

数据生产要素核算

——关于理论与方法的若干讨论

李静萍
中国人民大学
2023.8.25

内容

- 研究背景
- 数据为什么是生产要素
- 数据要素核算对象
- 数据要素核算方法
- 总结

研究背景

- 数字经济已经成为经济增长的“新引擎”。
- 数字技术支撑
- 顶层设计加持
 - 2019年十九届四中全会：健全劳动、资本、土地、知识、技术、管理、数据等生产要素由市场评价贡献、按贡献决定报酬的机制”。
 - 2021年3月发布的《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》：激活数据要素潜能，充分发挥海量数据和丰富应用场景优势，打造数字经济发展新优势的目标。
 - 2023年2月27日，《数字中国建设整体布局规划》：建设数字中国是数字时代推进中国式现代化的重要引擎，夯实数字基础设施和数据资源体系是数字中国建设的“两大基础”。

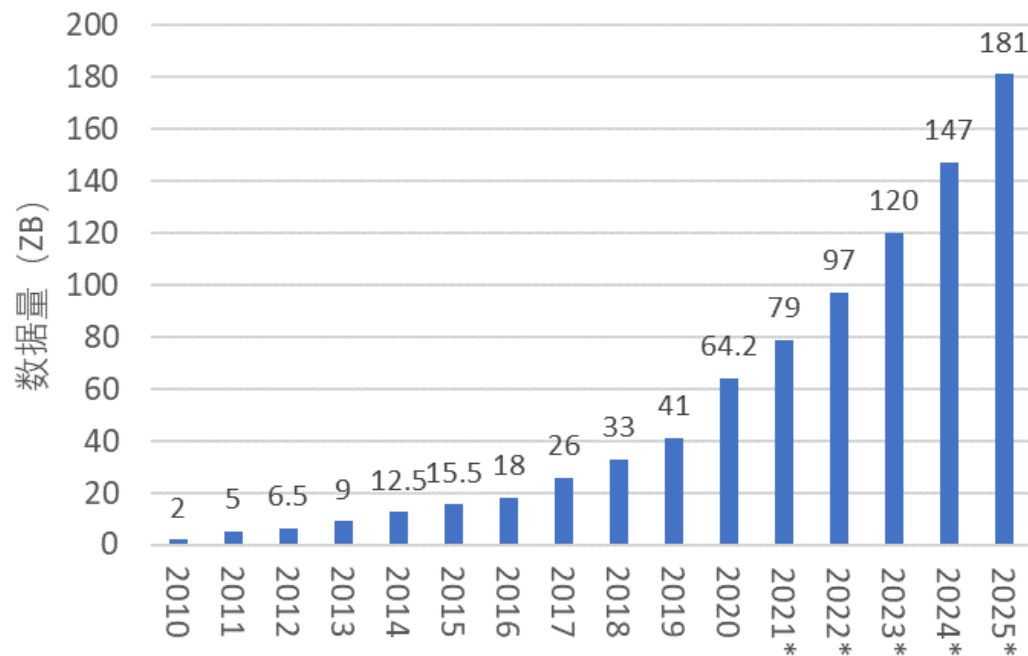


图1 全球创造、消费和存储的数据量

注：2021-2025年为预测数。

数据来源：<https://www.statista.com/statistics/871513/worldwide-data-created/>

研究背景（续）

- 在现行微观会计和国民账户体系中，尚未将数据生产要素纳入核算范围。
- 数据进入微观会计资产负债表的进程提速。
 - 2022年12月，财政部发布了《企业数据资源相关会计处理暂行规定（征求意见稿）》。
- 国际上数据核算研究与实践加速。
 - 加拿大统计局2019年对数据价值进行试算。
 - 2020年3月，联合国统计委员会秘书处间国民账户工作组（ISWGNA）将“数据如何纳入国民账户体系”明确列入SNA-08修订的研究议程。
- 国内数据生产要素宏观核算的进展滞后。
 - 目前的核算研究主要停留在数据资产的理论研究。
 - 从实践来看只有深圳市统计局等少数地区进行了数据生产要素统计核算试点工作。

数据生产要素核算现状极大滞后于数字经济发展的需要，应成为国民经济核算研究的优先问题。

从词源理解数据

- 剑桥词典：“被搜集的、经检验和考虑、用来**辅助决策**，或者以**电子形式**存在的、可被**计算机**存储和使用的**信息**（特别是**事实或数字**）。”
- 韦伯斯特词典：data既是datum的复数形式，又是独立于datum的单数词汇，是一个**有一定组织的整体**。

data (n.)

1640s, "a fact given or granted," classical plural of *datum*, from Latin *datum* "(thing) given," neuter past participle of *dare* "to give" (from PIE root **do-* "to give"). In classical use originally "a fact given as the basis for calculation in mathematical problems." From 1897 as "numerical facts collected for future reference."

Meaning "transmittable and storable information by which computer operations are performed" is first recorded 1946. *Data-processing* is from 1954; *data-base* (also *database*) "structured collection of data in a computer" is by 1962; *data-entry* is by 1970.

数据是对现象进行观察之后形成的记录，是多条记录形成的整体，可以特指已经完成电子化后的数据。

从数据价值链理解数据

- 三种核算口径：
 - 窄口径：把数据限定为原始数据。
 - 中口径：对数据与信息不做区分，如Jones & Tonetti (2020) 将数据定义为信息中不属于“创意”和“知识”的部分。
 - 宽口径：把整个数据价值链全部纳入数据核算范围，如许宪春等(2022) 构建了包含“数据收集、数据存储、数据分析和数据应用”4个阶段的数据价值链，来描述数据的生产过程。

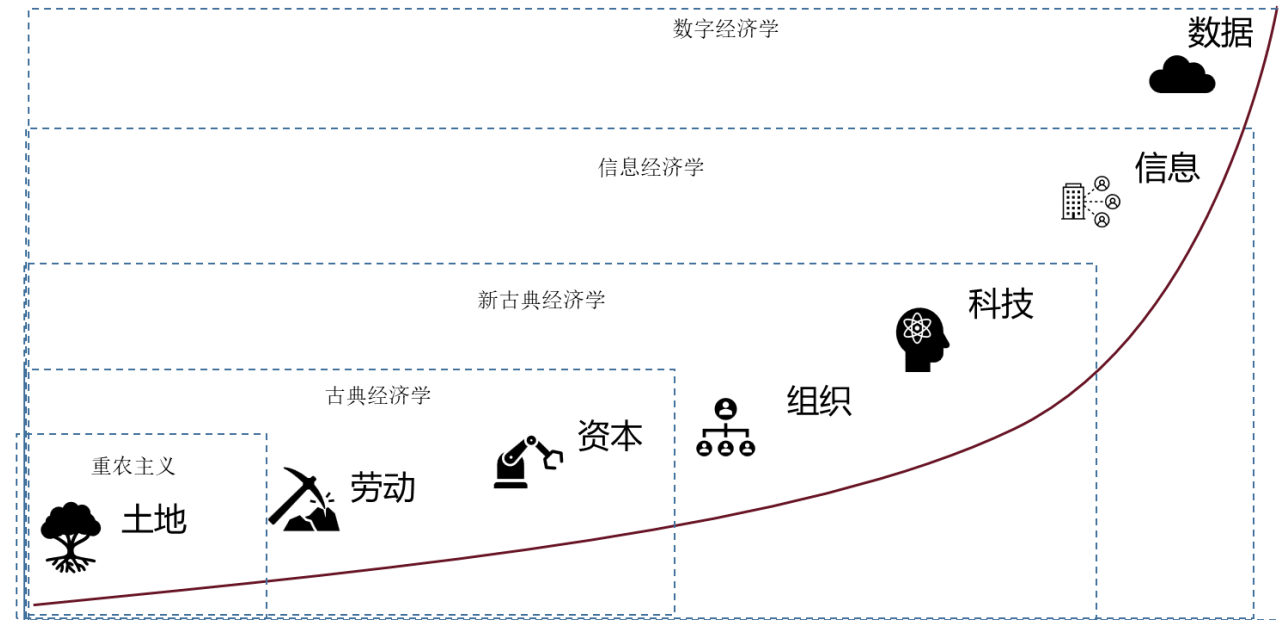


图2 数据价值链

资料来源：Curry等(2016)

生产要素

- 生产要素是指能够对生产过程有所贡献的东西，是对生产过程的投入。
- 从经济学意义上来看，只有具备稀缺性的投入才被视为生产要素。
- 在生产的循环中，不同的生产要素处于不同的等级。
- 劳动和土地是原始生产要素。
- 对各种次级生产要素的认识是在产业技术革命和经济发展的不同阶段，随着次级生产要素作用的凸现而不断扩展的。



数据具有稀缺性吗？

- 观点1：数据的产生是无尽的，因此具有非稀缺性。
- 观点2：从数据的实际供求关系来看，数据是稀缺的。
 - 有效的数据分析对样本量的要求。
 - 高质量数据瓶颈。高质量数据存量增长缓慢，我们可能会陷入训练数据的瓶颈：对于语言模型，将发生在 2030 年至 2040 年之间；对于图像模型，将发生在 2030 年至 2060 年之间（Villalobos 等，2022）。
 - 数据共享和交易存在很多障碍，流通不畅导致数据的稀缺性比其他生产要素更为突出。

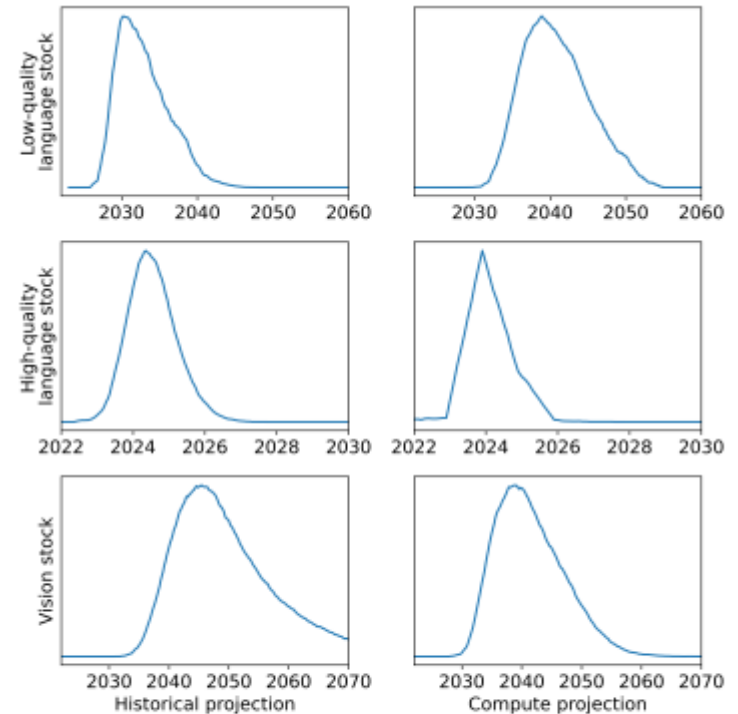


Fig. 6: Distribution of exhaustion dates for each intersection of the data availability trend and data consumption trend. Note that the time scale is different for each kind of data.

资料来源：Pablo Villalobos et. al., Will we run out of data? An analysis of the limits of scaling datasets in Machine Learning, arXiv:2211.04325v1[cx.LG] 26 Oct 2022

数据要素必须是数据资产吗？

- 根据生产要素的定义，任何在生产过程中使用的、具有稀缺性的投入都可以作为生产要素。
- 无论数据是短期或一次性用于生产过程，还是长期用于生产过程，都是生产要素，都应进入核算范围。
- 由于数据资产对经济增长的贡献更持久，同时由于数据资产核算还涉及对存量的核算，核算难度更大，因此数据核算的研究重点应该放在数据资产核算。

DZTT数据核算的研究进展

- DZTT是2019年4月成立的数字化问题核算的国际专家组，其研究成果和建议提交给国民经济核算的顾问专家咨询组（AEG）和国民账户秘书处工作组（ISWGNA），作为SNA修订的基础工作。
- 截至目前，DZTT已经在AEG的第13次、14次、15次、19次和20次会议上汇报了数据核算的研究进展。
- 关键问题
 - 数据的生命周期是多长？所有数据是否符合资产定义？
 - 数据属于生产资产还是非生产资产？
 - 对数据资产最合适估价方法是什么？
 - 在应用成本法时，应包括哪些成本？
 - 如何记录与持续更新数据库相关的支出？
 - 如何记录产生数据和使用数据的不同方式？

DZTT数据核算的研究进展（续）

明确了数据与可观察现象（OP）的关系

数据由OP的信息元素组成，一旦OP的信息元素以数字格式记录和存储，就成为数据。

更新了数据的定义

通过访问和观察现象产生的信息内容；以数字格式记录、组织和存储这些现象中的信息元素，这些信息元素在生产活动中使用时会产生经济效益。

确认了数据的生产资产属性

从分类看应归类为生产资产，是知识产权产品资产的一部分。如果可能，应当为数据单独创建一个新的资产类别。

明确了不同来源数据的属性

所有自给性生产的数据都被视为资本形成，通过市场交易购买的数据，如果打算在一年及以上的时间内使用，则应将其资本化，否则应作为中间消耗。

提出自给性数据生产核算的首选方法是成本法

规划、准备和开发数据生产战略的成本；

与访问、记录和存储OP中嵌入的信息相关的成本；

与设计、组织、测试和分析数据以从中得出信息和结论相关的成本。

明确了数据的经济所有者

以数字格式收集、记录、组织和存储OP的信息元素，从而创建数据的经济单位是数据的经济所有者。

OP所指的个人或实体并非数据的经济所有者。

数据生产要素核算对象

- 对何种数据进行核算？
 - 识别出以数据产出为主要产品的活动，其他与数据相关的活动则作为辅助活动，无需进行单独记录。
 - 建议将组织中专门的数据收集、数据存储和数据分析活动包含在内，而与组织的其他业务活动联系更为密切的数据应用排除在外，并将组织在其他业务活动中完成的辅助性数据活动排除在外。
 - 建议只把存储在数据中心（可以是本地的也可以是在外部托管的）或由数据专业人员建立的数据库中存储的数据纳入核算范围。
- 哪些机构部门应进行数据核算？
- 数据核算的结果是什么？

数据生产要素核算对象

- 对何种数据进行核算？
- 哪些机构部门应进行数据核算？
 - 企业部门是数据生产最重要的部门，也是从数据中获取经济价值的最主要的部门，建议在非金融企业部门中设置数据企业子部门，即依托数据资产提供主要产品的企业的集合。
 - 通过行政记录或统计调查收集数据是政府部门的基本职能之一，因此也是数据核算的主体。建议在政府部门中为国家统计系统（包括统计局系统和各部门的专业统计机构）单独设置一个子部门，对该子部门进行单独核算。
 - 住户部门虽然也有采集、存储和分析数据的行为，但是其数据活动的专业化程度较低，通常不会长期积累和存储数据，在当前的数据核算中可以不考虑该部门。
- 数据核算的结果是什么？

数据生产要素核算对象

- 对何种数据进行核算？
- 哪些机构部门应进行数据核算？
- 数据核算的结果是什么？
 - 在国民账户体系的核心账户中记录数据生产要素涉及到生产核算、收入分配核算、资本形成核算和存量核算等内容。
 - 主要核算指标包括数据生产的增加值、数据资产的形成额、数据资产存量。

数据产出的估价

- 成本法是优先推荐的方法。
 - 劳动投入
 - 中间投入
 - 与资本有关的投入：固定资本消耗，营业盈余或混合收入净额。
- 把数据专业人员与其他业务人员区分开。
- 把数据生产使用的资产与其他资产区分开。

数据资产存量估计

- 总成本
- 永续盘存法
 - 利用企业组织资本存量衡量数据价值 (Li等 , 2019)
 - 折旧率估计 : 前瞻性利润模型 (Li & Hall , 2018)

总结

- 数据具备成为生产要素的全部条件，应当被视为生产要素。
- 数据生产要素核算具有特有难题，值得核算研究者“探险”。
- 统计部门应与相关机构密切合作，奠定扎实数据资产核算的微观数据基础。
- 尽快设计和组织数据要素的试验性核算，为全面实施数据要素核算提供经验。

谢谢！