中国自动化学会通讯

COMMUNICATIONS OF CAA

第11期

2024年11月

第45卷 总第254期

主办：中国自动化学会　http: //www.caa.org.cn　E-mail: caa@ia.ac.cn　京内资准字2020-L0052号

2024中国自动化大会

China Automation Congress

CAA 中国自动化学会

扫描二维码 关注官方微信

扫描二维码 关注官方微博

中国自动化学会通讯
Communications of CAA

主管单位 中国科学技术协会
主办单位 中国自动化学会
编辑出版 中国自动化学会办公室

关注官方微信　关注官方微博

刊名题字 | 宋　健
地　　址 | 北京市海淀区中关村东路 95 号
邮　　编 | 100190
电　　话 |（010）8254 4542
传　　真 |（010）6252 2248
E-mail:caa@ia.ac.cn
http://www.caa.org.cn
印刷日期 | 2024 年 11 月 30 日
发行对象 | 中国自动化学会会员及自动化领域科技工作者

本刊声明

◆ 为支持学术争鸣，本刊会登载学术观点彼此相左的不同文章。来稿是否采用并不反映本刊在学术分歧或争论中的立场。每篇文章只反映作者自身的观点，与本刊无涉。

主编的话

郑南宁

中国自动化大会是由中国自动化学会创办的自动化、信息与智能科学领域顶级综合性学术会议，自 2009 年创立以来，已历经十五载春秋。十五年间，中国自动化大会的足迹遍布全国，包括杭州、北京、长沙、武汉、济南、西安、上海、昆明、厦门、重庆等城市，大会吸引了众多两院院士、Fellow、学科领军人物以及企业家等高层次人才的积极参与，已成为自动化领域内学术研究和产业发展的重要风向标。

2024 年 11 月 2–3 日，2024 中国自动化大会在青岛世界博览城国际会议中心举行。本届大会由中国自动化学会主办，青岛科技大学、国家橡胶与轮胎工程技术研究中心承办。大会以“新质发展 智控未来”为主题，邀请了 10 余位院士和 300 余位高校、科研院所校长、院长等学术精英与会，3000 余名来自学术界和产业界的代表参会，近千万人次在线观看了直播。大会共设 10 场主旨报告和 46 个平行会议，涵盖工业互联网、脑机智能、智慧教育、智能机器人、智慧能源等多个前沿热点领域，300 余位专家学者作会议报告，共接收会议论文 1500 余篇，10 余家行业知名厂商赞助支持。

本期专刊聚焦“2024 中国自动化大会”，旨在为读者呈现一场关于自动化、信息与智能科学领域的思想盛宴。我们特别分享了中国自动化学会理事长郑南宁院士，山东省教育厅总督学王浩，青岛市市委常委、常务副市长耿涛，青岛科技大学党委常委、副校长吕万翔的大会开幕式致辞。重点介绍了国家橡胶与轮胎工程技术研究中心副主任、软控股份有限公司董事长兼总裁官炳政的“轮胎行业智能制造与精益管控”和龙芯中科技术股份有限公司董事长、中国科学院计算所研究员胡伟武的“破解卡脖子问题，构建新发展格局”2 篇专题文章。

借此机会，我们向所有贡献智慧与见解的专家学者致以最诚挚的感谢。您们的真知灼见不仅丰富了本期专刊的内容，也为业界同仁提供了宝贵的学术资源和实践指导。

【目录】CONTENTS

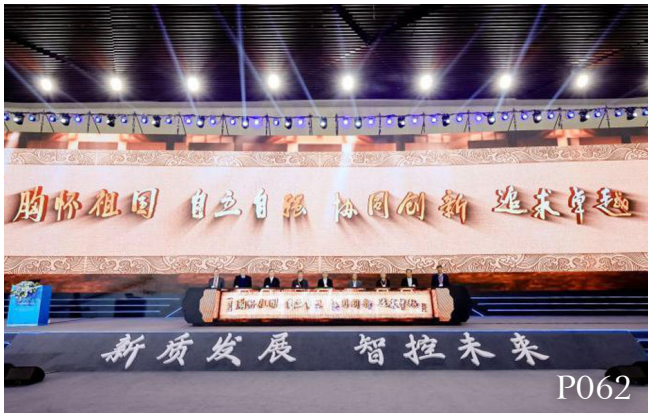
专题 / Column

观点 / Viewpoint

学术前沿 / Academic Frontier

科普园地 / Science Park

学会动态 / Activities



P062

P075

P072

党建强会 / Party Building

形势通报 / Voice

P073

在 2024 中国自动化大会开幕式上的致辞（内容节选）

文 / 中国自动化学会理事长　郑南宁

中国工程院院士、中国自动化学会理事长、西安交通大学教授郑南宁致辞

当前，人工智能正在使自动化技术以更大的规模和深度融入各行各业，渗透到我们生活的每一个角落。人工智能，特别是机器学习和深度学习算法，能够赋予自动化系统以学习和适应的能力，使其能在复杂、动态的环境下完成人类赋予的任务。我们的未来，正在被这些智能系统以一种前所未有的速度和复杂度重塑。

我们知道，控制的本质是反馈，无论是在工程、经济和生物系统中，反馈都是实现有效控制的核心要素。将人工智能、外部环境变化的表征与控制融合，我们就能实现一种面向任务的具身智能。具身智能是人工智能领域的一个重要方向，它强调智能体通过感知、认知和行动，与物理环境进行深度交互。简而言之，具身智能是指“身体”不是等待加载算法的机器，身体也参与算法的进化。具身智能将极大增强自动化系统的柔性能力，使得自动化系统不再局限于预设的程序，而是能够根据环境的变化进行自主决策和行动，包括对人的意图和行为动机的理解。这也是未来工业软件－具身软件发展的新趋势。

智慧从来都是激情的澎拜，创新必然都是思想的飞翔。各位同仁，让我们携手并进，把握自动化和人工智能带来的历史性机遇，共同应对这些挑战，用智慧和创新引领中国自动化技术的未来。

在 2024 中国自动化大会开幕式上的致辞（内容节选）

文 / 山东省教育厅总督学　王浩

山东省教育厅总督学王浩致辞

随着新一轮科技革命和产业变革深入发展，建设现代化产业体系成为推进中国式现代化的重要支撑。自动化作为科技创业与产业创新深度融合的关键抓手，在推进现代化产业体系建设中发挥着不可或缺的重要作用。

习近平总书记视察山东时指出，山东科技、教育、人才资源丰富，在推进科技创新与产业创新深度融合，发展新质生产力，完善现代化产业体系上大有可为。

近年来，山东省深入学习贯彻习近平总书记重要指示精神，发挥高等教育、职业教育大省优势，统筹教育、科技、人才一体发力，积极构建科技创新和产业发展需求牵引的学科专业调整机制和人才培养模式，持续强化科研平台建设，不断攻克重大科研成果，有效加速传统产业升级，新兴产业壮大，未来产业布局，为山东在全国发展中“走在前，挑大梁”提供了有力支撑。

目前，全省高校设有控制科学与工程等自动化相关领域博士点 47 个，硕士点 153 个，专业 217 个，拥有自动化领域全国重点实验室 4 个，教育部和山东省重点实验室 84 个。

十四五以来，累计获自动化领域国家和省部级科学技术奖 108 项，为绿色低碳高质量发展注入了新动能，塑造了新优势。

中国自动化大会，是我国自动化信息与智能科技领域规格最高、影响力最大的学术盛会，本次大会紧扣自动化领域前沿技术，与行业全面深化改革，以“新质发展，智控未来”为主题，以自动化技术创新与应用为主线，共同探讨行业未来发展方向，对推动我国自动化产业创新发展，加快形成适应高质量发展的新质生产力具有重要意义。

希望青岛科技大学以本次大会为契机，瞄准国际前沿，聚焦国家战略，着眼区域需求，积极深化与国内外顶级专家学者，科研院所、头部企业的交流合作，共同推动自动化领域产学研深度融合，努力产出一批国际领先的重大标志性创新成果，攻克更多卡脖子关键技术，为自动化产业的创新发展作出新的更大贡献。

在 2024 中国自动化大会开幕式上的致辞（内容节选）

文 / 青岛市市委常委、常务副市长　耿涛

青岛市市委常委、常务副市长耿涛致辞

中国自动化大会作为国内最高层次的自动化信息与智能科学领域的大型综合性学术会议，自2009 年创办以来，已成为国内自动控制、仪器仪表、航空航天、人工智能等领域重要的学术盛会，为推动学科发展进步，促进产学研用深度融合，实现高水平科技自立自强作出了重要的贡献。

作为全国重要人工智能创新应用先导区，青岛不断完善人工智能产业生态，持续推动人工智能在产品创新、产业集聚、应用赋能、综合支撑等方面的能力提升，积极拓展人工智能算力和算法在生产生活、社会治理等方面的应用场景，取得了积极的成效。

目前全市共有人工智能企业500 余家，在智能家居、智能交通、智能医疗、智能机器人、智能制造和智能硬件等六个应用领域具备先发优势。

本次大会以“新质发展，智控未来”为主题，为自动化相关领域的专家学者提供了展示创新成果，展望未来发展的平台。我们热切期待，通过本次大会与各位嘉宾交流思想，探讨合作，衷心希望大家走进青岛，了解青岛，为青岛发展多提宝贵建议，把更多新项目、新业务、新产品落地青岛，我们将持续打造一流营商环境，提供优质服务，促进自动化产业快速发展，共同创造合作共赢的美好未来！

在 2024 中国自动化大会开幕式上的致辞（内容节选）

文 / 青岛科技大学党委常委、副校长　吕万翔

青岛科技大学党委常委、副校长吕万翔致辞

青岛科技大学自 1950 年建校以来，始终秉承“明德、笃学、弘毅、拓新”的校训，致力于培养高素质人才、推动科技进步、服务社会发展。学校坚持开放办学、强化内涵发展，现已发展成为一所以工为主，理、工、文、经、管、法、教育、艺术等多学科协调发展的高水平大学。学校在智能制造、机器人技术、工业控制系统等领域不断取得重要进展，自动化专业更是历经数十年发展，成为国内知名的品牌专业，为社会培养了大批优秀的技术人才，为我国自动化产业发展贡献了积极力量。

作为中国自动化学会的重要合作伙伴，青岛科技大学很荣幸能够承办这一届全国性的学术盛会。本次大会以“新质发展、智控未来”为主题，旨在促进自动化、信息与智能科学技术的深度融合，助力我国自动化产业创新发展。这不仅是对青岛科技大学办学实力和科研水平的肯定，也是对我们服务国家战略、支撑地方经济发展能力的信任。此次大会的举办，将为我校师生提供一个与国内外专家学者交流学习的宝贵机会，也将进一步推动我校自动化及相关学科的建设与发展。

在此，我衷心希望：一、各位专家学者能够充分利用此次大会，分享最新的研究成果，激发创新的火花，为自动化领域的发展贡献智慧；二、希望通过大会，进一步加强青岛科技大学与各高校、科研机构及企业之间的交流合作，共同推动自动化技术的应用与发展；三、期待各位青年学者和学生们能够从大会中汲取知识与灵感，拓宽视野，激发科研热情，成长为未来自动化领域的栋梁之才；四、祝愿本次大会能够取得丰硕成果，为我国自动化事业的发展增添浓墨重彩的一笔。

破解卡脖子问题，构建新发展格局

文 / 龙芯中科技术股份有限公司董事长　胡伟武

导读：2024 年 11 月 2-3 日，由中国自动化学会主办，青岛科技大学、国家橡胶与轮胎工程技术研究中心承办的 2024 中国自动化大会在青岛召开。会议特别邀请了龙芯中科技术股份有限公司董事长、中国科学院计算所研究员胡伟武作了题为“破解卡脖子问题，构建新发展格局”的报告。报告指出，通过长期积累和迭代，龙芯中科已将自主化优势转化为性能和生态优势，跻身世界领先行列。新一代通用处理器龙芯 3A6000 已正式发布，性能达到市场主流产品水平，标志着龙芯已走出一条基于成熟工艺、通过设计优化提升性能的道路。胡伟武董事长表示，龙架构成为与 X86/ARM 并列的顶层开源生态系统，我国信息产业新发展格局正在加速构建中，我们应将自主进行到底，不依赖国外技术授权和供应链。

发展新质生产力关键在于构建具有自主创新能力的第三套体系。目前，我国信息产业发展受限于美西方主导的 X86 和 ARM 体系，在 X86 体系中，我国主要负责整机生产，而在 ARM 体系中，我国则负责芯片设计，均依赖外部技术和标准。这种依赖限制了我国的自主创新能力，降低了市场竞争力，并增加了安全风险。为打破这种局面，必须通过自主研发和积累知识产权，构建一个独立、创新的体系，既能够与现有体系兼容，又不再过度依赖外部，从而保障产业链的自主可控与安全，推动我国信息产业的发展。

一、掌握根技术，将自主进行到底

2023 年我国工业企业的利润率为 5.8%，而电子工业的利润率仅为 4.2%。同年，苹果公司单独的利润总额为 970 亿美元，远超我国整个电子工业的利润。这一现象的原因在于我国信息产业过度依赖外部技术与体系。习近平总书记对此进行了精准的诊断，指出如果仅在他人的基础上建设自己的产业体系，即使再大再美也无法经得起挑战。为了突破这一瓶颈，习近平总书记提出要加速推进国产自主可控计划，构建安全自主的信息技术体系。

我国在多个领域，特别是在电力、能源、交通等行业，正在推进自主化替代，但这一过程的最终目的是建立我们自己的体系。因此，摆脱传统的生产力发展路径，构建独立于 X86 和 ARM 的全新信息技术体系，成为我们发展的必由之路。正如中国有北斗、商飞等自主技术体系，我们也必须建立属于自己的信息产业体系，从而确保产业安全、可控。

在这一过程中，关键的技术包括 CPU、操作系统等“根技术”，这些构成了信息产业的底座，如图 1 所示。尽管我们在整机制造方面取得了很大进展，但底层技术仍然受制于国外技术授

图 1　信息产业体系根技术

权。例如，国内某些芯片企业无法获得 ARM 的 V9 指令集授权，且 ARM 的最高性能 IP 系列也因制裁原因无法提供给中国，这直接影响了芯片性能。此外，芯片生产所需的光刻机等设备也被禁运，进一步限制了我们在这一领域的自主发展。

为了突破这些瓶颈，我们必须实现指令系统、IP 核以及生产工艺的自主可控。龙芯公司作为自主创新的实践者，已经实现了完全不依赖外部技术授权的目标，展示了我国在这一领域的可行性。然而，尽管自主可控至关重要，仍然面临两个核心问题：一是自主技术的性能是否能与国际先进水平相竞争；二是我国自主的技术生态能否与 ARM 和 X86 相匹敌。

二、龙芯自主 CPU

从 2001 年开始，龙芯公司一直致力于通用 CPU 的性能提升，至今已完成了性能“补课”，如图 2 所示。CPU 的性能可以分为通用性能和专用性能，其中 GPU 和 AI 处理器是专用的，而 CPU 则是通用的，能够应对各种计算任务。相比专用处理器，CPU 具有更大的灵活性，可以用来执行图形计算、AI 算法等任务，尽管如此，单核性能依然是性能提升的核心，尤其在技术相对落后的情况下，单核性能比多核更加关键。

提升性能有两个主要途径：设计优化和工艺迭代。设计优化可以比作战术调整，工艺迭代则相当于武器装备的提升。龙

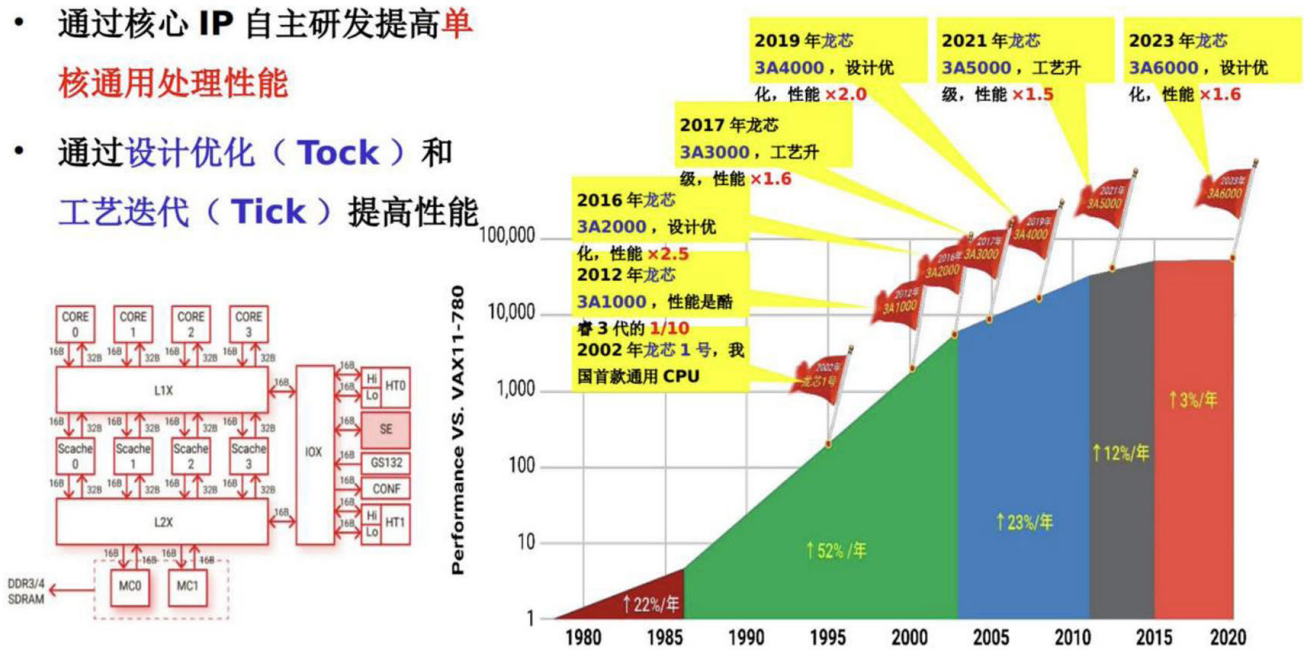


图 2　CPU 性能提升

芯通过不断优化设计弥补工艺差距，逐步提升了性能。2002 年推出的龙芯 1 号与 Intel 的奔 4 相比，其性能仅为 1/20；2012 年推出的 3A1000 性能为 Intel 酷睿第三代的 1/10。到 2016 年，3A2000 的性能相比前代提升了 2.5 倍，2017 年通过工艺调整提升了 1.6 倍，2021 年通过再次升级工艺，3A5000 的性能提升了 1.5 倍。

2023 年推出的 3A6000 达到 Intel 第 10 代酷睿水平，与前代产品 3A5000 相比，单核性能提高了 60%，多核性能也有显著提升，如图 3 所示。这一成绩表明，通过设计优化，即使工艺落后一代，依然可以取得可与国际竞争者相媲美的性能。此外，龙芯还推出了基于成熟工艺的服务器 CPU，虽然工艺没有变，但通用处理性能成倍提升，成本也有所降低。与 Intel 至强 6338 对比，龙芯的 32 核服务器 CPU 在处理能力上也达到了相同水平。

龙芯芯片在每 GHz 性能方面已经接近全球领先水平，根据国内微型计算机的评估，龙芯芯片在性能上预计将超过市场上大部分 PC CPU，可能超越当时市场上 70% 至 80% 的主流产品。尽管龙芯未能采用先进的 7nm 工艺，其通过性能优化策略实现了接近 X86 7nm 芯片的性能，确保了芯片的竞争力，如图 4 所示。

CPU 型号	SPEC CPU 2006				Stream		UnixBench	
	INT	FP	INT_rate	FP_rate	单线程（MB/s）	四线程（MB/s）	单线程（MB/s）	多线程（MB/s）
3A5000	26.6	28.5	76.4	76.7	15842	21969	1718	4224
3A6000	43.1	54.6	155	140	32211	42468	2303	7608

CPU 型号	SPEC CPU 2006				Stream		UnixBench	
	INT	FP	INT_rate	FP_rate	单线程（MB/s）	四线程（MB/s）	单线程	多线程
3A6000	43.1	54.6	155	140	32211	42468	2303	7608
i3-10100	43.6	57.4	129	123	25409	21336	2049	7302

图 3　龙芯 3A6000 桌面 CPU

图 4　龙芯 3B6600 桌面 CPU

图 5　龙架构

在指令系统方面，龙芯自主研发了“龙架构”，如图 5 所示，这一指令系统的构建涉及到 BIOS、内核及编译器的开发，且对硬件和操作系统的兼容性要求极高。指令系统是构建信息产业生态的核心技术之一，换指令系统的难度远高于更换工艺，因此龙芯选择自主研发指令系统，以确保生态系统的自主可控。尽管指令系统的更换具有挑战性，龙芯通过自有指令系统与全球应用的兼容，逐步构建起完整的软件生态。

在软件生态建设方面，龙芯实施了“三位一体”的发展战略：首先，夯实基础，建立与 X86、ARM 体系并列的 Linux 基础软件体系；其次，通过横向编译和二进制翻译技术，解决 X86 和 ARM 应用的迁移问题，实现与主流应用的兼容；最后，推动自主应用的开发，以期建立完整的本土化生态系统。目前，龙芯在 Linux 基础软件体系、国际开源软件支持、以及与商业企业的合作方面取得了显著进展，已经得到主流商业软件的支持。

通过上述举措，龙芯在构建自主生态、兼容国际生态的同时，不断提升其芯片的性能和应用兼容性，逐步解决了与全球主流体系的兼容性问题，并朝着自主应用和产业生态的最终目标迈进。

三、龙芯 CPU 生态面临的挑战

龙芯生态的建设面临多方面的挑战，尤其是在硬件、软件和产业三个核心生态体系的完善与发展方面。首先，硬件生态的构建是其发展的基础之一。与 ARM 架构类似，龙芯 CPU 的推广和应用离不开与之配套的硬件设备的完善。例如，GPU、网卡、显卡等外部硬件的研发与生产，这些硬件必须能够与龙芯的 CPU 平台无缝协作，从而保证系统整体性能的提升和用户体验的优化。随着技术积累和合作的深入，硬件生态的完善将成为支撑龙芯平台广泛应用的关键。

在软件生态方面，龙芯面临的核心挑战之一是操作系统与应用程序在跨平台兼容性上的问题，如图 6 所示。为了解决这一

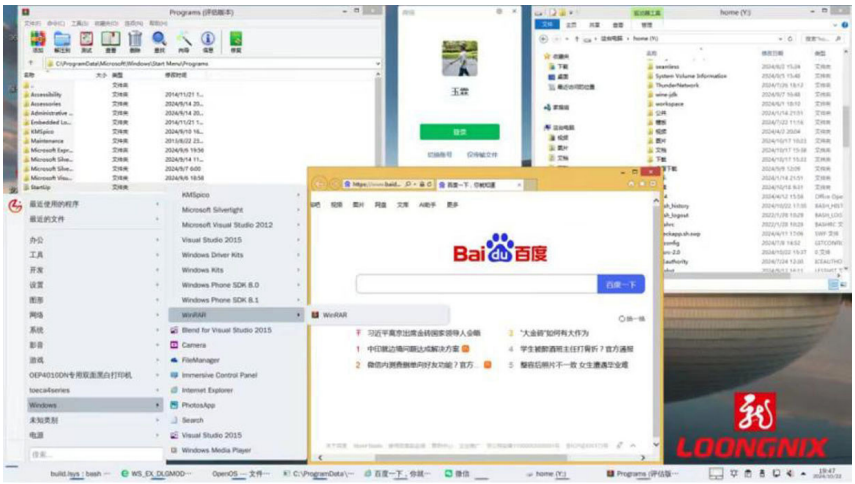


图 6　建成与 X86/ARM 并列的 Linux 基础软件体系

问题，龙芯在 Linux 平台上通过“窗口融合、文件系统融合和 I/O 融合”技术，提升操作系统与应用程序的兼容性。这些技术使得 Linux 平台可以支持 Windows 应用，并通过文件和设备的共享增强用户的跨平台体验。此外，龙芯正在推动自主操作系统的研发，包括基于 Debian 和鸿蒙的定制版本，这些操作系统不仅支持传统桌面应用，还适配了各种工控应用。通过构建应用兼容框架，龙芯可以解决操作系统升级后可能导致的应用不兼容问题，从而为开发者提供一个稳定的开发环境。

为了彻底解决应用与操作系统之间的兼容问题，龙芯还需要构建自主的编程框架，并推动本土应用的开发。该框架将成为操作系统的核心，确保应用程序在操作系统升级后仍能保持兼容性。此外，龙芯致力于通过提供符合本土需求的开发工具和资源，鼓励更多的开发者参与到应用开发中来，最终形成一个以龙芯为核心的自主应用生态。

产业生态的建设是龙芯战略实施的核心之一。除了技术突破，龙芯还需要通过与产业链上下游的合作，推动硬件设备、操作系统和应用程序的融合与创新，如图 7 所示。产业生态不仅涉及技术实现，还需在市场中形成广泛的应用基础。通过政策支持和市场引导，龙芯可以提升其市场份额，形成规模效应。龙芯不仅要强化硬件和软件的协调合作，还需要扩大其在工控、嵌入式系统等细分市场的应用推广，并加强与行业领先企业的合作，通过技术合作与创新推动产业生态的成熟。

在硬件、软件和产业生态的推进过程中，人才的培养和引进至关重要。龙芯需要通过提升国内教育和培训体系，培养更多的专业技术人才，尤其是在 CPU 设计、操作系统开发、应用程序开发等领域。通过完善的培训体系，龙芯不仅能够推动技术的自主创新，还能为产业发展提供源源不断的技术支持。在政策引导和技术创新的双重推动下，龙芯通过逐步完善这些生态体系，致力于形成一个自主、可持续发展的技术平台，并在全球竞争中占据一席之地。

四、龙芯 CPU 在工业领域应用案例

龙芯在工业领域的应用已经取得了显著进展，特别是在高性能桌面、服务器以及工控等多个领域。龙芯的产品覆盖了三个主要系列：3 号系列主要面向桌面和服务器，2 号系列则专注于工控和终端，而 MCU 系列则提供了价格较低的芯片，最低仅需一两块钱。这些产品具有广泛的应用场景，且由于龙芯的指令系统不依赖于境外授权，避免了地域和领域限制，使得龙芯在全球范围内具有更大的灵活性。

龙芯的技术在工业界的应用已覆盖多个关键领域。首先，龙芯在工控领域已经得到广泛应用，特别是在能源和交通等重要行业。龙芯 CPU 在能源领域的应用案例尤为突出，包括水电、风电等发电系统的使用，尤其是在变电和配电领域，已经开始实现批量应用。此外，龙芯还在高铁、地铁和公路等交通领域开展了自主化

图 7　产业生态建设

图 8　龙芯 CPU 系列

图 9　龙芯 CPU 在交通领域应用案例

研发和试点工作，涉及人机交互、车载网关、安全计算平台等多个关键技术领域，如图 9 所示。尤其是在地铁系统中，青岛地铁和其他城市的地铁线路都已开始使用基于龙芯 CPU 的 PLC 系统。

龙芯在自主先进制造领域的应用也正在逐步增加，涉及到数控、加工、PLC 等多个领域，并

图 10　龙芯 CPU 在先进制造领域应用案例

且已有数家公司在国内使用龙芯进行 PLC 的开发和应用，如图 10 所示。

在未来，龙芯的目标是建立一个独立的、自主可控的信息技术体系。与X86和ARM形成“三足鼎立”的局面，构建一个完全自主的信息化格局。这条道路虽然充满挑战，但龙芯团队对于未来充满信心。尽管目前仍面临诸多困难和阻力，但只要坚定不移地走下去，最终将克服所有困难，实现第三套自主信息技术体系的建设。

* 本文根据作者所作报告速记整理而成

作者简介

胡伟武，龙芯中科技术股份有限公司董事长、总经理，工学博士、中科院计算所研究员、博士生导师，第十一届全国人大代表，党的十八大、十九大、二十大代表。胡伟武先生是龙芯总设计师，从 2001 年起投身于龙芯处理器的研制工作。先后主持完成了龙芯 1 号、龙芯 2 号、龙芯 3 号系列处理器的研制，改写我国信息产业无芯历史。胡伟武博士作为首席科学家全面主持龙芯中科的技术工作，组织规划龙芯研发线路图，组织领导新产品的研发。承担国家项目 10 余项，发表论文 50 余篇，撰写 3 本专著。曾荣获“北京市突出贡献中关村奖”“CCF 王选奖”“全国五四青年奖章”“中国科学院杰出成就奖”“第五届光华工程科技奖”“全国首届优秀博士论文奖”等奖项。

轮胎行业智能制造与精益管控

文 / 国家橡胶与轮胎工程技术研究中心副主任　官炳政

导读：2024 年 11 月 2-3 日，由中国自动化学会主办，青岛科技大学、国家橡胶与轮胎工程技术研究中心承办的 2024 中国自动化大会在青岛召开。会议特别邀请了国家橡胶与轮胎工程技术研究中心副主任、软控股份有限公司董事长兼总裁官炳政作了题为“轮胎行业智能制造与精益管控”的报告。报告指出，轮胎作为安全产品，轮胎质量影响到交通出行安全，中国轮胎产量占到全球的 50% 以上，是典型的轮胎制造大国。轮胎制造过程为流程 + 离散的混合型制造模式，面临完备信息、机理经验、质量控制等难题，对高质量、高效能、高一致制造轮胎产品提出巨大的挑战。针对这些问题，官炳政董事长分享了轮胎智能制造的进阶之路与轮胎精益管控的实践探索等内容，他在总结中提出，中国橡胶轮胎正处于由跟跑、并跑到领跑，由大向强的发展过程，希望产业出题，学界答题，共同做好轮胎产业这篇文章，让中国轮胎服务好国家大战略、大工程，服务好每一个家庭。

众所周知，在现代文明中，通信和交通是两个不可或缺的关键因素。在交通领域，轮胎作为汽车的重要部件，其制造过程复杂且精密，甚至轮胎是否能够做到完美的圆形，都是一个极具挑战性的课题。

一、轮胎行业典型特征

轮胎行业在国家的重大战略和重大工程实践中发挥着重要作用。以橡胶材料为例，它在空间站建设、太空服研制、神舟飞船推进剂以及蛟龙号等项目中，发挥了不可替代的作用，在这些高精尖领域，橡胶材料和密封件是关键技术之一，特别是在航天与深海探测过程中，橡胶部件的应用至关重要。

在陆地上，轮胎同样服务于军工装备，例如装甲车、飞机以及特种车辆等，如图 1 所示。这些装备在战略核心任务中的作用不可或缺，正是轮胎使得这些设备能够顺利运行，缺乏轮胎的支持，这些装备将寸步难行。

除了其在战略意义上的重要性，轮胎还是一种关乎安全的重要产品。日常出行中，汽车轮胎是唯一与地面接触的部件，它承载着驾驶者及乘客的安全与幸福。

图 1　轮胎的广泛应用

因此，轮胎的质量与性能直接关系到每个人的生命安全。

随着中国新能源战略的推进，中国新能源汽车市场份额已经达到全球 64.5%。新能源乘用车的市场渗透率突破了 53.3%，如图 2 所示，这一切都推动了轮胎产业的快速发展。根据权威机构的预测，到 2029 年，全球轮胎市场的规模将达到 1.6 万亿人民币。中国轮胎产业在这一庞大市场中占据了重要地位，生产量超过全球 50% 的份额，并且在全球轮胎行业的 75 强企业中，中国企业占据了 38 家，中国已经成为了轮胎产业的全球领导者。

尽管中国轮胎产业在全球市场中占据了重要地位，但从品牌影响力和市场销售来看，仍然面临着“大而不强”的挑战。首先，特种轮胎领域，如航空轮胎和巨型轮胎等，其先进制造工艺仍然被国外企业垄断，这成为了“卡脖子”问题。中国国产轮胎在品质一致性方面存在差距，质量水平参差不齐。中国轮胎品牌的国际影响力较为薄弱，尚未能够在全球市场上形成强有力的竞争力。

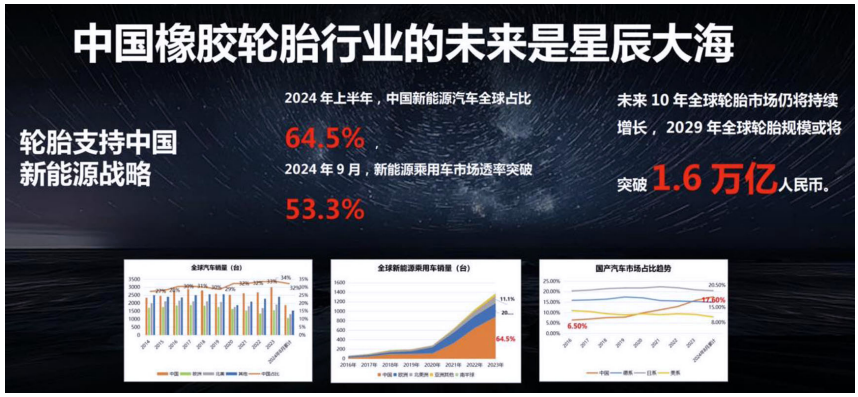


图 2　中国橡胶轮胎行业

图 3　国家橡胶与轮胎工程技术研究中心

其次，尽管中国企业在全球轮胎行业的 75 强中占据了 38 席，但其销售收入占比仅为 18%。从前 10 强企业来看，中国企业仅有两席，分别是中策橡胶和赛轮轮胎。因此，中国轮胎企业正面临着从“大”到“强”的历史性跨越阶段。

为了应对这一挑战，国家橡胶与轮胎工程技术研究中心应运而生，致力于解决轮胎行业中的“卡脖子”技术问题，并推动产业化创新。该中心由青岛科技大学、软控股份和赛轮集团共同承建，如图 3 所示。软控股份和赛轮集团均由青岛科技大学孵化，青岛科技大学被誉为“橡胶界的黄埔军校”，在橡胶领域具有重要地位，是亚洲规模最大、门类最齐全的橡胶类院校。

赛轮集团目前其产值规模位居全国第二，是国家首家智能制造试点示范企业；软控股份为全球领先的智能制造工厂整体解决方案供应商，其规模多年位居全球第一。

国家橡胶与轮胎工程技术研究中心自 2009 年由科技部批复以来，在推动中国轮胎产业技术创新和产业升级方面发挥了关键作用。

二、轮胎智能制造进阶之路

国家橡胶与轮胎工程技术研究中心积极推进轮胎产业的智能

化转型升级，探索产业从小到大、再由大到强的过程，分为三个主要阶段，如图 4 所示。首先，装备自动化阶段通过攻关核心工艺和装备，实现工厂的自动化，解决了部分制造效率和质量问题，重点是通过机器替代人工。接着进入全流程协同阶段，推动工艺、装备、材料等环节的上下游协同，虽然取得了一定成效，但制造一致性问题仍需解决。最后，进入全要素精益管控阶段，依托智能技术提升核心质控点的管理水平，实现生产过程的精细化控制，进一步提升质量一致性。

在装备制造的自动化进程中，第一阶段主要集中在物流领域，特别是通过现代化的物流方式替代传统的人工搬运工作，如图 5 所示。在轮胎制造过程中，胶块和胶片等材料的搬运和处理往往存在较高的劳动强度。自 2015 年以来，行业推动了“机器换人”计划，显著提升了物流效率和减少了人工劳动的重复性。

第二阶段的自动化目标是促进机器与工艺的协同，机器人技术与轮胎制造过程的节拍紧密结合，实现高效、零延迟的制造，如图 6 所示。在多品种、小批量生产的背景下，快速换模和智能控制技术的应用，使得生产过程中规格切换的时间缩短了 50% 以上，显著提升了生产的灵活性和效率。

图 4　中国轮胎智能制造进阶之路

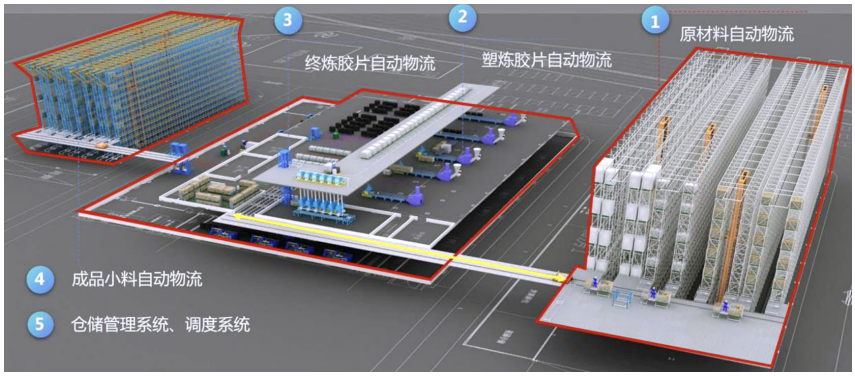


图 5　机器换人的物流自动化

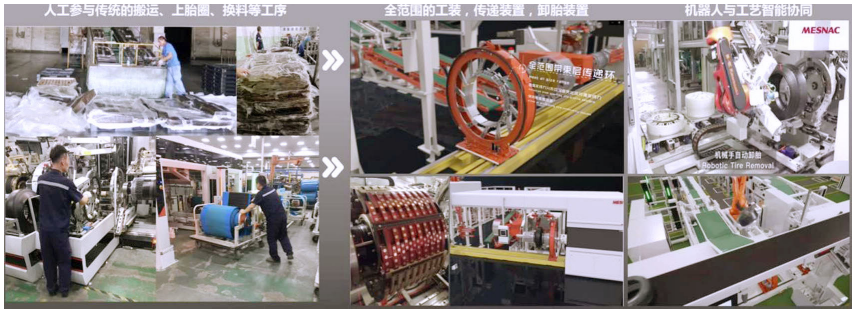


图 6　机器与工艺协同自动化

第三阶段则是将自动化与物料特性相结合，尤其是针对橡胶这种具有弹性体特性的材料，如图 7 所示。在轮胎生产过程中，通过控制算法的应用，解决了制造过程中的卡点和技术难题。例如，通过张力控制技术，确保在轮胎的贴合、卷曲和输送过程中，材料的变形保持在可控范围内。国橡中心针对这一问题进行了深入的研究，采用 PID 控制和模型预测控制等方法，在应对材料变化、生产速度调整以及未知扰动时，能够做出及时反应，优化控

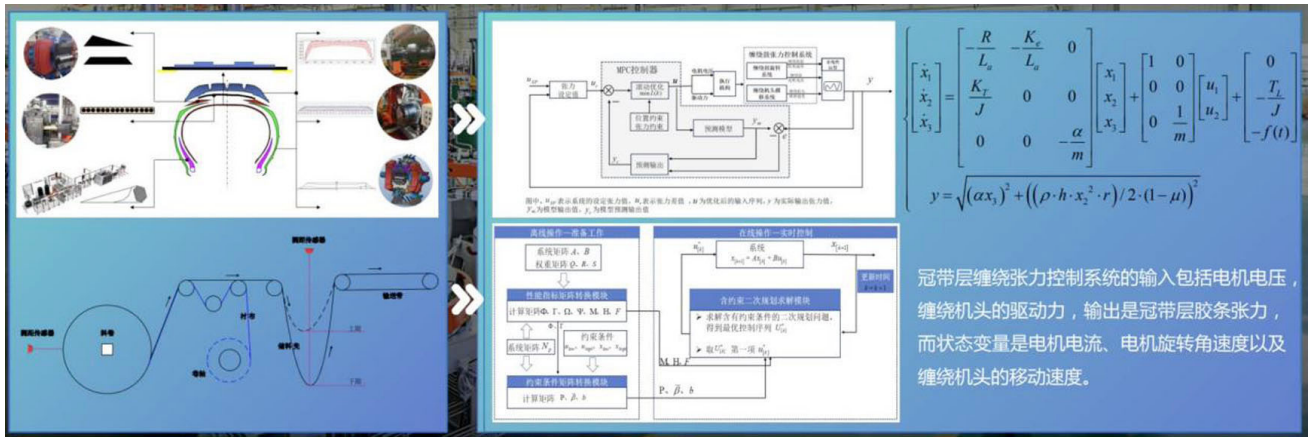


图 7　智能算法驱动设备控制自动化

制策略。

在装备自动化方面，国橡中心取得了显著的技术突破，涵盖了关键装备和技术的创新。例如，轮胎行业首个由中国企业主导的轮胎用 RFID 四项国际标准，联合国际领先企业如米其林、普利斯通等共同制定。此外，半钢子午线轮胎成型机获得了工业和信息化部制造业单项冠军的荣誉，而全自动称量系统则获得了国家科技进步二等奖。

国橡中心经过十年的研究，发明了被誉为“液体黄金”的新型橡胶材料，如图 8 所示，该材料显著提升了轮胎的耐磨性能、抗湿滑性能和滚动阻力性能。尤其是在滚动阻力和耐磨性之间的矛盾中，采用该材料打破了传统的“魔鬼三角”定律，达到国际先进水平。

为了支撑这些技术创新，国橡中心推动了自动化装备的进一步发展，强调数字化平台驱动下的制造高质量。通过模块化设计和软件自动化，提升了生产过程中设备的一致性和可靠性。工业视觉技术的应用，也有效解决了人工检查过程中可能出现的误差和疲劳问题，实现了轮胎生产的高精度检测，准确率超过 99.7%。

此外，国橡中心还开发了自主的工业软件 MCC，如图 9 所示，它优化了设备的操作系统，增强

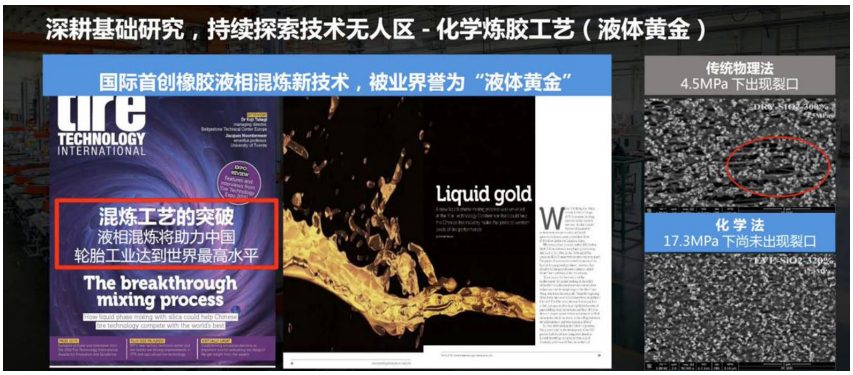


图 8　创新制造工艺

图 9　工业视觉提升智造一致性

了设备间的互联互通，并通过知识图谱技术实现了设备的智能运维。这一系列创新举措，不仅推动了轮胎制造过程中的技术进步，也为行业的高质量发展奠定了基础。

尽管如此，行业内仍然存在一些技术难题，需要通过专家的共同努力进行攻关。国橡中心也积极探索工业大模型在轮胎制造中的应用，致力于进一步提升制造过程中的一致性和精细管控，以推动中国轮胎行业从“跟跑”和“并跑”向“领跑”迈进。

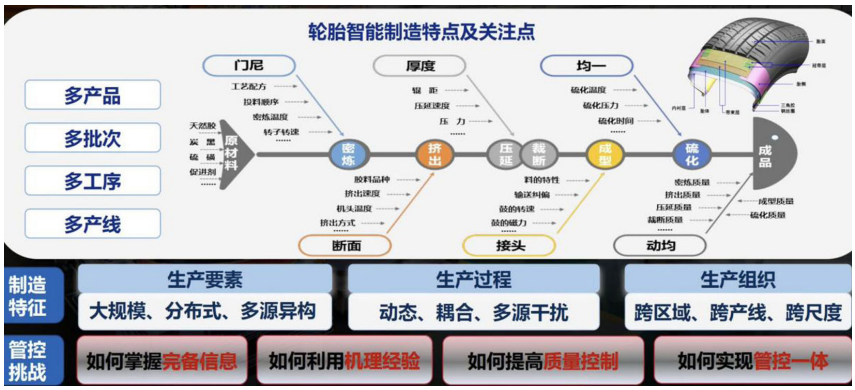


图 10　轮胎智能制造特点

三、轮胎精益管控实践探索

在轮胎制造过程中，涉及五大关键工序和 1100 个质量控制点，体现了其生产过程的高度复杂性，如图 10 所示。首先，轮胎制造采用大规模生产模式，涵盖原材料、半成品、成品的物理和化学转化过程，其中每个工序之间的连续性至关重要。其次，由于轮胎材料主要为弹性体，其生产过程中涉及粉料、固态颗粒、流体以及非牛顿流体等多种物态的精准控制。生产过程中，半成品和成品的质量要求尤为严格。

轮胎作为安全产品，其批量生产具有高度一致性要求，尤其是在大规模生产中，每批次的质量误差需要控制在极小范围内，并且要求具有较强的可追溯性。此外，面对小批量、多样化定制生产模式的需求，如何有效应对这些小批量、多样化的生产任务也成为一大挑战。因此，轮胎制造业需要通过高科技手段实现“流程制造”和“离散制造”的全价值链精益管理，从而提高生产过程的一致性、效率和质量。这一过程中，面临几个核心难题。

首先，数据价值的深度挖掘是提升质量预测准确性的重要方向，如图 11 所示。然而，轮胎制造过程中异常数据的缺乏，导致了过程参数间的低相关性和生产工艺参数的复杂性。因此，如何在小样本数据中识别出大数据规律，成为解决问题的关键。为此，我们正在联合数学、算法控制及

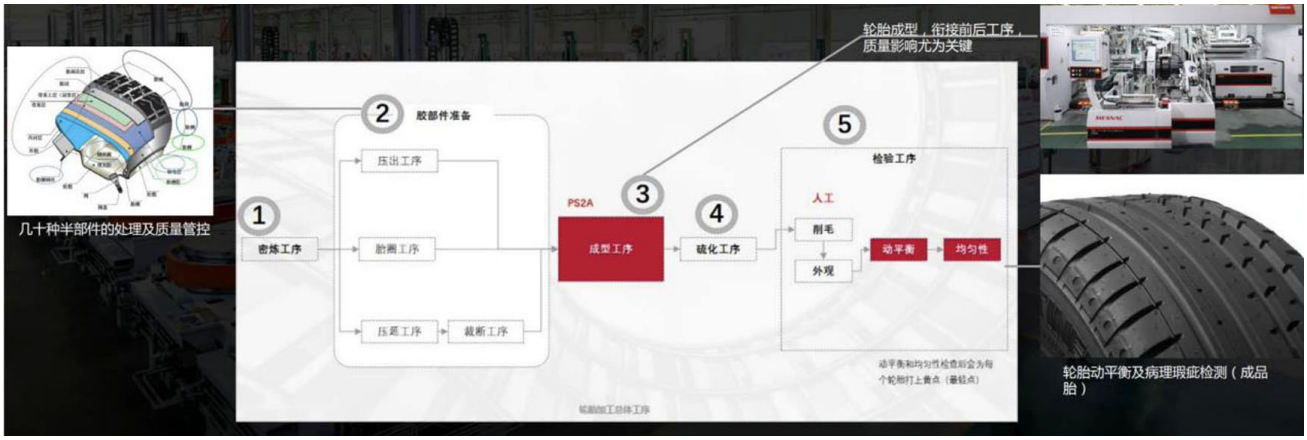


图 11　轮胎智造数据价值

图 12　工业视觉大模型

软件领域的专家，共同攻克这些技术难题，特别是在基于大数据的工艺过程优化，尤其是质量预测方面，已展开相关研究。

第二，工业视觉与大模型的融合，特别是图像识别技术的应用，成为提升自动化检测水平的关键，如图 12 所示。在传统轮胎检测中，大量人工劳动和时间投入未能有效确保最终产品的质量。为了克服这一难题，基于大数据技术的自动判级系统已在研发中。通过图像识别与大数据模型的结合，实现更为精准、高效的质量检测。

通过上述科技创新，中国橡胶轮胎产业将实现跨越式发展，从“跟跑”到“并跑”再到“领跑”，提升产业的整体竞争力，并为全球市场贡献中国智慧。在这一过程中，学界与产业界的紧密合作将是推动技术进步和产业升级的关键。最终目标是通过技术创新，为国家战略和社会需求提供强有力的支持，助力中国轮胎行业在全球占据领先地位。国家橡胶与轮胎工程技术研究中心将积极承担国家使命，凝聚行业智慧，推动中国橡胶工业向世界领先水平迈进！

* 本文根据作者所作报告速记整理而成

作者简介

官炳政，国家橡胶与轮胎工程技术研究中心副主任，软控股份有限公司董事长兼总裁，中国橡胶工业协会主席团主席、山东省橡胶行业协会副会长、青岛市人大代表、青岛拔尖人才、全国石油和化工优秀科技工作者。专利授权 61 项，作为项目负责人在研山东省重点研发计划 2 项，青岛市重点研发计划 1 项，参与国家自然科学基金重大研究计划 1 项，编制国家标准 3 项，团体标准 1 项。

张钹院士：具身智能——人工智能与机器人发展的重要推力

具身智能是指一种思想或者理论。它强调认知受智能体的感知与动作的影响，主张智能或智能行为来自智能体身体与周围环境的交互，而不仅仅是大脑的功能。这种思想与建立在笛卡尔身心二元论基础上的离身智能不同，后者没有考虑大脑与身体之间的联系。具身智能理论对认知科学、心理学、人类学、艺术等都产生了深刻的影响，由此诞生了具身认知、具身心理学等分支学科。当前，具身智能理论深入到大众关注的人工智能、机器人技术等领域，并产生重大影响。

推动人工智能迈向更高水平的通用人工智能

早期人工智能研究者认为，机器可以通过对符号的操作实现对人类思考（理性行为）的模拟，即让机器像人类那样思考，无需考虑身体和环境。他们把人工智能定义为模拟人类的智能行为，除了模拟思考之外，还包括感知、动作，以及情感与灵感等。由于他们认为思考和身体、环境是分离的，因此在模拟这些行为时，是各自独立进行，没有考虑行为之间的相互联系。人工智能的另一学派内在主义，主张人工智能应该模拟人类大脑的工作原理，我们通常称之为类脑计算。这个学派也没有考虑思考、感知与身体的关系。因此传统人工智能中的两大学派都属于离身智能。

具身智能则认为人类的智能行为不仅依赖大脑的思考，还需要通过与物理环境的交互来进行学习和决策，最后通过感知将执行的结果反馈到大脑，形成闭环，构成一个智能体，这样才能真正实现模拟人类智能行为的目标。具身智能这种思想的提出，与传统人工智能不同，将推动人工智能向更高的水平——通用人工智能发展。

人工智能的发展历史，经历过三个阶段。第一代人工智能提出以知识与经验为基础的推理模型，通过这一模型来模拟人类的思考。这一模型具有可解释性与可理解性的优点，但由于知识表示与获取的困难，该模型在应用和产业化上受到很大限制。第二代人工智能提出数据驱动的模型，通过基于大数据的机器学习，实现对人类感性和情感行为的模拟。由于模型的图像、语音等输入数据来自客观世界，因此具有很大的应用潜力，但由于模型本身存在不安全、不可靠、不可信（不可解释）、不可控和不易推广等缺陷，因此应用范围有限。虽然科学家在这期间也进行过模拟人类动作，比如手的操作与脚的步行等研究，但没有将这些研究与思考、感知等其他方面联系起来。总之，这两代人工智能都有很大的局限性，而且只针对特定领域、利用特定模型去解决特定任务。

2020 年大语言模型的出现，把人工智能推向新的发展阶段——第三代人工智能。大语言模型具有强大的语言生成能力，使机器能够在开放领域下实现与人类的自由交互，这表明机器已经完全掌握了人类的语言。此事意义重大。正如哲学家维特根斯坦所言：“我的语言界限，就是我的世界界限。”机器一旦掌握了人类的语言，也就在某种程度上理解了人类的世界。大语言模型向通用人工智能迈出了关键一步。

尽管人工智能在语言生成上取得成功，在许多其他任务上还不能实现领域的通用性。比如医疗诊断，目前我们还没有开发出适用于各种疾病的计算机医疗诊断系统。另外，大语言模型只会说（生成语言），不会干（行动）。还有大量的任务，特别是复杂的任务，人工智能目前还无法完成。我们离真正的通用人工智能仍有一定距离，其原因就在于，传统人工智能对人

类（大脑）的思考、感知与动作等方面的模拟是相互分开的，而且与环境隔离。具身智能则强调物理身体、环境感知与反馈的重要性，并通过它们实现与外部世界的交互。这正好弥补了传统人工智能的不足，使机器通过与环境的反复交互，逐渐学习适应环境并优化其决策和行动，不断迭代，完成更多、更复杂的任务，这就是新的学习范式——强化学习。由此可见，具身智能将使人工智能在更广泛的环境下，完成更多的任务，向通用人工智能迈进。

推动机器人领域三方面的技术进步

机器人的研究远在人工智能诞生之前。1954 年，科学家发明数字控制可编程的机械臂，为现代工业机器人打下基础，也预示着现代机器人的真正诞生。在人工智能的发展过程中，虽然也将模仿人类动作纳入研究范围，但由于当时研究的重点是思考与感知，动作（身体）并没有受到很大关注。后来，机器人研究慢慢形成一个新领域——机器人学，与人工智能中的动作研究同时存在。机器人学研究的范围更宽，也更多关注技术、实际应用和产业化。1986 年，移动机器人的先驱罗德尼·布鲁克斯提出了行为主义机器人学的主张，他认为传统的符号主义人工智能过于依赖复杂的推理系统，忽视了身体与环境的互动对智能行为的关键性贡献。可以说，这是在机器人领域首次提出具身智能思想。

总的来看，具身智能推动了机器人领域三方面的技术进步。

其一是推动机器人技术从单一感知（如视觉或触觉）向多模态感知发展。自动驾驶汽车需要在各种气候条件和复杂环境下行驶，全天候的环境感知极为重要。目前除了利用单一的摄像机之外，人们更多地考虑多传感器融合，即除摄像机之外，还要综合利用激光扫描仪、雷达、声呐、卫星定位等不同传感器，以保证感知的可靠性。同样，对机械手来讲，也需要利用触觉、力觉、力矩觉等多传感器的融合，以便在复杂环境下实现复杂操作，比如，装配、抢险救灾、排雷等。

其二是推动机器人技术更加关注软件与硬件的高度整合。为了适应不同工作任务的需求，我们需要设计不同类型的机器人，如为家庭服务的人形机器人，拥有灵巧双手的装配机器人，拥有特殊功效的仿生机器人等。不同类型的机器人需要不同的控制技术，也会导致不同的学习和感知方式，这反过来又影响机器人形态的设计。这些都属于机器人软硬件协同的问题。

其三，具身智能中的智能体概念对机器人技术也将产生重要影响。根据智能体这一概念，机器人将思考、感知和动作连为一体，既可以在数字世界中也可以在物理世界中完成任务，利用数字世界又可以对物理世界进行仿真，这些对推动机器人的发展意义重大。比如，无人车的实验如果在物理世界中进行，就要受到气候条件、突发事件等诸多因素限制，利用仿真模拟则会大大提高效率，节省时间与费用。利用智能体理论，我们可以让机器人相互之间进行博弈，实现机器的自我进化（迭代），使机器人的性能不断提高。

来源：人民日报

人物简介

张钹，中国科学院院士、中国自动化学会会士，清华大学教授计算机科学与技术专家，俄罗斯自然科学院外籍院士，清华大学人工智能研究院院长、教授、博士生导师。张钹从事人工智能理论、人工神经网络、遗传算法、分形和小波等理论研究，提出了问题求解的商空间理论，曾获国家科技进步三等奖。

高文院士：数字技术“底座”支撑低空经济高质量发展

低空经济的发展离不开数字基础设施的支持。中国工程院院士、鹏城实验室主任高文在采访中表示，从技术上看，低空网络还需要攻克很多难关，如果没有真正的低空飞行基础设施做保障，只用现在的4G、5G去做，离低空经济高质量发展目标差距很大，我们要补上这个技术短板。

前瞻布局低空基础设施

业界普遍认同，稳定、可靠的通信和感知网络是低空经济顺利运行的前提条件。

高文提出，推动低空经济高质量发展，要筑牢三个“底座”，即政策“底座”、数字技术“底座”和服务“底座”。其中，数字技术“底座”是低空经济能不能顺畅做起来的一个核心因素。

“仅依靠4G、5G网络还不足以支撑低空经济发展。”高文认为，低空网络除了要解决通信问题外，还要重点解决感知问题。所以它既要有网络通信功能，也一定要有感知能力，即通感都要有，这是一体的。

“现有的4G、5G通信网，主要覆盖的是地面和楼宇，空中并不是它的主要覆盖区域。比如，规划空中航道，到底是不是每一个点的信号都能满足要求，或者强度够不够、带宽够不够，但现在这些方面其实还是缺乏的。”高文说。

由于4G、5G网络缺乏必要的感知能力，无法有效监控空域内的活动。高文认为，低空网络的发展需要具备感知能力，能够识别和监控空域内的无人机和其他飞行器，这是当前低空经济发展的主要技术瓶颈，也是需要尽快解决的紧迫问题。

当前，针对低空通信感知网络已经取得一定突破。鹏城实验室构建的低空经济关键技术验证平台，可实现米级颗粒度的低空环境重构，具备1km以上感知距离和不低于50Mbps边缘通信速率。该平台可实现可靠、鲁棒、安全的低空立体网络覆盖，提高低空环境下飞行器的监控与管理水平。

“真正做好低空网络，其实还有很多课要补。”在高文看来，低空网络还处在初级阶段，而算力

图为首个全面自主可控的E级智能算力平台“鹏城云脑”网络智能重大科技基础设施外景图

网其实能帮助低空网络解决遇到的一些问题。

算力网将支持低空经济发展

在数字经济时代，算力网作为一项前沿性技术，对低空经济的发展将发挥支撑作用。

虽然无人机等飞行器和算力并不直接相关，但它们所依赖的低空网与算力关联紧密。“低空经济对算力也有需求，比如说，物流配送。未来低空经济的应用，它需要云计算支持，要非常强，非常准确，而且要实时。”高文说，保障低空畅通安全、实时洞悉，离不开算力网的支撑。

高文拿自动驾驶做类比。他说，由于路况随时变化，信号指示也在变化，所以自动驾驶需要依靠 AI 大模型训练，每天实时更新。而低空定位道理也一样，空中也会遇到新的建筑，航路调度管理也随时变化，这就需要通过算力网络的支撑，确保低空对公业务的一些数据能够及时获得准确回传。

高文预测，算力网对低空经济来说，开始先面向 ToB 的场景，然后慢慢扩大到对大客户的 ToC 业务也会支持。只要需要算力的地方，就可以通过算力网慢慢进行支撑。

“算力网本身就是两个关键词，一个是算力，一个是网，两者缺一不可。”高文认为，首先要有足够的算力，这些算力可能是众多算力的集合，也可能是单个算力。网络是算力之间的通信网络，或者是算力和最终用户之间的网络，它必须保证实时、安全、可靠。“算力网络和电力网络逻辑一致，但算力网要发展到电网那样，可能需要几十年的时间。”高文说。

目前，已经形成超算、智算和通用算力三个关键领域。高文认为，算力的供给需要根据市场变动进行及时调整，算力规模也会随着需求的变动而变动。

“通过算力网，将推动国家网络基础设施不断提升，算力通过网络传输，把数据送过去，把结果送回来。”高文认为，随着“东数西算”战略布局推进，算力网在推动产业转型升级方面将发挥重要作用。

来源：新华网

作者简介

高文，中国工程院院士，中国自动化学会会士、北京大学博雅讲座教授。1982 年于哈科大获得学士学位，1985 年于哈工大获得硕士学位，1988 年和 1991 分别获得哈工大计算机应用博士学位和东京大学电子工程博士学位。1991 至 1995 年就职于哈尔滨工业大学，1996 至 2005 就职于中国科学院计算技术研究所，2006 年至今就职于北京大学。IEEE Fellow、ACM Fellow、中国工程院院士。他的研究领域为多媒体和计算机视觉，包括视频编码、视频分析、多媒体检索、人脸识别、多模态接口和虚拟现实。他最常被引用的工作是基于模型的视频编码与基于特征的对象表达。他先后出版著作七本，发表 280 余篇期刊论文、700 余篇国际会议论文。先后多次获得国家科技进步奖、国家技术发明奖、国家自然科学奖等学术奖励。

动态系统的实时安全性评估技术

文 / 何　潇　刘泽夷　胡嵩乔　刘　畅　周东华

摘要：动态系统的实时安全性评估在防止潜在安全事故导致重大损失方面发挥着关键作用。随着系统功能和复杂性的日益增加，实时安全性评估技术面临着更大的挑战。该文阐述了动态系统实时安全性评估的概念定义，从环境的平稳性及评估模型的构建方式两个维度出发提出了一种分类框架，给出了相应的问题描述，较系统地回顾了动态系统实时安全性评估技术的现有进展，讨论了针对不同实际系统的部署策略，分析了现有技术的发展趋势，探讨了实时安全性评估中亟待解决的问题与未来的发展方向。

关键词：动态系统，实时安全性评估，安全事故，非平稳环境

动态系统是指一类状态随着时间推移而不断变化的模型或系统。随着现代科学技术的蓬勃发展，各类实际系统的复杂程度也随之不断提升。在工业 4.0 愿景下，高度复杂的动态系统往往具有部件耦合、相互影响等特点，对相关运行与维护技术提出了更高的要求。此类系统一旦发生事故，可能会导致巨大的财产损失和人员伤亡。因此，提高动态系统的安全性对于减少重大事故的发生具有重要意义。根据美国国防部所制订的第五代安全系统计划标图 1 描述了不同学术概念所对应的示意场景，以无人机飞行一段时间后去执行任务的场景为例准 MIL-STD-882E，安全性可被定义为：免除导致死亡、伤害、职业疾病、设备或财产损失，或对环境造成损害的条件。在我国《GJB900A-2012 装备安全性工作通用要求》中，安全性这一概念则被定义为：不导致人员伤亡、装备损坏、财产损失或不危及人员健康和环境的能力。该要求对系统安全性的标准进行了正式定义，为我国系统安全性研究工作提供了指引方向。而依据国际电工委员会（IEC）制作的电工百科（Electropedia）中针对控制技术领域的定义，安全性（Safety）则被描述为：免除对所考虑的外部功能和物理单元的不可接受的风险。而国际自动控制联合会（International federation ofautomatic control，IFAC）故障检测、监督和安全性技术委员会所采纳的安全性定义为：系统不对人员、设备或环境造成危险的能力，进一步突出了在描述动态系统安全性时考虑系统本身与外部单位之间交互关系的重要意义。

图 1 描述了不同学术概念所对应的示意场景，以无人机飞行一段时间后去执行任务的场景为例。若无人机可如图 1（a）所示在预期时间内从起飞点成功到达预期落点，且在运行过程中对周围人、环境及财产造成潜在安全威胁的能力较弱，该无人机系统具有高的安全性。相比之下，如图 1（b）所示，若无人机虽能按照预期轨迹在预期时间内从起飞点到达预期落点，但在运行过程中对周围人、环境及财产造成潜在的安全威胁的能力较大（如经过人口密集的城市，无人机在城市上空坠落的影响极大），则该无人机系统被视为具有低安全性。表 1 总结了几类相关学术概念的描述。参考以上标准，动态系统

的安全性在本文中可被定义为：动态系统在运行过程中不对人、财产及环境造成危害的能力。

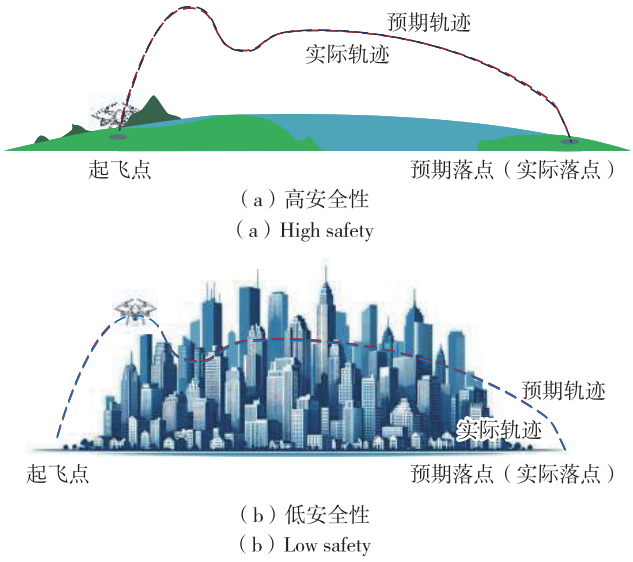


图1　不同学术概念对应的示意场景：以无人机飞行过程为例

自20世纪40年代提出系统安全的概念以来，安全性评估相关问题得到了国内外学者的广泛关注，针对不同工业领域中相关问题出现了不同的技术路线。自20世纪50年代起，美国研究人员开始尝试将故障模式及影响分析（Failure mode and effect analysis，FMEA）方法应用于航天系统的研发工作中。随着系统工程技术的不断发展，该方法被推广至不同的应用领域。在此基础上，研究者又提出并发展了确定性安全性评估（Deterministic safety assessment，DSA）与以概率论为基础的概率安全性评估（Probabilistic safety assessment，PSA）方法。从方法的视角看，采用包含初步危险分析（Preliminary hazard analysis，PHA）、故障树分析（Fault tree analysis，FTA）、事件树分析（Event tree analysis，ETA）、因果分析图法及运行危险分析（Hazard and operability analysis，HAZOP）等，试图发现系统设计的薄弱环节。Temp_Color 与此同时，部分研究者尝试将人的行为和习惯所致使的潜在失误纳入考虑范围，借助人为操作的可靠性分析技术降低安全隐患，以此形成了一套人因安全性（Human factors in safety，HFS）方法。上述研究已在核能、航空航天、化工及医疗等领域得到了广泛地应用，并在事故预防及保障维护人员安全等方面做出了举足轻重的贡献。

基于传统概率安全性评估框架的安全性评估技术通常无法捕捉系统状态随时间变化的动态性，需要大量的离线历史数据来估算概率和潜在事件的频率。因此，对于当下日趋复杂的复杂系统，概率安全性评估方法的分析将会变得非常困难。在过去几十年中，基于概率框架的安全性评估方法被陆续提出。持续性概率安全性评估（Living probabilisticsafety assessment，LPSA）是其中早期的研究工作。此类方法通过更新来反映系统的变化及当前整体系统相应的安全状态。在此基础上，研究者进一步提出了一类基于事故序列前兆（Accidentsequence precursor，ASP）数据的动态安全性评估（Dynamic safety assessment，DSA），可根据一定时间内获得的关于系统组件状态的知识及对危害情景的仿真结果，更新对恶化系统风险的估计，并依此捕捉系统风险概况的时间依赖行为。

表1　几类相关概念在国家军用系列标准下的描述及主要侧重角度

概念	描述	主要侧重	参考来源
安全性	不导致人员伤亡、装备损坏、财产损失或不危及人员健康和环境的能力	系统状态、环境状态	
可靠性	在特定时间段内，系统不发生故障且能够按要求执行任务的能力	系统状态	
测试性	具备及时准确地确定系统状态，并隔离其内部故障的设计特性	系统状态	

然而，设备状态反馈的实时性需求的增加、对未知危险情景认知的缺乏、以及事件关联性及级联效应增强等现实问题愈发明显。对于具有强耦合、非线性、时变性及高不确定性特点的复杂动态系统而言，迫切需要研究实时安全性评估技术，将系统实时状态监测数据及对人、财产与环境的潜在影响考虑进去，以对系统安全性状态进行及时分析与反馈。通常而言，动态系统的实时安全性评估技术在实时场景下利用对系统特征的在线监测，在危险发生早期能够对安全威胁进行快速响应，给出系统实时安全性状态的描述，从而降低事故风险、提高生产效率，以及保护人员、环境及财产的安全。经典安全性评估方法与实时安全性评估方法在几个重要维度上的区别与联系如表 2 所示。

表 2　经典安全性评估方法与实时安全性评估方法的概念辨析

	经典安全性评估	实时安全性评估
先验知识需求	较多	较少
计算资源需求	较少	较多
主体适用对象	系统级	部件级
主要侧重阶段	方案设计阶段	使用保障阶段
典型应用领域	核能、航空航天	电力、工程结构
现有理论成果	较多	较少
应用价值	较高	较高

本文首先对动态系统的实时安全性评估方法进行概述，明确其概念定义、分类框架及问题描述。随后以环境的平稳性与评估模型构建规则的视角，对现有实时安全性评估方法进行分类论述。之后介绍应用部署的情况，给出以往研究中一些典型实例，最后指出实时安全性评估研究中亟待解决的问题和未来的发展方向。

1　实时安全性评估技术概述

1.1　概念定义

动态系统的实时安全性评估技术，是一种在线监测和评估系统安全程度的方法，为及时的警报与决策提供支持。在工程实践中，动态系统在运行过程中的状态随时间不断变化，所在的环境也可能随时发生改变，实时安全性评估方法的结果也可能会相应地调整。动态系统的实时安全性评估过程通常包含以下三步：

1）设计评估指标：通过分析系统运行监测数据和影响安全性的关键因素之间的关系，选取系统运行过程中可表征运行安全的过程变量，得出系统实时安全性评估指标集合。

2）构建评估模型：基于对动态系统历史运行过程中事故演化机理的认知，综合考虑任务目标的具体要求，建立评估指标相关的监测数据与安全性状态之间的映射关系。

3）优化评估策略：依据系统运行状态的实时变化以及决策过程中的反馈信息，实施对评估模型的持续更新和优化机制，以确保评估结果的准确性与时效性。

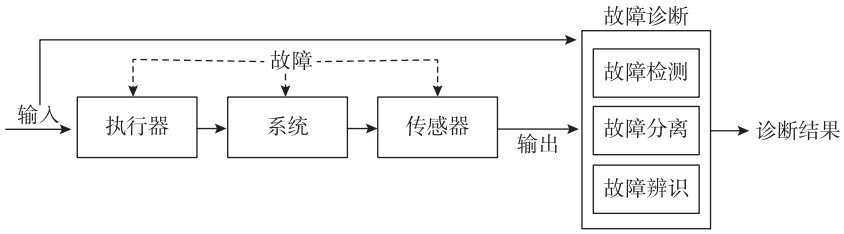
现有研究中对实时故障诊断与实时安全性评估之间的区别与联系讨论较少。尽管二者都是在系统的在线运行过程中实时得到结果，二者侧重点和目标仍然有所不同，其总体流程可被总结如图 2 所示。实时故障诊断技术通常利用系统的输入输出数据，使用不同方法对故障进行检测、分离与辨识。该技术的核心是检测和识别系统中的故障，侧重于发现可能导致系统性能降级或功能失效的自身问题。相较之下，实时安全性评估技术不仅追求系统的内部健康运行，还着重于系统对外部环境和人员潜在安全威胁的持续了解。

1.2　分类框架

随着理论研究的深入和相关领域的发展，各种实时安全性评估方法层出不穷。从评估模型的构建形式出发，本文将现有动态系统实时安全性评估技术进行了分类，如图 3 所示。其中，显式分析是指通过制定规则来直接分析系统的实时安全性状态，通常可得到明确的系统模型参数，而隐式分析侧重于识别出测量数据与安全性状态之间难以直接观察到的模式。前者

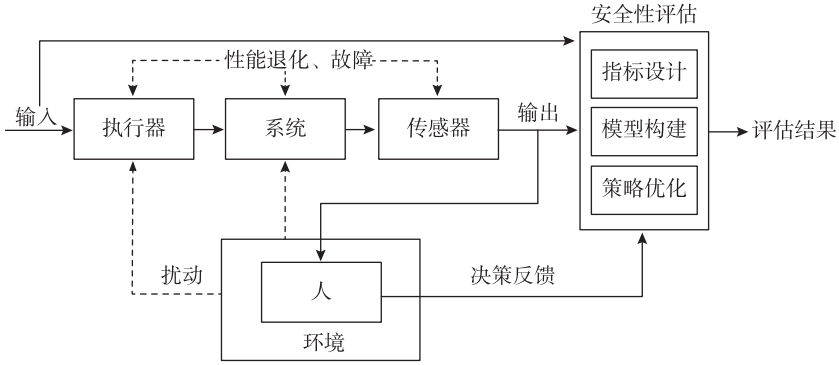
相对适用于规则清晰、数据结构化、可解释性需求高的场景，而后者则更适用于具有复杂程度高、数据非结构化或数据规模较大的场景。

此外，从环境的平稳性角度出发，可将实时安全性评估技术分为以下两个方面：1）平稳环境（Stationary environment）下的实时安全性评估技术；与 2）非平稳环境（Non-stationary environment）下的实时安全性评估技术。其中平稳环境下的实时安全性评估技术假设系统运行过程中处于平稳环境，该类技术通常适用于对系统的控制和监测具有高度可预测性的场景。而非平稳环境下的实时安全性评估技术假设系统运行过程处于非平稳环境。在这种情况下，



（a）实时故障诊断框架

（a）Framework of real-time fault diagnosis



（b）实时安全性评估框架

图 2　动态系统实时故障诊断与实时安全性评估示意图

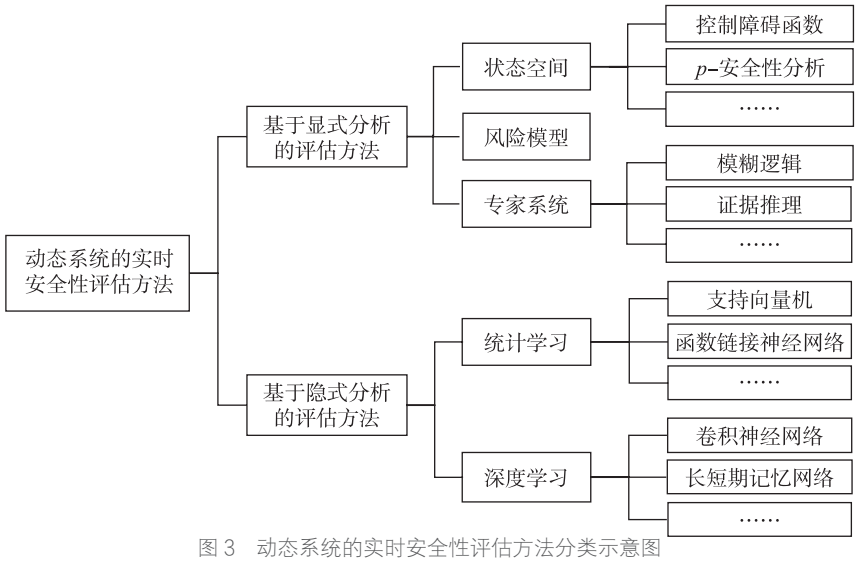


图 3　动态系统的实时安全性评估方法分类示意图

针对安全威胁的认知可能出现剧烈变化。因此，该类技术通常适用于对系统控制和监测具有不可预测性的情况。对现有文献进行调研，不同环境条件下实时安全性评估方法的研究分布情况可被总结如表 3 所示。

1.3 问题描述

如前文所述，系统安全性需要同时考虑系统状态及环境状态，令 $\mathcal{S}$ 表示系统的安全性状态，其可被描述为：

表 3 不同环境条件下实时安全性评估方法的研究情况

所依据的主要理论体系	平稳环境	非平稳环境
状态空间	极多	极少
风险模型	较少	较多
专家系统	较多	较少
统计学习	较多	较多
深度学习	极多	极少
信号处理	多	少

$$\mathcal{S} = (\boldsymbol{\mathcal{M}}, \mathcal{I}; \mathcal{K}) \tag{1}$$

其中 $\boldsymbol{\mathcal{M}}$ 表示系统测量，$\mathcal{I}$ 表示所构建的安全性指标集，$\mathcal{K}$表示完成安全性评估所基于的知识体系，包含了在特定条件下对系统对人、环境及财产影响程度的认知，同时外部环境的影响也可能作用于系统测量 $\boldsymbol{\mathcal{M}}$。该式表明安全性评估是以充分考虑专家知识为条件，关于相关量测及指标的函数。

由于环境的复杂性和多样性，实时安全性评估技术的建模问题极具挑战性。参考如图 2（b）所示的示意图，实时性安全性评估技术可定义如下。

定义 1. 对于一个动态系统，存在一个安全性指标集 $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2, \cdots, \mathcal{I}_d\}$，其由部署在系统中的多个传感器收集得到，通常包含用于监测系统内部的健康状况的系统状态监测传感器，以及用于监测系统外部的环境条件的环境感知传感器。

对于安全性指标集 $\mathcal{I}$，假设在离线阶段所收集到的历史测量数据集为 $\boldsymbol{\mathcal{M}}^{\text{off}} \in \mathbf{R}^{n\times d}$，其可被表示为：

$$\boldsymbol{\mathcal{M}}^{\text{off}} = \begin{bmatrix} \boldsymbol{\mathcal{M}}_1^{\text{off}} \\ \vdots \\ \boldsymbol{\mathcal{M}}_n^{\text{off}} \end{bmatrix} = \begin{bmatrix} \mathcal{M}_{1,1}^{\text{off}} & \cdots & \mathcal{M}_{1,d}^{\text{off}} \\ \vdots & \ddots & \vdots \\ \mathcal{M}_{n,1}^{\text{off}} & \cdots & \mathcal{M}_{n,d}^{\text{off}} \end{bmatrix} \tag{2}$$

在此背景下，可根据具体应用场景离散地得到 m 个安全性状态：$\mathcal{S}_k^{\text{off}} \in \{\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_m\}$，或连续地划分为一定范围内的安全性程度：$\mathcal{S}_k^{\text{off}} \in [0, 1]$。离线阶段所构建的安全性评估模型 Φ 可被定义为满足以下条件的映射：

$$\Phi : \boldsymbol{\mathcal{M}}^{\text{off}} \to \mathcal{S}^{\text{off}} \tag{3}$$

其可通过风险模型或专家系统等显式分析方法得到，也可结合统计学习或深度学习等隐式分析方法得到。安全性状态的表示形式通常取决于具体的应用场景，在一些场景下，回路中的专家可以根据经验直接实时给出系统的安全性状态，此时人的因素将直接作用于评估过程；而在另一些场景下，可直接依据专家制定的规则来实时分析系统的安全性状态，此时人的因素将间接作用于评估过程。

平稳环境所刻画的是数据特性和系统状态表现出相对的稳定性和可预测性。相比之下，系统状态在非平稳环境中是动态变化的。令 $(\boldsymbol{\mathcal{M}}^{\text{on}}, \mathcal{S}^{\text{on}}) = \{(\boldsymbol{\mathcal{M}}^{(t)}, \mathcal{S}^{(t)})\}_{t\in\mathbf{N}}$ 表示在线监测过程中的系统实时测量与实时安全性评估状态，$\boldsymbol{\theta}^{(t)}$ 表示 t 时刻评估模型的实际参数。评估模型需要依据系统的实时测量对系统的安全性实时地给出评估结果 $\hat{\mathcal{S}}^{(t)}$，其可表示为：

$$\hat{\mathcal{S}}^{(t)} = \Phi\left(\boldsymbol{\mathcal{M}}^{(t)}; \boldsymbol{\theta}^{(t)}\right) \tag{4}$$

令 $\mathcal{D}$ 表示描述分布之间差异性的距离度量，$\mathcal{P}(\boldsymbol{\mathcal{M}}^{(t)}, \mathcal{S}^{(t)})$ 表示 t 时刻下系统测量所对应的真实分布，$|\cdot|$表示时间序列的长度，可定义如下：

定义 2. 对于 $\forall$ 时间子序列 T_1 与 T_2，若满足 $|T_1|$，$|T_2| > \tau$ 个单位时间长度且 $\tau \in \mathbf{N}$，在满足以下假设的情况下，可称之为**平稳环境**：

$$\mathcal{D}\left[\mathcal{P}\left(\boldsymbol{\mathcal{M}}^{(T_1)}, \mathcal{S}^{(T_1)}\right), \mathcal{P}\left(\boldsymbol{\mathcal{M}}^{(T_2)}, \mathcal{S}^{(T_2)}\right)\right] = 0 \tag{5}$$

定义 3. 时间子序列 T_1 与 T_2，若满足 $|T_1|$，

$|T_2| > \tau$个单位时间长度且$\tau \in \mathbf{N}$，在满足以下假设的情况下，可称之为**非平稳环境**：

$$\mathcal{D}\left[\mathcal{P}\left(\mathcal{M}^{(T_1)}, \mathcal{S}^{(T_1)}\right), \mathcal{P}\left(\mathcal{M}^{(T_2)}, \mathcal{S}^{(T_2)}\right)\right] \neq 0 \tag{5}$$

定义 2 与定义 3 分别为平稳环境与非平稳环境的数学描述，在现有文献中，通常根据系统所处的实际场景对环境的平稳性进行判断，以采用合理的方式来构建评估模型。在一些应用场景中，若系统内部退化、外部环境干扰等因素所带来的影响相对较小，则通常可认为该环境具有平稳特征，若上述因素不可被忽视，则通常可认为该环境具有非平稳特征。以深海载人潜水器（Deep-sea manned submersible，DSMS）的实时安全性评估方法为例，受到如水流、水压、海洋生物等因素的影响，其在下潜过程中会遇到不同类型的安全威胁，实际深海环境为典型的非平稳环境。不同安全威胁所对应的安全性等级可被总结如图 4 所示。在一些情况下，故障等异常事件可能引发安全威胁，对人、环境及财产造成损害，但并非所有故障都可能会导致系统具有较低的安全性，不同异常事件可能会导致不同的安全性等级。环境带来的安全威胁及认为误操作等异常事件，也可能对系统安全性造成实时影响。因此，合理地设计安全性指标、构建评估模型及优化评估策略，对于实时得到准确地安全性评估结果具有重要意义。

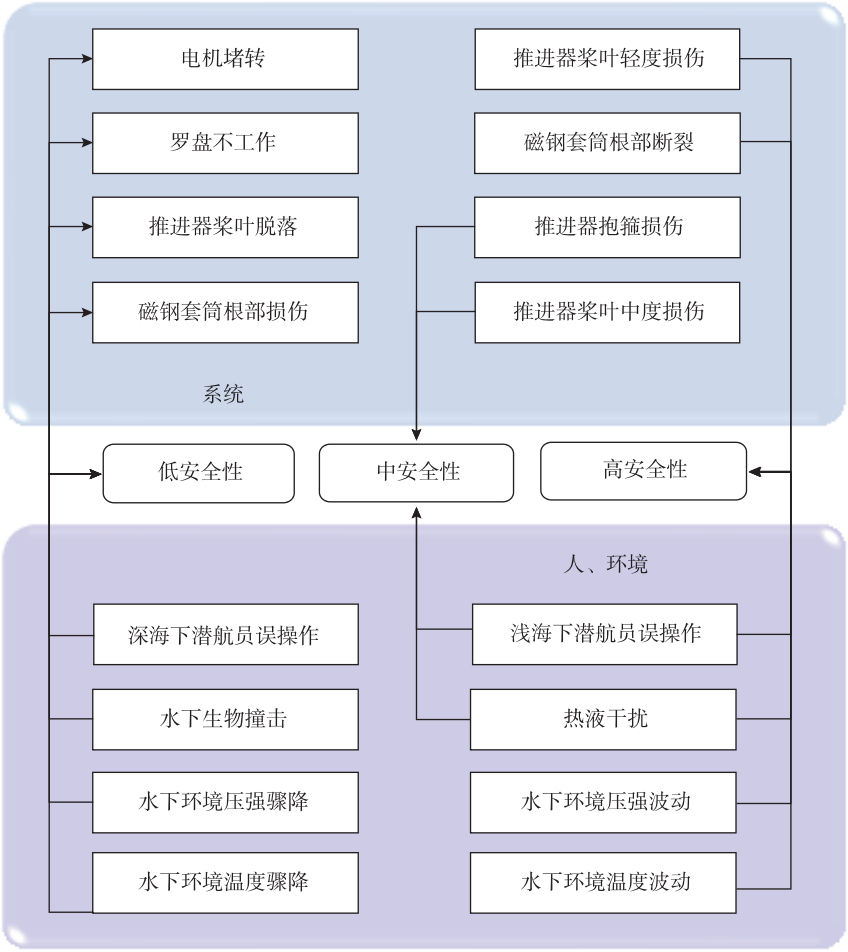


图 4　不同安全威胁对应的安全性等级示意图：以深潜器为例

综上，动态系统的实时安全性评估方法一般可被描述为：通过利用特定的监测数据和专家知识，寻求合理的映射关系，并将其用于实时估计系统安全性状态的过程。平稳环境下的实时安全性评估技术侧重于在离线阶段对如式（3）所示的映射规则进行设计。模型可对已有的历史模式进行学习，其可在监测过程中得到较高的评估准确性。而在非平稳环境下的安全性评估技术更侧重于对适应环境的变化。通常需要频繁根据最新的数据动态，在在线阶段中调整评估模型的内部参数或结构，以完成对评估逻辑的更新与优化。

2　现有方法与进展

从环境及分析方法特点出发，本章在图 3 所示方法的基础上，对现有动态系统的实时安全性评估方法进行分类总结与讨论，其主要逻辑结构总结如图 5 所示。

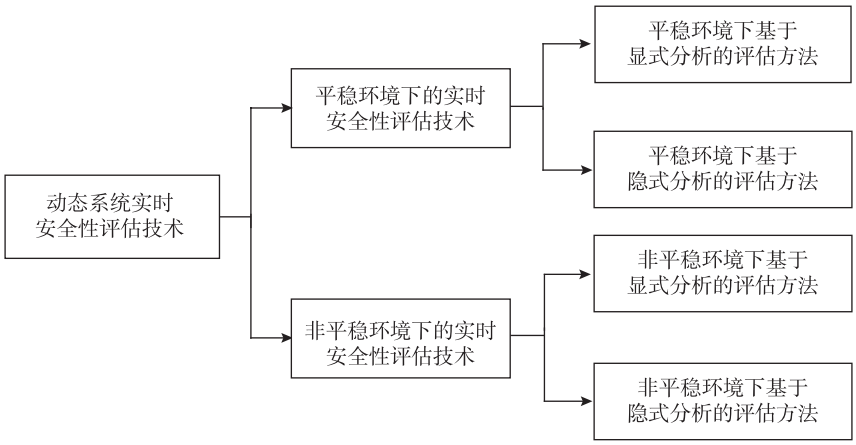


图 5　分类讨论逻辑结构示意图

2.1　平稳环境下的实时安全性评估技术

平稳环境下动态系统的实时安全性评估技术通常假设评估模型在离线阶段具有足够丰富的认识，且在线阶段相邻时刻的测量所对应的真实分布的变化是有限的。评估模型可基于此建立历史数据与安全性状态的映射关系，使其能够在实时在线监测中迅速识别异常或潜在安全威胁，并作出准确的实时安全性评估。该类方法可视为忽略人类基于对环境、财产所遭受的安全威胁认知对系统及执行器的干预及调整，同时在一定程度上忽略决策反馈因素对于评估器的影响。

2.1.1　平稳环境下基于显式分析的评估方法

对于动态系统而言，在平稳环境下构建显式的安全性评估模型有助于清晰地展示如何从系统特征、参数或数据中推导出安全性评估结果，使人能够追踪和理解评估的逻辑，提高技术的可解释性，并增强评估结果的可信度。因此，这种方法在学术界受到了广泛的讨论和关注。

随着近些年来控制理论技术的不断发展，一类基于动态系统的状态空间方程来分析系统安全性的方法在近几十年内引起了相关学者的重视，在一些文献中也被称之为安全性验证（Safety verification）问题，尤其是针对故障可能会导致严重安全事故的情况，比如人员伤亡、环境破坏或者重大财产损失的安全关键系统（Safety-critical systems），此类方法显得尤为重要。

对于有限状态系统的验证，Clarke 提出的模型检查（Model cheking）技术一经提出便引起了学术届的广泛重视，并促进了用于验证连续系统的类似研究。考虑到在非线性系统上的适用性，Prajna 等参考了 Lyapunov 函数的核心思想和稳定性分析理论，提出了一种基于障碍函数（Barrier functions）的分析方法 2，用于限制系统状态变量的变化范围，以确保系统的安全性得以满足。此类方法假设系统在初始时刻是安全的，将运行安全分析问题转化为对障碍函数的约束是否成立的判断问题。一些研究者将障碍函数推广成了控制障碍函数（Controlbarrier function，CBF），以此将系统状态上的约束映射到控制输入上的约束，随后发展出了指数控制障碍函数与高阶控制障碍函数（High-or-der CBF，HOCBF）。考虑到过于严格的约束易使得安全性评估结果过于敏感，近些年来，一些研究者试图通过寻找障碍函数的较弱约束以提高评估表现。Zhu 等开发了一类超球面方法，通过构造一个超球面来覆盖不安全集（Un-safe Set），使用代数几何方法建立了障碍函数。随后借鉴有限时间稳定性的稳定条件表达式的形式，一种指数 安全准则被提出以实现对动态系统运行安全的控制。一些研究者也试图利用神经网络技术对高维、非线性系统上进行推广，但分析方法仍待进一步发展。

与此同时，一类针对随机系统的 - 安全分析（ -safety）问题被广泛讨论，该类方法旨在从危险事件的视角出发，通过计算达到 - 避免概率（Reachavoid-

probability），对系统的初始状态进行分类，危险项（Hazard item）是导致概率大于 的不安全区域的初始状态。Wisniewski 等针对一般类别的马尔可夫过程完善了随机到达避免的概念，并通过使用了扩展生成器和多项式优化来近似搜索最大安全状态集，以此解决概率安全问题来确定随机系统的安全性。Girard 提出了一种基于近似双相似抽象的方法，用于有效计算在可达性规范约束下的安全控制器。文献 采用具有时间调度结构的多个 Lyapunov 函数来估计受到停留时间约束的切换系统的输出可达集，基于输出可达集估计结果，研究了不确定切换系统的安全验证问题。Schürmann 等提出了一种新的安全控制方法，将最优控制器和安全控制器结合使用。在大多数情况下，系统由最优控制器控制，而安全控制器则不断监测最优控制器的操作，在系统变得不安全时介入控制。安全控制器利用了向后可达集计算的新概念，并开发了一种描述可达集的新方法，可有效地最大化向后可达集的尺寸，这些分析方法在控制器设计过程中具有重要价值。

此类方法旨在以控制理论的视角对系统安全性进行建模，通常将人对系统安全性状态的理解表现在对状态约束或安全区域的划定上。通常无需预先推演危害事件，在对系统参数具有足够丰富的认知的条件下，此类分析方法表现出较强的适用性，可有效分析由操作失效或误操作导致系统安全状态变化的一类场景。然而，障碍函数需要随时间发生变化时，这类方法的有效性受到了挑战。

另一类主流方法基于专家先验知识来构建规则库，进而更全面地审视和分析系统的安全性3。在现有文献中，诸如证据推理（Evidential reasoning，ER）及模糊规则（Fuzzy rule，FR）等融合推理技术被广泛用来描述系统测量和安全性状态之间的关系，其在表示不确定性方面展现出了出色的能力。Li 等使用条件广义最小方差方法来选择关键特征，减少了计算复杂性并提高了评估准确性，并随后提出了一种用于复杂系统安全性评估的优化模型，进一步考虑了特征间的相关性和冗余性。Tang 等将传感器的可靠性和扰动因素引入到了评估过程。Liu 等使用专家群体交互关系对客观融合权重进行修正，强调了专家在评估过程中的影响作用。Zhou 等进一步发展了评估模型可靠性的计算方法，通过应用一阶局部敏感性方法推导出属性可靠性的敏感性分析，以定量分析输入可靠性对模型输出和模型可靠性的影响，在 JRC-7M 航空继电器的安全性状态评估上取得了很好的效果。

这类方法在信息不完整或不精确的情况下通常具有出色的性能，在系统处于平稳环境时，可以有效帮助领域内专家理解评估结果的原因。通常将安全性指标集所对应的传感器视作不同的信息源，利用系统测量进行显式融合推理，基于对安全性状态的置信度水平来描述系统的整体安全性。然而，专家知识成本昂贵、先验数据较难获取等问题在一定程度上仍较难有效解决。对于一些具有多个交互组件和相对复杂关系的系统，该类方法通常难以对复杂模式进行有效观测，这限制了其在一些具体工程领域的应用价值。

2.1.2　平稳环境下基于隐式分析的评估方法

随着系统复杂性的不断增加，安全性评估模型的显式构建变得愈发困难。近几十年来，针对主要应用领域的系统特性，逐渐发展出了各种不同的基于隐式构建安全性评估规则的方法体系。

一类方法通过引入统计学习模型来分析和理解系统安全性，其通常基于历史数据对发生过的安全事件进行研究。通过上述分析过程，评估模型能够学习到映射规则，进而有效地识别和评估潜在的安全威胁。在过去数十年里，使用决策树（Decisiontree，DT）、人工神经网络（Artificial

neuralnetwork，ANN）等技术构建评估模型已经取得了广泛研究。Yu 等提出了一种混合特征选择方案，用于自动筛选健康指标，以此提升其安全性评估效果。文献 提出了一种基于因果关系的特征选择方法，其利用系统物理学来学习特征间的因果结构并选择特征，从而实现更快的计算速度、更高的鲁棒性和更好的可解释性。这类方法需要在引入专家知识描述系统安全性状态的基础上，通过构建隐式模型来捕捉更加复杂的非线性关系，揭示系统测量与安全性状态之间不易察觉的模式和趋势。尽管此类方法进一步降低了对专家经验的依赖性，但也同时对安全性指标的设计方法提出了更高的要求，对数据中的噪声、缺失等异常也更加敏感。

进入 21 世纪以来，随着复杂装备的规模性急剧提升，用于分析和描述安全性状态的信息和数据得到了显著增长，并且这一趋势仍在持续。在此背景下，随着近年来人工智能技术的蓬勃发展，一些研究人员试图通过提升模型容量的方式来增强学习能力，进而捕捉数据中更复杂和更细微的模式和特征，以此更好地识别安全威胁。集成学习（Ensemble-learning）框架通常被视作有效的模型结构设计方式，以此提高评估模型的泛化性能。通常采用决策树、极限学习机（Extreme learning machine，ELM）、支持向量机（Support vectormachine，SVM）等作为基学习器。Xu 等采用集成极限学习机，在此基础上设计了决策规则。Sarmadi 等引入三个层次上的三种马氏距离度量，提出了一种应用于非生成式顺序算法的集成学习方法，提高了马氏距离的表现以及损伤检测能力。Dworakowski 等采用集成 ANNs 来处理单个 ANN 初始权重随机分配的问题，提出了一种自主集成设计方法，选择最适合解决特定任务的 ANN 子集。此外，基于无导数贝叶斯优化求解，Liu 等提出了一种基于贝叶斯神经网络的安全约束最优功率流模型，在 IEEE 39- 母线系统上说明了其有效性。He 等使用随机子空间方法离线训练多个小型决策树，并针对缺失数据设计了加权投票机制，以提高评估模型的整体鲁棒性。

随着深度学习（Deep learning，DL）在计算机视觉、自然语言处理和语音识别等领域取得卓越表现，其多层次、自动学习特征的能力被广泛认可。不少学者将深度学习引入安全性评估领域中，希望从原始数据中提取关键特征，识别复杂的模式和异常，进而提供更深层次、更全面的解决方案。

借助深度学习处理高维度特征的能力和广泛的适应性，Alawad 等充分融合天气、轨道状况、压载条件、车站人数等数据，利用卷积神经网络（Convolutionalneural network，CNN） 进行物体检测、行为分析，以评估铁路运行过程的安全性。Sun 等提出一种基于深度学习的特征提取框架，利用深度自编码器将传统状态变量数据转化至低维空间，从而降低安全性评估难度。Sarkar1 等应用深度学习技术来表征复合材料上裂纹形式的损伤，在大量实际图像数据中验证了方法的鲁棒性和准确性。尽管深度学习展现出卓越的能力处理高维度复杂数据，其潜在挑战之一在于对数据量的高需求。为了获取丰富多样的训练数据，以捕捉数据中的抽象特征和复杂关系。Azimi 等采用迁移学习技术将类似形式的数据迁移到结构健康监测中，由此预训练卷积神经网络，进而实现在实际大规模系统内进行损伤识别和定位。Ren 等在能源系统安全性评估中引入生成对抗网络（Generative adversarial network，GAN），利用生成器（Generator）产生符合历史数据分布的样本、鉴别器（Discriminator）区分生成数据和真实历史数据的差异，通过迭代训练达到平衡，用于生成能够准确替代缺失数据的新数据。除此之外，深度学习的可解释性也是限制其应用在安全性评估领域的一大挑战，

这一问题也激发了众多学者的研究兴趣。Warnecke 等提出了安全领域内对评价解释性方法的指标，如稳定性、效率、鲁棒性，根据这些指标评价现有解释性方法在安全性评估领域内应用的优劣性。更具体地，Guo 等提出了一种专为安全性应用设计的高保真度解释方法，其核心思想是通过扰动输入、梯度反向传播等简单的可解释模型来近似深度学习决策边界的复杂区域，在恶意软件分类器中得到了较好的应用。针对实际动态系统运行过程中可能出现的未知情景，Liu 等提出了一种基于主动增量学习的在线安全性评估方法，采用具有增量学习能力的分布外检测模型，有效降低了未知情景对评估过程的负面影响。进一步地，文献 针对实际无标注数据中可能存在未知情景的挑战，提出了一种主动标注辅助的半监督安全性评估模型，从而在更广泛的场景下缓解了未知情景所导致的局限性。

尽管在近些年来已经有了大量的研究成果，深度学习在安全性评估领域的研究仍然具有一些客观的局限性。深度学习模型在安全性评估中的表现显著依赖于数据的质量和数量。对于罕见异常状态，数据的缺乏可能导致模型无法准确识别或应对这些事件。此外，模型决策的不透明性可能引发最终用户的理解和信任问题。另外，此类架构通常对计算资源和内存需求较高，这些局限在资源有限的终端设备上尤其突出，在一定程度上限制了其应用范围并增加了部署难度。

2.2 非平稳下的实时安全性评估技术

非平稳环境是指在线监测过程中，系统特性随时间发生一定程度变化的情况。以工业系统为例，在设备磨损、环境条件变化、工艺变化、材料老化、外部干扰等多重因素的影响下，监测到的数据可能会呈现不规律的波动或漂移，使得离线阶段所构建的实时安全性评估模型的性能受到挑战。因此，评估模型需要具有极强的适应性，以确保对系统实时安全性进行准确评估。该类方法需要充分考虑人类对安全威胁的认知变化，并需要重点考虑决策反馈因素对于评估器的影响。

2.2.1 非平稳环境下基于显式分析的评估方法

考虑到在非平稳环境下环境条件随着时间不断变化，离线阶段中的系统模型参数可能不再适用，因此需要及时更新障碍函数来保持评估的准确性和时效性。此类方法的研究重点通常是对自适应控制障碍函数（Adaptive CBF，aCBF）的设计，通过对控制障碍函数进行自适应调整，以反映对系统安全性状态认知的变化。在文献 中，Lopez 等引入数据驱动的思想，提出了一种具有较少保守性的鲁棒自适应控制障碍函数，以实现系统的自适应安全性。通过在 HOCBF 的定义中引入惩罚函数，Xiao 等提出了参数自适应控制障碍函数和弛化自适应控制障碍函数，以此进一步提高系统的安全性和适应性。在文献［120］中，一种 BarrierNet 架构被进一步提出，使得控制障碍函数的参数能够自适应环境条件的变化，并可以从数据中学习。尽管这类方法有助于使得系统能够适应内部状态和外部环境的变动，建模复杂性的增加所导致的一系列问题仍然有待进一步研究。随着障碍函数的动态调整，对未知参数估计的难度也相应增加。考虑到对时变动态系统进行状态估计本身所具有的难度，aCBF 更新机制的引入会使得估计过程更加复杂，需要更高级的算法和计算资源，同时对数学分析能力提出了更高的要求。

近年来，对融合推理技术进行在线自适应调整的必要性和有效性引起了广泛的关注和讨论。这种自适应调整旨在更好地应对和管理环境带来的复杂性，通过维护专家知识库或修正置信规则库的方式改变映射规则，从而动态调整对系统当前所处安全性状态的判据，以此反映对系统安全性状态的认知变化。近年来，Hu 等提出了一种考虑预警功能的自

适应在线安全评估方法。Zhao 等认为由于动态系统在运行过程中的运行状态不断改变，系统状态的指标也应该随之改变，以此综合考虑不同阶段的系统安全性特征，提出了一种基于自适应权重求取模型的证据推理方法。在文献［123］中，安全性指标的可靠性也被视作提高评估的准确性的重要因素。考虑到实际场景通常具有缺乏观测信息、系统复杂度高以及环境干扰等特点，一种具有属性可靠性的置信规则库也被提出以考虑属性可靠性和全局不确定性。Zhao 等提出了基于最大偏差的权重系数计算方法，为每个安全性指标分配自适应权重。上述方法的主要思路是通过动态调整对安全性指标重要性估计的方式影响规则库。然而，对于一些相对复杂的动态系统，确定各信息源的可靠性和相应权重变得更加困难，同时由于系统中可能存在未知的相互作用和非线性关系，动态调整推理规则的方法可能难以有效预测和应对未知的系统响应。

随着测量技术的发展，一些研究者也开始尝试将经典动态风险分析方法的分析框架应用于处理动态系统实时安全性评估问题。然而，直接应用通常存在一定的局限性。传统动态安全性评估方法只使用事故序列前兆数据（即来自类似系统的事故的计数数据）来更新风险指标，该类方法必须等到事故或几乎失误（前兆）发生后才能更新风险指标的估计。且由于统计数据普遍从类似系统中收集，其结果仅能反映群体特征，而不能充分考虑目标系统的个体特征。

在此基础上，一类方法试图通过引入状态监测（Condition monitoring，CM）数据来对上述评估过程进行优化 4，以此提供有关目标系统个体退化过程的信息。因此，多种控制领域的先进技术被用作估计动态系统真实的退化状态，从而使得评估结果更加真实可靠。在现有研究中，Zadakbar 等将 Kalman 滤波器应用于上述过程来对真实的退化状态进行协同估计。为了应对非线性和非高斯特征，Yu 等基于状态监测数据开发了一种基于自组织映射的 CMBRA 方法。而 Wang 等通过监控过程中的多个关键变量来估计不良事件发生的概率，并使其随着系统的实时干扰而不断更新。Zeng 和 Zio 认为可联合使用状态监测数据与在类似系统运行期间发生事故的后果的统计数据，因此提供群体信息，而状态监测数据则可以描述系统特定动态系统的实时特征，被视作是对 ASP 数据的有效补充。

随着故障预测与健康管理（Prognostics andhealth management，PHM）领域在近些年的蓬勃发展，一些研究者将 PHM 中的剩余寿命（Remaininguseful life，RUL）预测及可靠性分析（Reliabilityanalysis）等核心技术应用于动态安全性评估过程，其可对动态系统所遭遇的潜在安全威胁进行辅助估计，具有重要研究价值。Zhao 等利用连续时间马尔可夫链来描述退化过程，并使用 Cox 比例风险模型以描述故障时间的危险率。Dehghanian 等设计了定量方法来定义高压断路器陷入每一类衰退 / 恢复状态的概率，并利用高压断路器（Highvoltage circuit breakers，HVCB）的控制电路条件监测数据对其衰退 / 恢复状态进行生命周期评估。Kim 等的研究着重探讨了随着操作和工厂环境的变化而引起失效概率的变化，旨在减少在危害时间仿真和频率估计中产生的保守性，但未考虑监测数据中可能存在的噪声。考虑到监测数据中的非平稳特性，BahooToroody 等将经验模式分解和分层贝叶斯模型相结合，根据对天然气分配系统的实时监测的数据更新超参数，进而更新失效概率分布函数。Xing 等用于将条件监测数据和检验数据（Inspection data）同时集成到评估过程中，并说明了其在核能发电系统上的有效性。Ni 等结合系统特点，将事故发生的可能性和严重性表示成反映概率风险的指标，并利

用实际能源管理系统说明了其方法的有效性。Li 等通过结合了母线 / 支线和节点 / 断路器网络模型执行在各种条件下的动态应急分析，计算实时传输限制，并提供在线控制建议以减轻运营风险。

尽管此类方法试图将状态监测信息引入到评估过程中，但绝大多数研究聚焦于将其用来对事故发生的可能性等参数进行调整，在理论层面仍需利用风险模型的分析框架。因此可被视作是利用事故发生概率与严重性之间的结合关系，基于系统测量对实时安全性状态进行显式分析。

2.2.2　非平稳环境下基于隐式分析的评估方法

随着系统日益复杂化，由更多异构元素（硬件、人员、数字）组成，这些元素组织成高度互联的结构，导致难以预测危险行为。在非平稳环境下，此类问题所带来的挑战显得尤其重要。隐式分析方法能更好地处理系统中不易观察或难以量化的因素，如人的行为、系统间的复杂相互作用以及难以预测的外部事件，因此在近年来得到了引起了研究人员的广泛关注。

为了进一步提升在线阶段下评估结果的准确性，具有增量学习（Incremental learning，IL）能力的方法在近些年来引起了越来越多的关注，此类方法试图让评估模型在系统运行过程中，利用所收集到的新数据进行实时更新及优化，来及时修正评估模型的映射规则。因此，该类方法能够适应之前未遇到的情况，提高对未知安全问题的识别能力，相对适用于状态变化频繁或需要对新情况快速响应的系统。

在 20 世纪 80 年代，Sobajic 等就已开始探索将自适应人工神经网络技术用于解决电力系统关键故障清除时间的预测问题，为这一领域的发展奠定了基础。此后，相关技术得到了进一步的发展和完善。Sun 等提出了一种利用相量测量和决策树的大规模互联电力系统安全性评估方法，该方案离线构建并定期更新决策树，以确定关键属性作为安全指标，进而提升评估结果的可靠性。Diao 等提出了一种使用同步相量测量和定期更新的决策树的在线电压安全评估方案，用以评估由电压崩溃（Voltagecollapse）所带来的安全威胁。He 等考虑电力系统在运行期间可能发生的运行条件变化和拓扑变化，提出了一种基于自适应集成学习的鲁棒在线安全评估方法，其在离线阶段使用提升算法（Boostingalgorithm）构建评估模型，在在现阶段中使用更新后的模型将当前运行条件下的实时系统测量映射到安全性状态上，并在实际系统上说明了其有效性。Zhang 等引入了一种基于在线序贯极限学习机的评估方法，在实际电力系统上进行了验证。文献 考虑了电力系统参数不确定性、参数变化以及高计算复杂性等问题，提出了一种基于高斯过程（Gaussian process，GP）的在线学习方法来评估电力系统的稳定性水平，并设计了迭代算法来更新评估模型参数。Singh 等采用基于粒子群优化（Particle swarm optimization，PSO）的加权平均方法来组合基学习器的预测结果，以此利用互补信息来提高模型的泛化能力，并在新英格兰 39 节点和 68 节点的测试电力系统上说明了这一策略的有效性。Liu 等提出了一种基于主动宽度学习系统（Broad learning system，BLS）的实时安全性评估方法，以降低决策反馈所带来的成本，并进一步考虑了有限计算内存的问题，在蛟龙号深海载人潜水器的生命支持系统上说明了其方法的有效性。在此基础上，进一步提出了基于子模性最大化理论的实时安全性评估方法，重点考虑了客观存在的极度类别不平衡问题。He 等基于动态模型解释结果，设计了一种在线主动学习策略，以此引入主动平稳性检测过程。

综合现有方法所表现出的特点，可以根据多个维度将不同环境条件下实时安全性评估方法所具有的特点和不足进行总结，如表 4 所示。

表 4　不同环境条件下实时安全性评估方法的研究情况

	平稳环境	非平稳环境
是否利用系统输入输出	是	是
系统特性变化	较少	较多
更新能力需求	低	高
外部因素影响程度	低	高
问题困难程度	较低	较高
实际应用范围	较小	较大
评估模型稳定性	较高	较低
决策反馈要求	低	高
现有理论成果	较多	较少

3　应用部署

3.1　典型实例

在当前研究中，动态系统的实时安全性评估仍处于发展早期，表 5 中列出了目前实时安全性评估方法有关的一些典型系统对象。

3.2　部署策略

在工程实际中，不同的动态系统通常展现出各自独特的特征。这些特征涵盖了系统的响应、行为、性能和特性等方面，取决于系统的类型、用途和环境条件等多种因素。因此，需要强调的是，为不同类型的动态系统制定标准化的实时安全性评估方法是不切实际的，这意味着安全性评估方法需要针对具体系统的特性和应用场景进行针对性部署，以确保准确性和有效性。本节将以几种典型系统为例，对实时安全性评估方法的部署策略进行简要讨论。

3.2.1　工程结构系统

工程结构系统通常是指包括建筑物、桥梁、管道等在内的工程和基础设施，其安全性与人员、财产及环境的安全紧密相关。通过研究和评估工程结构系统的安全性，可以及早发现潜在的问题，预防事故的发生，具有重要的社会意义和经济意义。

在现有文献中，一类代表性的方法被称之为结构化健康监测（Structural health monitoring，SHM）技术，其主要目标是监测和评估各种工程结构系统（如建筑、桥梁、管道等）的健康状况和性能，以确保它们在运行过程中的安全性。结构健康监测系统通常包括各种传感器、数据采集设备和数据分析算法，通过实时或定期收集有关结构物的信息，以此及时检测到结构物的异常行为、损伤或劣化，进而分析系统安全性。自 20 世纪末以来，针对工程结构系统的健康监测系统开始被陆续建立。Ye 等考虑了台风期间桥梁健康监测过程中所表现出的非平稳特征，采用经验模态分解（Empiricalmode decomposition，

表 5　几类典型动态系统在实时安全性评估框架下的现有进展

系统对象	参考文献			
	平稳环境下显式分析方法	平稳环境下隐式分析法	非平稳环境下显式分析法	非平稳环境下隐式分析法
工程结构系统	[159-163]	[74，84-85，94-95，107-108，154-158]	[126，129，164-166]	[200，202]
交通系统	[21，170]	[167-169]	[172-173]	[150-153，171]
电力系统	[178]	[80-82，86，88，91-93，96-97，174-177]	[135，136-140][143-149，201]	[143-149，201]
核能发电系统	[19，180-181]	[179]	[34，36，127，132，137，139，183][182]	[182]
化工系统	[187-188]	[184-186]	[44，130-131]	[189]
航空航天系统	[79，191-193]	[190]	[37，128，195]	[194]

EMD）等方法从原始风数据中提取出时变平均风速，作为非平稳风模型的趋势项。Liu 等提出了一种混合模型来处理非平稳环境下风速数据的高波动性、自相关性和随机波动性难以用单一模型预测的问题。Hua 等则研究了桥梁在车辆－桥梁相互作用下的随机非平稳特性。近些年来，为了应对环境作用和结构响应，研究人员建立了多种环境作用模型及动态性能监测方法，其通常以振动等监测信号作为系统测量，以进行对安全威胁的早期健康监测。通常情况下，系统测量数据具有规模大、复杂性高、维度高等特点，通常涵盖了如位移、应力、温度等多个方面的监测信息。如在文献［159］中，Li 等基于长期应变监测数据，采用断裂力学、损伤累积理论和无损检测等技术，开发了一种钢箱梁焊接细节的疲劳可靠性评估方法，得到了如图 6 所示的评估结果。

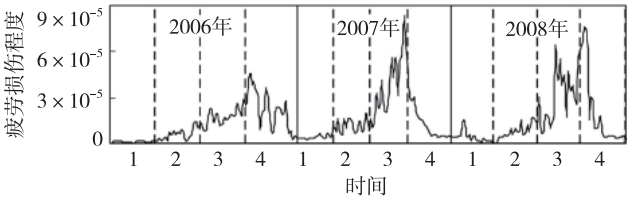


图 6　悬索桥焊接细节的疲劳损伤监测结果

在此背景下，一些学者尝试引入压缩采样（Compressive sampling）等技术来应对上述问题，取得了一定的进展。然而，工程结构系统的复杂性使得传统基于显式分析的评估模型较难捕捉数据中的非线性特征，对局部微损伤敏感性欠佳。对于一些大型工程结构系统，一类基于损伤检测（Damage detection）的安全性评估方法在近些年得到了广泛研究。随着近年计算机视觉（Computer vision）技术的飞速发展，研究人员已成功将具有高泛化性的深度学习方法应用于工程结构系统表面损伤或变化的分析，取得了显著的成果。而在工业物联网的推动下，深度学习方法得以更好地与大规模数据集和传感器网络结合，为工程结构表面的实时监测提供了更强大的支持。例如，Azimi 等以被称为 IASC-ASCE 结构健康监测基准问题的四层四分之一比例钢框架模型为应用对象，采用加速度传感器、电传感器收集应用数据，结合迁移学习技术训练卷积神经网络，并对该应用进行健康监测，平均准确率可达到 90% 以上。

因此，对于大型工程结构系统而言，当下研究主要集中在平稳环境下基于隐式分析的实时安全性评估方法。随着应用场景的发展，相关方法已经开始朝向非平稳环境下基于隐式分析的实时安全性评估方法过渡，在非平稳环境下对环境噪声的处理，以及更好地考虑人的决策反馈因素，具有较高的研究价值。

3.2.2　交通系统

一类在实际场景下被广泛讨论的应用对象是交通系统，其中替代安全性评估（Surrogate safety assessment，SSA）等方法被视为交通系统安全分析的重要组成部分，其通过使用诸如冲突严重性、交通行为模式、车辆轨迹和速度等替代安全性指标来实时评估交通系统的安全性状态。此外，要建立系统主动安全防护体系，对系统潜在或即将发生的风险进行感知，并从交通运营服务的稳定性和可靠性保障出发进行针对性风险响应。通过分析交通流中的潜在安全威胁，实时安全性评估方法可以在事故发生之前识别风险点，从而有助于采取预防措施。其不仅需要在充分利用道路传感器、视频监控、GPS 数据等数据源，且依赖于交通模拟软件和数据分析平台等先进的计算模型和分析工具。在该领域中，基于场景（Scenario）的方法是一个具有较高应用价值的研究方向，其可通过确定有用的场景来降低测试的工作量。在此背景下，一些数据驱动方法已经被提出来进行场景构建，进而验证车辆安全性。考虑到场景选择的随机性，此类方法具有较高的时间成本。因此一些研究人员试图更侧重于通过设计关键性指标来构建关键场景。在文献［173］中，采用了从欧洲收集的、由测试驾驶员在混合驾驶环境中进行的约 25 万公里的驾驶测试数据，利用碰撞时间（Time-to-collision，TTC）作为关键度量（Criticality metrics），并通过比较其与和人类

驾驶员引起的事故的实际距离来分析安全性，部分评估结果如图 7 所示。Wang 等根据物理驾驶实例，在 CarMaker 和 ROS 驱动程序中进行仿真模拟，根据车辆感知到的信息做出决策，并使用虚拟自动驾驶实例来感知周围环境，以便其了解对象的正确状态。随后将决策结果输入到触发器中，以评估行为安全性并识别关键场景。

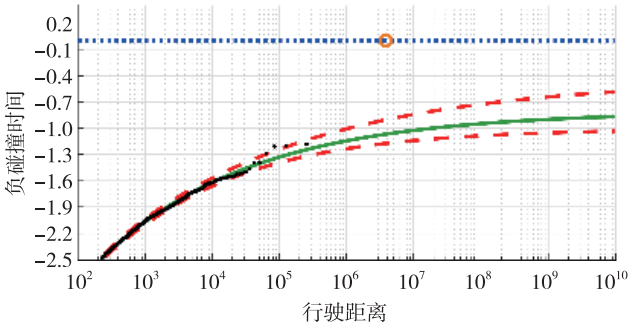


图 7　负 TTC 与实际行驶距离关系示意图

综上，考虑到由人类驾驶员的社会交互和与场景的物理交互与交通系统安全性之间的密切关联，现有交通系统领域的实时安全性评估方法更强调对安全性评估模型的可解释性。因此，通常采用的方法是基于显式分析的非平稳安全性评估方法。通过利用更先进的测量技术提升系统环境的感知能力，进而更好地分析潜在的安全威胁，具有重要的研究意义。

3.2.3　电力系统

考虑到具有多变性、多元性、规模性、互联性及非线性等特点，电力系统通常被视作是最复杂的动态系统之一。传统的电力系统安全性评估方法依赖于运行连续流和瞬态稳定性模拟，随着现代电力系统变得越来越复杂，此类方法难以在线评估其安全性。在电力系统操作参数和相应的安全性状态之间进行快速和准确映射对于提升评估性能具有迫切性。因此，一类被称之为动态安全评估（Dynamicsecurity assessment，DSA）技术引起了广泛关注。此类技术在过去几十年间取得了显著的研究进展 7，旨在通过分析对电力系统瞬态性能至关重要的属性和阈值，以此评估电力系统在各种动态条件下的安全性状态，可主要被分为：瞬态稳定性（Transientstability，TS）、电压稳定性（Voltage stability，VS）和频率稳定性（Frequency stability）等细分任务，相关研究通常在 IEEE 标准系统上进行效果验证。

过去几十年内，专家系统等基于显示分析的方法取得了一定的成功。然而随着电力系统的复杂化，理解系统的极限变得越来越困难，这限制了其应用空间。因此，该领域的研究人员在 20 世纪末即开始尝试引入决策树、人工神经网络等统计学习技术作为解决方案，以此学习安全性指标与安全性状态之间的映射关系。例如，Liu 等为预防动态不安全性，提出了基于贝叶斯神经网络的安全约束最优电力流模型，通过无推导贝叶斯优化来进行求解。该研究在 IEEE39 总线系统进行了应用分析，结果表明，与传统的点估计神经网络相比，贝叶斯神经网络能够正确可靠地检测电力系统的安全状态，能够避免过拟合问题。在文献 中，基于部分互信息和皮尔逊相关系数选择关键变量，进而采用迭代随机森林预测系统的 暂态稳定裕度（Transient stabilitymargin，TSM），以此提供电力系统的实时安全性信息，其原理示意图如图 8 所示。该方法在 IEEE39 总线系统进行了应用，并提供了区域 TSM 可视化结果，如图 9 所示。文献 提出了一种量化拓扑变化对评估模型准确性影响的度量方法，能够在具有高影响拓扑变化的系统中进行有效地安全性评估。

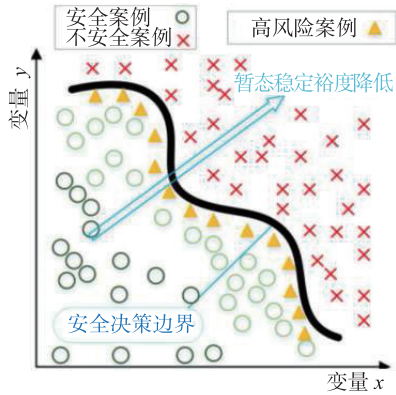


图 8　评估原理示意图

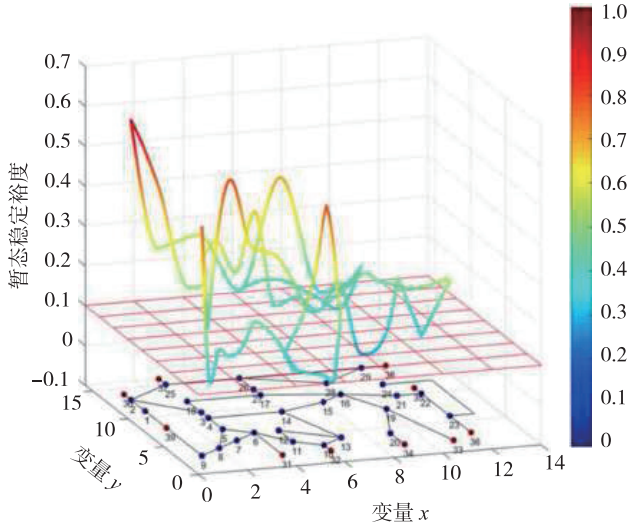


图 9　IEEE39 总线系统区域暂态稳定裕度分布图

随着同步相量（Synchrophasor）等测量技术的出现与发展，基于相量测量单元（Phasor measurementunit，PMU）等测量数据的安全性评估方法也随之取得了突破。考虑到电力系统在出现负荷变化、市场变化及政策影响等场景下可能处于非平稳环境，一类在线动态安全评估（Online dynamicsecurity assessment，ODSA）方法在近些年被用来对传统动态安全评估技术进行补充。此类方法利用更加先进的模型和仿真工具来分析电力系统的动态行为，考虑当前运行状态或即将发生的运行状态，试图确保系统能够在面临负荷变化等异常情况时仍能保持稳定运行。在此背景下，形成了较多基于隐式分析的非平稳安全性评估方法。近年来，一类由相量测量单元故障、通信丢失和相量数据集中器故障等原因所造成的数据缺失问题成为了研究热点，一些研究人员在近些年也尝试将半监督学习、主动学习、迁移学习 等先进的技术应用于解决数据不确定性及高计算复杂度等不同的细分问题。

因此，目前针对电力系统的实时安全性评估方法在基于隐式分析的非平稳安全性评估方面已经取得了显著进展，这一趋势显著地体现出了方法论的创新和应用的深化。尽管如此，如何在负荷变化或事故发生后，在电力系统状态监测过程中增强人类决策者的反馈仍是一个值得深入考虑的问题，其涉及到如何有效整合人工智能和自动化技术的分析结果与人类操作者的直觉和经验，以应对电力系统中的环境干扰及不确定性因素，实现更加准确和高效的决策过程。与此同时，针对电力系统实际问题的优化，尤其是在提高评估效率、增强系统适应性及处理复杂性方面，仍然存在着广阔的研究空间。

3.2.4　核能发电系统

核能发电系统具有较高的复杂性，为评估核能发电系统的安全性，传统的安全性评估方法通常基于事件树分析框架，以树形图的形式呈现系统可能发生的某种事故与导致事故发生的各种原因之间的逻辑关系，找出事故发生的主要原因，从而进行危险源辨识（如再循环泵、安全阀、自动减压系统、反应堆）。然而，事件树分析本质上是静态的，即在运行前对风险指标进行评估，以此分析系统运行过程中组件风险性。在日益发展的核电系统中，这种方法难以长期准确地评估安全性能。近年来，研究者试图将状态监测数据引入到风险模型的设计中，以此来解决传统方法所带来的局限。一种主要的方法是动态事件树分析（Dynamic ETA）框架，利用系统状态监测数据实时更新组件风险指数，在一些核电系统上的实验表明，动态事件树可以捕获系统特定的特征，从而提供对目标系统风险的更完整和准确的描述。例如，Zeng 等以核电站中的预期瞬态无补偿切断事故（Anticipated transientwithout scram accident）为应用背景，建立了一种用于条件监测数据建模的隐马尔可夫高斯混合模型，并利用贝叶斯网络将估计的退化状态与动态风险分析数据进行了集成。利用集成数据源更新事件树，并利用事件树评估系统安全性。该研究表明，将两个数据源集成到动态风险分析中会比使用两个单独数据源中的任何一个提供更准确和稳健的结果。

随着人工智能技术的迅速发展和一些核电系统组件中自主控制系统的发明，一些研究者开始尝试利用

数字孪生（Digital twin）等人工智能技术来辅助评估核电系统的安全性，以维护该控制系统的安全。上述的方法大都聚焦于核能发电系统的内部安全问题，然而，一些事故的发生让人们意识到除系统内部安全隐患外，核恐怖主义（即系统外部因素）的威胁也可能导致灾难性后果。于是有工作将层次分析法（Analytic hierarchy process，AHP）与人工神经网络相结合以研究核恐怖主义所带来的安全威胁，输出系统安全范围，为核能发电系统的实时安全性评估方法提供更全面的视图。

综上，针对核能发电系统的实时安全性评估方法仍然以显示分析方法为主，如何应对非平稳环境所带来的问题已经成为新的研究趋势，利用不断发展的人工智能技术以更深入地挖掘系统潜在的安全隐患仍有较大的研究空间。

3.2.5　化工系统

化工系统通常涉及复杂的化学过程、高压设备和危险品，存在潜在的安全威胁。在化工系统中，安全风险因素表（Safety risk factor table，SRFT）得到了广泛的应用。SRFT 用于确定影响系统整体安全的因素，如位置、能见度、财产、所有权、是否存在可作为大规模杀伤性武器前驱体的化学物质等，一些方法根据 SRFT 对系统的安全状态进行评级。除此之外，与其他系统类似，一些经典的概率统计方法也被广泛用于化工系统的实时安全性评估中。例如，Wang 等将事故发生的可能性建模为指数分布，该方法通过监测过程中的多个关键变量来估计不良事件发生的概率，并且通过考虑变量中的实时扰动使该概率不断更新。该方法在基准的田纳西－伊斯曼过程（Tennessee eastmanprocess，TEP）中进行了测试，验证了其可行性。近年来，不少研究者也引入多元线性回归、极限学习机、人工神经网络、支持向量机和集成技术等机器学习方法来进一步分析化工系统安全性。例如，Amin 等提出利用人工神经网络识别故障模式，将贝叶斯网络用于故障检测，利用事件树建模故障与系统安全性之间的关系。在此基础上，采用倒置正态损失函数评估一个与时间和故障场景相关的损失，并采用概率自适应机制来估计每个时间片中的条件概率，从而实时调整安全性评估模型。该研究也分别在基准的 TE 过程和仿真过程中进行了验证，展示了其适用性和灵活性。

综上所述，当下针对化工系统的实时安全性评估主要基于显式分析方法，但一些基于隐式分析方法的有关研究正取得越来越多的关注。目前，非平稳环境下基于隐式分析的实时安全性评估方法相对较少，通过先进的人工智能技术，结合监测数据实时挖掘、更新数据潜在模式，进而更快更准地评估系统安全性，具有重要的研究价值。

3.2.6　航空航天系统

航空航天系统的事故往往伴随着严重的财产损失，还直接关系到乘客、机组人员以及地面工作人员的生命安全。因此，对航空航天系统进行安全性评估是十分重要且必要的。现有文献中，针对航空航天系统的安全性评估方法仍然以传统的故障树分析、模糊逻辑、置信规则库、马尔可夫过程等传统分析框架为主，大多属于显式分析框架。例如，胡昌华等考虑到环境干扰下火箭监测数据存在不确定性的问题，提出了一种基于距离的监测指标不确定度求取方法，通过指标不确定度来反映噪声对于监测数据的影响程度。进一步考虑指标的不确定度，提出了一种输入数据与置信规则库之间的匹配度转换方法，并基于置信规则库融合不同指标的输入信息，以此得到火箭的安全性等级。如图 10 所示，在使用监测数据对初始模型中的参数进行调整后，蓝线能够更加准确地评估火箭的安全性状态，在专家无法准确判断的部分监测点能够做出准确的评估，有效克服了 专家知识存在的不确定性和局部无知性，达到了监测数据对专家知识的有效补充，实 现了监测数据与专家知识的有效融合。

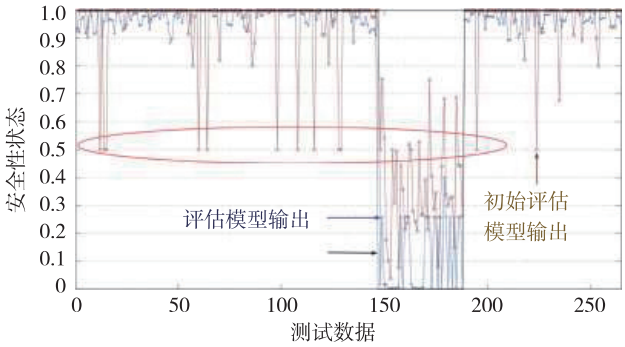


图 10　液体运载火箭结构安全性评估模型训练与测试

由于航空航天系统的特殊飞行属性，其周围环境常常发生变化。其安全性不仅受系统内部情况所制约，还与碰撞风险、电磁干扰、天气影响等外部环境因素息息相关，所以只适用于平稳环境下的评估方法具有较大的局限性。随着计算机视觉技术的发展，一类方法利用航空航天系统飞行过程中收集到的外部数据以评估系统的安全性。Wang 等利用无线通信技术，将无人机的飞行数据传输至本地电脑中，利用 MATLAB 平台进行数字孪生技术的仿真测试，采用卷积神经网络来识别无人机是否侵入禁飞区域，以此判断无人机及其周围的人或物是否存在安全威胁。

总体而言，目前航空航天系统的实时安全性评估方法主要为显式分析框架。考虑到该系统所处的环境复杂多变，对于非平稳环境下实时安全性评估方法的需求正在随着计算机视觉等技术的发展而不断提高。

4　总结与展望

动态系统实时安全性评估技术旨在通过高效地利用实时系统测量数据迅速识别潜在的安全威胁，此类技术使得数据处理和响应过程变得更加及时，从而在各类实际复杂应用场景下表现出显著的优势。本文首先对动态系统实时安全性评估技术的概念定义进行了概述，从环境的平稳性及评估模型的构建方式的角度给出了一类分类框架，并分别对问题描述进行了具体的形式化描述。随后对现有安全性评估技术进行了系统性回顾，并围绕实际系统的部署策略进行了归纳讨论。可以看到，当前的技术研究进展已经表现出了与应用系统高度相关的特点，这一研究趋势表明评估技术与其应用领域的结合程度在整体上不断提升。各个应用领域整体上从专注于平稳环境下评估技术的研究转向对非稳定环境下评估技术的深入探索。对于系统模型较简单、且对可解释性具有较高需求的场景，研究人员往往更偏好于基于显式分析的评估方法。而对于系统模型较复杂、且建模难度高的场景，基于隐式分析的评估方法获得了更多的关注。

近年来，随着测量技术及建模能力等方面的蓬勃发展，对实时安全性评估技术的要求也随之提升。鉴于此，本文对未来潜在的研究热点进行以下展望：

1）分布式动态系统的实时安全性评估技术：随着物联网技术在近年来的发展，系统复杂性的提升逐步成为了不可避免的未来趋势。然而，终端设备计算资源受限、隐私保护需求高等因素使得安全性评估变得更具挑战性。安全威胁的来源趋于多元化，其传播特性也呈现出高度的迅速性，实际终端设备及其配置可能会频繁变化，也进一步为有效的评估过程带来了客观难度。如何同步和分析跨多个节点收集的数据，从而减少系统停机时间和潜在损失，具有重要的研究意义，这要求开发新的算法和协议，以优化资源使用，并确保对敏感信息的保护。

2）高数据复杂性下的实时安全性评估技术：在实际场景下，实时收集到的系统测量可能表征出特征维度高、数据规模大、不确定性强等多重特点，数据之间可能存在复杂的关联和相互依赖性，这对安全性指标的设计及评估模型的构建都带来了不可忽视的难度。如何更有效地在高数据复杂性的条件下捕捉关键的特征信息，并针对相关关系、因果关系、时序关系等特点对模型架构及更新策略进一步优化，从而降低数据复杂性所带来的消极影响，值得进一步探索与研究。

3）具有未知情景的实时安全性评估技术：随着系统结构的日趋复杂，动态系统在实际运行过程中通常会不可避免地出现未知情景，使得现有基于完备认

知假设的评估方法容易出现误判现象，进而降低评估结果准确性。未知情景所固有的历史监测数据匮乏等特点又使得其难以被准确辨识，进一步为评估模型的构建带来了困难与挑战。因此，如何在开放的环境下考虑对未知情景的辨识与适应，对于推动实时安全性评估技术在工程实际上的应用，具有重要意义。

4）决策反馈约束下的实时安全性评估技术：如引言中所述，系统安全性的判断通常与人的认知水平密切相关，受到认知差异、场景变化等因素的影响。当下人、系统与环境之间的交互关系愈发紧密，使得在评估过程中引入人在回路（Human-in-theloop）特性成为至关重要的因素。在此背景下，有效应对决策反馈过程所存在的客观约束，具有重大现实意义。如何妥善解决由于标注受限、验证延迟等问题所带来的技术挑战，并对潜在的误行为进行协同分析，从而促进理解和预测动态系统的复杂行为，是一个潜在的研究热点。

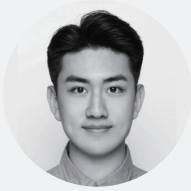
5）评估模型动态可解释性的理论研究：随着近些年来各类基于隐式分析方法的兴起，安全性评估技术的相关研究取得了显著的进展，但由此也带来了对模型可解释性的担忧，这为大规模模型部署带来了难度。同时非平稳环境的客观特点进一步加深了对评估结果的信任问题。因此，如何在确保隐式分析方法捕捉非线性映射关系能力的前提下，进一步探求针对评估模型的动态模型解释，从而使得评估过程更加透明和易于理解，具有重大研究价值。

6）针对环境非平稳性的主动检测策略设计：在多重复杂因素的共同作用下，考虑非平稳环境特点的实时安全性评估技术已经成为当下的研究趋势。环境对系统行为的影响不断扩大及测量技术的显著发展，使得监测数据呈现出底层分布变化加剧、变化行为趋于复杂等现象，评估模型对安全威胁的响应需求也进一步提高。研究如何主动检测环境的变化行为，基于预警信息及时对模型更新策略进行优化，从而提高安全性评估的准确性和适应性，值得进一步探索。

来源：自动化学报

作者简介

何　潇　清华大学自动化系长聘教授，2010 年获得清华大学博士学位，主要研究方向为动态系统的故障诊断与容错控制，本文通讯作者。



刘泽夷　清华大学自动化系博士研究生，主要研究方向为动态系统的故障诊断与安全性评估。

胡嵩乔　清华大学自动化系博士研究生，主要研究方向为动态系统的故障诊断与安全性评估。

刘　畅　清华大学自动化系博士研究生，主要研究方向为动态系统的故障诊断与安全性评估。

周东华　山东科技大学和清华大学教授，主要研究方向为动态系统的故障诊断与容错控制，故障预测与最优维护技术。

基于大语言模型的交通异常事件检测与辅助决策

文 / 李炎英　王新宇　王　晓　孙长银

摘要：大语言模型卓越的数据分析和逻辑推理能力为实时交通管理和辅助决策提供了新思路。ChatGPT 能够高效地处理和分析公开的社交媒体数据，检测数据中包含的城市、道路信息及发生的交通事件，可用于辅助交通管理人员进行实时询问、追溯起因以及探讨应对措施。构建了一个融合社交媒体与 ChatGPT 的智能问答框架 TMGPT（traffic management GPT），研究如何基于大语言模型进行交通异常事件的快速检测并辅助交通管理部门进行决策。该框架通过对社交媒体数据的获取、处理与分析，精准检测交通异常事件并生成针对性应对策略。该框架通过反馈机制持续优化系统性能，为交通管理和政策制定部门提供决策依据，从而提升城市交通运行效率和安全性。实验结果表明，相较于传统方法，TMGPT 框架显著提高了交通异常事件检测的准确率并缩短了响应时间，展现出大语言模型在复杂城市交通管理中的应用潜力。

关键词：ChatGPT；交通事件检测；社交媒体；辅助决策

0　引言

随着我国机动车数量的持续增长以及新型交通参与者（如共享单车及外卖员等）加入现有城市交通系统，城市通行量显著增加，交通拥堵问题日益突出，导致交通事故频发。据《上海市交通行业发展报告（2022）》显示，截至 2021 年底，上海市注册机动车保有量（含“沪 C”牌照机动车）494.2 万辆，较上年增长 5.4%，而 2021 年上海市发生各类道路交通事故共计 52.3 万次，较上年增加 19.8%。尽管城市智能化水平不断提高，交通事故的发生频率并未降低。现有智能交通系统对交通异常事件的检测具有一定的滞后性，交通出行人员无法有效规避，进而造成更大范围的交通拥堵以及更多的社会损失。因此，提升异常交通事件检测效率和异常交通事件的即时反馈，对于城市交通管理具有重要的经济价值与应用意义。

近年来，人工智能技术的发展对社会和民众的生活方式产生了深远影响。例如，社情民意以舆情的方式反馈给管理部门，而管理部门通过抓取和分析社会媒体数据，提高交通管控效率以及民众的出行满意度。基于此，本文提出了一种融合 ChatGPT（chat generative pre-trained transformer）和社交媒体的交通异常事件检测与决策辅助框架——TMGPT。TMGPT 框架充分利用 ChatGPT 在自然语言处理上的优势，基于社交媒体数据的发布时间、文本内容及上下文信息，精确检测交通异常事件的发生地点、时间等关键信息，掌握事件的详细情况并生成应对措施，从而为交通管理者的决策提供更有力的科学支撑。该框架能够显著提高交通事故的处理效率，缩小事故影响范围，并增强城市交通管理系统的健壮性和稳定性。本文的主要贡献如下。

（1）首次提出基于 ChatGPT 对社会媒体数据进行交通异常事

件检测的方法，突破了传统交通异常事件检测方法的限制，为智能交通管理带来了新的研究思路。

（2）基于社交媒体发布的事件内容和背景信息，使用提示词引导 ChatGPT 关注事件的时间、地点等重要信息，有效提升交通异常事件检测的灵敏度和即时性。这种创新性的结合，不仅能够更快速地检测交通异常事件，还能够提供更加详细和准确的事件描述。

（3）通过深入分析交通异常事件信息，ChatGPT 能够计算交通异常事件引发交通事故的概率和危害性，并为交通管理者决策提供智能辅助，从而提升交通管理的效率和安全性。

1 相关工作

ChatGPT 凭借其强大的自然语言分析和反馈能力，在全球引起了广泛关注，并在教育、金融、视觉等多个领域进行应用。交通管理与控制领域的研究者和从业者也开始探索 ChatGPT 在复杂城市交通管理与控制方面的应用。例如，Villarreal 等通过大规模用户实验测试 ChatGPT 能否协助解决复杂的混合交通控制问题。结果表明，ChatGPT 在交叉路口和瓶颈环境下生成策略的成功率显著提高，甚至超过了专家水平。此外，Da 等提出的 PromptGAT 框架将大语言模型与基于提示的动态交通信号建模方法相结合，用于交通信号控制。该框架利用领域知识和实时交通数据来预测交通信号系统动态，并通过强化学习优化控制策略，提升了从仿真到现实的迁移效果。Zheng 等结合大语言模型和跨模态编码器提出了一个基于智能手机的框架。该框架基于新闻文本、图像和音频信息自动生成和分析碰撞报告，使交通事故报告的准确性和效率大大提升。研究表明，大语言模型在城市交通管理方面具有巨大的应用潜力。

目前，传统的交通异常事件检测方法通常依赖有限的数据源，如摄像头和传感设备等，这些数据受到时间和空间限制，无法提供全面、最新的信息。相比之下，社交媒体作为社情民意的感知与反馈平台，以人在环内的方式对交通异常事件的态势演化和社会影响进行补充。社交媒体数据丰富且多样，包括文本、图片和视频等，已被广泛应用于交通异常事件检测。例如，Xu 等基于社交媒体平台上的地理位置信息检测交通事件，利用关联规则过滤无关帖子，再利用时空聚类方法检测交通事件并自动推断事件内容。Ali 等使用本体、潜在狄利克雷分布和双向长短时记忆网络等方法，并结合查询引擎、主题建模、情感分析等技术，从社交媒体数据中提取交通信息、自动识别事件并分析交通事件的情况。Jonnalagadda 等关注交通事件的即时检测，通过分析用户的社交媒体数据精确识别有关拥堵、事故的交通信息。每天有数以百万计的城市居民在社交媒体上分享交通异常事件的信息，为城市交通异常事件检测提供了重要的实时数据。

2 TMGPT 框架

本文提出了一种基于 ChatGPT 的交通异常事件检测与决策辅助框架 TMGPT，旨在提升城市交通管理的智能化水平和决策准确性。TMGPT 框架如图 1 所示，其以社交媒体数据为主要数据源，通过 ChatGPT 与用户的对话式交互，增强对这些数据的分析和处理能力，从而检测交通异常事件，并生成针对性的应对策略。此外，该框架还可提供警力分配、舆情监控和风险分析等多方面的决策支持。

如图 2 所示，TMGPT 的执行流程由 3 个主要部分组成：社交媒体数据获取与处理、交通异常事件检测以及辅助决策。首先，利用网络爬虫技术定向爬取微博评论数据，提取其中的时间和文本内容，并筛选与交通异常事件相关的信息。其次，进行数据预处理，确保数据的质量和准确性。然后，利用 ChatGPT 对预处理后的社交媒体数据进行事件检测，设计合适的提示词以引导

ChatGPT 关注事件的关键要素，如时间和地点。最后，ChatGPT 基于检测到的交通状况生成具体的应对策略，不符合实际情况的建议通过反馈机制进行修正，形成专家知识与 ChatGPT 知识的迭代闭环，从而提高系统的准确性和实用性。

2.1 社会媒体数据

本文利用网络爬虫定向爬取与交通相关的微博评论数据，获取微博发布时间和评论文本。首先利用 Python 的 Request 模块向新浪微博平台发起 HTTP 请求，获取响应中的 JSO 数据。其次，解析 JSON 文件，提取微博评论的发布信息，并将其存储为文本数据。然后，运用 Python 的正则表达式模块执行信息筛选操作，包括但不限于 URL、数字、字母、用户名以及特殊符号的去除。最后，使用 Python 的自然语言处理工具包 Spacy 进行中文分词处理。获取的社会媒体数据包含交通异常事件发生的时间点和其他文本信息，为后续交通异常事件的检测提供了必要信息，保证了 ChatGPT 输入数据的质量。

2.2 事件检测

在事件检测阶段，本文借助 ChatGPT 的文本理解能力，检测异常交通事件。这一阶段的关键在于确保 ChatGPT 准确捕获交通异常事件的关键信息，如事件发生的地点和时间。本文设计了合适的提示词用于引导 ChatGPT 关注交通异常事件的时间和地点等信息，从而提高对交通异常事件的检测准确度。需要强调的是，ChatGPT 的输出结果由专业人员核对，以确保其与实际情况保持一致。对于与实际情况不符的输出，采取迭代反馈机制，将这些信息反馈给 ChatGPT 并重新检测。这种反馈循环机制能够提升 ChatGPT 的性能，尽可能准确地检测交通异常事件。

2.3 辅助决策

在辅助决策阶段，ChatGPT 基于检测到的交通情况，为交警提供具体可行的应对措施。这一阶段通过 ChatGPT 的智能分析和应对措施推荐来支持交通决策。ChatGPT 可以快速分析大量数据，并基于上下文情境提供实用的建议。此外，为了提高方法的可靠性和实用性，本文采用反馈迭代机制，将与实际情况不符的建议反馈给 ChatGPT 进行修正和优化。这种反馈机制不仅提升了 ChatGPT 的性能，还构建了一个有关交通异常事件处理的专家知识与 ChatGPT 知识的迭代闭环。此外，ChatGPT 还可用于交通安全知识宣传、舆情分析以及交通数据优化分析等，这些应用为城市交通管理提供了更多的可能性。

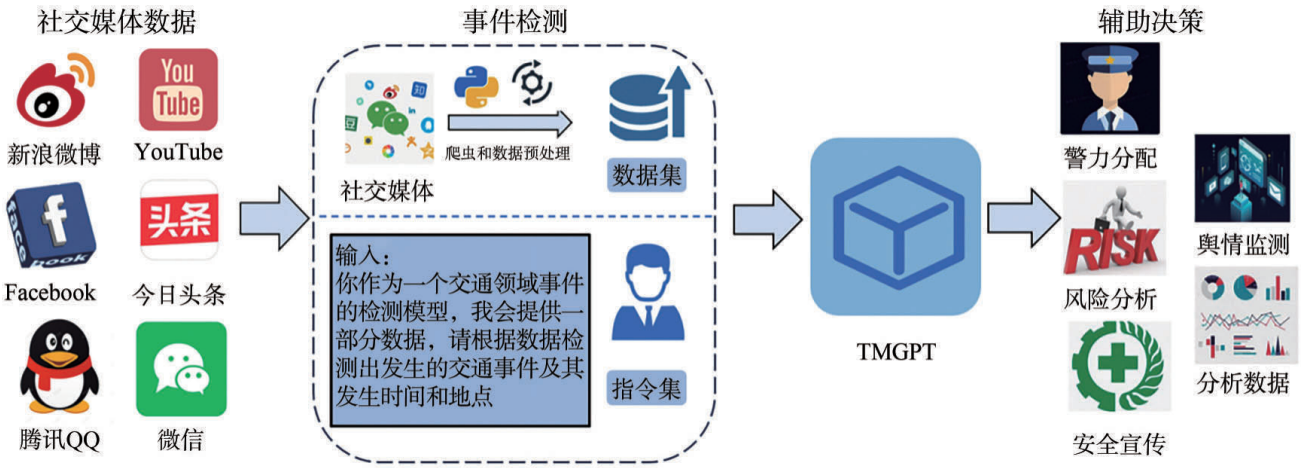


图 1 TMGPT 框架

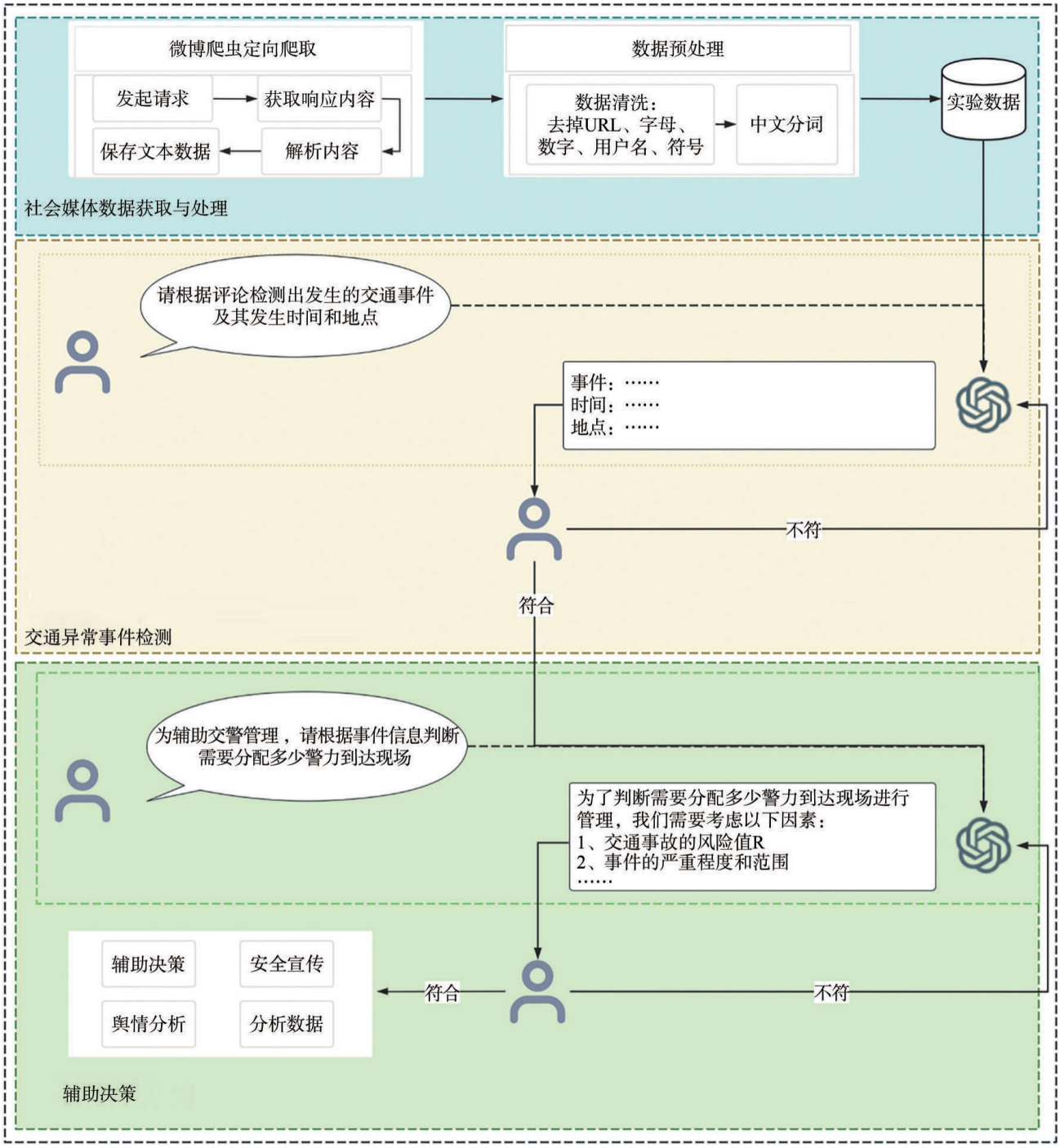


图 2　TMGPT 的执行流程

3　实验设计

本文以上海市交通异常事件检测为例，介绍如何利用 TMGPT 检测交通异常事件以及辅助交通管理部门进行决策。

3.1　微博数据收集和预处理

本文采用 Chen 等提出的方法进行微博数据采集。首先，定义搜索条件，包括关键词“微博”以及“微博内容”，并明确定义时间范围和地理坐标边界。然后，

向新浪微博平台发送 HTTP 请求，获取特定的微博数据。这些数据包括时间戳、用户 ID、微博 ID、文本内容等，如果用户分享了经纬度信息，也包括其地理位置信息。最终，我们收集了 2012 年 4 月 1 日至 2013 年 11 月 30 日期间在上海发布的 550 726 条微博，再从中筛选与交通相关的微博。每条相关微博至少包含一个与交通相关的关键词。我们使用与事故相关的关键词（如车祸、事故、剐蹭、追尾、相撞）和与拥堵相关的关键词（如堵、拥堵、阻塞、塞车）进一步筛选微博数据。我们随机抽取了 1 000 条与交通事故和拥堵相关的微博，按照 8∶2 的比例将其划分为训练集和测试集。

在数据收集完成后，进行数据清洗操作，去除微博文本中的 URL、特定符号、数字、用户名以及停用词等。数据清洗旨在消除噪声和不必要的信息，确保 ChatGPT 输入数据的准确性。删除 URL 和特定符号有助于规范文本内容，去除数字、用户名和停用词有助于减少干扰，提高数据的纯净度，从而提高实验结果的可靠性。数据预处理结果样例如表 1 所示。

3.2 ChatGPT 模型

本文使用的模型为 OpenAI 开发的 ChatGPT 大语言模型，其是基于 GPT-3.5 架构构建的。ChatGPT 的底层技术是 Transformer 模型，该模型已在自然语言处理领域取得显著成效。ChatGPT 采用基于人类反馈的强化学习（reinforcement learning fromhuman feedback，RLHF）进行训练，通过预训练和人工监督相结合的方式进行模型调优。这种方式使 ChatGPT 具有更强的语境理解能力，能够更好地理解上下文语义，显著提高了问题反馈的准确度。作为一种通用的大型预训练模型，ChatGPT 在面对特定领域问题时通常性能不佳，因此引入提示学习以提高其性能。提示学习旨在引导模型在不同特定任务上进行学习，通过为模型设计差异化的提示，使其快速适应不同领域。提示学习的优势在于可以将大规模预训练模型具备的通用知识应用于具体任务上。与训练特定模型相比，提示学习更为高效。此外，提示的设计，可根据需要在不同任务中进行优化。

3.3 提示设计与学习

提示包括 ChatGPT 的角色解释、数据说明、交通信息说明、检测要求以及来自用户的特定输入。本文以交通异常事件的检测为例，具体设计如下。

（1）解释 ChatGPT 扮演的角色，为其提供任务的上下文。这有助于明确 ChatGPT 在整个过程中需要承担的职责和功能。

（2）对模型将要处理的数据进行详细说明，确保 ChatGPT 能够更好地理解和处理这些数据，从而提高分析的准确性。

（3）明确交通事件的具体内容，包括事件的类型、地点和时间。这对于确保 ChatGPT 准确地检测和分析交通信息至关重要。

（4）引导 ChatGPT 特定关注交通事件的时间和地点，并对这些要素进行约束，以确保检测的可靠性。

（5）输入具体的交通数据，ChatGPT 根据前面的提示和输入的数据进行反馈输出。

提示设计见表 2。

表 1 数据预处理结果样例

微博发布时间	微博正文	纬度	经度	是否有经纬度信息
Sat Aug 18 14：08：07 + 0800 2012	沪嘉现在是露天大型停车场，堵的没脾气。	None	None	No
Thu May 23 08：19：32 + 0800 2013	南浦大桥堵车堵到令人发指！！！我能用走的吗	121.499701	31.208492	Yes

表 2　提示设计

你作为一个交通领域的事件检测模型，我将给你提供一部分数据，请你根据要求检测出数据中的交通异常事件及其地点和时间信息，并输出结构化的数据表示。输出格式为“交通异常事件：{事件信息}；地点：{地点信息}；时间：{时间信息}”。除此之外，请不要输出任何多余的文字。 数据说明：提供的数据是发生交通异常事件时，用户在微博平台上发布的相关帖子，包括微博发布时间、微博正文、帖子经纬度等信息。 交通信息说明： 交通异常事件：与正常交通流模式显著不同的情况，这些异常情况可以通过监测交通流量、速度、密度等参数来识别。交通异常事件通常由外部或内部因素引起，可能包括非正常的高峰时段交通流量、突发的交通拥堵、特殊天气情况以及大型公众活动等。 地点：发生交通异常事件的地点。 时间：发生交通异常事件的时间。 检测要求： （1）检测地点信息，并转换成对应的经纬度。如果数据中的地点信息缺失或者模糊不清，则识别地理标签数据作为地点。 （2）如果数据中的时间信息缺失或者模糊不清，则结合发帖时间进行推理，并说明是发帖时间。 （3）如果不同文本表达同一事件，则统一为一个结果输出，时间用时间段表示。 给定数据：…… 检测结果：……

3.4　评价指标

本文使用准确率（Precision）、召回率（Recall）、F1、错误检测率（Error）、平均响应时间来评估模型的有效性，各项指标的计算式如下：

$$\text{Precision} = \frac{\text{correct}}{\text{out}} \tag{1}$$

$$\text{Recall} = \frac{\text{correct}}{\text{number}} \tag{2}$$

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

$$\text{Error} = \frac{\text{error}}{\text{out}} \tag{4}$$

其中，correct 表示系统正确检测出的异常事件数，number 为数据集中异常事件的总数，out 为数据集中手动标注的异常事件数，error 为系统错误检测的异常事件数，平均响应时间是指系统从接收到检测请求到返回检测结果所需的平均时间。

4　实验结果和分析

4.1　事件检测

以上海市南浦大桥的交通拥堵事件为例，进行案例分析，以此展示 TMGPT 的具体执行流程。案例的检测结果如表 3 所示。首先，将包含数据的提示输入 ChatGPT 模型。然后，该模型依据输入信息进行事件检测，精确辨识发生的交通异常事件，并确定事件发生的地点以及时间。地点的经纬度信息有助于交警部门直接定位事发地点，从而开展应急管理工作。

本文在数据集上分别利用 TMGPT 框架和 MASCA 算法进行实验，其中 MASCA 算法是一种传统基于词相关性特征的社交网络突发事件检测方法。实验结果如表 4 所示，本文方法的准确率为 0.76，召回率为 0.8，F1 值为 0.7794，单条数据的平均响应时间为 1.2 s。与 MASCA 算法相比，本文提出的 TMGPT 框架在准确率、F1 值以及平均响应时间方面均表现出明显优势。这种提升表明，TMGPT 框架在交通异常检测中能够更精确地捕捉突发事件，并显著提升了处理效率。这也进一步验证了大语言模型在数据处理上的强大优势，特别是在应对复杂场景时的高效性和准确性。

4.2　决策辅助分析

在检测到上海市南浦大桥的交通拥堵事件后，利用 TMGPT 框架进行事故风险评估，风险评估结果可

表 3　案例的检测结果

用户	你作为一个交通领域的事件检测模型，我将给你提供一部分数据，请你根据要求检测出数据中的交通异常事件及其地点和时间信息，并输出结构化的数据表示。输出格式为“交通异常事件：{事件信息}；地点：{地点信息}；时间：{时间信息}”。除此之外，请不要输出任何多余的文字。
	数据说明：提供的数据是在发生交通异常事件时，用户在微博平台上发布的相关帖子，包括微博发布时间、微博正文、帖子经纬度等信息。
	交通信息说明：
	交通异常事件：与正常交通流模式显著不同的情况，这些异常情况可以通过监测交通流量、速度、密度等参数来识别。交通异常事件通常由外部或内部因素引起，可能包括非正常的高峰时段交通流量、突发的交通拥堵、特殊天气情况以及大型公众活动等。
	地点：发生交通异常事件的地点。
	时间：发生交通异常事件的时间。
	检测要求：
	（1）检测地点信息，并转换成对应的经纬度。如果数据中的地点信息缺失或者模糊不清，则识别地理标签数据作为地点。
	（2）如果数据中的时间信息缺失或者模糊不清，则结合发帖时间进行推理，并说明是发帖时间。
	给定数据：Thu May 23 08：19：32+08002013 南浦大桥堵车堵到令人发指！！！ 我能用走的吗 121.499701 31.208492 Yes
ChatGPT	交通异常事件：南浦大桥堵车堵到令人发指！！！ 我能用走的吗；
	地点：121.499701，31.208492；
	时间：Thu May 23 08：19：32 +0800 2013。

表 4　交通异常事件检测结果

模型	准确率	F1	召回率	错误检测率	响应时间（s/ 条）
MASCA	0.64	0.77	0.6615	0.3385	2.2
TMGPT	0.76	0.80	0.7794	0.2206	1.2

为交通管理者提供决策支持。TMGPT 框架根据事件信息构建了一套基础的风险计算模型，风险评估结果见表 5。根据事件描述，假设事故发生的概率为 P（如 0.8），事故的危害度为 D（如 0.7），通过 $R=P\times D$ 计算出事故的风险值 R 为 0.56。风险值的取值范围为 0 到 1，用以推测事故发生的可能性。除此之外，TMGPT 提供推测过程的判断依据，指出交通事故的发生与交通拥堵、车辆状况及驾驶员的精神状态密切相关，这与已知的交通事故诱发因素一致。

根据 TMGPT 的风险评估结果，对南浦大桥的交通拥堵事件进行辅助决策分析，如确定现场所需的警力部署。如表 6 所示，TMGPT 综合考虑了交通事故的风险值、事件的严重程度和影响范围以及现场处理需求，进而得到警力的最优分配方案。在此过程中，TMGPT 构建了一个基础模型，用于计算警力配置。例如，通过以下计算式确定警力数量：警力数量 = 基础警力 +（风险值 × 调整因子），以此确保警力调配与事故风险相匹配。

表 5　TMGPT 风险评估结果

用户	交通异常事件往往会引发交通事故，因为它们扰乱了正常的交通流动模式和规律，显著增加了交通事故发生的风险。请根据事件信息计算该地区发生交通事故的概率 P 和危害性 D，并计算出交通事故的风险值 R，并给出判断依据。
ChatGPT	要计算发生交通事故的风险值 R，我们需要先定义事故发生的概率 P 和事故的危害度 D 的计算方法。可以假设一些基础计算模型。
	概率 P：
	根据事件信息“堵车堵到令人发指”，我们可以推测此类严重拥堵事件通常会显著增加交通事故的风险。假设此类严重堵车导致事故的概率较高，我们设定 P 为一个较高的值，例如 0.8。
	危害度 D：
	交通事故的危害度可以根据事故的潜在影响来估计。严重拥堵可能会导致连环追尾、多车碰撞等严重事故。假设危害性 D 在一个较高的范围内，我们设定 D 为一个较高的值，例如 0.7。
	风险值 R：
	风险值 R 可以通过 P 和 D 的乘积来计算，即 $R=P\times D$。
	根据上述假设和计算方法，我们可以得出：
	概率 P：0.8
	危害性 D：0.7
	风险值 R：$R=P\times D=0.8\times 0.7=0.56$
	判断依据：
	● 严重的交通拥堵显著增加了交通事故的风险，因为车辆的速度差异、驾驶员的焦虑和急躁等因素都会影响交通安全。
	● 历史数据和交通研究表明，严重的交通拥堵通常会增加交通事故发生的可能性。
	● 交通事故的危害度评估考虑了可能发生的事故类型及其潜在影响，例如多车连环相撞可能导致严重的财产损失和人员伤亡。

表 6　TMGPT 辅助决策分析结果

用户	此时，为了辅助交警管理交通道路，请根据该交通异常事件信息判断需要分配多少警力到达现场，并给出判断依据。
ChatGPT	为了判断需要分配多少警力到达现场，我们需要考虑以下因素。
	交通事故的风险值 R：根据上一步计算，得到风险值 R=0.56。
	事件的严重程度和影响范围：严重拥堵可能涉及较大范围，并且有较高的事故风险。
	现场处理需求：通常包括交通疏导、事故处理、救援和秩序维护等。
	假设我们有一个基础模型来分配警力数量 N，计算式为警力数量 = 基础警力 +（风险值 × 调整因子）。
	假设：
	基础警力数量为 5 名警力（用于基础的交通管理和事故处理）。
	调整因子与风险值 R 成正比，假设为 $R\times 10$。
	那么，$N=5+(0.56\times 10)=5+5.6=10.6$。
	根据实际情况，警力数量应取整，即需要分配 11 名警力到现场。
	判断依据：
	● 风险值 R：由于风险值接近中高（0.56），表明现场情况较为严重，需要较多的警力。
	● 事件的严重程度：事件描述为“堵车堵到令人发指”，表明拥堵程度极为严重，可能涉及大量车辆，需要较多警力进行现场管理和疏导交通。
	● 综合处理需求：处理严重拥堵和潜在的交通事故，需要考虑现场秩序维护、事故处理和救援等多方面需求，应合理增加警力数量。
	综上所述，建议分配 11 名警力到达现场进行交通管理和事故处理。

5 总结和展望

本文提出了一种基于 ChatGPT 的交通异常事件检测与决策辅助框架 TMGPT。TMGPT 框架不仅能帮助交警部门检测交通异常事件，还能提供决策支持。通过大数据分析和自然语言处理技术，TMGPT 为交警提供决策建议，如警力及其他资源的分配。此外，TMGPT 在促进公众参与方面也具有一定潜力，公众在社交媒体分享的交通信息可为交警提供参考，交警亦能通过这些平台与公众互动，及时传达重要信息。

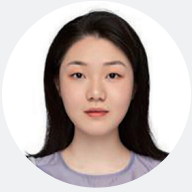
该框架仍处于初步探索阶段，未来的研究方向包括：①将大语言模型与交通数据处理相结合，如整合交通流量、天气数据及实时监控视频等多源数据，充分挖掘其潜在价值，为城市管理提供更有力的支持；②实现交通异常事件的预测，极大改善城市交通环境。

为了充分发挥 TMGPT 框架在交通领域的潜力，我们希望与交通管理部门建立紧密的合作关系。交通管理者的反馈对于 TMGPT 框架的输出结果至关重要，有助于提升模型的性能，提高其在实际应用中的可靠性。此外，在使用 TMGPT 框架处理敏感数据和信息时，必须采取严格的安全措施确保数据的安全性。我们希望通过增强 TMGPT 框架在交通领域的处理能力，加强与交通管理部门的合作，以及持续关注安全隐私保护问题，为城市交通的智能化和科学化发展提供有力支持，为城市交通的安全和畅通作出贡献。我们期待与各方共同努力，积极推动智慧城市、智慧交通和自动驾驶等领域的发展，为城市智能交通系统提供更多参考。

来源：智能科学与技术学报

作者简介

李炎英（2000— ），女，安徽大学人工智能学院硕士生，主要研究方向为社会计算、生成式人工智能、交通事件检测。



王新宇（2003— ），女，安徽大学人工智能学院本科生，主要研究方向为交通事件预测、生成式人工智能。

王 晓（1988— ），女，安徽大学人工智能学院教授，主要研究方向为社会计算、群体行为建模、无人自主系统及其平行测试。

孙长银（1975— ），男，安徽大学人工智能学院教授，主要研究方向为智能控制与优化、强化学习、神经网络和数据驱动控制。

集值系统的辨识与控制导语

文 / 中国科学院系统科学研究所　赵延龙

随着信息化、大数据和智能化的发展，集值系统大量涌现，对系统建模和控制产生了新的需求。集值系统的特点是无法测得输出、状态等系统数据的精确值，只知道它是否属于某个或某些集合。例如：大量医学数据如白血病数据只知道“健康”或“疾病”，大量工程数据如卫星观测数据只知道“百米级”“千米级”等，智能决策问题如雷达目标识别关心的是目标的“真”或“伪”。与传统的精确测量系统相比，集值系统具有可利用信息少和非线性强两大特征，其辨识和控制具有本质困难，需要新的研究视角和理论方法。

一、引言

集值系统是指无法测得系统数据的精确值，只知道它是否属于某个集合中的系统。它是随着信息化、生物技术和大数据的发展而诞生。产生集值系统的原因主要有如下四点：

1）生物体的固有特征或工程设备的局限

神经元是生物体的基本元素，它由一个细胞体和突触两部分组成。突触分轴突和树突。轴突把本神经元的输出发送至其他相连接的神经元；树突与其他神经元的轴突相连，以接收来自其他神经元的生物信号。轴突和树突共同作用，实现了神经元间的信息传递。神经元把所有接收到的信息加权求和得到一个总效应。如果总效应大于该阈值电平，那么神经元表现为兴奋状态，反之则为抑制状态。可见，神经元就是一个典型的集值系统。

2）数据传输、储存和保密的要求

在数据传输和存储之前首先要将模拟信号转化为数字信号。量化后的数据经过编码后，得到的是低精度的集值数据。众所周知，在数据传输过程中，越小的集值步长会产生越大的数据量，而数据传输会受到传输带宽的限制，在一定带宽下，过多的数据会造成信道的拥塞，使数据传输产生很大的延时甚至失真。在这种情况下，即使每个传感器得到的数据很精确，由于受传输带宽的影响，在传输之前也不得不将信号进一步转换成精度非常低的集值数据。

3）关注的对象是集值

食物和光照是影响绒孢菌的主要因素，研究表明绒苞菌厌光并且喜欢食物丰富的环境。为了考察两者中哪一个对绒孢菌的影响更大，设计决策实验，选取两个区域，一个是光照强但食物丰富的区域，另一个是光照弱但食物缺乏的区域。随后，将绒孢菌均匀分散到两个区域中，经过一段特定时间之后，考察两边绒孢菌的个数，从而反映食物和光照对绒孢菌的影响强弱度。绒孢菌对食物丰富的环境感兴趣，当关心它最终选择哪个区域，而不是它选择的快慢，那么关心的就是一个集合信息。

4）追求精确数据存在弊端

精确数据对人们更好地了解系统和控制系统是很有帮助，但为了测量精确数据有可能会对系统造成破坏，比如光合作用实验和量子系统。为精确地测量叶片在光合作用的内部数据，现有技术是对叶片进行切片处理，测量叶片内部的反应物和生成物的浓度。不难看出，这种实验对叶片是一种破坏且不可恢复。为了测量每个时间的实验数据，需要给很多叶片相同的培养环境，在所需时间的节点对其中一部分做切

片，从而描绘出反应的变化过程。

然而，考虑到叶片个体的差异，这种做法的实验效果一直不理想，且无法对同一叶片重复实验。再如量子系统，为了测量量子系统的精确状态，需要对系统进行测量，然而每次测量都会引起量子态的变化，也就是遵循海森堡测不准原理，因此是否测量，如何测量一直是研究关注的焦点。可以看出，精确测量对系统造成了破坏，如果仅利用叶片毛孔的开关、叶片吸入氧气还是呼出二氧化碳、量子处于哪一类状态等集值数据就可以分析系统内部特征的话，那就会减小这一破坏。

低精度的集值数据在实际应用中起到了越来越重要的作用，从研究的角度而言，以上诸多例子都可以归结为图 1 中的系统。

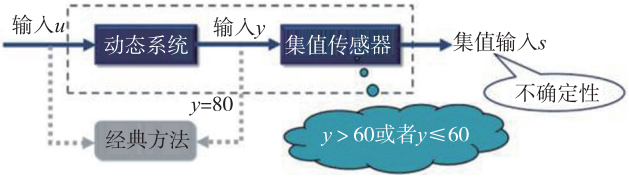


图 1　集值系统结构

二、主要研究内容

从结构上来看，集值系统相比传统系统最大的不同是多了集值传感器部分，它可以是实际工业中的氧传感器、智能传感网络，也可以是虚拟的量化器以及生物中的比较器。集值传感器使得我们在辨识的时候，无法获取精确输出值，比如 y = 80，只能获得一些粗糙的集值数据，比如 y > 60 或者 y ≤ 60。

在粗糙的集值数据下，如何实现高精度的系统辨识是实际中关心的基本问题。集值测量数据与系统输出之间的误差，既不是有界的，也不具备良好的统计特性。集值数据可用信息少，而且测量器件的非线性强。因此，无法将集值系统放到已有的理论框架下研究。另一方面，已有的辨识的方法，如最小二乘、卡尔曼滤波等算法，都是基于系统输出构建的，而集值数据下，系统输出是不知道的。所以，要研究集值系统，就必须提出新的理论和方法。

该研究方向的核心问题是如何利用粗糙的集值数据完成对目标系统的精确辨识或控制？所涉及的主要科学问题包括：如何做到集值系统的辨识、滤波以及适应控制？如何融合不同类型的数据，达到更好的辨识或控制效果？基本科学贡献点在于集值系统的可辨识性、辨识的最优性、控制的最优性、数据精度与辨识或控制精度的定量关系等。

集值系统辨识与控制研究具有理论和实际的双重重要意义。从理论角度来讲，集值数据下的系统辨识和控制研究可以建立一套与现代控制理论平行的系统建模、辨识和适应控制的完整理论体系，这是控制论在新时代背景下的发展。同时，可以定量描述数据精度与系统辨识和控制的相互关系，从控制的角度研究数据精度，给出最优测量方法以提高数据的使用效率。从实际应用角度来讲，集值信息将极大减弱系统辨识和控制对数据精度的依赖，从而可以有效提高数字传输的效率，降低测量元件的成本，为与集值系统相关的新兴技术提供新的理论支持。

2.1　集值系统的辨识

针对不同的系统噪声、不同的阈值信息，可分别构造几种典型系统的辨识算法，并分析算法的收敛性、时间复杂性以及最优性。

2.1.1　基于二集值输出的线性系统辨识及输入设计方法

针对集值数据，团队提出了满值输入设计方法和经验分布函数法，给出输入信号充分丰富的条件，并且构造了强一致收敛的辨识算法。进而，研究了其他更复杂的情形，例如，系统不仅有输出噪声还有输入测量噪声以及输入激励噪声，二集值传感器的阈值未知，噪声的分布函数未知等。

对最为基本带输出随机噪声的二集值输出的线性系统，辨识的基本困难在于二集值的输出信息少，团队提出满值输入设计和经验分布函数法，给出输入信号充分丰富的条件，并可构造强一致收敛的辨识算法。另外，对于系统不仅存在输出噪声，而且存在输入噪

声和输入量测噪声的情况下，以及阈值未知或噪声的分布函数未知时，均可以构造强一致收敛的辨识算法。

2.1.2　二集值输出的 Wiener 系统辨识

对具有二集值输出的典型非线性系统—Wiener 系统，团队提出了比例满秩信号和联合可辨识性的方法，克服了二集值输出信息少以及非线性给系统辨识带来的困难，构造了辨识算法，并分析了算法的最优性。考虑图 2 中的 Wiener 系统，系统输出被阈值为 C 的二集值传感器测量，通过传感器的输出只知 y（k）≤ C 还是 y（k）> C。

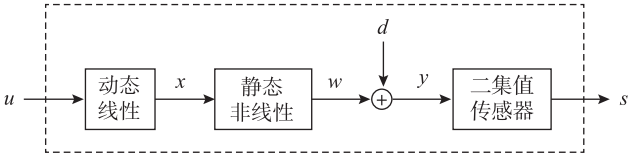


图 2　基于二集值输出的 Wiener 系统

针对非线性，团队提出比例满秩信号和联合可辨识性的概念。输入信号的比例特性使得系统的中间变量具有相同的比例性质。结合成比例的状态和二集值的输出信息就构成了关于非线性部分辨识的“核心辨识问题”，而联合可辨识性则保证了核心辨识问题的解决，可以证明关于核心辨识问题的算法是渐近最优的。这样，就把非线性部分的参数辨识出来了。利用已有的非线性部分参数的估计，结合输入信号的满秩特性，对线性部分的参数进行估计。

2.1.3　集值输出的 Hammerstein 系统辨识

对如图 3 所示的具集值输出以及多项式形式静态非线性的 Hammerstein 系统，在不要求非线性部分单调的前提下，针对系统的结构特点，提出比例强满秩信号的概念，构造了辨识算法，分析了算法的最优性。

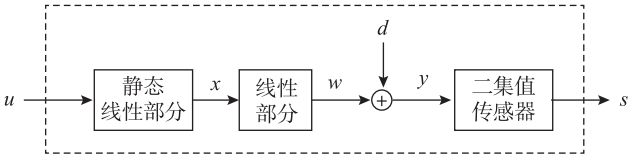


图 3　集值输出 Hammerstein 系统

在二集值输出观测的前提下，辨识 Hammerstein 系统比辨识 Wiener 系统还要难，这是因为 Hammerstein 系统与二集值传感器这种非线性元件结合之后，非线性同时存在于输出端和输入端，这为系统辨识带来了一定的难度。为了研究二集值输出的 Hammerstein 系统辨识，就需要探索新的思路和方法。考虑图 3 中的 Hammerstein 系统，团队重点研了 Hammerstein 系统具有多项式形式的非线性部分，且系统的输出被二集值传感器测量。用强满秩周期输入，可以有效克服两端非线性给辨识带来的困难。

2.1.4　递推投影辨识算法

由于二值系统的适应控制问题要求一种在线辨识算法，现有文献［7］提出了递推投影辨识算法，给出了相比已有的递推投影辨识算法更快收敛速度的结果。在算法的收敛性分析时，估计的均方误差包括状态项、中间项和新息项，新息由二值化信息和噪声分布函数构成，且投影确保了估计的有界性。由于中间项和新息项的这些特征，可以得到整个估计的均方误差收敛速度的数量级。

2.1.5　适用更一般条件的递推辨识算法

上述的递推投影辨识算法存在算法设计依赖于噪声分布、仅适用于 FIR 模型、要求输入是持续激励的、待辨识参数的取值范围事先知道等限制，如何放宽上述限制并设计有效的递推辨识算法是值得研究的问题。以下文献［9-11，13，14］从更广的噪声、更一般的系统、更宽的输入、更少的先验等四方面介绍团队在递推辨识算法方面的研究进展：

更广的噪声条件：目前几乎所有量化辨识的文献仅在随机噪声或确定性噪声的前提下研究基于量化数据的参数辨识，缺乏统一两种框架的递推辨识算法。团队基于一种自适应阈值的设计，克服了量化系统在确定性框架中缺乏激励而难以辨识的难点，构造了首个同时适用于确定性噪声和随机噪声的递推在线辨识算法，并分别给出了算法的收敛性分析。

更一般的系统：团队针对在持续激励和固定阈值

下具有二值观测的 ARMA 模型参数辨识问题，克服了回归项信息太少的难点，提出了一种仅使用二值观测值的在线识别算法，包括参数估计以及系统输出值的预测，同时利用耦合的 Lyapunov 函数，最终证明在一定条件下该算法是收敛的，收敛速度可以达到与迭代次数的倒数同阶，与之前 FIR 系统参数辨识结论的阶次相同，并且当 ARMA 模型退化到 FIR 模型，相应收敛速度达到的条件也相同。

更宽的输入条件：目前集值辨识算法大多要求输入数据的持续激励性，但这在反馈控制系统中很难满足，为此，团队基于时变投影的设计了拟牛顿型递推辨识算法，其步长由标量变为矩阵形式，使得其数据信息更充分的融合到步长设计中，在最弱激励条件下建立了算法的几乎处处收敛性。

更少的先验信息：针对具有二值观测的 FIR 系统辨识问题，团队首次在无需参数先验信息前提下提出了无投影在线量化辨识算法。创新性地利用算法尾分布的性质，攻克了压缩系数与估计耦合的难点，并在非周期信号条件下给出了目前最优收敛阶的几乎处处收敛速度和均方收敛速度。在此基础之上进一步考虑了衰减激励输入、未知噪声分布和缺少参数位置先验信息情形，分析了集值系统的可辨识性问题，并设计了“分解－重构”型联合辨识算法，实现了对集值系统未知参数及噪声分布参数的联合在线辨识。

2.1.6　最优递推辨识算法

随着集值辨识理论体系的初步建立，一个自然的疑问是如何实现集值辨识算法的性能极限，即如何设计最优辨识算法。由于集值观测的强非线性，其均方损失函数呈现非凸特性，无法通过求解最小损失函数给出最优算法设计。但注意到 Cramer-Rao 下界可作为估计好坏的评判规则，因此团队在文献［8，12］中建立了集值系统的 Cramer-Rao 下界来刻画集值数据的信息量。通过分析 Cramer-Rao 下界的特性，基于自适应权重的设计构造了集值观测下的基于信息矩阵的拟牛顿型加权辨识算法。利用时变投影的设计和高阶矩“加速”思想，攻克了不可交换随机矩阵乘积和步长参数与估计存在强耦合的分析难点，建立了算法的最优收敛速度和渐近有效性。

2.2　集值系统的控制

在集值系统的控制方面，主要有适应控制和同步控制这两方面的研究。

2.2.1　适应控制

集值输出系统的适应控制问题与单纯的参数辨识和参数已知时的镇定控制相比要困难得多，原因之一是适应控制与参数辨识算法相互依赖，已有文献中关于集值输出系统的辨识算法的良好性质是以周期输入为前提的，而适应控制中的控制律是根据控制目标和参数估计值而确定的，因此，一般来讲是随机的且不具周期性；原因之二是适应控制下闭环系统更为复杂，表现为（即使对简单的线性定常系统）闭环系统的高度非线性和输入、输出的复杂随机性，输出端的集值信息和估计值的关系也相应地更为复杂。

随着集值系统辨识方面的研究积累和成果系统化，集值系统逐步受到国内外学者的关注，代表性的工作有：1）扩展的极大似然方法解决集值系统的辨识问题，提出了固定数据长度下的无穷步迭代算法；2）量化滤波算法，该方法是以量化误差服从正态分布为基础的；3）鲁棒辨识的方法研究有界噪声下的参数估计问题。然而已有针对集值系统方法由于自身限制均无法做到适应控制。

针对集值系统的适应控制，一个可能可行的思路是：首先将已有工作中的周期输入推广为一般的持续激励输入，在不限制集值信息与参数估计值的相关性的前提下，直接构造在线的面向控制的辨识算法，然后以这些算法为基础去设计适应控制器，进而借鉴传统的自校正调节思想，完成集值输出系统的适应控制设计和闭环系统的性能分析。

对于一般的集值输出系统，设计这样的控制器是非常困难的。作为这方面的初步性工作，现有文献仅考虑一类单参数的二集值输出系统的适应跟踪控制。

我们用一类投影递推算法辨识系统的未知参数，用参数估计和控制输入来调整系统输出的阈值，根据“必然等价原则”构造适应控制律，在较一般的未知参数先验信息、噪声统计特性和被跟踪信号等条件下，给出了算法的收敛性和收敛速度，证明了闭环系统的稳定性和跟踪的渐近最优性。现有文献将该方法推广到一般线性系统中。

2.2.2　同步控制

多个体的同步控制问题近年来引起了广泛的关注，因为它的应用范围非常广泛，比如多处理器的负载平衡、移动车辆的协同控制，以及无人地面、空中和水下的交通工具的控制。同步问题旨在设计一种分布式协议，使所有个体都渐近地达到一个共同的状态，这在合作控制的背景下是非常重要的。

信息是局部的是多个体系统的主要特征之一，每个个体只能获取其邻居的状态信息，但精确状态值的获取需要非常高质量的网络通信能力，而实际通信过程中往往存在噪声和带宽的限制。因此，在研究多个体同步问题时，有必要考虑除了局部性限制以外的其他因素，即噪声和量化。

考虑含有 n 个体的多个体系统，为了实现同步，每个个体需要与邻居个体传递信息，但是由于带宽等因素的影响，邻居个体之间只能传递如下二值信息。为了实现同步，每个个体都需要通过使用其邻居的状态来设计同步控制。但是邻居的状态是未知的，所以每个个体都需要估计邻居的状态，然后根据邻居状态的估计值来设计同步控制。由于每个个体只能获得邻居的二值信息，一种方法是每个个体等待一段时间来收集足够的二值信息，然后利用无截断辨识法对其邻居的状态进行估计；另一种方法是利用递推投影辨识法对邻居节点进行估计，该算法是在线辨识算法，估计实时更新，个体状态不需要保持一段时间不变，那么控制也就可以进行实时更新。

1）二时间尺度同步算法：在小的时间尺度上，每个个体先估计其邻居的状态；在大的时间尺度（一些跳跃的时间点）上，个体根据状态的估计值设计控制率，在该控制下个体的状态会进行更新。重复上述估计和控制过程，就形成了二时间尺度同步算法。从这个算法的结构可以看出，当时间增大到一定程度时，个体用于估计的时间越长，估计会越来越准。但与此同时，在估计上花费的时间就越多，这使得个体状态的更新速度会很慢，可能会影响算法整体上的收敛速度。

2）递推投影同步算法：该算法在对邻居节点状态进行估计时，利用递推投影辨识算法，每个个体可以实时地估计邻居个体的状态，然后利用平均原则，用估计值代替邻居状态真值，并取增益步长 1/（t＋1）来，实时设计同步控制律。虽然递推投影辨识算法是对固定参数估计的一种算法，状态实时变化时估计性质不一定很好，但是同步控制可使得状态趋于一致，个体的状态在后期就不会有太大变化，这样递推投影辨识算法依然可以实现一个好的估计。最后，通过估计和控制的合作，多个体系统就可以实现同步。

在算法的收敛性分析时，我们分别分析了估计和状态更新的性质，然后联合考虑这两者之间的相互作用关系，最后证明了算法的收敛性和收敛速度。同时，也在理论上证明了递推投影同步算法相比二时间尺度同步算法具有更快的收敛速度。目前，团队将事件驱动机制引入控制器设计中进一步减少通信成本和控制成本。

2.3　集值系统的应用

2.3.1　复杂疾病建模

复杂疾病系统就是一类典型的集值输出系统，样本属性只是“疾病”或“健康”，样本的观测维度极高，而且样本量相对较少。要建立复杂疾病的模型，从医学的角度研究人类疾病与遗传的关系，核心问题在于如何利用海量遗传数据找出各种遗传性疾病的遗传规律、发病机制、诊断和防治措施。基因测序技术的不断发展使得研究者可以获得更全面更准确的遗传信息，比如现下流行的全基因组关联分析（GWAS）

能够检测数以百万计的单核苷酸多态性（SNPs）作为单个样本的遗传标记。然而，高昂的测序成本极大地限制了测序的样本数量。数以千计的样本数量与数以百万计的观测维度使得许多传统方法不再有效。

在复杂疾病建模和统计验证方面，我们与美国圣朱迪儿童研究医院合作，基于实际数据成功建立了白血病的集值模型，并构造了新的、更为有效的统计验证方法。

2.3.2　雷达目标识别

雷达目标识别是军事领域的一项重要任务。伴随着弹道导弹制造技术在全球的扩散，出现了反弹道导弹系统，为了突破反导系统，现代弹道导弹都具备先进的防突设施，其中一项重要的措施就是释放多种形式的诱饵，增加反导系统的拦截负荷。弹道导弹经过攻防之间不断的交互竞争过程，现面临的核心问题就是如何从目标群中识别出真弹头。对于雷达目标识别，出于成本与安全的考虑，不可能通过大量实际试验来获取观测样本。为了提高识别精度，如何利用好小样本获取尽可能精确的估计结果成为一项重要难题。

团队与相关单位合作完成了对真假目标的集值建模。首先对雷达数据进行预处理，将大通量数据降维得到一些特征信息，使得这些特征信息可以尽可能多体现目标与其他物体的差异；然后利用集值辨识等方法对目标进行分类识别。与已有方法相比，得到了更高的识别率，尤其是在低信噪比时，识别率提高得更为明显，可高达 10%。

2.3.3　无拖曳卫星控制

无拖曳控制技术是重力梯度测量卫星的关键技术之一，其目的是为了补偿在轨卫星受到的干扰力及力矩，使得卫星在地球重力场的作用下运行。无拖曳控制的理论显得非常直接，但真正实现高精度或者极低“剩余”加速度的无拖曳控制还面临诸如模型信息缺失、重力梯度仪测量信息饱和受限、对控制精度要求极高等难题。

团队与相关单位合作以重力梯度卫星为对象，研究了无拖曳控制中的干扰估计和位姿联合控制等问题。在重力梯度仪的量程范围内插入多个阈值，从而将精确与量化混杂的饱和约束观测数据近似为多值量化数据，因而可以引入递归投影随机逼近型多值量化数据参数辨识算法实现对阻力增益系数等未知参数的辨识。算法估计值在闭环系统中强一致地收敛到真实值，并得到与精确数据下参数估计算法同阶的几乎处处收敛速度。基于得到的阻力增益系数等参数的估计值，进一步设计带投影的自适应控制器，实现了渐近最优的自适应跟踪控制。

三、总结与展望

集值系统在网络化和信息化大环境下涌现出来并得到了广泛应用，其系统辨识和控制研究可以建立一套与现代控制理论平行的系统建模、辨识和适应控制的完整理论体系，是控制论在新时代环境下的发展。从应用角度来看，集值信息能够极大减弱系统辨识和控制对数据精度的依赖，从而可以有效提高数字传输的效率、降低工业中测量元件的成本，为生物实验和量子系统等与集值信息相关的新兴技术提供新思路。

集值系统的辨识与控制是一个新兴的研究领域，有很多有价值的问题亟待解决。比如：1）对集值系统进行辨识以后，如何利用辨识得到的信息来解决一般信号条件下的适应控制问题？2）除了考虑集值量测限制外，通信过程中往往会存在时延和丢包，如何在考虑这些因素的情况下实现同步？3）进一步，如何刻画数据与辨识和控制任务的定量关系，即如何在给定数据条件下实现最优的辨识或控制效果，抑或是如何刻画实现辨识和控制任务所需的最小数据条件？这些问题都具有一定的难度，需要众多科研工作者的广泛关注和积极参与。本文的工作还只是抛砖引玉，相信随着大家对集值系统研究重要意义认识的逐步深入，此方向的研究必将得到更快的推进和发展。

来源：集智俱乐部

2024 中国自动化大会圆满落幕

贤者云集，相聚青岛。2024 年 11 月 2-3 日，2024 中国自动化大会在青岛世界博览城国际会议中心举行。本届大会由中国自动化学会主办，青岛科技大学、国家橡胶与轮胎工程技术研究中心承办。大会以“新质发展 智控未来”为主题，邀请了 10 余位院士和 300 余位高校、科研院所校长、院长等学术精英与会，3000 余名来自学术界和产业界的代表参会，近千万人次在线观看了直播。大会共设 10 场大会主旨报告和 46 个平行会议，涵盖工业互联网、脑机智能、智慧教育、智能机器人、智慧能源等多个前沿热点领域，300 余位专家学者作会议报告，共接收会议论文 1500 余篇，10 余家行业知名厂商赞助支持。

大会总主席，中国工程院院士，中国自动化学会会士、理事长，西安交通大学教授郑南宁；中国工程院院士，中国自动化学会会士、特聘顾问，浙江大学教授孙优贤；中国工程院院士，中国自动化学会会士、特聘顾问，中国科学院沈阳自动化所研究员王天然；中国工程院院士，鹏城实验室主任，北京大学教授高文；中国工程院院士，中国自动化学会会士、副理事长，中南大学教授桂卫华；中国工程院院士，中国自动化学会会士、监事，华东理工大学教授钱锋；中国科学院院士、中国自动化学会会士、哈尔滨工业大学教授段广仁；中国科学院院士，中国自动化学会会士、副理事长，中国空间技术研究院研究员杨孟飞；中国工程院院士，中国自动化学会会士、副理事长，中国科学院沈阳自动化所研究员于海斌；欧洲科学院院士、中国自动化学会会士、澳大利亚斯威本科技大学副校长韩清龙；中国橡胶工业协会会长徐文英；英国皇家工程院院士、剑桥大学教授 David Cebon；瑞典皇家理工学院教授 Karl H. Johansson；山东省教育厅总督学王浩；青岛市市委常委、常务副市长耿涛；青岛科技大学党委副书记沈文青；青岛科技大学党委常委、副校长吕万翔；青岛市科协党组书记、主席王崇江；青岛市教育局总督学王元泳；中国自动化学会会士、副理事长，青岛科技大学副校长，上海交通大学教授李

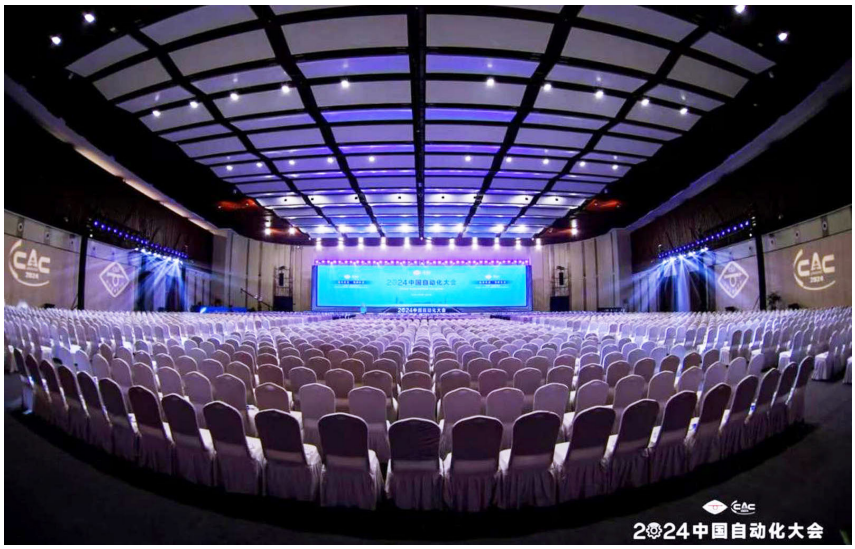


图 1

图 2　周东华教授主持

图 3　郑南宁院士致欢迎辞

少远；中国自动化学会会士、副理事长王成红；欧洲科学院院士，中国自动化学会会士、副理事长，华南理工大学教授陈俊龙；中国自动化学会会士、副理事长，山东科技大学教授周东华；中国自动化学会会士、副理事长，中国钢研科技集团有限公司原党委副书记张剑武；中国自动化学会会士、副理事长，中国科学院自动化研究所研究员侯增广；中国自动化学会会士、常务理事，山东大学教授张承慧；清华大学教授刘云浩；国防科技大学教授胡德文；龙芯中科技术股份有限公司董事长、中国科学院计算所研究员胡伟武；国家橡胶与轮胎工程技术研究中心副主任、软控股份有限公司董事长兼总裁官炳政等出席会议。

大会开幕式由周东华教授主持。中国工程院院士，中国自动化学会会士、理事长，西安交通大学教授郑南宁；山东省教育厅总督学王浩；青岛市市委常委、常务副市长耿涛；青岛科技大学党委常委、副校长吕万翔分别致辞。

郑南宁院士指出，人工智能正在使自动化技术以更大的规模和深度融入各行各业，渗透到我们生活的每一个角落。控制的本质是反馈，无论是在工程、经济和生物系统中，反馈都是实现有效控制的核心要素。通过结合人工智能、外部环境变化的表征与控制，可以实现一种面向任务的具身智能。具身智能是人工智能领域的一个重要方向，它强调智能体通过感知、认知和行动，与物理环境进行深度交互，其将极大增强自动化系统的柔性能力，使得自动化系统不再局限于预设的程序，而是能够根据环境的变化进行自主决策和行动，包括对人的意图和行为动机的理解。这也是未来工业软件——具身软件发展的新趋势。我们应把握自动化和人工智能带来的历史性机遇，共同应对这些挑战，用智慧和创新引领中国自动化技术的未来。

王浩总督学表示，山东省将深入发挥高等教育、职业教育大省优势，统筹教育、人才、科技一体发力，积极构建科技创新和产业发展需求牵引的学科专业调整机制和人才培养模式，持续强化科研平台建设，不断攻克重大科研成果，有效加速了传统产业升级、新兴产业壮大、未来产业布局。本次大会紧扣自动化领域前沿技术与行业全面深化改革，对推动我国自动化产业创新发展，加快形成适应高质量发展的新质生产力具有有重要意义。希望青岛科技大学能够以本次大会为契机，瞄准国际前沿，聚焦国家战略，积极深化与国内外顶尖专家学者、科研院所、头部企业的交流合作，共同推动自动化领域产、学、研深度融合，攻克更多卡脖子关键技术，为自动化产业的创

图 4　王浩总督学致辞

图 5　耿涛副市长致辞

图 6　吕万翔副校长致辞

新发展作出新的重大贡献。

耿涛副市长指出，青岛作为全国重要人工智能创新应用先导区，在智能家居、智能交通、智能医疗、智能机器人、智能制造和智能硬件等六个应用领域具有先发优势。青岛始终致力于完善人工智能产业生态和持续推动人工智能在产品创新、产业集聚、应用赋能等方面的能力提升，并积极拓展人工智能在生产生活、社会治理等方面的应用场景。希望能够以本次大会为契机，诚挚邀请各位专家学者深入了解青岛，为青岛的发展献计献策，将更多创新项目、业务和产品落地青岛，青岛将持续提供优质服务，以促进自动化产业快速发展，携手共创一个合作共赢的美好未来。

吕万翔副校长指出，在经济全球化、信息化和智能化加速推进的大背景下，大数据、人工智信息技术赋能的自动化技术已成为推动新质生产力发展的关键力量。自动化学科是青岛科技大学的传统优势学科，近年来，学校加快推动学科交叉融合，在深空、深海、低空飞行等未来产业和战略性新兴产业领域开展一系列前瞻性研究，并在智能制造、工业控制系统等领域取得了重大进展，为培育和发展新质生产力的新动能提供了有力支持。青岛科技大学将以本次大会为契机，全面提升学校在人才培养、学科建设、科学研究、社会服务、国际交流等方面的能力与水平，加快建设特色鲜明、走在前列的高水平大学，为推动学术进步、科技发展、产业繁荣，实现高水平科技自立自强贡献智慧和力量。

开幕式期间重磅发布了“胸怀祖国、自立自强、协同创新、追求卓越”的自动化人精神。这一精神的提出，不仅是对自动化领域精神内涵的深刻诠释，也将推动广大科技工作者把爱国之情、报国之志转化为抢占科技制高点、奋勇攀登自动化高峰的不竭动力，将激励每一位自动化人以胸怀祖国为魂，以自立自强为骨，以协同创新为翼，以追求卓越为梦，共同书写科技发展的辉煌篇章。

在主旨报告环节中，中国工程院院士，中国自动化学会会士、副理事长，中南大学教授桂卫华

图 7　自动化人精神发布仪式

为大家带来了题为“商用大飞机全电刹车控制技术的探索与思考”的报告。桂院士首先介绍了国产大型客机的重要战略地位，提出商用大飞机是一个国家工业、科技水平的综合实力的集中体现，C919 实现商飞标志着我国航空高端装备的快速发展。全电刹车作为多电飞机关键技术之一，是服务绿色航空、落实国家“双碳”战略的重要手段，但仍面临精细建模、精准控制、精确协同等方面挑战。桂院士提出，全电刹车技术的应用前景广阔，拥有自主知识产权的全电刹车系统可以减少对海外供应链的依赖，保障国产大飞机项目的独立性和稳定性。

中国工程院院士，鹏城实验室主任，北京大学信息与工程科学部主任、博雅讲席教授高文作了题为“中国算力网与鹏城脑海大模型”的报告。高院士指出，在智能时代，算力规模是国家竞争力的核心要素，算力将成为社会发展的新质生产力。算力优势不在于算力体量，而是更高的计算效率、广泛的新兴技术应用以及健全的基础设施支撑。算力网是算力基础设施发展的必然趋势，面向国家“东数西算”战略工程，鹏城实验室提出了“中国算力网”创新计划，通过超级直连光网络将全国区域内的大型绿色算力以超宽带低时延的体验输送给用户，实现算力与能源的绿色高质量协同发展。最后，高院士在总结中提出，中国已成为世界算力发展涨幅的领跑者，中国人工智能需要自己的大模型底座，以便承载以中文语料和中国历史的文明传承，脑海是其中的一个尝试。

瑞典皇家理工学院教授 Karl H. Johansson 带来了题为“Machine Learning Components in Cyber-physical Transport Systems”的报告。Johansson

图 8　桂卫华院士作报告

图 9　高文院士作线上报告

图 10　Karl H. Johansson 教授作报告

图 11　David Cebon 教授作报告

教授指出，随着传感、连接、计算和电气化技术的飞速发展，人类出行和货物运输的基础设施正在经历深刻的变革。在优化和提升交通系统韧性的研究中，控制技术对出行方式的广泛影响日益凸显。在此背景下，学习型信息物理融合系统作为一个新兴领域，正逐渐受到学术界的广泛关注。报告中，Johansson 教授详细阐述了如何利用物理信息机器学习技术，使联网车辆成为移动传感器和执行器，从而实现前所未有的大规模交通预测与控制。他通过具体实例，展示了这种技术在智能交通领域的广泛应用前景，如提高交通效率、减少拥堵和事故等。此外，Johansson 教授还深入探讨了自动驾驶车辆与人工驾驶车辆之间安全交互的复杂性，并强调了形式化推理方法与远程操作技术的结合应用。他认为，这些技术的融合将为未来智能交通系统的发展提供有力支撑。

英国皇家工程院院士、剑桥大学教授 David Cebon 带来了题为“Heavy Goods Vehicle Safety，Productivity and Environmental Impact”的报告。Cebon 教授指出，重型货车在现代城市化社会中扮演着至关重要的角色，它们负责运输食物等维持社会生活的重要物资，同时清除废物并提供工业运作所需的货物和材料。然而，重型货车具有复杂的动力学特性，包括多个耦合的车辆单元、高重心和非线性悬挂系统，导致它们行驶时可能产生大幅运动和高动态轮胎力，对道路和桥梁造成损害。报告中，Cebon 教授详细列举了重型货车面临的安全挑战，如翻车、折叠、气动制动导致的紧急制动性能不佳、操纵性差以及与弱势道路用户的复杂交互等。同时，他还指出重型货车的高碳排放量和非排气排放对人类健康和环境造成的负面影响。针对这些问题，Cebon 教授分享了他四十年来在重型货车动力学领域的研究成果，旨在提高重型货车的安全性、效率和环境友好性。他的研究涵盖了车辆动力学建模、安全性能优化、节能减排技术等多个方面，为重型货车的未来发展提供了宝贵的参考。

中国自动化学会会士、常务理事，山东大学教授，“新能源优化控制”国家级工程中心主任张承慧带来了题为“新能源系统控制技术发展”的报告。张承慧教授指出，新能源大规模开发利用已成为全球共识和我国国家战略，而信息控制技术直接决定能源转型成败。不同于传统能源，新能源需控制系统并阀，才能实现大规模高效利用。控制系统关乎新能源利用核心指标，直接决定系统容量、性能、发电量，应以控制简拓扑，以软件代硬件，以智能增性能。然而，新能源系统本质是一类高动态、非线性、强耦

图 12　张承慧教授作报告

图 13　于海斌院士作报告

图 14　段广仁院士作报告

合的复杂巨系统，面临高效率发电优化难、高可靠并网控制难、大容量化控制难等问题。张承慧教授在报告最后展望了元宇宙与新能源系统的深度融合前景，为新能源数智化控制提供了新范式。

中国工程院院士，中国自动化学会会士、副理事长，中国科学院沈阳自动化研究所研究员，机器人学国家重点实验室主任于海斌为大家带来了题为“人工智能驱动的智能制造新范式及前景展望”的报告。于院士表示，制造业是国家经济发展的支柱和动力引擎，在制造业由数字化深化迈向智能化转型的关键时期，面临创新设计能力弱、智能自动化水平低、产业供应链韧性差等技术挑战。而制造业智能化发展前景和空间巨大，应发挥海量制造数据的优势，实现知识和业务流程自动化和智能化，满足设计制造一体化、柔性制造、绿色制造等关键需求。最后，于院士展望了工业人工智能全面和体系化赋能先进制造业前景，并提出人工智能科技必将遵循“从数字走向实体”的发展规律，才能真正成为新质生产力的基石。

中国科学院院士，中国自动化学会会士，哈尔滨工业大学教授段广仁为大家带来了题为“全驱系统控制理论——诞生背景、发展现状与应用进展”的报告。段院士对全驱系统方法的诞生背景、研究现状和应用进展做了一个较为系统的概述。他提出，虽然今天控制工程界仍然很繁荣，但控制理论体系却早已陷入了困境。就在基础控制理论走向沉寂的时候，受物理上全驱系统的启发，我们提出了一个全新的全驱系统理论体系的基本框架，得到了国内外众多控制学者的认同与参与。目前，在全体全驱系统方法研究者的共同努力下，全驱系统理论走过了“从无到有、从小到大、由弱变强”的成长历程，一个全新的全驱系统理论体系已经形成，它在处理非线性、时变性、滞后特性、非光滑特性等一系列复杂问题上显示了极其强大的优越性，为控制理论发展开辟了一条崭新的途径。

清华大学教授刘云浩为大家带来了题为“物联网：从实现到现实”的报告。刘云浩教授在报告中提到，物联网是实现“感知智能”的重要载体，使人类可以通过网络化的设备与自然界对话。在人工智能的加持下，今天的物联网迈入了新阶段，更加注重其可赋能性，特别是在工业互联网蓬勃发展的背景下，物联网从服务消费端逐步延伸到服务生产端。通过工业互联网系统的赋能，企业可以逐步在数字世界中对生产过程进行预测和推演，实现“数字先行”。这使得数字孪生企业能够先于实体生产进行主动思考、推演和论证，从而更好地指导实际生产过程，提升生产效率和决

图 15　刘云浩教授作报告

图 16　胡伟武董事长作报告


图 17　官炳政董事长作报告

策水平。刘云浩教授还表示，这种转变标志着物联网正在从单纯的感知层面，向着赋能产业智能化转型的方向迈进。

龙芯中科技术股份有限公司董事长、中国科学院计算所研究员胡伟武为大家带来了题为“破解卡脖子问题，构建新发展格局”的报告。胡伟武董事长表示，通过长期积累和迭代，龙芯中科已将自主化优势转化为性能和生态优势，跻身世界领先行列。日前，新一代通用处理器龙芯 3A6000 正式发布，性能达到市场主流产品水平，标志着龙芯已走出一条基于成熟工艺、通过设计优化提升性能的道路。龙架构成为与 X86/ARM 并列的顶层开源生态系统。胡伟武董事长在总结中提出，我国信息产业新发展格局正在加速构建中，我们应将自主进行到底，不依赖国外技术授权和供应链。

国家橡胶与轮胎工程技术研究中心副主任、软控股份有限公司董事长兼总裁官炳政为大家带来了题为“轮胎行业智能制造与精益管控”的报告。官炳政董事长提出，轮胎作为安全产品，轮胎质量影响到交通出行安全，中国轮胎产量占到全球的 50% 以上，是典型的轮胎制造大国。轮胎制造过程为流程 + 离散的混合型制造模式，面临完备信息、机理经验、质量控制等难题，对高质量、高效能、高一致制造轮胎产品提出巨大的挑战。针对这些问题，官炳政董事长分享了轮胎智能制造的进阶之路与轮胎精益管控的实践探索等内容。他在总结中指出，中国橡胶轮胎正处于由跟跑、并跑到领跑，由大向强的发展过程，希望产业出题，学界答题，共同做好轮胎产业这篇文章，让中国轮胎服务好国家大战略、大工程，服务好每一个家庭。

此外，本届大会热点话题精彩纷呈，“脑机接口与脑机智能创新发展”“智慧环保前沿理论与技术”“ 新模式赋能的智慧诊疗”“教育大模型技术前沿发展与挑战”“大模型与网联智能”“复杂装备智能运维与高端控制”“边缘具身智能”等 46 个平行会议也陆续在大会期间举行，300 余名专家学者分享领域最新学术进展，共话产业新未来。

本次大会还特设“智向未来”智能科学成果展，龙芯中科、华龙迅达、度量科技、施普林格、元客视界、朝元时代、史河机器人、青瞳视觉、宇树科技等 10 余家企业聚焦创新成果和未来集群布局，展出了光学运动捕捉系统、四足机器人、双足机器人等“亮眼”新品，特色展区“青少年优秀科普作品展”，展示了中小学生的科普成果，为中国未来自动化及人工智能创新发展蓄积后备力量。

大会闭幕式由中国自动化学会会士、副理事长，青岛科技大

图 18　专题论坛

图 19　展览展示

学副校长，上海交通大学教授李少远主持，中国科学院院士，中国自动化学会会士、副理事长，中国空间技术研究院研究员杨孟飞致辞，对大会各项成果的进行了充分肯定和对未来提出了殷切期望。青岛科技大学副校长李少远作 2024 中国自动化大会总结，从筹备情况、现场盛况、精彩瞬间、论文接收情况等角度回顾了本场学术盛宴。

大会同期进行了“CAC 杯”交接仪式，2024 中国自动化大会组委会代表——青岛科技大学副校长李少远，将沉甸甸的“CAC 杯”郑重转交给 2025 中国自动化大会承办单位代表——哈尔滨工程大学智能科学与工程学院院长蔡成涛。2025 中国自动化大会将于 2025 年 11 月在哈尔滨召开，续写自动化大会新篇章。

余音未绝，未来可期。CAC2024 虽已落幕，但日新月异的自动化技术正以前所未有的融合创新、前沿突破迸发出全新活力，在新一轮科技革命和产业变革中迅猛发展，赋能千行百业、赋能未来世界。

中国自动化大会是中国自动化学会主办的大型综合性学术会议。自 2009 年创办以来，历经十五载，现已经成为我国自动控制、仪器仪表、航空航天、人工智能、网络安全、智能制造等领域最具影响力的学术盛会之一。大会充分彰显了自动化科技对全球进入智能时代的重要作用；并以前瞻视角和超前思维推进多领域交叉汇聚，建立各方向有机联动、相互促进的良性循环关系；有效贯通学科链、创新链与产业链，推动自动化领域创新全方位融入国家发展需求。

学会秘书处　供稿

图 20　杨孟飞院士致辞

图 21　李少远副校长作总结

AI for Engineering · 工业软件赋能新型工业化——2024 国家工业软件大会在上海盛大开幕

2024 年 11 月 9 日，由中国自动化学会主办，中国系统工程学会等 8 家全国学会联合协办，上海市松江区人民政府、华东理工大学、国家流程制造智能调控技术创新中心等 30 家单位联合承办的 2024 国家工业软件大会在上海开幕。大会以“AI for Engineering · 工业软件赋能新型工业化”为主题，20 余位院士、千余位工业软件领域专家学者和行业精英齐聚一堂，共同探讨中国工业智能领域的战略布局、技术创新应用和产业融合发展等前沿热点，共谋工业软件发展新篇章。

参加大会开幕式的嘉宾有：中国工程院院士，中国自动化学会特聘顾问、会士，浙江大学教授孙优贤；中国科学院院士，中国科学院过程工程研究所研究员李静海；中国工程院院士，中国自动化学会理事长、会士，西安交通大学教授郑南宁；中国工程院院士，中国自动化学会特聘顾问、会士，东北大学教授柴天佑；中国工程院院士，中国自动化学会特聘顾问、会士，中国科学院沈阳自动化研究所研究员王天然；中国工程院院士、中国石油化工股份有限公司原总裁王基铭；中国工程院院士，中国自动化学会副理事长、会士，中南大学教授桂卫华；中国工程院院士，国家智能制造专家委员会副主任，中国自动化学会监事、会士，华东理工大学教授钱锋；中国工程院院士，中国自动化学会常务理事、会士，湖南大学教授王耀南；中国科学院院士、东华大学教授朱美芳；中国工程院院士，中国自动化学会副理事长、会士，中国科学院沈阳自动化研究所研究员于海斌；中国科学院院士、哈尔滨工业大学教授李惠；中国工程院院士、华中科技大学教授陈学东；加拿大工程院院士、香港中文大学（深圳）副校长朱世平；挪威工程院院士、挪威皇家科学院院士、挪威科技大学教授陈德；加拿大工程院院士、香港中文大学（深圳）教授崔曙光；中国科学技术协会专职副主席、书记处书记孟庆海；上海市人民政府副市长陈宇剑；国家工业和信息化部高新技术司副司长陈彦丞；国家自然科学基金委员会交叉科学部原副主任潘庆；上海市人民政府副秘书长庄木弟；上海市经济和信息化委员会书记程鹏；上海市科学技术委员会副主任屈炜；上海市松江区委副书记、区长王靖；华东理工大学校长轩福贞；上海市松江区人民政府副区长陈容。

大会开幕式由钱锋院士主持。郑南宁院士，陈宇剑副市长，陈彦丞副司长，轩福贞校长，王靖区长分别致辞。

郑南宁在致辞中表示，随着人工智能技术的飞速发展，工业

图 1　国家智能制造专家委员会副主任钱锋院士主持

图 2　中国自动化学会理事长郑南宁致辞

图 3　上海市人民政府副市长陈宇剑致辞

软件的技术深度和应用广度都在不断拓展，为实现生产效率提升、资源优化配置、绿色可持续发展提供了强有力的支撑。工业软件的进步是科技与工程智慧的结晶，我们应携手并肩，共同推进工业软件的发展，不断强化我国在全球工业领域的技术竞争力。

陈宇剑在致辞中表示，上海依托自身在科技创新、产业基础、市场资源等方面的独特优势，积极推动工业软件与制造业的深度融合，努力在关键技术攻关、应用场景拓展、产业生态构建等方面实现新突破、展现新作为，为推动我国工业软件产业迈向更高水平、实现制造业强国目标贡献智慧和力量。

陈彦丞在致辞中表示，工业软件在我国新型工业化进程中发挥着不可或缺的作用，在人工智能等新技术加持下，正经历深刻变革，向全方位、全过程、全链条的高阶智能化转型，支撑制造业高端化、智能化、绿色化发展。我们要推动科技创新与产业创新深度融合，提升产业科技创新能力，形成以产业为导向、以技术创新为引领、以企业为主体的工业软件协同创新格局，全面提升我国工业软件自主创新能力。

轩福贞在致辞中表示，近年来，华东理工大学以“智能化、数字化”为重要抓手，积极推动国家战略需求牵引的学科调整机制和人才培养模式，加速推进“双一流”建设。希望以此次会议的召开为契机，持续推进国有工业软件自主可控健康发展，为加快发展新质生产力，推进新型工业化作出更大贡献！

王靖在致辞中表示，松江制

图 4　国家工业和信息化部高新技术司副司长陈彦丞致辞

图 5　华东理工大学校长轩福贞致辞

图 6　上海市松江区委副书记、区长王靖致辞

造业基础扎实、创新要素集聚、高能级市场主体集聚，拥有工业软件企业发展亟需的丰富应用场景、算力资源支撑和产业基础。诚挚邀请各位专家学者和企业家关注松江软件信息产业发展的新机遇，链接工业软件产业的高端资源来松集聚，共同推动上海工业软件产业向深向新发展。

大会期间，孙优贤院士、郑南宁院士、李静海院士、柴天佑院士、谭建荣院士、桂卫华院士、杨华勇院士、丁烈云院士、王耀南院士、陈德院士分别作主旨报告，深入剖析工业软件发展趋势，为推动智能制造技术创新与产业升级带来新的启示。

大会聚焦中国工业智能领域发展，设置研发设计类、生产制造类、经营管理类、AI 赋能类、数字孪生与控制装备类、土木工程类、EDA 软件应用类等 15 场专题论坛，众多专家学者分享领域最新学术进展，共同探索工业软件领域新未来。大会同期举行“流程工业软件发布”仪式、“工业智能操作系统”白皮书发布仪式、工业操作系统联盟聘任仪式和 AI 赋能千行百业暨企业产业大脑需求发布仪式，旨在引领工业领域的智能化变革，为新型工业化注入新动能。

此外，大会特设工业软件成果展，30 余家企业聚焦创新成果和未来集群布局，集中展示中国在工业软件领域的最新模型、先进算法和软件成果。

本届大会由中国自动化学会主办，中国系统工程学会等 8 家全国学会联合协办，上海市松江区人民政府、华东理工大学、国家流程制造智能调控技术创新中心、国家智能设计与数控技术创新中心、国家数字建造技术创新中心、国家集成电路设计自动化技术创新中心、全国 / 国家重点实验室以及行业龙头企业中国石油化工集团有限公司、国家能源投资集团有限责任公司、中国机械工业集团有限公司等 30 家单位联合承办。

学会秘书处　供稿

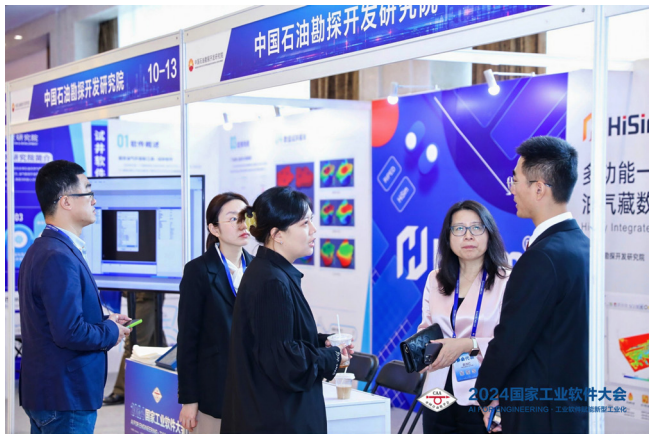


图 7　展览展示

胸怀祖国、自立自强、协同创新、追求卓越
——自动化人精神发布仪式在青岛举行

2024 年 11 月 2 日，2024 中国自动化大会在青岛世界博览城国际会议中心盛大开幕，会上隆重举行了“自动化人精神”发布仪式，以“胸怀祖国、自立自强、协同创新、追求卓越”为核心，深刻诠释了自动化人精神的丰富内涵。这一仪式不仅象征着自动化领域精神的传承，更预示着未来发展的新篇章。

中国工程院院士，中国自动化学会会士、理事长，西安交通大学教授郑南宁；中国工程院院士，中国自动化学会会士、特聘顾问，浙江大学教授孙优贤；中国工程院院士，中国自动化学会会士、特聘顾问，中国科学院沈阳自动化所研究员王天然；中国工程院院士，中国自动化学会会士、副理事长，中南大学教授桂卫华；中国工程院院士，中国自动化学会会士、监事，华东理工大学教授钱锋；中国科学院院士、中国自动化学会会士、哈尔滨工业大学教授段广仁；中国工程院院士，中国自动化学会会士、副理事长，中国科学院沈阳自动化所研究员于海斌；英国皇家工程院院士、剑桥大学教授 David Cebon；瑞典皇家理工学院教授 Karl H. Johansson 出席发布仪式，共同按下启动按钮，宣布“自动化人精神”正式发布。

发布仪式现场

六十多年前，我国的科技基础极为薄弱亟待改善。为积极响应国家“向现代科学进军”的号召，推动自动化等新兴科学的发展，自动化界老一辈科学家冲破重围，回到祖国，投身到新中国的建设当中。人总是要有精神的。60 多年来，一代代自动化人不忘初心、努力奋斗，为我国自动化事业的发展做出了不灭的贡献，也培育铸就了“胸怀祖国、自立自强、协同创新、追求卓越”的自动化人精神。这是我们自动化人立命之本，是为全面建设社会主义现代化国家、全面推进中华民族伟大复兴的精神力量。

“胸怀祖国”，是我们自动化人的坚定信仰。一代代自动化人以中华民族的伟大复兴为己任，书写着对国家的热爱与奉献。这份情怀激励着我们自动化人不断追求科技创新，推动国家发展，实现民族复兴。

“自立自强”，是我们自动化人应对挑战的坚强毅力。唯有自立自强，方能攀登科技之巅，立于不败之地。它犹如夜空中最亮的星，照亮自动化人前行的道路。

“协同创新”，是我们自动化人应对挑战的有效途径。在科技报国的道路上，自动化人携手并进，碰撞思想的火花，激发创新的灵感，共同绘制科技发展的宏伟蓝图，滋润着科技创新的土壤，结出更加丰硕的智慧果实。

“追求卓越”，是我们自动化人永恒的目标。自动化人不断挑战自我，超越极限，坚持追求着科技与创新的极致。这份对卓越的执着与追求，使自动化人在科技创新的征途上破浪前行，不断攀登新的高峰。

自动化人精神是一首平凡筑就伟大的赞歌。它激励着每一位自动化人，以胸怀祖国为魂，以自立自强为骨，以协同创新为翼，以追求卓越为梦，共同书写科技发展的辉煌篇章。

学会秘书处　供稿

第六届中国认知计算与混合智能学术大会圆满落幕

2024 年 11 月 1 日，第六届中国认知计算与混合智能学术大会在山东青岛成功落下帷幕。本次大会由国家自然科学基金委员会信息科学部、中国自动化学会和中国认知科学学会联合主办，吸引了众多国内外知名专家学者参加，共同探讨了人工智能领域的前沿技术和未来发展方向。

本次会议大会主席为中国科学院院士、中国认知科学学会理事长、中科院生物物理研究所研究员陈霖，中国工程院院士、中国自动化学会理事长、西安交通大学郑南宁教授。郑南宁院士在开幕式致辞中指出，本届大会旨在促进人工智能领域的前沿研究与发展，希望各位专家、学者能在本次会议上相互启发，共同推动我国新一代人工智能的发展。

CCHI 2024 主论坛由中国自动化学会副秘书长、西安交通大学辛景民教授主持，邀请了国家电网有限公司郭剑波院士、武汉大学刘胜院士、国防科技大学胡德文教授分别作“人在回路的电网调控混合增强智能技术”“低成本高可靠车规级传感器对人形机器人的启示”“人与机器人智能融合技术”的报告。

CCHI 2024 会议分论坛“受神经科学启发的计算架构与器件”由中国自动化学会副秘书长、武汉大学张俊教授主持，邀请了中国南方电网有限责任公司郑外生教授级高工、上海交通大学王贺升教授分别作“基于神经网络的电力系统‘驭电’智能仿真器”和“基于视觉的机器人定位导航与控制”的报告。

CCHI 2024 会议分论坛“认知计算的基础理论”由香港科技大学（广州）郑心湖助理教授主持，邀请了西安交通大学兰旭光教授、清华大学李勇教授分别作“具身智能的挑战：世界模型构建与因果推理”和“城市认知与决策智能”的报告。

本次会议通过主题报告、分论坛讨论及互动环节为参会嘉宾和听众提供了一个高水平的认知计算和混合智能研究领域知识共享平台，报道了该领域最新的学术成果，为推动认知计算和混合智能领域内学术研究和技术发展作出了贡献。

学会秘书处　供稿

图 2　中国自动化学会理事长郑南宁院士致辞

图 3　主论坛主持人辛景民教授

《2024 年控制科学与工程学科发展报告》座谈会在青岛召开

为全面跟踪国内外控制科学与工程学科前沿发展情况，科学评价近年来控制科学与工程学科的新进展、新成果等，2024 年 11 月 1 日，中国自动化学会在青岛成功召开了《2024 年控制科学与工程学科发展报告》座谈会。中国自动化学会副理事长王成红，中国自动化学会副理事长、中国科学院自动化研究所研究员侯增广，中国自动化学会秘书长张楠，中国自动化学会副秘书长王坛，来自高校科研院所和青托代表共 60 余人参加会议。

中国自动化学会副理事长、中国科学院自动化研究所研究员侯增广在致辞中指出，学科发展报告的编制工作要充分发挥学会的平台优势和专业力量，广泛汇聚行业智慧，深入调研分析，确保报告内容全面、准确、前瞻，既能反映当前学科发展的最新成就，也能揭示存在的问题与挑战，更能提出具有针对性和可操作性的对策建议，为行业发展和学术研究提供有力支撑。随后，学会秘书处介绍了《2024 年控制科学与工程学科发展报告》项目要求和实施方案，详细阐述了往年的经验做法和今年的计划安排，明确任务分工和工作进度。

会上，专家学者分组围绕报告大纲等进行了深入交流和研讨，就学科分类和撰写规划发表了看法和见解，还就如何进一步撰写报告内容、提高报告质量等问题并提出了许多宝贵的意见和建议。

图 1　座谈会现场

王成红副理事长特别强调，今年的报告编写工作，应充分借鉴以往的经验，结合过去的不足之处，力求在内容、结构、深度等方面实现全面提升。报告不仅要全面展现国内在控制科学与工程领域取得的最新研究成果，还要加强与国际先进水平的对比分析，明确学科目前存在的优势与短板，为未来的学科发展指明方向。

侯增广副理事长在总结中提出，人工智能与自动化之间存在着不可分割的紧密联系，在编制学科发展报告时，必须紧密融合人工智能元素，还应高度关注涉及交叉领域及交叉学科的新方向，丰富学科内涵。

此次《2024 年控制科学与工程学科发展报告》座谈会的成功召开，不仅为项目的顺利完成指明了方向，也为广大专家学者提供了一个学术交流重要平台。相信在专家学者们的共同努力下，《2024 年控制科学与工程学科发展报告》能够成为引领学科发展、服务国家战略的精品力作。

学会秘书处　供稿

第十四届“中国智能车未来挑战赛”在江苏常熟成功举办

为落实科教兴国和创新驱动发展战略，不断塑造人工智能发展新动能、新优势，第十四届“中国智能车未来挑战赛”于2024年11月22日至23日在江苏省常熟市成功举行。本届比赛由国家自然科学基金委员会信息科学部、中国自动化学会主办，常熟市人民政府承办，中国（常熟）智能车综合技术研发与测试中心、人机混合增强智能全国重点实验室共同协办。

23日上午，大赛组委会主席、中国工程院院士、中国自动化学会理事长、西安交通大学郑南宁教授，常熟市委常委、常熟高新区党工委书记王建国，国家自然科学基金委员会科学传播与成果转化中心彭杰主任，大赛总裁判长、中国自动化学会监事长王飞跃研究员，大赛仲裁委员会主任、中国自动化学会副理事长王成红研究员，常熟国家大学科技园党工委副书记、管委会主任王奇峰以及常熟国家大学科技园党工委委员、管委会副主任潘伟等相关领导以及参赛队200余人参加大赛开幕式。

大赛开幕式由国家自然科学基金委员会信息科学部三处处长赵瑞珍研究员主持。常熟市委常委、常熟高新区党工委书记王建国、国家自然科学基金委员会科学传播与成果转化中心彭杰主任分别代表常熟市政府、国家自然科学基金委员会信息科学部致辞。

23日上午9时许，大赛组委会主席郑南宁院士隆重宣布第十四届“中国智能车未来挑战赛”开始。

本届比赛设置“面向真实交通环境的全自动驾驶挑战赛”和

图1　开幕式出席嘉宾

图 2　参赛车辆

“复杂交通环境下通 – 感算法线上比赛”。“面向真实交通环境的全自动驾驶挑战赛”针对当前开放动态环境下全自动驾驶技术面临的痛点问题，重点考察自然交通场景中恶劣天气、意外抛物、非合作目标、道路施工等极端情况下智能车的全自动驾驶能力，并考察自动驾驶智能车的人机自然交互能力和复杂交通场景的周密感知和智能决策能力。同时，模拟“最后一公里”内自动驾驶汽车从公开道路进入未建图封闭住宅 / 商业区停车场的常见生活案例，鼓励通过大语言模型、计算机视觉大模型和通用智能体等人工智能创新克服现有技术局限，实现精细化、个性化的无缝自动驾驶服务。“复杂交通环境通 – 感算法线上比赛”包含车联网场景信道大尺度衰落预测、高效波束预测方案设计、协同定位追踪和自主代客泊车智能决策等算法赛题。

本届大赛有 16 支车队参加“面向真实交通环境的全自动驾驶挑战赛”经过为期一天的激烈角逐，西安交通大学“Pioneer 先锋号”以总分排名第一的成绩获得一等奖；香港科技大学（广州）悟空号首次参加大赛即以总分第二名的成绩获得二等奖的好成绩。北京理工大学精工智能号以总分第三名的成绩也获得二等奖。本届“真实交通环境挑战赛”比赛设一等奖 1 名、二等奖 2 名和三等奖 3 名。

“复杂交通环境通 – 感算法线上比赛”有 60 个团队报名参赛。来自山东大学、西北工业大学、北京邮电大学、浙江大学等 4 支参赛队分别获得不同算法赛题的第一名。

作为全球历史最悠久的自动驾驶赛事，自 2009 年首届比赛以来，十五年来，“中国智能车未来挑战赛”培养和输送了大量自动驾驶领域专家和人才，成为自动驾驶领域人才培育的摇篮。大赛涌现出许多标志性成果，孵化了一批创新企业，对我国自动驾驶产业形成了有力支撑，实现中国自动驾驶领域多项“从 0 到 1”的突破，引领和推动了中国自动驾驶的跨越式发展。

学会秘书处　供稿

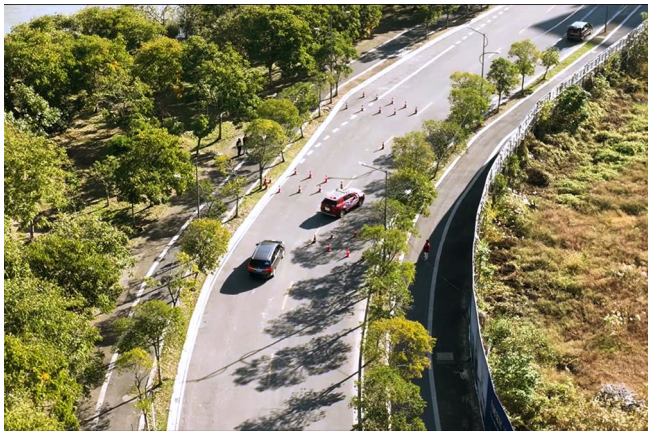


图 3　参赛车辆比赛中

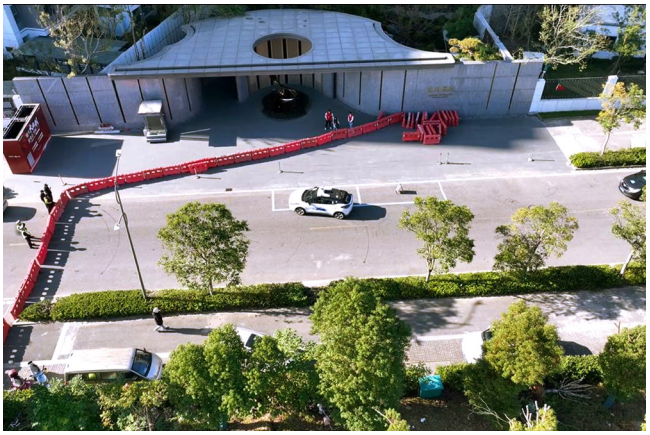


图 4　参赛车辆比赛中

中国自动化学会会士荣誉称号授予仪式在青岛隆重举行

中国自动化学会会士是学会会员在学会的最高荣誉，旨在表彰和奖励对自动化、信息与智能科学技术发展做出卓越贡献、在自动化及相关领域取得杰出学术创新成果或为学会服务做出突出贡献的学会会员。2024 年 11 月 2 日，中国自动化学会隆重举行了 2024 年中国自动化学会会士荣誉称号授予仪式，中国工程院院士，中国自动化学会会士、理事长，西安交通大学教授郑南宁与中国工程院院士，中国自动化学会会士、特聘顾问，浙江大学教授孙优贤共同为获奖者颁发奖牌。2024 年共有十位卓越的科技工作者获此荣誉。

会士名单

按姓氏拼音排序

蔡伯根
北京交通大学教授

陈　谋
南京航空航天大学教授

程　翔
北京大学教授

胡昌华
中国自动化学会常务理事，火箭军工程大学教授

胡庆雷
北京航空航天大学教授

刘妹琴
西安交通大学教授

夏元清
北京理工大学教授

辛景民
中国自动化学会副秘书长，西安交通大学教授

钟麦英
山东科技大学教授

周　杰
中国自动化学会副会秘书长，清华大学教授

持续深化党的纪律教育

加强纪律建设是全面从严治党的治本之策。通过对党员开展纪律教育提升党性、净化党风、凝心聚力，是党百余年来取得伟大成就的一条重要经验。新征程上，我们要把纪律建设摆在更加突出的位置，总结运用党纪学习教育的好经验好做法，常态化推进学纪知纪明纪守纪，使广大党员干部在遵规守纪的前提下勤奋工作、放手干事、锐意进取、积极作为。

建立经常性和集中性相结合的纪律教育机制。党的十八大以来，我们党先后开展一系列主题教育和党内集中教育，特别是今年在全党开展的党纪学习教育，聚焦解决一些党员干部对党章党规党纪不上心、不了解、不掌握等问题，强化党员干部学纪知纪明纪守纪的自觉性。《中共中央关于进一步全面深化改革、推进中国式现代化的决定》指出："建立经常性和集中性相结合的纪律教育机制，深化运用监督执纪'四种形态'，综合发挥党的纪律教育约束、保障激励作用。"要清醒认识到，党纪学习教育是一个循序渐进的长期过程，是一项基础性、经常性工作，不可能毕其功于一役。党员干部的纪律意识、规矩意识和遵规守纪的自觉性，必须通过持续的思想引导、严格的党性锻炼、深入的学习领悟才能逐步达成。要把集中性教育的经验和成果推广并运用到党组织经常性教育工作中，并针对集中性教育的遗留问题和薄弱环节进一步整改，采取党委理论学习中心组学习、上纪律党课、学典型通报、看警示教育片等多种方式，把纪律教育融入党员干部日常教育管理监督，推进党纪学习教育常态化长效化。

实现自律与他律的良性互动。纪律的根本在于执行与遵守，纪律教育的目的是让党员干部把纪律规矩铭记于心，知道能干什么、不能干什么。如果只靠外力强制确保党员干部不触犯纪律规矩，以惩戒发挥刚性约束作用，监督执纪的成本就会很高。习近平总书记强调："一个人的清正廉明，从根本上讲不能完全靠外部约束，而要靠自觉自律。自觉自律是人向上向善的内在动力。"我们党历来强调要培养"自觉的纪律"，认识到"党的纪律是带着强制性的；但同时，它又必须是建立在党员与干部的自觉性上面"。要把他律要求转化为自律行为，就必须坚持用理想信念、初心使命、党性原则筑牢遵规守纪的思想根基，引导党员干部正确处理个人与集体、民主与集中、纪律与自由等的关系，时刻以党纪国法为镜，处处与纪律规矩对标，注重平时点滴养成，真正把外在的纪律规矩转化为内在的政治自觉、思想自觉、行动自觉。

坚持党性党风党纪一起抓。"党性、党风、党纪是有机整体，党性是根本，党风是表现，党纪是保障。"党性党风党纪的内在一致性，决定了必须以系统性思维、全周期管理理念将三者一起抓才能见效。一方面，一起抓不是割裂开来分别抓。从目标任务来看，党性党风党纪对应思想建设、作风建设、纪律建设，统一于党的建设中，三者相互联系、相互作用、相互促进，在贯彻落实新时代党的建设总要求中形成合力。纪律教育不仅是纪律建设的手段，更是思想建设和作风建设的重要载体，如果将三者割裂开来，就会陷入"头痛医头、脚

痛医脚”的局面。另一方面，一起抓不是不分主次一把抓。党性是根本，说到底是立场方向问题，是实现三者有机贯通、协同推进、同向发力的“牛鼻子”；党风是表现，党风的好坏能够反映出一个地区、一段时间内的政治生态，能够暴露不足和问题，使得强党性更有针对性、严党纪更好对症下药；党纪是保障，整肃党风必须让“用制度治党管权治吏”成为全党共识，锤炼党性必须让党的纪律内化为日用而不觉的言行准则。全面从严治党越是往纵深推进，越需要坚持党性党风党纪一起抓，深刻把握不正之风和腐败问题互为表里、同根同源的规律性认识，通过锤炼坚强党性固本培元、弘扬优良党风激浊扬清、严明党章党规党纪拒腐防变，在提高觉悟、涵养风气、惩治震慑上一体发力。

把开展纪律教育同推动高质量发展结合起来。习近平总书记指出：“要巩固深化党纪学习教育成果，坚持融入日常、抓在经常，把党纪学习教育成果持续转化为推动高质量发展的强大动力。”巩固深化党纪学习教育成果，关键是突出一个“实”字，落于一个“干”字，同解决问题、改进工作、推动发展结合起来，为实现新时代新征程党的中心任务提供坚强纪律保障。要进一步强化党员干部遵守政治纪律和政治规矩的自觉性，切实把思想和行动统一到党中央决策部署上来，做到党中央提倡的坚决响应，党中央决定的坚决执行，党中央禁止的坚决不做；进一步树牢党员干部造福人民的政绩观，把纪律教育同持续整治群众身边不正之风和腐败问题等重点工作贯通起来，以典型案例为鲜活教材，用身边事教育身边人，引导党员干部加强自我约束、提高免疫能力，算好经济账、名誉账、亲情账、自由账，力戒“不会查到我”的侥幸、“温水煮青蛙”的麻痹、“就我说了算”的任性，做到清清白白为官、干干净净做事、老老实实做人；进一步培塑党员干部求真务实的工作作风，把开展纪律教育同整治形式主义为基层减负结合起来，紧盯形式主义老问题和新表现，用党的纪律规范履职行为、激励担当作为，切实增强党员干部忠诚干净担当的主动性和自觉性，营造干事创业良好氛围。

来源：中国共产党新闻网

喜报 XI BAO

热烈祝贺中国自动化学会多位会员、理事入选科睿唯安 2024 年度“全球高被引科学家”名单

2024 年 11 月 19 日，科睿唯安发布 2024 年度“全球高被引科学家”名单，来自全球 59 个国家和地区 1200 多个机构的 6636 名科学家入选。

热烈祝贺中国自动化学会66位专家入选科睿唯安2024年度“全球高被引科学家”名单！详情请查看：https://mp.weixin.qq.com/s/NwPHLRfSVxb4hYq90YWSfw

中国自动化学会一届十九次党委工作会议召开

2024 年 11 月 2 日，中国自动化学会一届十九次党委工作会议在青岛召开。会议由中国自动化学会党委书记、理事长郑南宁院士主持。

会议期间，各位委员围绕增设学会会员中心、杨嘉墀科技奖改革方案、IAPR 理事会代表候选人等议题进行审议。会上，委员们积极发言，对中国自动化学会的各项工作提出了宝贵的意见和建议。

会议最后，郑南宁理事长在总结发言中指出，中国自动化学会作为推动我国自动化领域发展的重要力量，应当紧扣国家战略需求和科技发展趋势，发挥其在科技创新和产业创新融合发展中的关键作用，搭建学术交流和技术合作平台，促进科技创新与成果转化，强化服务科技工作者的使命，为实现科技强国的目标贡献力量。

学会办事机构党支部　供稿

通知

关于推荐中国自动化学会科普教育基地的通知

各推荐单位：

为贯彻《中华人民共和国科学技术普及法》，加强自动化领域科普教育基地的建设，鼓励社会力量参与自动化科普工作，推动自动化、信息与智能科学知识普及，依据中国科协《全国科普教育基地认定办法（试行）》和《中国自动化学会章程》，中国自动化学会现面向全国（暂不包括港、澳、台地区）开展“中国自动化学会科普教育基地”推荐工作，详情请查看：https://www.caa.org.cn/article/192/5018.html

人民日报整版关注人工智能安全

确保人工智能安全、可靠、可控，有利于人类文明进步，是人工智能发展必须解决的重要课题。党的二十届三中全会《决定》作出“建立人工智能安全监管制度”“完善生成式人工智能发展和管理机制”等重要部署。如何加强人工智能治理，有效防范化解人工智能发展带来的各类安全风险，不断提升人工智能安全监管的制度化、法治化水平？本期学术版围绕这些问题进行探讨。

提升生成式人工智能三大风险治理能力

习近平总书记指出：“人工智能是新一轮科技革命和产业变革的重要驱动力量，将对全球经济社会发展和人类文明进步产生深远影响。”生成式人工智能是指基于算法、模型、规则生成文本、图片、声音、视频、代码等内容的技术。在海量数据与强大算力支撑下，听得懂、说得出、能互动的生成式人工智能快速迭代升级，呈现出良好互动性、高度通用性、智能生成性等特征，并正与各行各业形成更加刚性、高频、泛在、深度的联结，也导致其潜在风险更多更真实。党的二十届三中全会《决定》科学把握人工智能发展规律和特点，提出“建立人工智能安全监管制度”“完善生成式人工智能发展和管理机制”，体现了更好统筹发展和安全的客观需要，为推动人工智能领域的技术进步、产业发展与安全保障指明前进方向。

生成式人工智能在技术运行上可分为三个阶段，即前置性学习训练及人工标注辅助算法升级的准备阶段，输入数据进行算法处理得出生成物的运算阶段，生成物进入社会加以运用的生成阶段。我们要深入分析生成式人工智能的运行机理，把握各阶段安全风险形成与发展的特征，运用法治手段加强系统性治理，确保生成式人工智能所蕴含的巨大力量始终在法治轨道上发挥作用。

在生成式人工智能的准备阶段，数据安全风险易发多发、较为突出。生成式人工智能通过数据训练、数据处理分析来提炼信息、预测趋势。这就必须对数据进行适当分类，确立不同类型数据的利用模式和保护方式，以妥善应对相关数据安全风险，避免数据违规利用或者不当泄露，产生侵权方面的纠纷。比如，在政务处理流程中形成的政务数据是数字政府的核心要素。生成式人工智能为了得出相对准确的结论，不可避免地要收集分析政务数据。应当明确生成式人工智能获取和利用政务数据的法律规则，既满足利用政务数据服务社会的需求，有力支持人工智能政务服务大模型开发、训练和应用，提高公共服务和社会治理智能化水平；又规范其加工方式，避免利用政务数据得出的成果侵害个人权益、破坏社会公共秩序。对于个人数据而言，生成式人工智能通过组合分析挖掘其潜在价值，其对个人数据的收集利用及其成果可能对公民权利造成侵害。实践中，生成式人工智能倾向于过度收集个人数据以提升结论准确性，比如，通过分析医疗健康数据来挖掘个人行踪、预测个人生活轨迹。为此，必须坚持依法收集，按照技术所需的最小范围收集个人数据，设置合理的数据处理深度，避免过度挖掘潜在信息。综上，应将分类分级的数据安全监管要求嵌入生成式人工智能的准备阶

段，避免数据安全风险演化为具体的法益损害后果。

在生成式人工智能的运算阶段，内生于人工智能大模型的算法偏见风险值得警惕。生成式人工智能对于数据的分析和处理主要通过算法模型。不同于传统算法模型，生成式人工智能在进行机器学习的同时，还会以大量的人工标注来校正机器学习的结论，推动人工智能进化。但“机器学习＋人工标注”作为算法技术内核，也会使人类的意志与偏好所产生的影响比单纯的机器学习更大。个人偏好的影响叠加在算法模型本身的偏见之上，将导致算法偏见的负面效应倍增，算法偏见的产生更加难以追溯和预防。防范化解算法偏见风险，应根据算法偏见的产生原理与产生场域进行针对性治理。要将法律规范的要求深度嵌入生成式人工智能的算法模型之中，推动技术向善，消除算法偏见，确保合理利用生成式人工智能算法并分配算力资源。基于技管结合理念，加强对算法的全周期安全监管，将法律规范的要求落实到生成式人工智能运行的全流程之中。在设置算法之初就要遵循相关法律规则与技术标准，落实“机器学习＋人工标注”的规范要求，审查存在风险的算法模块，更好发现生成式人工智能算法模型中的技术风险；当发现先天性算法偏见时，依据法律要求从生成式人工智能的算法内部进行纠正，确保修改后的算法能正常运行；事后出现问题时，对人工智能算法进行溯源治理，实现精准归责并加以纠正，推动完善生成式人工智能的算法监管标准，填补事前预防审查的不足，以技术手段与法律手段并行做到发展与管理并重。

在生成式人工智能的生成阶段，存在着与生成物相关的知识产权风险、生成物滥用风险等多种风险。由于生成式人工智能的智能化程度很高，可以实现内容自动化编纂、智能化润色加工、多模态转换以及创造性生成，直接改变了内容的生产方式与供给模式，相较于以往的人工智能系统产生了颠覆性变化，由此引发了生成式人工智能的生成物知识产权归属、知识产权保护等问题。有的人认为生成式人工智能生成物是数据算法的结论，其本质上是计算与模仿，而非智力劳动，无法成为知识产权的客体。反对者则认为生成式人工智能模拟人脑神经网络的构造来获取与输出数据，通过卷积神经网络控制自身的设计与制造，其具有独创性与创新性的生成物应当受知识产权法保护。同时，生成式人工智能还增加了知识产权纠纷风险和保护难度，一些生成物可能含有侵犯他人知识产权的内容，或者经过加工等手段被包装成个人拥有完全知识产权的原创作品，引发相关知识产权争议。为及时化解相关问题，应对生成式人工智能的技术模式、技术原理按照知识产权法的标准开展实质分析，如果技术上需要人类意志介入，使生成物能够产生独创性与创新性，应赋予知识产权并明确其归属，强化生成式人工智能领域知识产权的系统性保护；同时要合理确定对生成物知识产权保护的范围，避免保护范围无限扩张，妨碍生成式人工智能的推广运用和技术发展。还要加强对生成物滥用风险的治理。比如，要求作品清楚标识生成式人工智能在作者创作中发挥作用的情况，加强对可能涉及违法犯罪的深度伪造、AI 换脸等生成物的精准化、常态化监管，等等。

生成式人工智能在社会应用中产生的扩散影响还有很多，除了上述风险还有很多其他类型的风险，比如加剧信息不对称、扩大数字鸿沟、损害数字弱势群体的利益等。要根据实际情况作出应对，尽量降低新技术给社会发展带来的不良冲击。

守护好人工智能时代的隐私安全

习近平总书记强调：“坚持以人为本、智能向善”。当前，人工智能技术日新月异，既深刻影响改变着人们的生产生活方式、加

速了经济社会发展进程，也对法律规范、道德伦理、公共治理等造成冲击。其中，对隐私权、个人信息安全等的威胁是值得关注的重要问题。党的二十届三中全会《决定》对“建立人工智能安全监管制度”作出重要部署，保护隐私权和个人信息安全是人工智能安全监管的题中应有之义。必须加强人工智能时代的隐私权保护，确保个人信息安全。

人工智能时代隐私权面临严峻挑战。隐私是自然人的私人生活安宁和不愿为他人知晓的私密空间、私密活动、私密信息。民法典规定：“自然人享有隐私权。任何组织或者个人不得以刺探、侵扰、泄露、公开等方式侵害他人的隐私权。”隐私权作为人格权的核心要素，是构筑人格尊严的重要基础。不被公开、不被知晓是隐私权的核心诉求。当前，人工智能以悄无声息的方式介入人们生产生活的各领域各方面各环节，产生智能医疗、智能交通、智能推荐等众多应用场景，技术本身存在的某些缺陷和规则的不完善，不可避免带来侵害隐私权的问题。比如，非法收集和使用个人信息，利用分析这些个人信息频繁推送所谓“个性化”的“精准广告”，泄露个人信息给第三方，导致私人生活频频受到垃圾信息侵扰；利用个人信息进行“大数据杀熟”，实现“一客一价”的精准价格歧视，令公民遭受财产损失；已脱敏个人信息被重新识别，因保护措施不当导致数据泄露，非法买卖个人信息现象屡见不鲜，侵害个人信息安全；借助个人信息进行深度伪造，通过声音仿真、AI 换脸等手段，实施诈骗等违法犯罪行为；等等。这说明，侵害隐私权不仅侵犯了公民的人格尊严，也会造成其他严重社会后果。

去私密化技术特征加剧个人信息安全风险。以大数据为基础的人工智能在应用之初，许多人是抱着观望、怀疑的心态看待这种新技术的。随着人工智能以拟人化的外在形式、个性化的服务提供、沉浸式的互动过程不断改善使用者的产品体验和心理感受，越来越多的人逐渐成为人工智能的忠实用户，享受着人工智能给自己带来的各种便捷。随着人机互动、万物互联的物联网技术普及，智能家居、智能办公、智能工厂、智能驾驶等人工智能应用场景也不断拓展，个人能够以数字人的存在形式在数字空间提出需求、获得服务，也在不知不觉中向人工智能源源不断地输送着个人信息。

个人在数字空间留下的任何痕迹都被数字化，形成个人信息，并作为人们“联系世界的介质”发挥着重要功能。与此同时，人工智能为了改善服务质量，也倾向于过度收集使用个人信息。这都使得人工智能具有鲜明的去私密化技术特征。也正是在那些人工智能使用者习以为常的个人信息流动中，混合着公共数据和私人数据的大数据被挖掘、整合、分析、利用，人们难以凭自己的感官察觉到隐私权被侵害，个人信息安全面临着更高的风险。

尊重个体选择，坚持知情同意。不同的人对个人信息被知悉、被利用的接受程度不同，应尊重个人意愿，科学合理地执行“知情同意”原则。知情同意原则包括知情和同意两方面，同意必须以知情为前提，没有充分的知情和理解，就不可能有真正意义上的同意。信息、理解和自愿，是知情同意原则的三要素。在全面“知情”的基础上，个人可以自主作出怎样“同意”的意思表示。这就需要在用户使用人工智能时，作出通俗易懂而又清晰明了的提示说明，征得用户对个人信息收集使用的同意。如果个人信息会在不同平台之间流动，需要将流动范围、目标、使用边界让用户知晓。为了良好和流畅的用户体验，也可以给用户提供一次性或分阶段进行授权的选择。要告知用户收集个人信息的范围、方式和用途以及与谁共享个人信息，用户也应当可以选择随时退出。在进行个人信息分析时，应以弹窗或其他形式提示用户注意并实

时授权。设置数据生命周期、按时删除个人信息也是保护个人信息安全的有效方式。

完善技术手段，确保智能向善。技术导致的问题，要善于从技术层面确立解决问题的思路。人工智能时代隐私权面临挑战，其直接的触发因素是技术的演进。从分析式人工智能到生成式人工智能，人工智能技术每一次迭代升级，都可能对隐私权带来新的冲击。因此，技术解决方案必须置于关键位置，应通过完善数据库安全、核心数据加密、个人数据脱敏等技术，建立保护隐私权和个人信息安全的防火墙。个人信息一般会经过收集、存储和使用三个阶段，而这三个阶段都可能存在侵害隐私权和个人信息安全的风险。应根据不同阶段个人信息所处的不同状况，从技术上进行有效保护。在个人信息收集阶段，加强匿名化技术推广运用。收集个人信息虽然不可避免，但只要匿名化，不把个人信息与身份对应，隐私权就不会受到侵害。个人信息存储阶段，要完善加密技术。当前，数据存储主要有数据库存储和云存储两种方式。外部入侵窃取和内部人员未经授权的查看、使用、泄露是存储阶段个人信息安全的主要威胁。要强化数据加密，同时严格数据访问权限控制。个人信息使用阶段，要从技术上加强对个人信息违法使用的实时介入、干扰、阻断，为隐私权和个人信息安全多添一层保护。随着我国法律规则日益完善、保护力度持续加强，特别是民法典、个人信息保护法对隐私权和个人信息保护作出详细规定，明确了个人信息处理活动中权利和义务的边界，人工智能时代我国对隐私权和个人信息安全的法律保护必将迈上更高水平，为人工智能健康发展、更好造福人民群众提供坚强法律保障。

探索人工智能体的模块化治理框架

科技兴则民族兴，科技强则国家强。党的十八大以来，我国高度重视人工智能发展，积极推动互联网、大数据、人工智能和实体经济深度融合，培育壮大智能产业，加快发展新质生产力，为高质量发展提供新动能。习近平总书记指出："要坚持促进发展和依法管理相统一，既大力培育人工智能、物联网、下一代通信网络等新技术新应用，又积极利用法律法规和标准规范引导新技术应用。"习近平总书记的重要论述为我国人工智能发展提供了根本遵循和行动指南。大力发展人工智能，提高人工智能安全治理水平，要把党的二十届三中全会《决定》提出的"建立人工智能安全监管制度"重要部署不折不扣贯彻落实好，准确把握人工智能发展动向，重点关注人工智能前沿技术及其带来的风险挑战，加强前瞻性思考，不断探索人工智能治理的创新方案。

当前，生成式人工智能开创了人机交互新范式，凭借其强大的交互、理解和生成能力，为发展以大型自然语言模型为核心组件，集记忆、规划和工具使用于一体，具备感知和行动能力的人工智能体开辟了广阔前景。人工智能体已成为通用人工智能最重要的前沿研究方向和科技企业竞相布局的新赛道。它以大型自然语言模型为"智慧引擎"，具有自主性、适应性和交互性特征，可显著提高生产效率，增强用户体验，提供超越人类能力的决策支持，已能够应用于软件开发、科学研究等多种真实场景。尽管大规模商业化落地仍在初步探索和孵化阶段，但人工智能体所代表的虚实融合、人机深度互动等趋势对经济社会发展具有重要引领意义。然而，由于技术局限，人工智能体也可能引发复杂的、动态的、不可预见的风险与隐忧。

从设计逻辑看，人工智能体需要通过控制端获得认知能力，通过感知端从周围环境获取和利用信息，最终在行动端成为基于物理实体进行感知和行动的智能系统。

在控制端，大型自然语言模

型作为人工智能体的“大脑”，通过学习海量数据形成知识，构成人工智能体控制系统中的记忆模块，但其在生成内容的可靠性和准确性方面存在风险。比如，模型生成的内容可能不遵循信息源或者与现实世界的真实情况不符，产生所谓“机器幻觉”；由于训练数据中的人类偏见，影响人工智能体的公平决策；等等。

在感知端，为充分理解具体情境下的显性信息和隐性信息，准确感知人类意图，人工智能体将感知范围从纯文本拓展到包括文本、视觉和听觉模式在内的多模态领域。这虽然提升了决策能力，却在融合和分析不同渠道和类型的多源数据时可能引发一系列隐私泄露和数据安全风险。比如，不当使用和分享人脸信息、指纹、声纹等高度个性化、具有永久性的生物特征数据，导致长期甚至永久性的隐私风险。为更好地处理复杂任务，部署多个人工智能体进行规划、合作甚至竞争，以完成和提高任务绩效的多智能体系统将成为主流和常态。多个人工智能体的系统交互就可能引发不可预见的系统性安全风险。即使每个算法在单独操作时看似安全和合理，但组合和交互之后仍可能产生完全不同且难以预测的风险，并迅速演化升级。比如，在股票市场中，如果人工智能被广泛应用，多个算法自动识别股票价格微小变化，同时大量执行高频交易进行套利，就可能引发股票市场闪崩的系统性安全事件。

在行动端，部署于真实物理环境的人工智能体将可能以更为立体、拟人的形象呈现。与虚拟空间不同，现实空间依赖交互式学习方法，人工智能体需要丰富的、全方位的信息感知来观察、学习和行动，通过基于反馈的学习优化能力，这可能对个人隐私构成全面性、侵入性和隐蔽性的风险。比如，解读用户的肢体语言并感知更加复杂的用户活动，未经用户授权持续隐秘地收集数据，一旦系统存在安全漏洞，可能引发巨大的数据安全风险。此外，随着人工智能体自主性不断提升，不仅可能干预和影响人类的认知和情绪，也挑战着人类作为独立决策者和独立行动者的能力与地位。比如，一些聊天机器人在与用户的交互过程中就出现了影响用户情感的输出，有时是负面并且具有操纵性的。

面对人工智能体带来的风险和挑战，要让人工智能体的行为符合人类的意图和价值观，需要探索创新性的治理方案，保证人工智能安全监管制度切实管用。人工智能体的发展正处于“从零到一”的关键期。治理方案应具备以不变应万变的能力，确保该技术的发展与应用始终处于可控的轨道上。人工智能体的开发、训练、部署、运行和服务等环节经过高度专业化的分工，形成了复杂的分层结构。每一层均有不同的参与者、利益相关方和潜在风险因素，使人工智能体具有“模块”化的产业链特质。因此，可以构建一种能够覆盖整个产业链和各个端层的模块化治理框架，从数据模块、算法模块、模型架构等关键节点出发，设计相应的治理模块。例如在部署环节，就可根据应用场景和部署模式的特性，灵活选择、协同组合不同的治理模块，构建与之相匹配的治理方案。模块化治理框架提供了一种具有可操作性的分解方法，通过将治理目标拆解为相对独立但关联耦合的治理模块，渐序推动治理体系形成，不仅提高了治理的灵活性和针对性，还能够适应技术的快速迭代。在构建基于数据、算法、模型和场景等维度的治理模块时，应以技术赋能监管，打造与人工智能体模块化治理框架相适配的、智慧化的治理工具，从而弥合风险动态性与监管静态性之间的张力，实现对特定高风险场景的精准化治理。

要构建面向人工智能体的交互式治理生态。人工智能体具有深度交互性、高度互联性以及动态适应性。相应地，治理方式应当超越传统的以个体为中心的治

理，推动形成广泛互联、多方参与、多层次协作的治理生态。其中，技术开发人员、运营维护人员等技术社群对于人工智能体的治理将起到至关重要的“吹哨人”作用。应更好发挥技术社群的监督优势，在人工智能企业内部构建有效的约束机制。还应积极提高广大用户的数字素养，增强其依法、安全、负责任使用人工智能体的意识，实现与人工智能体的良性交互，推动形成向上向善的运行状态。

来源：人民日报

《国家自然科学基金条例》公布修订

国务院总理李强日前签署国务院令，公布修订后的《国家自然科学基金条例》（以下简称《条例》），自2025年1月1日起施行。《条例》共7章45条，修订的主要内容如下。

一是坚持党中央集中统一领导，明确工作原则。强调国家自然科学基金工作坚持党中央集中统一领导，应当面向世界科技前沿、面向经济主战场、面向国家重大需求、面向人民生命健康。明确国家自然科学基金用于资助基础研究，支持人才培养和团队建设，遵循公开、公平、公正的原则。确定基金资助项目时既要鼓励自由探索，又要坚持目标导向。

二是健全管理体制，适应科技创新发展新趋势新要求。规定基金管理机构、有关部门以及依托单位的工作责任。明确基金主要来源于中央预算拨款，同时鼓励多元化投入，支持更多社会力量参与基础研究，鼓励企业和其他组织投入资金开展联合资助，建立科技创新合作机制。加强信息化建设，完善科研诚信管理信息共享、基金资助项目成果共享等机制。

三是完善资助制度，发挥基金促进基础研究发展的作用。扩大依托单位范围，除高等学校、科学研究机构、有关公益性机构以外，其他符合国家规定的单位可以注册成为依托单位。明确设立专项资金，用于培养青年科学技术人才，支持在科学技术领域取得突出成绩且具有明显创新潜力的青年人才。优化重大原创性、交叉学科创新等基金资助项目的申请与评审程序，完善基金绩效评价制度，要求定期开展评价，根据评价结果调整完善政策、改进管理、优化预算安排。

四是加强科研诚信制度建设，营造良好创新环境。在基金资助项目申请与评审阶段强化科研诚信审核力度，落实科研诚信承诺要求，规定申请人、参与者、依托单位以及基金管理机构工作人员、评审专家等均应当签署科研诚信承诺书，并对有科研诚信严重失信行为的个人、组织，按照国家有关规定实施联合惩戒。

全文如下：

国家自然科学基金条例

（2007年2月24日中华人民共和国国务院令第487号公布

2024 年 11 月 8 日中华人民共和国国务院令第 796 号修订）

第一章　总则

第一条　为了规范国家自然科学基金的使用与管理，提高国家自然科学基金使用效能，加强基础研究，培养科学技术人才，增强自主创新能力，实施创新驱动发展战略，推动实现高水平科技自立自强，根据《中华人民共和国科学技术进步法》，制定本条例。

第二条　国家自然科学基金工作坚持党中央集中统一领导。国家自然科学基金工作应当面向世界科技前沿、面向经济主战场、面向国家重大需求、面向人民生命健康。

第三条　国家设立国家自然科学基金，用于资助基础研究，支持人才培养和团队建设。国家自然科学基金资助工作实行尊重科学、发扬民主、提倡竞争、促进合作、激励创新、引领未来的方针，培育和践行社会主义核心价值观，遵循公开、公平、公正的原则。

第四条　国家自然科学基金主要来源于中央预算拨款。国家鼓励地方人民政府、企业和其他组织投入资金开展联合资助，建立科技创新合作机制。国家鼓励社会力量向国家自然科学基金捐赠。国家自然科学基金的资金应当全额纳入预算管理。

第五条　国家自然科学基金资助项目（以下简称基金资助项目），应当推动基础研究自由探索和目标导向有机结合，根据基金发展规划和年度基金项目指南予以确定。确定基金资助项目应当充分发挥专家的作用，采取宏观引导、自主申请、平等竞争、同行评审、择优支持的机制。

第六条　国家自然科学基金管理机构（以下简称基金管理机构）依法管理国家自然科学基金，制定国家自然科学基金资助与管理制度，负责资助计划、项目设置，以及评审、立项、监督等组织实施工作，提升国家自然科学基金资助效能。国务院科学技术主管部门依法对国家自然科学基金工作进行宏观管理、统筹协调和监督检查。国务院财政部门依法对国家自然科学基金的预算、财务、会计进行管理和监督，对基金的运行情况开展绩效评价，根据评价结果调整完善政策、改进管理、优化预算安排。审计机关依法对国家自然科学基金的使用与管理进行审计监督。

第七条　基金管理机构应当加强国家自然科学基金工作信息化建设，推动完善科研诚信管理信息共享、基金资助项目成果共享等机制，加强基金资助项目与其他科学技术计划项目的衔接与协调。

第二章　组织与规划

第八条　基金管理机构应当根据国民经济和社会发展规划、科学技术发展规划以及科学技术发展状况，制定基金发展规划和年度基金项目指南。基金发展规划应当明确优先发展的领域，年度基金项目指南应当规定优先支持的项目范围。国家自然科学基金应当设立专项资金，用于培养青年科学技术人才，支持在科学技术领域取得突出成绩且具有明显创新潜力的青年人才。基金管理机构制定基金发展规划和年度基金项目指南，应当广泛听取高等学校、科学研究机构、学术团体和有关国家机关、企业的意见，组织有关专家进行科学论证。年度基金项目指南应当在受理基金资助项目申请起始之日 30 日前公布。

第九条　基金管理机构应当确定依托单位作为基金资助项目及其资金管理的责任主体。中华人民共和国境内具有独立法人资格、具备基础研究能力的高等学校、科学研究机构和其他公益性机构，以及其他符合国家规定的单位，可以在基金管理机构注册为依托单位。基金管理机构应当定期公布已注册的依托单位名单。依托单位应当建立和完善基金资助项目管理、科研诚信、科技伦理管理等制度，遵守科学技术活

动管理规范。

第十条 依托单位在基金资助管理工作中履行下列职责：（一）组织申请人申请国家自然科学基金资助；（二）审核申请人或者项目负责人所提交材料的真实性、完整性和合法性；（三）提供基金资助项目实施的条件，保障项目负责人和参与者实施基金资助项目的时间；（四）跟踪基金资助项目的实施，管理和监督基金资助资金的使用；（五）配合基金管理机构、科学技术主管部门和相关行业主管部门对基金资助项目的实施和资金的使用进行监督、检查和绩效评价；（六）加强基金资助项目研究成果的知识产权管理，支持研究成果开放获取，促进研究成果转化；（七）对违背科研诚信和科技伦理的行为进行调查处理；（八）完成基金管理机构委托的其他工作。基金管理机构、科学技术主管部门和相关行业主管部门对依托单位的基金资助管理工作进行指导、监督。

第三章 申请与评审

第十一条 依托单位的科学技术人员具备下列条件的，可以申请国家自然科学基金资助：（一）具有承担基础研究课题或者其他从事基础研究的经历；（二）具有高级专业技术职务（职称）或者具有博士学位，或者有2名与其研究领域相同、具有高级专业技术职务（职称）的科学技术人员推荐。符合条件的从事基础研究的科学技术人员无工作单位或者所在单位不是依托单位的，取得依托单位的同意，可以依照本条例规定申请国家自然科学基金资助。申请人应当是申请基金资助项目的负责人。

第十二条 申请人申请国家自然科学基金资助，应当以年度基金项目指南为基础确定研究项目，在规定期限内通过依托单位向基金管理机构提出书面申请。申请人申请国家自然科学基金资助，应当提交证明申请人符合本条例第十一条规定条件的材料。申请人申请基金资助的项目研究内容已获得其他资助的，应当在申请材料中说明资助情况。申请人应当对所提交申请材料的真实性、完整性和合法性负责。申请人、参与者及其依托单位应当遵守科研诚信和科技伦理要求，签署科研诚信和科技伦理承诺书，承诺不从事危害国家安全、损害社会公共利益、危害人体健康、违背科研诚信和科技伦理的科学技术研究开发和应用活动，没有虚构、伪造、剽窃、篡改等行为。

第十三条 基金管理机构应当自基金资助项目申请截止之日起45日内，完成对申请材料的初步审查。有下列情形之一的，不予受理，通过依托单位书面通知申请人，并说明理由：（一）申请人不符合本条例规定条件的；（二）申请材料不符合年度基金项目指南要求的；（三）申请人、参与者因有违背科研诚信和科技伦理等行为被禁止承担或者参与财政性资金支持的科学技术活动的；（四）申请人、参与者申请基金资助项目超过基金管理机构规定的数量的。

第十四条 基金管理机构应当聘请具有较高的学术水平、良好的职业道德的同行专家，对已受理的基金资助项目申请进行评审。评审专家聘请、评审活动管理和相关监督保障机制的具体办法由基金管理机构制定。

第十五条 基金管理机构对已受理的基金资助项目申请，应当先从同行专家库中随机选择3名以上专家进行通讯评审。基金管理机构应当根据通讯评审评分等情况确定进入会议评审的项目申请，组织专家进行会议评审。对因国家经济、社会发展特殊需要或者其他特殊情况临时提出的基金资助项目申请，可以只进行通讯评审或者会议评审。评审专家对基金管理机构安排其评审的基金资助项目申请认为难以作出学术判断或者没有精力评审的，应当及时告知基金管理机构；基金管理机构应当依照本条例规定，选择其他评审专家进行评审。

第十六条 评审专家按照国家自然科学基金资助导向，对基

金资助项目申请从科学价值、创新性、社会影响以及研究方案的可行性等方面进行独立判断和评价，提出评审意见。评审专家对基金资助项目申请提出评审意见，还应当考虑申请人和参与者的研究经历、基金资助资金使用计划的合理性、研究内容获得其他资助的情况、申请人实施基金资助项目的情况以及继续予以资助的必要性。会议评审提出的评审意见应当通过投票表决。

第十七条 为了鼓励原创性基础研究工作，对重大原创性、交叉学科创新等基金资助项目，基金管理机构可以制定专门的申请与评审规定。

第十八条 基金管理机构根据本条例的规定和专家提出的评审意见，决定予以资助的研究项目。基金管理机构不得以与评审专家有不同的学术观点为由否定专家的评审意见。基金管理机构在作出资助决定前，应当按照国家有关规定，通过科研诚信管理信息系统，及时开展严重失信行为数据比对、核查等工作。基金管理机构决定予以资助的，应当及时书面通知申请人和依托单位；决定不予资助的，应当及时书面通知申请人和依托单位，并说明理由。基金管理机构应当整理专家评审意见，并向申请人提供。

第十九条 申请人对基金管理机构作出的不予受理或者不予资助的决定不服的，可以自收到通知之日起 15 日内，向基金管理机构提出书面复审请求。对评审专家的学术判断有不同意见，不得作为提出复审请求的理由。基金管理机构对申请人提出的复审请求，应当自收到之日起 60 日内完成审查。基金管理机构认为原决定符合本条例规定的，予以维持，并书面通知申请人；认为原决定不符合本条例规定的，撤销原决定，重新对申请人的基金资助项目申请组织评审专家进行评审、作出决定，并书面通知申请人和依托单位。

第二十条 在基金资助项目评审工作中，基金管理机构工作人员、评审专家有下列情形之一的，应当申请回避：（一）基金管理机构工作人员、评审专家是申请人、参与者近亲属，或者与其有其他关系可能影响公正评审的；（二）评审专家自己申请的基金资助项目与申请人申请的基金资助项目相同或者相近的；（三）评审专家与申请人、参与者属于同一法人单位的。基金管理机构根据申请，经审查作出是否回避的决定。基金管理机构作出是否回避的决定，应当说明理由。基金管理机构发现其工作人员、评审专家有本条第一款规定的需要回避情形的，可以不经申请直接作出回避决定。基金资助项目申请人可以向基金管理机构提供 3 名以内不适宜评审其申请的评审专家名单，基金管理机构在选择评审专家时应当根据实际情况予以考虑。

第二十一条 基金管理机构工作人员、评审专家均应当签署科研诚信承诺书。基金管理机构工作人员不得申请或者参与申请国家自然科学基金资助，不得干预评审专家的评审工作。评审专家应当遵守科研诚信要求和评审行为规范，独立、客观、公正开展评审工作，不得由他人代为评审，不得有接受请托、说情干预等不公正评审行为，不得利用工作便利谋取不正当利益。基金管理机构工作人员、评审专家均应当依法承担保密义务，不得以任何方式披露未公开的评审专家的基本情况、评审意见、评审结果等与评审有关的信息。

第四章 资助与实施

第二十二条 申请人收到基金管理机构的基金资助通知后，即作为基金资助项目的负责人组织开展研究工作。依托单位和项目负责人自收到基金资助通知之日起 20 日内，按照评审专家的评审意见、基金管理机构确定的基金资助额度填写项目计划书，报基金管理机构核准。依托单位和项目负责人填写项目计划书，除根据评审专家的评审意见和基金管理机构确定的基金资助额度对

已提交的申请书内容进行调整外，不得对其他内容进行变更。

第二十三条 基金管理机构对本年度予以资助的研究项目，应当按照《中华人民共和国预算法》和国库集中支付制度等有关规定，办理基金资助项目拨款。项目负责人应当依法依规使用基金资助资金，依托单位应当对项目负责人使用基金资助资金的情况进行管理和监督，在其职责范围内及时审批项目资金调整事项。项目负责人、参与者及其依托单位不得以任何方式虚报、冒领、贪污、侵占、挪用、截留基金资助资金。基金资助资金使用与管理的具体办法由国务院财政部门会同基金管理机构制定。

第二十四条 项目负责人应当根据项目计划书组织开展研究工作，做好基金资助项目实施情况的原始记录，通过依托单位向基金管理机构提交项目年度进展报告。依托单位应当审核项目年度进展报告，查看基金资助项目实施情况的原始记录，并向基金管理机构提交年度基金资助项目管理报告。基金管理机构应当对项目年度进展报告和年度基金资助项目管理报告进行审查。

第二十五条 基金资助项目实施中，依托单位不得擅自变更项目负责人。项目负责人有下列情形之一的，依托单位应当及时提出变更项目负责人或者终止基金资助项目实施的申请，报基金管理机构批准；基金管理机构也可以直接作出终止基金资助项目实施的决定：（一）不再是依托单位科学技术人员的；（二）不能继续开展研究工作的；（三）其他客观原因导致项目无法继续实施的。项目负责人调入另一依托单位工作的，经所在依托单位与原依托单位协商一致，由原依托单位提出变更依托单位的申请，报基金管理机构批准。协商不一致的，基金管理机构作出终止该项目负责人所负责的基金资助项目实施的决定。

第二十六条 基金资助项目实施中，研究内容或者研究计划需要作出重大调整的，项目负责人应当及时提出申请，经依托单位审核后报基金管理机构批准。

第二十七条 自基金资助项目资助期满之日起60日内，项目负责人应当通过依托单位向基金管理机构提交结题报告；基金资助项目取得研究成果的，应当同时提交研究成果报告。依托单位应当对结题报告进行审核，建立基金资助项目档案。依托单位审核结题报告，应当查看基金资助项目实施情况的原始记录。

第二十八条 基金管理机构应当及时审查结题报告。对不符合结题要求的，应当提出处理意见，并书面通知依托单位和项目负责人。基金管理机构应当按照国家科学技术报告制度的要求，将结题报告、研究成果报告和基金资助项目摘要予以公布，促进基金资助项目成果的传播与共享。

第二十九条 发表基金资助项目取得的研究成果，应当注明得到国家自然科学基金资助。基金资助项目所形成的科学数据，应当由依托单位按照国家科学数据管理的有关规定汇交到相关科学数据中心。

第五章 监督与管理

第三十条 基金管理机构应当对基金资助项目实施情况、依托单位履行职责情况进行以抽查为主要方式的监督检查，确保监督检查全面覆盖。抽查时应当查看基金资助项目实施情况的原始记录和资金使用与管理情况。抽查结果应当予以记录并公布。基金管理机构应当建立项目负责人和依托单位的基金资助项目信誉档案。

第三十一条 基金管理机构应当定期对评审专家履行评审职责情况进行评估，评估时应当参考申请人的意见；根据评估结果，建立评审专家信誉档案。

第三十二条 基金管理机构应当在每个会计年度结束时，公布本年度基金资助的项目、基金资助资金的拨付情况以及对违反本条例规定行为的处罚情况等。基金管理机构应当配合国务院财

政部门和科学技术主管部门开展基金绩效评价，并应当按照全面实施预算绩效管理的要求，建立健全绩效管理制度，定期对基金使用情况开展绩效评价，强化绩效评价结果应用，加大绩效信息公开力度。

第三十三条 任何单位或者个人发现基金管理机构及其工作人员、评审专家、依托单位及其相关工作人员、申请人或者项目负责人、参与者有违反本条例规定行为的，可以检举或者控告。基金管理机构应当公布联系电话、通讯地址和电子邮件地址。

第三十四条 基金管理机构依照本条例规定对外公开有关信息，应当遵守国家有关保密规定。基金管理机构应当健全科技安全制度和风险防范机制，加强科学技术研究、开发与应用活动的安全管理，支持国家安全领域科技创新，增强科技创新支撑国家安全的能力和水平。

第三十五条 对违反本条例规定，有科研诚信严重失信行为的个人、组织，按照国家有关规定记入科研诚信严重失信行为数据库，实施联合惩戒。

第六章 法律责任

第三十六条 申请人、参与者有下列行为之一的，由基金管理机构给予警告，正在申请基金资助的，取消其当年申请或者参与申请国家自然科学基金资助的资格；其申请项目已实施资助的，撤销原资助决定，追回已拨付的基金资助资金；情节较重的，1 至 3 年不得申请或者参与申请国家自然科学基金资助；情节严重的，3 至 5 年不得申请或者参与申请国家自然科学基金资助：（一）虚构、伪造、剽窃、篡改申请材料的；（二）以请托、贿赂等不正当方式干预评审工作的；（三）有其他违背科研诚信和科技伦理行为的。申请人、参与者有前款所列行为的，由依托单位按照人事管理有关规定取消其在一定期限内参加专业技术职称评审的资格。

第三十七条 项目负责人、参与者有下列行为之一的，由基金管理机构给予警告，暂缓拨付基金资助资金，并责令限期改正；逾期不改正的，撤销原资助决定，追回已拨付的基金资助资金；情节较重的，1 至 5 年不得申请或者参与申请国家自然科学基金资助；情节严重的，5 至 7 年不得申请或者参与申请国家自然科学基金资助：（一）擅自变更研究方向或者降低申报指标的；（二）不依照本条例规定提交项目年度进展报告、结题报告或者研究成果报告的；（三）提交弄虚作假的报告、原始记录或者相关材料的；（四）成果发表署名不实或者虚假标注资助信息的；（五）不配合监督、检查或者绩效评价工作的；（六）虚报、冒领、贪污、侵占、挪用、截留基金资助资金的；（七）虚构、伪造、剽窃、篡改研究数据或者结果的；（八）有其他违背科研诚信和科技伦理行为的。项目负责人、参与者有前款所列行为的，由依托单位按照人事管理有关规定取消其在一定期限内参加专业技术职称评审的资格。

第三十八条 依托单位有下列情形之一的，由基金管理机构给予警告，责令限期改正；逾期不改正的，核减基金资助资金，并可以暂停拨付或者追回已拨付的基金资助资金；情节较重的，t1 至 3 年不得作为依托单位；情节严重的，向社会公布其违法行为，3 至 5 年不得作为依托单位：（一）不履行保障基金资助项目研究条件的职责的；（二）不对申请人或者项目负责人提交的材料的真实性、完整性和合法性进行审核的；（三）不依照本条例规定提交项目年度进展报告、年度基金资助项目管理报告、结题报告和研究成果报告的；（四）组织、参与、纵容、包庇申请人或者项目负责人、参与者弄虚作假的；（五）擅自变更项目负责人的；（六）不配合基金管理机构监督、检查基金资助项目实施的；（七）虚报、冒领、贪污、侵占、挪用、截留基金资助资金的；（八）违反保密规定或者管理严重失职，造成不良影响或者损失的；（九）以请托、

贿赂等不正当方式干预评审工作的；（十）不履行科研诚信和科技伦理相关管理职责的。

第三十九条 评审专家有下列行为之一的，由基金管理机构给予警告，责令限期改正；情节较重的，2 至 7 年不得聘请其为评审专家；情节严重的，向社会公布其违法行为，不得再聘请其为评审专家：（一）不履行基金管理机构规定的评审职责的；（二）未依照本条例规定申请回避的；（三）披露未公开的与评审有关的信息的；（四）对基金资助项目申请有接受请托、说情干预等不公正评审行为的；（五）利用工作便利谋取不正当利益的。

第四十条 基金管理机构工作人员有下列行为之一的，依法依规给予组织处理、处分：（一）未依照本条例规定申请回避的；（二）披露未公开的与评审有关的信息的；（三）干预评审专家评审工作的；（四）利用工作便利谋取不正当利益的；（五）有其他滥用职权、玩忽职守、徇私舞弊行为的。

第四十一条 违反有关财政法律、行政法规规定的，依法依规予以处罚、处分。

第四十二条 申请人或者项目负责人、参与者违反本条例规定，情节特别严重的，或者从事危害国家安全、损害社会公共利益、危害人体健康、违背科研诚信和科技伦理的科学技术研究开发和应用活动的，终身不得申请或者参与申请国家自然科学基金资助。违反本条例规定受到处理的科学技术人员，在处理期内不得聘请其为评审专家；情节特别严重的，终身不得聘请其为评审专家。

第四十三条 违反本条例规定，构成违反治安管理行为的，依法给予治安管理处罚；构成犯罪的，依法追究刑事责任。

第七章 附 则

第四十四条 基金管理机构在基金资助工作中，涉及项目组织实施费和与基础研究有关的学术交流活动、基础研究环境建设活动的基金资助资金的使用与管理的，按照国务院财政部门的有关规定执行。

第四十五条 本条例自 2025 年 1 月 1 日起施行。

来源：中国政府网

喜报 XI BAO

《自动化学报》（英文版）和《自动化学报》入选中国科技期刊卓越行动计划二期领军期刊

2024 年 11 月 28 日，中国科技期刊卓越行动计划办公室发布了中国科技期刊卓越行动计划二期项目入选名单，IEEE/CAA Journal of Automatica Sinica（《自动化学报》（英文版），简称 JAS）入选中国科技期刊卓越行动计划二期英文领军期刊，《自动化学报》入选中国科技期刊卓越行动计划二期中文领军期刊，详情请查看：https://mp.weixin.qq.com/s/jI9emdOn1MJWKQ_s_-paeQ

中国自动化学会

中国自动化学会（Chinese Association of Automation，缩写CAA）于1961年成立，是我国最早成立的国家一级学术团体之一，是中国科学技术协会的组成部分，是发展我国自动化科技事业的重要社会力量。学会现有个人会员8万余人，团体会员单位300余个，专业委员会64个，工作委员会11个，30个省、自治区、直辖市设有地方学会组织，覆盖了我国自动化科学技术领域的各个层面。

中国自动化学会在改革中求发展，不断加强群众组织力、学术引领力、社会公信力和国际影响力。近年来，中国自动化学会重点从学术交流与应用推广、组织建设与会员服务、科技评估与人才评价、课题研究与决策支撑、科学普及与继续教育等方面开拓创新，推动中国自动化科学和事业的发展和壮大，成为连接政府、产业、学术、科研、会员的重要纽带，致力于成为国内外有影响力的现代社会团体组织。

学会品牌学术活动

·中国自动化大会 ·中国认知计算与混合智能学术大会 ·中国自动化与人工智能教育大会
·国家智能车发展大会 ·国家机器人发展大会 ·国家智能制造大会 ·国家工业软件大会
·青年菁英系列活动 ·智能自动化学科前沿讲习班 ·钱学森国际杰出科学奖系列讲座
·中国控制会议 ·中国过程控制会议 ·青年学术年会

学会奖励奖项

·CAA科技进步奖 ·CAA自然科学奖 ·CAA技术发明奖 ·CAA青年科技奖
·CAA科技成就奖 ·CAA论文卓越行动工程 ·CAA青年托举工程
·CAA教育教学成果评价 ·CAA科学普及成果

学会主办期刊

·中国自动化学会通讯 ·自动化学报 ·自动化学报（英文版）
·信息与控制 ·机器人 ·模式识别与人工智能 ·电气传动
·自动化博览 ·国际智能控制与系统学报(英文)

官方微信　　官方微博

地址：北京市海淀区中关村东路95号自动化大厦
网址:http://www.caa.org.cn/
电话：010-62522472
传真：010-62522248
邮箱：caa@ia.ac.cn
邮编：100190