标准对话：
标准化助力人工智能健康发展

于欣丽
中国标准化协会理事长

尊敬的各位专家，大家好！欢迎参加由《中国标准化》杂志社举办的标准高端对话活动，我是此次活动的主持人于欣丽。

首先，我介绍一下参加活动的专家，中兴通讯标准战略部部长田力、中科院软件所副研究员孟令中、蚂蚁集团技术战略发展部总监彭晋、中国电子技术标准化研究院信息技术研究中心人工智能研究室主任/高工徐洋、宝武中央研究院数智中心/宝钢股份数据AI部主任张伟。

田 力

孟令中

彭 晋

徐 洋

张 伟

2024年《政府工作报告》指出：“要大力推进现代化产业体系建设，加快发展新质生产力。”在深入推进数字经济创新发展方面，提到深化大数据、人工智能（AI）等研发应用，开展“人工智能+”行动，打造具有国际竞争力的数字产业集群。这是“人工智能+”首次被写入《政府工作报告》中。随着后续 GPT-4 Turbo、Sora等产品的持续推出，人工智能作为新质生产力的典型代表，成了2024年产业政策的聚焦点。为此，《中国标准化》杂志社举办了此次主题为“标准化助力人工智能健康发展”的标准对话活动。

1956年夏季，以麦卡赛、明斯基、罗切斯特和申农等为首的一批远见卓识的年轻科学家在一起聚会，共同研究和探讨用机器模拟智能的一系列有关问题，并首次提出了“人工智能（Artificial Intelligence，简称AI）”这一术语，会上提出：人工智能就是要让机器的行为看起来就像是人所表现出的智能行为一样。另一个定义认为，人工智能是人造机器所表现出来的智能性。总体来讲，人工智能的定义可划分为四类，即机器“像人一样思考”“像人一样行动”“理性地思考”和“理性地行动”。这里“行动”应广义地理解为采取行动，或制定行动的决策，而不是肢体动作。

“人工智能”是计算机科学的一个分支，它是一门边缘学科，属于自然科学、社会科学、技术科学三向交叉学科。专注于研究、开发和应用模拟人类智能的理论、方法、技术和应用系统。它的目的是创造出能够独立思考、决策、学习和适应环境变化的智能机器或软件代理，使之能够在复杂环境中执行通常需要人类智能才能完成的任务。涉及学科包括：哲学和认知科学、数学、神经生理学、心理学、计算机科学、信息论、控制论、不定性论、仿生学等。

作为21世纪最具颠覆性的技术之一，人工智能已经在各个领域取得了显著的成果。从自动驾驶汽车到智能家居，从虚拟助手到医疗诊断，人工智能的应用无处不在。

2017年7月，国务院印发《新一代人工智能发展规划》，推进办公室由科技部、发改委、工信部、中科院、工程院、军委科技委、中国科协等15个部门构成，负责推进新一代人工智能发展规划和重大科技项目的组织实施。该规划提出了我国人工智能在2020年、2025年、2030年三个阶段的发展目标，为后续一系列落地政策提供了依据和指导。同年11月15日，科技部召开新一代人工智能发展规划暨重大科技项目启动会，宣布成立新一代人工智能发展规划推进办公室，并公布首批国家新一代人工智能开放创新平台名单。将用13年的时间，将我国打造成世界主要人工智能创新中心。

2023年12月，中央经济工作会议聚焦“发展新质生产力”和“加快推动人工智能发展”，提出

要大力推进新型工业化，发展数字经济，加快推动人工智能发展。

2024年2月19日，国资委召开央企人工智能专题推进会，要求中央企业把发展人工智能放在全局工作中统筹谋划，深入推进产业焕新，加快布局和发展人工智能产业，推动央企在AI领域实现更好发展、发挥更大作用。

2024年7月，工业和信息化部联合中央网信办、国家发展改革委、国家标准委共同发布《国家人工智能产业综合标准化体系建设指南（2024版）》，从基础共性标准、基础支撑标准、关键技术标准、智能产品与服务标准、赋能新型工业化标准、行业应用标准、安全/治理标准七方面明确标准化体系建设的重点方向。

2024年7月4日，国务院总理李强在上海出席2024世界人工智能大会暨人工智能全球治理高级别会议开幕式并致辞。李强指出，中国始终积极拥抱智能变革，大力推进人工智能创新发展，高度重视人工智能安全治理，实施了一系列务实举措，发布了《全球人工智能治理倡议》，向第78届联合国大会提出了加强人工智能能力建设国际合作决议并获得一致通过，为全球人工智能发展和治理作出积极探索，贡献了建设性思路和方案。其中，在加强协同共治，确保智能向善方面，李强指出，加强发展战略、治理规则、技术标准等对接协调，推动建立普遍参与的国际机制，形成具有广泛共识的治理框架和标准规范，确保人工智能发展安全、可靠、可控，始终符合人类的根本利益和价值标准。

目前，我国已出台《智慧城市 人工智能技术应用场景分类指南》《信息技术 神经网络表示与模型压缩 第1部分：卷积神经网络》《人工智能大规模预训练模型 第2部分：评测指标与方法》《人工智能 机器学习系统技术要求》等多项人工智能领域相关标准。

随着人工智能的快速发展，算法偏见、数据安全、隐私保护等问题日益凸显，成为其健康发展的瓶颈。此时，标准化便显得尤为重要。它如同一把尺子，衡量着人工智能技术的成熟度与安全性，确保技术的每一次跃进都能稳健前行。

在今天这次标准对话中，我们首先请田力先生谈谈我国在人工智能标准化领域的国际地位及未来发展方向；请孟令中先生谈一谈如何建立灵活、动态的标准化机制以适应人工智能技术的快速发展；请彭晋先生谈一谈针对数据隐私和安全问题，标准化应如何发挥作用以保障用户权益；请徐洋女士谈谈如何通过标准化构建健康、可持续的人工智能产业生态；请张伟先生谈一谈标准化将如何推动人工智能技术在智慧城市、智慧医疗、智能制造等领域的深度应用。

我国在人工智能标准化领域的国际地位及未来发展方向

田 力　中兴通讯标准战略部部长

我将基于中兴通讯在人工智能领域的实践经验展开分享，为后续嘉宾的深入探讨做个铺垫。当前，人工智能已成为新一轮科技革命与产业变革的核心驱动力，在全球科技竞争中占据关键地位。自大模型出现以来，其对全球经济、社会和治理体系产生了深远影响。中兴通讯积极响应，致力于将自身打造为“极致AI公司”，并将自身在信息通信技术（ICT）行业长期积累的基础能力进行拓展，秉承开放的理念，与产业伙伴共同推动人工智能技术的标准化和模块化发展，赋能千行百业。

标准化在人工智能领域的前沿竞争中至关重要，它是推动人工智能技术进步和产业应用落地的关键支撑。一方面，通过构建规则与规范，能够促进人工智能创新与安全的协同发展。另一方面，标准化是技术成果转化为产业应用的重要桥梁，有助于提升产业效率，为人工智能技术在各产业的有效落地保驾护航。

在国家政策层面，人工智能发展规划办公室已经成立，国家经济规划中也包含了人工智能相关的部署内容。尤其值得一提的是，2024年7月，四部委联合发布了《国家人工智能产业综合标准化体系建设指南》，该指南内容详尽，涵盖7个板块、涉及50个细分领域，充分体现了国家在这方面的顶层规划与细致部署。同时，国家网信办很早就开展了人工智能监管立法工作，于2023年7月发布了《生成式人工智能服务管理暂行办法》，2024年9月又出台了文件，要求对人工智能生成的信息（如图片、视频等）添加标识，以便于监管识别。此外，数据要素与人工智能的发展密不可分，国家专门成立了国家数据局及全国数据标准化委员会，促进国家数据供给、流通、应用、安全保障等制度体系建立，支撑数据基础设施建设，构建数据产业生态。在行业层面，工信部发起筹备人工智能技术标准化委员会，并于近期召开了成立大会，正式启动相关的行业标准化工作。

从国际形势来看，在人工智能技术标准化领域，中美竞争态势较为突出。欧洲在监管方面走在前面，2024年7月发布了第一份人工智能法案，这是非常有借鉴意义的。我国在国际标准组织贡献显著，以ISO（国际标准化组织）、IEC（国际电工委员会）、ITU（国际电信联盟）为例，在ISO/IEC JTC 1/SC 42 人工智能技术

委员会中，中国专家担任第五工作组的召集人与秘书职位，注册专家一共55人，各工作组专家覆盖率达100%，且主导研制了8项人工智能相关标准。在国际电信联盟（ITU）中，虽没有专门针对人工智能的研究组，但由于人工智能的高热度，各个研究组都有所涉及。例如：SG 13，它是一个跟未来网络相关的研究组，中国专家担任未来网络与机器学习需求和架构报告人组的报告人，以及人工智能原生网络（AI-Native Network）焦点组的副主席和4个子组的组长；在多媒体研究组（原SG 16，现SG 21），中国专家担任人工智能赋能多媒体应用报告人组的组长。在 2022 年至 2024 年的研究周期内，中国牵头制定了70余项人工智能基础技术标准和10余项网络智能化相关标准。同时，中国还在国际电信联盟理事会层面积极推动成立人工智能专家组，这与在联合国系统中推行人工智能治理理念相契合。同时，在一些行业组织，如3GPP（第三代合作伙伴计划），在5G/6G标准制定中，人工智能相关的讨论非常活跃，中兴通讯专家担任RAN3工作组主席，对网络智能化、智能运维、人工智能数据采集和交互等方面的标准化发挥着重要作用。在 IEEE（电子与电气工程师协会） 中，涉及人工智能基础技术治理和行业应用，其中中国一共牵头制定了 40 多项技术标准。

关于人工智能的未来发展方向，我依据国家发布的《人工智能产业综合标准化体系建设指南》，从底层到顶层依次分析。其中最关键最重要的是基础设施，在智算基础设施方面，特别是算力相关领域，虽然

GPU 本身侧重于性能竞争，但算力互联标准更为关键。构建大规模计算中心需要大量 GPU 卡进行互联，还有不同算力节点之间的互联，其互联性能直接影响推理和训练效率。在这方面，中美呈现出竞争态势。美国英伟达构建封闭系统，而AMD、英特尔等企业联合打造开放生态，如芯片间互联的UAlink（超高速互联联盟），以及服务器间互联的UEC（超以太网联盟）。从国际上来看，美国这两派的阵营是非常明显的。中国有些类似，多数企业对互联持开放态度，但目前头部企业都在各自构建自己的生态，标准不统一，未来需要加强协作，形成合力，以推动智算基础设施的规模产业化。往上就是大模型和人工智能算法层面，重点在于建立分级分类标准，支撑大模型的推理能力、多模态支持、幻觉、安全等方面的评测和评估，以满足各类应用需求。在应用层面，标准和产业组织可以汇聚各类优秀实践案例并进行推广，如ITU每年举办的 AI for Good（人工智能向善峰会）活动，人工智能产业发展联盟发布的先锋案例合集等，通过将实践经验应用在各行各业进行规模复制，实现赋能新型工业化的效果。最后，在安全伦理治理方面，考虑新技术的两面性，需要鼓励和监管并举。期望政府、企业、高校和科研机构等各方协同合作，引导人工智能向善发展，推动中国的人工智能治理理念和框架走向国际，形成符合我国价值观的全球治理规则。

于欣丽点评：

田力先生跟我们分享了我国在国际人工智能标准化领域做出的成就以及发挥的作用，更重要的是，田力先生对我国发布的《人工智能产业综合标准化体系建设指南》进行了分层的详细介绍，感谢田力先生。

如何建立灵活、动态的标准化机制以适应人工智能技术的快速发展

孟令中 中科院软件所副研究员

我们团队长期参与人工智能标准的制定工作，承担了一些工信部、科技部及国防的人工智能相关项目，在工作中积累了一定经验，并形成了建立适应AI技术发展的标准化机制的思考。

在探讨如何建立灵活和动态的标准化机制来契合AI技术的发展需求时，需着重考量以下几个方面：首先，AI技术发展速度超乎寻常。例如，前两年专注于国标神经网络算法评估规范的制定，而近两年间，生成式人工智能以及大模型等新兴技术已迅速崛起。其次，AI技术的迅猛发展极大地推动了其在各领域应用的快速拓展。从早期的智能家居和智能网联汽车领域，延伸至如今新涉足的钢铁行业、电力行业等诸多领域，AI解决方案的需求呈井喷式增长，且各类应用场景对相关解决方案的需求极为迫切。

AI技术发展对标准制定的影响就是标准的更新快，这体现在两个方面。第一，ISO标准工作组的变化。从2018年ISO/IEC JTC1/SC 42成立至2024年，ISO工作组从最初的5个工作组发展到10个左右的工作组和联合工作组，且工作组涵盖的标准范围从基础共性技术拓展到应用、知识图谱、大模型、生成式人工智能及算力等方面，发展极为迅速。第二，国标数量与要求的变化。2022年人工智能标准主要集中于术语方面，如今国标数量逐步增多。例如风险管理国标，2023年12月立项时要求18个月完成，今年上半年被要求缩短至12个月，这充分体现了对标准制定的迫切需求。

灵活的标准化机制主要表现在灵活输入、灵活编制、灵活输出和灵活展示4个方面。

（1）灵活输入。国标制定的输入来源不应局限于ISO/IEC JTC 1/SC 42人工智能标准采标，还应涵盖美国的ANSI、欧盟等。除ISO/IEC SC 42外，SC 7软件工程领域、SC 27智能网联汽车领域以及欧盟智能网联汽车的SAE标准等都可作为输入。例如ISO/IEC SC 27在智能网联汽车的预期功能安全合作及AI安全引入方面的工作，均涉及AI相关内容，均可为我国国标制定提供参考。此外，国内标准输入最初主要在工作组内论证，今后可扩大范围，逐步从产业界和学术界广泛征集标准需求，使标准能更好地反映实际

应用中的迫切要求。

（2）灵活编制。标准编制通常由牵头单位组织、联合单位参与。建议增加标准应用单位参与编制，在标准发布前开展应用工作，实现边编制、边应用、边反馈、边落地、边完善，形成敏捷编制标准的模式，促进标准更好地落地实施。

（3）灵活输出。输出内容不应仅面向单一领域标准，应打通团标、行标和国标，同时向AI应用方面倾斜。输出对象除技术外，更应注重人工智能标准的应用落地。

（4）灵活展示。可依托中国电子标准化研究院的平台，进一步展示标准编制过程与内容，尤其是标准应用情况，包括典型案例与培训等，使应用单位更好地理解标准，推动标准落地。

动态的标准化编制，主要表现在动态立项和动态参编等两个方面。

（1）动态立项。我国团标和国标立项可灵活参考ISO的SC 42、SC 7或SC 27等内容同步研究。鉴于知识产权问题，在保障前提下，针对重要标准在国内成立专门对应的工作组，由多位国内专家组成动态工作组，支撑对口专家向国际标准输入，形成良好的输入输出动态，推动立项工作。

（2）动态参编。标准编制过程中，应根据AI技术在不同领域的应用差异进行动态调整，及时引入新技术领域的应用单位和技术单位，以适应AI标准时效性要求。同时，考虑参与国际标准制定的国内专家利用熟悉国际标准的优势，动态牵头相关国标落地工作，能提高

标准制定效率。

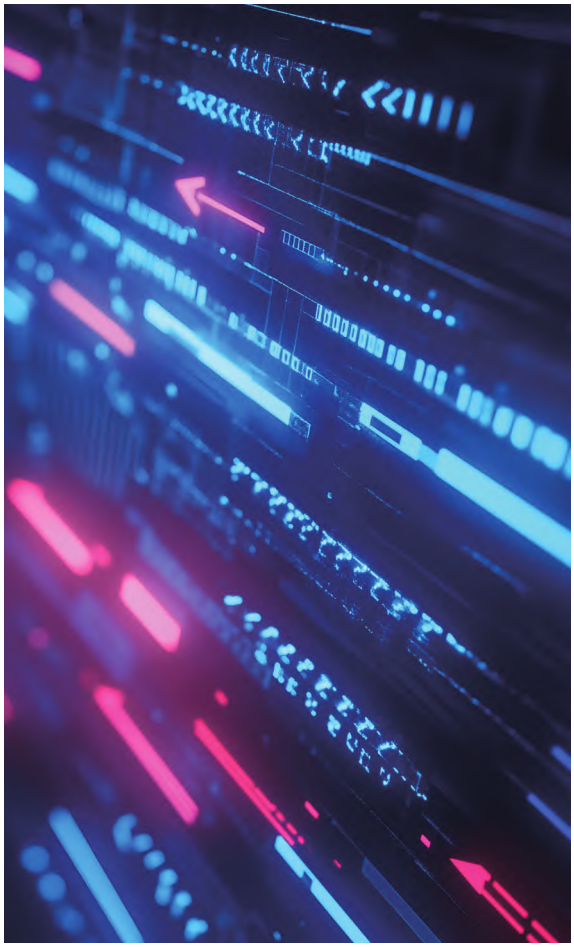
可持续的标准化编制机制，建议包括以下4个方面。

（1）可持续的标准规划。人工智能分委会的十几个工作组每年制定标准规划，规划应动态调整，工作组间可联合编制标准。这既能体现国际标准发展趋势，又能突出各工作组重点，持续输出标准，确保标准制定工作的可持续性。

（2）国际标准的可持续支撑。在60多位国内专家参与国际标准制定的优势下，国内应形成动态研究工作组共同支撑。利用秘书长单位角色，持续向国际标准输出成果，紧跟国际标准发展趋势，做好相关立项工作。

（3）国内标准框架的持续性演进。AI技术发展迅速，团标内容易过时，需做好团标、行标到国标的持续演进，解决标准更新快的需求，建立可持续发展的方式与途径。

（4）标准生态的可持续建设。标准工作涉及研制、应用、评估等多单位，是一个生态系统。应明确各角色职责，做好标准共享，通过平台建设，使更多人使用标准，发挥标准最大作用，形成共享发展、可持续的局面，适应AI技术快速发展。



于欣丽点评：

孟令中先生基于丰富的实践经验，在标准化管理体制运行机制方面的认识极具价值，对标准化管理部门意义重大。其涉及灵活输入、编制、输出、展示与动态立项等内容，类似英国BSI的flash标准，该标准6个月一版，速度快于我国国标研制目标（24个月缩至18个月），其特点是颁布即修订且未牺牲程序，靠协商机制，既使未能协商一致也可出台标准。孟令中先生的观点与之相似，强调立项、参编、牵头等工作的动态性。其可持续性观点包括要有规划、持续支撑国际标准、推动国内标准框架演进及重视生态建设，以实现标准化工作可持续发展。

人工智能中的隐私保护和安全标准化

彭 晋　蚂蚁集团技术战略发展部总监

当我们关注人工智能的隐私保护和安全相关话题时，标准是一个非常重要的框架和工具。标准的作用之一是明确产品服务应当遵从的不同层次的要求。对于数字产品和数字服务，数据安全、网络安全、系统安全，以及包括隐私保护在内的用户权益保护，都是各类产品和服务必须要考虑的一个部分。标准在这些方面形成了一个很好的共识机制，提升了公开透明性。安全要求来自多个方面，主要包括国家安全保障的需要，行业与企业健康良性发展的需求，用户权益保护的需要等，这些需求和实现这些需求的方式也随着技术的进步和社会的发展不断演进。不断演进的标准体系和应时而生的各级标准则将有利于整个行业遵循底线要求，以及不同上下文、不同层次的更高要求。

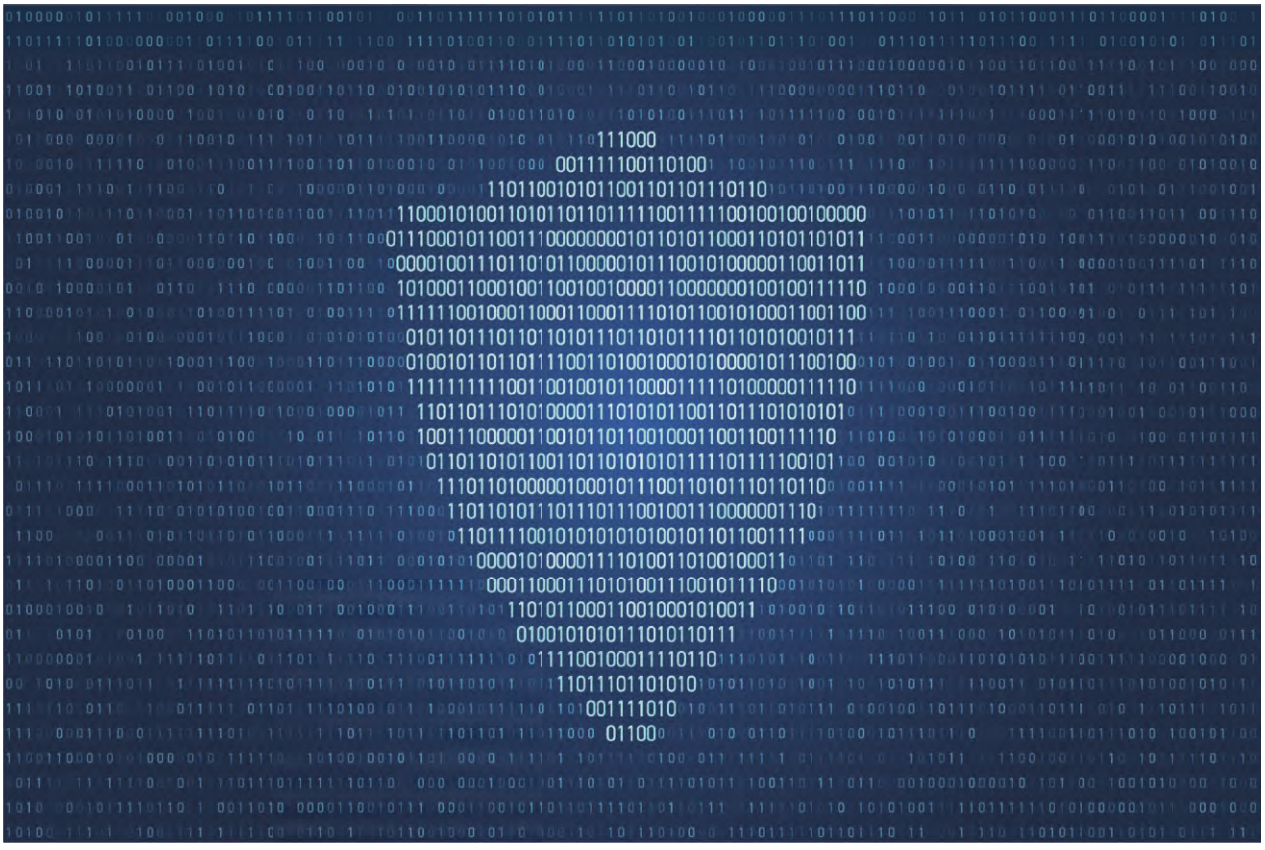
关于隐私保护和信息安全的标准，并不是人工智能系统特有的。在互联网发展进程中，各级标准组织已经制定了包括个人信息保护、信息系统安全等级保护等一系列的国家标准。不同的行业，从不同的维度、不同的切入层测和切入点对信息系统、数据安全、网络安全、云安全、APP安全、应用安全、产品安全、服务安全、开发过程安全、个人信息保护等建立了国标、行标、团标等各级标准，来规范治理要求、技术要求，并提供实施指南等。例如GB/T 35273—2020《信息安全技术 个人信息安全规范》就是个人信息保护的通用基线型国标,GB/T 41817—2022《信息安全技术 个人信息安全工程指南》从产品全生命周期的视角对个人信息保护给出具体指导，GB/T 22239—2019《信息安全技术 网络安全等级保护基本要求》是关于网络安全等级保护的一些基本要求等。人工智能、大模型的开发、部署、服务也需要遵循这些标准，对大模型都适用。但是人工智能中所涉及的新的范式，如机器学习、深度学习、大模型等智能体，对标准会有新的要求。

我国于2022年正式实施的《互联网信息服务算法推荐管理规定》，对算法推荐服务提出向上向善、不得设置诱导用户过度消费、加强内容管理、建立完善人工干预和用户自主选择机制等要求。该规定还特别针对用户权益保护提出7条要求，包括保障算法知情权、保障算法选择权、不得利用算法推荐服务诱导未成年人沉迷网络、便利老年人安全使用算法推荐服务等。又比

如国标GB/T 42888—2023《信息安全技术 机器学习算法安全评估规范》，针对机器学习算法从设计开发到退役下线全生命周期，提出了具体的安全要求；并且针对生成合成类、个性化推送类、排序精选类、检索过滤类、调度决策类五类算法给出了针对性的安全要求。

在国家新一代的人工智能标准体系里，安全和伦理是八个独立的部分之一。网络安全标准化技术委员会正在推进针对这一波大模型浪潮的生成式人工智能的安全治理标准，制定工作包括《网络安全技术 生成式人工智能服务安全基本要求》，在语料安全、模型安全、数据标注安全等方面提出明确要求，并描述了实施机制和评估方法。《网络安全技术 生成式人工智能预训练和优化训练数据安全规范》则规定了生成式人工智能预训练和优化训练数据及其处理活动的安全要求和对应的评价方法。《网络安全技术 生成式人工智能数据标注安全规范》规定了生成式人工智能训练的数据标注基础安全要求、数据标注规则安全要求、标注人员要求、数据标注核验要求和标注安全测试方法；《网络安全技术 人工智能生成合成内容标识方法》则规定了人工智能生成合成内容显式标识和隐式标识的方法。

在一些新的场景下，在人工智能具备越来越多的类人特性和能力的时候，其与人交互，以及执行任务的方式会发生改变，比如copilot的形态和Agent的形态。人工智能可以作为人的代理，帮助或者代替用户完成一系列的人机交互的操作，在自身智能成熟之后，这种操作还将从用户与APP、小程序之间的操作的基础上，延



伸到与物理世界的互操作。人工智能与用户之间的交互边界，决策边界会涉及用户的自主权的问题，这个边界的界定原则，以及实际操作中的约束，沟通，协调等都可能需要新的标准。

另外，因为大模型的系统非常复杂，在各个技术层次上增加了新的风险敞口，从而也面临很多新的攻击方式，也需要新的标准来定义系统的安全要求、安全定级、测评测试方法等，这也是后面需要进一步开展的工作。

在美国，NIST制定了NIST AI 100-1 《AI风险管理框架1.0》（Artificial Intelligence Risk Management Framework，AI RMF）。主要是帮助组织去评估和管理人工智能在设计开发，部署和使用过程中面临的风险。核心的目标是促进负责任的人工智能的开发和使用,确保伦理隐私和安全贯穿于人工智能系统的整个生命周期。基于这样的一个顶层设计，针对大模型，NIST又相继发布了NIST AI 600-1《人工智能风险管理框架：生成式人工智能轮廓》（Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile ），包括在知识产权算法偏见误导性信息、有害信息这方面的一些关注；NIST AI 800-1《管理双重用途的误用风险基础模型》（Managing Misuse Risk for Dual-Use Foundation Models ）:管理双重用途的误用风险的基础模型的一些要求，就是对具有更加强大的算力，达到一定规模的人工智能，对它的安全性、保密性和可信度提出了一些新的指导方针，加强对它的监管。

欧盟则在2024年通过了《人工智能法案》，把人工智能分成四个风险等级，基于不同的风险等级进行监管，进行分类的程序要求，包括对透明度、数据保护等方面都会有一些差异性的要求。

国际标准化组织ISO/IEC JTC1 SC 42是制定人工智能ISO/IEC国际标准的子委员会，在2023年发布了非常重要的42001:2023《信息技术 人工智能 管理体系》,规定了在组织内建立、实施、维护和持续改进人工智能管理系统（AIMS）的要求。此外还有ISO/IEC TR 24028:2020《信息技术 人工智能 人工智能的可信赖性概述》国际标准研究报告，它是对人工智能可信度的一个概述。安全隐私可以看作是更为广泛的人工智能，因为可信会超越安全和隐私，会引入更多的可用性、弹性、可靠性、准确性这方面的一些要求。作为一个可信的人工系统，至少是安全和可信的。

目前，国际国内政策和技术要求的标准层面，在宏观和中观层面相关的文件布局越来越完善，而在微观的操作层面上，在可操作规范方面急需更多的标准建设，在技术要求类的标准上面可以形成更多可供产业参考实现的一些指南性的标准来更好地支撑大模型的构建和服务拓展。

一是建议在隐私保护和大模型的数据利用之间建立技术指导性的标准，比如针对个人信息保护中提到的匿名化的技术手段，在大模型训练中如何实现？匿名化是指通过技术处理使得个人信息无法识别特定自然人且不能复原的过程。在大模型训练这种将海量数据通过巨大算力进行压缩的场景里面，怎样才能做到去标识化的数据不能被复原呢？需要采取哪些控制手段把去标识化的个人信息做到匿名化的级别，从而能合规地利用？这个可以有更加细致的标准来进行明确。

二是在安全方面，可以结合针对大模型的安全攻击的演进变化，进一步完善安全要求，并建立安全分级的标准。所有的安全措施都需要投入成本，也无法在任何环境下做到绝对的安全，那么在到底什么样的环境下，我们要抵抗什么样的安全攻击，是一个安全分级的问题。安全分级的标准是在安全性与可用性以及成本之间的匹配的共识，更有利于在 To B 的生态里面更好地去进行标准的评估和采购，也更有利于保护用户的信息安全。

三是在大模型的可解释性和透明度上可以建立可行的、可操作、可度量的标准。针对深度学习的可解释性仍然是一个学术难题，但做到一定程度的可解释性和透明性是有可能的，恰当的透明度既有利于保障用户权益又有利于在To B生态里面的模型产品选型，即使是简单的模型信息披露也是有所助益的。

四是模型应用结合行业需求所需的标准。包括对垂类模型的专业度评价、伦理评价等方面的标准。从专业度的角度，如果用大模型来写作文或者画画，对它的可信度的要求会相对低一些，甚至有时候我们会利用大模型的幻觉来搞一些新的创作，但是在一些严肃的行业应用里面，大模型的输出是否专业准确就有不同的意义了，应有面向专业应用的度量，比如金融、医疗等应用。而且在不同的行业里面，伦理、合规的范畴也变得专业，需要有专门的标准去规制。

最后，我们刚才谈到安全和隐私方面，大部分都是从治理的角度出发的。从发展的角度来看，其实大模型的应用也可以促进用户权益的保护，比如用人工智能的算法来实现数字应用的普惠，比如通过人工智能应用来做数字应用的适老化以及帮助残障人士使用数字化应用，消除数字化鸿沟，将来也可以在这些方面推进标准制定来支持业务发展，做到更加广泛的用户权益的保障。

于欣丽点评：

彭晋先生为我们分享了数据隐私和安全问题。在隐私保护和大模型的应用语境里，需要标准来指引和推进，还有在安全方面要综合安全攻击的演进，要结合安全攻击的发展，根据不同的演进、不同的环境来进行标准的制定，要分等、分级、分类来要求，才能使标准更加具有适用性、才能更加好用。用户的隐私保护标准的制定，要与行业实际相结合，要落地，标准的落地能够引导行业的发展。另外是安全的隐私，从发展的角度看，大模型的应用也可以促进用户权益的保护。

标准化引领高质量人工智能产业生态建设和发展

徐 洋　中国电子技术标准化研究院信息技术研究中心人工智能研究室主任/高工

当前，党中央、国务院高度重视人工智能发展，围绕人工智能产业发展、赋能应用、安全治理等方面做出重要部署，推动人工智能与实体经济深度融合、赋能新型工业化。标准化在人工智能产业发展过程中发挥着基础性和引领性作用，对推动技术进步、加快产品落地、牵引生态建设、加强行业自律具有重要意义。如何以标准促发展成为各国关注的主题。

一、世界主要国家人工智能标准化建设情况

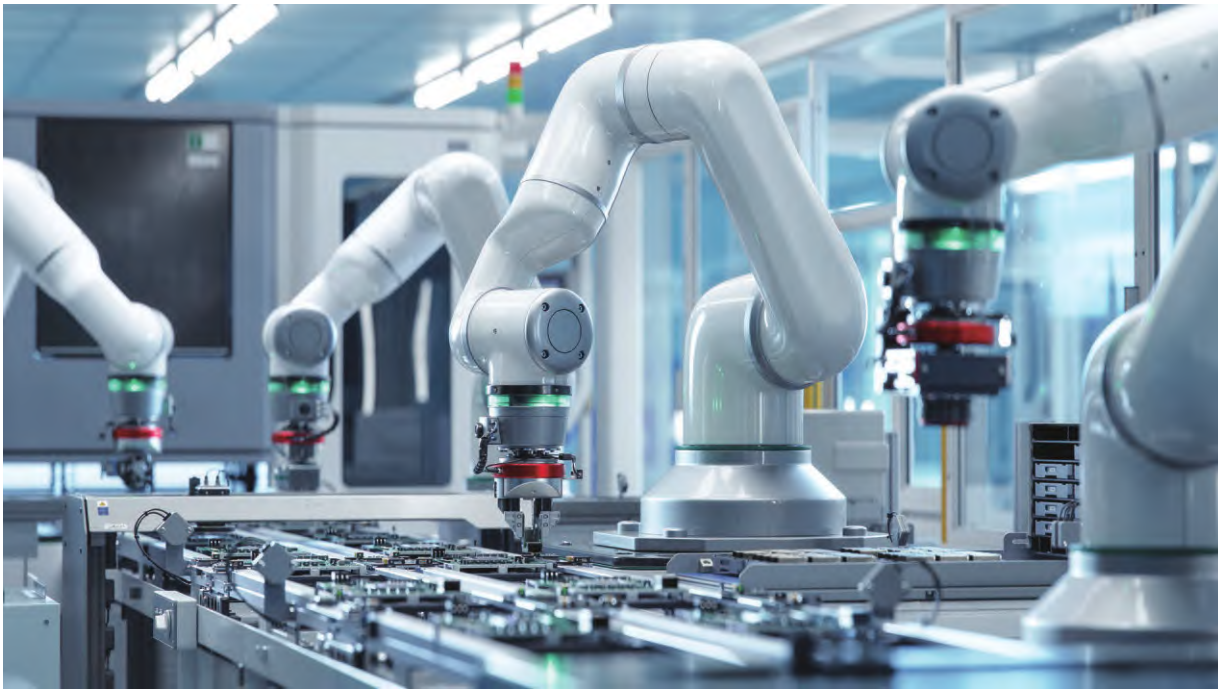
从国际方面看，各国正以前所未有的速度加紧人工智能政策布局，产业发展热度空前高涨。与此同时，各国对人工智能标准的重视程度也达到了前所未有的高度。近两年，美国发布了《关键新兴技术标准战略》和《战略路线图》《关于安全、可靠、可信赖地开发和使用人工智能的行政命令》《关于提升在标准制定中的参与度和领导力的建议》等4项与AI标准紧密相关的政策文件，着重强调要加强对标准预研的投资，以促进技术创新、前沿科学和转化研究，从而推动其在国际标准发展中的领导地位。同时，授权国家标准与技术研究所（NIST）等机构开展人工智能风险管理（AI RMF）系列标准研制、建立基准、开发测试环境、制定全球参与计划，以构建研究—标准—认证生态，为美国人工智能产业引领全球市场做准备。

欧盟今年正式发布《人工智能法案》，基于《人工智能法案》，授权CEN/CENLEC共同开展风险管理、数据质量、透明性等10项协调标准研制，并计划于2025年12月前交付。通过标准将法案中较为模糊的基本要求转化为具体的技术要求，推动人工智能产业的健康、有序发展。

由此可以看出，美国和欧盟虽然在标准化工作的方向和手段各有不同，但都是以标准为抓手，来抢占人工智能产业的全球主导地位和话语权。

二、国内产业现状及存在问题

目前，在国家各相关部门的大力推动下，我国人工智能产业发展成效显著。技术层面。围绕算力、算法和



数据等关键核心技术不断突破，构建从智能芯片到算法框架到大模型的全栈式产业链，大模型技术迭代和产品不断迭代升级。

产业生态层面，我国人工智能核心产业规模不断提升，企业数量超过4500家，一批生态主导型龙头企业和“专精特新”企业茁壮成长。依托人工智能创新应用先导区，形成了上海模速空间、北京全域人工智能之城、深圳人工智能先锋城市等产业聚集区。

赋能应用层面，智能网联汽车、人形机器人、医疗装备等智能产品加速迭代，智能工厂、智能车间等新模式新业态不断涌现。目前，我国已培育421家国家级智能制造示范工厂、万余家省级数字化车间和智能工厂，成为推动人工智能赋能新型工业化的关键力量。

从全局上看，我国人工智能产业已经取得全面的发展。但同时，还存在产业生态小而散等问题。标准作为各利益相关方协调一致的产物，对整合产业资源，推动技术路线收敛，服务产业生态建设，增强技术和产品的适用性，提高人工智能产业发展质量发挥重要作用。

三、几点考虑

（一）以标准汇聚产业发展强大合力。据统计，截至2023年底，我国人工智能核心产业规模接近5800亿元，同比增速13.9%，核心企业数量超过4400家。解决供给方案不足、各行业智能化升级路径不清晰等共性问题。从供给侧和应用侧统筹布局，系统化推进标准研制。聚焦人工智能能力、智能化场景、系统性解决方案等方面，通过标准的研制和应用推广，规范产业制度、提升供给能力、推动典型应用场景落地。

（二）以标准引领前沿技术创新。目前大模型从单模态向多模态发展，以及向行业大模型发展。围绕大模型训练、推理、部署等全流程研制技术要求标准，并面

向电力、钢铁、教育、医疗、交通等行业布局行业大模型标准，形成全面的大模型标准体系。二是要响应人工智能技术产业化需求，研制智能家居、智能个人计算机（AI PC）等智能产品标准，以及模型即服务（MaaS）、部署工具链等智能服务标准，推动大模型产品落地应用。三是积极响应人工智能技术发展趋势，面向智能体、具身智能、群体智能等下一代人工智能创新技术，适度超前布局相关标准研制工作。

（三）以标准促产业智能化转型升级。人工智能是发展新质生产力的重要引擎，与传统产业、新兴产业、未来产业深度融合，赋能千行百业数智化转型。标准通过凝练行业共识，确定通用技术方案、统一接口规范，减少重复研发投入，节约行业智能化升级成本，支撑“人工智能+”高水平赋能。选择重点工业行业开展标准研制和应用试点，将人工智能标准应用于研发设计、中试验证、生产制造、运营管理等全流程关键环节，以关键场景应用为需求，牵引标准研制。通过典型案例的应用和推广，形成跨行业、跨领域、跨大中小融通发展的人工智能标准推广模式，高水平赋能工业制造体系。

（四）以标准规范产业发展安全治理和秩序。人工智能技术发展速度快于安全能力和治理机制建设，引发信息泄露、模型漏洞、算法歧视等多重挑战，恐成为数字世界的“脱缰之马”。标准对规范行业秩序、设立市场门槛具有重要作用，通过布局人工智能安全规范，能够划清技术红线，保障安全底线。

（五）以标准促进全球国际交流合作。全球主要国家高度重视人工智能产业发展，将标准作为国际交流的通用技术语言、国际贸易的基本规则、国际市场的通行证。人工智能国际标准化促进了人工智能国际交流合作，促进国内先进的人工智能技术、产品走出去，为全球人工智能产业发展贡献中国智慧。

四、几点建议

（一）创新标准化组织建设模式。人工智能赋能传统行业的应用过程中，打破了各行业之间的界限边界和配套关系。在跨领域的融合过程中，标准化的协作成为关键。如何发挥好国家人工智能标准化总体组，统筹好各组织协作，更好发挥作用，成为关键。

（二）强化标准链和创新链协同。一是建立高效联动机制，促进科技创新与标准创制供需对接，推动将先进适用的技术攻关成果快速转化成标准。二是编制人工智能领域标准图谱，开展国内外标准比对分析，明确标准长短板和发展规划。三是加强全产业链标准化协作，推动人工智能技术研发、标准研制与专利布局同部署、同推动，促进产业链锻长补短。

（三）强化重点标准供给保障和宣贯应用。一是瞄准产业急需，面向数据质量、评测基准、互联互通等重点问题，“急用先行”研制相关标准规范。二是突出应用导向，加强跨行业、跨领域标准化技术组织的协作，协同推进人工智能与重点行业融合应用。三是加强示范推广，推动行业协会、标准化技术组织等开展人工智能标准体系、重点标准的宣贯。

（四）强化国际交流竞合与标准互通。一是支持相关单位积极参与ISO、IEC、ITU等国际标准化活动，贡献中国方案，传递中国理念。二是凝聚产学研合力，组建多元化、高水平的人工智能国际标准专家团队，不断提升我国国际标准提案数量和质量。三是通过金砖国家、“一带一路”倡议等国际合作机制，推动区域内标准互联互通，提高中国标准影响力。

于欣丽点评：

徐洋主任跟我们分享了很多，我国人工智能发展情况是大模型多而不精，徐洋主任对此给出了非常好的建议，包括标准化组织建设模式创新、标准链与技术链有效衔接和协同、人工智能与传统产业融合中体现标准的作用、加强国家标准之间的互联互通和认证等等，以此来促进人工智能产业的发展。

标准化将如何推动人工智能技术在智慧城市、智慧医疗、智能制造等领域的深度应用

张 伟　宝武中央研究院数智中心/宝钢股份数据AI部主任

一、人工智能应用正在不断深化，规模正在快速扩大

人工智能是引领新一轮科技革命和产业变革的基础性和战略性技术，正成为发展新质生产力的重要引擎。在智慧城市、智慧医疗、智能制造等领域，AI技术的应用已逐步渗透并展现出巨大的经济贡献。2017年国务院印发的《新一代人工智能发展规划》中预测，2025年新一代人工智能在智能制造、智能医疗、智慧城市、智能农业等领域得到广泛应用，人工智能核心产业规模将超过4000亿元。根据《中国新一代人工智能科技产业发展报告2024》的数据，2023年我国人工智能核心产业规模为5784亿元，发展速度远远超过规划。这组数据说明人工智能技术在智慧城市、智慧医疗、智能制造等领域的应用正在不断深入、规模不断扩大，这是标准化推动深度应用这一话题的前提和基础。

马里兰大学杰出教授、工业人工智能中心主任李杰指出，无论中美，都在加大利用落地的人工智能技术解决疾病治疗、交通堵塞、投资决策等实质问题。比如人工智能赋能医疗，在美国一款药品从开发到最终上市大约要18年时间，最后7年是临床试验期，无法缩短。前面11年可以在人工智能的帮助下缩短开发、验证的过程，帮助很多人更早摆脱对应疾病的威胁。

在智慧城市建设中，推动智能基础设施和高效治理。AI技术被广泛应用于交通管理、能源管理、环境监控和公共安全等方面。例如，通过AI分析交通数据，可以智能调控交通信号，缓解城市交通拥堵；通过大数据与AI的结合，可以优化能源使用，提升资源利用效率。然而，智慧城市的构建涉及不同的系统与设备，如何确保数据共享、设备兼容、系统协同工作，是AI技术落地的关键问题。标准化在这一过程中发挥着重要作用。统一的数据格式、通信协议、接口标准等能够确保不同设备与系统的互联互通，推动智慧城市的高效运营与智能化治理。

在智慧医疗领域，提升医疗服务的质量与效率。AI技术已经在疾病诊断、个性化治疗、远程医疗等方面取得了显著进展。例如，AI可以通过分析医学影像辅助医生诊断，或者通过算法预测疾病风险。然而，医疗行业涉及大量敏感数据和多种技术平台，AI应用的标准化同样至关重要。医疗数据的标准化可以确保各医疗机构之间的信息互通与共享，避免数据孤岛；而诊疗算法的标准化能够保证不同系统中AI模型的有效性与准确性，避免因技术差异导致的诊断误差。制定

统一的标准，将帮助构建一个更加高效、可靠的智慧医疗生态。

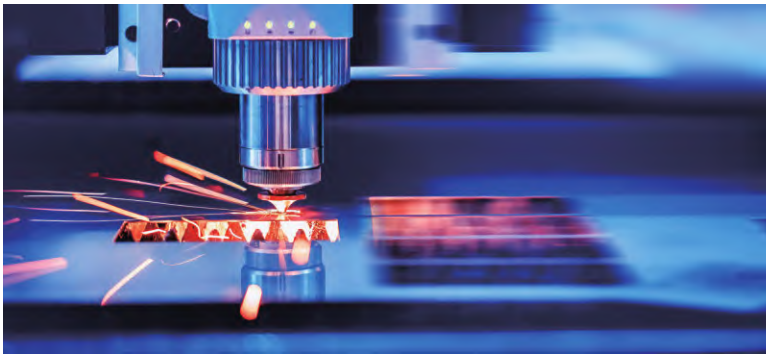
在智能制造领域，提升生产效率与智能化水平。通过机器学习和深度学习技术，AI能够分析生产过程中的大量数据，优化生产线的效率，提升产品的质量。然而，智能制造往往涉及不同类型的设备与技术平台，如何确保这些设备间的协作与兼容，这就需要依赖于一套清晰的标准化体系。标准化的工作应包括AI应用场景标准、工业数据治理标准、AI算法的验证标准等，这将提升制造过程中的自动化水平与生产效率。特别是钢铁、石化化工、有色金属、建材等具有流程制造属性的原材料工业处于既要守住传统优势阵地，又要构建新材料等新的增长引擎，加快培育新质生产力的发展新阶段。原材料工业流程复杂、工序长、耦合性强，利用人工智能技术解决跨工序、跨界面的应用场景多、潜力大。以钢铁行业领军企业中国宝武钢铁集团有限公司为例，作为国资央企，中国宝武对人工智能技术高度重视，在新一轮规划中明确重点开展人工智能及冶金行业大模型应用，打造一批大数据、人工智能与钢铁深度融合的典型示范项目，塑造钢铁行业智能化新优势。

AI赋能行业不断深化、应用规模不断扩大，急需标准化的引导与支撑。标准化不仅能够为AI技术提供统的规范与框架，还能够促进技术互联互通、数据共享和系统兼容性，从而推动人工智能技术的应用。

二、标准化推动人工智能深度应用的进展

人工智能技术赋能行业要求新一代信息通信技术与领域技术深度融合，贯穿到业务活动的各个环节，具备感知、学习、决策、执行、适应等新型功能。融合的模式、路径和应用的场景涉及大量全新实践和经验总结，从标准化角度正在开展3方面工作：（1）厘清概念、统一语言、形成共识；（2）顶层设计、构建体系、形成架构；（3）需求驱动、形成标准、推广应用。

国际标准化组织（ISO）、国际电工委员会（IEC）等标准化组织先后成立智能制造或组建全球先进工



业系统组，推进全球智能制造技术体系的统一和协调。国家标准化管理委员会落实《中国制造2025》中明确的制造业标准化提升计划，推进智能制造技术标准研制。

2024年6月，工业和信息化部等四部门印发《国家人工智能产业综合标准化体系建设指南（2024版）》。指南从五大方面阐述了人工智能产业标准体系的总体要求、建设思路，特别发布了人工智能标准体系结构图和框架图。指南指出，人工智能产业链包括基础层、框架层、模型层、应用层4个部分。其中，基础层主要包括算力、算法和数据，框架层主要是指用于模型开发的深度学习框架和工具，模型层主要是指大模型等技术和产品，应用层主要是指人工智能技术在行业场景的应用。人工智能标准体系结构包括基础共性、基础支撑、关键技术、智能产品与服务、赋能新型工业化、行业应用、安全/治理7个部分。这份指南既体现了标准体系建设的重要性，又为后续人工智能标准化提供了总体方向。

另一方面，业务端在人工智能技术应用方面也在快速发展。以中国宝武为例，在制造管理、供应链决策、关键制造工序等开展了大量人工智能技术研发，开展了较广泛的人工智能技术应用，包括服务钢铁产品质量的机器视觉、深度学习技术应用，服务钢铁关键工序控制的神经网络、随机森林技术应用，服务“一总部、多基地”供应链的组合优化、决策树技术应用，同时开展了一批冶金行业专用机器人系统研发。未来3年全面布局人工智能模型能力，形成决策型、控制型、感知型钢铁行业人工智能。这对标准制定提出了紧迫需求，也为标准制定提供海量应用场景。

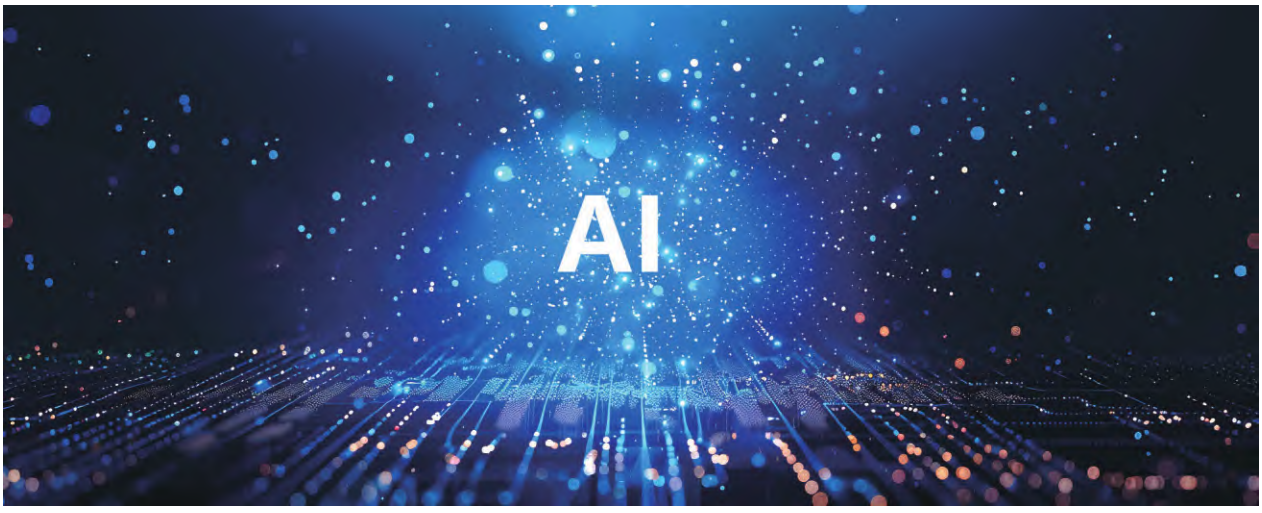
为此，全国信息技术标准化技术委员会人工智能分委员会创新工作模式，成立钢铁、电力、石化等行业应用工作组，从应用端和技术供给端双向发力，推动标准化。宝钢股份作为钢铁应用工作组组长，推动研制了《人工智能 钢铁大模型技术要求》《人工智能 钢铁生产过程知识图谱构建技术要求》和《人工智能 钢铁生产冶炼过程智能决策的数据处理与模型构建技术要求》等多项行业标准，已获立项，正在研制《人工智能钢铁大模型数据集》等标准，推动“人工智能+钢铁”统一语言，形成架构。同时，开展《人工智能赋能钢铁行业典型场景图谱》《人工智能钢铁大模型数据集技术要求》《人工智能 炼钢生产计划排程与调度优化系统技术要求》等指导应用的标准研制。

据统计，“十三五”以来，在智能制造领域已累计发布国家标准394项，其中基础共性标准109项、关键技术标准280项、行业应用标准5项，牵头制定了48项国际标准（欧阳劲松，《从国内外标准化实践看智能制造

技术发展》，中国仪器仪表，2024年）。这些标准为推动智能制造，提升国际影响力起到了重要作用。

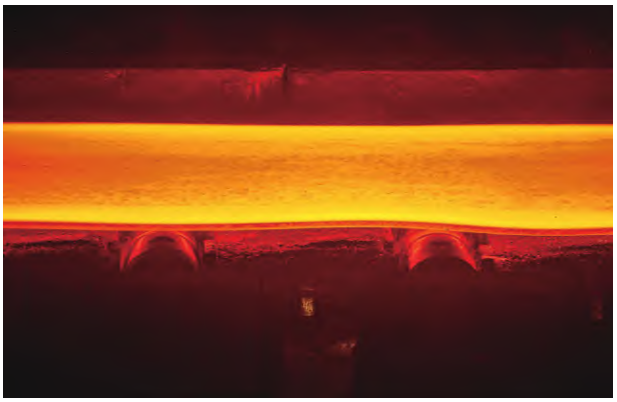
三、标准化推动人工智能深度应用的思考与建议

（1）跨界面，体现标准的系统性、融合性。人工智能技术赋能行业的标准对象广，技术交叉多（刘斓冰，《钢铁行业智能制造标准化成效与展望》，工业技术创新，2023年）。需要重点围绕跨领域、跨层次、跨界面的系统集成开展标准化，形成数据、算力、算法和应用融合的标准体系。首先，数据标准化是AI技术成功应用的基础，统一的数据格式与交换标准能够保证数据在不同平台间的流动与共享。其次，算法标准化也至关重要，特别是AI算法的透明性和可解释性，以确保AI决策的可信度与公正性。此外，安全性标准是标准化工作中的关键内容。随着AI技术的应用范围日益扩大，如何确保AI系统的安全性、数据隐私保护以及防止算法滥用，需要制定严格的安全标准。以钢铁行业为例，具有流程长、工序界面多、生产体系复杂、单/多基地管控模式多样等特点，跨工序、跨界面的协同管控与优化是人工智能应用的重大场景，也是标准化工作的重大潜力。可以围绕新一代信息技术应用，制定工业大数据平台、数据治理、检测方法、评价模型、网络安全等规范/规程/指南类标准；围绕铁钢轧大工序边侧智能工厂、单基地及企业多基地协同管控等，制定质量、物流、能源环保、安全、设备等分层集成优化管控等规范或规程类标



准；围绕上料、生产、下料等环节，制定铁钢轧各工序的生产过程智能化控制、辅助工序智能化、数字孪生及工序衔接等规范或规程类标准；围绕质量缺陷检测、性能检验及工序界面转运等场景，制定智能检测装备、工业机器人及无人运输装备等智能装备类规范/规程/指南类标准；制定面向行业的智能工厂评价、能力评估等实施指南标准。

（2）强布局，体现标准的时效性、引领性。人工智能、大数据和信息化技术迭代快，发展快，要跟上技术进展的步伐，既避免炒热点跟热点影响标准的权威性



准确性，又避免过于保守错失新技术新应用的标准化。提前研究技术成熟度及其与工业要求的适配性，制订一批新技术应用制造过程的“技术要求”标准，以科学引导和强化新技术对制造业提质增效的“倍增器”作用（欧阳劲松，《从国内外标准化实践看智能制造技术发展》，中国仪器仪表，2024年）。

（3）重实效，体现标准的实用性和价值性。坚持场景和需求驱动，从应用中来，回到应用中去，让标准具有生命力。应用场景的标准化，通过制定AI赋能行业场景图谱相关标准，识别并引导高价值应用场景，提升AI赋能行业质量。2024年，工业和信息化部科技司开展《人工智能赋能新型工业化典型应用》案例征集遴选工作，151项案例入选。工信部电子标准院组提出将这些重点场景优先孵化为国家标准、行业标准，充分体现了需求驱动标准化这一原则。再具体而言，钢铁工作组2025年工作计划中纳入了《炼钢生产计划排程标准》和《钢铁企业转底炉智能控制系统》两项具体工作，其原因就是可以帮助钢铁企业创造显著的经济效益，同时应用了较多的人工智能技术，处于国际领先，未来有望申报国际标准，体现中国的引领地位。

于欣丽点评：

张伟先生从应用层面对人工智能的一些标准进行了评论，提出了需求推动标准这一原则，并且张伟先生也提出了行业应用要与标准的供给端结合起来、融合起来，共同推进人工智能的标准化工作。