数据资产评估

数据资产评估是指通过第三方评估机构或企业对数据所有者在生产、运营过程中所产生的数据进行内在价值和使用价值的评估，以便为数据要素流通交易提供基础性参考。例如，中关村数海数据资产评估中心与 Gartner 合作推出全球首个数据资产评估模型，为企业提供数据资产登记确权 and 评估服务。

目前，数据资产评估多采取市场法、收益法、成本法等资产评估方法。由于数据资产存在无形化、虚拟化的特性，因此数据要素流通交易直接产生的收益或所需的成本存在较难核算的问题。同时，我国数据要素市场仍处于培育期，交易规模尚小，给数据资产评估提供的范例较为有限。因此，

数据资产评估整体上存在标准模糊、执行困难的问题，亟须建立数据要素市场资产评估机制架构。

## （2）登记结算

登记结算是指在数据作为资产的前提下，为数据所有者及采购方提供名册建立与维护、数据交易结算等服务。登记结算机构为数据交易双方建立数据持有及交易资质名册，提高数据交易效率，规避数据交易风险。同时，登记结算机构通过现金结算、票据转让及转账结算等方式为数据所有权变更、API技术服务采购等交易过程提供交易场所。例如，贵阳大数据交易所创立大数据登记确权结算服务，将数据视为和房产、股标等一样的实际资产，通过交易所的数据平台，尝试登记数据所有权，然后对数据的使用权、运营权等进行公开竞价，以实现数据的登记确权及变现。

由于数据要素存在容易获取、高复制性、交换或散布成本低等特点，使得国内数据登记结算机构面临着数据被攻击、窃取或盗用的风险，一旦产生数据泄露将直接造成数据所有者或购买方的利益损失。因此，数据登记结算机构需进一步完善“数据可用不可见”等技术的落地应用，相关部门亦需制定数据登记结算监管机制，保障数据登记结算过程的安全性。

## （3）交易撮合

交易撮合是指为数据交易双方提供给交易信息查询、信息匹配、供需对接及交易竞价等第三方服务。交易撮合机构将数据持有者及采购者等多方信息整合起来，并通过分析与

评估将有效信息反馈给交易双方，以促进数据交易的达成。例如，华东江苏大数据交易中心在其官方网站开通供需撮合平台，通过咨询匹配、需求分析、定制方案与双方协作等四个步骤为各类经济企业提供一站式服务，搭建起数据所有者与采购方的沟通桥梁。

囿于数据要素市场规模，目前我国数据交易撮合机构存在着数量较少、单打独斗的问题。同时，仍有相当数量的企业对数据资源重要性的认识程度不够，参与数据要素市场交易的积极性有待提高。因此，我国数据要素市场需继续提升数据开放程度，扩大交易撮合机构规模，盘活数据要素市场资源。

#### （4）争议仲裁

争议仲裁是指当数据交易双方针对数据交易过程及结果产生民事争议时，对数据交易争议进行裁决并协调双方矛盾。数据交易争议双方在争议发生前或发生后依自愿原则将争议提交至争议仲裁机构，争议仲裁机构依据相关法律法规对争议的是非曲折作出判断，争议双方有义务执行争议仲裁结果。例如，深圳大数据仲裁中心为数据交易提供电子证据固化、在线公证保全和网络裁判等权威数据证明及争议解决方案，为数据要素市场创造法制化、公平化、便利化的交易环境。

作为新兴领域，我国数据要素市场在解决争议仲裁问题是还面临着法律法规存在空白、监管制度不健全、判例较少等困境。同时，数据要素无形化、虚拟化所带来的权属界定



困难等问题也给数据交易争议仲裁从业人员带来挑战。因此，行业主管部门需进一步完善数据要素市场监管体制机制，培育数据交易争议仲裁专业人才，为数据要素流通交易提供保障。

#### （5）跨境流动监管

跨境流动监管是指对跨越国界或产生第三国访问的数据传输、处理及存储过程进行监督管控，以维护本国数据安全。跨境数据流动监管机构通过构建系统化制度体系，开展数据跨境流动双方信息采集及分析，对数据跨境流动进行审查和管理。例如，北京国际大数据交易所依托数字贸易试验区积极探索跨境数据安全流通，吸引跨国企业和国际机构加入，愿景构建立足中国、面向国际的国家级数据资源流通与监管生态体系。

由于各国数据立法进展不同及对数据要素重视程度存在差异，因此数据跨境流动监管面临着标准不统一、执行困难等问题，给国家数据安全带来挑战。因此，我国数据要素市场应完善数据跨境流动监管法律法规，扩大数据开放共享程度，积极融入国际数据跨境流动体系，并参与数据跨境流动国际法规制度制定，提升国际话语权。

#### （四）运行机制

数据要素市场需要在政府及第三方机构的监管下，制定合适的市场定价机制及收益分配机制，以实现数据从数据所有者到数据购买方的流通交易，并使数据所有者、购买方、平台方均获得一定收益，从而保证数据要素市场实现可持续

发展。

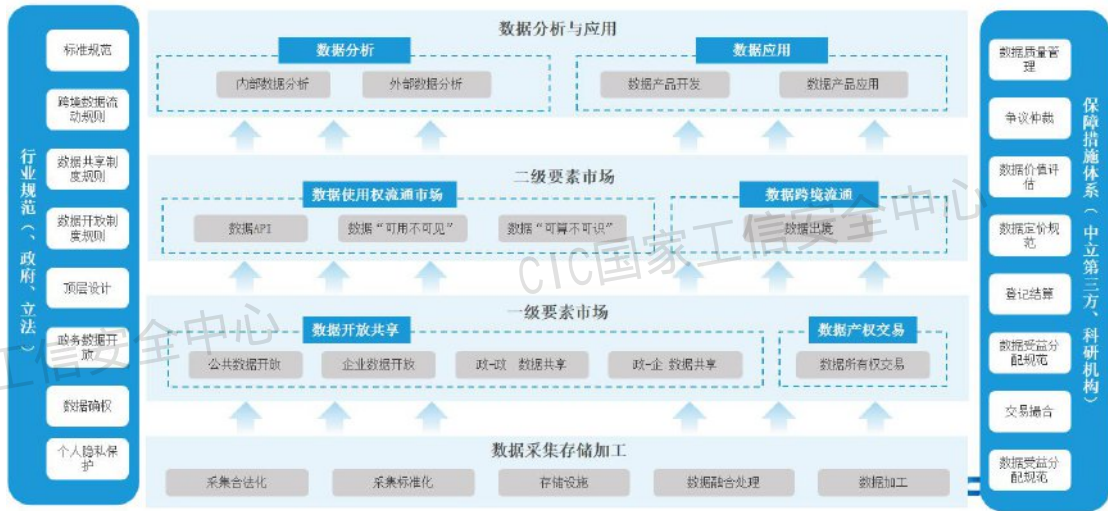


图 13 数据要素市场机制一览

来源：国家工业信息安全发展研究中心

## 1. 定价机制

数据要素市场定价基于对数据自身价值的评估。目前资产价值评估主要包括市场法、收益法及成本法等方法，而由于数据自身的无形化、虚拟化等特性使得上述资产价值评估基础理论方法存在局限性。

对于市场法，其基于数据资产在市场中的交易价格计算作为交易对象的数据所代表的价值，从而进一步为数据要素市场交易提供价格参考。市场法的优势在于通过交易价格易于得到数据价值判断的依据，且数据价值与交易价格呈正相关。然而，数据交易作为新兴市场交易量尚小，往往不能为市场定价提供指导。同时，数据价值评估反作用于市场交易定价，新兴市场存在的不规范交易等情况使得数据要素市场定价机制陷入“先有鸡还是先有蛋”的问题中。对于收益法和成本法，其基于数据要素市场中由于数据交易而带来的利润或需要的成本。收益法与成本法的优势在于通过利润或成

本可以体现出数据价值的本质，并为数据价值提供具象化的表征。然而，数据要素持有方往往难以界定由于数据交易所带来的利润或创造数据价值所需的成本，利润或成本的量化过程浮动范围较大，影响数据交易定价参考。

基于上述资产价值评估理论，同时考虑到数据资产的无形化及虚拟化特性，国内数据要素市场在实践中探索定价机制，目前主要存在数据所有权交易定价和数据使用权交易定价两大类，又可细分为第三方平台预定价、协议定价、拍卖定价、按次计价（VIP 会员制）及实时定价等五种大数据交易定价机制。

#### （1）数据所有权交易定价

数据所有权交易类型是指数据交易双方直接产生数据所有权属变更的交易，如数据集的交易等。数据所有权交易定价可参考资产评估方法，确定交易过程中数据资产的价值。

**一是第三方平台预定价。**如果数据卖方无法确定数据产品的具体价格，大数据交易平台可以委托大数据交易相关第三方专业人员进行评估定价。第三方专业人员基于大数据交易平台特点，利用数据质量评价指标（数据量、数据种类、数据完整性、数据时间跨度、数据实时性、数据深度、数据覆盖度、数据稀缺性等）给出评价结果，并根据评价结果和同类同级数据集/产品的历史成交价给出一个合理的价格区间，数据卖方基于此价格区间在交易前对交易数据进行再定价。

**二是协议定价。**即数据买卖双方协商交易价格。大数据交易平台用于促进数据交易双方的沟通，使数据交易双方达成对该交易数据交易价格的一致认可，并完成数据交易的最终成交。这种定价方式目标性强，在不违反政策限制的情况下，双方对价格的把控自由度大，沟通机会多，成交率高。但这有可能会使整个协议过程形成一个漫长的博弈过程，增加时间成本。

**三是拍卖定价。**即在一个卖方和多个买方之间经过拍卖而确定价格。现今各大数据交易平台交易的数据至少都经过了脱敏等预处理，这样数据交易双方无法对所成交数据的最终使用价值进行准确的定价。于是，交易双方会倾向于通过拍卖方式对交易数据进行最终的定价。拍卖定价方式是强调对数据产品及服务的一次性交易，是针对数据产品或技术所有权的直接交易。

## (2) 数据使用权交易定价

数据使用权交易类型是指数据交易双方不产生数据所有权属的变更，而主要通过调用数据集达到训练算法模型等目的，如 API 技术服务等。数据使用权交易定价更多地将数据交易视为服务形式，并参考服务业定价机制。

**一是按次计价机制。**基于数据调用模式，数据买方每调用一次数据就付费一次。这种方式实际上出售的仅仅是数据的使用权。这一定价方式的计价标准由大数据的提供者定价，大数据交易平台或大数据技术服务提供商作为中介对数据进行传输。部分企业在按次计费的基础上延伸了 VIP 会员

制，即购买 VIP 会员即可获得免费接口一定时间范围内的调用次数。按次计价定价方式强调对 API 的多次调用，是针对数据产品或技术使用权的多次交易，因此此定价方式与权利归属中的交易权方式结合，能更好地体现大数据交易平台的盈利模式。

**二是实时定价机制。**依据数据的样本量和单一的数据指标项价值，通过交易系统自动定价，价格实时浮动。采用实时定价的数据商品价格受市场环境和市场供求关系的影响，当市场供需实时变化时，数据价值也实时波动。此外，数据所包含的商品价值和使用价值会随着时间变化出现波动，将会直接影响交易数据的最终交易定价。若交易数据处于市场需求低，数据价值低的时段，数据交易价格也会较低。

## 2.收益分配机制

收益分配机制是指基于数据权利归属和定价方式的数据价值实现机制，大数据交易平台和数据卖方的价值实现是大数据交易的关键。

### （1）数据交易平台收益分配机制

目前我国典型政府类大数据交易平台，大多数都扮演着数据交易中介的角色，主要交易来源于不同数据所有者提供的数据。我国大数据交易平台的收益分配机制主要有交易分成和保留数据增值收益权两种。

**一是交易分成收益分配机制。**在数据交易完成后大数据交易平台与数据卖方按约定好的比例分成。大数据交易平台作为数据交易中介会在促成数据所有权或使用权交易后收

取相应的中介费用。如贵阳大数据交易所 4:6 与数据供应商分成，同时视具体数据价值，适当对数据买方进行收费。大数据交易分成机制是目前国内大数据交易平台普遍采用且符合市场规律的收益分配机制。

**二是保留数据增值收益权分配机制。**即大数据交易平台对数据保留增值收益权并以此为基础收费的方式。数据包含丰富的价值，大数据交易平台作为数据中介机构需要在交易前准确预测数据交易后能否产生增值价值并保留数据增值收益权。

## (2) 大数据交易卖方收益分配机制

大数据交易卖方是数据所有者，根据权利归属和定价方式的不同，其收益分配机制主要包含一次性交易所有权、多次交易使用权和保留数据增值收益权三种机制。

**一是一性交易所有权收益分配机制。**即在数据交易中一次性转移数据占有权、使用权、处分权、收益权。这一模式主要适用于协议定价、拍卖定价方式。协议定价方式能够形成数据交易双方讨价还价的博弈，协调得出一个交易双方认同的交易价格。在拍卖定价方式下，数据卖方虽然根据自身对数据价值的评估给出了起拍价及加价幅度等相关拍卖规则，但是实际最终定价的权利属于参与竞拍的多个买家。所以面对协议定价和拍卖定价方式下的一次性交易所有权收益分配机制，数据卖方对最后定价权利很被动，相应地压缩了利润空间。

**二是多次交易使用权收益分配机制。**即不将数据所有权

一次性转移，只针对数据使用权进行反复多次的交易，进而带来更多的收益。数据交易双方约定只针对数据使用权进行交易，数据卖方能够反复对数据进行交易以获取更多的利益，尤其是在按次计价定价方式或 API 技术服务模式下。因此，多次交易数据使用收益分配机制是目前数据服务商进行数据交易的首选。但由于数据产品的低成本可复制性、便捷可传递性，在该模式下，数据卖方如何对交易数据进行安全、保密、可控传递，避免数据被大规模复制使用成为这一收益分配机制实现的关键。

三是保留数据增值收益权分配机制。数据卖方更清楚数据的来源和数据采集、处理、分析过程，因此更能直接准确地评价数据的价值，并预测数据交易后是否有增值收益的可能性。基于相关优势，数据卖方能更准确地判断是否需要保留对收益权的占有，并按多少比例进行合同约定。

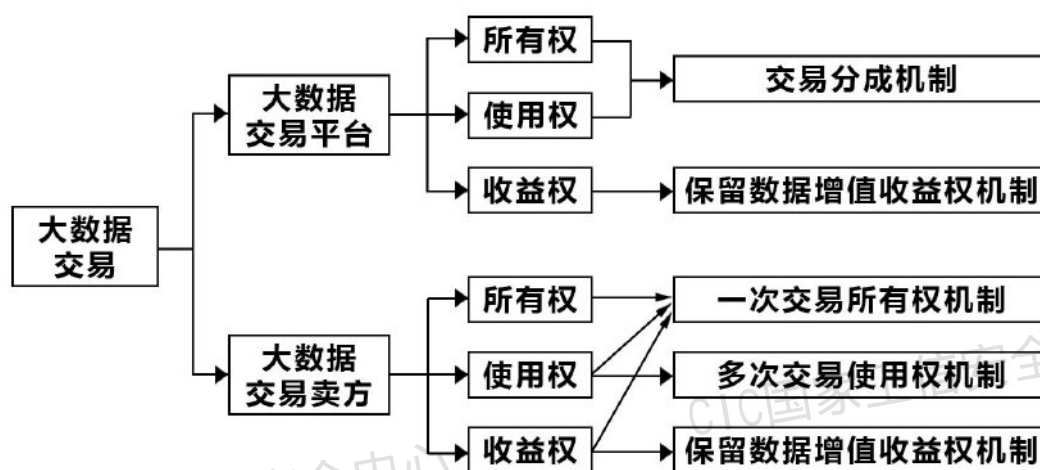


图 14 数据要素市场收益分配机制

来源：国家工业信息安全发展研究中心

### 3. 政企合作机制

现阶段我国 90% 的可用数据均是政府所有，因此政府如

何激活数据要素、推动数据服务，是数据要素行业繁荣的关键。各地在数据服务方面，也进行了多种尝试，但由于政府机关与国企对于市场敏感程度较低、对于数据行业了解程度不高、技术积累薄弱，因此以纯政府行为推动的数据服务项目，成功案例极少，因此当前政府在推动数据服务方面，仍是以企业合作为主。目前主流政企合作模式可根据企业参与程度不同，分为技术服务、企业代理运营模式与政企合资模式。

### （1）以技术服务为主的合作模式

以技术服务为主的合作模式是指企业根据政府的需求通过计算机与大数据技术为其提供数据服务产品（例如搭建平台和系统），系统交付后即合作终止，不参与后续的数据运营，仅作为技术合作方进行产品技术的更新，以及平台的运维。这种模式的优点在于企业的整体模式较轻，可以在多地进行大规模复制，二次开发成本较低，利润率可观。但该模式缺点为单一项目额度较低，同时后期企业参与程度较低，产品与运营分离导致数据服务产品的实际效用也较低，很难形成示范级的项目。

### （2）企业代理运营模式

长期合作中的企业代理运营模式是指企业根据政府的需求为其提供产品、运营等一套解决方案，但不占有股权，在管理上收政府或相关单位上的指导，类似“官督民办”形式。该模式在早期大数据交易所建立时，有大量的案例，如华东大数据交易所，就采取该模式。但伴随着该模式效率较



低，各地大数据交易所，以及数据运营公司都在积极吸引社会资本参与，并且有一些数据交易所，已经完全私有化。

### （3）政企合资模式

政企合资模式是指企业以技术入股方式占有股份，负责产品、运营等，提供一整套解决方案，政府以资金入股，占主要股权，企业与政府双方强绑定。其优点是可以盘活数据资产，将是未来合作的主流方向。投资为亿级规模，该模式可以在政府对数据进行掌控的情况下，保证企业的活力与积极性。现阶段已经有一些企业，通过与地方合资成立公司的形式，通过运营经验赋能的方式服务于当地业务。

## 四、我国数据要素市场发展面临的问题与挑战

### （一）数据权属界定问题尚不明晰

习近平总书记在中共中央政治局集体学习上指出，“要制定数据资源确权、开放、流通、交易相关制度，完善数据产权保护制度”。近年来，西方国家出台了多部与数据相关的专门法规，如欧盟《通用数据保护条例》、日本《人工智能、数据利用相关签约指南》等，相关法规均对数据权属等问题进行了系统界定。我国《中华人民共和国民法典》等法律虽然规定了须对个人信息和数据进行保护，但相关立法对数据要素市场中的数据权属问题一直未作出正面回应。

数据权属界定不明确造成数据在流通、交易、使用过程中的可解释空间大，导致市场规范性变差。例如，根据《腾讯微信软件许可及服务协议》，微信帐号的所有权归腾讯公司所有，而用户只享有使用权，不得赠与、借用、租用、转让或售卖微信帐号。而在重庆新世纪百货与腾讯公司的名誉权纠纷一案中，腾讯公司又解释称账号所有权属于用户，腾讯公司本身并没有占有处分权和所有权。数据权属不确定问题使得针对微信账号数据的挖掘、开发等市场行为处于司法实践的灰色领域。此外，数据权属不确定问题还给相关执法带来困难，间接造成数据交易违法成本降低。例如，大量数据集在黑市进行交易，数据隐私泄漏问题屡见不鲜。因此，相关部门应尽快建立数据权属界定方法体系，分离数据所有权与使用权，为数据要素的流通交易奠定基础。

## （二）数据要素市场存在安全风险

近年来，数据隐私泄漏等公共安全事件频发，折射出数据要素市场存在的技术安全风险问题。数据要素在采集、存储、加工和流通过程中，数据安全防护更加困难，容易受到非法势力攻击和窃取，在进行市场化运作时，数据要素流通的环节复杂，与传统要素相比更加容易发生泄漏的问题，因此数据要素流通风险防治一直是影响制约要素流通的关键问题，对数据安全和隐私保护的要求更高，甚至直接影响数据要素的流通以及流通时的价值。

数据要素流通过程中，各种基础性、人为性或合规性的原因均可能导致数据发生泄漏，而一旦发生泄漏，以目前的技术，较难对数据进行追溯。区块链等技术虽然在一定程度上可以解决数据溯源，但可用性与落地性仍有一定距离，存在数据被多次盗用的风险。安全多方计算及联邦学习等新兴技术在促进要素交易流通方面虽然有巨大的潜力，但企业对其接受度还有待提高。未来，针对数据要素流动核心风险的防护技术依然有待加强。

## （三）数据流通交易机制有待完善

根据数据所有权的不同，可将流通数据分为政府数据和企业数据。一是政府数据方面，其流通多为“一把手”工程。数据流通需要地方党政领导充分意识到大数据重要性，协调各部门关系，合理赋予大数据相关部门权利，利用刚性制度约束，促进各部门数据打通、开放。目前，贵州、浙江等以信息技术立省的地方，由于党政一把手重视，相应刚性制度

完善，数据流通工作开展迅速。以贵州为例，贵州省大数据管理局掌控各平级部门的信息化预算，以此推动政府数据流通。但仍有部分地方大数据主管部门缺乏配套的刚性制度约束，无法实现政府内部的数据流通管理。二是企业数据方面，我国企业数据流通方式以“数据交易所+数据公司”模式为主，但根据相关统计数据，各地的数据交易所已经有 14 家处于停止运营或半停止运营状态。贵州大数据交易所作为最大的数据交易所，年交易额仅 3 亿。这主要是因为数据是非标品，因此数据交易所作为平台交易商，从基因上无法解决数据质量和数据定价问题。虽然数据公司通过提供标准化的数据产品，可以在一定程度上解决数据质量和数据定价问题，但是仍存在数据安全性风险、企业定制化采标数据交付周期长且数据成本高等问题。

#### **（四）数据监管法律法规存在空缺**

目前我国有关数据方面的法律法规、部门规章主要集中在政府数据开放、个人信息保护和数据交易流通等方面，但在实际操作中缺乏具体细则。一是在政府数据公开方面，出台了《政府信息公开条例》等系列政策文件，但并没有针对政府数据公开的范围、数据质量评估等方面的具体细则，各省市结合自身实际制定的政府数据资源共享条例等相关规章制度也并不统一，部门间共享数据的权利责任问题也没有厘清，这些都制约了公共数据资源的进一步开放共享。二是在跨境数据流动管理方面，在《网络安全法》中提到“关键信息基础设施的运营者在中国境内运营中收集和产生的个

人信息和重要数据因业务需要，确需向境外提供的，应当按照国家网信部门会同国务院有关部门制定的办法进行安全评估”，但没有出台数据跨境安全评估细则和操作方法。三是在数据交易和流通方面，《网络安全法》要求网络运营者“未经被收集者同意，不得向他人提供个人信息”，但规定了例外情形，即“经过处理无法识别特定个人且不能复原的除外”，这为个人数据的流通提供了法律上的依据，但依然没有制定出台数据交易和流通的专门性法律法规。这种法律法规的不清晰导致很多有意义有价值的数据要素流通被“授权墙”阻断。比如，有些数据分析任务在实际操作过程中无法取得所有单个数据主体的授权，在此类场景下，使用去标识化和防身份关联技术可以有效保证数据主体不被识别或关联，但目前的法律法规尚未明确这些技术是否可以在适当情况下替代数据主体授权。

## 五、我国数据要素市场未来的发展趋势

### （一）数据要素市场主体趋于多元化

随着数据要素市场不断培育壮大，政府、企业、数据交易机构、科研院所及产业孵化场所等数据要素市场主体逐步投入市场运作，进一步促进数据要素市场发展。

一是政府将提升数据开放共享水平，促进数据要素市场交易流通。由于缺乏主动开放数据的意识和动力以及部分信息存在安全性要求，社会数据占比高达 80% 的政务数据尚未充分体现公共服务属性<sup>8</sup>。未来，政府部门将合理推动高价值数据（如信用、卫生、医疗、企业登记、行政许可、交通、就业、社保等）在市场中开放共享，为企业、个人、社会提供数据资源，进一步促进数据红利充分释放。同时，作为公共服务及行业管理部门，政府也将加强数据安全监管法律法规建设，设立政府数据开放共享负面清单，构建个人信息保护制度框架，加强跨境数据流动监督管理，统筹推进各数据要素市场主体协同发展。

二是企业将充分发挥市场主体作用及技术优势，推动数据要素市场及大数据产业蓬勃发展。一方面，企业尤其是拥有数据资源的互联网巨头更多地参与数据交易环节，将数据集或数据产品通过所有权或使用权变更的方式进行交易，实现企业盈利与发展。另一方面，企业通过提供数据产品和服务，更好地盘活数据要素市场资源，加快产业数字化转型进程。大数据企业收集经济部门生产、运营、管理过程中

<sup>8</sup> 李纪珍，钟宏等. 数据要素领导干部读本[M]. 国家行政管理出版社，2021 年.

的数据，通过大数据技术产生智能化生产工具，提高企业管理与协作效率，促进产业生产过程可视化、精细化与个性化；大数据企业为政府部门和各类市场主体提供数据采集、数据加工及纸质档案数字化等数字化服务，助力打通地区间数据壁垒，促进国家治理体系与治理能力现代化。

三是数据交易机构、科研院所及产业孵化场所等第三方机构将为数据要素市场良性运转提供保障。大数据交易所、数据经纪商等数据交易平台将积极发挥作用，参与构建数据价值评估体系，通过数据资产评估、登记结算、交易撮合及争议仲裁等形式保障市场行为自发有序进行；科研院所将积极研发数据交易流通过程核心技术，探索新型交易模式及安全保障制度，为数据要素市场运转提供技术支撑；产业孵化场所等将培育更多合格的市场主体，丰富大数据供给侧企业规模，实现数据要素市场多元化发展。

## （二）“数据可用不可见”成为技术趋势

随着大数据、人工智能、隐私计算、区块链等技术的不断发展完善，数据要素市场与新技术的融合发展逐渐深入，数据流通与交易过程更加安全有序。

一是数据交易流通过程中更加注重隐私信息保护。基于“数据可用不可见”技术的数据交易流通模式是未来大数据交易的主要发展方向。该技术通过隐私计算对信息进行加密与共享，实现数据按用途与用量进行使用且在使用过程中不被泄露。隐私计算解决了长久以来在数据交易流通中较难规避的保护敏感信息（包括个人隐私、商业机密）不被泄漏、

不可反推问题。例如，阿里云与蚂蚁集团联合通过联邦机器学习在医疗行业中与医疗机构在联合运营、联合诊断增强等业务应用场景中展开了试点实践。

二是数据交易流通中避免数据资产被数据需求方二次流转以保障数据产权。“数据可用不可见”中的安全多方计算、联邦机器学习技术旨在解决“数据孤岛”和数据共享中的隐私安全这一两难问题。与传统的多中心合作模式相比，

“数据可用不可见”能有效帮助多个机构在满足用户隐私保护、数据安全和政府法规的要求下，进行数据使用和机器学习建模，避免数据的二次流转。例如，蚂蚁集团通过安全多方计算技术在金融行业中与金融机构在联合信用风险评估、联合交易反欺诈/反洗钱风险评估等业务应用场景中均展开了较大规模的实践。其中，蚂蚁隐私计算在 2021 年助力浦发银行获得“IDC 国际金融创新大奖”。

三是新模式、新技术的应用逐渐为数据交易流通提供整体安全保障。“数据可算不可识”模式在大规模或实时性要求较强的数据分析场景下为数据去除标识，构建集中式可信计算环境，为数据交易流通中的个人隐私权保障提供了安全保障。区块链技术具有去中心化等安全度较高的技术信任特性，可以为数据写入唯一的数字摘要码，已经被多家数据交易平台用于数据确权。然而，区块链并非实现数据可信存证的唯一技术解决方案，在未来需要鼓励探索更多样、有效的综合性技术方案，从场景出发解决数据交易流通中的可信存证问题。



### （三）数据要素市场行业应用不断拓展

中共中央、国务院《关于构建更加完善的要素市场化配置体制机制的意见》指出，要培育数字经济新产业、新业态和新模式，支持构建农业、工业、交通、教育、安防、城市管理、公共资源交易等领域规范化数据开发利用的场景。

当前，5G、大数据、人工智能、区块链等技术加速向各行业融合渗透，数据赋能、赋值、赋智作用日益凸显，数据要素市场应用场景不断拓展。在农业农村领域，车间农业、认养农业、云农场等新业态新模式方兴未艾，农业物联网、病虫害数字化防控、智能育种等新技术加快应用。在工业生产领域，智能硬件、可穿戴设备、智能网联汽车等新业态层出不穷，工业互联网、智能工厂等新模式蓬勃发展。2019年我国规模以上工业企业生产设备数字化率已达47.1%，未来数字化转型加快推进。在服务消费领域，“数据+”不断激发消费市场活力，居民消费习惯加速向线上迁移。2020年，我国实物商品网上零售额比上年增长14.8%，占社会消费品零售总额的比重为24.9%，增速明显高于同期社会消费品零售总额，未来服务消费数据将发挥更加重要的作用。在货币金融领域，移动支付全面推进，数字人民币试点提速，金融服务中小微企业精准性显著提升。在公共治理领域，数据要素在疫情监测分析、病毒溯源、防控救治、资源调配、复工复产等方面发挥了重要支撑作用，数字政府建设持续推进，政府管理和社会治理加快转型。

#### （四）数据要素市场生态体系日趋完善

应用场景的丰富提供了大量垂直领域的的数据需求，促进数据要素市场生态体系建设。

一是在行业发展方面，数据采集、数据标注、数据服务的行业属性也会伴随着应用场景的丰富不断加深。各行各业也将出台相关政策，不断释放数据红利，进而释放技术红利。未来，在智慧医疗行业，需要相关的病例数据、医疗影像数据等，进行自然语言处理、计算机视觉等技术的数据训练，进而利用相关技术深度应用；在工业领域，数据要素的重要性更加明显，工业企业拥有较长的产业链，工业数据的全产业链整合在未来生产中发挥的作用会更加显著，同时，工业场景大量的应用场景会用到人工智能、大数据技术，如利用图像识别技术进行的缺陷识别等，不同的零件、产品都需要进行单独的标注与训练，会使数据要素的行业属性更加明显。

二是在区域发展方面，各地充分发挥自身优势，优化地区产业布局。数据流通交易与技术研发依然集中在北上广深等大城市，而数据标注、清洗等人力密集型产业，逐渐向围绕中心经济带的欠发达地区转移。以北京为例，北京拥有大量大数据、人工智能研发企业，数据需求庞大，而北京周围如保定、太原等地，基于本地的人力成本优势，进行了如数据标注、数据审核相关产业的发展。同时，中心研发城市的产业辐射能力不断增强，整体产业开始走多地协同路线。例如，山西太原利用自身人力成本，发展数据标注产业，引入

百度公司，形成“北京研发、太原标注”的发展局面。

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

## 六、对策与建议

### （一）重塑数据价值流动规则

避免数据权属争议是数据要素发挥其市场价值的根本保障。

一是完善法律法规，构建适应数字经济发展需求的规则体系。各地也可以根据结合本地数据要素市场发展实际情况来制定适应性的法律法规与政策，探索创新型数据立法模式，构建完整数据法法律体系，通过法律形式对数据进行有效保护，并对数据进行合理确权，使市场运行可以在有法可依的条件下进行。在地方立法中，深圳市率先于2020年发布《深圳经济特区数据条例（征求意见稿）》，首次提出了“数据权”的概念，针对不同的数据类型明确数据权归属：自然人对其个人数据依法享有数据权，公共数据属于新型国有资产，数据权归国家所有，数据要素市场主体对其合法收集的数据和自身生成的数据享有数据权。而在《深圳经济特区数据条例（送审稿）》中，又删除了“数据权”，代之以“数据权益”，这也是全国首创。

二是考虑重构数据产权制度。数据作为生产要素并不是要打破现有知识产权制度，而是要构筑新的数据资源有效利用秩序。有形物体排他性的所有权不适用于数据要素，可以考虑以反垄断和隐私保护为框架，将数据使用决定权赋予数据形成和价值创造者，让数据创造者获得数据产生的收益，适应数据强调的访问和利用权利的需求。此外，还可以开展数据公有制研究，探索分离数据的“所有权”和“使用权”，

通过数据国家所有，保证数据不被龙头企业垄断，能够在社会中充分流通。

三是借助技术手段来实现数据产权界定。例如，目前的可验证计算、同态加密、安全多方计算等密码学技术可以支持“数据可用不可见”，某种程度实现了数据资产的不可辨别，保护数据主体权利不受损害，提供了“不影响数据主体所有权的前提下交易数据的使用权”的可能性。

## （二）扩大公共数据开放程度

一是加速政府数据流通管理方法论研究，构建政府数据标准化体系。结合实际对政府数据分类分级标准、数据目录标准、数据格式标准、数据平台标准、数据访问接口标准、数据分析模型标准、核心元数据编制等进行统一编制，同时制定政府数据采集、汇聚、交换、共享、开放、存储、计算、分析、安全、质量、平台对接等关键共性技术标准和管理规范，建立数据质量管理体系，支撑公共数据的共享和开放。

二是建立数据分级清单体系，将数据根据重要程度、保密程度，进行数据进行分级分类。根据不同分级水平，将数据进行差异化管理。高等级数据由国家控制，不对外进行开放共享；中等级数据利用“可用不可见”技术、区块链技术进行有条件的开放共享；低等级数据无条件开放共享，以数据集或 API 方式开放共享。

三是建立一站式开放数据平台，政府可通过平台将经济、医疗、教育、环境与地理等方面的数据以各种可访问的方式发布，并将分散的数据整合，开发商还可通过平台对数

据进行加工和二次开发，以此推动公共部门数据开放共享、科研数据共享、私营企业数据分享。

四是推动政府数据开发合理化，建立新兴的政企合作模式。政府将自身数据进行梳理后，将数据授权给独资的国资公司，国资公司对数据进行进一步管理及运营，从而促进公共数据的应用共享。

### **（三）促进数据要素交易流通**

一是加强建设数字信息基础设施，解决目前城乡之间、地区之间、行业之间存在的“数字鸿沟”问题，进而构建统一开放、竞争有序的数据要素大市场。发挥数据管理公共服务平台、数据资源服务及流通交易平台等设施的桥梁纽带作用，进一步释放数据资源价值。

二是规范数据利用行为，严格限制数据控制者滥用数据行为及不合理垄断行为，创造公平有序的数据要素交易环境。

三是大力培育数据要素市场主体，充分发挥大数据交易所、数据经纪商等市场中介作用。

四是培育多种类型的数据要素商业化模式，创新交易模式、数据收益分配模式和交易定价模式，建立行业级数据流通利用平台，明确规定各参与主体的权利及业务。

五是推动建立企业数据流通体系，主要方式为建立自动驾驶数据资源池、智能语音数据资源池、人脸识别数据资源池等专有领域数据资源池。

六是推动数据交易主体对于多源数据汇集、非结构化处

理、数据清洗、数据建模等技术和工作的应用，从而推动数据要素资产的估值和定价，提高数据要素的交易和流转效率。

#### （四）提升数据安全保障能力

一是健全数据安全法律法规，明确各方主体的数据安全法律责任。完善打击大数据犯罪相关立法，加强数据安全治理，提升数据安全防护水平。

二是加强数据安全防护体系建设，促进数据安全协同治理。政府、企业、行业组织需要有效配合，发挥各自优势，建立适应数据要素市场发展的协同治理模式。积极推动数据管理领域国家标准（DCMM）贯标评估工作，引导企业逐步建立完善的数据管理体系。

三是培养一批大数据领军企业和多层次、多类型的大数据人才队伍。促进核心技术领域取得新突破，推动数据安全技术在数据流通过程中的应用，降低已知风险的同时减少对数据要素流通的干扰。

四是创新数据要素流通技术手段。研发隐私计算、区块链、动态加密、零知识证明、群签名、环签名、差分隐私、数据标识等新技术，促进“数据可用不可见”“数据可算不可识”等模式推广应用，保障数据在交易流通过程中的安全性。

五是推动数据安全技术在数据流通过程中的应用。在国家 and 行业标准中，加入安全技术指标，持续推动数据质量标准建设，全面提升数据治理及管理能力。通过行业标准倒逼

各主体在数据流通管理、数据流通等环节应用新技术。

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心



CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

CIC国家工信安全中心

国家工业信息安全发展研究中心 人工智能所

地址：北京市石景山区鲁谷路 35 号

邮政编码：100040

联系电话：010-88692450

网址：<http://www.cics-cert.org.cn/>

CIC国家工信安全中心