



郑曦 男，1974年10月出生，高级工程师。现任中国电子数据产业有限公司副总经理，曾任中国电子信息产业集团有限公司规划科技部主任助理及数字办主任助理，中国电子信息产业集团有限公司第六研究所生产与资产运营部主任，北京国际系统控制有限公司总经理助理等职务。长期从事信息化系统、行业数字化和大数据等相关工作，承担多项国家和省部级项目，荣获多次省部级奖励，发表多篇技术成果。

加快数据基础设施建设 释放数据要素创新活力

编者按：为推动数据产业健康快速发展，促进数据产业生态良性循环，近日，由中国经济体制改革研究会、中国电子信息产业集团有限公司（以下简称中国电子）、郑州市人民政府、中国经济改革研究基金会联合主办的中国“数据要素×”生态大会在郑州国际会展中心成功召开。《网络安全与数据治理》期刊作为特邀媒体，对中国电子数据产业有限公司副总经理郑曦进行专访，就数据要素治理、数据基础设施建设等各方关注的议题进行解读。

问题一：数据元件的核心理念在产品体系中是怎么体现的？

郑曦：首先，作为连接数据供需两端的“中间态”，数据元件实现原始数据与数据应用“解耦”，破解了“安全与流通对立”的难题。数据元件在技术上更贴近源端，有效消除了数据应用的针对性，更利于促进数据流通。因此，数据元件为打通数据资源侧、数据供给侧和数据应用侧提供了技术路径。

其次，数据元件在数据要素流通过程中实现了安全性、价值属性和品质属性的统一。原始数据通过数据“解耦”实现风险隔离的同时，被加工提纯为数据元件，使数据价值密度和数据品质得到大幅提升，利于数据低成本、高效率的流通，为数据资产化和数据市场化流通交易提供了有利条件。

最后，对数据要素的全生命周期管理势在必行。在产品体系中，数据要素加工交易中心负责实现数据元件开发生产和交易，并围绕数据元件构建全流程智能化监管平台，实现“数据元件可控可计量，数据元件可信可溯源”，推动数据要素高效、集约、安全生产和规模化流通。数据金柜是数据金库中的存算单元，承担关键数据的安全存储和数据元件的存储计算任务。安全可信数据空间通过数据元件实现高效率、多主体在线供需对接、多对多数据安全融合共享。智能政务一体化平台的数据引擎对接数据元件，实现数据组织的高效利用与安全保障。

中国电子数据产业有限公司将继续坚持理论与研究

产品研发双轨并行，将数字平台与数据平台在数据元件理论基础上深度结合，不断推动数据产业做强做优做大。

问题二：请介绍一下系列产品目前的落地实践情况。

郑曦：中国电子数据安全与数据要素化工程理论研究与实践始于2020年，已经分别在德阳、大理、徐州、温州、郑州等省市落地试点。其中，德阳市试点项目是中国电子数据安全与数据要素化工程实践最早的试点示范工程。作为“全国城市数据治理工程”首批试点城市，德阳的数据安全与数据要素化工程实践形成“1+4+N”制度体系，涵盖主体、设施、数据、市场四方面制度设计，交付数据要素加工交易中心，初步构建起数据要素三级市场，目前上架涉及电力、能源、金融等领域的数据元件超过1600类，开始全面释放数据要素价值。

国家数据局提出实施“数据要素×”行动，其中重要的一项就是要加快推进数据基础设施的建设工作。中国电子先行先试，聚焦“关键数据过于分散、安全保障不足”等问题，自主研发的“数据金库”可广泛应用于核心数据、重要数据、敏感数据和数据元件的存储计算和互联互通，目前已经成为数据资源存储和数据要素流通的重要自主安全数据基础设施。徐州市数据金库（试点）作为中国电子在全国落地建设的首个数据金库，目前已顺利通过验收委员会专家验收，一致认为该项目统筹解决数据安全和数据要素化协同发展难题，在全国具有较强的示范引领作用。目前郑州、温州等地的数据金库建设工作正在紧锣密鼓的展开。

中国电子数据安全与数据要素化工程进入构建“全

程全网”数据要素流通和市场配置体系的高阶阶段，郑州、温州、徐州、德阳、大理州五地“数据要素互联网”组网运行今天正式启动。作为数据基础设施的建设的的重要组成部分，全国首个基于“数据元件”的数据要素网的建成标志着数据元件规模化流通和产品化交易的图景开始变为现实，数据要素乘数效应即将释放。同时，安全可信数据空间、数据分类分级管理系统、智能政务一体化平台正在相关部委、重点行业和地区交付和试运行中。

问题三：数据要素流通面临哪些挑战？面向政府的数据要素服务，数据要素加工交易中心的技术优势是什么，如何落地？

郑曦：目前数据要素流通存在主要四个方面的难题：一是数据产权方面，数据的主体多元交织，造成确权难；数据的权属关系复杂交错，造成析权难。二是流通交易方面，缺少标准化的交易标的物，缺少稳定安全的交易机制，缺少高效可信的交易平台，难以规模化高效流通。三是收益分配方面，受制于产权制度不清晰、交易标的明确，导致实践过程中数据价值评估难、数据定价难。四是安全治理方面，数据流通交易的过程安全问题日益严峻，数据安全和数据要素是一体两面，安全不解决，数据无法流通。

当前，面向政府的数据要素服务仍存在供需不顺畅、标准不统一、数据孤岛多、共享应用难等问题。数据元件可有效破解政务数据难题，实现数据要素的高效配置，让数据“供需两端”真正贯通，有效应对这种开放、复杂、多变的供需结构。基于数据元件打造的数据要素加工交易中心是由“一套操作系统、五大支撑体系、五大业务平台和五类数据工具”等构成，是大规模、全流程、自动化的数据元件加工生产流水线，可对数据要素化流程以及平台软硬件资源、数据资源进行调度管理，实现全生命周期的数据要素开发与管控。

数据要素操作系统是整个工程系统的核心，向上可支撑“场景域”等政务数据创新应用，向下可配置数据金库和算力中心等基础设施资源，适配数据归集、数据处理、元件开发、元件维护、元件交易等各类数据治理的相关工具，并根据清洗处理、资源管理、元件开发、元件交易、监管等管理平台的指令进行统一任务调度和流程管控，实现政务数据要素服务的自主安全。

总之，数据要素加工交易中心可有效提升政府履职能力，增强数字政府治理效能，为推进国家治理体系和治理能力现代化提供有力支撑。同时，数据资源要成为数据资产，需要形成标准化产品，并能够在市场中形成流通和交易。通过搭建数据要素加工交易中心，利于数据的确权、定价和大规模流通，可打通数据要素的上下

游产业链，逐步构建数据要素产业生态，对于实现政府资产的价值增值具有现实意义。

问题四：安全可信数据空间产品可以应用在哪些领域或场景，可以解决哪些问题？怎么样保障数据在数据空间中使用是安全合规的？

郑曦：长期以来，行业数据被分割在各个市场主体内部，行业数据壁垒阻碍了数据价值的最大化开发，政府机构及大型企业内部同样面临着“数据孤岛”问题，而数据跨主体共享流通又面临着安全、合规、信任等重重难题，数据存量优势难以转化为高质量发展优势。

为推动行业及机构企业数据“活起来、动起来、用起来”，我们创新打造基于数据元件的安全可信数据空间。安全可信数据空间以数据金库和数据元件为基础，结合数据智能识别能力、数据安全防护能力和数据安全监测能力，保障多主体数据安全融合共享，助力金融、能源、航空、双碳等领域解决空间内数据规模化加工和融合问题，解决行业数据的安全、可信和高效流通的问题，促进数据空间的数据价值化长效运营。以法律法规及行业规范为依据，内置行业场景规则及空间运行规范，通过规章制度策略化、使用规范合约化、合约履行智能化，支持字段级定向场景细粒度的合规管控与安全控制。安全可信数据空间还提供全链条的流通溯源控制、全生命周期的过程安全管控和全栈式国产化基础设施技术支撑，系统化解决空间内安全与合规问题。

问题五：数据分类分级对于保障数据安全流通具有重要意义，请问你们的分类分级产品目前支持哪些行业？具有哪些优势？

郑曦：金融、政务、能源、医疗、教育等行业对于数据的安全性、隐私保护和合规性有着极高的要求，数据分类分级管理系统针对行业特点和需求，提供了定制化的解决方案。通过深入了解各行业的业务场景和数据特性，构建并内置了多个行业知识库，确保分类分级的准确性和适用性，很好地满足了行业需求。

在分类分级的精准度方面，系统适配数十种常见数据库和主动元数据管理，支持 AI 算法进行智能分类分级和自动化更新运维，通过持续的学习和优化，系统的分类分级准确性不断提高。在数据安全方面，系统采取了多重防护措施。除了数据脱敏加密、身份验证、功能权限控制、安全审计等常规手段外，还提供了文档隐写水印功能和数据权限控制，以确保数据在传输和存储过程中的安全。在操作易用性方面，支持用户以思维导图或知识图谱形式查看分类法分级统计成果，做到心中有数，用户无需具备深厚的技术背景，也能快速上手。

(下转第 18 页)

- 息保护法》第 27 条第 1 分句 [J]. 法律科学 (西北政法大学学报), 2023, 41 (3): 62-75.
- [10] 李振林, 潘鑫媛. 生成式人工智能背景下数据安全的刑法保护困境与应对——以 ChatGPT 为视角的展开 [J]. 犯罪研究, 2023 (2): 25-33.
- [11] 全国信息安全标准化技术委员会大数据安全标准特别工作组. 人工智能安全标准化白皮书 (2023 版) [R]. 2023: 5-7.
- [12] 张凌寒. 生成式人工智能的法律定位与分层治理 [J]. 现代法学, 2023, 45 (4): 126-141.
- [13] 张惠彬, 肖启贤. 人工智能时代文本与数据挖掘的版权豁免规则建构 [J]. 科技与法律, 2021 (6): 74-84.
- [14] STROWEL A. ChatGPT and generative AI tools: theft of intellectual labor? [J]. IIC-International Review of Intellectual Property and Competition Law, 2023, 54 (4): 491-494.
- [9] 刘练军. 个人信息与个人数据辨析 [J]. 求索, 2022 (5): 151-159.
- [15] KELLER D. Comments on the guidelines on transparency under regulation 2016/679 [J]. 2018.
- [16] CHARI B R, MCKINNEY S A, GREENSTEIN G R, et al. Dos and don'ts for developing, extending and using generative AI models [J]. Computer and Internet Lawyer, 2023, 40 (7): 3

- 10.
- [17] HACKER P. A legal framework for AI training data - from first principles to the Artificial Intelligence act [J]. Law, Innovation and technology, 2021, 13 (2): 257-301.
- [18] CARLINI N, HAYES J, NASR M, et al. Extracting training data from diffusion models [J]. arXiv preprint arXiv: 2301.13188, 2023.
- [19] HELBERGER N, DIAKOPOULOS N. ChatGPT and the AI Act [J]. Internet Policy Review, 2023, 12 (1).
- [20] 郭春镇. 生成式 AI 的融贯性法律治理——以生成式预训练模型 (GPT) 为例 [J]. 现代法学, 2023, 45 (3): 88-107.

(收稿日期: 2023-12-14)

作者简介:

邢露元 (1998-), 女, 硕士研究生, 主要研究方向: 国际经济法、数字法学。

沈心怡 (2000-), 女, 硕士研究生, 主要研究方向: 信息科技法、公司金融法。

王嘉怡 (2002-), 女, 本科生, 主要研究方向: 数字法学、计算法学。

(上接第 2 页)

只有更好地了解数据的价值和重要性, 才能制定出更有针对性的策略。系统为用户提供了全面的数据资产管理功能, 可以对数据进行标签化、评级、评估、封装、整合等操作, 集灵活性、智能化、安全性于一身, 以满足行业客户在数据安全流通领域的各类亟需。

问题六: 大模型等人工智能技术方兴未艾, 您认为将对数据要素治理的产生哪些启示和影响?

郑曦: 数据治理与大模型技术的发展密不可分, 目前不断生成、优化、迭代符合中国国情的大模型技术的任务非常紧迫。大模型技术和自身的优化迭代要依托高质量的数据集和语料库, 目前数据主要集中在政府、行业企业或者头部互联网平台企业手里, 如何让政府、企

业或平台能将、敢将数据共享出来, 供我国的大模型技术使用是一个基础性问题。而数据要素与数据安全治理工程恰恰可以提供解决方案, 通过数据元件的中间态, 保障数据安全、提高数据价值、提升数据品质, 为人工智能的训练和优化迭代提供支撑, 为大模型演化提供高质量的数据集和语料库。通过数据元件可将不同行业的数据融合提供给大模型进行训练, 通过优秀行业的人工智能应用促进整个数字经济和智慧社会的发展。数据元件的中间态不仅解决了数据质量问题, 同时也解决了多场景数据融合的问题, 是数据治理未来发展的一个重要方向。

(本刊记者: 范赫男)

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com