

数据经济学第 1 次作业评讲

陈希路

暨南大学经济学院

2026 年春

1. 数据与想法的区别及其排他性差异

简答题 1

简述在数据经济学中数据与想法的区别，并结合自动驾驶领域的现状，说明两者在排他性程度上的差异。

核心要点解析

- **经济学本质区别：**
 - **想法：**相当于**生产函数**，是一组指导如何制造商品的指令或配方
 - **数据：**相当于**生产要素**，是用于产生新想法或直接投入生产的原始投入
- **自动驾驶领域的排他性差异：**
 - **作为想法的算法逻辑：**往往通过开源社区、学术论文公开，排他性较低
 - **作为要素的驾驶数据：**包含复杂多变的路测记录，属于核心商业机密，企业通常会通过加密、封闭生态等手段对数据进行严密监控
 - **结论：**想法具备非排他性倾向，而数据在实际应用中往往呈现出**有限排他性**

2. 理解数据的三个主要维度与基本要求

简答题 2

目前学界对数据尚无统一共识，但通常可以从三个主要维度来理解数据的基本定义。请简述这三个维度，并说明真实、准确为何是对数据的基本要求。

核心要点解析

• 理解数据的三个主要维度：

- ① 数据是对现实世界的抽象，是现实世界的模型
- ② 数据本质上是一种表示方式，需要遵循特定的规范和标准
- ③ 数据可支持分析、推理、计算和决策

• 为什么要求真实、准确？

- 在科学、商业或政府治理领域，数据是推理和计算的基础
- 只有保证数据的真实与准确，加工提炼出的信息才具有指导价值，否则基于错误数据做出的分析和决策将带来严重的经济损失或政策偏离

3. 数据要素与传统生产要素的核心差异

简答题 3

简述数据要素与传统生产要素在形态特征与稀缺性两个方面的核心差异。

对比分析矩阵

比较维度	传统生产要素	数据要素
形态特征	实体形态，伴随不可逆的物理磨损	虚拟形态，无物理损耗，可无限次、不限期重复使用
稀缺性	具有排他性	具备可无限复制与可共享性

4. 数据要素投入生产时的边际收益变动

简答题 4

解释数据要素投入生产时，其边际收益如何变化，并说明数据要素为什么能在特定阶段产生边际收益递增的特征。

核心要点解析

- 边际收益的变化轨迹：
 - 初期：数据规模不足，可能因噪声带来负效应
 - 中期：当数据量累积超越临界规模后，信息价值加速释放，边际收益**递增**
 - 后期：随着数据覆盖维度趋近饱和，边际收益增速逐渐放缓
- 特定阶段边际收益递增的内在成因：
 - 单一维度数据价值较低，但当不同维度的数据结合时，会产生**乘数效应**
 - 数据的边际成本接近于零，流转规模越大，其协同与溢出价值越高

5. 数据的自我迭代效应：以大语言模型训练为例

简答题 5

传统生产要素如土地和资本往往面临总量限制或折旧问题。相比之下，数据作为新型生产要素具有怎样独特的产生方式？请结合人工智能大语言模型的训练过程，详细说明上述特性。

核心要点解析

- **独特的产生方式——自我迭代效应：**
 - 数据在积累和运用的过程中，能够通过自身的反馈循环，逐步提升自身质量、丰富度与经济价值
- **大语言模型训练中的演进过程：**
 - ① 研发人员收集并过滤公开文本，作为模型的初始无监督训练语料
 - ② 随着用户接入模型并产生交互，用户输入指令及人类反馈成为**新数据**
 - ③ 平台将新生成的对话记录再次喂给算法，实现算法模型的自进化

6. 数据分级授权机制与交易成本的优化

简答题 6

平台企业在收集和处理用户产生的大量行为记录时，往往面临极高的确权协商成本。请简要论述数据分级授权机制的内涵，并解释标准化授权条款是如何降低交易成本并促成共识的。

核心要点解析

- **数据分级授权机制的内涵：**
 - 将数据资产流转、交易及使用的控制权限进行**模块化、标准化分级**
 - 信息流转的多方在明确预设的标准化选项中快速匹配并达成初始产权契约
- **标准化授权条款降低成本、达成共识的路径：**
 - 规避了平台与数亿级个人用户逐一进行高成本、零散化谈判的物理壁垒
 - 用户可以根据透明的对价机制，自主高效做出决策
 - 标准化条款能够显著降低后续潜在的司法裁决不确定性

案例导入：传统服装制造 vs. 数字快时尚电商

案例分析题背景

通过对比传统大型服装企业“老裁缝”与新兴跨境快时尚电商平台“数智风”的经营模式，深度剖析数据要素对现代产业链和核心竞争力的重塑。

核心双模态对比

- **“老裁缝” (传统经验驱动)**：依赖核心设计师、历史销售预测、提前半年下达大批量订单 → 产生**库存积压与滞销危机**
- **“数智风” (新型数据驱动)**：网络抓取多维素材、系统自动提炼流行趋势、柔性供应链 → 实现**“小单快反”极低库存与爆发式增长**

1. 生产范式归属与数据要素的变革推动

案例分析题 (1)

“老裁缝”与“数智风”分别代表了历史上的哪种生产范式？数据要素是如何推动这种生产组织方式发生变革的？

核心要点解析

- 生产范式归属界定：
 - “老裁缝”：代表了石油革命带来的“福特制大规模生产范式”，其核心特征是标准化、大批量、经验驱动
 - “数智风”：代表了信息革命带来的“数字经济范式”，其核心特征是数据驱动、小批量、敏捷反应

1. 生产范式归属与数据要素的变革推动

案例分析题 (1)

“老裁缝”与“数智风”分别代表了历史上的哪种生产范式？数据要素是如何推动这种生产组织方式发生变革的？

核心要点解析

- **数据要素推动变革的内在机制：**
 - 数据作为新型生产要素，打通了前端消费端与后端生产端的信息鸿沟
 - 数据的实时流动推动了柔性生产，使供需匹配效率提升，生产组织方式从厂家主导转变为**消费者实时需求主导**

2. 数据与信息的界定及转化为核心竞争力的路径

案例分析题 (2)

“数智风”抓取的海量图片、热词和点赞量属于数据还是信息？它们是如何转化为企业核心竞争力的？

核心要点解析

- **概念界定：**
 - 平台抓取的社交图片、热词搜索量、用户点赞与评论等，均属于未经加工、无序的原始事实
 - 在经济学与信息学上，它们属于典型的**非结构化数据**

2. 数据与信息界定及转化为核心竞争力的路径

案例分析题 (2)

“数智风”抓取的海量图片、热词和点赞量属于数据还是信息？它们是如何转化为企业核心竞争力的？

核心要点解析

- **转化为核心竞争力的演进路径：**
 - 原始数据本身无法直接指导生产，平台通过算法模型对其进行清洗、分类和聚合处理
 - 从中提炼出有经济价值的流行趋势，完成从数据到信息的转化
 - 提取出的信息直接支持了“小单快反”的柔性生产与精准追加订单，跨越价值鸿沟，固化为企业的**可持续核心竞争力**

3. 数据的石油属性与公共产品属性分析

案例分析题 (3)

《经济学人》杂志曾提出，数据具有像石油和像公共产品的双重属性。请结合本案例场景，分别谈谈“数智风”的数据是如何表现出这两种属性的。

核心要点解析

- 石油属性：
 - 独家抓取的图文趋势数据及实时的销售、购买转化率反馈，是驱动其生产闭环精准运转的燃料
 - 这类数据具备高商业价值，企业通常会将其作为独占核心资产严密保护，表现出强排他性与增值潜力

3. 数据的石油属性与公共产品属性分析

案例分析题 (3)

《经济学人》杂志曾提出，数据具有像石油和像公共产品的双重属性。请结合本案例场景，分别谈谈“数智风”的数据是如何表现出这两种属性的。

核心要点解析

- 公共产品属性：
 - 假设这些宏观时尚趋势数据进行脱敏并向行业开放，它们可以被无数个服装企业（如“老裁缝”）同时且反复利用
 - 数据在被多方共享和反复消费时，其本身的物理与经济价值**并未减少**，能有效消除全行业的盲目生产，降低社会总资源浪费

案例背景：智医科影与医疗影像大模型

案例材料

- 智医科影投入数亿元购买高性能计算集群，并高薪聘请医疗专家标注了大量肺部 CT 影像初级数据
- 辅助诊断系统上线后，随着越来越多基层医院接入平台，每次新增辅助诊断的网络分发计算成本极低
- 各地医院上传的病例数据不断反哺给中央云服务器，使得模型诊断算法的准确率大幅提升

1. 边际成本趋于零的特性与例外情况

案例分析题 (1)

结合数据产品与服务边际成本趋于零的特性，分析智医科影辅助诊断系统在推广过程中的成本优势，并指出在何种例外情况下系统仍需付出高昂的额外成本。

核心要点解析

- **辅助诊断系统在推广过程中的成本优势：**
 - 边际成本趋于零是指生产额外单位的数字产品不增加总生产成本
 - 智医科影将处理好的云端诊断服务分发给新增的基层医院时，网络传输与复制的边际成本几乎为零
- **系统仍需付出高昂额外成本的例外情况：**
 - 例外情况包括不确定性带来的成本以及时效性带来的成本
 - 为了向医院提供标准可信和紧跟前沿的医疗诊断结果，企业必须持续投入大量人力和算法资源，对异质化严重的原始医疗数据进行脱敏与清洗核验

2. 数据要素规模报酬递增的理论逻辑

案例分析题 (2)

运用数据要素规模报酬递增的理论逻辑，解释各地基层医院接入系统后如何帮助智医科影实现数据要素规模报酬递增。

核心要点解析

- **平均成本效应层面：**

- 智医科影前期采购计算集群与标注原始影像付出了巨额固定成本
- 大量基层医院接入后迅速摊薄了单次问诊的平均硬件成本

- **网络效应层面：**

- 各地医院上传的差异化病例数据形成互补，提升了诊断模型的通用能力
- 系统积累的数据特征越丰富，输出的诊断质量就越高，进而吸引更多大型医院入驻并带来优质新数据

3. 数据要素价值非独立性与决定因素

案例分析题 (3)

原始的 CT 影像数据如果没有辅助诊断系统的处理，对于普通患者而言完全无法理解且价值极低。请从数据要素价值非独立性的角度分析上述现象，并说明决定数据要素价值的两个因素在该案例中是如何体现的。

核心要点解析

- **数据要素价值非独立性的角度分析：**
 - 数据要素价值非独立性：数据要与其他生产要素结合实现自身价值
 - 原始 CT 影像脱离了特定算法和医学场景，仅会消耗硬盘存储资源
- **决定数据要素价值的两个因素在案例中的体现：**
 - **科技应用范围：**智医科影对影像进行特征提取，在医疗诊断场景中创造价值
 - **数字基础设施：**超高清医学影像的实时传输与大模型训练依赖超级计算中心与高速宽带网络，这是保障医疗数据价值成功变现的基础物理支撑

案例背景：智悦行与数据交易所的合作探索

案例材料

- 用户每天佩戴智悦行的智能手环，产生步数、心率、睡眠质量等体征指标，起初仅用于优化手环自身的计步算法
- 研发部门将脱敏后的群体健康指标进行清洗、交叉分析与建模，打造出城市人群健康指数报告并上架至数据交易所
- 商业保险公司订阅指数报告用于调整健康险产品的精算与定价策略，向新增保险公司开通权限的物理开销几乎为零

1. 信息科学角度的数据生产步骤

案例分析题 (1)

从信息科学角度看，智悦行开发出城市人群健康指数报告经历了哪些数据生产步骤？

核心要点解析

- **数据生产的四个步骤：**
 - **数据采集环节：**手环记录步数与心率，构成数据采集
 - **数据清洗环节：**研发部门对原始记录进行脱敏处理并剔除异常值
 - **数据存储环节：**清洗完毕的指标被存储在云端数据库中
 - **处理与解读环节：**业务团队开展交叉分析与精算建模，凝练出对保险机构有价值的指数报告

2. 数据权属理论与三权分置运行机制

案例分析题 (2)

结合数据权属理论，说明体征指标在手环用户、智悦行以及商业保险机构之间是如何体现数据资源持有权、数据加工使用权和数据产品经营权的分置运行机制的？

核心要点解析

- **各种权利在不同利益主体间的分离：**
 - **手环用户：** 在日常生活中产生体征指标的原始信息
 - **智悦行（平台运营方）：**
 - 拥有脱敏和清洗汇聚后底层数据库的数据资源持有权
 - 投入算力和算法对底层指标进行清洗建模，行使数据加工使用权
 - 将生成的指数报告上架至交易所并收取订阅费用，行使数据产品经营权
 - **商业保险机构：** 购买加工后的结果与参考价值，不掌握原始底层明细

3. 数据供给侧的成本构成理论与成本结构

案例分析题 (3)

根据数据供给侧的成本构成理论，简析城市人群健康指数报告作为一种数据产品所表现出的特殊成本结构。

核心要点解析

- **初始创作成本高与再生产边际成本趋于零的结构：**
 - **初始投入与生产成本：**智悦行在研发初期面临硬件部署、云端存储、算法研发以及合规审查等生产成本和初始投入
 - **再生产边际成本趋于零：**指数报告在交易所挂牌上架并吸引商业保险公司订阅读时，运营方只需开放系统访问权限，新增客户带来的物理复制和分发成本微乎其微，体现了数据要素的非竞争性特质