

创新中国从『互联网+』挺进『数据要素×』

本报记者 张保淑

近日,《“数据要素×”三年行动计划(2024—2026年)》(以下简称《行动计划》)正式发布。《行动计划》由国家数据局会同中央网信办、科技部、工业和信息化部等共17个部门联合印发,这是国家数据局揭牌两个多月以来会同相关部门联合印发的首个文件。该文件从2023年12月15日向社会征求意见到当月29日相关媒体吹风会举行,再到2024年1月4日印发实施,历时仅20天,可谓迅速、高效。

《行动计划》提出一个引人注目的表述——“数据

要素×”,大力部署实施“数据要素×”行动并列出重点推进的12个领域,要求充分利用中国海量数据资源、丰富应用场景等优势,以数据流引领技术流、人才流、物资流等,实现知识扩散、价值倍增,催生新产业、新模式,为推动高质量发展、推进中国式现代化提供有力支撑。《行动计划》明确提出到2026年底数据要素应用总体目标,其中包括打造300个以上示范性、显示度高、带动性广的典型应用场景,数据产业年均增速超过20%。



工业级流体仿真大模型“东方·御风”助力中国C919大型客机早日翱翔蓝天。图为2023年12月9日,中国东方航空接收的第三架C919大型客机从上海浦东国际机场起飞。 新华社发

实现科创效能倍增

科学技术是第一生产力,是先进生产力的集中体现和主要标志;创新是引领发展的第一动力,是推动生产力发展的关键因素。在《行动计划》中,“科技创新”被列为“数据要素×”应用的12个重点领域之一。由此可见,数据要素被赋予实现科技创新效能倍增,推动社会生产力更快发展的重大使命。

《行动计划》鲜明地提出以“数据驱动”科技创新,要求“充分依托各类数据库与知识库,推进跨学科、跨领域协同创新,以数据驱动发现新规律,创造新知识,加速科学研究范式变革”。

在中国工程院院士王坚看来,以“数据要素×”实现科技创新效能倍增,与当今时代“数据驱动型创新”特点高度契合。王坚把20世纪形成的传统创新模式称为“电气化驱动创新”。正是在电气化的强有力推动下,与激光、互联网、计算机等有关的科技发明不断涌现,塑造了人们的生产生活方式和社会面貌,可以说20世纪是电气化驱动科技创新的世纪。进入21世纪,以美国数据科学专家、图灵奖得主詹姆斯·格雷为代表的科学家提出“数据密集型科学发现”,凸显数据对科技创新的重大价值。

那么数据是如何驱动科技创新的呢?对此,王坚解释,科学家可以根据相关数据,提出新的假设,并借助高性能计算技术、人工智能技术等,对假设进行验证,从而获得新的科学发现。

与王坚的上述看法类似,中国科学院院士、北京科学智能研究院院长鄂维南认为,在当今大数据时代,科研环境从依靠图书馆、电子显微镜等传统基础设施转变为以大数据、人工智能为代表的新型基础设施。

对《行动计划》提到的“加速科学研究范式变革”,中国科学院院士、复旦大学校长金力表示,传统的科研范式经历了4个阶段演化:从通过实验描述自然现象的“实验范式”,到通过模型或归纳进行研究的“理论范式”,再到应用计算机仿真模拟解决学科问题的“计算范式”,再到通过大数据分析研究事物内在关系的“数据范式”。今天,随着数据资源不断增长和需要求解的科学问题日益复杂,科学研究的第五范式应运而生,就是在“数据范式”基础上,引入了智能技术,强调人的决策机制与数据分析的融合,将数据科学和计算智能有效地结合起来,利用AI技术学习、模拟、预测和优化自然界和人类社会的各种现象和规律,以解决各种科研问题,从而推动科学发现和创新,这就是科学研究“人工智能驱动范式”(AI For Science)。

2022年底以来,以ChatGPT为代表的生成式人工智能大模型大放异彩。科学研究“人工智能驱动范式”迫切需要类似ChatGPT的人工智能大模型提供支撑。为此,《行动计划》提出打造服务于中国科技创新事业的人工智能大模型并作出相关部署,要求“以科学数据支持大模型开发,深入挖掘各类科学数据和科技文献,通过细粒度知识抽取和多来源知识融合,构建科学知识库底座,建设高质量语料库和基础科学数据集,支持开展人工智能大模型开发和训练”。

对科学研究“人工智能驱动范式”的前景,金力进行了乐观展望,认为这一范式是实现科技原始创新的利器,比如,在大数据支撑下,人工智能在生命科学领域可以高效、精准地预测蛋白质结构。无独有偶,在近日举行的一个论坛上,国家数据局局长刘烈宏在演讲中也举了这个例子。他说,长期以来,如何准确、快速确定蛋白质三维空间结构一直是一个难题,而基于海量、多元生物数据构建起的人工智能算法模型,能在几天甚至几分钟预测出以前要花费数十年才能得到的、具有置信度的蛋白质结构,颠覆了生命科学领域的研究范式。

科学数据重在应用,《行动计划》指明了科学数据两大应用方向:一是聚焦科研本身,以科学数据助力前沿研究,面向基础学科,提供高质量科学数据资源与知识服务,驱动科学创新发现;二是聚焦服务产业,以科学数据支撑技术创新,聚焦生物育种、新材料创制、药物研发等领域,以数智融合加速技术创新和产业升级。

广泛赋能千行百业

由《行动计划》,人们很自然地联想到“互联网+”,联想到2015年7月印发的《关于积极推进“互联网+”行动的指导意见》。两个文件虽然时隔8年多,但是仍有诸多可以比较的地方。两者分别以“互联网+”和“数据要素×”为关键词,都引入了数学符号,给人非常直观的感受。同时,两者都对部署的措施冠以“行动”并都详细列举了开展“行动”的重点领域。前者列举了“创业创新”“协同制造”“现代农业”等11个领域,后者列举了“科技创新”“智能制造”“智慧农业”等12个领域。

如何解读“互联网+”和“数据要素×”之间的关系呢?对此,国家数据局副局长沈竹林表示,从前者到后者可以用3个转变来概括:一是从连接到协同的转变。“互联网+”强调的是连接,即各行各业拥抱互联网,实现基于数据生成和传递的互联互通;而“数据要素×”强调的是数据协同,是基于数据有效应用的全局优化。二是从使用到复用的转变。“互

联网+”强调的是千行百业利用互联网技术;而“数据要素×”强调的是基于行业间数据复用的价值创造,拓展经济增长新空间。三是从叠加到融合的转变。“互联网+”强调的是通过叠加到融合的转变。“互联网+”强调的是通过叠加到融合的转变;而“数据要素×”强调的是融合多来源、多类型的数据,驱动创新,培育经济增长新动能。

中央财经大学中国互联网经济研究院副院长欧阳日辉是《行动计划》的起草人之一。他近日撰文对“互联网+”和“数据要素×”之间的关系进行了阐释,指出互联网和数据是数字技术的统一体,前者是“露出水面的冰山”而后者是“深藏在水下的主体”,实施多年的“互联网+”行动为采取“数据要素×”行动奠定了坚实基础,“数据要素×”是“互联网+”的升级和升华,二者本质上都是数字技术与经济社会的深度融合、双向赋能。

纵观《行动计划》所列的“数据要素×”12个重点领域,基本覆盖了一、二、三产业的各行业及重点细分领域,涉及千行百业。比如,在工业制造领域,《行动计划》要求,支持工业制造类企业融合设计、仿真、实验验证数据,培育数据驱动型产品研发新模式,提升企业创新能力。比如,在现代农业领域,《行动计划》提出,支持相关方面融合利用遥感、气象、土壤、农事作业、灾害、农作物病虫害、动物疫病、市场等数据,加快打造以数据和模型为支撑的农业生产数智化场景,实现精准种植、精准养殖、精准捕捞等智慧农业作业方式。

诠释新质生产力

数据是经过收集、记录、测量、统计、描述所得到的信息或信息的载体,可以是数字、文字、图像、声音等形式。随着网络技术的发展,数字技术创新成果与经济社会各领域深度融合,不同领域的数据被生产、记录、传输、存储和消费,数据规模以摩尔定律的速度快速增长。数据的充分挖掘和有效利用,可以优化资源配置和使用效率,对价值创造和生产力发展有着广泛影响。

欧阳日辉表示,在数字经济时代,数据的重要性不言而喻,但是数据并不是从一开始就被作为生产要素,而是经历了一个过程。

2017年12月,习近平总书记在中共十九届中央政治局第二次集体学习时强调“构建以数据为关键要素的数字经济”。2019年10月,中共十九届四中全会明确了劳动、资本、土地、知识、技术、管理和数据等7种生产要素。这是首次正式明确数据的生产要素地位。之后,中央作出改革和加快培育数据要素市场、更好发挥数据要素作用等一系列重大部署,而《行动计划》就是落实这些部署的最新举措。

值得注意的是,《行动计划》把培育和发展新质生产力作为实施“数据要素×”行动的重要目标之一。对此,南开大学经济与社会发展研究院教授杜传忠指出,新质生产力的培育有多种推动力量,其中数据要素作为国家基础性战略资源和关键性生产要素,与算力相结合,将衍生形成强劲的数据生产力,成为新质生产力的重要组成部分。杜传忠特别提到类似ChatGPT这样的生成式人工智能大模型,他认为,AI大模型以智能算力作为底座支撑,通过深度学习算法和万亿级别数据的训练,不断迭代,从而形成新质生产力。此外,数据要素在人工智能技术加持下,推出大量具有深度学习、自主学习能力的机器人,可以进一步深化劳动分工、优化劳动力供给结构,从而形成新质生产力。

中国首个工业级流体仿真大模型“东方·御风”被认为是新质生产力的典型代表。2022年9月,该大模型由中国商飞公司联合有关方面共同发布,在相关流体仿真软件的支持下,有效提高了对复杂流动的仿真能力,将仿真时间缩短至原来的1/24。同时,该模型能对流场中变化剧烈的区域进行精准预测,流场平均误差降低到百分之一量级,达到工业级标准。正是有了大模型“东方·御风”的助力,中国C919大型客机才得以缩短研发周期,才能早日翱翔蓝天。

据新华社电(记者熊琳)“当前正值全球技术革命、产业变革与生物经济发展的重大交汇期,合成生物制造成为各方竞相抢占的战略科技高地。”北京市委常委、副市长于英杰6日表示,北京市正加紧推进市级合成生物制造产业发展规划、行动计划和相关措施,积极推进生物制造产业创新发展。

1月6日,北京市科委、中关村管委会联合昌平区政府在昌平未来科学城举办北京市合成生物制造产业创新发展工作推进会。会上,北京市合成生物制造技术创新中心(筹)和中关村合成生物制造产业集聚区揭牌。

于英杰表示,北京市高度重视合成生物制造产业发展,拥有比较雄厚的合成生物创新资源、比较突出的合成生物产业基础、比较完善的合成生物顶层设计,要紧紧抓住机遇,着力打造具有全球影响力的合成生物产业创新高地。

北京市科委、中关村管委会主任张继红表示,未来将以产业需求牵引科技创新,切实强化创新转化、资源统筹、区域联动和示范先行,促进北京合成生物制造的创新资源集聚、科技成果影响力、新兴产业引领力、区域辐射带动力全面提升,把合成生物制造产业打造为生物经济发展新引擎和未来产业发展新标杆,为北京加快建设国际科技创新中心提供有力支撑。

据介绍,目前北京市级相关单位和部门正在抓紧联合研究制定“三个一”政策,即北京市合成生物制造产业发展规划、行动计划和相关措施,推进会上还介绍了北京市加快合成生物制造创新发展行动计划编制情况。

昌平区委书记、未来科学城党工委书记甘靖中表示,昌平是北京建设国际科创中心的重要承载区,集聚了一批重点高校、科研机构和一批成长性、科技型企业,初步形成了产学研用、协同发展的良好态势。北京市合成生物制造技术创新中心、中关村合成生物制造产业集聚区启动建设,将为昌平和全市合成生物制造产业创新发展注入强劲动力。

江西瑞昌:

科技订单生产忙



近日,江西省瑞昌经开区智造小镇内的江西瑞康技术有限公司车间内,工人正在生产线上赶制订单。该公司是一家专注提供智慧教育系统及智能终端解决方案的高科技企业,产品不仅大量供应国内市场,还远销韩国、日本、俄罗斯、美国等国家。 魏东升摄(人民视觉)

湖南怀化:

小课程点亮大梦想



湖南省怀化市通道侗族自治县引导中小学开发科普特色课程,一些集科普性、竞技性、趣味性于一体的课后服务内容深受学生欢迎。

图为近日,通道侗族自治县第二完全小学的学生参加课间手掷飞机活动。 栗勇主摄(人民视觉)

云遥一号星座3颗星入轨

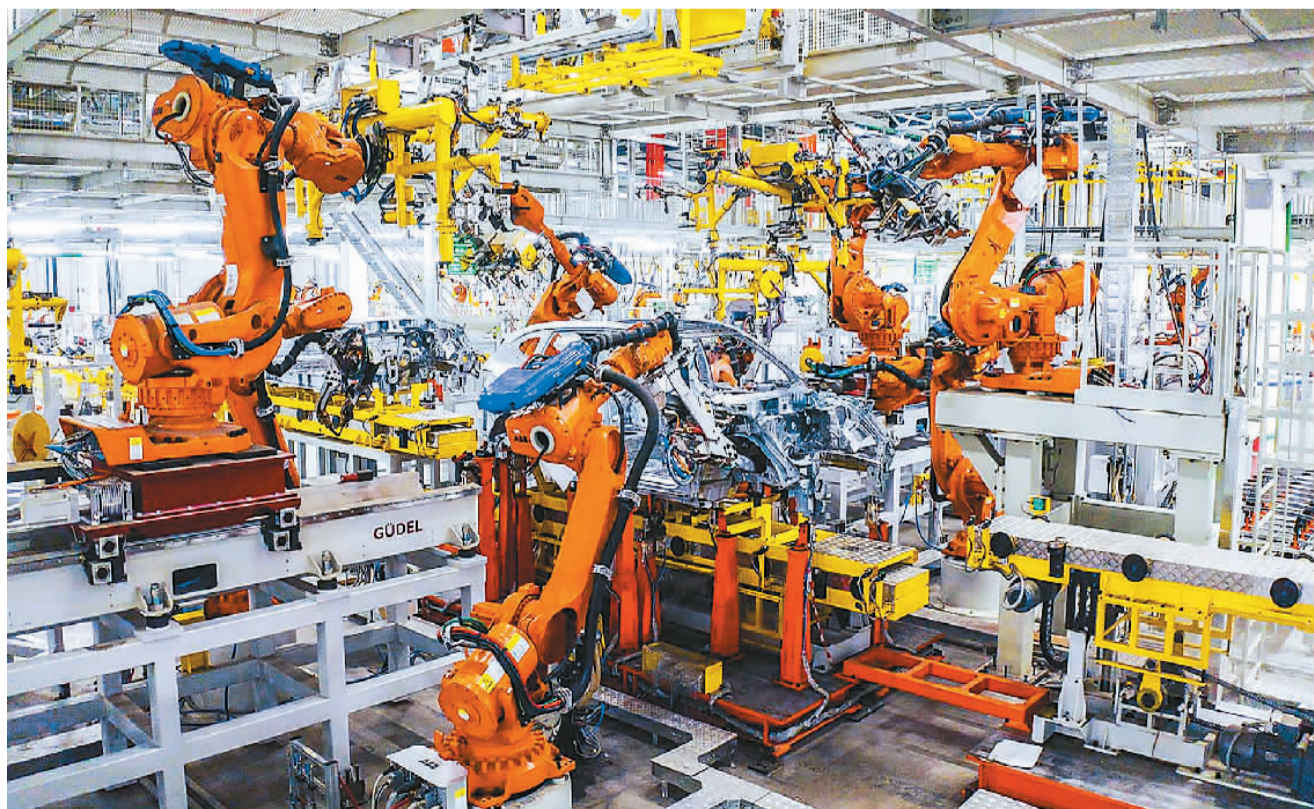
本报电(记者张保淑)近日,云遥一号商业气象空间探测星座3颗卫星在山东烟台附近海域顺利发射入轨。

此次发射入轨的云遥一号星座卫星是吉林长光卫星技术股份有限公司与天津云遥宇航科技有限公司、联信盈达有限公司联合研制的新型气象空间观测小卫星,是“云遥一号”批产卫星的第三批组星,具有研制周期短、集成度高、重量轻、体积小、性价比高等特点,是落实自动化测试和流水线式总装的典型应用,践行批量化生产理念的重要成果。

云遥一号商业气象空间探测星座配置全球导航卫星系统(GNSS)掩星接收机、长波红外相机等载荷,主要产品可用于气象实况分析、数值天气预报、空间环境预警及重大灾害预警等研究及应用。

吉林省科协党组书记、副主席陈耀辉莅临发射现场,向参与云遥一号卫星研制的长光卫星的一线科技工作者表示热烈祝贺和诚挚慰问,希望长光卫星科技人员继续大力弘扬“两弹一星”精神和载人航天精神,为推动吉林省高质量发展注入新动能,为推动航天强国建设作出新贡献。

北京推进生物制造产业创新发展



基于数据要素的智能制造大大提高了生产效率。图为一家中国车企的智能制造车间生产场景。

百度图片