



工业互联网产业联盟
Alliance of Industrial Internet

Trusted Industrial
Data Matrix | 2022

可信工业数据空间系统架构1.0

声明

Statement

本报告所载的材料和信息，包括但不限于文本、图片、数据、观点、建议，不构成法律建议，也不应替代律师意见。本报告所有材料或内容的知识产权归工业互联网产业联盟所有（注明是引自其他方的内容除外），并受法律保护。如需转载，需联系本联盟并获得授权许可。未经许可，任何人不得将报告的全部或部分内容以发布、转载、汇编、转让、出售等方式使用，不得将报告的全部或部分内容通过网络方式传播，不得在任何公开场合使用报告内相关描述及相关数据图表。违反上述声明者，本联盟将追究其相关法律责任。



牵头编写单位

工业互联网产业联盟

中国信息通信研究院

参与编写单位

北京交通大学

中国电信集团有限公司

华为技术有限公司

阿里云计算有限公司

东方电气集团科学技术研究院有限公司

百度在线网络技术（北京）有限公司

中车青岛四方机车车辆股份有限公司

中车株洲电力机车研究所有限公司

沈阳鼓风机集团测控技术公司

国能信控互联技术有限公司

中国航空工业集团信息技术中心

中信戴卡股份有限公司

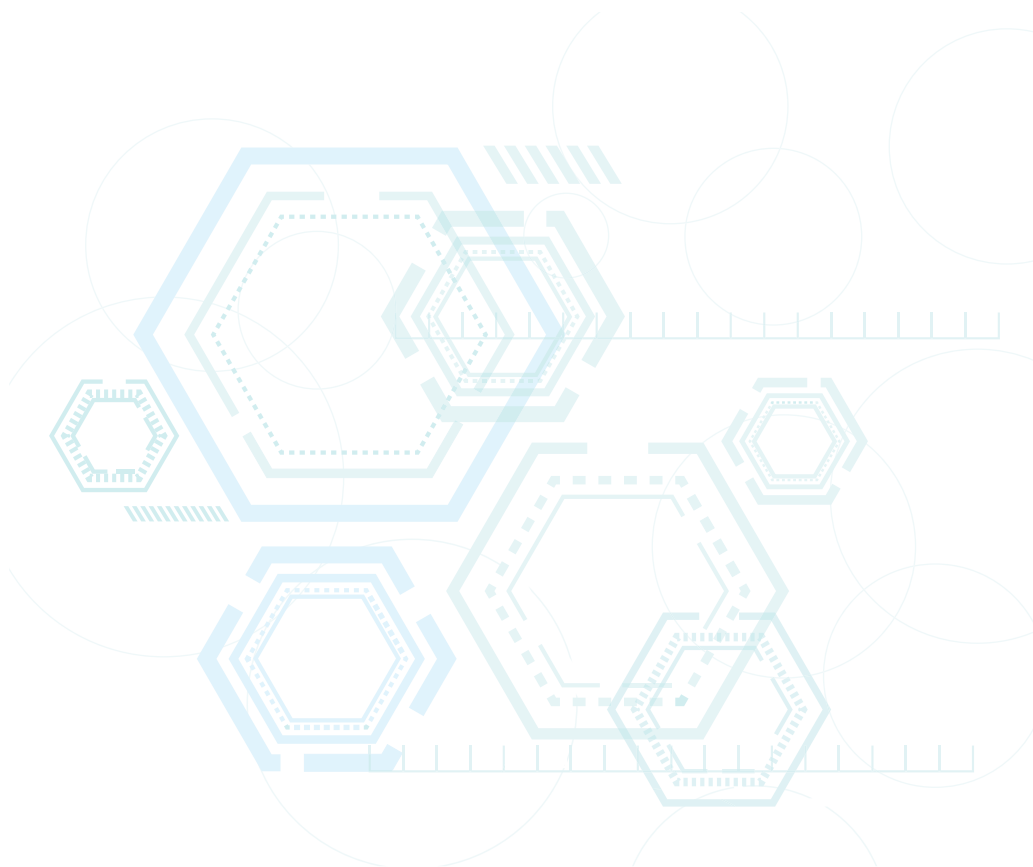
重庆工业大数据创新中心

华控清交信息科技有限公司

上海光华冠群软件有限公司

智能云科信息科技有限公司





目录

Contents

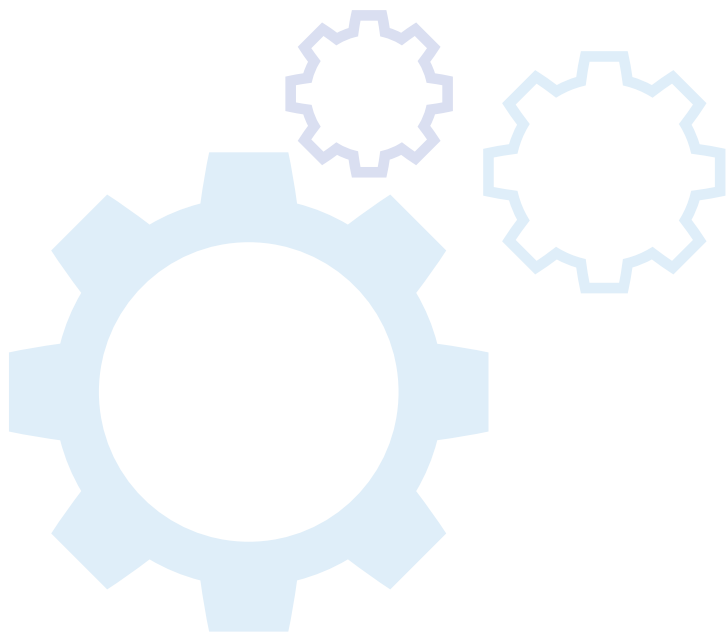
前言	01
Introduction	
第一章 全球发展现状	02
Chapter 1 Global Development Status	
第二章 产业需求分析	08
Chapter 2 Industrial Demand Analysis	
第三章 系统架构	14
Chapter 3 System Architecture	
第四章 技术视角	22
Chapter 4 Technical Perspective	
第五章 标准体系	32
Chapter 5 Standard System	
第六章 产业案例分析	36
Chapter 6 Industry Case Analysis	



前言

Introduction

随着新一代信息技术与制造业的深度融合发展，全球工业数据应用已经进入纵深发展的新阶段，数据作为新型生产要素和重要战略资源，正在制造业数字化转型过程中发挥出更大的作用。在这一进程中，工业数据的流通共享受到广泛关注。顺应新发展形势，我国积极营造多方主体参与的数据共享流通生态，国务院先后发布《关于构建更加完善的要素市场化配置体制机制的意见》、《要素市场化配置综合改革试点总体方案》，明确提出在确保数据安全的前提下，分级分类、分步有序推动部分领域数据流通应用。在此背景下，工业互联网产业联盟联合中国信息通信研究院提出建立可信工业数据空间，并将其作为实现工业数据开放共享和可信流通的新型基础设施，发挥数据要素禀赋。



第一章 全球发展现状

Chapter 1 Global Development Status

（一）为什么需要工业数据的可信交换和流通

随着数字经济的深入发展，数据要素的支撑作用变得愈发重要。在工业领域，工业数据正逐渐从制造过程的副产品转变为企业和供应链环节带来新价值的战略资源，成为提升制造业生产力、竞争力、创新力的关键要素，有力推动了制造业在更大范围、更深层次实现更有效率、更加精准的资源配置。然而，在极大促进了全社会要素资源的网络化共享、集约化整合、协作化开发、高效化利用的同时，工业数据也面临着流通不畅、信息泄露、过度利用等风险。

和其他生产要素一样，数据要素流通能力直接影响数字产业化和产业数字化的产出效率，可信、安全、透明、可计量的数据共享、流通、交换和交易已成为大数据时代下一步发展共识。然而，当前数据要素流通仍不充分、不完善，根据 Gartner 预测，到2022年，只有不到5%的数据共享活动能够依赖可信数据和数据源^{*}。

* Gartner-Data Sharing Is a Business Necessity to Accelerate Digital Business

造成上述现象的一个重要原因是，传统数据流通方式已难以完全满足工业应用需求。目前数据共享流通多为集中式的双边信息传输模式，一种是通过数据交易平台进行共享流通，达成交易的各方通过数据包传递达成交易，在政务、消费、金融、证券等行业已取得显著成效。但对具有复杂来源的工业数据而言，难以保护数据所有者利益，易导致涉及企业核心竞争力的信息暴露以

及数据被使用方二次利用甚至滥用，适用性不强。另一种是基于明文数据API接口进行流通，将加工处理完的单方结果数据以API形式输出，一定程度保护了用户隐私信息以及降低二次利用可能性，但同时也降低了数据价值融合的可行性，难以释放数据协同共享的应用价值。因此亟需寻找一种适合工业场景的新型数据共享流通解决方案，在数据价值释放及数据安全需求中取得平衡。



（二）发达国家加快数据共享流通顶层战略布局

从德国探索构建工业数据空间架构模型，到《欧洲数据战略》明确提出建设以工业为代表的九大行业数据空间，再到日本提出建设“互联产业开放框架”，发达国家展现出抢抓工业数据共享流通主导权的战略意图。

一是工业数据空间实践从德国走向欧盟，企业界积极促进工业数据共享流通。为促进数据安全保护和数据资源释放，2015年德国在工业4.0项目下启动工业数据空间研究，2017年开始向其他行业扩展，并在欧盟达成共识，2020年，欧盟委员会先后发布《欧洲数据战略》和《欧洲数据治理条例》，提出在保证欧洲公共利益和数据提供者合法权益的条件下，构建工业、绿色政务、

出行移动、健康、金融、能源、农业、公共管理、技能等九大数据空间，实现更广泛的数据资源释放和国际数据共享。以此为契机，依托雄厚的制造业基础，欧洲企业界积极投身工业数据空间建设，并形成两条典型的应用路径，一方面从供应链切入，提升透明化管控水平；另一方面从研发端切入，实现大范围的协同研发制造。工业数据空间在制造业、医疗等诸多行业的关键环节显现出巨大的应用价值。

二是欧盟基于工业数据空间，打造GAIA-X（盖亚X）框架，构建数据生态系统。为给数据空间提供统一的基础设施底座，以德、法为代表，欧盟积极推动欧洲统一





的云计算基础设施建设。2019年德国正式提出“盖亚计划”，旨在为欧洲打造一个具备竞争力、安全与可靠的数据基础架构，包括数据生态系统，联邦服务以及云计算、边缘计算和数据存储等基础设施生态系统。其中，联邦服务主要包括身份识别、联邦服务目录、主权数据交换和合规认证等内容。该计划已应用到欧盟的能源、金融、健康、工业4.0、农业、交通、公共服务、智慧生活等领域。

三是日本工业价值链促进会打造互联产业开放框架，聚焦底层数据资源流通。为发挥机器人等工业装备产业的领先优势，实现产业数据的深度共享，打造超智能的“社会5.0”时代，2019年日本工业价值链协会发

布《互联产业开放框架》（以下简称“CIOF”），标志着日本工业数据共享流通布局正式开始。CIOF具有三大特征：一是采用数据流通合同的形式定义工业数据的使用许可范围、供需双方的权利义务以及数据使用过程中衍生的其余问题；二是采用自下而上的方式，由数据提供方和数据使用方直接进行对接，逐步形成双方均认可的数据字典，保障数据流通的独立性；三是现场数据和相关数据资产由CIOF系统直接管理，并通过区块链技术保证数据按需加密和传输，以防止数据伪造。当前日本工业价值链协会正在推进一项包含日本100强企业的工业数据共享项目，将使用区块链技术降低数据泄露风险和运营成本。



（三）我国工业数据共享流通发展现状

中国作为制造大国，同时也是数据资源大国和应用大国，2025年，中国数据总量将跃居世界第一，全球占比有望达到27%^{*}以上，将成为数据量最大、数据类型最丰富的国家之一。与此同时，我国工业数据共享流通正处于起步阶段，亟需充分吸取发达国家在顶层设计、交易共享方面的先进实践经验，构建符合我国发展实际的工业数据共享流通框架，进一步促进数据要素市场化。

一方面，逐步形成战略和立法相互交融的顶层设计。

相关政策密集出台，引导工业数据有序利用。习近平总书记指出，要“构建以数据为关键要素的数字经济”“系统推进工业互联网基础设施和数据资源管理体系建

设，发挥数据的基础资源作用和创新引擎作用。”2021年发布的《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》中明确提出统筹数据开发利用、隐私保护和公共安全，加快建立数据资源产权、交易流通、跨境传输和安全保护等基础制度和标准规范，推动数据跨境安全有序流动。此外，2020年工信部印发《关于工业大数据发展的指导意见》和《工业数据分类分级指南（试行）》，提出建设工业数据空间的重点任务。2021年工信部印发《工业互联网创新发展三年行动计划（2021-2023年）》，再次提出在重点行业建立工业数据空间。引导数据有序开放共享。

^{*} 中国国家信息中心主任刘宇南在国家公共数据开放平台建设地方专题会的发言

法律法规不断完善，保障工业数据合规流动。 国家积极推动数据安全流通相关法律法规制定，2021年6月份《数据安全法》审议通过，强调在加强对重要数据的保护，保障数据有序流动的同时，鼓励塑造数据自由流动的市场秩序，提高数据流动性。《数据安全法》提出支持市场主体、行业协会乃至政府机构共同搭建数据交易平台，探索建立适当的数据交易程序、交易担保机制。通过创制规则，减少不必要的数据要素交易成本，最大程度地加快社会数据交易的频次。可以看出，国家正积极发挥上位法律对数据交易指引的指引作用，减少实践中“不会”“不当”和“不敢”交易数据的现象。

另一方面，工业数据流通需求旺盛同时面临三大难题。

为促进数据共享、流通、交易，各地开展了积极探索，近年来，贵阳、上海、武汉等多地相继成立了数据交易中介机构，不断探索推出相关数据交易方案和规则。然而，当前工业数据共享流通面临质量、主权、交易等难题。一是质量管理难。高质量数据是共享流通的重要前提，但工业数据的采集记录标准、频率和时间千差万别，导致数据质量参差不齐和资源浪费。二是确权难。工业数据涉及主体众多，主权边界难以界定。三是交易难。工业数据的安全要求远比消费数据高，企业担心泄露商业机密或暴露客户隐私。

在本白皮书第三章中，我们将针对上述问题提出面向我国的可信工业数据空间解决方案。



第二章 产业需求分析

Chapter 2 Industrial Demand Analysis

为了梳理工业企业在可信工业数据共享、流通中的典型场景、采用的技术和管理手段以及存在的问题，工业互联网产业联盟开展了《工业数据流通与使用状况》问卷调查，共回收有效问卷125份，覆盖电子信息、制造、能源、物流、工业服务、工业信息技术等行业。

问卷显示，96%的工业企业存在数据流通场景，覆盖研发、生产、物流、销售、服务等产品全生命周期。其中，研发、生产、服务的覆盖率超过了7成，如图2.1所示。

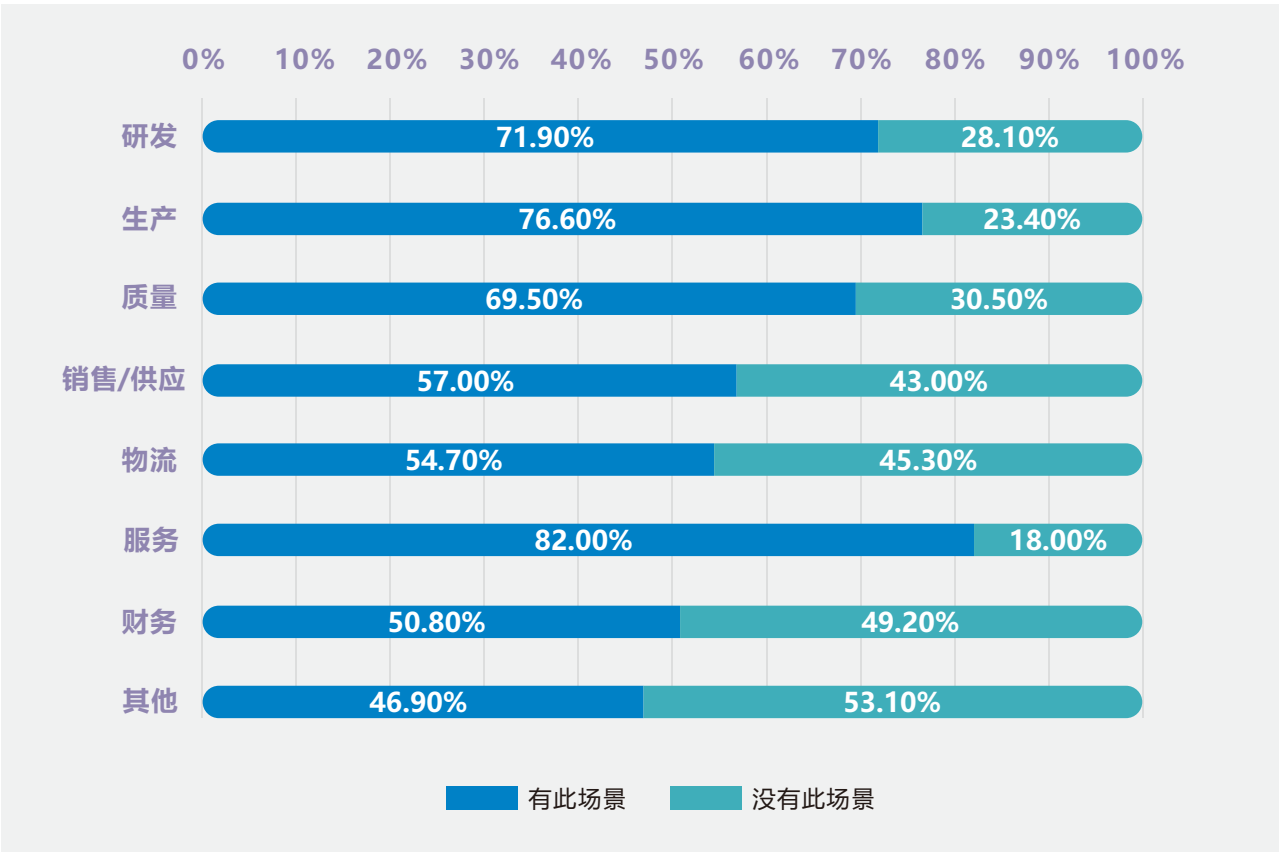


图2.1 数据流通场景

当企业作为数据提供方时，其主要顾虑依次包括提供的数据被用于合同目的之外、商业情报随数据泄露、技术随数据流出，以及接收方对数据保管不善，如图2.2所示。

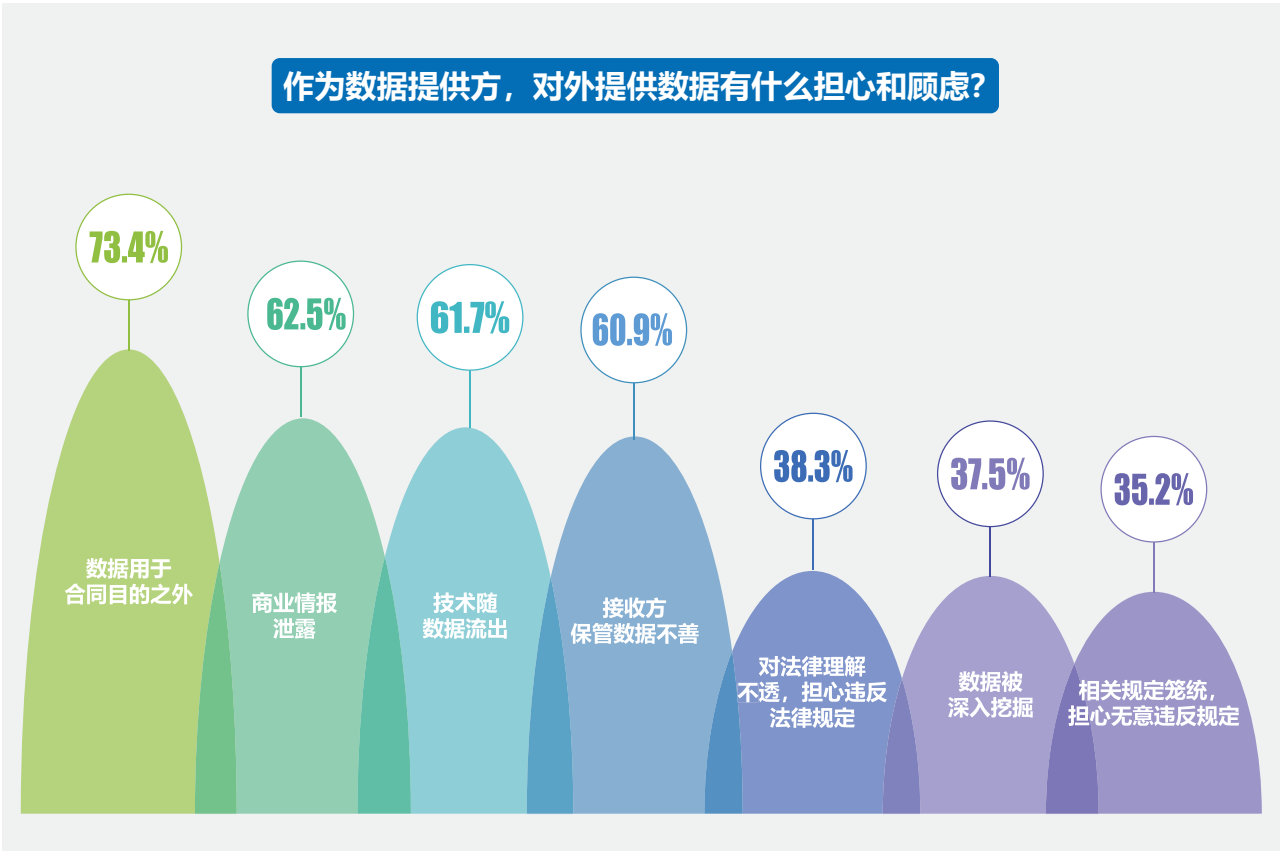


图2.2 数据提供方的主要顾虑

企业目前采取的应对措施主要有义务性约束和技术性约束两类，如图2.3所示。其中，在“义务性约束”中，85%的企业通过合同、协议、规定等对数据使用方进行约束。在“技术性约束”中，近8成的企业主要采用

身份认证（78.9%）、访问日志（70.3%）、数据库日志（64.1%）等手段。目前，数据脱敏、区块链、隐私计算等技术的使用率还不高，如图2.4所示。

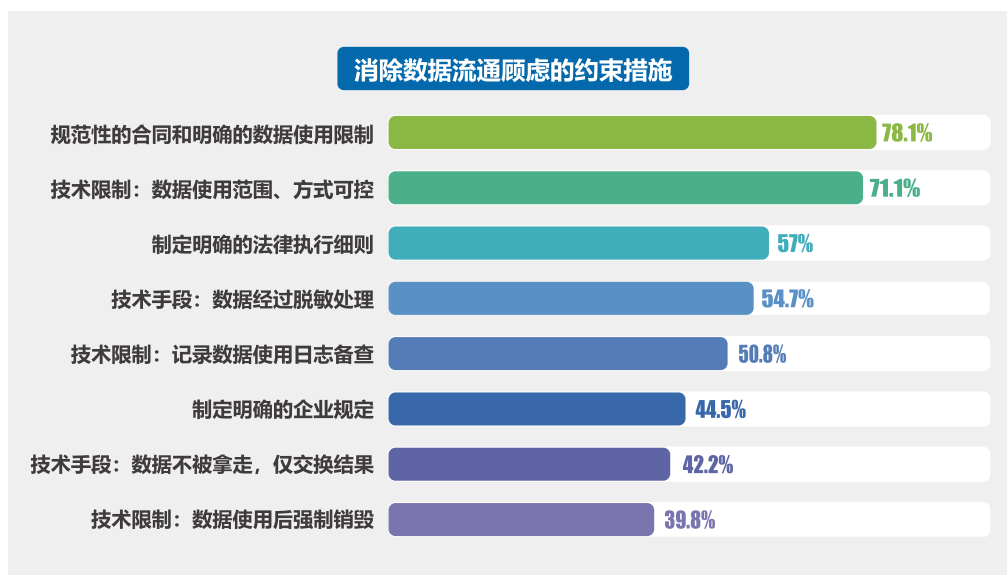


图2.3 消除数据流通顾虑的约束措施

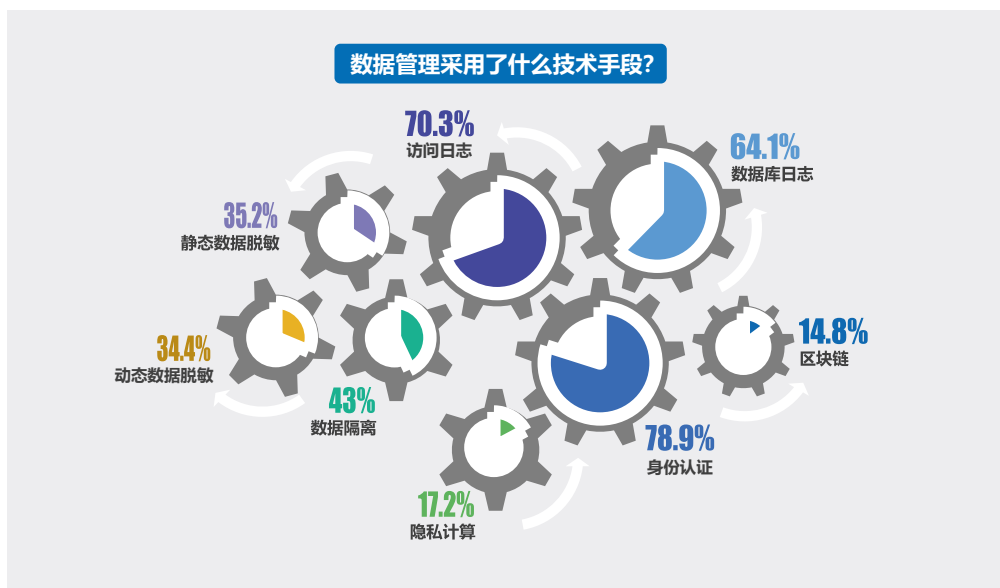


图2.4 数据管理的技术手段

尽管八成的企业采用了义务性约束，但实际的约束效果欠佳。作为数据接收方，有4成企业没有建立针对外部数据使用的相关制度来保障数据管理义务的落实，近半数企业对合同执行缺乏监督手段。因此，企业期待能有更加完善的技术手段来解决工业数据流通管控的问题。

另外，数据提供方和数据使用方两个角色对可信数据分享和流通关注不同的技术性约束手段。其中，作为数据提供方，企业最期待的是实名身份认证、对数据使用方式的控制、使用过程留有存证、数据用后销毁、采用自动执行的电子合约并有可信的内外部监管机构进行监督，如图2.5所示。

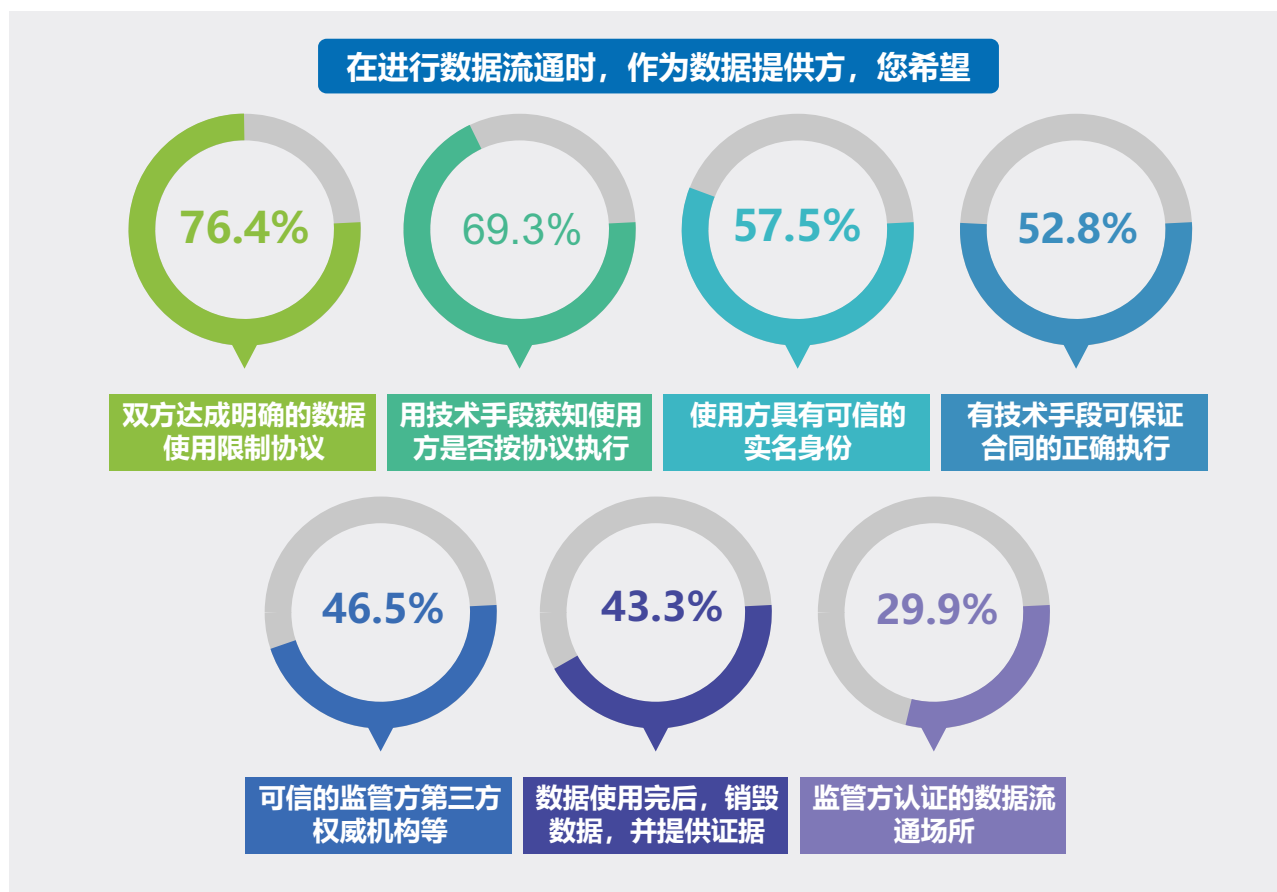


图2.5 数据提供方的期望

作为数据使用方，同样期待数据提供方的身份可信、数据可溯源、数据质量有第三方保障，并期待数据中间机构来提供数据目录，以便更方便地找到想要的数



图2.6 数据使用方的期望

基于上述调研，数据提供方和使用方在可信数据流通方面，主要存在以下三方面的需求：一是对数据使用施加控制，即限制使用范围和方式；二是，对使用进行日志采集存证，以实现溯源、过程存证和有效监管；三是，支撑数据流通的基础服务，如实名身份、供需对接、数据质量、电子合约保障等。



第三章 系统架构

Chapter 3 System Architecture

(一) 概念内涵

基于当前国内外相关政策、法律法规和产业需求，本白皮书提出了一种面向工业数据可信、安全共享和流通的新型基础设施——可信工业数据空间（Trusted Industrial Data Matrix）。

可信工业数据空间是实现工业数据开放共享和可信流通的新型基础设施和技术解决方案，基于“可用不可见、可控可计量”的应用模式，为工业数据要素市场化提供了实现路径。其主要功能有三：一是为数据拥有者

提供数据使用对象、范围、方式的控制能力，满足了企业对工业数据可用不可见、可用不可存、可控可计量的需求，消除流通顾虑；二是为数据处理者提供数据流通处理的日志存证，提供内外部合规记录，实现数据资源有效管理；三是为数据供需双方提供中间服务，便利供需对接，促进工业数据要素资源的价值转换。

本章节将从业务视角、功能视角和技术视角构建可信工业数据空间的系统架构。



（二）系统架构 —— 业务视角

业务视角从数据共享流通各参与方的需求出发，基于各参与方之间的业务关系形成的数据共享流通模式，主要分为三类：点对点模式、星状网络模式以及可信工业数据空间融合模式。

一是点对点模式，该模式在工业数据共享流通场景中最常见。以离散制造行业中的配件生产为例，企业甲将配件图纸及配套数据交付配件厂乙生产配件。在此过程中，**数据提供方**（企业甲）提供图纸数据，**数据使用方**（配件厂乙）需要图纸数据进行配件生产，两家企业内部的存证部门作为**存证方**对数据的使用进行监督（如图3.1所示）。

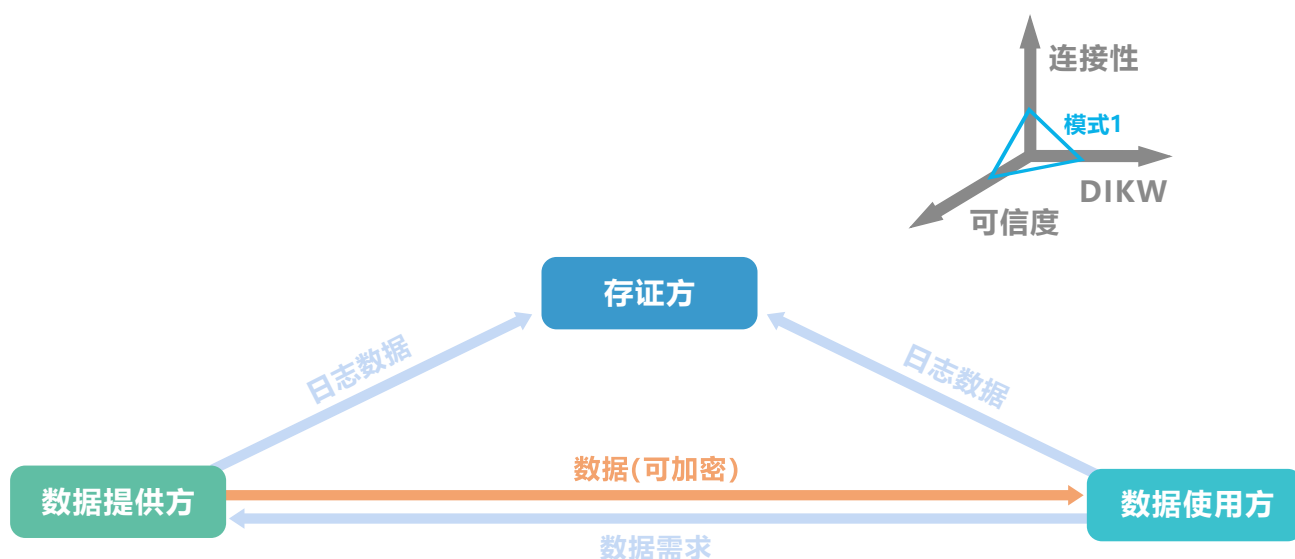
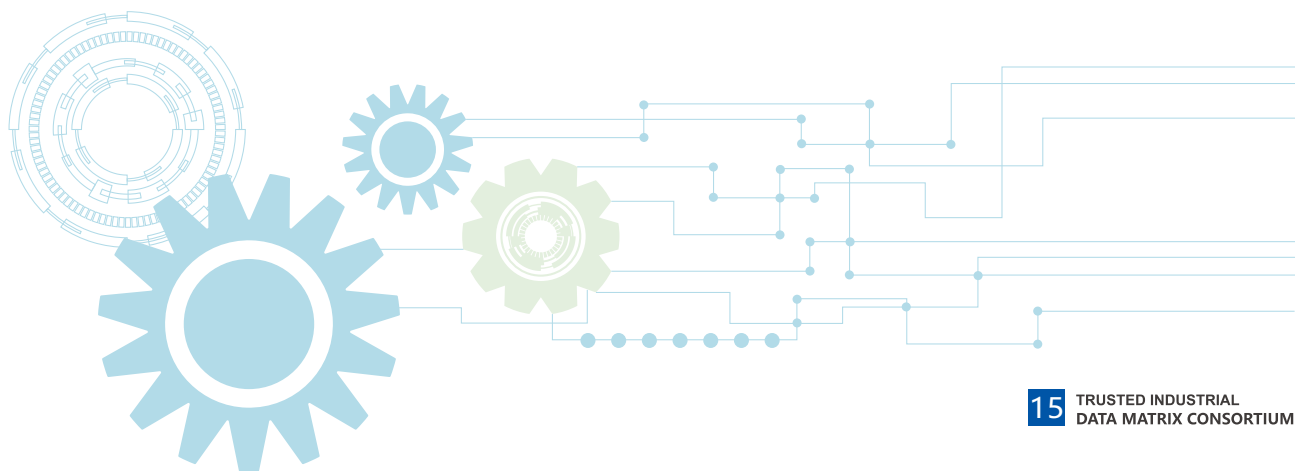


图3.1 模式一：点对点模式



二是星状网络模式。随着数据提供方和使用方数量增多，以及双方对数据的使用形式和深度提出了不同需求，点对点的数据共享流通方式难以满足用户需求，星状网络结构因此逐渐出现（如图3.2所示）。例如：多家零件企业使用多方安全计算，共享零件数据样本，共同训练神经网络模型。除了多方安全计算，数据汇聚、数据沙盒和联邦学习也是星状网络模式中三种常见的数据共享流通方式。星状网络结构使得数据的共享与流通在连接性、可信度以及应用深度上均有提高。

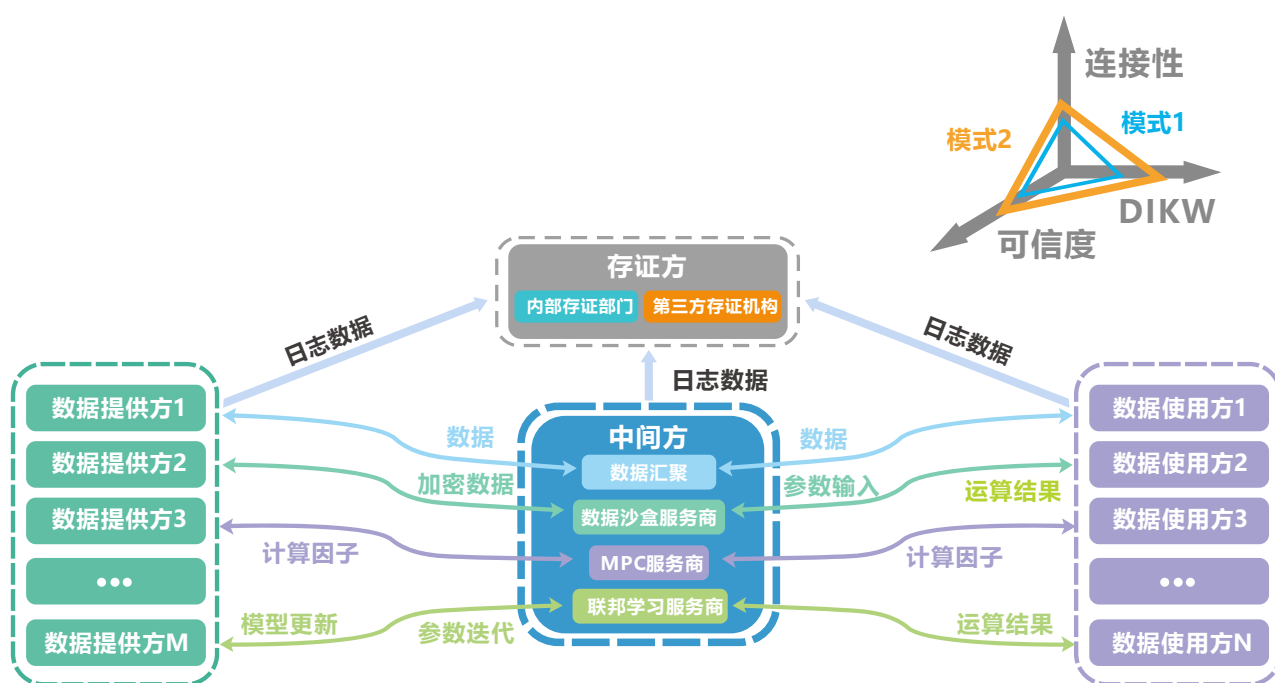


图3.2 模式二：星状网络

三是可信工业数据空间融合模式。主要基于模式一和模式二中各利益相关方对数据使用范围、深度和可信的不同要求，在模式三中，定义了五种主要参与方，包括数据提供方、数据使用方、存证方、中间服务方和IT基础设施提供方。该类模式覆盖的角色和业务流程相对完整，也构成了可信工业数据空间的业务视图（如图3.3所示），后续的功能视角和技术视角将依此视图进行展开。

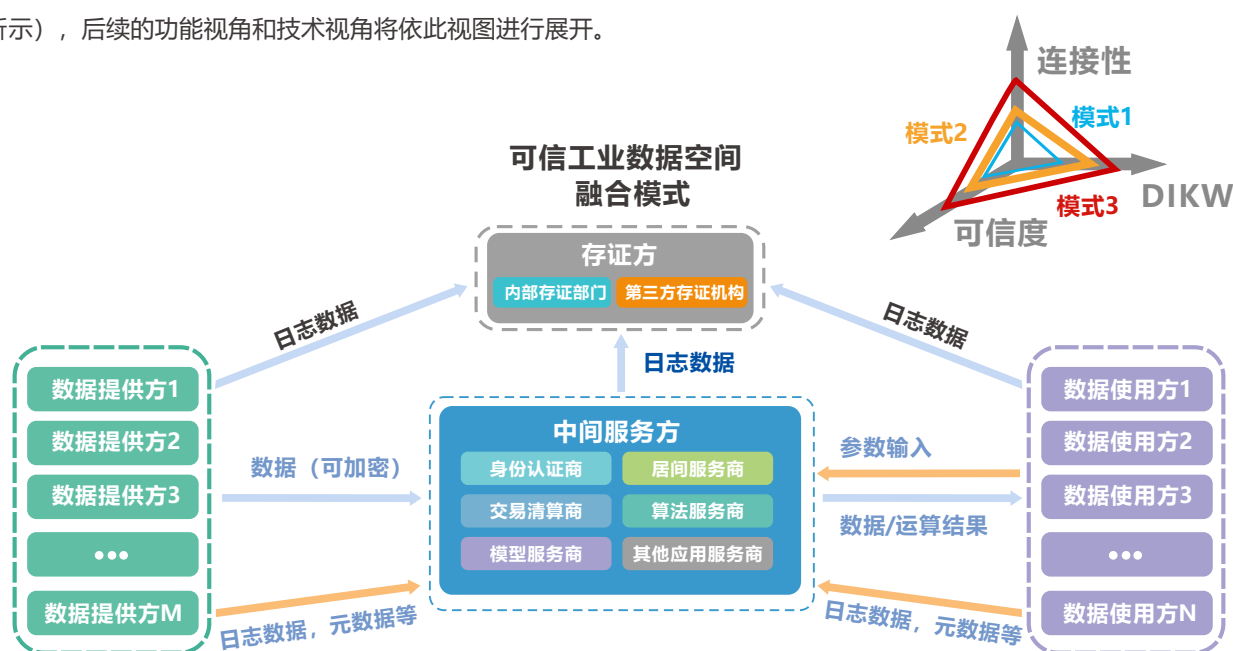


图3.3 模式三：可信工业数据空间—业务视图



数据提供方

数据、加密数据、计算因子以及部分运算结果的提供方。



数据使用方

数据或运算结果的需求方，需求参数的提供方。



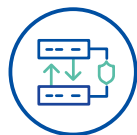
存证方

企业内部的存证部门以及第三方的存证机构。存证日志可为政府审计部门提供支持。



中间服务方

参与方与接入的身份认证、目录推送与资源检索、供需对接、合约达成与执行、交易清算与用后评价等基本功能服务，以及算法、模型等其他应用服务的提供方



IT基础设施提供方

数据传输储存等IT基础设施功能的提供方。

(三) 系统架构 —— 功能视角

根据业务视图中各方主体扮演的角色以及之间的供需关系，可以归纳出一个功能的供需矩阵（如图3.4所示）。例如：纵轴[数据提供方]与横轴[数据使用方]对应的[数据使用控制]单元意为[数据提供方]需要[数据使用方]执行[数据使用控制]。

功能实现 供应需求	数据提供方	数据使用方	存证方	中间服务方	IT基础设施 提供方
数据提供方	N.A	数据使用控制	日志存证 溯源	身份认证 目录推送 服务对接等	数据传输储存 数据处理计算
数据使用方	提供数据	N.A	日志存证 溯源	身份认证 目录搜索 服务对接等	数据传输储存 数据处理计算
存证方	提供日志信息	提供日志信息	N.A	身份认证 提供服务数据	提供日志信息
中间服务方	提供身份信息 提供目录信息	提供身份信息 提供检索信息 提供需求信息	提供身份信息 日志存证 溯源	N.A	提供身份信息 提供传输储存 数据处理计算
IT基础设施 提供方	提供服务信息	提供服务数据	日志存证 溯源	身份认证 提供服务数据	N.A

图3.4 供需矩阵

结合供需矩阵形成可信工业数据空间的功能视图（如图3.5所示），自下而上分别为：数据接入层、传输处理层、中间服务层、数据控制层以及数据应用层。其中蓝色的三层是数据共享流通的基础与载体，黄色的两层实现了数据共享流通的可信与透明，也是可信工业数据空间的核心部分。

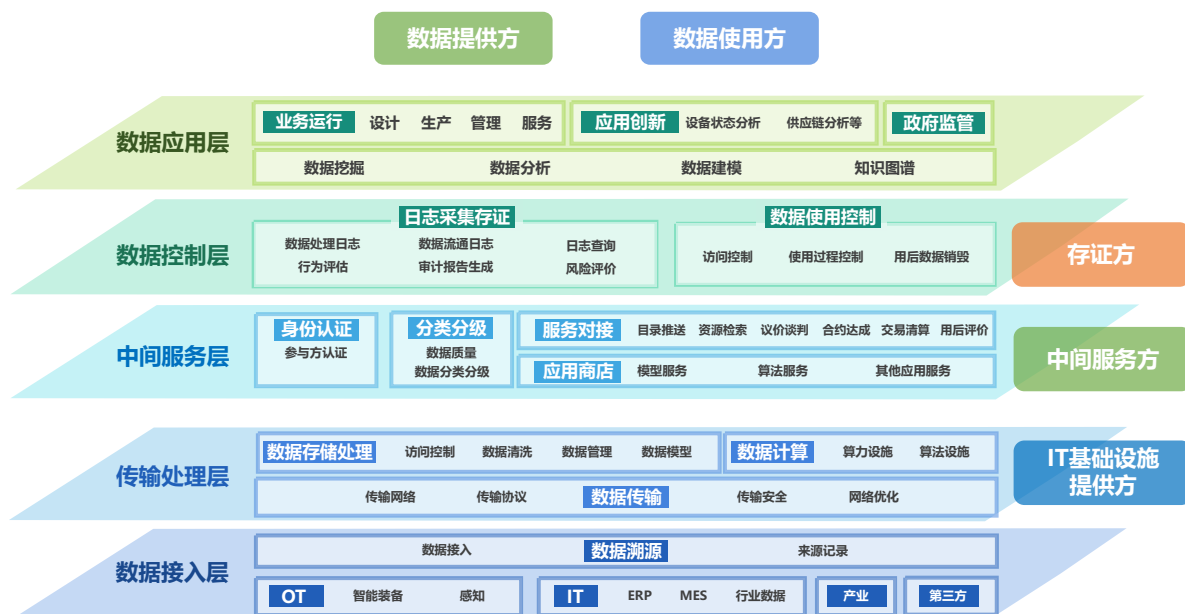


图3.5 可信工业数据空间—功能视图



1.数据接入层

数据接入层是工业数据的来源。数据接入层主要包括OT层的智能装备、感知设备等，IT层的ERP、MES等。产业数据以及第三方数据来源，数据的来源记录和数据溯源也是数据接入层的主要功能。



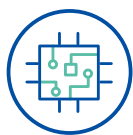
2.传输处理层

传输处理层主要考虑的是对数据的传输、处理以及计算。传输的网络、协议与安全、网络性能优化、数据的访问控制、清洗储存、管理处理以及提供的算力算法服务都是传输处理层的核心组成单元，该层的主要功能由IT基础设施提供方提供。



3.中间服务层

中间服务层主要由中间服务方提供的第三方服务组成。如参与方认证与接入认证、数据的分类分级管理与价值评估、目录推送与资源检索、供需方的议价谈判、电子合约的达成与执行、交易清算与用后评价，以及第三方提供的算法和模型等应用商店服务。



4.数据控制层

数据控制层由两大功能组成。一是为用户提供提供的日志采集存证功能，该功能采集了数据流通与处理日志，并能够进行行为评估、风险评价和审计报告生成。二是数据全生命周期的接入控制与使用控制功能，包括数据的使用过程控制和用后的数据销毁过程。



5.数据应用层

数据应用层主要包括企业业务运行、应用创新相关功能，以及政府的监管应用。例如通过数据挖掘、数据分析、数据建模及知识图谱等方式。

(四) 系统架构——技术视角

为了实现功能视图中的上述功能，可信工业数据空间的实现主要需要以下七大类技术，包括安全技术、隐私计算技术、存证溯源技术、数据控制技术、管理技术、计算处理技术以及OT技术（如图3.6所示）。

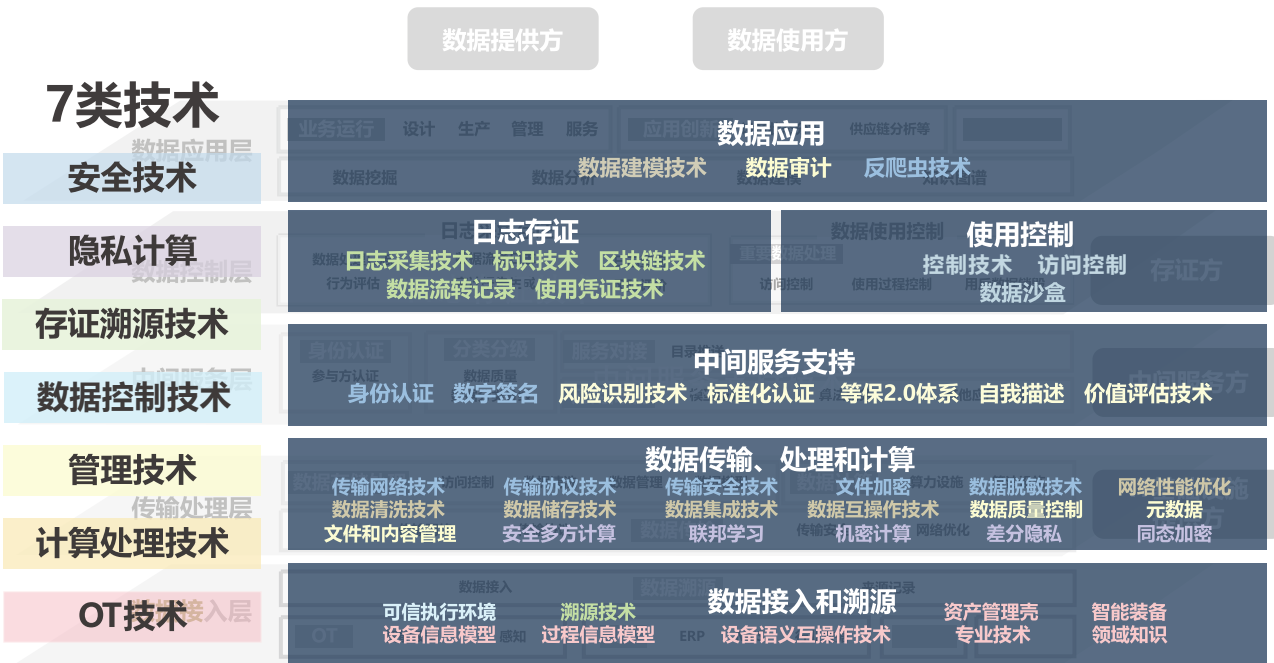


图3.6 可信工业数据空间—技术视图



1.安全技术

安全技术是保障数据安全重要基础。主要包括文件加密、身份认证、数字签名、数据脱敏、反爬虫技术、传输网络、传输协议、传输安全以及可信执行环境等技术。



2.隐私计算技术

隐私计算类技术可以在原始数据不出本地的情况，发挥数据的价值，保护用户的数据隐私。主要包括安全多方计算、联邦学习、机密计算、差分隐私、同态加密等技术。



3.存证溯源技术

存证溯源技术主要负责对数据全生命周期进行日志存证与溯源。主要包括日志采集技术、标识技术、区块链技术、数据流转记录技术、使用凭证技术以及数据溯源等技术。



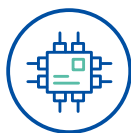
4.数据控制技术

数据控制技术实现了数据提供方对数据全生命周期的掌控，例如数据撤回、使用次数与时间限制、以及用后即焚。使用控制技术主要包括：控制技术、访问控制以及数据沙盒等技术。总体而言数据控制技术是对传统访问控制技术的丰富与革新。



5.管理技术

管理技术主要用于实现中间服务层和计算处理层的功能。管理技术主要包括数据安全审计、风险识别技术、标准化认证、等保2.0体系、自我描述、数据质量控制、元数据技术、文件和内容管理以及价值评估等技术。



6.计算处理技术

计算处理技术负责对数据的清洗、储存、计算与处理提供支持。主要包括网络性能优化、数据清洗技术、数据建模和设计技术、数据存储技术、数据集成技术以及数据互操作等技术。



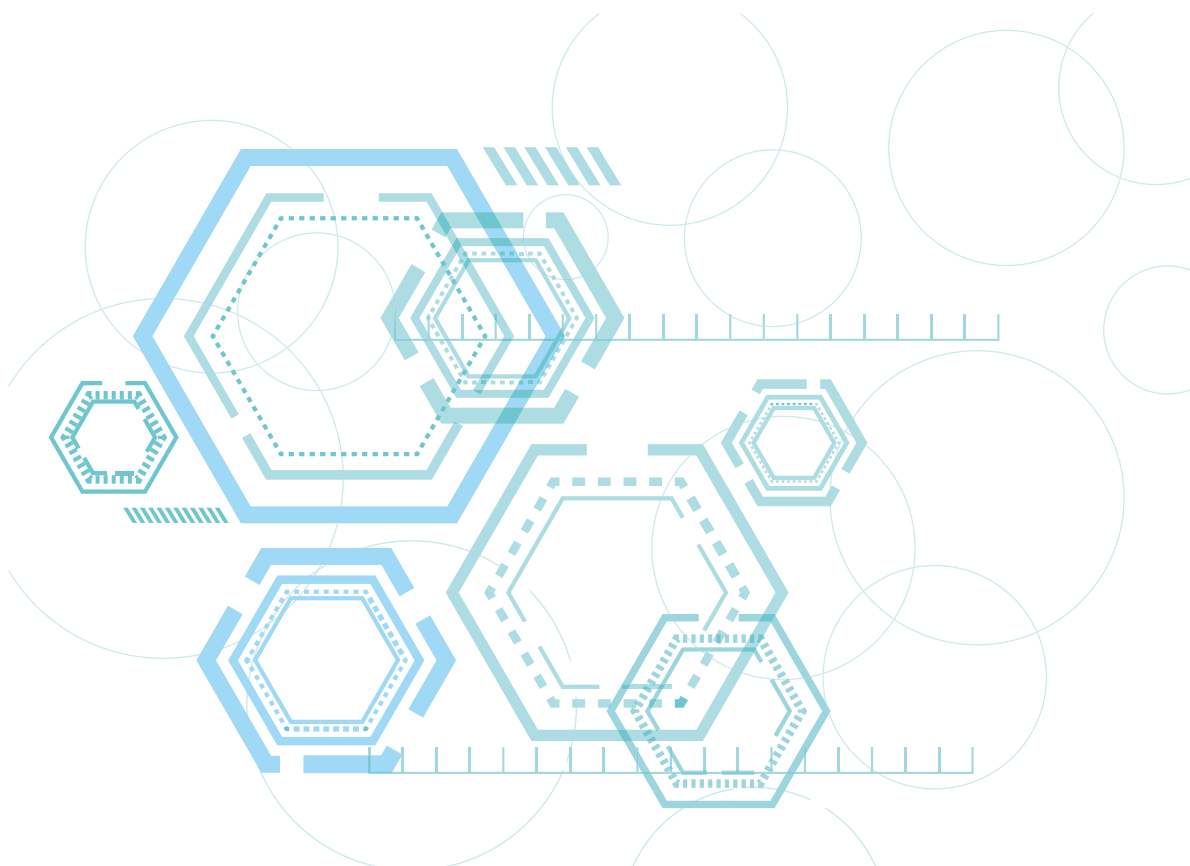
7.OT技术

OT技术为可信工业数据空间架构提供支撑。主要包括资产管理壳、智能装备、信息模型、设备语义互操作技术、相关专业技术及领域知识等技术。

第四章 技术视角

Chapter 4 Technical Perspective

根据可信工业数据空间系统架构-技术视角，可信工业数据空间的功能实现主要需要以下七大类技术，包括安全技术、隐私计算技术、存证溯源技术、数据控制技术、管理技术、计算处理技术以及OT技术。其中安全技术、隐私计算技术、存证溯源技术以及数据控制技术为可信工业数据空间的核心技术，将在下文白皮书中重点讨论。



（一）技术概述

1. 安全技术类

① 数据脱敏

A. 数据脱敏的定义

数据脱敏是指对某些敏感信息通过脱敏规则进行数据的变形，实现敏感隐私数据的可靠保护。目前数据脱敏较常见应用在涉及客户安全数据或者一些商业性敏感数据的情况下，在不违反系统规则条件下，对真实数据进行改造并提供测试使用，如数据安全领域的隐私数据身份证号、手机号、卡号、客户号等个人信息都需要进行数据脱敏。工业领域的数据敏感性场景较为复杂，需要开发适合不同场景的数据脱敏方法，并根据实际场景加以应用。数据脱敏类别可分为结构化数据脱敏和非结构化数据脱敏，这两类数据都普遍存在于工业数据空间中。

B. 数据脱敏的作用

数据脱敏可以将数据漂白，抹去数据中的敏感内容，同时保持原有的数据特征、业务规则和数据关联性。保证开发、测试、培训以及大数据类业务不会受到脱敏的影响，达成脱敏前后的数据一致性和有效性。

C. 对应可信工业数据空间的功能

数据脱敏技术可应用于可信工业数据空间功能视角中的数据应用层、传输处理层以及数据接入层。





② 差分隐私

A.差分隐私定义

差分隐私是对统计数据库的隐私泄露问题提出的一种隐私保护技术，通过加噪声的方式避免原始信息外露，实现了统计意义上的保密。即提供一种当从统计数据库查询时，最大化数据查询的准确性，同时最大限度减少识别其记录的机会，简单来说，就是保留统计学特征的前提下去除个体特征以保护用户隐私。工业领域的数据敏感性场景较为复杂，需要根据实际场景加以应用。

B.差分隐私的作用

可以给出数据使用方总体需要的信息，保留源数据统计学特征。但可以去除个体特征保障数据提供方敏感及隐私数据的安全。

C.对应可信工业数据空间功能

可应用于可信工业数据空间数据应用层和传输处理层。

2. 隐私计算技术

1 可信执行环境

A. 可信执行环境的定义

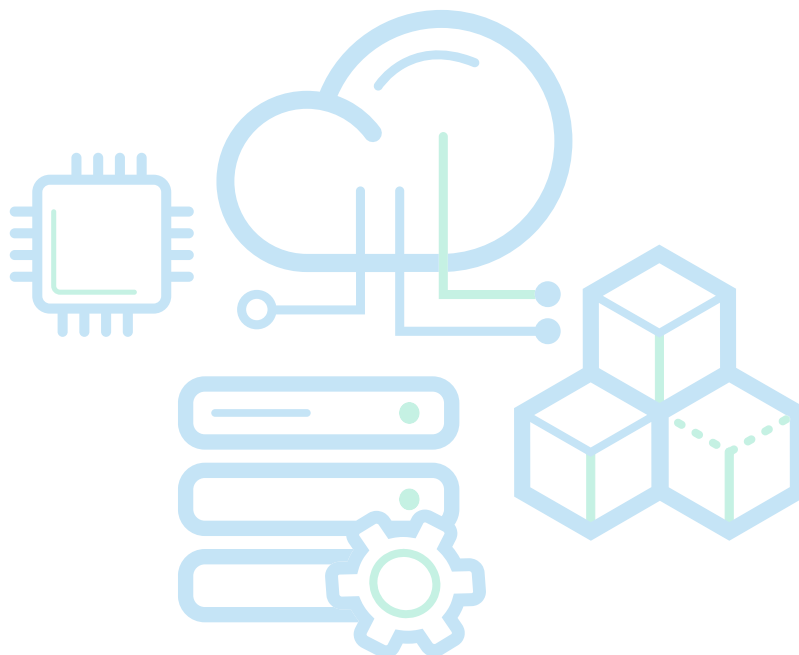
可信执行环境（以下简称TEE）是主处理器内的安全区域。它运行在一个独立的环境中且与操作系统并行运行。这个并行系统相对于传统系统更加安全，它确保加载的代码和数据的机密性和完整性都得到保护。在TEE中运行的受信任应用程序可以访问设备主处理器和内存的全部功能，而硬件隔离保护这些组件不受主操作系统中运行的用户安装应用程序的影响。TEE通过加密和隔离保护不同的受信任应用程序。

B. 可信执行环境的作用

可信执行环境在中央处理器上构建一块受保护的计算执行空间与Rich OS 隔离，保证在该环境下所加载的数据和执行的程序的安全性、完整性。在可信执行环境内部，应用运行也是相互独立的，不能无授权互访问。

C. 对应可信工业数据空间功能

可信执行环境可应用于可信工业数据空间数据控制层的控制功能模块，为数据访问控制和使用控制提供基础可信环境。





② 安全多方计算

A.安全多方计算定义

安全多方计算问题是在一个分布式网络上计算基于任何输入的函数，每个输入方在这个分布式网络上都拥有一个输入，而这个分布网络确保输入的独立性，计算的正确性，而且除了各自的输入外，不透露任何可用于推导其他方输入和输出的信息。主要有混淆电路、同态加密、秘密分享三种形式。安全多方计算主要面向的是在多个参与方的环境下，每一个参与方都拥有自己的私密信息，同时又希望利用其它的信息来共同完成计算一个函数的过程。

B.安全多方计算的作用

安全多方计算提供了一种基于密码学的解决方案，重点解决工业数据在计算过程中的隐私安全问题，在不泄露原始数据的前提下，实现多方协同计算，使得参与方无法得到除计算结果之外的其他信息，保障工业数据的安全共享、流通、计算和交易。

C.对应可信工业数据空间功能

可应用于可信工业数据空间的数据应用层、数据控制层和传输处理层。

③ 联邦学习

A. 联邦学习定义

联邦学习是一个机器学习框架。在多参与方或多计算结点之间开展高效率的机器学习。联邦学习做到各个参与方的自有数据不出本地，而后通过加密机制下的参数交换方式，在保障数据隐私的情况下，建立一个虚拟的共有模型。这个虚拟模型相当于聚合在一起建立的最优模型。在建立模型的过程中，各个参与者的身份和地位相同，而联邦系统帮助大家建立了安全共享的策略。另外根据原始数据的分布规律，常用的联邦学习模式主要有横向联邦学习、纵向联邦学习和迁移联邦学习。

B. 联邦学习的作用

能够保证参与各方在保持独立性的情况下，进行信息与模型参数的加密交换，并同时让各自的模型获得成长。

C. 对应可信工业数据空间的功能

可应用于可信工业数据空间架构数据应用层、数据控制层和传输处理层。



3.存证溯源技术

① 区块链存证

A.区块链存证定义

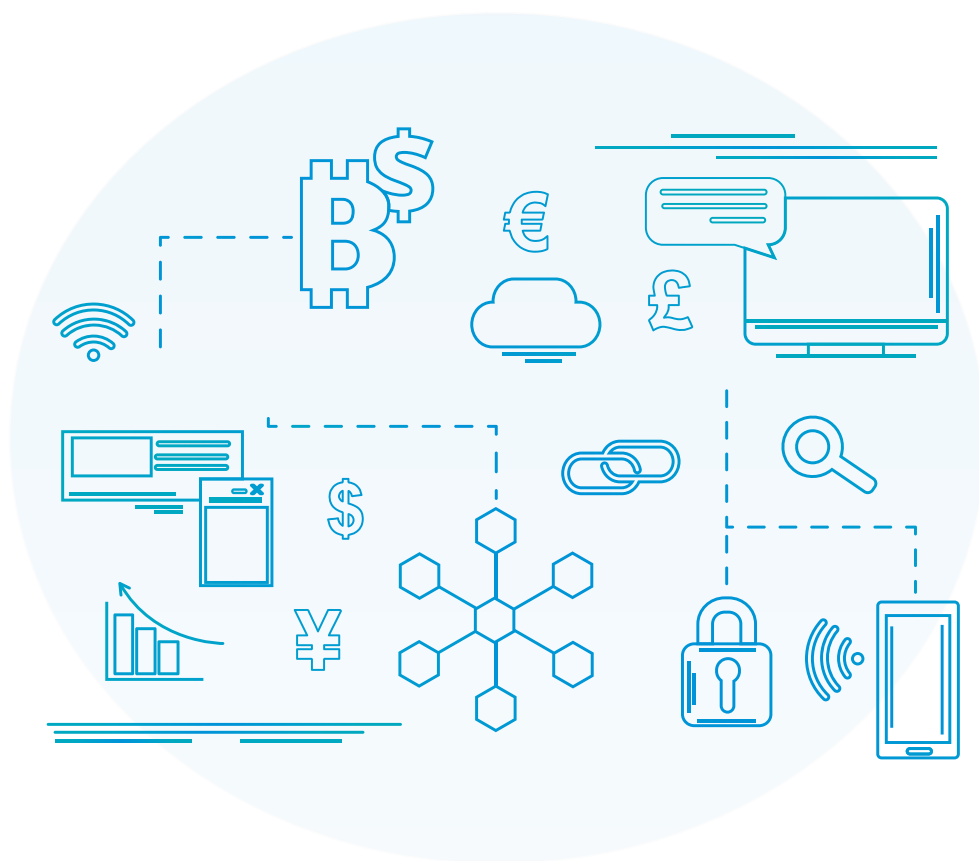
区块链数据存证，就是把数据标识（哈希值）存到区块链上，达到防篡改、可追溯、数据来源可信的目的。一般情况下，采用链上链下协同工作，采用文件与哈希值分离的方式，链上只保存文件的哈希值，原文件保存在链下。只要计算出文件的哈希值，与链上的哈希值比对，就知道文件是否被篡改。

B.区块链存证的作用

区块链存证最大作用就是防篡改。尤其针对电子合同，区块链存证功能加强了电子合同存证期间的不可篡改性。

C.对应可信工业数据空间功能

区块链存证可用于可信工业数据空间数据控制层和中间服务层。





② 一体化智能合约技术

A. 一体化智能合约技术的定义

将传统合约数字化，并以信息化方式传播、验证或执行的计算机协议。与传统合约相比，智能合约不依赖第三方执行合约，让合约验证和执行过程更加便捷。合约允许多方用户共同参与制定一份智能合约，通过P2P网络扩散并存入区块链，并自动执行，从而能够保证合约条款的安全可靠，让双方进行可信交易，并且这些交易可追踪且不可逆转。在工业数据交易过程中，通过智能合约技术能够缩短交易流程。

B. 一体化智能合约技术的作用

智能合约不需要第三方执行，并且可以自动执行。另外智能合约可以实现自动化交易托管，减少交易流程中耗费的时间成本。

C. 对应可信工业数据空间功能

可用于可信工业数据空间数据控制层和中间服务层。



4.数据控制技术

① 访问控制技术

A.访问控制定义

对具有价值的信息资源的访问都应该通过认证。用户在进行访问之前，通常会被要求提供正确的认证。最常见的认证方法是提供用户名和密码的组合。然而用户名和密码是静态认证方式，很容易被破解。对于具有高价值的信息，应采用强度更高的认证方法。访问控制通过组合多种认证要素可以有效的提高认证强度。按用户身份及其所归属的某项定义组来限制用户对某些信息项的访问，或限制对某些控制功能的使用的一种技术。

B.访问控制的作用

通过对身份认证和确认，实现控制只允许哪些人可以访问数据和哪些人不可以访问数据。

C.对应可信工业数据空间功能

可应用在可信工业数据空间数据控制层。

② 使用控制技术

A.使用控制技术定义

访问控制技术仅仅是在某个指令执行前发挥作用，一旦操作完该指令，访问控制便再也不会对数据有任何作用。使用控制是将数据控制权限始终保持在数据提供方这里。使用控制基于访问控制对身份的确认并将数据使用控制延伸到数据使用方，例如控制数据在使用方使用一定时长或一定次数后自动执行删除。通过使用控制技术将数据控制权始终控制在数据提供方手中。

B.使用控制技术的作用

在执行交换关键和敏感数据交易中，使用控制技术强制执行数据提供方加载的数据限制，始终把源数据控制权把握数据提供方手中。使用控制技术通过将数据控制权限、知情权限和拒绝权限始终保持在数据提供方手里，打消了数据提供方数据流通利用的顾虑。

C.对应可信工业数据空间功能

可应用在可信工业数据空间数据控制层。



第五章 标准体系

Chapter 5 Standard System

为贯彻落实《中华人民共和国国民经济和社会发展的第十四个五年规划和2035年远景目标纲要》和《“十四五”国家信息化规划》中关于加快建立数据交易流通和安全保护等标准规范的部署要求，加强国家可信工业数据空间标准化工作顶层设计，指导可信工业数据空间标准研制工作有序推进，满足技术进步和工业数据可信共享、流通、交易发展的需要，制定了可信工业数据空间标准体系。



可信工业数据空间标准体系结构包括“A基础共性”、“B关键技术”、“C行业应用”等3个部分，主要反映标准体系各部分的组成关系。可信工业数据空间标准体系结构图如图5.1所示。

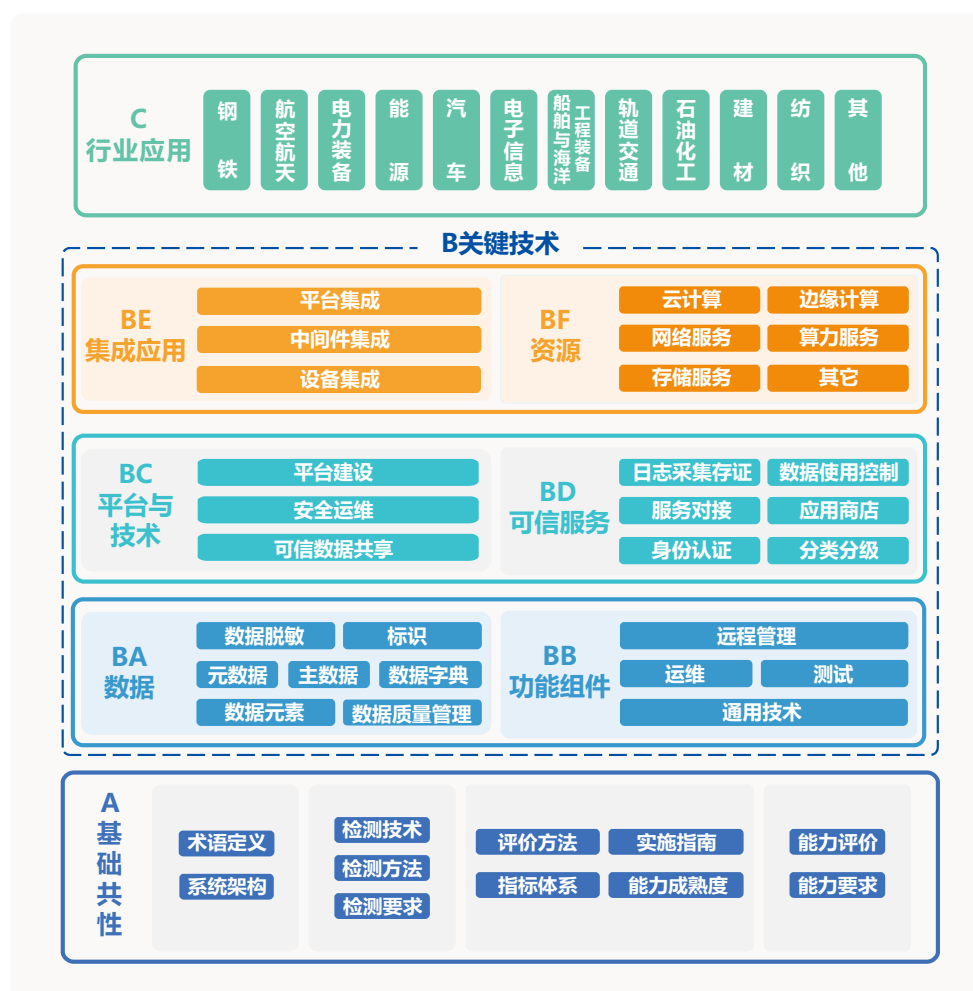


图5.1 可信工业数据空间标准体系结构图

具体而言，A基础共性标准包括通用、检测、评价、人员能力等4大类，位于可信工业数据空间标准体系结构图的最底层，是B关键技术标准和C行业应用标准的支撑。B关键技术标准是根据可信工业数据空间系统架构功能视角映射而形成的，BA数据标准主要聚焦于功能视角各层对数据的要求，包括数据元素、数据质量管理、元数据、主数据、数据字典、数据脱敏和标识等相关标准；BB功能组件标准主要聚焦于在各利益相关方部署的分布式功能组件标准化要求，包括通用技术、检测、运维和远程管理等相关标准；BC平台与技术标准主要聚焦于中间服务层和数据控制层对应的平台要求，包括平台通用技术要求、建设规范、流程规范、运维管理和接口规范等相关标准；BD可信服务标准主要聚焦于中间服务层和数据控制层提供的各类可信服务，包括身份认证、数据分类分级、需求交互、交易撮合、应用商店、日志采集存证和数据使用控制等相关标准；BE集成应用标准主要聚焦于可信工业数据空间可集成的各类平台、中间件和设备要求，包括设备集成、中间件集成和平台集成等相关标准；BF资源标准主要聚焦于可信工业数据空间所需的各类资源要求，包括云计算、边缘计算、网络服务、算力服务、存储服务等相关标准。C行业应用标准位于可信工业数据空间标准体系结构图的最顶层，面向行业具体需求，对A基础共性标准和B关键技术标准进行细化和落地，指导各行业推进可信工业数据空间应用。



根据可信工业数据空间当前的产业应用现状，根据“共性先立，急用先行”的原则，拟优先制定以下重点标准：

(一) A基础共性标准



(二) B关键技术标准

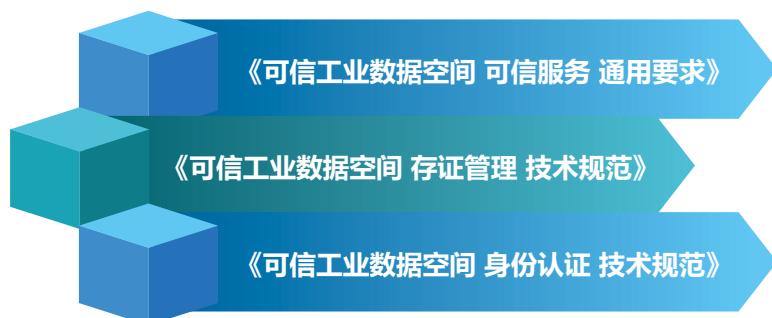
1、BB 功能组件：



2、BC平台和技术：



3、BD可信服务



第六章 产业案例分析

Chapter 6 Industry Case Analysis

当前，我国工业数据的共享与流通仍处于探索阶段，形成了多样化的切入点和路径。本白皮书选择了若干代表性案例，他们或是采用了新的架构和技术理念，或是基于通用技术实现了数据共享流通中的某个功能，旨在为理解可信工业数据空间应用价值和实践路径提供参考。



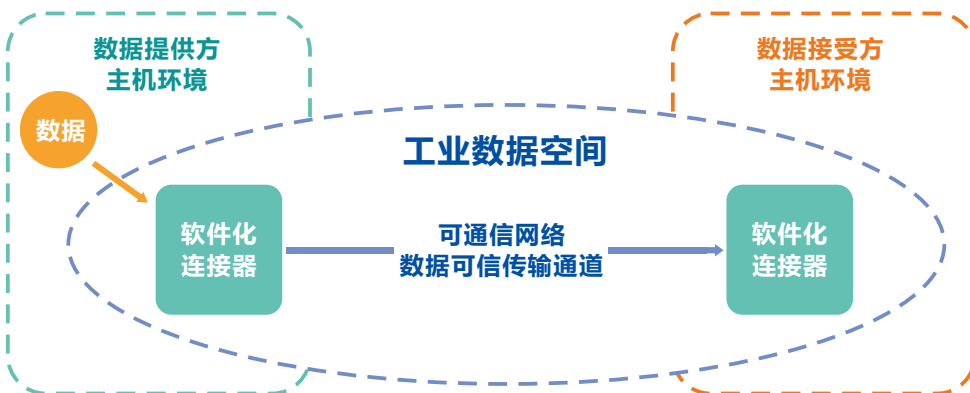
（一）东方电气—针对工业3D模型数据的可信共享流通

1.现状需求

针对装备研发，CAD模型数据的安全可信传输是一个关键难题。通常，CAD数据解密出库后,就进入了技术上不受控的阶段，为了保障安全，目前经常采用U盘进行物理传输，由专人进行U盘的配送监督与数据的用后删除，亟需开发一个既满足保密需求又能供提高研发效率的技术方法。



2.解决方案



中国信息通信研究院、东方电气和中国电信联合开发了可信流通测试床，实现了工业CAD模型数据在可通信系统之间的安全可信传输。实现了以下三个目标：

- 1 数据提供方可控制数据不被发送给第三方；
- 2 数据提供方可根据需求在一定时间撤回数据；
- 3 数据提供方可以监控数据被谁调用、被谁存根等。

3.实施效果

测试床提供了异地企业之间工程模型数据可信传输的一种可行性方案，在保障数据可信使用的前提下，替代人力解决数据不受控的问题，简化传输流程，降低工业数据跨企业传输的成本。



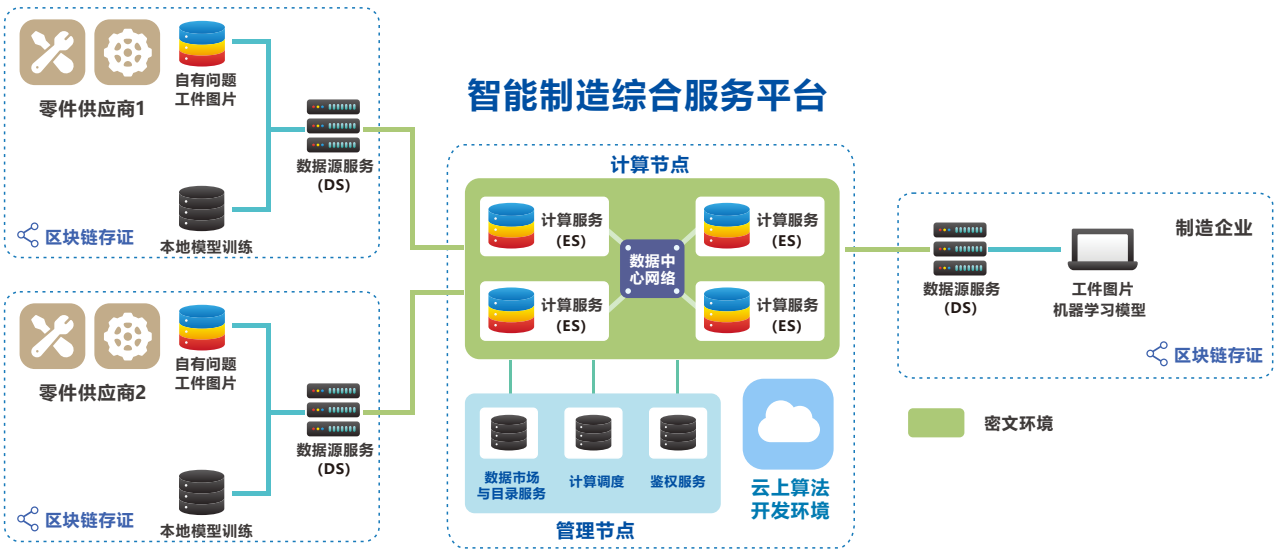
(二) 华控清交—工业自动化质检过程中的多方数据联合训练

1.现状需求

在离散型工业制造供应链上，通常由多个零件生产商为下游企业供应同一规格零件，零件的批量较大，一般采用人工抽样检测的方式来进行工件质检。这样会造成两个问题：一是随机抽样方式不覆盖所有工件；二是检测完全依赖检验员的业务经验和工作态度，质检效果

波动大、效率低。生产商一方面需要基于其他生产商全量样本数据进行模型训练，但另一方面又不希望将零件数据本身的信息透露给其他生产商。亟需建立一种原始数据不出本地、基于跨企业数据共享的分析挖掘方式。

2.解决方案



华控清交帮助零件生产商企业建立的能够实现以上需求的数据共享流通平台。该平台采用基于多方安全计算（MPC）的数据流通架构，利用服务平台功能，将质检员在流水线每个环节采集到的问题工件图片，通过安全多方计算进行共享。计算节点使用共享数据集进行机器学习联合训练，生成并使用问题工件预测模型，给企业进行全量自动化质检。平台主要实现了多方安全计算、区块链存证、数据传输以及供需对接功能。

3.实施效果

通过集成以上功能，该平台实现了数据可用不可见、数据使用的可控可计量、模块化与易开发，在保障数据安全的前提下，提高了计算性能与数据的价值。参与实施的两个零件供应商A和B各自提供带有划痕和缺陷的工件图片，通过数据共享，进行联合模型训练，经测试大幅提升了模型准确率。企业通过服务平台获得最终的结果模型，并利用模型进行自动化工件质检，减少了成本、优化了检测效果并提高了检测效率。

1.现状需求

已经失去了风险阻断的时效性。处于产业生态圈中的企业，迫切希望能够建立一种依赖于程序自实现的对数据主权保护的数据交换平台。

The diagram illustrates the Data Flow Model (DFM) with the following components and flows:

- Stakeholders:** 数据提供者 (Data Provider), 数据所有者 (Data Owner), 数据消费者 (Data Consumer), 数据使用者 (Data User), 服务提供商 (Service Provider), 中间代理商 (Intermediate Agent), 清算中心 (Clearing Center), 应用商店 (App Store), 应用开发者 (App Developer), and 词汇提供者 (Vocabulary Provider).
- Data Flows (Green Arrows):**
 - 数据所有者 → 数据提供者 (授权 - Authorization)
 - 数据提供者 → 服务提供商 (数据传递 - Data Transfer)
 - 服务提供商 → 数据消费者 (接收数据 - Receive Data)
 - 数据提供者 → 中间代理商 (发布元数据 - Publish Metadata)
 - 中间代理商 → 数据消费者 (搜索元数据 - Search Metadata)
 - 数据提供者 → 数据消费者 (数据传递 - Data Transfer)
- Metadata Flows (Orange Arrows):**
 - 数据提供者 → 清算中心 (日志数据 - Log Data)
 - 数据提供者 → 应用商店 (日志数据 - Log Data)
 - 数据提供者 → 应用开发者 (提供词汇库 - Provide Vocabulary)
- Software Flows (Blue Arrows):**
 - 应用开发者 → 应用商店 (发布应用程序 - Publish Application)
 - 应用商店 → 数据消费者 (使用应用程序 - Use Application)
 - 数据提供者 → 应用商店 (使用应用程序 - Use Application)

Legend:

- 数据流 (Data Flow) - Green Arrow
- 元数据流 (Metadata Flow) - Orange Arrow
- 软件流 (Software Flow) - Blue Arrow

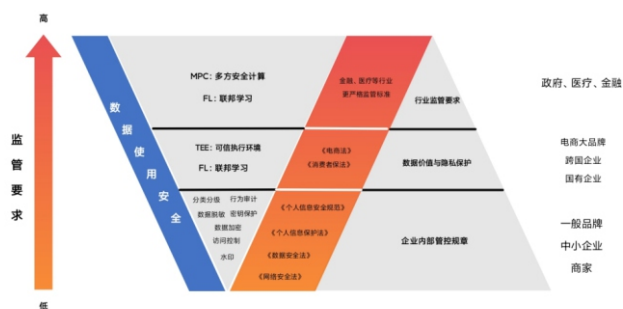
3.实施效果

目前，华为已经实现了点对点的数据共享交换过程实例，为参与方企业提供了一对一的数据共享流通平台，企业可以通过该平台进行依赖程序自实现的跨企业数据可信流通。

(四) 阿里云—解决数据隐私安全的DataTrust隐私增强计算平台

1.现状需求

目前，企业在数据流通领域主要面临五个问题。标签丰富度欠缺，准确度无法验证；用户洞察维度较少，无法有效赋能运营与内容推广活动；新标签沉淀较慢，数字化建设需要时间；数据输出建模担心隐私与数据安全问题；缺乏用户标签，影响数据算法结果。针对其中的隐私与数据安全问题，阿里云通过多种安全技术的组合，研发DataTrust平台，完成了数据隐私与共享需求的平衡。



零售



品牌方通过联动平台、第三方等全域数据，在保护个体隐私及数据安全的前提下，构建品牌数智化运营能力，优化人货场的配置，拉动业务增长。

政务



通过隐私增强计算加速各部门数据的互联互通，提高协同效率，并赋能公共事业：实现交通数据，水电燃气数据互联网数据的融合利用，提高城市的公共管理，公共服务的整体水平。

金融



满足行业或企事业单位数据不出自身环境下，通过可用不可见技术，实现与多方数据的联合风控，提高风控识别有效性，助力业务良性增长。

医疗



满足行业数据不出自身环境的要求，通过可用不可见技术，实现多个医疗机构样本的联合建模，解决样本量少的问题，帮助医生提升诊断的准确率，实现智能诊断。

2.解决方案

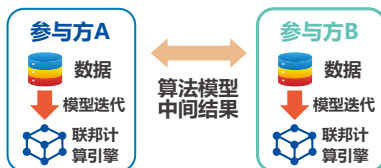
DataTrust技术内核解析，多种安全技术组合

可信执行环境TEE



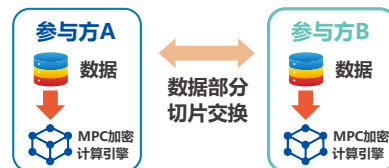
- 基于SGX2.0，数据加密出用户环境
- 性能高、提供联合分析SQL功能
- 通常用于对复杂条件下的数据分析工作

可联邦学习Federated Learning

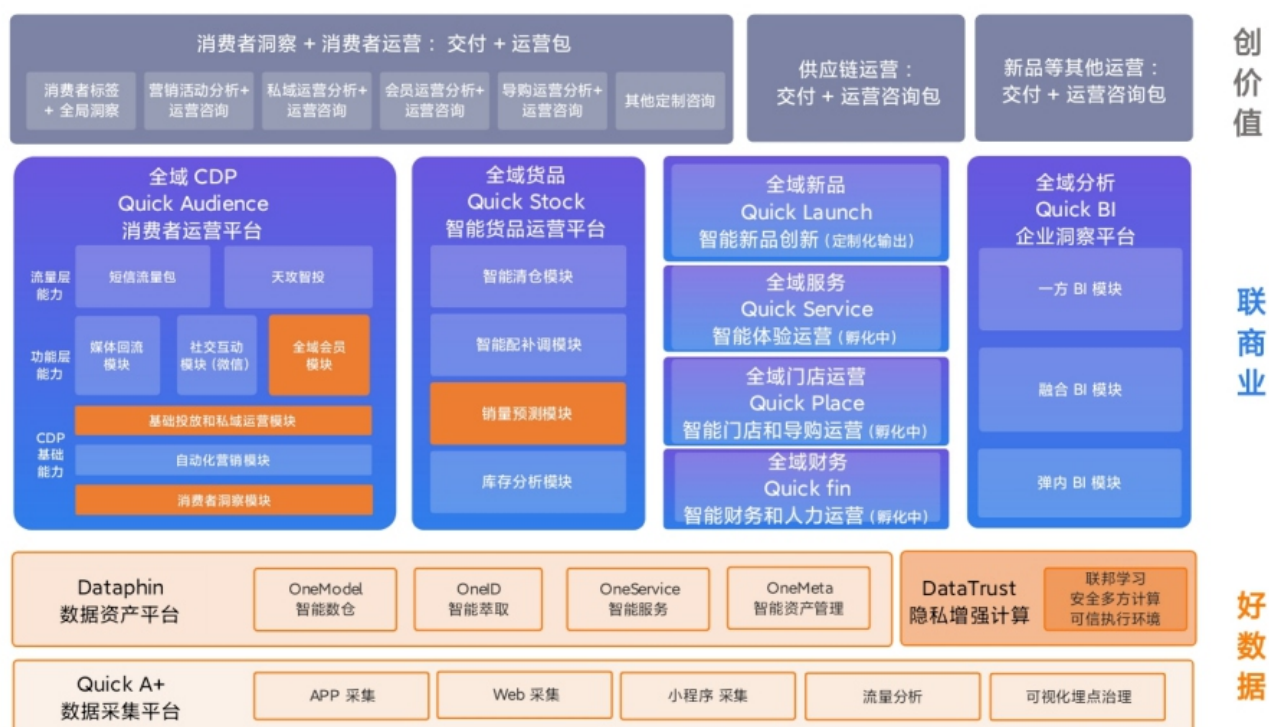


- 原始数据不出用户环境
- 支持十亿级别的联合预测
- 主要用于联合建模、联合预测通常用于搜索推荐、营销场景

可多方安全计算MPC



- 原始数据不出用户环境
- 支持十亿级别的ID匹配
- 多用于联邦学校的样本对齐，企业间用户重合度分析



DataTrust平台主要包括三大功能模块：可信执行环境、联邦学习与多方安全计算。可信执行环境主要用于复杂条件下的数据分析工作。联邦学习和安全多方计算结合，共同实现原始数据不出本地的数据共享流通。此外，平台还实现了数据接入、帐号权限管理、存证与审计等功能服务。

3.实施效果

DataTrust平台通过帮助企业建立数据精细运营方案，实现了跨企业间的数据可信流通。首先，平台能够安全接入各品牌的自有数据，使用多种隐私计算技术，在原始数据不出本地的情况下进行联合训练，实现了数据计算结果的共享流通，在满足了各企业对数据的需求同时，也解决了企业对数据隐私的担忧。

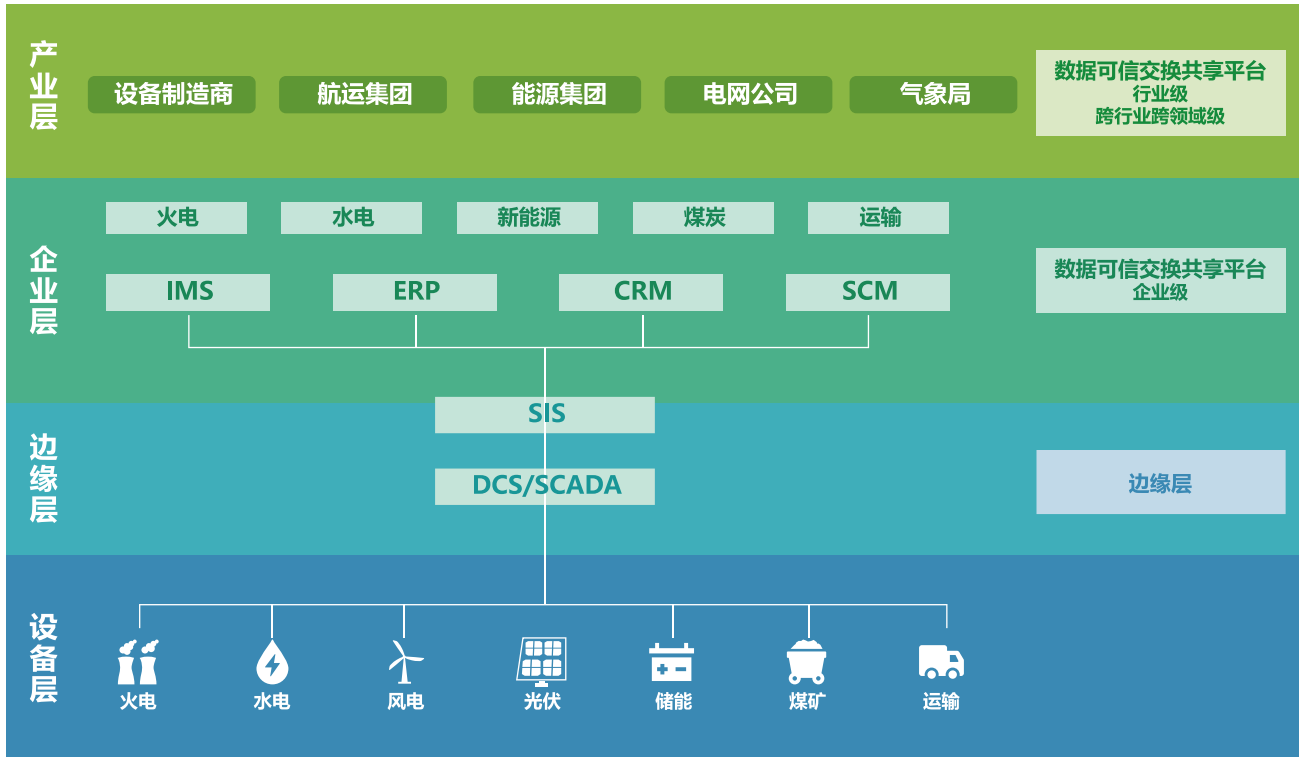
(五) 国能信控—能源行业数据可信交换共享案例

1.现状需求

国家能源集团业务涵盖煤炭、火电、新能源、水电、运输、化工等板块，涉及业务种类复杂。在集团数字化转型需求的背景下，企业面临业务协同效率低的问题。国能信控着眼于集团不同层级企业以及外部数据之间的复杂数据流通共享交换场景，搭建企业内、企业间的工业数据安全可信交换共享平台，实现了数据驱动的多业务协同应用。



2.解决方案



一方面，电力、煤炭、运输等不同业务的电厂/场站作为数据提供方，向数据平台提供工业实时生产及业务数据；外部数据源将交易数据、天气数据、企业数据等提供给数据平台；区域公司和集团则提供企业计划经营数据；数据平台作为中间服务商对数据进行汇聚接入，并分类分级储存和管理。

另一方面，数据可信交换共享平台也扮演着中间服务方的角色，为各层次企业提供电力发售协同、协同调运、多能互补、共享储能等数据应用服务，以实现各层级、各板块企业之间业务协同，以及煤炭、运输及电厂上下游之间的供应链协同，实现价值最大化。

3.实施效果

结合能源集团具体业务，平台主要应用在智慧燃料调运、电力协同营销以及供热/固弃物优化管理场景中，实现了集团内部的数据流通的可信与安全。

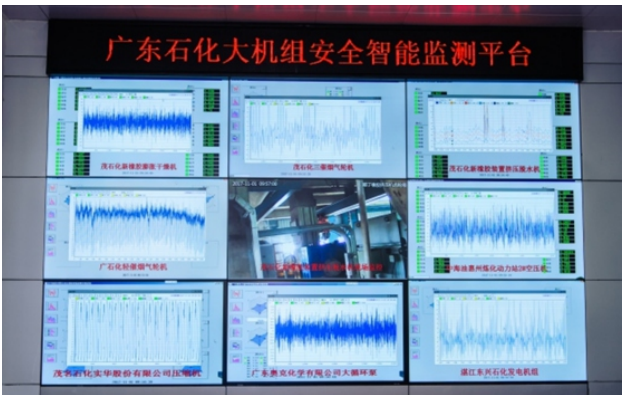
- ① 智慧燃料调运通过与煤炭、运输等行业数据互通，把握煤价、运费、港口库存等变化趋势，形成优选调运计划，为采购决策提供量化数据支撑。
- ② 电力协同营销通过与外部客户的工商信息、风险信息互通，形成用户画像，支持企业决策分析的应用，最终达到对业务闭环、决策支持和区域协同营销的要求。
- ③ 与煤炭市场行情、周边客户及大基建、天气情况数据互通，同步供热及固弃物客户信息等，为子分公司供热/固弃物优化管理提供决策数据支撑。



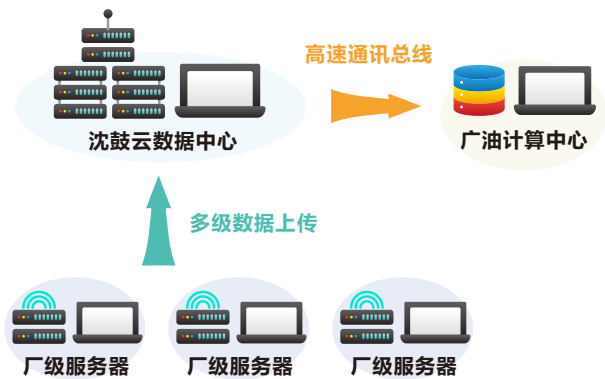
(六) 沈鼓测控—基于石化大机组数据的可信安全集成

1.现状需求

据统计，在石油化工工业生产过程中，89%的非计划停机是由随机故障造成的，无法通过定期维护保养预防。通过汇聚各机组的数据，对全量数据进行模型分析，能够进行早期故障筛查及预警。但由于石化数据具有敏感性，亟需在企业内部构建一个可信流通的环境，保障在不同厂区的石化大机组间实现数据安全的集成共享。



2.解决方案

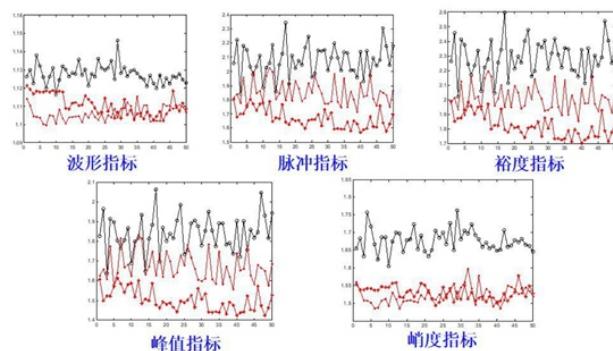


沈鼓对石化大机组的多源信息进行集成汇聚，建立了厂间数据安全流通平台，主要实现了以下功能：

- ① 基于多源异构的信息整定和互无量纲指标的特征精细提取技术的数据接入与语义互操作功能；
- ② 数据存储与处理功能；
- ③ 数据安全传输功能。针对机组数量多、数据量大的场景，采用了基于数据缓冲队列的接口方式，使得数据传输吞吐量、稳定可靠。

3.实施效果

序号	项目	2018年	2019年	2020年
1	服务的用户总数	127	165	193
2	服务的机组总数	663	811	970
3	在线消除故障的机组台数	27	35	50
4	立即停机避免更大损失的机组台数	12	21	31



2018-2020年，平台安全汇集了不同厂间的石化机组数据，以全量数据进行模型训练，在线消除故障的机组台数累计112次，避免了用户非计划停机，累计为生产企业增加效益约44800万元。



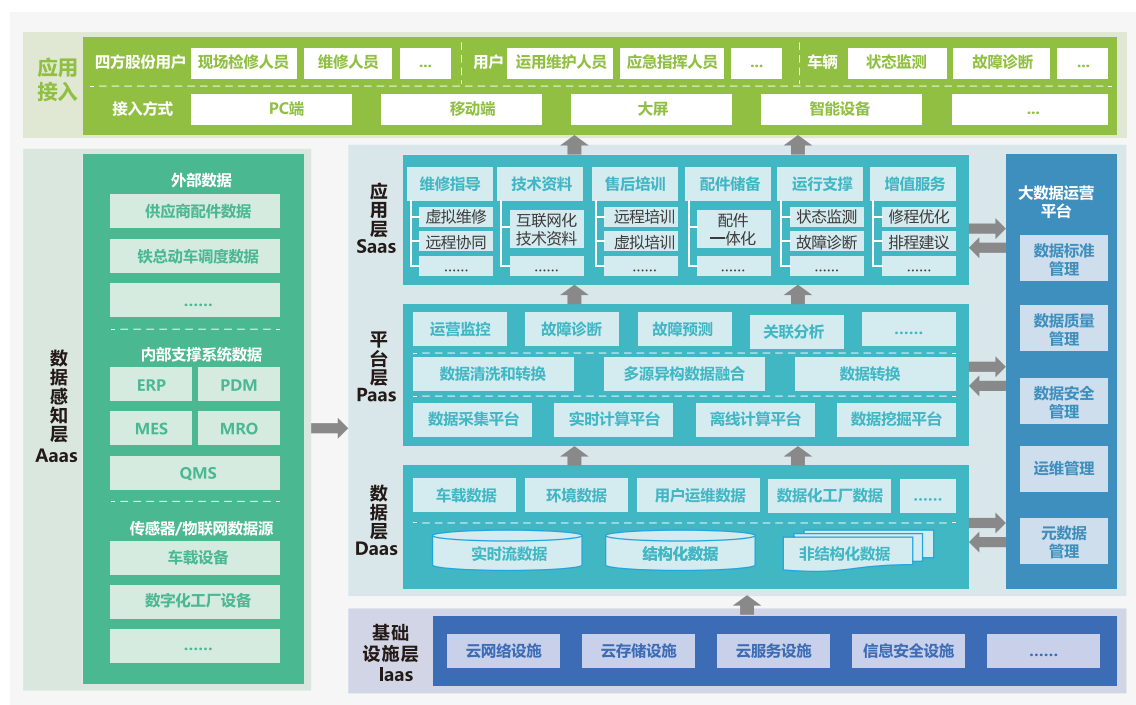
(七) 中车四方—动车产业链上下游间的数据共享流通

1.现状需求

轨道交通装备全球化、网络化格局对运营安全保障提出严峻挑战：列车数量多、运量大、分布地域广、服役环境复杂等综合特征，对运维体系的集成化、协同化、时效性提出更高要求。亟需建立一套上下游产业链间的数据共享流通平台，汇集上游配件商数据，下游已售列车运行数据，构建高品质、高效率的运维服务新模式。



2.解决方案



中车四方基于该架构形成平台+应用理念，主要实现数据接入、多源异构数据融合、数据加密传输、数据清洗与储存、数据计算处理、访问控制、日志采集存证以及相应管理条列。通过技术和管理手段的结合，解决了上下游产业链间的数据需求与数据隐私的平衡。

3.实施效果

中车四方平台为轨道交通装备安全、可靠运营提供了保证。平台实现了主厂数据、供应商数据以及已售列车运行数据等数据的可信共享流通，达成了协同化生产与交通运输装备的实时维护。



设备利用率OEE

车轴单件加工时间分析

月度异常问题统计

关键尺寸与工序能力CPK分析

一次交检合格率



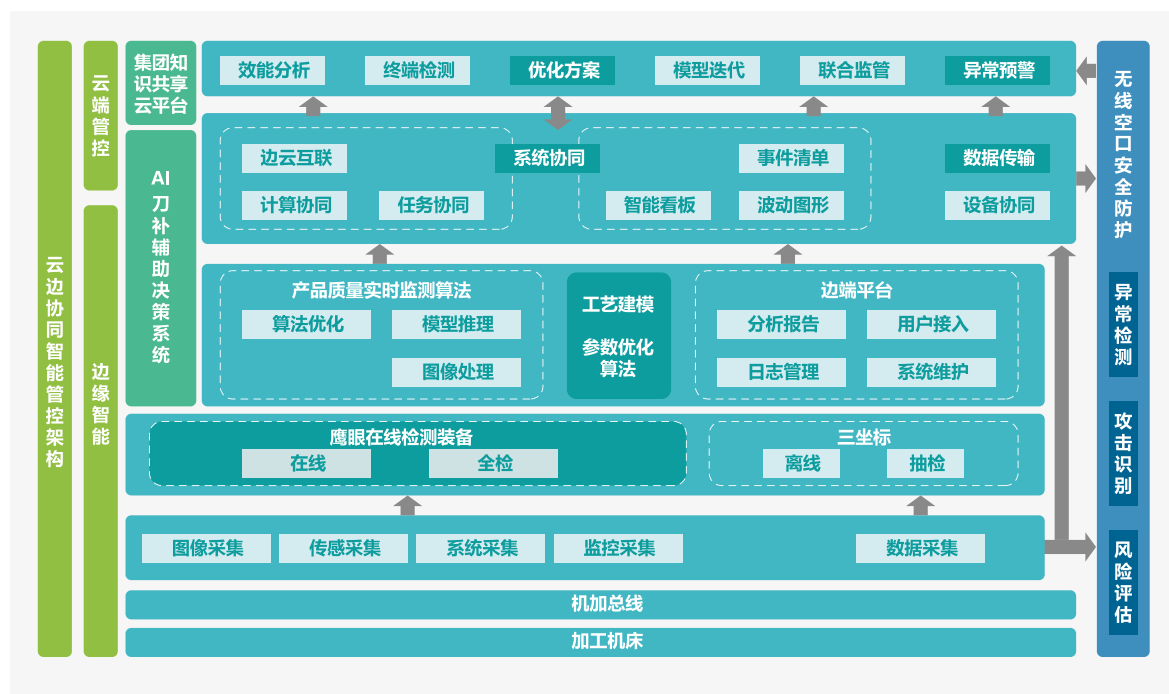
(八) 中信戴卡—基于联邦学习的数据共享与联合训练

1.现状需求

轮毂生产机加工序中，技术人员根据毛坯基本状态和质检结果进行停机工艺调整。由于信息没有实现互联互通，调机效率较低，各子公司之间的工艺水平也无法保持一致，公司内部经常发生的重复性工作。亟需建立一套实现公司内部信息透明的数据共享方案。



2.解决方案



中信戴卡与精诺数据联合攻关，建立了以数据+模型双驱动的数据共享方案。平台采集接入机床本身数据以及检测装备数据，然后将样本数据传输至云端进行联邦学习，实现了多条产线、各子公司间的联合训练与模型共享，以此完成集团铝合金车轮生产的自动化。

3.实施效果

在原始数据不出本地的情况下，该方案实现了企业间数据的可信流通。不仅实现检测人员以及调刀工艺人员减少80%的目标，更重要的是实现了刀具调整知识的量化，沉淀知识库。解决传统轮毂制造中存在的刚性生产、经验式工艺、人工离线检测、高能耗和集团模型难以共享、行业知识难以沉淀等难以克服的问题。





工业互联网产业联盟
Alliance of Industrial Internet

📍 北京市海淀区花园北路52号

☎ 010-62305887

🌐 <http://www.aii-alliance.org>