



基于“智简”范式的6G通信技术的发展

□ 张 平



张平，北京邮电大学教授，中国工程院院士，长期致力于移动通信理论研究和技术创新，提出了宽带4G TDD移动传输理论与方法，创建了多域认知的无线异构组网架构，攻克了测试仪器动态可重构技术，建立完整的语义信息理论框架，为我国自主技术成为国际主流作出了基础性的贡献。先后获国家科学技术进步奖特等奖等多项奖励，获首届“全国创新争先奖”奖章、光华工程科技奖、何梁何利基金科学与技术进步奖。

“智简”范式的三大突破

以“智简”范式为指导，语义通信颠覆了传统通信技术体系，为通信与人工智能的深度融合开辟了新的技术路径。团队在基础理论研究、核心技术攻关、工程试验验证等方面均实现了突破性进展。

一是基础理论的突破。建立了语义信息理论体系，这一体系是对经典香农信息理论的拓展，为我国信息理论创新奠定了坚实基础。随着经典信息论指导下的通信技术逼近理论极限，如何突破成为业界焦点。与语法通信关注通信符号的精确传输不同，语义通信主要关注信息含义的准确传递，通过深入挖掘信息的语义层次内涵，探索语义信息表征、度量、知识库构建、弹性编码及协同传输等核心方法，推动了通信与人工智能融合的方法论创新与技术体系变革，从基本原理上改变了传统信息通信体系。语义通信的数学理论揭示了语义信息的本质，构建了完整理论框架，拓展了通信系统的理论极限，将成为通信与人工智能融合的新指导理论。这一成果有望助力我国6G技术突破瓶颈。目前，该理论已得到国内外专家的广泛认可，认为语义通信是未来信息通信技术的重要方向，具有显著的理论与应用价值。

二是核心技术突破。构建了自洽的通信与AI一体化融合的语义信息通信技术体系，为智能泛在化、可持续发展创设了一条“新道”。当前，AI与通信融合发展成为技术趋势，但两者的内在规律迥异：通信系统依赖理论指导，具有可解释性，性能受能耗、芯片、成本等工程应用的多重制约；而AI系统以实验为指导，不具有可解释性，虽在模拟人类智能方面取得了巨大成功，但以大算力、大数据和大模型（算法）带来的成本剧增和资源消耗为代价。通信和AI堆叠式、浅层次融合，难以打破各自内在规律，特别是从网络到用户物理空间的割裂，使得智能泛在化面临智能传递与分发的“最后一公里”鸿沟。为此，以“智简”范式指导，团队提出了通信与AI一体化融合的载体、AI构建的信息通信表征单元——语义基；

赋能AI一体化融入通信系统，在成本和功耗等制约条件下为系统提供额外增益，不再是通过堆叠式增加效能，而是采用简约设计、智能地做减法提升性能，这是与常规AI方法最大的不同。

基于语义基的一体化融合纽带，语义通信基础理论及在通信领域引发的突破带来了传统通信系统设计的变革，通信系统的设计不再是传统通信系统在工程实现时采用堆叠式、模块分离延长线的碎片化创新，而是采用AI的系统化优势并与通信系统模块化、低成本、可解释的一体化融合，获得了巨大的系统性能增益。团队从基础理论、设计思想、核心方法等层面全面革新语法通信。在设计思想上，以语义信息（语义基作为表征单元）为对象，深度融合通信与人工智能，以语义知识库为关键支撑，重新塑造通信系统的信息提取、表征、编译码、综合和重建等环节；在核心方法上，针对语义任务、信道能力，实现弹性信息压缩，并通过信源—信道—信宿的联合设计，提出模分多址等方法，达到多用户高效组网的目的。因此，语义通信通过理论、思想、方法层面的升华，达到“锻铁成钢”的目的。

三是工程试验验证。继理论证明后，我国建成的世界首个面向6G的智能通信融合外场试验网，首次验证了基于智简语义通信的智能泛在化与可持续发展能力，是支持6G可持续发展的拐点技术，突破了经典信息论的框架限制。该技术采用低制程芯片，替代高精密封程工艺，有效解决了可持续发展难题。语义通信的基础理论和核心技术是我国“从0到1”原始技术创新的成功典范，证明了我国不仅可以在制造、应用等方面取得成就，也可以在基础理论和核心技术方面位居世界一流，发挥引领作用。

团队在智简网络、可持续发展等方面前瞻布局，针对新的系统理论验证与关键技术验证需求，建设语义通信科学基础设施，开展语义通信关键技术研发与验证，支撑智简语义通信领域创新链的建设，为智简语义通信创新发展提供一个真实外场试验验证环境。

该试验面向6G通信与人工智能融合，展示了语义通信在智慧模型传输、沉浸式视频传输、无人车联、无

人机网络、泛在连接卫星通信及大规模物联网等场景中的优异性能，验证了语义通信在4G/5G链路上可以达到6G传输能力，频谱效率更是突破了香农极限，同时在容量、覆盖、效率三项通信核心基础指标上同时实现了多倍性能提升。试验结果还表明，在极限通信状态下的语义通信，打破了传统通信理论的局限，对未来的各行各业，尤其是国防通信、内生安全等具有重大的意义。

重大意义

语义通信理论与技术体系，有望成为开启新一轮科技革命和产业变革的一把钥匙，为我国智能时代的信息通信发展赢得先机。AI和通信技术的深度融合，开辟了一条革命性、颠覆性“新道”。这条创新之道的重要意义主要有三点。其一，基础理论的突破。这是中国在国际竞争激烈的信息通信领域首次提出并系统阐述了自己的基础理论——语义信息论，揭示了语义信息的本质，构建了完整模型，并证明了最优语义编码的存在性定理，为这项技术的发展指明了方向。其二，考虑到智能泛在化及可持续发展的需求，这无疑是对传统通信技术的极大挑战。在未来的网络中，流动的不再仅仅是冷冰冰的符号，而将是丰富的知识。AI通过实时通信的桥梁，将构建出更为庞大的知识体系。借助语义信息的深度应用，通信与AI的融合将成功跨越“最后一公里”鸿沟，极大地推动无线AI、具身智能等前沿应用的成熟与普及，加速通信与AI技术的快速进步。其三，通过30年的努力，中国在移动通信领域实现了从跟随到引领的跨越，改变了全球移动通信格局。如今，面对第三次信息科技中心转移的机遇与挑战，中国正通过基础理论创新，开辟新一代通信技术路径，构建新的信息通信体制与生态。



牵住深化科技体制改革的“牛鼻子”

□ 王亚强

党的二十届三中全会审议通过《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》，紧紧围绕推进中国式现代化，明确提出构建支持全面创新体制机制的战略任务，并对深化科技体制改革作出系统部署，全面擘画了深化科技体制改革的新蓝图。

深化科技体制改革是全面深化改革的重要内容和关键环节。深化科技体制改革，本质上是通过调整生产关系，使之适应生产力的发展，有效调动科技人员的积极性、主动性、创造性，从而解放和发展科技第一生产力，充分释放创新第一动力。深化科技体制改革必须坚持走中国特色自主创新道路，不断完善科技创新基础性制度和中国特色国家创新体系，聚焦加快实现高水平科技自立自强，更好促进科技与经济紧密结合，推进科技治理体系和治理能力现代化，为建成科技强国、推进中国式现代化提供强大动力和制度保障。

党的十八大以来，以习近平同志为核心的党中央部署实施创新驱动发展战略，把科技领域作为最需要不断改革的领域，坚持目标导向和问题导向相结合，引领科技体制改革不断谱写新篇章。近年来，科技体制改革全

面发力、多点突破、纵深推进，在加强基础研究、扩大科研自主权、完善科技成果评价机制、促进科技成果转化、强化企业科技创新主体地位等方面，解开了束缚科技创新的“粗索细绳”，充分激发了创新活力。改革只有进行时，没有完成时。面对推进中国式现代化和实现高水平科技自立自强的时代需要，科技体制改革在支持全面创新、强化改革攻坚、促进改革协同等方面面临新的更高要求，必须抓住关键环节、牵住“牛鼻子”，推进科技体制改革不断取得新突破。

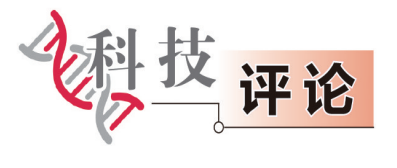
坚持以鼓励创新为原则，加强科技体制改革与政策统筹。中国式现代化要靠科技现代化作支撑，实现高质量发展要靠科技创新培育新动能。在制定科技创新相关政策时，坚持以鼓励和支持创新作为基本原则和政策取向，助推新质生产力培育壮大，审慎出台限制新技术、新产业、新模式发展的政策措施。支持科技创新不完全是科技部门的事，要建立跨部门的科技创新政策统筹协调机制，确保各部门制定政策符合鼓励创新的基本原则，符合科技创新的基本规律。建立完善科技创新政策一致性评估制度，定期评估出台科技创新政策的实施效果、实际影响，及时修订或废止对科技创新

产生负面影响的政策内容。进一步增强改革系统性、整体性、协同性，协同推进科技领域与经济社会领域改革，改革优化医药健康、人工智能、商业航天等行业监管政策，加大政府采购自主创新产品力度，营造创新友好型制度环境。

聚焦束缚创新的痛点卡点，持续为科研人员松绑赋能。加强企业主导的产学研深度融合，建立“企业出题、高校院所答题”机制，探索支持企业牵头的创新联合体申报重大科技项目、承接重大攻关任务，进一步强化对高校科研人员参与横向科研项目的激励措施。要创新不问出身，英雄不论出处，积极改革科技项目立项和组织管理方式，避免将论文、职称、学历、奖项等作为科研人员申报科技项目的限制性条件。针对承担核心技术攻关、基础研究、应用研究和技术开发、社会公益研究等创新活动的科技人员，分类制定科学合理的评价标准，避免评价标准“一刀切”。建立健全具有操作性的科技创新容错机制，激励科研人员积极承担风险性较高的科技项目，大胆探索、挑战未知，消除科研人员的后顾之忧。

紧抓高水平科技自立自强和高质量发展，深化教育科技人才体制机制

一体改革。根据科技创新发展态势，聚焦人工智能、集成电路、智能制造等关键领域，动态调整优化高校学科和专业设置，大力加强“卡脖子”技术和学科交叉复合型人才培养。加快建立健全科教协同育人机制，支持高校与科技型企业联合建设特色研究院、产教融合共同体等，深化工程硕博培养改革，加大工程技术人才培养力度，探索职业技能等级与企业技能认证证书互认机制。加强创新创业教育，注重培养学生创新思维、批判性思维 and 创新能力。积极改革完善社保、职称、档案相关制度，打通高校、科研院所和企业人才交流通道，鼓励高校聘请企业高层次人才担任兼职研究人员或者兼职教师，鼓励高校院所科技人员到企业兼职、挂职、参与项目合作，依法取得报酬。着力构建具有国际竞争力的引才用才机制，下放外籍人才工作许可审批权，加大对急需紧缺科技创新人才、科技服务人才的引进力度。



AI尽头是电力，随着AI技术的广泛应用和深入发展，其背后的算力需求激增，进而引发了全球对数据中心及计算设备能耗激增的广泛关注。据权威统计数据 displays，AI技术的普及与应用已显著推动了全球数据中心能耗的攀升，成为能源特别是电力消耗的一大推手，这也是人们在追求科技进步的同时，需要亟待解决的重要议题。

全球AI耗电现状

总体来看，伴随着人工智能技术的迅猛发展，以及产业指数级增长，作为重要支撑的数据中心耗电总量和占比大幅上升。据测算，AI大语言模型GPT-3一次训练的耗电量就达1287兆瓦时，大概相当于3000辆特斯拉电动汽车共同开跑、每辆车跑20万英里所耗电量的总和。ChatGPT按每天响应约2亿个请求计算，这一过程消耗超过50万度电。国际能源署2024年1月发布报告称，2022年全球数据中心的总耗电量约460太瓦时（1太瓦时=1×10⁹千瓦时=10亿度），约占全球用电量的2%；到2026年，这一数据将超过1000太瓦时，大约是整个日本2022年全年的用电量。与此同时，大量的碳排放和数百万加仑淡水的消耗也不容小觑。

欧美等国家AI耗电形势严峻。以跨国公司数据中心“扎堆”的爱尔兰为例，数据中心的耗电量已超过该国所有城市家庭用电量的总和，占比约为21%。波士顿咨询集团则预计，到2030年，美国数据中心的用电量将是2022年的三倍，而这一增幅主要来自人工智能。OpenAI首席执行官萨姆·奥特曼和特斯拉首席执行官马斯克都曾多次表达了对电力供应的担忧。

从中国来看，AI耗电问题也是日益严重。2022年全年，我国数据中心耗电量达到2700亿千瓦时，占全社会用电量约3%。预计到2025年，这一占比将提升至5%，到2030年全国数据中心耗电量将接近4000亿千瓦时，数据中心的节能降耗迫在眉睫。数据中心是数字经济发展的基石，从全国一体化大数据中心体系想到部署“东数西算”，数字中国战略持续深化落地，算力不断提高，必然需要更高的能源电力，7×24小时连续运行，电力成本占运营成本成本的60%—70%，规模增长迅速。伴随着我国产业转型升级和数据中心规模不断扩大，耗电量会持续攀升。

总之，人工智能及支撑它的数据中心所带来的耗电问题困扰着全世界，是制约人工智能技术飞速发展的主要障碍之一。

AI为何成为耗电大户

AI成为耗电大户的原因比较复杂，涉及数据处理、计算需求、芯片技术、算力中心等多个方面。

首先，从芯片上分析，大语言模型需要巨大的算力支撑，特别是拥有并行计算能力的GPU（图形处理器）。在大模型训练过程中，需要多块GPU接连不断运转，一块GPU的能耗比CPU（中央处理器）高出10—15倍。如一块英伟达A100 GPU功耗为400瓦。GPT-3训练用到了1024块A100芯片，而GPT-4更攀升至25000块，后续将更新至H100、B200，型号的提升和数量的激增导致能耗显著增加。到2027年，英伟达将推出150万台A100服务器，其中95%被应用于AI行业。每一台服务器如果内置8块A100芯片，以1240万块A100芯片的耗电量进行估算，耗电量将高达85—134太瓦时，这个已接近瑞典或荷兰全年总用电量，相当于当前全球用电量的0.5%。

其次，从运行机制上分析。GPU架构在AI计算中虽具备并行处理优势，非常适合AI算法，可以节省计算时间，但其当前利用率较低（32%—36%），且故障率高，导致训练周期长，如GPT-4训练耗时长达90—100天，进而引发长期的能源消耗。相比之下，CPU历经近80年发展，技术成熟且节能高效；而GPU架构出现不到30年，尚存利用率低与差错率高的挑战，这导致了能源消耗的增加。因此，优化GPU架构以提高其效率和稳定性，对于减少能源消耗至关重要。

最后，大语言模型遇从规模效应，随着参数和数据规模的提高，大模型的智能表现将出现跃升。因此，各大公司都在不断地扩大模型规模，这给电能消耗带来了巨大的压力。如果大模型所用到的算力集中在一个数据中心里，在有限时空范围内进行训练，会给局部电网带来非常大的用电负荷，甚至导致整个电网的崩溃。

人工智能为何如此耗电

□ 孙松林

AI耗电问题的解决方案

虽然AI的耗电量在社会总用电量中的占比不大，但与家庭用电量相比，已经非常可观了，而且增速太快，如果不加以控制，将导致巨大能源压力。因此如何应对人工智能带来的巨大能耗，是业界一直高度重视的问题。

首先，在技术层面，通过各种方法降低对芯片的功耗需求是业界努力的方向。英伟达、谷歌公司等做了很多的尝试与努力，通过优化算法模型，如牺牲一点性能以大大减少计算资源和时间，进而带来能耗效率的大幅提升；用光纤连接芯片，减少原有芯片外部的电力驱动和电光/光电转换环节，从而减少能耗；研发高效的芯片级散热技术将散热效率提升百倍；积极研发新型计算和传输架构以提升计算机处理速度并降低能耗；等等。

其次，数据中心的能耗可以通过精心布局加以优化。如针对能耗中的散热问题，从最初的风冷，到现在的液冷，各家企业都在尝试各种方法为数据中心散热：微软曾尝试部署海底数据中心，Facebook数据中心选址在北极圈附近，阿里云千岛湖数据中心使用深层湖水制冷等，我国很多企业则把数据中心布局在水电资源丰富的西南地区。“东数西算”的原则也是对数据与算力进行合理匹配，将大量的算力资源部署到西部地区，以降低能耗。

再次，在新能源层面，建立多样化的能源利用体系。集中力量开发太阳能、风能等可再生能源，扩大应用并加强储能技术研发，确保稳定供电。鼓励企业探索建设分布式光伏发电、燃气分布式供能等配套系统，引导新型数据中心向新能源发电侧建设，就地消纳新能源，推动新型数据中心高效利用清洁能源和可再生能源、优化用能结构，助力信息通信行业实现碳达峰碳中和目标。同时，探索核聚变等前沿技术作为长远清洁能源。

最后，在政策层面，制定严格的节能政策与标准，对AI系统的能效实施有效监管，以激励企业采用高效低能耗的计算技术与硬件设备；积极推广绿色数据中心建设，倡导采用高效冷却、能源回收技术，并鼓励数据中心采用可再生能源供电；建立健全电力消耗监管机制，实时监控并数据分析AI系统的能耗情况，提高能源管理水平，减少不必要的浪费；加强国际合作与交流，共同应对AI技术带来的能源挑战，分享成功经验、技术与资源，推动全球范围内节能标准的制定与实施，以实现能源利用的优化与升级，减少对传统电力的依赖，全面促进人工智能产业的绿色低碳发展。

