



5. 数据要素市场

宫汝凯

gong.rukai@dhu.edu.cn

2025.04.22

身边的数字竞争

- 微信读书是腾讯运营的一款手机阅读应用，用户可以使用该软件阅读书籍。用户黄某在使用微信读书时发现，在自己并未授权的情况下，微信将其好友关系数据提供给微信读书，使后者“关注”栏目下出现了该用户的微信好友名单。微信读书未经用户授权，默认向关注用户的好友公开读书想法等信息。2020年7月30日，北京互联网法院对该案作出一审判决，认定微信读书构成对个人信息权益的侵犯。
- 如何从经济学角度认识隐私权和个人信息权？数据要素市场发展对产权和隐私规制提出了哪些目标？本讲将在介绍数据要素市场和数据产业链的基础上，分析和理解这些问题。

2025年4月

2

主要内容

- 数据要素与数据市场
- 数据要素产业链
- 数据确权
- 隐私与安全规制

2025年4月

3

1. 数据要素与数据市场

- 作为关键要素的数据
- 数据能够形成交易市场的基础之一：其能够作为生产要素，帮助企业与平台在各类生产经营活动中生产商品或提供服务。
 - ✓ 以大数据为基础的个性化推荐和精准营销有助于提升市场效率，优化市场供需的匹配效果。
 - 大数据技术会显著降低追踪消费者行为的追踪成本，使平台和企业能够更准确地根据消费者偏好投放广告，进行针对性的个性化推荐。
 - ✓ 大数据有助于推动技术进步和商业模式创新。
 - 企业通过大数据更高效、更准确地了解用户的需求，更有针对性地进行产品设计与创新，使开发周期有效缩短，生产过程更有效率；
 - 借助大数据技术，企业更容易鉴别现有生产流程与产品设计中的问题，有助于促进产品与服务的改进，也有助于商业模式的创新。

2025年4月

4

1. 数据要素与数据市场

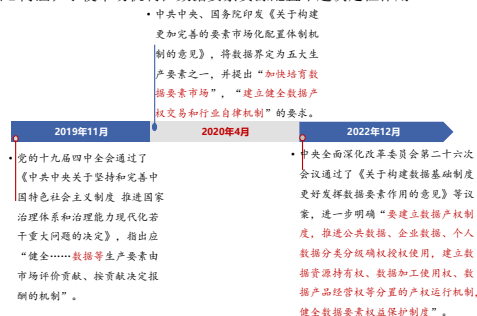
- 作为关键要素的数据
- ✓ 大数据有助于优化企业内部管理，提高管理效率和组织绩效。
 - 数据规模越大，包含的信息量就越大，所传达的内容就越精确，帮助管理者做出更理性的决策；
 - 为使组织绩效达到某种特定水平，决策任务越复杂，决策行为所要依据的信息也就越多，以大数据为基础的信息管理系统能够更好地收集信息并促进组织内更有效地分发信息，帮助管理者降低决策成本、提高管理效率。
- ✓ 大数据有助于促进产业链协同。
 - 大数据技术与企业数字化改革推动了企业之间的信息共享，将进一步提高企业间的协同创新能力，深化产业分工，帮助产业链上下游整合，形成以大数据平台为基础的完整统一的产业生态系统。

2025年4月

5

数据要素市场

- 数据要素市场的内涵
- 核心内涵在于使市场机制在数据要素资源配置中起决定性作用。



6

数据要素市场

► 数据要素市场核心内涵：使市场机制在数据要素资源配置中起决定性作用。

表5.1 数据要素市场规模

	2019年	2020年	2021年	2022年	2023年
大数据市场	8000亿元	10000亿元	——	1.58万亿	1.74万
数据要素市场（部分）	375亿元	545亿元	704亿元	——	1332亿
数据交易	23.93亿美元	——	——	——	74.1亿

数据来源：中国信息通信研究院、国家工业信息安全发展研究中心、前瞻产业研究院。

2025年4月

7

数据要素产业链

► 数据并非天然就是生产要素，而是需要经过若干个产业链环节：

► 数据采集→数据储存→数据加工

► →数据流通→数据分析→数据应用→生态保障

2025年4月

8

数据要素产业链

► 数据要素产业链的各环节

- 1. 数据采集
- 数据采集是数据要素市场的**根本**，也是数据产生的起点。
 - 关注数据采集的准确度和全面性；
 - 企业在生产经营中，需要在不同场景中采集不同的数据，采集的对象主要分为企业内部数据、企业外部数据以及定制化数据。

表5.2 数据采集对象及采集目的

采集对象	采集目的
企业内部数据	经营活动数据 实现数字化生产经营，以优化生产经营过程，增强抗风险能力和研判能力。
	生产环节数据
企业外部数据	舆情数据 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。
定制化数据	非结构化数据 满足对目标领域、场景的特定需求而进行数据采集。

资料来源：国家工业信息安全发展研究中心

9

数据要素产业链

- 2. 数据存储
- 数据存储是完成数据采集后，对有价值数据进行有效存储以便进一步加工处理的环节。
 - 关注数据存储的安全性和调用的实时性、便利性；
 - 企业在数据存储过程中，综合考虑数据的敏感度、时效性、企业硬件水平等选择合适的存储服务。
 - 现阶段，我国数据存储主要有公有云、私有云以及混合云。

表5.3 数据存储方式及优缺点

存储方式	优缺点
公有云	存储能力强，市场成熟度高，技术发展快；但数据安全性低，流量峰值期间的云端访问易出现性能问题（如网络拥堵）。
私有云	可保证数据的隐私性，可避免在流量峰值期间云端访问出现性能问题，数据传输、下载更快；但搭建和维护成本较高，较难通过远程访问，广泛部署存在限制。
混合云	分类分级混合采用公有云、私有云，兼具数据存储的安全性及灵活性，降低了存储成本；但提高数据维护的复杂性，可能面临兼容性等问题。

资料来源：国家工业信息安全发展研究中心

数据要素产业链

- 3. 数据加工
- 数据加工是对企业采集和存储的数据进行筛选和处理，提高数据可用性的环节，为数据资源的挖掘和分析奠定基础。
 - 关注数据的加工精度；
 - 具体包括数据清洗、数据标注、数据审核以及数据融合处理等方式。

表5.4 数据加工方式及作用

加工方式	作用
数据清洗	对数据进行校验的过程，目的在于删除重复信息、纠正存在错误，提升数据质量。
数据标注	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。
数据审核	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。
数据融合处理	将多源、多模态（异构）数据互相整合成数据集，以备进一步挖掘分析。

资料来源：国家工业信息安全发展研究中心

11

数据要素产业链

- 4. 数据流通
- 数据流通是数据要素市场的**核心环节**，在发挥数据要素作用、提高公共决策效率、扩展商业应用场景等方面起关键作用。
 - 根据**技术方式的不同**，可细分为：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。
 - 其中，数据开放共享主要涉及数据的两大所有权主体——**政府、企业**，各自以及之间的数据共享。

2025年4月

12

数据要素产业链

5. 数据分析

➢ 数据分析是指最大化开发数据的功能，发挥数据的要素作用。

关注数据的深度分析挖掘，根据分析对象可分为企业内部数据分析和企业外部数据分析。

表5.5 数据分析对象及目的

分析对象	分析目的
企业内部数据	主要针对企业内部生产经营分析，进行检测预警，以进一步落实企业内部数据采集的目的。
企业外部数据	主要针对企业外部数据，进行同业竞争分析、营销投放检测等。

资料来源：国家工业信息安全发展研究中心

2025年4月

13

数据要素产业链

6. 数据应用

➢ 数据应用是将数据作为生产要素，以在合理充分的应用中创造价值，降低传统生产要素(资本、劳动力、土地等)的获取成本、使用效率，以赋能企业各项业务运营。

例如，在娱乐场景中，应用软件根据用户点赞收藏的歌曲、视频等信息做进一步的精准推送。购物平台通过统计分析用户浏览过、购买过的商品，进行商品推荐和针对性的广告投放。

7. 生态保障

➢ 生态保障强调为数据要素市场运转保驾护航、构建良好的数据要素市场生态，以搞活、融通数据要素，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。

2025年4月

14

数据要素产业链

表5.6 数据要素生态保障

保障环节	措施
数据资产评估	通过第三方评估机构或企业对数据所有者在生产、运营过程中产生的数据进行内在价值和使用权价值的评估，为数据要素流通交易提供基础性参考。
登记结算	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。
交易撮合	为数据交易双方提供信息检索、信息匹配、供需对接及交易竞价等第三方服务。
争议仲裁	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。
跨境流动监管	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。

资料来源：国家工业信息安全发展研究中心

➢ 要素产业链所需的技术

数据要素全产业链的每一个环节从采集、存储到加工、流通、分析以及应用无不需要借助一定的技术手段。

2025年4月

15

数据要素产业链

表5.7 数据要素产业链各环节及所需技术

产业链环节	所需技术	应用技术
数据要素	企业内部数据：经营活动数据、生产环节数据 企业外部数据：舆情数据、定制数据 结构化数据：数据库、数据湖、ETL (Extract-Transform-Load) 技术、数据仓库、数据中台、数据工厂、数据集市、数据沙箱等 非结构化数据：图像、音频、视频、文本、日志、传感器数据、物联网数据、PLC传感器等	数据开放共享：数据交易平台 数据交易：数字交易平台
数据存储	公有云、私有云、混合云 数据清洗：大数据平台提供自动化的数据清洗功能，简化数据加工过程 数据标注：人工标注	数据流通：API技术服务 “数据可用不可见”“数据可控不可篡改”“数据可证去标识”
数据加工	数据审核：自然语言处理 (NLP)、计算机视觉等技术 数据库技术	数据分析：企业内部数据：离线：云计算平台 (HDFS文件系统、MapReduce运算框架)、EXCEL、Oracle以及Yonghong Z-Suite BI套件等国产专业软件在线：一般构建在云计算平台的NoSQL系统上 企业外部数据：爬虫技术

资料来源：国家工业信息安全发展研究中心

数据确权

➢ 数据要素的主要特征

➢ 数据确权的目标

➢ 数据确权中的权衡取舍tradeoff

2025年4月

17

数据要素主要特征

➢ 1. 非竞争性

✓ 数据本质上是一种信息的载体和表现形式，具有可复制、可无限分享的特性，即使用者增多不影响数据资产本身的价值。

✓ 非竞争性：可被多个市场主体同时使用，且不会被消耗掉，打破资本、劳动等传统要素有限供给对增长的制约。

✓ 非竞争性+“质变”特征：

与技术和制度等非竞争性要素相似，

*数据对生产活动的影响不具有明确的投入/产出比；

资本、劳动力等传统竞争性要素：随投入规模变化；

数据对产出的影响往往会经历一段漫长的积累期，随后迎来生产效率的指数级增长。

2025年4月

18

数据要素主要特征

➤ 2. 衍生性

- ✓ **第一个层次**：数据之间相互关联，可以产生更多的价值。
 - 涉及单个数据与多维海量数据之间的关系。例如，
 - 对于淘宝商家而言，要想分析哪种产品更加畅销，单单观察某一个消费者的消费历史无法推断，只有把足够多的消费者数据汇聚和进行大数据分析，才能产生可以用于生产经营决策有价值的信息。
- ✓ **第二个层次**：加工处理**原始数据**后，能够分析出更有价值的信息。
 - **加工处理**是使数据真正成为生产要素的重要前提：**原始数据**所蕴含的价值非常低，其价值只有在采集、加工、分析之后才能充分显现，例如，
 - **数据和信息的生产要素化**：一份自动记录的账表往往不能直接作为数据指导生产经营，而对其进行挖掘、清洗、处理、整合、转换与分析，符合分拆使用要求后，可能成为**数据资本**。

19

数据要素主要特征

➤ 3. 隐私负外部性

- 任何消费者的“**用隐私换便利**”交易行为都可能带来隐私负外部性的问题。
- 现实生活中，看似分散独立的信息之间普遍存在相关性：消费数据可能透露出该消费者的消费习惯(***显示性偏好**)及其关联其他用户的信息，给消费者带来一系列负面效用，带来了隐私负外部性。
- **数据安全与隐私保护视角**：频繁的数据流转和交换将使得数据泄露不再是一次性事件，众多**非敏感数据**通过大数据聚合和分析可以形成**敏感数据**和有价值的衍生信息。

2025年4月

20

数据确权目标

➤ 1. 数据充分使用的效率要求

- 非竞争性在效率上的**引申含义**是数据**应该**充分流通和共享。
 - 非竞争性意味着同一**数据**所蕴含的**信息**和**知识**可以同时被多人使用；
 - 数据对于提高生产效率具有***乘数效应**；
 - 社会效率最大化角度：数据要素应被市场上所有相关主体充分获取与使用。
 - 数据是数字经济时代最重要的生产要素，只有广泛地交易流通才能充分发掘数据要素的价值。

2025年4月

21

数据确权目标

➤ 1. 数据充分使用的效率要求

- 然而，**数据处理主体**与**数据要素需求方**存在竞争关系，便缺乏数据流通的动力。
- 设想两个相互竞争的平台A和B，A通过数据采集和处理获得了1单位的数据要素，该要素可应用于帮助平台改进用户体验；
- 如果平台A将数据要素提供给平台B，那么会提高后者的服务质量，强化其市场竞争能力，对平台A是不利的；
- ***若将数据产权配置给数据处理主体**，出于竞争的考虑，它们将缺乏数据流通和共享的动力，将导致数据的非充分使用；
- ****若将数据产权配置给消费者**，消费者可以对需要数据的企业进行授权，有利于实现数据要素的充分流通。

2025年4月

22

数据确权目标

➤ 2. 数据处理主体的激励

- 衍生性的**引申含义**是数据产权配置应当为数据处理主体提供激励。
 - 数据的衍生性意味着数据要素的形成需要经过采集、加工、分析等一系列过程，而每个过程都需要数据处理主体进行大量资本、技术等方面的投入；
 - 为了使数据处理主体有动力进行相关投入，需要对数据权益进行肯定和保护。

2025年4月

23

数据确权目标

➤ 2. 数据处理主体的激励

- 衍生性的**引申含义**是数据产权的配置应当为数据处理主体提供激励。
 - 为了激励参与者对数据进行开发处理，需要奖励投入和鼓励创新；
 - **投入**：对数据的收集和存储；
 - **创新**：数据开发和使用过程中的优化改进；
 - **奖励投入**：数据处理主体需要产权激励才有动力生产足够多的数据，不合理的产权配置可能导致**搭便车 (free-riding)** 行为；
 - **鼓励创新**：数据产业链包含采集、存储、加工、流通、分析、应用、评价、交易等多个环节，每个环节都有创新优化的空间，而创新优化离不开参与者的投资。相比之下，平台企业相比个人更有能力进行相关投资，但需要有足够的激励，保证它们有权通过投入来获取相应的收益。

2025年4月

24

数据确权目标

- 3. 隐私保护目标
- 消费者数据交易带来的**隐私负外部性**是消费者对隐私保护问题的主要担忧。具体来说，隐私负外部性问题表现在三个方面：
 - 个人天然有对隐私的偏好，一旦隐私被侵犯将出现**效用**损失；
 - 个人数据被厂商利用所可能导致的经济损失，如厂商基于消费者个人信息进行价格歧视（“**大数据杀熟**”）；
 - 个人信息**滥用**的风险，如垃圾短信、骚扰电话和诈骗等。

2025年4月

25

数据确权中的权衡取舍

- 数据确权的**三重目标**(数据使用效率、对数据处理主体激励与隐私保护)之间往往难以兼顾，构成数字要素市场的一个“**不可能三角**”
- 数据确权时，需要考虑三大目标之间的平衡：
 - ✓ 若考虑**使用效率最大化目标**：数据在使用过程中充分流通，消费者隐私泄露的风险必然增加；
 - ✓ 若考虑**激励目标**：将数据产权分配给数据处理主体，会产生两个问题：
 - 出于**竞争的考虑**：数据处理主体可能不愿意把数据提供给市场上其他竞争对手，反而策略性地阻碍数据流通，不利于效率最大化；
 - 若数据产权全部归属于数据处理主体，在数据处理过程中将不会考虑隐私保护问题。
 - ✓ 考虑**隐私保护目标**：让数据产权全部归属于个人，数据处理主体的权益难以得到保障，抑制后者进行数据投资的动力。

2025年4月

26

数据确权中的权衡取舍

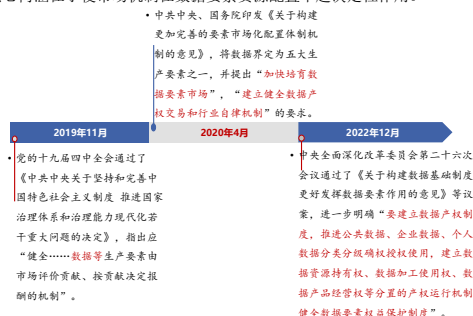
- 数据确权的多重目标是相互冲突的，引申出数据产权的界定与配置问题：
 - ✓ 在单一数据产权结构下，问题的焦点是如何进行数据产权的配置来兼顾多重目标。
 - Dosis和Sand-Zantman(2023)主张，数据产权到底属于消费者还是厂商，应取决于数据的价值。该理论的特点在于强调了**数据使用上的不可验证性**，即消费者和第三方都无法验证厂商的数据使用情况。结论来自于两方面因素的权衡取舍。产权归属于厂商将导致数据的过度使用和对隐私的过度侵害，而归属于消费者则使厂商缺乏在数据处理上进行投入的动力。对于价值相对较低的数据，数据处理不足的效率损失被隐私成本所抵消，所以产权配置给消费者是合意的；高价值数据的情形则相反。
 - ✓ 多重目标需要多重的数据产权体系。
 - 申卫星（2020）认为应当区分数据的**所有权和用益权**；所有权归属于数据原发者，用益权则可以归属于数据处理者。数据用益权既可以基于数据所有权人授权和数据采集、加工等事实行为取得，也可以通过共享、交易等方式继受取得，以保障数据处理者通过数据获取收益的权利。

2025年4月

27

Recall:数据要素市场

- **数据要素市场的内涵**
- 核心内涵在于使市场机制在数据要素资源配置中起决定性作用。



28

隐私与安全规制

- 支持规制观点
- 反对规制观点

2025年4月

29

支持规制观点

- 提倡隐私规制的原因主要来自于两个方面：
 - ✓ 数据确权对隐私保障的**无效性**；
 - ✓ 隐私可能带来的**外部性问题**。
- 1. 在厂商具有市场势力的情形下，个人数据产权的界定无法形成对隐私的有效保护；
- 个人**数据商品化**将会松动个人对隐私的控制；
- 在个人数据市场上，市场结构能够显著影响数据确权的效果和福利。

2025年4月

30

支持规制观点

➤ 支持规制的观点

- 比如，有足够市场势力的厂商能够以拒绝提供任何商品的方式强迫消费者提供与自己偏好相关的信息，导致隐私权不能为消费者带来任何额外的**信息租金**；
- Acquisti等(2016)通过一个包含基于行为的价格歧视垄断模型研究发现，在消费者保留个人数据情况下，垄断厂商就不能对其进行价格歧视，只能统一定价。在统一价市场上边际消费者并未获得正的剩余。一个很小的报价就能使该边际消费者**卖出数据**。按照这个逻辑递归，均衡时厂商就会无成本地获得全部消费者的隐私数据。由此得出结论：即使界定了清晰的个人数据产权，也无法有效保护个人隐私。
- 综观上述结论，政策含义在于，仅靠数据确权不能防止个人数据被过度收集，也不能改善消费者福利，所以保护隐私需要施加规制。
- **Pre-paper:** Acquisti A, Taylor C, Wagman L. *The economics of privacy*[J]. *Journal of Economic Literature*, 2016, 54(13): 442-492.

31

支持规制观点

➤ 2. 由于**信息的外部性**，仅靠数据确权不能实现有效率的数据采集。

- 虽然有些消费者会隐藏个人信息，但大数据分析技术发展让企业能够利用**其他消费者**提供的信息推测出这些消费者的类型；
- 即便每一位消费者对其个人信息的公开可能导致的后果都有充分的认识，数据采集总量也会偏离**社会最优水平**；
- 消费者会对此做出策略性反应：既然个人信息可能因其他消费者的数据交易或许可而泄露，每个消费者便会缺少足够的动机保护自己的隐私，导致**过低的数据采集成本**和**过度的数据采集**。

2025年4月

32

支持规制观点

➤ 3. 在个人隐私权问题上，如果**策略性行为**会引起新的**外部性问题**，则需要对其进行恰当的隐私规制。

- 例如，就某些敏感医疗记录（如戒毒治疗或艾滋病治疗）而言，如果没有良好的隐私保护措施，那么预期到信息泄露所可能导致的社会歧视，人们就可能缺乏接受治疗的动力，从而危害到私人 and 公众的利益。
- 不同消费者之间也存在外部性问题。
- 例如，低支付意愿消费者的信息隐瞒行为将导致企业将营销努力更多地转向高支付意愿的消费者，导致他们遭受额外的隐私侵害（Hann和Katz, 2008）。这种外部性问题很难单纯通过基于产权的市场交易来内部化，这就进一步加强了隐私规制的必要性。

2025年4月

33

反对规制观点

➤ 关于隐私规制的反对意见主要包括四种观点：

- ✓ 1. 隐私规制会阻断对市场有益的信号传递机制。
 - 芝加哥学派从20世纪70到80年代早期就基于此原因反对隐私规制；
 - 由于隐私保护而产生的信息隐藏会造成负外部性问题。
 - **Ex:** 求职者可能隐藏一些对自己不利的个人信息，而隐私保护会让企业不能对应聘者进行全方位的审查。因此，保护应聘者的隐私是以牺牲企业的盈利能力为代价的，并最终将负面信息特征的成本转移到其他市场参与者身上。
- ✓ 2. 隐私规制会对商业模式的创新造成负面影响。
 - 电子商务对网络广告和个性化推送的依赖性很大，而严格的隐私规制可能会阻碍平台在个性化广告上的商业模式创新。

2025年4月

34

反对规制观点

➤ 关于隐私规制的反对意见主要包括四种观点：

- ✓ 3. 隐私规制会加强垄断企业的市场势力。
 - 如果隐私监管主要依靠强制要求消费者的知情和同意，那么用户就不会选择不太成熟的或新厂商提供的服务，使隐私规制成为一种事实上的进入壁垒。
- ✓ 4. 企业的策略性行为导致规制无法达到预期，甚至会产生新的问题。
 - 以“知情同意”关键原则为例。企业可以通过“包装”隐私权协定来引诱用户接受侵犯隐私的条款；
 - 隐私权规制可能带来新的外部性问题，Aridoret等(2023)以欧盟《通用数据保护条例》为例，研究发现，《通用数据保护条例》的执行使消费者可以通过直接拒绝提供数据的方式来取代以往不完美的信息隐藏措施(如删除网络痕迹)。不完美的隐私保护手段会在厂商采集数据中制造噪音，使其更难精确地追踪用户的行为，而《通用数据保护条例》所引致的替代行为使得厂商能够得到更加“干净”的数据环境，从而**使用户更容易被追踪和识别**。

2025年4月

35

本讲小结

- 了解数据要素与数据市场
- 掌握数据要素产业链
- 掌握数据确权
- 了解隐私与安全规制

2025年4月

36