

基于物理-事理-人理的数据要素市场影响 因素分析

黄朝椿¹, 魏云捷^{2,3}

(1. 中国科学院大学经济与管理学院, 北京 100190; 2. 中国科学院数学与系统科学研究院, 北京 100190;
3. 中国科学院预测科学研究中心, 北京 100190)

摘要 数据要素作为新型生产要素, 在数字经济发展中起着基础性、关键性作用。本文基于 356 份问卷调查所得的数据, 从物理-事理-人理角度探究了数据要素市场的影响因素。首先, 通过理论分析和文献梳理, 基于 WSR 系统方法论, 将影响数据要素市场建设的 42 个因子分为物理维、事理维和人理维; 其次, 基于物理、事理、人理分别进行主成分分析, 从中提取影响三个维度的重要因素; 最后, 根据所提取的重要因素, 通过验证性因子分析, 构建数据要素市场影响因素模型。实证研究表明, 数据要素的供给是关键, 数字技术的突破是基础, 数据价值的测度是前提, 数据收入的分配是动力, 数据安全的保护是底线, 专业人才的培养是重点。最后, 提出了突破关键数字技术、加强市场制度设计、制定数据产权法、加快专业人才培养等政策建议。

关键词 WSR 系统方法论; 数据要素市场; 主成分分析; 验证性因子分析

Analysis of Data Elements Market Influencing Factors: Evidence from WSR Methodology

HUANG Chaochun¹, WEI Yunjie^{2,3}

(1. School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China;
2. Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;
3. Center for Forecasting Science, Chinese Academy of Sciences, Beijing 100190, China)

Abstract As a new type of production factor, data plays a fundamental and critical role in the development of digital economy. Based on the data obtained from 356 questionnaires, this paper explores the influencing factors of data elements market

收稿日期: 2023-04-25

基金项目: 国家自然科学基金 (72171223, 71801213, 71988101)

Supported by National Natural Science Foundation of China (72171223, 71801213, 71988101)

作者简介: 黄朝椿, 博士研究生, 研究方向: 数字经济, E-mail: 105042223@qq.com; 通信作者: 魏云捷, 中国科学院数学与系统科学研究院副研究员, 研究方向: 经济预测理论与方法、经济政策分析, E-mail: weiyun-jie@amss.ac.cn.

from the WSR methodology perspective. Firstly, by theoretical analysis and literature review, based on WSR methodology, the 42 factors influencing the construction of data elements market are divided into Wuli, Shili and Renli; secondly, principal component analysis is conducted based on Wuli, Shili and Renli respectively, from which the important factors influencing the three dimensions are extracted; finally, based on the extracted factors, we construct a data elements market impact model through validation factor analysis. The empirical study shows that the supply of data elements is the key, the breakthrough of digital technology is the foundation, the measurement of data value is the premise, the distribution of data income is the driving force, the protection of data security is the bottom line, and the cultivation of professional talents is the focus. Finally, policy suggestions such as breakthroughs in key digital technology, strengthening the design of market system, formulating data property rights law, and accelerating the construction of professional talents are put forward.

Keywords WSR system methodology; data elements market; principal components analysis; validated factor analysis

1 引言

数字经济是推进中国式现代化的重要引擎。数据是数字经济的关键要素。数据要素市场的建设,就是要推动数据要素的优化配置,从而实现在资源优化配置基础上的经济高质量发展。2022年12月,中共中央、国务院发布《关于构建数据基础制度更好发挥数据要素作用的意见》,对数据要素市场建设作出了明确部署。十四届全国人大一次会议通过的新一轮国务院机构改革方案,明确组建国家数据局,负责协调推进数据基础制度建设。加快数据要素市场建设,从政策、制度、技术层面都已经具备了良好的基础。

近年来,学界、业界对于数据要素市场的研究明显增多(洪永淼和汪寿阳(2021),邱东(2021),萧政(2021))。国内外学者关于数据要素市场的研究主要集中在以下三个领域(冯珏(2023)):一是数据要素市场中数字平台发展规则研究(Liang et al. (2018))。二是企业生产模式转变研究,例如数字化转型、金融科技等(Sharon and Zandbergen (2017))。三是政府行为与相关政策性的研究,例如政府信息化、电子政务等(Rubinfeld and Gal (2018))。黄朝椿(2022)探究了基于供给侧的数据要素市场建设的三个问题,提出了设立国家数据管理局、建立国家数据中心等建议。刘金钊和汪寿阳(2022)回顾了国际数据要素市场化建设的战略布局和我国数据交易市场发展进程中应对困境的关键举措,立足于我国数据要素市场建设实践,提出了四点对策建议。目前,学者们对数字政府治理战略以个案经验为主。如英国致力打造“平台驱动的数字政府”、丹麦打造无缝隙互动型的电子政务,也有学者着眼于多国数字政府战略的比较视角研究(Agion and Howit (1998))。根据要素市场的现状,学界普遍认为,目前数据要素市场主要体现为确权难、定价难、入场难、互信难、监管难“五难”问题。解决这些问题,需要深入研究数据要素市场的本质特征和影响因素。

数据要素市场是一个新型市场,具有很强的专业性、复杂性,总体上还处在一个起步探索阶段,无论是国内还是国外,从实践上讲,没有形成成功的模式和建设经验;从理论上讲,还没有成熟的数据要素市场理论。数据要素市场是一个复杂系统,需要运用新的思想和方

法,特别是从系统方法的角度进行深入研究.本文运用“物理-事理-人理”系统方法(WSR方法论),基于问卷调查数据对数据要素市场进行分析和研究.WSR方法论最初的思想源于钱学森,1994年由顾基发和朱志昌正式提出,并高度凝练地概括为“懂物理、明事理、通人理”的实践准则(寇晓东和顾基发(2021)).WSR作为一种思想,其核心是物理、事理和人理是系统实践中需要综合考察的三个方面,在处理复杂问题时既要考虑对象的物的方面(物理,W),又要考虑这些物如何更好地被运用的事的方面(事理,S),还要考虑到认识问题、处理问题和实施管理决策都离不开人的方面(人理,R).WSR方法论被广泛地用于经济学、管理学相关问题的研究中.朱晓武和张雪祺(2021)运用“物理-事理-人理(WSR)”方法论探究分布式自治组织运行机制与监管策略研究.

本文基于356份问卷调查所得的数据,从物理-事理-人理的角度探究了影响数据市场要素的主要因素,具体步骤如下:首先,进行理论分析和文献梳理,深入分析数据要素市场相关的法律制度、政策文件、数据资料和发展现状,根据学术期刊和实践报告的研究梳理,提取专家学者和实践工作者关注的要素市场的影响因素;其次,基于WSR方法论,将影响数据要素市场建设的42个因子分为物理维、事理维和人理维;再次,基于物理、事理、人理分别进行主成分分析,并分别从中提取影响三个维度的重要因素;最后,基于所提取的重要因素进行验证性因子分析,在主成分分析的基础上,构建数据要素市场影响因素模型.

本文的创新点主要有以下四点:第一,基于调研问卷数据,通过实证研究,首次定量探究数据要素市场建设的影响因素,为相关政策制定提供了重要参考;第二,成功将WSR系统方法论用于数据要素市场影响因素的研究,分析了数据要素市场物理维、事理维、人理维因素的具体影响机理,填补了这一方法论运用于数据要素市场研究的空白;第三,综合运用主成分分析和验证性因子分析方法,构建了数据要素市场影响因素模型,得出了物理、事理、人理三个维度因素对于数据要素市场建设的影响系数及各自内部影响因素系数;第四,从理论上解释了数据要素市场建设的基本着力点和关键因素,提出了下一步数据要素市场建设的政策建议.

2 基于 WSR 方法论的数据要素市场影响因素分析

本文利用中文知识文献库中国知网(<https://www.cnki.net/>)、万方数据知识服务平台(<http://www.wanfangdata.com.cn/>)、超星期刊数据库(<http://qikan.chaoxing.com/>)、维普资讯中文期刊服务平台(<http://m.qikan.cqvip.com/>)和外文知识文献库 WOS 数据库(<http://webofscience.com/>)、Springer 数据库(<https://www.springer.com/cn>)和 Elsevier 数据库(<https://www.sciencedirect.com/>)进行搜索,截止2022年9月30日,对于有关数据要素市场文献进行检索并适当进行数据清洗之后,得到61篇中文文献和38篇外文文献,一共99篇中外文文献.

根据每个因素至少有两个文献提到的原则,一共提取了455个因子.经过适当整合归纳,剔除其中重复因素,筛选出42个数据要素市场影响因子(表1).

根据以上文献调研提取的影响因子,运用WSR系统方法论思想,对于影响数据要素市场影响的因素,从物理、事理、人理三个方面进行分析.

表 1 文献调研数据要素市场影响因素提出因子汇总表

编号	影响因素	编号	影响因素	编号	影响因素
b1	市场交易机构	b15	数据供求机制	b29	加强政府职责
b2	市场体系建设	b16	限制场外交易	b30	支持数据融资
b3	数据交易技术	b17	培育数商体系	b31	统一质量标准
b4	市场交易成本	b18	限制数据垄断	b32	数据产业生态
b5	隐私保护技术	b19	加强数据执法	b33	市场主体性质
b6	公共数据开放	b20	建设交易场所	b34	算法歧视行为
b7	规范产品形态	b21	数据科技投入	b35	公众认知水平
b8	数据定价机制	b22	数据资产入表	b36	市场信用体系
b9	数据采集技术	b23	数据专项基金	b37	基础理论研究
b10	平台公司数据	b24	数据价值计量	b38	经济发展水平
b11	个人数据入市	b25	数据产权清晰	b39	普及数据知识
b12	企业数据入市	b26	数据参与分配	b40	保护个人隐私
b13	数据分析技术	b27	统一数据定价	b41	数据专业人才
b14	市场基础设施	b28	强制信息披露	b42	交易失信行为

2.1 物理维 —— 物质技术因素

根据 WSR 系统方法论思想, 物理维因素主要是指客观物质世界、法规、规则以及物质运动的机理, 通常要用到自然科学知识, 主要回答研究对象是什么的问题。

综合相关研究, 物理因素的基本内涵有: ① 物理是象征着本体论 (Ontology) 的客观存在 (自然、社会; 具体、抽象), 包括物质及其组织结构 (顾基发和唐锡晋 (2000)). ② 物理是阐述自然客观现象和客观存在的定律、规则, 通过数据、方程、描述、及其他方式表达出来 (高飞 (2000)). ③ 物理就是指物质运动和技术作用的一般规律, 是一种客观存在, 不以人的意志为转移。许国志 (1981) 谈到: 我有意的把运筹学和系统工程中的“硬”模型和技术工具与自然科学的方法和规则归并在一起统称为物理, ……简言之, 物理是指管理过程和管理对象中可以并应该由自然科学、工程技术和“硬运筹方法”描述和处理的层面。

具体到数据要素市场, 物理维因素包括以下三个方面的内容。

1) 数据要素市场的构成要件。主要是指一个市场的主要构成要件, 由可供交换的商品、提供商品的卖方、人格化 —— 买方组成。这里的商品既包括有形的物质产品, 也包括无形的服务, 以及各种商品化了的生产要素, 如劳动力、土地、资本、技术等。商品是物, 需要由它的所有者 —— 出卖商品的当事人即卖方通过市场进行交换, 卖方或商品所有者就成为向市场提供商品的代表者, 并作为市场供求中的供应方成为基本的市场构成要素。卖方向市场提供一定量的商品, 作为既有需求又具备支付能力的购买者, 就成为商品交换的另一方形成市场需求, 最终决定商品交换能否实现。此外, 市场还有其他一些促成交易、降低交易成本的客观主体存在, 如数据交易机构、交易中介机构。

2) 数据要素市场的运动规律。市场机制作为市场特有的调节方式、调节功能和特殊的运动过程, 离不开供求这一基本要素。同时, 市场供求也不是孤立存在, 其运动变化直接受市场价格及市场竞争状况的制约。因此, 构成市场机制运动的三大基本要素是价格、供求、竞争,

不论市场性质、规模、范围如何,这三大直接要素不会变.这三大要素的组合及交互运动就是市场经济的基本规律即价值规律、供求规律、竞争规律、平均利润率规律、货币流通规律等共同作用于市场的结果.在供求、竞争机制的作用下,价格总是围绕价值上下波动.数据要素市场也一样,存在着市场运行的一般规律.运行机制就是一种客观存在,无论什么市场,都以市场机制有效发挥作用为基本前提.

3) 数据要素市场的法规规则.这主要是指对于物理因素的描述和研究方法,特别是指数学、物理、化学等自然科学方法,比如通过数据、方程、公式、定理描述物质运动规律和基本原理.具体到数据要素市场,算法(Algorithm)就是典型的法规规则,通过算法提出对于所要解决的问题进行准确而完整的描述,表现为一系列解决问题的清晰指令,描述算法的方法有自然语言、结构化流程图、伪代码和PAD图等.算法能够对一定规范的输入,在有限时间内获得所要求的输出.如果一个算法有缺陷,或不适合于某个问题,执行这个算法将不会解决这个问题.需要指出的是,这里的法规规则不是指一般政治的、法律的、社会的规则,不包括国家法律制度、市场规则等,因为这些法律制度并不是客观存在,而是一种反映客观存在的上层建筑.

对以上三个方面的内容进行具体分析,根据文献调研中提取的影响因子,可以列出42个影响因素中,有15个因子属于物理因素(表2).

表2 数据要素市场物理维影响因素

维度	因子数量	包含内容
物理维	15个	b1(市场交易机构); b2(市场体系建设); b3(数据交易技术); b4(市场交易成本); b5(隐私保护技术); b6(公共数据开放); b7(规范产品形态); b8(数据定价机制); b9(数据采集技术); b10(平台公司数据); b11(个人数据入市); b12(企业数据入市); b13(数据分析技术); b14(市场基础设施); b15(数据供求机制)

2.2 事理维 —— 制度政策因素

事理是指做事的道理,也即涉及某项系统项目(问题)处理过程中人们面对的客观存在及其规律时介入的机理(张彩江和孙东川(2001)),或者说,就是如何去理解、发现事物内在机理及运动法则的最好办法和最佳途径,主要解决如何去安排谋划,通常用到运筹学与管理科学方面的知识,主要回答怎样去做的问题.

归纳相关论述,事理因素的基本内涵有:①物有常规,事有定理,这些基本规律可以称为事理.②人类某些活动的理性行为就是事理.③如物理学一样,事理学要区分事理现象和事理本质,事理现象是指人们在事理活动中直观把握的一切外部形态的东西及联系,……构成办事过程的各种事实、现象,就是事理现象.事理的本质指隐藏在这些现象背后的办成办好事情应遵循的道理、规律等(苗东升(1998)).④事理意味机理(Mechanisms),强调世界的联系和过程(顾基发和唐锡晋(2000)).⑤事理是指方法,或那些帮助人们基于世界和客观存在的机理之上的有效处理事务的方法(顾基发和高飞(1998)).⑥事理是指管理者介入和执行

管理事物的方式和规律, 包括如何感知、看待、认识、思考、描述和组织管理对象和管理过程(朱志昌(2000)).

根据对于事理内涵的界定, 数据要素市场事理维因素可以概括为以下三个方面:

1) 构建数据要素市场需要遵循的道理. 就是数据要素市场的形成、发展需要遵循什么样的一般规则, 有哪些事理现象和事理本质? 无论是马克思主义市场理论, 还是微观经济学, 对于市场的构成要件, 即市场交易什么、如何交易、谁来交易的问题, 在理论上都有共识. 在实践中, 无论商品市场, 还是要素市场, 构成市场的基本要素都十分清楚和明确. 但是, 现实的情况是, 市场千差万别, 并没有因为只要有完整的市场构成要素, 就一定会出现一个繁荣的市场. 根据市场设计理论, 一个市场的形成, 需要考虑市场厚度、市场拥堵、市场行为的安全性和简易性, 数据要素市场的构建, 必须满足这三个条件, 不然数据要素市场将无法形成和发展.

2) 构建数据要素市场的有效方法是什么? 针对数据要素市场存在的问题, 为了推动数据要素市场发展, 采取了什么样的思路、理论、方法来解决这些问题. 针对数据要素市场存在的确权难、定价难、入场难、互信难、监管难等问题, 国家出台了一系列的法律规定和政策文件, 地方政策在推动国家政策的落实中, 进行了大量的探索, 包括出台地方性条例, 努力解决数据要素市场的一系列问题. 学界针对数据要素市场存在的问题, 进行了大量开创性、探索性研究, 回答了构建数据要素市场的诸多理论问题. 市场主体在实践中大胆创新, 从交易模式、交易方式、数据产业链的形成等方面, 形成了一些有效的思路和办法.

3) 构建数据要素市场的介入管理. 如何正确发挥政府与市场的作用, 是市场经济条件下的一对永恒关系. 一个市场的形成, 会根据其本身的特点, 有多种不同的模式. 有的市场大致满足完全竞争假设前提, 比如简单的劳动力市场, 一些土地市场, 民间借贷等, 买卖双方具有对称的信息、平等的地位, 通过市场机制就能完成生产要素的流通交易; 有的市场满足完全竞争市场的部分假设条件, 由于信息不完全、不对称, “逆向选择” “道德风险” 的存在, 需要政府这只看得见的手发挥作用, 通过法律和制度来弥补市场缺陷; 有的市场除了具备市场的要件外, 满足完全竞争的大多数假设条件都不满足, 这就需要政府发挥积极作用, 用一种科学的制度设计, 引导培育一个制度市场. 数据要素市场就是后两种市场, 需要发挥有为政府的作用, 通过政策指导和引导形成和发展数据要素市场.

根据上述对于事理维因素的界定, 从文献调研提取的 42 个影响因子中, 属于事理维因子的有 18 个(表 3).

中国特色社会主义市场经济体制最为本质的特点, 就是统筹发挥“有为政府”和“有效市场”的作用, 使市场在资源配置中发挥决定性作用, 更好发挥政府作用, 这就包括各级政府及其管理部门. 政府发挥作用的领域很多, 包括加强宏观调控政策, 实施产业政策, 消除要素瓶颈, 降低交易成本, 建设基础设施, 提供公共服务, 推动实现规模经济效应、范围经济效应、聚集经济效应创造条件等. 实践中, 生产要素市场的形成都有一个从自发产生到培育发展的过程, 这就需要政府发挥积极有效的推动作用. 对于市场进行有效监管, 这是弥补市场失灵、有效发挥市场作用的必然要求. 为了保证数据要素市场对于数据要素资源的高效配置, 必然同其他生产要素一样, 加强对于数据要素市场的必要监管. 因此, 事理维因子可以作为发挥有为政府作用的基本切入点和着力点.

表3 数据要素市场事理维影响因素

维度	因子数量	包含内容
事理维	18个	b16(限制场外交易); b17(培育数商体系); b18(限制数据垄断); b19(加强数据执法); b20(建设交易场所); b21(数据科技投入); b22(数据资产入表); b23(数据专项基金); b24(数据价值计量); b25(数据产权清晰); b26(数据参与分配); b27(统一数据定价); b28(强制信息披露); b29(加强政府职责); b30(支持数据融资); b31(统一质量标准); b32(数据产业生态); b33(市场主体性质)

2.3 人理维 —— 环境认知因素

人理指做人的道理,通常要用人文与社会科学的知识,是一个人文环境,主要回答最好怎么做的问题。

归纳相关论述,人理因素的基本内涵有:①人理是关注涉及系统项目中所有团体相互之间的主观关系,包括顾客、权力当局、组织者、专家、潜在业主、使用者、操作者、受益和受损者(顾基发和唐锡晋(2000)).②人理是指基于心理学、社会学、组织行为学,结合文化、传统、价值、观念等来把人组织在一起有效地开展工作的方法(顾基发和高飞(1998)).③人理是指管理对象和过程中的人与人之间的关系,包括管理主体与主体之间的关系.研究人理包括研究管理过程中管理主体之间如何相互沟通、学习、调整、谈判、或者排斥甚至损害,涉及人与人之间的得与失、爱与恨、信任与怀疑、扶持与操纵,等等(朱志昌(2000)).

根据对于人理内涵的界定,数据要素市场人理维因素,可以概括为以下几个方面:

1) 市场主体之间的相互关系.市场主体是市场上从事交易活动的组织和个人,既包括自然人,也包括以一定组织形式出现的法人;既包括营利性机构,也包括非营利性机构,企业、居民和其他非营利性机构构成了市场主体的诸要素.市场主体的关系,既包括交易形成的经济关系,也包括由此带来的社会关系.在数据要素市场上,除了因数据的交易形成的经济关系外,更为重要的关系是由此产生的数据伦理、数据主权、数据安全问题,这些关系的复杂程度远远大于交易形成的经济关系.

2) 对于数据要素市场的认知和接受程度.任何一个市场,要满足市场厚度的条件,最为重要的就是市场主体的认知和接受程度.数据要素作新型生产要素,数据要素市场作为新型市场,之所以出现有数无市、有市无数的现象,原因之一,就是市场主体对于这一市场还不了解、不熟悉.在问卷调查中,有相当部分的人认为,就是想出售数据,也不知道应该去哪里卖、如何卖、卖给谁.因此,数据要素市场的形成和发展,需要加快普及相关知识,营造良好的人文环境和社会环境.

3) 数据要素市场的专业人才.任何一个专业的市场,专业的人才是基础.从市场的视角看,专业的人才能够有效组织市场,通过专业知识设计市场,弥补市场失灵,降低市场交易成本,促进市场的繁荣发展.比如,《证券法》就规定,从事证券业务的人员应当品行良好,具备从事证券业务所需的专业能力.相比证券市场,数据要素市场更为复杂、更为专业.数据要

素市场的数据交易, 涉及国家和广大市场主体的切身利益, 关系经济发展、数据安全、个人隐私, 具有较强的公共属性. 因此, 培养打造一支专业的数据要素市场人才队伍, 对于加快构建数据要素市场具有重要的基础性作用.

根据以上三个方面的界定, 可以从文献调研提取的 42 个因素中梳理归纳数据要素市场的人理维因素 (表 4), 一共是 9 个.

表 4 数据要素市场人理维影响因素

维度	因子数量	包含内容
人理维	9 个	b34(算法歧视行为); b35(公众认知水平); b36(市场信用体系); b37(基础理论研究); b38(经济发展水平); b39(普及数据知识); b40(保护个人隐私); b41(数据专业人才); b42(交易失信行为)

3 数据采集与分析

基于所开发的数据要素市场影响因素调查问卷, 本研究于 2023 年 1 月 7 日进行了网络问卷调查, 选取中国联通数据部门作为调查对象, 共发放问卷 365 份, 回收有效问卷 356 份.

3.1 描述性统计

本次调研有效样本的描述性统计分析结果如下:

从表 5 至表 8 可以看出, 发放问卷 356 份, 其中男性 249 人, 女性 107 人, 未填性别 27 人. 年龄结果以 30 岁以下和 31~40 岁的为主, 占总样本比例分别为 34.83% 和 40.45%. 样本职业分布以数据主管部门和数据企业为主, 占总样本比例分别为 42.98% 和 43.26%. 样本职务分布以普通员工为主, 占总样本比例为 94.38%.

表 5 有效调研样本性别分布

性别	问卷填表人数	占比
男	249	69.94%
女	107	30.06%

表 6 有效调研样本年龄分布

年龄	问卷填表人数	占比
30 岁及以下	124	34.83%
31~40 岁	144	40.45%
41~50 岁	79	22.19%
51~60 岁	8	2.25%
60 岁以上	1	0.28%

表 7 有效调研样本职业分布

职业	问卷填表人数	占比
数据主管部门	153	42.98%
数据企业	154	43.26%
专家	19	5.34%
网民	30	8.43%
交易机构	0	0

表 8 有效调研样本职务分布

职务	问卷填表人数	占比
高层领导	3	0.84%
中层管理	15	4.21%
普通员工	336	94.38%
其他	2	0.56%

3.2 问卷信度检验

根据问卷调查数据,对所有 42 个变量进行信度检验(表 9).

1) 运用克隆巴赫系数(Cronbach's alpha)对调查问题进行信度测试.一般来说, α 系数愈高,则工具的信度愈高.在基础研究中,信度至少应达到 0.80 才可接受;但是,在探索性研究中,信度只要达到 0.70 就可接受,介于 0.70~0.98 均属高信度;如果低于 0.35 则为低信度,必须予以拒绝.根据 SPSS 统计软件的计算,本文针对所有变量数据计算得到的 Cronbach's α 系数为 0.9830,对于探索性研究而言,属于高信度量表,说明调查问卷所得数据的信度很好.

2) 对调查问卷变量矩阵进行 KMO 检验. KMO 检验用于检查变量间的相关性和偏相关性,取值在 0~1 之间, KMO 统计量越接近于 1, 变量间的相关性越强, 偏相关性越弱, 因子分析的效果越好. 实际分析中, KMO 统计量在 0.7 以上时效果比较好; 当 KMO 统计量在 0.5 以下时, 一般就不考虑应用因子分析法, 而应考虑重新设计变量结构或者采用其他统计分析方法. 本文所有变量 KMO 统计量为 0.973, 完全符合主成分分析条件.

3) 进行 Bartlett's 球状检验. Bartlett 检验用于检验相关阵中各变量间的相关性, 检验各个变量是否各自独立. 在因子分析中, 若拒绝原假设 H_0 , 则说明可以做因子分析; 否则, 则说明这些变量可能独立提供一些信息, 不适合做因子分析. 由 STATA 检验结果显示 p 值 = 0.000, 说明符合标准, 数据呈球形分布, 各个变量在一定程度上相互独立. 本文所有数据的 Sig. 值接近 0, 适合做因子分析.

表 9 所有变量 KMO 和 Bartlett 检验

KMO 和 Bartlett 的检验		
取样足够度的 Kaiser-Meyer-Olkin 度量		0.973
	近似卡方	15226.16
Bartlett 的球形度检验	自由度	861
	p 值	0

表 10 物理维 15 个变量 KMO 和 Bartlett 检验

KMO 和 Bartlett 的检验		
取样足够度的 Kaiser-Meyer-Olkin 度量		0.958
	近似卡方	3925.829
Bartlett 的球形度检验	自由度	105
	p 值	0

表 11 事理维 18 个变量 KMO 和 Bartlett 检验

KMO 和 Bartlett 的检验		
取样足够度的 Kaiser-Meyer-Olkin 度量		0.957
	近似卡方	5268.154
Bartlett 的球形度检验	自由度	153
	p 值	0

表 12 人理维 9 个变量 KMO 和 Bartlett 检验

KMO 和 Bartlett 的检验		
取样足够度的 Kaiser-Meyer-Olkin 度量		0.926
	近似卡方	2482.845
Bartlett 的球形度检验	自由度	36
	p 值	0

4 基于主成分分析的数据要素市场影响因素分析

4.1 数据要素市场构建: 物理维

通过对物理维 15 个数据变量做主成分分析, 删去特异性方差占比均不超过 0.6, 因此 15 个变量均可以进行主成分因子分析, 测试结果见表 13.

表 13 对物理维内部数据主成分因子分析的结果

因子	特征值	差值	方差贡献比例	方差累计贡献比例
Wu_Factor1	9.0248	7.9757	0.6017	0.6017
Wu_Factor2	1.0491	0.3978	0.0699	0.6716
Wu_Factor3	0.6513	0.0581	0.0434	0.715
Wu_Factor4	0.5932	0.0735	0.0395	0.7546
Wu_Factor5	0.5197	0.0399	0.0346	0.7892
Wu_Factor6	0.4797	0.0701	0.032	0.8212

注: 展示方差累计贡献率高于 80% 的因子.

根据测试结果, 选取了方差累计贡献率大于 80% 的 6 个因素, 这也是对于数据要素市场物理维影响因素最大的六个因子, 方差累计贡献率为 82.12%, 这意味着这 6 个因子表征和解释了物理维所有数据变量的大部分影响因素.

对于以上主成分分析得出的 6 个因子进行旋转因子载荷分析, 得出萃取因素与测量变量之间关系的因素负荷量矩阵 (factor loading matrix, 以 A 表示) (表 14).

因子载荷类似于回归系统, 载荷数值反映的是各潜在变量对于测量变量的影响程度. 根据 Tabachnicat and Fidell (2007) 的研究, 负荷量大于 0.71 是非常理想的状态, 大于 0.63 属于非常好的状态, 大于 0.55 则属于好的状态, 大于 0.45 属于普通状态, 而小于 0.32 时, 该因素解释不到 10% 的观察变量的变异量, 这些因素就可以删除.

根据上述判断标准: Wu_Factor1 对 b1、b2 和 b5 的负荷量分别为 0.802、0.729 和 0.576, 表明 Wu_Factor1 对这 3 个因子具有较强的解释力; 同样, Wu_Factor2 对于 b12(0.711)、

表 14 旋转后因子载荷系数表

变量	Wu_Factor1	Wu_Factor2	Wu_Factor3	Wu_Factor4	Wu_Factor5	Wu_Factor6
b1	0.802	0.081	0.26	0.11	0.335	0.118
b2	0.729	0.366	0.081	0.3	0.103	0.18
b3	0.487	0.247	0.178	0.36	0.456	0.382
b4	0.212	0.227	0.215	0.19	0.198	0.858
b5	0.576	0.281	0.166	0.583	0.007	0.083
b6	0.227	0.17	0.219	0.795	0.269	0.183
b7	0.331	0.511	0.123	0.501	0.244	0.282
b8	0.51	0.336	0.275	0.286	0.326	0.341
b9	0.268	0.258	0.208	0.218	0.751	0.196
b10	0.179	0.217	0.764	0.212	0.353	0.067
b11	0.175	0.217	0.865	0.113	0.049	0.205
b12	0.238	0.711	0.421	0.2	0.159	0.141
b13	0.101	0.563	0.346	0.37	0.424	0.201
b14	0.461	0.582	0.18	0.204	0.248	0.296
b15	0.345	0.536	0.254	0.117	0.491	0.174

注: 载荷系数绝对值大于 0.55 已加粗.

b13(0.563)、b14(0.583) 和 b15(0.536) 具有较强的解释力; Wu_Factor3 对于 b10(0.764)、b11(0.865) 具有较强的解释力; Wu_Factor4 对于 b5(0.583)、b6(0.795) 具有较强的解释力; Wu_Factor5 对于 b9(0.751) 具有较强的解释力; Wu_Factor6 对于 b4(0.858) 具有较强的解释力.

根据上述分析结果, 可以列出六个因子主要解释的数据变量关系 (表 15).

表 15 六个因子主要解释的数据变量

Wu_Factor1	Wu_Factor2	Wu_Factor3	Wu_Factor4	Wu_Factor5	Wu_Factor6
b1 市场交易机构; b2 市场体系建设 b5 隐私保护技术	b12 企业数据入市; b13 数据分析技术; b14 市场基础设施; b15 数据供求机制	b10 平台公司数据; b11 个人数据入市	b5 隐私保护技术; b6 公共数据开放	b9 数据采集技术	b4 市场交易成本

4.2 数据要素市场构建: 事理维

通过对事理维的 18 个数据变量做主成分因子分析, 结果如下.

表 16 展示了对事理维数据变量做主成分因子分析的结果. 对这些数据变量影响最大的六个因子对方差的累计贡献率为 80.27%, 这意味着这 6 个因子表征了事理维所有数据变量的影响因素的大部分.

表 17 汇报了旋转因子载荷和特异性方差占比的结果. 根据载荷值判断标准, Shi_Factor1 对于 b24、b26、b27 和 b33 的载荷值分别为 0.68、0.683、0.64、0.616, Shi_Factor2 对于 b31 和 b32 的载荷值分别为 0.727 和 0.737, Shi_Factor3 对于 b20、b23 和 b30 的载荷值分别为 0.786、0.588、0.565, Shi_Factor4 对于 b19 和 b25 的载荷值分别是 0.732、0.721, Shi_Factor5 对于 b16 的载荷值为 0.803, Shi_Factor6 对于 b18 的载荷值为 0.871, 这些载荷值表明, 潜在变量对于测量变量都有较强的解释力.

根据载荷分析, 列出六个因子主要解释的数据变量表 (表 18).

表 16 对事理维内部数据主成分因子分析的结果

因子	特征值	差值	方差贡献比例	方差累计贡献比例
Shi_Factor1	11.0283	10.1754	0.6127	0.6127
Shi_Factor2	0.8529	0.0558	0.0474	0.6601
Shi_Factor3	0.7971	0.1002	0.0443	0.7043
Shi_Factor4	0.6969	0.1381	0.0387	0.7431
Shi_Factor5	0.5588	0.0449	0.031	0.7741
Shi_Factor6	0.5139	0.0411	0.0286	0.8027

4.3 对人理维内部数据主成分因子分析的结果

表 19 展示了对人理维数据变量做主成分因子分析的结果. 对这些数据变量影响最大的三个因子对方差的累计贡献率为 80.6%, 这意味着这 3 个因子表征了人理维所有数据变量的影响因素的大部分.

表 17 旋转因子载荷和特异性方差占比

变量	Shi_Factor1	Shi_Factor2	Shi_Factor3	Shi_Factor4	Shi_Factor5	Shi_Factor6
b16	0.229	0.224	0.165	0.267	0.803	0.109
b17	0.243	0.484	0.108	0.262	0.517	0.319
b18	0.199	0.122	0.191	0.222	0.082	0.871
b19	0.193	0.251	0.295	0.732	0.302	0.183
b20	0.206	0.135	0.786	0.31	0.08	0.198
b21	0.493	0.32	0.483	0.436	0.083	0.134
b22	0.267	0.465	0.503	0.356	0.187	0.173
b23	0.377	0.387	0.588	0.054	0.206	0.227
b24	0.68	0.421	0.271	0.25	0.128	0.169
b25	0.406	0.26	0.128	0.721	0.195	0.299
b26	0.683	0.272	0.212	0.37	0.201	0.25
b27	0.64	0.341	0.255	0.218	0.223	0.281
b28	0.155	0.455	0.333	0.087	0.361	0.499
b29	0.261	0.493	0.27	0.452	0.408	0.01
b30	0.446	0.247	0.565	0.032	0.396	0.149
b31	0.35	0.727	0.268	0.304	0.202	0.1
b32	0.332	0.737	0.216	0.237	0.244	0.223
b33	0.616	0.13	0.4	0.201	0.356	0.036

表 18 六个因子主要解释的数据变量

Shi_Factor1	Shi_Factor2	Shi_Factor3	Shi_Factor4	Shi_Factor5	Shi_Factor6
b24 数据价值计量; b26 数据参与分配; b27 统一数据定价; b33 市场主体性质	b31 统一质量标准; b32 数据产业生态	b20 建设交易场所; b23 数据专项基金; b30 支持数据融资	b19 加强数据执法; b25 数据产权清晰	b16 限制场外交易	b18 限制数据垄断

表 19 对人理维内部数据主成分因子分析的结果

因子	特征值	差值	方差贡献比例	方差累计贡献率
Ren_Factor1	6.1207	5.5034	0.6801	0.6801
Ren_Factor2	0.6173	0.1016	0.0686	0.7487
Ren_Factor3	0.5157	0.049	0.0573	0.806

运用 SPSS 软件, 对于以上 3 个因子进行旋转因子载荷分析, 得出萃取因素与测量变量之间关系的因素负荷量矩阵 (factor loading matrix) (表 20).

表 20 呈现了旋转因子载荷的结果. 根据因子载荷判断标准: Ren_Factor1 对 b36、b40、b41 和 b42 的负荷量分别为 0.761、0.774、0.685 和 0.731, 表明 Ren_Factor1 对这 4 个因

子具有较强的解释力; 同样, Ren_Factor2 对于 b38(0.8)、b39(0.784) 都具有较强的解释力; Ren_Factor3 对于 b34(0.82)、b35(0.669) 也具有较强的解释力.

根据上述分析结果, 可以列出 3 个因子主要解释的数据变量关系 (表 21).

表 20 旋转因子载荷和特异性方差占比

变量	Ren_Factor1	Ren_Factor2	Ren_Factor3
b34	0.217	0.339	0.82
b35	0.42	0.359	0.669
b36	0.761	0.078	0.5
b37	0.469	0.497	0.525
b38	0.266	0.8	0.379
b39	0.386	0.784	0.319
b40	0.774	0.44	0.158
b41	0.685	0.505	0.262
b42	0.731	0.364	0.303

表 21 三个因子主要解释的数据变量

Ren_Factor1	Ren_Factor2	Ren_Factor3
b36 市场信用体系;	b38 经济发展水平;	b34 算法歧视行为;
b40 保护个人隐私;	b39 普及数据知识	b35 公众认知水平
b41 数据专业人才;		
b42 交易失信行为		

5 验证性因子分析

根据主成分分析的结果, 物理维的第一个和第二个因子的特征值大于 1, 事理维和人理维仅有一个特征值大于 1. 因此我们选择特征值大于 1 的因子中载荷系数高于 0.5 的因子作为该维度的主要影响因素, 共有 15 个因子: 物理维的 b1(市场交易机构)、b2(市场体系建设)、b5(隐私保护技术)、b12(企业数据入市)、b13(数据分析技术)、b14(市场基础设施)、b15(数据供求机制); 事理维的 b24(数据价值计量)、b26(数据参与分配)、b27(统一数据定价)、b33(市场主体性质); 人理维的 b36(市场信用体系)、b40(保护个人隐私)、b41(数据专业人才)、b42(交易失信行为).

基于这些因素, 将调研得到的 356 个样本的数据要素市场影响因素数据采用验证性因子分析进行拟合运算和分析. 为了验证所构建数据要素市场影响因素的验证性因子分析的合理性, 选择以下三个样本组: 1) 将所有 42 个影响因素数据共同进行验证性因子分析拟合 (样本 A) (表 22); 2) 将样本 (1) 中因子载荷系数低于 0.76 的删除 (表 23), 再进行验证性因子分析拟合 (样本 B) (表 24); 3) 使用主成分分析筛选得到的 15 个影响因素数据共同进行验证性因子分析拟合 (样本 C) (表 25); 4) 将样本 (3) 中因子载荷系数低于 0.76 的因子剔除 (表 26), 再进行验证性因子分析拟合 (样本 D) (表 27).

表 23 显示了 42 个变量的标准载荷系数, 其中一些变量的因子载荷系数较低. 因此我们

删除标准载荷系数在 0.76 以下的因子 (b1, b2, b4, b5, b6, b9, b10, b11, b16, b18, b20, b28, b33), 形成样本 B, 并重新进行验证性因子分析.

表 22 样本 A 模型拟合指标

常用指标	χ^2	df	p	卡方自由度比 χ^2/df	GFI
判断标准	-	-	>0.05	<3	>0.9
值	3702.062	816	0	4.537	0.625
常用指标	RMSEA	RMR	CFI	NFI	NNFI
判断标准	<0.10	<0.05	>0.9	>0.9	>0.9
值	0.1	0.039	0.809	0.768	0.798
其它指标	TLI	AGFI	IFI	PGFI	
判断标准	>0.9	>0.9	>0.9	>0.9	
值	0.798	0.585	0.809	0.565	
其它指标	PNFI	SRMR	RMSEA 90% CI		
判断标准	>0.9	<0.1	-		
值	0.728	0.049	0.089~0.105		
Default Model: $\chi^2(861) = 15934.874, p = 1.000$					

表 23 样本 A 因子载荷系数表

Factor (潜变量)	测量项 (显变量)	Factor 非标准载荷 系数 (Coef.)	标准误 (Std. Error)	z (CR 值)	p	标准载荷系数 (Std. Estimate)
物理维	B1	1	-	-	-	0.687
物理维	B10	1.214	0.097	12.531	0	0.691
物理维	B11	1.2	0.109	11.049	0	0.605
物理维	B12	1.294	0.092	14.127	0	0.785
物理维	B13	1.241	0.085	14.524	0	0.809
物理维	B14	1.242	0.082	15.21	0	0.851
物理维	B15	1.225	0.084	14.623	0	0.815
物理维	B2	1.068	0.079	13.528	0	0.75
物理维	B3	1.226	0.083	14.83	0	0.828
物理维	B4	1.132	0.093	12.18	0	0.671
物理维	B5	0.981	0.077	12.761	0	0.705
物理维	B6	1.063	0.083	12.827	0	0.708
物理维	B7	1.198	0.082	14.557	0	0.811
物理维	B8	1.253	0.084	14.997	0	0.838
物理维	B9	1.178	0.088	13.331	0	0.738
事理维	B16	1	-	-	-	0.689
事理维	B25	1.08	0.076	14.272	0	0.788
事理维	B26	1.226	0.083	14.858	0	0.822
事理维	B27	1.236	0.084	14.767	0	0.817
事理维	B28	1.184	0.091	12.95	0	0.711

表 23 (续)

Factor (潜变量)	测量项 (显变量)	Factor 非标准载荷 系数 (Coef.)	标准误 (Std. Error)	z (CR 值)	p	标准载荷系数 (Std. Estimate)
事理维	B29	1.121	0.079	14.169	0	0.782
事理维	B30	1.23	0.089	13.844	0	0.763
事理维	B31	1.181	0.078	15.186	0	0.842
事理维	B32	1.174	0.077	15.175	0	0.841
事理维	B33	1.146	0.086	13.4	0	0.737
事理维	B17	1.062	0.077	13.872	0	0.764
事理维	B18	0.938	0.088	10.601	0	0.577
事理维	B19	1.075	0.078	13.85	0	0.763
事理维	B20	1.294	0.102	12.692	0	0.696
事理维	B21	1.264	0.085	14.849	0	0.822
事理维	B22	1.165	0.079	14.657	0	0.81
事理维	B23	1.249	0.09	13.95	0	0.769
事理维	B24	1.257	0.085	14.773	0	0.817
人理维	B34	1	-	-	-	0.764
人理维	B35	0.983	0.058	16.936	0	0.818
人理维	B36	0.894	0.056	15.91	0	0.777
人理维	B37	1.015	0.058	17.481	0	0.839
人理维	B38	0.989	0.061	16.317	0	0.793
人理维	B39	0.999	0.059	16.943	0	0.818
人理维	B40	0.883	0.056	15.65	0	0.766
人理维	B41	0.96	0.056	17.094	0	0.824
人理维	B42	0.895	0.054	16.555	0	0.803
数据要素市场	物理维	1	-	-	-	0.989
数据要素市场	事理维	1.048	0.084	12.436	0	1
数据要素市场	人理维	1.288	0.095	13.598	0	0.99

根据表 24 的结果可知, 样本 B 的拟合效果不理想, 诸多判断指标不显著. 因此, 我们进一步将主成分分析的结果引入, 使用主成分分析筛选得到的 15 个影响因素数据共同进行验证性因子分析拟合.

表 25 显示了样本 C 验证性因子分析拟合效果, 可以看出, 相较于样本 A 和 B, 样本 C 的 CFI、NFI、NNFI 等指标有了明显的改善. 但是根据表 26 的结果可知, 15 个变量的标准载荷系数中一些变量的因子载荷系数较低. 因此我们删除标准载荷系数在 0.76 以下的因子 (b1, b2, b5, b33), 形成样本 D, 并重新进行验证性因子分析.

根据表 25 和表 27 的结果可以发现, 样本 D 的拟合度指标 IFI、TLI、CFI 都大于 0.90, RMSEA 小于 0.1, 表明模型拟合度较好, 拟合结果优于样本 A、样本 B 和样本 C. 因此, 数据要素市场建设的影响因素图 (图 1), 即数据要素市场受到物理维、事理维和人理维的共同作用, 其中物理维主要受到稳定供求关系、基础设施完善度、企业数据入市、数据分析水平的影响, 占到 34.7%; 事理维主要受到数据价值计量、数据参与分配和统一数据定价因素的影

表 24 样本 B 模型拟合指标

常用指标	χ^2	df	p	卡方自由度比 χ^2/df	GFI
判断标准	-	-	>0.05	<3	>0.9
值	1909.796	374	0	5.106	0.705
常用指标	RMSEA	RMR	CFI	NFI	NNFI
判断标准	<0.10	<0.05	>0.9	>0.9	>0.9
值	0.108	0.032	0.858	0.83	0.846
其它指标	TLI	AGFI	IFI	PGFI	
判断标准	>0.9	>0.9	>0.9	>0.9	
值	0.846	0.657	0.859	0.606	
其它指标	PNFI	SRMR	RMSEA 90% CI		
判断标准	>0.9	<0.1	-		
值	0.765	0.042	0.096~0.113		
Default Model: $\chi^2(406)=11247.407, p = 1.000$					

表 25 样本 C 模型拟合指标

常用指标	χ^2	df	p	卡方自由度比 χ^2/df	GFI
判断标准	-	-	>0.05	<3	>0.9
值	407.675	87	0	4.686	0.855
常用指标	RMSEA	RMR	CFI	NFI	NNFI
判断标准	<0.10	<0.05	>0.9	>0.9	>0.9
值	0.102	0.031	0.927	0.91	0.912
其它指标	TLI	AGFI	IFI	PGFI	
判断标准	>0.9	>0.9	>0.9	>0.9	
值	0.912	0.8	0.928	0.62	
其它指标	PNFI	SRMR	RMSEA 90% CI		
判断标准	>0.9	<0.1	-		
值	0.754	0.043	0.092~0.112		
Default Model: $\chi^2(105)=4514.469, p = 1.000$					

响, 占 33.1%; 人理维主要受到市场信用体系、保护个人隐私、数据专业人才和交易失信行为的影响, 占 32.2%。

根据图 1, 我们可以得出以下结论:

结论 1 数据要素的供给是关键。数据要素市场建设物理维影响因素中, 稳定供求关系和企业数据入市合计占 50.3%, 这都属于数据要素的供给问题。市场机制有效运行的前提是具有完整而对称的信息。而在现实市场中, 信息总是不完整、不对称的。一般来讲, 商品市场的信息不对称表现为卖方信息占优, 卖方掌握比买方关于产品质量、性能、使用寿命、生产成本等信息, 以“柠檬市场”最为典型, 如果任由市场机制作用, 将是市场萎缩甚至消失, 这就是需求侧引致的市场失灵; 要素市场则正好相反, 买方占据信息优势, 买方掌握比卖方更多关于要素用途、未来收益、风险程度以及会否二次开发利用等信息, 以数据要素市场最为典

表 26 样本 C 因子载荷系数表

Factor (潜变量)	测量项 (显变量)	Factor 非标准载荷 系数 (Coef.)	标准误 (Std. Error)	z (CR 值)	p	标准载荷系数 (Std. Estimate)
物理维	B1	1	-	-	-	0.679
物理维	B2	1.094	0.082	13.29	0	0.759
物理维	B5	0.987	0.08	12.361	0	0.701
物理维	B12	1.334	0.096	13.937	0	0.8
物理维	B13	1.24	0.089	13.918	0	0.799
物理维	B14	1.246	0.085	14.597	0	0.843
物理维	B15	1.225	0.087	14.019	0	0.805
事理维	B24	1	-	-	-	0.844
事理维	B26	0.98	0.048	20.356	0	0.853
事理维	B27	0.996	0.049	20.4	0	0.854
事理维	B33	0.9	0.054	16.646	0	0.751
人理维	B36	1	-	-	-	0.786
人理维	B40	1.054	0.061	17.332	0	0.826
人理维	B41	1.11	0.061	18.293	0	0.861
人理维	B42	1.036	0.058	17.709	0	0.84
数据要素市场	物理维	1	-	-	-	1
数据要素市场	事理维	1.292	0.093	13.824	0	0.949
数据要素市场	人理维	1.116	0.086	13.037	0	0.949

表 27 样本 D 模型拟合指标

常用指标	χ^2	df	p	卡方自由度比 χ^2/df	GFI
判断标准	-	-	>0.05	<3	>0.9
值	179.362	41	0	4.375	0.911
常用指标	RMSEA	RMR	CFI	NFI	NNFI
判断标准	<0.10	<0.05	>0.9	>0.9	>0.9
值	0.097	0.026	0.958	0.946	0.944
其它指标	TLI	AGFI	IFI	PGFI	
判断标准	>0.9	>0.9	>0.9	>0.9	
值	0.944	0.856	0.958	0.566	
其它指标	PNFI	SRMR	RMSEA 90% CI		
判断标准	>0.9	<0.1	-		
值	0.706	0.035	0.083~0.112		
Default Model: $\chi^2(55)=3351.940, p = 1.000$					

型, 如果任由市场机制作用, 将是市场萎缩甚至消失, 这就是供给侧引致的市场失灵. 正是数据要素市场信息逆向不对称, 现实中出现了“有数无市, 有市无数”的现象, 因此, 如何解决市供给不足的问题, 将是数据要素市场建设的关键.

结论 2 数字技术的突破是基础. 数据分析水平和基础设施完善程度, 在物理维影响因

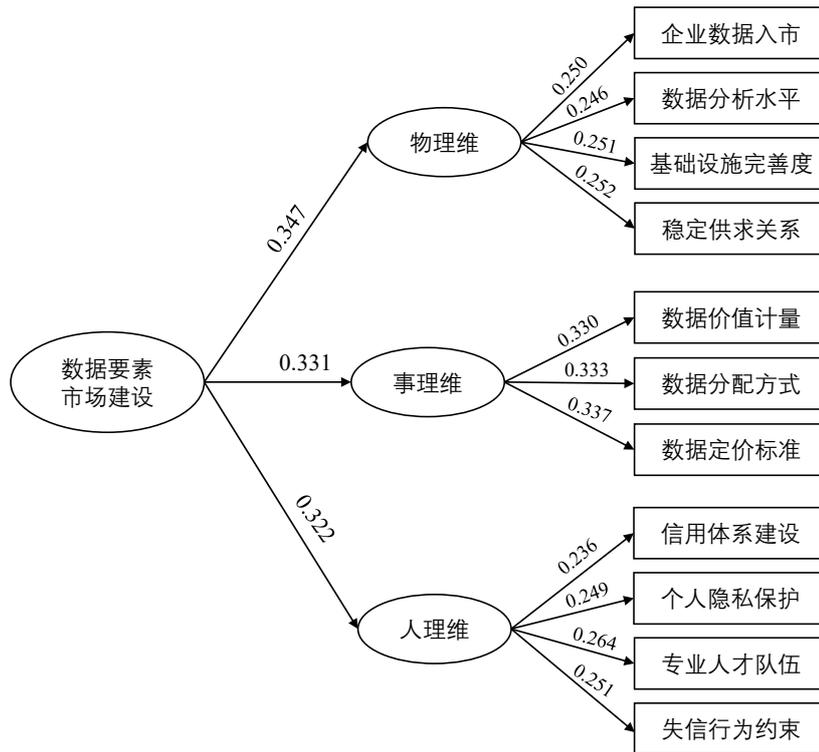


图 1 数据要素市场影响因素模型

素中合计占了 49.6%, 这可以归结于数字技术因素. 数据是一种客观存在, 但其物质形态并不像有形要素那样看得见、摸得着, 通过技术手段解决由此带来的问题, 将是解决数据要素市场构建的关键. 即使有了完美的理论逻辑和法律规定, 数据的权属界定、定价方法、价值测度、隐私保护等, 都要有赖于数据技术的支撑. 现实中, 数据交易所尽管已有相应实践, 也认识到了确定难、定价难等问题, 但并未从根本上解决数据交易效率、交易安全保护等问题, 从根本上讲, 就是数据交易技术没有实现完全突破. 从这个意义上讲, 作为一种特殊的生产要素, 数字技术对于数据要素市场的构建具有决定性意义, 只有数字技术的发展, 才能实现数据交易权属清晰、追溯便捷、监管有效, 推动形成涵盖多行业、多领域、多平台、多人群、多机构的数据交易体系.

结论 3 数据价值的测度是前提. 交易的前提是要确定要素价值. 在数据要素市场建设事理维影响因素中, 数据定价标准占 33.7%, 数据价值计量占到了 33%, 这两个因素可以归结为数据价值的测度. 根据 Miller and Mork (2013) 提出的“数据价值链”的概念, 数据价值创造过程是一个从原始数据到数据产品的整体耦合过程. Curry (2016) 将数据价值链分解为数据采集、数据分析、数据管理、数据存储和数据使用等五个环节. Corrado (2019) 认为, 原始数据开始只具有最低的价值, 通过聚合、分析、应用, 数据的价值持续增加, 通过对价值链的分解, 可以对数据价值链的各个环节进行估值. Curry (2016), Corrado (2019) 的研究包含了数据要素价值链两个阶段的形成过程, 第一阶段是数据要素自身价值的形成; 第二阶段是数

据要素参与生产的价值形成。而要完整地体现数据要素的价值,就必须科学测算两个阶段的价值,这也就成为数据要素市场建设的重要环节。

结论 4 数据收入的分配是动力。生产要素按贡献参与分配,是社会主义市场经济的内在逻辑。任何市场行为,基本的驱动力量是物质利益及其分配,前提是产权归属的界定。在数据要素市场事理维影响因素中,数据参与分配占 33.3%。根据维基百科的定义,数据(Data)是指通过观测得到的数字性的特征或信息。根据数据产生及权利主体的性质,可以把数据分为个人数据、企业数据、公共数据。无论哪一种数据,都存在“两个主体”,一个是数据事实主体,一个是数据持有主体。相应地,作为一个客观事物,数据存在“财产权”;同时,数据背后存在“人格权”,或者相应的名称权、名誉权、荣誉权等。因此,数据要素包含“两种权利”,既有与其他生产要素一样的财产权,也有其他生产要素不具有的人格权。因此,数据要素要在“两个主体”之间分配“两种权利”,比起其他要素参与分配更为复杂。如何解决数据作为要素参与分配的问题,成为数据要素市场发展的内生动力。

结论 5 数据安全的保护是底线。在数据要素市场建设人理维影响因素中,市场信用体系占 23.6%,保护个人隐私占 24.9%,交易失信行为占 25.1%,这三个因素可以归结为数据安全问题。作为财产权的数据,与其他要素一样,具有财产安全的问题。同时,数据还有人格权,需要综合考虑国家制度、法律、文化、习俗等多种因素。数据安全还关系国家安全,也关系到社会第一个人的敏感神经。从物质形态看,由于数据具有无形性、不稀缺、不消耗、可复制、低成本、难测度等特点,因此,市场“策略性行为”(Strategy-proof)将会广泛存在。从经济性质看,根据是否具有权利上的排他性和消费中的竞争性两个特点,数据具有公共物品、准公共物品、私人物品等多重经济属性,产权性质的模糊性决定了数据安全保护的难题大大高于其他要素。能否解决好数据安全问题,将是数据要素市场建设成败的关键。

结论 6 专业人才的培养是重点。在数据要素市场建设人理维影响因素中,数据专业人才的培养占 26.4%。与商品市场相比,要素市场最大的特点之一,就是需要一批专业的市场经营管理人才。像证券交易所、产权交易所等,都形成了一批专业的会计师、评估师、经纪人等专业人才队伍。数据要素市场作为现有要素市场中最复杂、最专业的市场,之所以出现“有数无市、有市无数”的现象,这与数据专业人才的缺乏密不可分。专业人才队伍的建设,将随着数据要素市场的发展而越来越重要。

6 结论及政策建议

本文基于 356 份问卷调查所得的数据,从物理-事理-人理的角度探究了影响数据市场要素的主要因素,具体步骤如下:首先,通过理论分析和文献梳理,一方面,获得数据要素市场相关的法律制度、政策文件、数据资料、发展现状等,在宏观上对数据要素市场进行全面了解;另一方面,根据学术期刊和实践报告的研究梳理,提取专家学者和实践工作者关注的要素市场的影响因素;其次,基于 WSR 方法论,将影响数据要素市场建设的 42 个因子分为物理维、事理维和人理维;再次,基于三个维度分别进行主成分分析,并分别从中提取影响三个维度的重要因素;最后,基于所提取的重要因素,通过验证性因子分析,构建数据要素市场影响因素模型。

实证研究表明,数据要素的供给是关键,数据要素市场建设物理维影响因素中,稳定供求

关系和企业数据入市合计占 50.3%; 数字技术的突破是基础, 数据分析水平和基础设施完善程度, 在物理维影响因素中合计占了 49.6%; 数据价值的测度是前提, 在数据要素市场建设事理维影响因素中, 数据定价标准占 33.7%, 数据价值计量占到了 33%; 数据收入的分配是动力, 在数据要素市场事理维影响因素中, 数据参与分配占 33.3%; 数据安全的保护是底线, 在数据要素市场建设人理维影响因素中, 市场信用体系占 23.6%, 保护个人隐私占 24.9%, 交易失信行为占 25.1%; 专业人才的培养是重点, 在数据要素市场建设人理维影响因素中, 数据专业人才的培养占 26.4%。

基于数据要素市场影响因素模型及上述结论, 我们提出了以下政策建议: 1) 集中力量加快关键数字技术的突破。一方面, 数字技术的突破, 有利于解决数字经济发展中的问题, 特别是确保数据的安全, 促进数据要素市场的稳定繁荣; 另一方面, 数字技术的突破, 将带动数字经济以至国民经济的整体跃升。2) 加强数据要素市场的制度设计。由国家数据局牵头, 基于市场设计理论, 综合总结各地的实践, 设立国家数据交易所, 组建国家数据中心, 围绕公共数据的开发利用催生和培育一批数据市场主体, 增加数据要素市场的供给, 带动数据要素市场的发展。3) 制定数据产权法。市场解决不了数据产权问题, 地方政府没有相应立法权, 只能由国家立法解决。现在, 数据立法在中国具备了深厚的实践基础, 可以采用民事特别法的立法方式, 实行数据产权单行法律方式单项立法, 为数据参与分配、数据价值的测度创造条件。4) 加快数据要素市场专业队伍队伍建设。实行数据交易从业资格考试制度, 培养一批数据交易经纪人、数据价值评估师、数据安全工程师等, 通过专业人才为数据要素市场发展注入稳定性、专业性、安全性。

参 考 文 献

- 冯珏, (2023). 数据要素市场研究的热点主题与演变趋势 [J]. 大数据时代, 71(2): 24-37.
- Feng J, (2023). Hot Topics in Research on Data Elements Market and Evolution Trends[J]. Big Data Time, 71(2): 24-37.
- 顾基发, 高飞, (1998). 从管理科学角度谈物理-事理-人理系统方法论 [J]. 系统工程理论与实践, 18(8): 2-6.
- Gu J F, Gao F, (1998). To See Wuli-Shili-Renli Systems Approach from the View of Management Science[J]. Systems Engineering — Theory & Practice, 18(8): 2-6.
- 顾基发, 唐锡晋, (2000). 从古代系统思想到现代东方系统方法论 [J]. 系统工程理论与实践, 20(1): 89-92.
- Gu J F, Tang X J, (2000). From Ancient System Thoughts to Modern Oriental Systems Methodology[J]. Systems Engineering — Theory & Practice, 20(1): 89-92.
- 洪永淼, 汪寿阳, (2021). 大数据、机器学习与统计学: 挑战与机遇 [J]. 计量经济学报, 1(1): 17-35.
- Hong Y M, Wang S Y, (2021). Big Data, Machine Learning and Statistics: Challenges and Opportunities[J]. China Journal of Econometrics, 1(1): 17-35.
- 洪永淼, 汪寿阳, (2023). 人工智能新近发展及其对经济学研究范式的影响 [J]. 中国科学院院刊, 38(3): 353-357.
- Hong Y M, Wang S Y, (2023). Impacts of Cutting-edge Artificial Intelligence on Economic Research Paradigm[J]. Bulletin of Chinese Academy of Sciences, 38(3): 353-357.
- 黄朝椿, (2022). 论基于供给侧的数据要素市场建设 [J]. 中国科学院院刊, 37(10): 1402-1409.
- Huang C C, (2022). On Building Data Market Based on Supply Side[J]. Bulletin of Chinese Academy

- of Sciences, 37(10): 1402–1409.
- 寇晓东, 顾基发, (2021). 物理-事理-人理系统方法论 25 周年回顾 —— 溯源、释义、比较与前瞻 [J]. 管理评论, 33(5): 3–14.
- Kou X D, Gu J F, (2021). A Twenty-five-year Review of WSR Methodology: Origin, Connotation, Comparison and Outlook[J]. Management Review, 33(5): 3–14.
- 刘金钊, 汪寿阳, (2022). 数据要素市场化配置的困境与对策探究 [J]. 中国科学院院刊, 37(10): 1435–1444.
- Liu J Z, Wang S Y, (2022). Dilemmas and Suggestions on Market-based Data Allocation[J]. Bulletin of Chinese Academy of Sciences, 37(10): 1435–1444.
- 苗东升, (1998). 系统科学辩证法 [M]. 济南: 山东教育出版社.
- Miao D S, (1998). Dialectics of Systems Science[M]. Ji'nan: Shandong Education Press.
- 邱东, (2021). 数据科学在社会经济领域应用的重心 —— 兼评《十字路口的统计学, 谁来应对挑战》[J]. 计量经济学报, 1(2): 250–265.
- Qiu D, (2021). The Focus of Data Science Applications in the Socio-economy: On Statistics at a Crossroads: Who is for the Challenge?[J]. China Journal of Econometrics, 1(2): 250–265.
- 萧政, (2021). 大数据时代关于预测的几点思考 [J]. 计量经济学报, 1(1): 1–16.
- Hsiao C, (2021). Some Thoughts on Prediction in the Presence of Big Data[J]. China Journal of Econometrics, 1(1): 1–16.
- 张彩江, 孙东川, (2001). WSR 方法论的一些概念和认识 [J]. 系统工程, (6): 1–8.
- Zhang C J, Sun D C, (2001). Some Concepts and Understandings about WSR[J]. Systems Engineering, (6): 1–8.
- 朱晓武, 张雪祺, (2021). 基于“物理-事理-人理方法论”的分布式自治组织运行机制与监管策略研究 [J]. 管理评论, 33(5): 114–122.
- Zhu X W, Zhang X Q, (2021). Distributed Autonomous Organizations Modeling and Governance on the Perspective Wu-li Shi-li Ren-li Approach[J]. Management Review, 33(5): 114–122.
- 朱志昌, (2000). 物理事理人理方法论国际交流的启示 [C]// 系统工程, 系统科学与复杂性研究: 第十一届系统工程学术年会论文集: 135.
- Zhu Z C, (2000). The WSR Approach in the International Systems Community[C]// Systems Engineering, Systems Science and Complexity Research — Proceeding of 11th Annual Conference of Systems Engineering Society of China: 135.
- Agion P, Howit W P, (1998). Endogenous Growth Theory[M]. Cambridge: MIT Press: 6.
- Corrado C, (2019). Data as an Asset: Expanding the Intangible Framework[C]// Presentation at the EMAEE 2019 Conference on the Economics.
- Curry E, (2016). The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches[M]// Cavanillas J M, Curry E, Wahlster W. New Horizons for a Data-driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe. Springer Open: 29–37.
- Miller H G, Mork P, (2013). From Data to Decisions: A Value Chain for Big Data[J]. IT Professional, 15(1): 57–59.
- Liang F, Yu W, An D, Yang Q, Fu X W, et al. (2018). A Survey on Big Data Market: Pricing, Trading and Protection[J]. IEEE Access, (6): 15132–15154.
- Rubinfeld D L, Gal M S, (2017). Access Barriers to Big Data[J]. Arizona Law Review, 59(2): 339–381.
- Sharon T, Zandbergen D, (2009). From Data Fetishism to Quantifying Selves: Self-tracking Practices and the Other Values of Data[J]. New Media & Society, 11(19): 1695–1709.

附录

数据要素市场影响因素调查问卷

(一) 调查问卷函

尊敬的受访者:

首先,感谢您对本次问卷调查的支持,对您抽出宝贵的时间填写问卷表示敬意!

“加快培育数据要素市场”,这是以习近平同志为核心的党中央关于发展数字经济的重要战略部署。近年来,随着数字经济快速发展,推进数据资源化、资产化、资本化、要素化、市场化成为学界和业界的共识。然而,现实中数据要素市场发展并不顺利,呈现出有数无市、有市无数的现象。为此,本研究围绕数据要素市场的影响因素,以问卷形式开展调研,期待您在这一研究课题中贡献智慧。

问卷分为两个部分,第一部分是受访人基本信息,我们已充分考虑了保护个人隐私问题,请放心填写;第二部分是数据要素市场影响因素重要程度量表,请您根据自己的理解或工作实践给出答案。

问卷只用于本课题研究,不会用于任何其他目的,原始数据也不会提供给其他任何用途,我们承诺为您参与本次问卷的任何个人信息安全负法律责任。

最后,祝您工作顺利、身体健康!感谢您对本研究贡献智慧和力量。

中国科学院大学经济与管理学院

在读博士研究生:黄朝椿

2023年1月7日

(二) 受访者基本信息

(请在选项上直接打√)

1、您的性别?

A. 男 B. 女

2、您的年龄?

A. 30岁及以下 B. 31~40岁 C. 41~50岁 D. 51~60岁 E. 60岁以上

3、您的职业?

A. 数据主管部门 B. 数据企业 C. 专家 D. 网民 E. 交易机构

4、您在本单位的职务?

A. 高层领导 B. 中层管理 C. 普通员工 D. 其他

5、您的学历?

A. 博士 B. 硕士 C. 大学本科 D. 专科 E. 专科以下

6、您认为数据交易市场的前景如何?

A. 很好 B. 好 C. 一般 D. 说不准

7、您是否会在确保数据安全的前提下出售个人数据?

A. 会 B. 不会 C. 不知道怎么卖 D. 没有想过

(三) 数据要素市场影响因素量表

请您根据自己的工作实践、直观理解和思考,判断表中因素对于数据要素市场的影响程度。您只需要在对应的表格中打“√”,并请尽量都作出判断。

表格 A1 中,“1”表示影响程度低,“2”表示影响程度较低,“3”表示影响程度一般,“4”表示影响程度较高,“5”表示影响程度高。

表 A1

编号	影响因素	重要性程度				
		低	较低	一般	较高	高
		1	2	3	4	5
b1	良好运营的交易机构					
b2	数据市场体系的完善程度					
b3	提升数据交易技术水平					
b4	降低市场交易成本					
b5	提升隐私保护技术水平					
b6	公共数据开放共享程度					
b7	数据产品形态的标准化程度					
b8	形成有效的数据定价机制					
b9	提升数据采集技术					
b10	平台公司实质掌控数据					
b11	推动个人数据入市交易					
b12	推动企业数据入市交易					
b13	提升数据分析技术水平					
b14	数据市场基础设施的完善程度					
b15	市场稳定的供求关系					
b16	有效限制场外交易					
b17	形成完善的数商体系					
b18	数据垄断程度					
b19	加强数据执法					
b20	建设更多数据交易场所					
b21	数据科技投入力度					
b22	实现数据资产入表					
b23	建立数据专项基金					
b24	解决数据价值计量问题					
b25	清晰界定数据产权					
b26	数据按要素贡献参与分配					
b27	实行统一的数据定价标准					
b28	强制披露市场交易信息					
b29	加强政府主管部门的职责					

表 A1 (续)

编号	影响因素	重要性程度				
		低	较低	一般	较高	高
		1	2	3	4	5
b30	支持数据作为资产融资					
b31	统一数据产品的质量标准					
b32	形成完整的数据产业生态					
b33	市场主体的所有制性质					
b34	有效限制算法歧视行为					
b35	社会对于市场重要性的认识					
b36	加强市场信用体系建设					
b37	数据基础理论研究水平的提升					
b38	经济总体发展水平的高低					
b39	数据知识的宣传普及程度					
b40	社会对个人隐私的实际保护程度					
b41	数据专业队伍的状况					
b42	对于市场交易失信行为的制度约束					