

产业级数据治理白皮书

出品方：广州信安数据有限公司、DAMA 中国

编者：高伟、汪广盛、陈韩霏、张艳红

2023 年 11 月

版权声明

本白皮书版权属于广州信安数据有限公司，并受法律保护。转载、摘编或利用其它方式使用本报告文字或者观点的，应注明“来源：广州信安数据有限公司”。违反上述声明者，本司将追究其相关法律责任。

前言

随着数据二十条的颁布与全国数据要素市场的建设与发展,数据的资产属性、要素属性不断被挖掘、激发,围绕数据进行治理的范围,也从传统的企业领域,扩大到了整个数据产业生态内。为建立安全、高质量的产业数据流通环境,解决数据跨域融合分析过程中的分散、不一致、权属确认、隐私保护等问题,产业级数据治理应运而生。

产业级数据治理以数据要素在产业层面的顺畅流通和价值释放为最终导向,立足于产业级视角,为数据要素市场构建、产业数字化转型、数字经济发展提供“底座式”的支撑,起到破除阻碍数据要素供给、流通、使用的障碍,推动大规模、跨领域数据的价值再提升、再创造的作用。

纵观各国的大数据发展态势和实践,由政府、企业、组织、个人多方协同,注重各主体数据主权维护、数据要素市场发展秩序塑造,依托于新兴智能化技术的产业级数据治理格局正在形成当中。在此治理格局中,有别于传统的数据治理模式,产业级数据治理关注的重点已经不再仅是数据本身,而是对“数据”、“生态环境”、“参与主体”、“市场活动”的多维规范和管控。

本白皮书在各国、各领域产业级数据治理成熟实践和新趋势萌芽案例的基础上,结合信安数据多年数据治理实践经验以及 DAMA 中国成熟方法论体系,深度解读产业级数据治理产生的时代背景与趋势、分析传统数据治理方法面临的局限和挑战、明确产业级数据治理的内涵并进一步拆解其框架内容。期望本白皮书的论证成果可以为社会各界认识新时代下的产业级协同数据治理模式、参与数据要素市场建设与实践提供有价值的参考和指导。

目录

版权声明	2
前言	3
目录	4
一、 时代背景与趋势	5
(一) 时代背景	5
(二) 发展趋势	6
二、 问题与挑战	8
(一) 当前发展现状	8
(二) 存在的局限性和挑战	9
(三) 挑战突破与解决思路	11
三、 产业级数据治理的内涵	11
(一) 产业级数据治理的定义	12
(二) 产业级数据治理的价值	13
(三) 产业级数据治理的特征	14
(四) 产业级数据治理的框架	16
四、 产业级数据治理框架详述	19
(一) 政策法规	19
(二) 组织形式	21
(三) 连通机制	24
(四) 互信机制	26
(五) 执行机制	29
(六) 监督机制	31
(七) 技术支撑	34
五、 总结与展望	37

一、时代背景与趋势

（一）时代背景

在数字技术带来越来越多红利的时代发展趋势下，欧盟、美国、英国、澳大利亚、俄罗斯等国家和组织不断细化和更新数字经济发展战略，以实现更加优化的数字经济¹发展环境和深化数据流通合作机制。数字经济不仅成为推动经济发展的重要引擎，也成为各国新一轮科技革命²的重要战略基础。

全球范围内，数据互联互通和合作共赢正在成为国际趋势。2023年7月7日，全球数字经济大会“数字经济全球治理”论坛在北京召开，针对“数字经济治理愿景、治理框架、创新合规、争端解决”四个专题开展方案讨论。2023年8月19日，二十国集团数字经济部长会议在印度班加罗尔举行，聚焦数字经济、互联互通等重点领域探讨合作。可见，全球数字产业的发展正呈现出持续蓬勃的态势，国际关于数字经济发展的政策部署和合作交流日益密切。

在我国，数据被认为是发展数字经济的核心基础和重要载体，直接关系到数字经济的稳定发展和持续创新。2019年10月，党的十九届四中全会首次提出“数据是生产要素³”的科学论断；2020年3月中共中央、国务院发布《关于构建更加完善的要素市场化配置体制机制的意见》，在全球范围内率先将数据正式定义为第五大生产要素；2022年12月，党中央、国务院印发《关于构建数据基础制度更好发挥数据要素作用的意见》（简称“数据二十条”），为最大化释放数据要素价值、推动数据要素市场化配置提出了最新指引。此外，各地政府也纷纷出台了相关政策措施，鼓励和支持数据要素市场的发展，为数据要素的管理和应用提供了有利环境。

¹ 数字经济是以数字化的知识和信息作为关键生产要素，以数字技术为核心驱动力量，以现代信息网络为重要载体，通过数字技术与实体经济深度融合，不断提高经济社会的数字化、网络化、智能化水平加速重构经济发展与治理模式的新型经济形态。具体包括数字产业化、产业数字化、数字化治理、数据价值化四大部分。

² 科技革命是指在科学技术起决定作用下而实现的社会生产力的根本变革，与数字经济关联紧密的是人工智能、大数据、量子信息、移动通信、物联网、云计算、区块链为代表的新一代信息技术。

³ 生产要素是指进行社会生产经营活动时所需要的各种社会资源，是维系国民经济运行及市场主体生产经营过程中所必须具备的基本因素，生产要素包括土地、劳动力、资本、技术、数据五种。

而在技术层面，大数据、云计算、人工智能、区块链等技术正在迅猛发展，各组织和企业主体通过更加先进智能的综合性技术体系，实现数据的安全、灵活、高效存储与处理，开展多源异构、大规模数据的快速分析挖掘，甚至让计算机在海量训练数据集的基础上，模拟人类智能并执行复杂的业务任务，基于此，各行各业数据的管理、流通和应用效率得到了大幅提升。可以说，技术的进步为数据要素市场的快速发展提供了有力支持。

综上所述，在环境、政策和技术的支持下，数据作为新型生产要素，与其他生产要素发生了广泛而深入的融合应用，重新定义和调整了社会的生产力和生产关系，大规模激活了制造、零售、医疗、能源等传统行业的沉默价值，推动了不同行业数据赋能实践百花齐放，加快了全行业的管理变革与数字化转型⁴升级，正以不可逆转的趋势改变人类社会。

（二）发展趋势

整个数据产业作为推动经济社会发展的重要力量，发展迅猛，渐趋成熟。随着产业规模的指数级增长，产业业态进入了一个全新的发展阶段。围绕海量数据要素的综合管理、协同治理和价值挖掘等领域，业务模式和形态日渐多样化，在技术创新、价值共赢、权益维护等方向形成了三大行业发展趋势。

在技术创新方向，数据的互联网络、数智化劳动模式正在成形。一是数联网的广泛链接。如同工业领域的物联网一样，依托庞大而广泛的数据要素基础设施，一张链接国家、行业、企业、个人的数据互联网正在编织成形，即将形成一个庞大的网络体系，实现物理世界与数字世界互相融合，推动世界的进一步智能化演进。**二是人、数、机协作日益密切。**人类、数据和机器之间的合作与协同越来越密切，在数据作为底座和基础要素的前提下，立足于劳动规则的洞察模仿、知识的积累与总结，智能化机器、生成式 AI⁵等创新技术从体力劳动与脑力劳动两个层面，协助人类实现更高效、轻松地完成工作任务。

⁴ 数字化转型指利用数字化技术来推动组织转变业务模式，组织架构，组织文化等的变革措施，以追寻新的收入来源、新的产品和服务、新的商业模式。

⁵ 生成式 AI 是一种人工智能技术,利用机器学习模型和深度学习技术,通过研究历史数据的模式来生成新内容,内容形式包括文本、图像、音频或视频等。

在价值共赢方向，广泛合作、利益共享的数据生态⁶正在成形。一是产业级数据生态的形成。数据产业正逐渐从独立的企业活动向产业链的多个环节延伸，形成了产业级数据生态。各种数据相关的企业、组织和个人在数据共享、合作、交换和创新方面展开更广泛的合作。数据生态的形成促进了数据资源的优化配置和协同创新，推动了数据产业的协同发展。二是数据资本化的探索尝试。为促进数据要素配置效率的大规模提升，各数据市场参与主体创新通过数据保险、数据托管、数据证券化、数据股权化等途径，将数据资源和数据资产转化为数据资本，实现数据交易流通过程中的风险分摊，和数据回报与收益在多级市场上的更大范围流动和分配。

在权益维护方向，从主权到伦理的权益保障环境正在成形。一是数据主权成为重要关注点。随着数据的重要性和价值的不断提升，数据主权越来越被重视，数据主权意味着对数据的掌控、管理、限制和保护，国家重视数据主权以维护国家安全、公民权益，企业组织重视数据主权以维持竞争平衡，避免垄断，个人重视数据主权以保护个人隐私、维护个人权益。二是数据伦理⁷的重要性凸显。数据产业的快速发展也引发了对国家、个人对数据伦理的关注。数据伦理强调在数据收集、处理和使用过程中考虑个人隐私、公正性、透明度⁸和道德原则。建立良好的数据伦理框架和准则，标准化数据使用的行为，维护数据的合法性和公平性，成为数据产业发展的必然要求。

数据要素的市场化带动了数字经济的增长，带领我们走向数据被广泛利用的新时代，为人类带来了前所未有的机遇和福利，通过充分利用大数据的潜力，人类社会实现了向更智能、更进步方向的发展。相对应的，我们也需要积极应对数据产业发展、数据技术与模式创新之下需要面临的数据治理挑战，确保数据被安全、得当、合理使用，以实现大数据时代的可持续发展和共享价值。

⁶ 数据生态是指由数据及其相关技术、组织、政策、法律等环境构成的生态系统。数据生态的目标是构建多元、开放、可持续且可信的数据环境，以满足生态内各主体利益的平衡和数据价值的最大化。

⁷ 数据伦理是指以符合伦理准则的方式获取、存储、管理、使用和销毁数据，伦理准则通常包括尊重他人、行善、公正、责任、可靠性、透明度等。

⁸ 透明度是指组织应向个人和其他主体提供有关其数据的数据管理、数据处理的制度和实践信息。

二、问题与挑战

（一）当前发展现状

良好的、高质量的数据治理是数据产业创新升级的关键基础，而要实现高质量数据治理离不开成熟的配套资源保障、清晰的治理规则和方法、丰富的治理过程活动和举措等等。当前，支撑数据产业发展的数据治理仍然存在一定的局限性，究其原因主要在于，纵观所有参与数据治理的政府、行业、企业、个人等相关主体中，在资源保障、方法论与实施举措等层面都比较成熟的主要是企业。

在数据治理资源保障方面，一是国内大部分企业在基础业务系统运维、数据分析挖掘、数据共享开发运营等过程中，逐步形成和建立了相应的数据治理组织，包括设立数据治理委员会⁹、管理团队、执行团队等，以确保数据治理的责任和权力划分清晰。**二是**搭建数据资产管理平台、数据治理平台、数据安全平台等技术平台能力，实现数据全生命周期的标准化、质量和安全保障。其中，部分企业数据治理能力实践在行业内树立了标杆，例如，工商银行，南方电网、国家电网、华为、联通、移动等企业的数据治理发展水平在行业内处于前沿水平。

在数据治理规则方法方面，相当一部分的企业在《数据管理能力成熟度评估模型》（GB/T 36073-2018）（简称“DCMM”）¹⁰的基础上，结合企业自身需要，形成符合自身特色的数据治理规则和流程，明确数据治理的目标、原则、责任和流程，并将其纳入企业的日常运营和管理中，通过清晰、精细化的治理规则，指引企业内部数据治理相关工作的有序开展。其中，部分企业对其特色的数据治理规则体系和方法以专业书籍、白皮书等形式进行总结推广，进一步带动了其他企业数据治理规则的制定和探索。例如，华为发布的《华为数据之道》、南方电网发布的《南方电网数据资产管理体系白皮书》等。

在数据治理过程和实施举措方面，企业围绕常规的数据盘点与梳理、主数据识别、数据标准制定、数据质量提升、数据安全防护等实施路径和举措，对数据进行分类和标记，建立数据目录和数据地图；识别和确定企业中最重要和核心的

⁹ 数据治理委员会是企业为管理数据治理规划、协调和升级处理企业数据问题而设置的管理组织。

¹⁰ 《数据管理能力成熟度评估模型》（GB/T 36073-2018）英文简称：（Data management Capability Maturity Model），是我国首个数据管理领域正式发布的国家标准。

数据实体，避免在不同系统和部门中存在重复和不一致的数据；统一数据的命名规则、数据格式、数据单位等规范，明确数据内容的标准约束要求；对数据进行清洗、验证和纠错，以确保数据的准确性和可靠性；在数据流动的过程中采取了一系列的安全措施，以保护数据的机密性和完整性。

企业层面在数据治理资源保障、规则和方法制定以及实施举措等方面取得的成果实践，为推动数据的流通应用、统一数据语义、提升数据基础质量、保障数据安全等数据治理需要起到了强力、稳健的支撑作用。

（二）存在的局限性和挑战

然而即便是数据治理能力最成熟的企业，其采用的一系列举措，仍然更多沿用传统数据治理的方法论，解决的是企业内部的数据质量、数据安全和数据供给的问题。而数据要素时代大背景下，数据治理的内涵已在逐步超出传统范畴，需要在更大的领域范围内支撑构建数据生态、优化数据环境、加速数据融合、赋能产业发展。在这个过程中会遇到的**数字鸿沟、数据确权¹¹、数据隐私和安全保护、数据处理行为监管、跨主体数据协同引导、互信建立和透明度保障、跨主体数据质量维护**等问题，已经不是单一企业能够解决的问题：

一是数字鸿沟问题。数据垄断¹²或不对称、数字素养¹³差异化等情况下，极易造成不同群体间的数据福利不均衡，产生数字鸿沟，进一步加剧社会差距。单一企业仅能从一定范围内实现接触人群的数字素养提升，从自身做起规避数据垄断和不对称问题，但无法实现社会层面数字鸿沟问题的杜绝。

二是数据权属确定问题。数据权属与权益的确定必须依赖法律法规的确认，单一企业仅能从现有的中央发布的“数据二十条”等政策中明确的数据产权、收益分配准则，进行进一步的权属管理规则研究细化，但企业研究遵循的权属确定和授权规则仅能在小范围内形成共识和机制确认，无法在更大范围公共场合应用。

¹¹ 数据确权是指确定数据的权利主体，以及确定权利主体所享有的权利内容的过程，即确定谁对数据享有权利，享有什么样的权利。

¹² 数据垄断是指少数数据持有者或者组织拥有大量独特、有价值的信息，并且限制其他人获取该数据的情况。

¹³ 数字素养是指数字社会公民学习工作生活应具备的数字获取、制作、使用、评价、交互、分享、创新、伦理道德等一系列素质的集合。

三是数据隐私和安全合规问题。随着数据的广泛互联和数据流通交易活动的大量开展,国家安全的维护、企业机密的保护、个人隐私的保护越来越难以维系,数据被滥用、被违规使用、被泄露等风险难以杜绝。数据隐私保护和安全合规,尤其合规方面,需要立法、律所等法律领域单位的参与,不再是单一企业内部的问题,而是要营造一个安全合规有保障的良好环境和机制。

四是数据处理行为监管问题。虽然目前国家出台了《数据安全法》、《个人信息保护法》等多部法律,对数据处理行为进行了约束规范,但缺乏监管机制,难以保障大部分数据处理主体的法律遵循程度。由于数据的爆炸增长和跨境流动,如果不通过国家主导推动相关法律政策的制定细化和相关监管机制的设立,单一的企业或者行业难以有效监管和管理数据使用、存储和传输活动的规范性、正当性。

五是跨主体数据协同引导问题。跨域数据的统筹治理必然需要更多企业从某种程度上达成共识,并通过配套的管理和技术手段,实现在跨主体高效协同下的治理合作,但由于国家、社会层面的制度问题,企业往往无力牵头和引导跨主体协同的数据治理工作,或在实际工作中存在多方阻力,也无法发布有效力的约束制度和标准。

六是互信建立和透明度保障问题。在数据的使用和共享缺乏透明度,算法规则、数据流通链路等未直观披露的情况下,企业间、个人自身无法掌握自己的数据将被用于何种目的,无法信任其他主体对自身数据的收集和使用行为。在此情况下的信任机制的建立、数据透明度的保障,依靠单一的企业的自律和努力是无法做到的。

七是跨主体数据质量维护问题。涉及多主体之间数据共享开放和融合的情况下,数据在广泛流通的过程中,其准确性和真实性不再是一家之言,需要满足更广泛数据需求方的质量要求,各参与者、消费者对其他主体数据的准确性、完整性和可靠性难以保证,协同制定大量主体需要共同遵守的数据标准也不是单一的企业能具备能量推动的事情。

由此可见,当前数据产业的新阶段下,传统的依赖企业自主开展数据治理,部分行业以监管协作推动数据治理的模式已经远远不能适应新阶段的发展要求,需要整个产业乃至国家层面高度协同,对部分领域现状勇于变革,为数据“量身定制”全新社会治理体系,才能推动数字经济加速发展。

（三）挑战突破与解决思路

面对以上众多挑战和要求，需要突破传统的以企业为主的数据治理模式，将数据治理延展到更大的范围，推动政府、组织、企业和个人协同参与数据治理，以服务于大规模数据流通为目标，建立一个多类主体互相协作、保障数据互信可用、保护各主体自主权益的数据治理机制。

该数据治理机制应当满足以下几个要求，以解决传统单一企业数据治理方法和机制无法解决的困境问题：

一是广泛聚合专业数据治理保障资源。依托高度专业化的数据治理资源保障，以满足宏观和微观层面的数据治理需求。通过聚合产业生态内数据治理相关的专业技术能力、理论知识和实践经验，培养参与主体的数据素养，在深入的数据治理行业理解的基础上，解决数据规模化流通下跨行业数据治理的需求和挑战，创新先进的智能化技术工具成果和理论方法，形成全新的数据治理模式成果，再利用这些成果来提供高质量的数据治理实践。专业化资源保障是弥补数字鸿沟差距、塑造数据互信环境、提供高效参与渠道的重要手段。

二是形成规模化运作的数字治理规则。针对海量数据设计满足各群体利益的、需要广泛遵循的数据治理规则，推动规则在社会层面的规模化运作，形成数据治理制度和规范的广泛共识和认知，消除数据要素市场参与主体在数据权属、数据合规、数据隐私保护、数据安全等层面的顾虑，提高数据互联互通和供给运营的效率，确保过往在主体繁多、规则不明问题下而形成的各类复杂的数据相关问题可以快速得到解决。

三是引导多主体协作实施治理活动。推动大量政府单位、行业组织、企业和个人协同合作，促进多样化的数据基础技术平台互通协作。这些协作的主体覆盖数据治理政策的制定者、数据治理共识的倡议者、数据行为的监管者、数据治理活动的实践者、数据治理服务的供应商等多种角色，通过边界清晰、协作有序的组织机制，协调数据治理的实施流程，满足不同层次主体的数据权益保护需求，解决传统单一企业无法解决的跨主体数据协同引导、透明互信建立、数据质量维护等问题。

三、产业级数据治理的内涵

对于以上提到的满足广泛聚合专业数据治理保障资源、形成规模化运作的数

据治理规则、引导多主体协作实施治理活动的据治理机制，我们将其定义为“产业级数据治理”。接下来的第三章中，我们将详细阐述产业级数据治理的定义、特征和蓝图框架，描摹产业级数据治理的全貌和内在机理，给出更清晰的产业级数据治理解决方案说明。

（一）产业级数据治理的定义

1. 什么是数据治理

对于什么是数据治理，目前在行业内还没有一致且通用的定义。国际数据管理协会（DAMA）的数据管理知识体系（DMBOK）将数据治理定义为“对数据资产管理行使权力和管控和共享决策（规划、监测与执行）的系列活动”；《数据管理能力成熟度评估模型》（GB/T 36073-2018）认为数据治理主要解决顶层设计问题；国内主流认知中，数据治理是“保证数据的可信可靠可用，满足业务对数据质量和数据安全的期待的系列举措，既包括了顶层设计，也包括了具体的数据标准管理、元数据管理、数据质量管理、主数据管理等落地执行事项”。

我们认为，数据治理是一件体系化的复杂工程，既包含数据管理的顶层设计，又包含需要落地执行的具体工作事项，其核心目标就是为了保障数据的质量和安全，从而让更多人、更方便、更便捷的获取数据、使用数据。

2. 什么是产业级数据治理

产业级数据治理是在大规模、跨行业的大数据发展趋势下对数据治理理论方法的进一步发展和创新，将数据治理的范围放大到产业层面，实现产业数据生态下的“环境治理”。

我们认为，**产业级数据治理是以数据要素在产业层面的顺畅流通和价值释放为最终导向，开展建设政策法规、建立协同组织、完善监督、管理与执行机制、创新技术支撑等体系化工作，从而实现跨行业数据高质可信、安全可控、连接融合、高效流通、价值倍增的一系列活动过程。**

3. 企业级-行业级-产业级数据治理的内涵演进

产业级数据治理与企业级数据治理、行业级数据治理从内涵上是层层递进的。企业级数据治理是指在单一企业范围内，由企业最高层直接主导或推动的、

企业内部多部门协同开展的数据治理活动，解决的是打破企业内部数据壁垒¹⁴，确保数据可信可用的问题。

行业级数据治理是指在同一行业内，多个企业为了实现数据互通和互利，由行业监管单位推动开展或由行业骨干企业或单位发起，基于统一的标准规范，统筹开展数据治理活动，保证行业内数据共享流通顺畅的过程，解决的是行业内的信息共通问题。

产业级数据治理则是站在产业级视角，对同一产业领域内的跨行业的企业数据进行归集和整合，打破行业壁垒，实现数据自由流通，促进产业级生产力提升，最终推动数字经济的高速发展，从而让整个人类社会受益，解决的是数据跨行业流通、融合和价值倍增的问题。

（二）产业级数据治理的价值

产业级数据治理来源于实践需要，是数据要素理论突破后，在数据治理领域的“投影”，也是必然趋势，是对数据治理理论体系进一步的创新和拓展，是中国特色大数据理论实践体系的一部分，其价值体现在以下几个方面：

第一，产业级数据治理可以营造一个互信、开放的数据流通环境。产业级数据治理的基础是彼此认同、彼此信赖的各行业数据主体，依托一套信用体系为纽带联结在一起。在满足条件的情况下，现有的主体可以退出，更多的主体可以加入进来。数据业务层面也是开放式的，只要不违法违规，所有能带来价值收益的业务和服务形式都可以开展。

第二，产业级数据治理可以提供一个公平、透明的数据交易机制。当前数据交易机制没有通用的标准。产业级数据治理在传统数据治理的基础上，根据跨行业、跨主体的特点，建立了执行和监管两层机制，从而确保所有数据主体根据约定好的规范、标准和规则开展活动，违规的主体受到惩罚，保障一个优良的数据交易流通氛围和风气。

第三，产业级数据治理可以催生更丰富的数据治理产业生态模式。产业级数据治理是对传统数据治理范围和能力的延展和变革升级，为实现产业级数据治理，

¹⁴ 数据壁垒指限制企业内部不同部门间数据共享和流动的阻碍，包括技术、观念等阻碍因素，这些会客观导致企业整体数据的难整合、难互通，利用效率低等问题。

诸多新的组织和个体角色参与到了数据治理的工作中来。例如，律所和律师第一次参与到数据治理的核心工作内容中来，独立的第三方代加工产商也成为整个治理体系中重要的产业“拼图”，而这些都会从某种程度上改变现有的数据治理业务模式，形成形式更多样、角色更丰富、价值点更多样化的新型产业生态。

第四，产业级数据治理可以深化促进数据要素融合赋能产业产能。产业级数据治理的最终目标是保障产业内数据高质量供给，促进数据要素高效流通和价值释放。通过一系列治理手段，各行业、企业数据可以实现深度融合，从而从多个层面为各行业、业务领域发挥数据要素价值叠加倍增效应，从产业链上下游逐步拓展到更广产业面，从而支撑产业产能和社会生产力的全面提升。

第五，产业级数据治理可以强力拉动全球数字经济引擎高速运转。产业级数据治理是数据要素市场构建，乃至全社会数字化转型的基础。产业级数据治理所打造的坚实的社会级高质量数据基础，为数字经济的发展提供了优质“原料”，贯穿和打通数字经济“脉络”，保障数字经济的“引擎”高效运转，从而推动整个人类社会加速迈进全新时代。

（三） 产业级数据治理的特征

我们说，企业数据治理是企业治理的一部分，行业数据治理是行业治理的一部分，那么产业级数据治理就是产业治理乃至社会治理的一部分。相对于传统数据治理，产业级数据治理具有其明显的特征。

从本质上看，产业级数据治理的核心特征在于，其关注的重点已经不再是数据本身，而是对数据流通融合所处的“生态环境”以及“数据-主体-活动”关系的理顺和重塑。具体特征体现在以下几个方面。

第一，数据来源更加多元化。一是数据来源的行业、企业多元化，不同的行业、企业往往都有自己的标准或治理要求，产业级数据治理需要兼容并蓄，以“最大公约数”满足跨行业数据互通融合的问题；二是数据自身格式、类型多元化，半结构化数据、非结构化数据等需要更丰富的治理手段和技术，确保能够相互关联融合并被用户使用；三是数据质量参差不齐，不同行业、企业数据治理的水平和阶段不同，能提供的数据质量也存在较大差异，而产业级数据治理要接受并平衡这种差异，才能确保产业级数据应用的有效性。

第二，组织机制更加松散灵活。一是在产业级数据治理体系范围内，各主体

之间本质上是协同合作的关系，既没有上下级之分，也没有行政强制性，每一个主体都能合理提出诉求，更多需要通过共同的目标和相互之间的互利共赢来维持组织机制的健全和日常运作；**二是**组织内部的运作机制也需要相对灵活，在各主体“权责利”¹⁵平衡的前提下，除必须满足统一的公约、标准、规范以及监管要求以外，生产方式、市场化对接等工作合理放开，收益分配灵活，从而有效促进数据价值有效释放。

第三，安全合规成为核心关注点。一是各主体内部需要更加重视数据安全合规问题，只有将自身需求和安全合规结合，才能将正确的数据以正确的方式提供出来，放心参与数据要素市场；**二是**产业级数据治理的基本环境是公共的，面临的数据安全合规的风险更大、潜在威胁更多、防范难度也更高，而数据安全合规无小事，从整个产业数据生态角度看，必须将安全合规作为治理的焦点，提供一个安全可靠的环境，才能吸引更多主体参与，创造更大价值。

第四，互信机制成为核心基础。一是产业级数据治理重要目标之一就是要建立起产业数据流通的互信机制，只有各主体之间存在信任关系，并且有一个公正的第三方能够“主持正义”，数据流通才具备了基础，数据融合才具备了可能性；**二是**除了一些共识性规范外，各主体应该从意识层面注重对其他合作主体的权益保护，不做损害其他主体权益的事情，并在自身产业级数据治理能力建设的过程中，主动提供其他主体行使权利的环境、渠道和条件，在不损害其他主体利益的前提下，最大程度实现价值共创和收益共赢。

第五，先进技术成为必备管理手段。在传统数据治理中，技术是辅助管理实现目标的支撑，在没有技术的情况下，依赖人的管理方式也能达到目标。但产业级数据治理不同，由于其数据分散、主体众多等特点，隐私计算、区块链等技术已经不仅仅是一种先进技术，而是本身已经成为了一种无法通过人工来代替的管理手段，少了这些技术就无法实现对产业级数据的有效治理和管理。可见，技术与管理的深度融合是产业级数据治理的重要特征。

第六，公共价值是核心驱动力。产业级数据治理并不是为单一主体服务的，而是为了数据对于社会公共价值的最大化。**一是**产业级数据治理更关注的是数据

¹⁵ 权责利是指“权力、责任、利益”这三大组织管理实施的关键要素，权责利是相互关联的，产业级数据治理中的组织机制需要保证主体间“权责利”的平衡，以建立公平、和谐、可持续的治理秩序。

的公共共享性，即在安全合规前提下高效流通，不断融合而产生新的价值，这个价值是单一主体或行业无法提供的；**二是**产业级数据治理为整个社会数字化转型建立了基础，并通过提升产业产能的方式表现出来，体现的整个行业的生产力水平，也是整个社会数字化水平的不断提高。

（四） 产业级数据治理的框架

基于以上观点，信安数据提出产业级数据治理体系框架如下：



1. 行业级、企业级数据治理与产业级数据治理的关系

行业级、企业级数据治理与产业级数据治理之间并不是完全割裂的关系，都属于广义的数据治理范畴，只是各自的关注点、目标、实施范围以及解决的问题不同，三者之间并非相互包含的关系，而是相互促进、相互依存的关系。

一方面，行业级与企业级数据治理是产业级数据治理的基础。企业级和行业级数据治理主要解决企业、行业内部数据治理问题，目的是全面提升数据质量、保障数据安全、加速数据流通。产业级数据治理并不关注企业、行业内部数据治理，但在对数据流通环境进行治理的过程中，每一个企业自身数据治理成果的好坏，却是产业级数据治理能够有效推进的关键基础。试想一下，如果一个产业级生态中各个主体自身数据都没有治理好，质量极差，产业级数据治理的过程将极其痛苦，层出不穷的数据问题将破坏主体间的信任，阻滞数据流通融合，最终毁

掉整个数据生态。所以需要借助企业级和行业级数据治理提升企业、行业内部数据质量，保障数据安全，进而为产业级数据治理打好基础。

另一方面，产业级数据治理是行业级和企业级数据治理在数据要素时代的必然演进要求。产业级数据治理并不是凭空出现的，而是在数据成为生产要素并正在形成数据要素市场的大环境下，当企业级和行业级数据治理所覆盖的职能范围已经无法满足实际发展的要求时应运而生的。产业级数据治理确立了更高层级、更大范围的数据治理规则和规范，提供了有别于现有形式的数据治理组织形式与技术创新，也将使突破数据现有权属、合规、安全等困境成为可能。所以说，产业级数据治理来源于实践需要，是数据要素理论取得突破后，在数据治理领域的“投影”，是对数据治理理论体系进一步的创新和拓展，是一种必然的演进趋势，也是中国特色大数据理论实践体系的有机组成部分。

2. 产业数据链与产业级数据治理的关系

产业数据链是政府、企业、组织、个人持有的数据从流通进入产业数据生态到被消费使用与评价的整个过程，是发生在不同企业、不同行业间的一项项组织有序的数据生产流通活动，类似于工厂里面的“生产流水线”。与产业级数据治理中的具体内容活动不同的是，产业数据链中的具体活动属于生产性、价值创造活动，产业级数据治理的活动则属于管理管控、秩序维护活动。从关系上来讲，产业级数据治理与产业数据链是治理与被治理，保障与被保障的关系。

首先，产业数据链是产业级数据治理的核心治理对象。一方面，产业级数据治理通过建立一系列的管控要求，并将这些要求附加到产业数据链的相应环节，从而实现对产业数据生产过程的管理和控制，起到维护产业数据链内相关主体的合法权益，引导产业内相关数据生产活动有序开展，防止与产业数据流通有关的安全、合规、质量问题发生等作用。另一方面，作为被治理对象，产业数据链也是驱动产业级数据治理演进迭代的重要驱动力，产业数据链中的各主体针对产业级数据治理实施过程中存在的缺陷不足进行分析总结，向产业级数据治理提出改善意见，促进产业级数据治理方针的持续优化与改进。

其次，产业级数据治理是产业数据链高效运转的关键保障。产业级数据治理通过向产业数据链输送多维度的资源支撑和建立各主体需要共同遵守的秩序规

则，实现对产业数据链效能的持续提高。一方面，产业级数据治理从政策、组织、技术、流程、环境等多个方面，实现对产业数据链内相关资源要素的统筹配置，优化和畅通产业数据链的各环节流程和组织方式，为产业数据链提供支撑保障。另一方面，产业级数据治理推动产业数据链上下游达成数据生产流通协作的规则共识，破除产业数据链中相关主体的合作与交易顾虑，提升产业数据链的整体效率和价值产出，从而保障产业数据链高效运转。

最后，必须注意的是，行业级和企业级数据治理与产业数据链之间并没有直接关系，而是为产业数据链提供“原料”。打个比方，如果产业数据链是“生产流水线”的话，产业级数据治理就是所在厂房的一系列“配套管理设备和人力资源”，而行业级和企业级数据治理只是在厂房外源源不断地为生产线提供满足质量要求的“原材料”。

3. 产业级数据治理七大板块的内涵定义

产业级数据治理整体框架为“124”体系，即“一套政策，两大保障，四大机制”。**一套政策**是指产业级数据治理的政策和法规；**两大保障**是指组织和技术两个层面的保障，包括组织规范和技术支撑；**四大机制**是指连通机制、互信机制、执行机制、监督机制四个层面的运作机制。

板块一：政策法规

定义：政策法规是指国家和政府为推动产业数据规模化流通而制定的相关法律法规与政策要求。

目标：打造有利的产业级数据治理法律和政策环境，以加快产业内数据的流通和利用，保障数据的安全和个人的隐私保护。

板块二：组织形式

定义：组织形式是指产业生态内不同数据主体组织之间，实现联合、合作、沟通与协同的多种形式形态。

目标：通过不同形态的组织方式，聚拢产业数据链上下游的相关主体并进行规范管理，提高不同参与主体间协作和互通的工作效率。

板块三：技术支撑

定义：技术支撑是指创新的大数据相关技术，通过创新技术的应用，实现对

产业数据的高效治理。

目标：通过创新能力，从技术的维度满足数据的存储交互、隐私保护、信息整合等治理需求，提高产业级数据治理的效率和便捷性。

板块四：连通机制

定义：连通机制是指实现产业内多主体间数据的互相连接，促进数据语义融通和共享利用的相关机制方法。

目标：从底层数据的层面打通不同行业企业数据关联融合的“通道”，为产业级数据治理的其他管理职能提供实现基础。

板块五：互信机制

定义：数据互信机制是指为了实现产业内各数据主体之间互信、互认，塑造平等互利合作环境而建立的一系列机制和体系。

目标：通过建立互信机制，形成不同角色之间互相信任、通力合作的基础，让各方可以放心地共享和使用数据，增强参与主体的参与信心，扩大数据生态的范围和规模。

板块六：执行机制

定义：执行机制是指围绕产业数据链全过程的管理需求和管理目标，从操作层面执行针对产业数据链相关活动和产出物的具体管理动作的机制。

目标：从实质上施行数据流通过程中的治理手段，实现产业数据链各环节的环环把控和质量、安全保障，提升数据流通过程中需求方的整体体验。

板块七：监督机制

定义：监督机制是指为了确保数据的正常流通和应用，防止数据滥用和违规行为而建立的监管及跟踪机制。

目标：通过监督机制保障数据流通整体环境和活动的合规性，防范和减少数据风险，保护数据要素市场竞争中相关方特别是弱势方的合法权益，促进产业数据经济的健康运行和社会经济的稳定发展。

四、产业级数据治理框架详述

（一）政策法规

通过法律法规的制定和实施、发展方针的指导、产业政策的支持和公共倡议

的推进，可以为建立健全数据治理体系提供依据，促进产业级数据的安全、合规和可信使用，实现数据价值最大化，推动数字经济的发展和社会的进步。

1. 法律法规

法律法规是指产业级数据治理所需遵循的各层级的法律和规章。它们为产业级数据治理提供了明确的法律框架，规定了数据的收集、存储、处理、传输和使用的规范和要求。法律法规可以保护数据的隐私和安全，维护公平竞争和数据主体的权益，确保数据治理的合法性和可行性。

例如，比较典型的有欧盟的数据法规体系。欧盟注重理念，精于立法，先后颁布了多部法律法规，以保障良好的数据生态环境。《通用数据保护条例》（简称“GDPR”）保护欧盟公民个人数据隐私权；《数据市场法案》（Digital Market Act）明确大型互联网平台企业关于数据、算法、反垄断的义务清单；《数据法案》（Data Act）规定数据提供者、使用者和第三方之间的数据访问和利用的权利义务关系；《数据治理法案》（简称“EDGA”）建立更安全可靠和透明的数据治理框架。欧盟通过一系列不断迭代的法规体系和政策文件，不断优化各领域产业数据的治理和合作方式，规范和促进数据产业的发展。

2. 发展方针

发展方针是指国家或地区对数据产业和数据治理发展的总体指导思想和目标。通过制定发展方针，可以明确数据产业的发展方向、发展重点，提供有利的政策措施，推动数据治理向高质量、高效益方向发展，促进数据经济的健康发展和社会进步。

例如，比较典型的是中国的数据要素改革发展方针。中国在 2020 年 4 月下发了《关于构建更加完善的要素市场化配置体制机制的意见》，在劳动力、土地、资本、技术四个生产要素的基础上，首次将数据作为第五生产要素，将其纳入国家发展战略的核心。2022 年 12 月又下发了《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》（简称“二十条”），进一步从数据产权、流通交易、收益分配、安全治理四方面提供了发展指导方向和方针指引。以上文件的发布极大促进了数据要素市场生态的加速形成，并不断完善国内的数据要素

治理相关法律、制度、组织、标准、规范，在政府引导下加强数据的跨部门、跨领域、跨地区的流通合作。

3. 产业政策

产业政策是指针对数据产业的发展所制定的经济、技术和组织上的具体政策。产业政策可以涵盖数据产业的规划、建设、培育和升级等方面，鼓励创新和投资，提供政策支持和激励，推动数据产业的发展和竞争力的提升。

例如，在数据要素改革的总体发展方针下，国内各省市数据管理政府机构先后出台了一系列数据产业相关的政策，包括上海市《立足数字经济新赛道推动数据要素产业创新发展行动方案（2023-2025 年）》、浙江省《推进产业数据价值化改革试点方案》等，立足大数据产业发展层面，引导推动数据产业的创新和发展，增强数据治理的一体化和整体效能。

4. 公共倡议

公共倡议是社会各界对数据治理的理念倡导和共同呼吁。公共倡议可以由政府、企业、非营利组织和公众等共同发起，旨在强调数据治理的重要性、促进数据合规和公平使用，促进数据的开放共享和社会共荣。公共倡议可以引导和推动数据治理的良好实践，形成社会共识和共同行动。

例如，为与各国合力推动全球数字经济健康有序发展，加快新型数字基础设施建设，促进数字技术与实体经济深度融合，中国提出《全球数据安全倡议》，呼吁共同探讨制定反映各方意愿、尊重各方利益的数字治理国际规则，积极营造开放、公平、公正、非歧视的数字发展环境。2023 年 5 月 18 日至 19 日，中国同中亚五国在西安举行“中国—中亚峰会”，各方支持在《全球数据安全倡议》框架内构建和平、开放、安全、合作、有序的网络空间，共同落实好《“中国—中亚五国”数据安全合作倡议》。

（二） 组织形式

产业级数据治理需要不同行业和组织之间的合作和协同，组织间在共同的发展目标驱使下，通过形式多样的跨界合作模式，实现数据互联互通和资源优化配

置，达成利益共享与发展共赢。具体的组织形式包括合营、共同体、联盟、协会和交易市场等。

1. 合营

合营是指两个或多个组织共同投资或拥有一个数据要素型企业¹⁶，各方可以共同经营或指定特定方经营，目的是实现数据资源、利益和责任的共享和互补，以及经营风险的分担，以创造更高的价值收益。合营是所有组织形式中，联合的独立组织间利益绑定和联系最紧密的组织形式。在建立了合营企业的前提下，组织间可以围绕合营企业的经营范围，形成以合营企业为核心连接对象，注资企业为利益共享体的跨界协作。

例如，国内比较活跃的数字广东公司、数字广西集团有限公司、云上贵州大数据（集团）有限公司、上海数据集团公司等，都是由多个单位或企业投资持股，以数据为核心业务的具有功能保障属性的市场竞争类企业，在实际的企业运营过程中，促进了投资单位间数据互联互通、业务经验共享、标准规范共创、资源优化配置的典型案例。

2. 共同体

共同体是某个领域内的多个组织在一定范围内形成的，以数据为核心业务对象和载体的，具有共同利益和共同责任的组织形式。共同体以达成特定领域或产业链条的数据可信互联为目标，在公约、协议或共识的框架之下，通过开展多层次数据合作来解决共同面临的产业发展问题，实现产业内共同的发展目标。

例如，比较典型的有欧盟能源数据共同体。欧盟通过建设共同数据空间，提供数据使用和交换的配套工具和基础设施。基于共同数据空间，能源领域的相关企业和组织在数据和能源发展相关的倡议下，以产业发展需求为导向，以客户为中心，以提供安全、可靠的绿色能源供应为目标，包含配电系统运营商、能源服

¹⁶ 数据要素型企业是指将直接参与数据资源要素化的企业。数据要素型企业按照数据供求关系可分为需求型数据要素型企业、供给型数据要素型企业、混合型数据要素型企业三类，按照数据业务特点可分为技术型数据要素型企业、服务型数据要素型企业、应用型数据要素型企业。

务公司、输电系统运营商、行业协会、市场运营商和研究机构在内的多种类主体基于数据空间形成利益共同体，即欧洲能源数据共同体，从而增强能源数据可用性和可靠性，促进数据跨行业共享。

3. 联盟

联盟是多个组织或企业以实现共同的目标或利益而建立的合作伙伴关系。数据联盟通过协商合作、资源互助和信息共享来达到共同目标，联盟成员之间通常保持相对的独立性，主要在特定的数据领域或议题上合作密切，互相支持。

例如，美国的“医药数据共享联盟”(Accumulus Synergy)，由辉瑞、强生、罗氏、赛诺菲、武田、礼来、Amgen、Astellas、Bristol Myers Squibb、GS K等10家知名生物制药企业组成，成员企业向联盟提供资金支持和初始应用开发，共同开发数据共享平台，推动全球医药产业和卫生部门之间的数据实时生产、交换与协作，使医药监管过程高效化、透明化，促进医药数据的大规模汇聚融合，支持药物模型的精准、快速训练，降低研发成本，为患者带来安全有效的药物。

4. 协会

协会是相同行业、专业或利益一致的个体或组织自愿形成的组织形式。成立协会可以为协会成员提供数据相关的行业交流、知识共享和资源互助的平台，达成其倡导和促进自身数据领域业务发展的目的，并通过共同的行动、交流和合作来推动共同利益的实现。

例如，比较典型的有日本官民数据活动共通平台协会(Japan Public-Private Data Utilization Platform Association)。协会的重点工作是指导构建官民数据共通平台，通过制定平台运营规范，研究促进平台运营的政策措施，提供开发、测试环境和宣传活动，吸纳在日本开展业务的微软、亚马逊和谷歌等超大型互联网外企入会等，促进公共部门和私营部门之间的数据共享和合作、促进社会机构和个人使用官民数据共通平台进行数据利用与创新、促进外企与国内企业数据的互联互通。

5. 交易市场

相较于其他组织形式，交易市场是结构最为松散、成员间利益绑定最弱的一种组织合作形式。数据相关主体通过商业化的数据交易行为联接在一起，在遵守数据交易市场的规则、流程、规范的前提下，实现数据的输送、资源的配置和利益的交换。

例如，中国的北京、上海、深圳等地的数据交易所，以及其他基于数据交易所实现场内交易的相关数据主体，在具体的市场交易过程中，形成的就是属于交易市场类型的合作形式关系。另外，以美国为例，其数据交易市场，既存在由微软 Azure、Datamarket、Factual、Infochimps 等数据中间平台撮合实现数据买卖的数据合作主体，也存在安客诚(Acxiom)、Corelogic、Datalogix、eBureau、ID Analytics 等以数据经纪商身份，收集和共享数据，实现数据买卖交易的不同合作对象，同样都属于交易市场的类型。

（三） 连通机制

连通机制解决的是数据互连互通的渠道和方式的问题。在产业级数据流通汇聚的过程中，来自不同主体、不同行业的数据的格式、内容、标准、结构等都存在不同程度的差异，会对数据的融合分析带来阻碍。因此，产业级数据治理需要通过一系列的手段和方法，解决差异问题，实现不同数据源的连接和贯通，以促进数据的流动和整合。

1. 转换协议

转换协议解决的是产业数据链内不同主体之间数据的“握手”问题。数据转换协议是用一种各方认可并共同遵守协议方式，为不同主体在数据合作意向范围内开展共享数据交换建立统一渠道，实现不同主体间数据的快速集成和高效交互。

例如，基于数字对象架构的数联网，以及数据空间技术体系内的 IDS 连接器等，都属于协议化的数据转化连接方式。通过数联网协议可以进行不同主体数据平台之间的配置式组网互联，支持平台间数据传输和信息交换，实现互操作性和互联互通，提高互联互通能力和效率。通过 IDS 连接器，不同主体间可以建立标

标准化通信协议下的数据连接接口，通过获得唯一身份标识、连接器部署配置、证书安全设置等，建立起与其他组件间安全可靠的互操作通信，实现数据空间内分布式数据之间的安全交换。

2. 统一语义

统一语义解决的产业范围内各主体在数据内容层面的语义统一问题。统一语义通过建立多方共同认可的数据语义描述，确保不同数据源中的数据能够获得一致的解释和理解。基于统一语义，可以建立符合参与产业级数据流通的不同主体视角的，自洽兼容、准确清晰、逻辑完备、理解一致的语义环境，消除数据语义不一致导致的误解和错误，实现数据的准确匹配、流转、融合和应用。

实现产业间数据语义统一的方法有很多，既包括从传统数据语义标准化领域衍生的跨主体、跨领域的统一元数据、统一数据范式、统一业务术语、统一模型与接口定义、统一指标与标签定义等等，也包括创新的面向产业级数据流通形势的新型统一语义手段，如数据编织、多模态统一模型、知识图谱构建等。

例如，美国白宫围绕开放数据制定了统一开放数据的元数据标准 POD v1.1，保障各方（数据提供者、发布者、中介和使用者等）对数据内容、含义和格式等的共同理解。这套规范不仅在美国地方政府得到了广泛应用，也推广到英国、澳大利亚和爱尔兰等国家。

3. 全局标准

全局标准解决的是产业数据链内的数据能够全局管理、标准统一的问题。围绕全局标准，需要制定应用于产业级数据链全局的统一、一致的数据标准规范并落标贯彻。全局标准可以涵盖跨行业、跨企业的数据的命名规则、编码规则、格式转换规范、数据质量标准、数据分类标准、数据安全标准等方面，通过建设和遵循这些广泛的数据标准，可以减少跨组织数据治理与维护的成本，促进数据共享、互操作和安全。

例如，中国的银行业金融机构监管发布 EAST5.0（银行业金融机构监管数据标准化规范），制定了其关注的监管报送数据的质量标准和规范要求，涉及规则共计 3283 条，督促银行业金融机构及理财公司搭建数据检核平台，在数据报送

前参照检核规则对当期报送数据质量进行评估，提高监管数据的准确性和合规性。

4. 分布地图

分布地图解决的是产业数据链内的数据能够分布可视、直接触达的问题。分布地图是面向产业内数据消费者的产业数据聚合视图，为数据消费者提供产业级数据目录、数据可视化展示、数据信息查询、数据检索等地图式服务。分布地图是连接机制中重要的一环，相比转换协议、统一语义、全局标准等偏中后台的促进数据连通的活动，分布地图是前台的，直接面向用户的，为产业数据链的运作提供重要载体，确保跨领域聚合数据的可观、可查、可触达。

例如，英国的政府数据开放平台，既可以提供综合的开放数据的数据目录，又能提供数据的快速检索与筛选功能，除按主题词进行快速匹配外，还可按发布部门（及部门类型）、主题、文件格式、数据集类型、内容标签等维度进行筛选，检索结果可按热门程度、发布时间、最近修改时间等标准显示。此外，还支持布尔逻辑检索、字段限制检索及字段与短语检索。

5. 共享模型

共享模型解决的是产业数据链内各主体的数据内容能够集成融合的问题。在实现产业级数据互连互通的过程中，支持数据进行共享和交换的产业级共享数据模型是基础底座。产业级共享数据模型可以包含产业级数据本体、产业级主数据、产业级整合模型、跨领域联合模型、大规模语言模型等多种实现方式。

例如，浙江省经信厅为推动企业的智能化改造和集群协同化发展，深入推进产业数据价值化改革，探索构建浙江省产品主数据生态系统，牵头打造浙江省工业产品主数据管理平台，推动企业产品主数据系统-行业数据仓（行业产品主题库）-产业数据仓（工业产品主数据管理平台）三级协同工作机制的形成。促进各行业产业链上下游企业产品主数据的归集接入、数据融合和数据开发等，打通行业企业数据壁垒。

（四）互信机制

不同身份、不同认知、不同能力的多个数据持有主体难以在复杂数据市场形

势下建立互信，是制约产业数据链内的相关活动的重要因素之一。互信机制通过公约共识、信用证明、公证服务、仲裁调解、授权赋能和行权保障，可以促进公平、公正、规范、可靠的互信产业环境的形成，消除各主体间的信任顾虑，为实现数据跨行业加速流通融合建立基础。

1. 公约共识

公约共识是指在不同的组织形式下，不同数据主体通过签署或发布公约、协议、共识书、行为准则等，确立数据使用、共享和管理的原则和准则，基于共同的原则和规范达成围绕产业级数据治理的共识和合作，促进各组织的自律和互相监督。

例如，由人民数据管理有限公司、中国电信集团、湖北数据集团（湖北省数据交易所筹）、数字安徽有限责任公司、浙江大数据交易中心等 10 家单位共同发起数据要素联盟的同时，发布了全国首部《数据要素联盟公约》，明确了联盟的管理原则、会员制度、组织管理要求、行为约束规定，促进联盟成员单位的深度合作。

2. 信用证明

信用证明是指通过评估和验证参与方的信用记录和商业行为，对其进行信用评级和证明，是产业级数据治理中确认数据提供方或使用方的信用状况和可靠性的一种方式，以便其他参与方更加信任和依赖信用证明主体的数据和服务。

例如，深圳数据交易所在深圳市发展和改革委员会、深圳市市场监督管理局（信用办）的支持下，启动全国首个数据交易信用体系建设，面向律师、律师事务所及数据商三大主体构建信用评级体系。通过制定信用评价及管理制度、科学设置信用评价指标体系，构建数据交易信用评级服务系统，鼓励行业自律组织开展信用共建共治，打造和建立诚信、互信、可信的数据交易生态。

3. 公证服务

公证服务是指公证机构为各数据持有主体提供的与数据相关的专业、权威的认证服务。公证服务可以包括对数据主体组织、数据相关处理行为、数据治理及

交易文件、数据服务及产品等对象的认证背书服务，通常从合规性、质量、安全性、真实性、可信度等方面考察对所认证对象的认可程度，通过颁发证书、证明、公证书、标签等确立公证服务机构的认可态度。

例如，欧盟推出的欧盟通用数据保护条例（GDPR）认证体系—Europrivacy（欧洲隐私）。Europrivacy 作为第一个符合 GDPR 规定的官方认证机制，用于评估、记录、认证和评价企业对 GDPR 等数据保护法规的合规情况，经评估后符合这一标准的认证对象，证书中可以获得欧洲数据保护印章，增加企业的价值和对其服务的信任。

4. 仲裁调解

仲裁调解是在产业级数据治理中解决纠纷和争议的一种方式。通过建立独立的仲裁机构或调解机构，对数据相关的争议进行仲裁和调解，以解决分歧、维护公正，并促进各方的合作和和解。

围绕数据的仲裁和调解，国内已有许多探索和实践。例如，温州成立数据资源仲裁院，致力于平等保护数据提供方、使用方和数据交易各方的合法权益，完善多元化数据资源民商事纠纷解决机制；深圳国际仲裁院与深圳交易集团、深圳数据交易所签订合作协议，探索共建数据交易领域“调解+仲裁”的多元化争议解决机制。

5. 授权赋能

授权赋能是指个人、企业、组织、政府将自身的数据授权给其他主体，用以赋能其他主体的业务发展和服务提升。数据授权赋能的过程需要有完备的数据授权管理流程和许可机制，在发挥授权主体将数据共享出去“利他”的主动性的同时，确保数据授权过程的合法性，建立授权双方的互相信任，推动数据的合规共享与利用。

例如，在美国，个人可以通过数据合作社，一种由个人组成的社群，基于 DSpark、Data Republic、Ocean Protocol、Dawex 和 Enigma 等数据合作社运营机构建立的数据平台，分享个人的数据资源，按照个人意愿将自己的数据存储在共享的平台上，并与其他成员共享和访问数据，以实现更广泛的社会、科研或商

业目的。

再如，在欧盟，由欧盟认证、申请机构自愿登记的、独立且非营利性的数据利他主义组织，推行“数据利他主义”，面向不同行业积极推进数据共享，从而为企业、组织和个人提供个性化定制的数据利他服务。

6. 行权保障

授权赋能与行权保障两个活动之间是存在一定关联的。在不同数据主体自主、自愿授权数据，以赋能其他组织机构的同时，也需要建立不同主体行使自身数据权利的保障机制，保证对数据主体合法权利的充分尊重，即行权保障。行权保障包括权利管理保障、权利控制保障、行权结果查看保障、权利反馈保障等多个方面。

例如，广东省的个人数字空间、法人数字空间，为个人和企业提供了数据的携带、授权、存证、溯源的边界安全服务，让其掌握和管理自身授权给政府公共服务部门的数据，为个人和企业提供了更多选择和控制权，促进了公平、透明和负责任的数据管理。

（五） 执行机制

执行机制主要关注的是产业数据链操作层面的具体执行机制，包括了环节性质的需求对接、登记确权等活动，也包括准入准出、服务评价、过程保障等日常性活动。在产业数据链范围内存在平台、数据交易所等主导运营方的情况下，执行机制中的具体活动由主导运营方承担；在不存在运营角色时，执行机制中的具体活动由产业数据链中的相关主体分摊操作，并引入必要的第三方机构和组织进行支持。

1. 登记确权

登记确权是指产业数据链中的确权环节中具体开展数据相关权属登记和确认的过程活动。通过对数据的登记和确权，可以明确不同主体间数据持有权、使用权等权属关系归属，防止数据滥用和侵权现象，保证数据被合法和正当使用。

目前在数据登记确权领域，国内已经有众多的实践。例如，广东省政数局颁

布了公共数据资产凭证，河南省数据要素确权与可信流通平台（河南根中心）发布了数据要素登记证书，北京国际大数据交易所数据资产登记中心发布了数据资产凭证及数字交易合约，贵阳大数据交易所建立了数据要素 OID 登记子节点授权机制等等，这些实践都为产业数据链中具体的数据运营和交易等彰显价值的环节提供了底层保障。

2. 需求对接

需求对接是指根据数据使用方的需求与数据提供方进行沟通和协商，完成需求匹配，达成双方合作的过程。通过需求对接，可以明确数据的具体需求和使用场景，确保数据的精准提供，满足实际需求。

例如，华中大数据交易所的《华中大数据交易平台规则》中，在设计数据交易流程时，将数据交易流程分成了数据出售和数据需求两个类型，数据出售流程中设计了需求提交、数据审核、服务应答、服务提交等对接活动，数据需求流程中设计了数据提交、数据审核等对接活动，这些活动作为必要的需求对接活动，实现了供需双方在需求描述、需求提交、需求沟通、需求确认等层面达成一致的共识。

3. 准入准出

准入准出是对数据产业链中参与流通交易的数据产品进入或退出产业数据链进行审核和管理的过程。准入主要是审核确认数据产品是否满足相关的合规、安全、信息描述规范，是否达到进入交易市场的基本条件；准出主要是审查数据产品是否满足质量、业务场景、技术规范等数据交易相关的要求，是否达到从数据交易市场售出到数据消费对象使用环境的条件。

例如，深圳数据交易所围绕数据交易前、中、后全流程构建了“1+4+7”的规则制度体系，建立了内外结合、专家委员会为主的审核机制，建立负面清单、谨慎清单及交易标的准入规则，提供在数据交易所的交易平台上架流通数据产品的全流程指引。

4. 过程保障

过程保障是指围绕产业级数据链内完成对接的两个或多个主体，以需求实现为目的，执行相应的质量、安全保障手段，规避和防范由质量和安全管控不当导致的风险和问题。

例如，深圳交易所为更好实现各市场参与主体在场内数据交易过程中的安全保障，从安全评估和技术创新两方面落实措施到位。在安全评估层面，通过引入数据安全评估的第三方服务机构，为数据交易双方提供数据安全能力和数据安全技术评估服务，把控双方交易的安全风险；在技术创新层面，以自主可控可信数据空间（Trusted Data Matrix，简称 TDM）架构为基础，融合区块链、物联网和零信任等先进技术，形成可信空间技术解决方案，有效解决了数据协同策略难统一、数据使用难监控、发生事件难溯源等诸多数据安全管控具体问题，实现了消费电子产品、新能源锂电池等生产全过程质量数据的可信共享。

5. 服务评价

服务评价是对产业数据链中提供数据服务的机构和组织的具体服务质量和效果进行评估和反馈的过程。通过开展服务评价，评估服务的满意度和改进空间，以提供更好的服务和满足用户需求，并且甄别出存在服务质量问题的机构和组织。

例如，英国的政务数据开放平台为用户提供反馈表单，通过含有 8 个问题的短问卷调查用户寻找该网站的便捷度、访问原因、数据搜索效果、数据使用目的、数据处理工具等信息，以更全面地了解用户的使用效果和开放数据服务评价。

（六） 监督机制

监督机制是产业级数据治理的重要且必不可少的组成部分。被赋予监督职责的机构需要强有力的公信力和执行力，一般会由各级政府单位或行业监管机构来承担。监督机制下，监督机构既会主动开展监督审计，发现产业数据链中各组织的违规行为，监督整改，也会作为投诉举报的受理机构，调查由群众和组织反馈的违规行为，进行通报和整改督办。

1. 行为审计

行为审计是指监督机构通过对产业数据链的各个环节活动进行事后审查，确认所有参与者的行为符合规范要求，并能够追溯行为的合规性和合法性。行为审计可以从数据采集、数据处理、数据传输等各个环节入手，分析行为是否存在违规操作。

例如，国家互联网信息办公室发布的《个人信息保护合规审计管理办法（征求意见稿）》，对个人信息保护合规审计机制进行了细化，明确了监管部门发现个人信息处理活动存在较大风险或者发生个人信息安全事件时，可以要求个人信息处理者进行合规审计的强制行为审计规定。强制审计必须由个人信息处理者委托专业机构进行审计，且个人信息处理者应当在收到监管部门通知后尽快按照要求选定专业机构。在强制审计的情形下，专业机构应当在 90 个工作日内完成合规审计。情况复杂的，报监管部门批准后可适当延长。

2. 举报受理

举报受理是指监督机构可以通过建立数据违规问题举报受理渠道，接收相关人员对产业数据链中违规行为的举报。监督机构在接收到举报后，需要进行初步审核，对举报内容进行调查核实。

举报依据可以是国家和地区内数据相关的法律法规制度和相关政策要求。例如，在欧盟《一般数据保护条例》（GDPR）生效后，欧洲各国相关投诉案件大幅增加。英国信息专员办公室（ICO），三周内收到了 1106 项数据保护投诉；法国国家信息自由委员会（法国数据保护局，简称“CNIL”）收到的投诉数量与前一年同期相比增加了 50%。

3. 调查验证

调查验证是指监管机构对数据违规行为的举报内容进行深入调查和验证，确认是否存在违规行为。调查验证阶段需要收集证据、了解相关情况，可能需要调取数据、询问当事人或第三方等。

例如，英国税务海关总署（Her Majesty's Revenue and Customs，简称“HMRC”）

被英国隐私与公民自由组织（Big Brother Watch）投诉。该组织认为，税务海关总署在未经同意的情况下收集超过 500 万个生物识别语音 ID 的行为违反了 GDPR 的规定。HMRC 表示自己“鼓励那些打电话利用语音 ID 服务的客户，而不是让客户明确表示同意，但他们可以选择退出并继续按照常规方式使用 HMRC 的服务”。英国国家数据保护局受理了投诉并开展调查发现，客户在呼叫英国税务海关总署的帮助热线时，并不存在可以选择常规方式进行身份认证的选项，并且在被允许访问 HMRC 服务之前，呼叫者需要重复“我的语音是我的密码”（“My voice is my password”）这一短语。该录音被输入到当局的生物识别语音 ID 数据库中，这显然违反了 GDPR 规则，也验证了英国隐私与公民自由组织投诉的真实性。

4. 违规查处

违规查处是指在确认发生数据违规行为后，监督机构采取必要的追责、罚金、教育、清退等处罚措施。对于数据泄露、数据篡改、未经授权的数据收集和访问等违规行为，进行查处的过程需要依据相关的法律法规和数据治理政策，确保程序的公正和透明。

例如，以下是欧盟国家监管机构依据 GDPR 对违规企业进行惩处的案例。2019 年，法国监管机构对 Google 未经用户同意的情况下擅自收集用户个人数据的违规行为处以 5000 万欧元罚款。2019 年，英国信息委员会对英国航空公司 British Airways 因公司网站和移动应用程序存在安全漏洞，导致超过 50 万名用户的个人数据被盗取的违规问题处以 2030 万英镑的罚款。严格而公正透明的违规查处结果可以促进企业加强对个人用户数据隐私和敏感数据的保护，同时提升企业数据处理行为的合规性。

5. 监管通报

监管通报是指监督机构将重大违规行为和处理结果进行通报，向相关部门和公众公开，起到警示和震慑作用，营造公众参与监督的良好氛围。监管通报可以增加产业级数据治理的透明度和公正性，进一步提高监管效果，促使企业或组织

加强自身数据治理体系建设，提高数据保护水平。

例如，《数据安全法》自 2021 年 9 月 1 日实施以来，工业和信息化部信息通信管理局已通报 880 余款侵害用户权益行为的移动互联网应用程序（App）及第三方软件开发工具包（SDK）。其中涉及“违规收集个人信息”占比最高，为 32.5%；其次是“App 强制、频繁、过度索取权限”，占比为 24.8%。

6. 整改督办

整改督办是指监督机构对存在违规行为的组织或个人改正问题、优化机制的进度进行监督和跟踪的过程，确保问题得到解决和纠正。整改督办包括制定整改计划、跟踪整改进度、评估整改效果等环节，以确保问题不再发生，并对整改不力的情况采取相应的处理措施。

例如，前面提到的英国税务海关总署（HMRC）违规收集用户录音的案例，英国信息专员办公室（ICO）在确认海关总署的录音的行为构成“重大违反数据和隐私规则的行为”后，给出了相应的整改督办要求，英国税务及海关总署（HMRC）被勒令删除了 500 万纳税人的录音。

（七） 技术支撑

在数据主体众多、数据链路复杂、传统技术无法实现高度互信和隐私保护的情况下，需要运用先进、创新的技术，通过构建和应用产业级语义技术、数据编织架构、智能化技术方案、隐私计算保护、区块链安全存储等方式，将技术与治理活动内容深度融合，形成基于底层技术能力的支撑保障，实现对产业级数据的有效管理和长效治理。

1. 语义技术

语义技术是通过对数据及其内容进行语义解析，从中提取核心语义信息，并进一步进行语义计算、语义处理、语义推理，实现所有信息组合和连接的技术。语义技术可以实现产业级数据链内各主体与主体之间的数据语义互通，构建通用性、行业性的数据语义网络，解决产业数据分散、语义不统一等问题，支持产业

内各主体更好地开展业务领域数据的搜索、共享与分析。

例如，美国国家医学图书馆研发的一体化医学语言系统 UMLS，利用语义技术构建了医学领域的语义网络。UMLS 涵盖了临床、基础、药学、生物学、医学管理等医学及相关学科，收录了约 200 万个医学概念，500 多万个医学词汇。UMLS 超级叙词表是一个非常庞大的术语库，集成了生物医学和健康方面的本体、叙词表、分类表、疾病编码集、专家系统、词汇表中的术语及相关信息，如 MeSH、SNOMED、CT 等，其语义网络为超级叙词表中的概念提供统一的组织和分类，并揭示概念之间的语义关系。UMLS 语义网络的构建和应用便利了全世界各地的医务人员的医学资源语义拓展查询和生物医学信息获取工作。

2. 数据编织

数据编织是一种数据架构类型，它通过打造数据的虚拟网络，将来自不同数据源的数据运用统一的方法进行虚拟整合和连接，把正确的数据，在正确的时间，传送给正确的人。数据编织可以实现产业数据链内的数据在管理集中、物理分散、逻辑统一的情况下，进行跨领域、跨数据主体的数据发现、语义理解和数据交互，保证所有数据源以灵活且业务可理解的方式交付给产业内数据需求方。

例如，美国陆军采用数据编织技术构建 Rainmaker（造雨者）数据架构，开发通用标准，在数量众多的不兼容的作战指挥系统中实现大量数据转换，解决了“烟囱式”的系统之间缺乏互操作能力，各梯队和职能之间存在信息共享障碍的问题，实现了联合网络链接。Rainmaker 数据结构不仅将美陆军、其他军种和美国盟友之间的数据联系起来，也将过去和未来系统之间的数据联系起来，并支持美陆军和其他机构建立广泛人工智能服务，实现从智能武器到训练 AI 的所有数据在作战网络上的共享。

3. 人工智能

人工智能是一种利用算法和数学模型让计算机实现与人类智能类似反应的技术，具体的技术能力范畴涵盖语音识别、视觉感知、自然语言处理和机器学习等特定方面。运用人工智能技术，特别是生成式 AI 的能力，可以实现产业级数

据链内的语义一致、过程追溯等跨域数据问题的智能化解决，并促进产业级数据治理中各机制内相关业务场景的智能化实现。

例如，Informatica 推出生成式 AI 驱动的数据管理平台 ClaireGPT，让用户通过简单的自然语言提示来消费、处理、管理和分析数据，智能化实现数据发现、数据流程创建和编辑、元数据探索、数据质量和关系探索以及数据质量规则生成。Salesforce 推出 AI 辅助数据分析工具 TableauGPT 和 TableauPulse，通过由生成式 AI 驱动的新方法，为用户提供自动数据分析和个性化分析体验。

4. 隐私计算

隐私计算是一种保护数据隐私的计算方法，它在数据处理或共享过程中，通过使用密码学和隐私保护技术，保证数据的隐私性和安全性，同时提供可计算功能。产业数据链内的各方可以通过隐私计算平台，进行敏感数据的协同计算，并且在进行计算时不暴露原始数据，以满足隐私法规和监管要求，为产业级数据治理提供了隐私保护的可行方案。

例如，国家健康医疗大数据研究院、微众银行、华链医疗科技有限公司三方协作，引入微众银行场景式隐私保护解决方案 WeDPR，应用隐私计算技术实现在保护数据隐私安全的前提下，开展医疗数据分析研究，解决了医疗领域的疾病数据合法合规共享的难题。解决方案利用联合统计、联合建模、联合预测、匿踪查询、隐私求交等隐私计算能力，构建了一套与外部医院、保险公司、体检中心等医疗机构进行多中心科研的隐私协作机制。各方的原始数据均无需出库，通过 WeDPR 进行多方大数据隐私计算，实现应对多场景规模化的医疗数据再利用。

5. 区块链

区块链是一种去中心化的分布式账本技术，通过将数据存储在多个节点上，并使用密码学技术确保数据的安全性和不可篡改性。区块链具有去中心化、透明和可追溯的特点，可以在产业级数据治理中提供安全可信的数据存储和交换机制，构建去中心化的数据共享网络，提高数据的可信度和共享效率。

例如，上海数据交易所建设了数据交易链，充分利用区块链中数据唯一性、不可篡改、去中心化等技术特点，构建了数据产品挂牌目录上链功能，实现挂牌

数据产品全量上链，做到数据产品“一地挂牌全链流通”。相关数据交易参与方只要加入交易链联盟节点，便可查看上海数据交易所的全量挂牌数据产品。

五、 总结与展望

我们正处在一场席卷全球的数字化变革浪潮之中，数据正在深刻影响和深度参与到国家治理、经济运行以及社会生活的方方面面。随着数据成为第五大生产要素，数据要素市场的构建和培育正在“快马加鞭”。对于所有认识到数据价值，并希望通过数据带来价值收益的各类主体来说，积极参与数据要素市场建设、加快数据要素跨行业流通、不断创新业务形态和商业模式的同时，更需要突破传统数据治理的局限性，以更宏观的视角加入产业级数据治理的“社会体系”，夯实高质量产业级数据基础，优化数据要素流通环境，保障基本权益不受损害，从而为更好实现数据价值变现建立坚实基础。

产业级数据治理并不是“一蹴而就”的。作为一套产业数据生态下的“环境治理方案”，要形成相对成熟的机制体制还需要相当长时间的探索、实践和总结，也会面临很多问题需要应对和解决。例如，需要强有力的牵头单位“挺身而出”，组织模式有待不断尝试和磨合，产业级标准体系不够完善，全新的授权和信用体系仍需权威机构参与，数据登记、服务评价等机制有待完善，监督审计尺度和粒度仍待摸索等等。这些问题和挑战需要政府、法律以及各行各业的实践者共同努力，以非凡的智慧攻坚克难，才能逐步实现完备的产业级数据治理理论和实践体系框架，保障和促进数据要素市场的高速发展和高度繁荣。

信安数据聚焦数据管理及数据要素流通领域，致力于成为国内最好的产业级数据治理及要素运营服务商，以“新理念、新技术、新模式”打造极致专业化的产品和服务，帮助客户管理数据资产释放数据要素价值。“风正劲足自当扬帆破浪，任重道远更需快马加鞭。”信安数据作为赛道中的探索者，始终秉承“专业、合作、创新”基本理念，期望与全国客户和合作伙伴一道，持续探索产业级数据治理这一全新课题，携手共进、互利共赢、共谋发展，为全社会数字化转型和数字经济发展贡献“信安智慧”。



 020-87581616

 广东省·广州市

 <http://www.faithindata.com.cn>