



我国工业软件的突围与重构

□ 高 亮



工业软件作为现代工业体系的“大脑”与“神经”，其自主可控性直接关系到我国制造业的高端化发展和产业安全。当下，我国工业软件发展面临诸多挑战，部分高校及企业因美国出口管制政策遭遇软件MATLAB禁用困境，华为海思也因电子设计自动化（EDA）软件断供致芯片设计停滞，暴露出我国95%的研发设计类工业软件依赖进口、高端CAD/CAE/CAM市场超90%被欧美巨头垄断的严峻现实。党的二十届三中全会强调，抓紧打造自主可控的产业链供应链，健全强化基础软件、工业软件等重点产业链发展体制机制，全链条推进技术攻关、成果应用，为我国工业软件的技术突围提供了方向指引。在此背景下，以国家智能设计与数控技术创新中心为代表的产学研力量，正通过技术攻坚、产业化转化与生态构建的系统性探索，推动我国工业软件从技术突围迈向生态重构，助力我国从制造业大国向制造业强国迈进。

核心技术攻坚：国产工业软件突围的第一战场

工业软件的技术突围，需在系统建模、结构设计、全生命周期管理等关键领域，从底层技术到应用场景实现全面突破，集中力量攻克“卡脖子”关键核心技术。

在系统建模仿真领域，国内团队自主研发的科学计算与系统建模仿真平台MWorks对标MATLAB，突破多领域统一建模与信息物理融合仿真技术，成为全球第四个具备完整自主知识产权的系统建模仿真平台。该技术已深度服务于嫦娥工程、空间站、C919大飞机等国家重大工程，实现从“依赖进口”到“自主可控”的跨越。在结构设计与优化技术方面，基于共形几何的复杂曲面设计与拓扑优化技术取得突破，研发的优化引擎使复杂结构设计效率提升40%以上，为运载火箭尾段、船舶推进器叶片等关键部件实现15%以上的轻量化。这一进展打破了AltairInspire/Hyperworks、Optistruct等国外软件在高端结构设计中的垄断局面。

在产品全生命周期管理（PLM）领域，对标美国PTC的Windchill、达索3DE、西门子Teamcenter，研发IntePLM产品全生命周期管理软件，攻克平台分层解耦、微服务架构、基于模型的系统工程正向产品研发过程支持、供应链跨域数字化研发协同和制造协同、基于几何参数的模型相似性搜索等关键核心技术。IntePLM在华为全面实现对美国Windchill的国产化替代，全面融入华为的工业软件生态，并进一步完成了30多家半导体和电子制造行业客户的试点应用，这标志着国内在工业软件生态融合领域迈出关键一步。

（上接1版）在推进马克思主义中国化时代化的历史进程中，我们党始终坚持解放思想、实事求是、与时俱进、求真务实，取得了毛泽东思想、邓小平理论、“三个代表”重要思想、科学发展观、习近平新时代中国特色社会主义思想等重大理论成果，为党和人民事业发展提供了科学理论指导。理论联系实际，让中国共产党在解决实际问题上不断提高执政本领、赢得人民拥护和支持，指引百年大党始终保持清醒、不断走向成熟。

推动实现民族独立、人民解放、国家富强。习主席强调，学习的目的，就是更好认识国情，更好认识党和国家事业未来发展大势，更好认识历史发展规律，更加能动地推进各项工作。在前进的道路上，中国共产党始终注重建设马克思主义学习型政党，牢固树立辩证唯物主义和历史唯物主义的世界观、方法论，坚持在实践上大胆探索、在理论上不断突破，独立自主地解决中国革命具体问题，不断适应社会生产力发展调整生产关系，不断适应经济基础发展完善上层建筑，在中华大地上全面建成了小康社会，顺利实现了第一个百年奋斗目标。历史证明，理论联系实际是马克思主义的基本原则，是我们党在革命、建设和改革中形成的优良作风，是我们党永葆生机活力、创造历史伟业的重要法宝。

推动新时代人民军队全面重塑、

高端数控系统的突破更具标志性。最新发布的华中10型智能数控系统，深度融合DeepSeek行业大模型与神经网络算法，成为全球首款具备自主学习能力的智能数控系统，突破多轴联动、高速高精、现场总线、开放式平台及基于指令域大数据的智能化等核心技术，实现对西门子、发那科、海德汉等国际巨头从“跟跑”到“领跑”的战略转折。

在生产系统仿真领域，FactorySimulation对标西门子PlantSimulation，攻克基于离散事件系统与智能体理论的多层次混合建模、高性能仿真引擎等关键技术，并在仿真性能、开放性及优化能力上实现全面超越，成功应用于中车、航天科工、航天科技、中国电科、中国船舶等10余家央企及大型民企。

在生产系统优化领域，对标日本Asprova开发的智能生产调度引擎，通过算法组件化与低代码化降低应用门槛，其核心调度技术已赋能数十家核心科技企业。

这些技术突破印证了一条核心路径：唯有以“从0到1”的底层技术突破构建自主可控的工业软件体系，才能打破国外垄断格局，为中国制造高端化提供核心支撑。

产业化转化：破局“实验室到市场”的生态壁垒

工业软件的价值最终体现在产业端的深度应用与价值创造，其产业化进程关乎技术创新的生死存亡。当前国产工业软件面临“实验室成果丰硕、工程场景落地艰难”的严峻挑战，其产业化瓶颈不仅在于技术成熟度，更源于应用生态的结构性锁闭。跨国巨头通过数十年深耕，已构建“产品、服务、标准”三位一体的“护城河”，致使制造企业在替换国产软件时面临数据迁移成本高、工艺适配风险大、技术支援体系弱三重阻碍。

破解这一困局，需以重大工程牵引技术迭代。如MWorks在空间站、载人登月等国家重大工程中，首次采用自主研发的工业软件，不仅验证了软件在极端工况下的可靠性，更通过工程实践积累了宝贵的应用数据。这种“以重大工程牵引技术迭代”的模式，已成为国产工业软件实现产业化转化的关键路径。通过参与国家重大工程，企业不仅能够获得宝贵的应用数据，还能在工程实践中不断优化产品性能，提升技术实力。同时，国家也可以通过政策引导，鼓励企业参与重大工程，为国产工业软件提供更多的应用机会。这种“以重大工程牵引技术迭代”的模式，已成为国产工业软件实现产业化转化的关键路径。

理论联系实际是打好建军百年攻坚战的重要保证

如期实现建军一百年奋斗目标，是党的意志、人民的期盼，是一项具有很强开拓性的事业，更加需要我們弘扬理论联系实际的马克思主义优良学风，善于用马克思主义之“矢”射新时代强军事业之“的”，不断提高分析问题、解决问题的能力，使各项工作更加符合客观实际、科学规律的要求。

现自顶向下的建模要求下发与自底向上的模型数字交付，纵向打通型号跨专业、跨阶段协同研发，纵向打通型号跨层级模型融合，并探索实践形成“建评管用”数字化工程研制标准化流程与新模式。这种“以应用促创新”的路径，有效加速了技术成熟度提升。

专利运营与市场对接机制同样关键。以“参数化水平集法结构拓扑优化技术”为例，8项发明专利构成完整技术体系，通过市场化协商实现技术价值提升，既保障国有资产规范管理，又实现了科研团队与企业需求的精准对接。这种模式为科研成果转化提供了可复制的经验。

生态重构：从单点突破到“开源—标准—协作”的体系竞争

工业软件的竞争本质是生态体系的竞争，单一技术突破难以撼动根深蒂固的全球产业格局。我国亟须构建“开源筑基、标准立信、全球协作”的生态网络，重塑产业规则话语权。

开源生态成为技术迭代加速器。MWorks开放系统内核模块，联合华为等企业构建开发者社区，并在300多所高校推广应用，发布60多本专业教材，将自主软件融入人才培养、创新实践体系。这种“技术开源+人才培育”的双轮驱动，打破了国外软件对开发生态的垄断。

标准体系建设是掌握话语权的核心。远程运维系统对标国际领先标准，建立机床全生命周期健康保障体系，主持参与制定6项行业标准，推动机床后服务数字化转型。标准的自主化，为国产工业软件在全球产业体系中争取规则制定权。

开放平台架构激活产业智慧。新型智能数控系统集成高性能控制平台、端一边一云协同架构、二次开发平台与智能APP平台于一体，为数控机床融合大数据与AI技术提供了算力与架构支撑，形成“基础系统+工业知识插件”生态，显著提升了数控机床的数字化、网络化和智能化水平，为智能制造装备的规模化应用奠定了平台基

础。这种平台化思维，正推动工业软件从单一产品竞争转向生态体系竞争。

未来之路：从自立到自强的战略跃迁

技术攻坚与生态构建虽能推进工业软件发展，但要真正突破认知和制度的深层障碍，还需开展全方位的思维与体系改革。目前，我国工业软件虽在技术和生态方面取得一定突破，但仍面临诸多深层挑战，例如，高端人才短缺、跨单位协同效率低下、科研成果快速工程化与产品化能力不足以及基础原始创新不够等问题。究其根源，往往不是技术本身，而是长期形成的思维定式，如对国外技术的依赖惯性及对国产软件成熟度的固有疑虑。同时，现有制度体系也存在束缚，包括成果转化机制不完善、跨组织协同壁垒较高以及人才评价激励机制不合理等问题。因此，工业软件的自立自强绝非单一机构能够完成，而是需要政产学研用协同推进的系统工程。

从MWorks支撑国家重大工程，到IntePLM融入华为生态，再到智能数控系统实现全球领跑，这些实践进展印证了一条清晰路径：以“有组织科研”模式集中突破核心技术，以重大工程与市场需求牵引产业化，以开源生态与标准体系构建产业话语权。当前，我国工业软件已从“跟跑追赶”进入“并跑竞争”阶段，但要实现真正的“自强引领”，仍需在基础研究、人才培养、制度创新等层面持续发力。这不仅技术的突围，更是产业生态与发展范式的重构。唯有如此，才能让工业软件这颗“工业大脑”摆脱外部束缚，为中国制造的高质量发展注入核心动力。

（作者系华中科技大学党委常委、副校长）



聚焦实践要求，提高打赢本领。“为学之实，固在践履。”一切学习都不是为学而学，学习的目的在于应用。虽然我军在不同时期担负的具体任务不同，但作为战斗队的根本职能始终没有改变。不断提升打赢能力，是新时代革命军人最直接的使命担当，也是弘扬理论联系实际学风最直接的实践担当。必须扭住主责主业，集中精力研究军事、研究战争、研究打仗，提高战略素养、联合素养、指挥素养、科技素养，把战斗力这个唯一的根本的标准立起来落下去，带头在重大军事斗争实践和军事演训活动中磨砺本领，确保在党和人民需要的时候拉得出、上得去、打得赢。

发扬整风精神，永葆纯洁光荣。学风影响党风。好学风才能形成好作风，伪学风就会导致伪忠诚。必须发扬彻底的自我革命精神，引导官兵常态检视自身思想作风和精神状态，牢固树立正确的世界观、人生观、价值观和权力观、政绩观、事业观，使自己的思维方式和精神世界更好适应事业发展需要。必须深入贯彻依法治军要求，不断提高一体推进“三不腐”能力和水平，重拳整治发生在官兵身边的“微腐败”和不正之风，以踏石留印、抓铁有痕的韧劲狠劲把作风建设引向深入，为官兵干事创业营造风清气正的良好环境，努力把强军事业推向前进。

（作者系武警黑龙江总队少将政治委员）

人工智能语料库是大模型训练的基础性资源，其规模和质量对模型的性能起决定性作用，是提高模型准确性和泛化能力的核心驱动力。目前，国际主流大模型训练语料以英文为主，占比超过90%。近年来，国内10亿参数规模以上的大模型数量以及参研单位持续增加，人工智能中文语料库供给量难以满足快速增长的大模型训练需求，部分模型不得不使用外文语料进行训练。然而，外文语料通常以西方价值观为主导，可能包含种族歧视、文化和意识形态偏见等问题。因此，加快建设大规模、高质量人工智能中文语料库势在必行，这是推动我国人工智能领域健康发展、提升国际竞争力的重要支撑。

建设现状

人工智能中文语料库是以中文文本为对象，经系统收集、整理、标注后形成能够用于模型训练的结构化数据集。它既包含新闻资讯、学术论文等正式文本，也涵盖社交媒体对话、网络评论等非正式语言素材，是帮助模型理解中文语义、语法规则、语义逻辑和句式结构的核心要素。近年来，我国积极推进人工智能中文语料库建设，并取得显著进展。

一是通用语料库和专业领域语料库规模持续扩大。2025年1月，在中央网信办指导下，集成27个数据集、总量达2.7TB的中文互联网语料资源平台正式发布，涵盖基础语料、专业语料和权威出版物样本。中译语通发布的西部AI语料库包含100多种语言，多语言多模态高质量数据达到PB级，内容覆盖农业、科技、金融、工业等领域。上海交通大学研究团队创建涵盖6种语言，包含255亿个token的医学语料库，以提高医疗诊断模型的准确性。

二是技术创新促进人工智能中文语料质量提升。智源研究院发布的中文互联网语料库CCI3.0，从语法、句法、教育程度等维度对原始数据进行分类和标记。DeepSeek-LLM（V1）通过数据去重、过滤和混洗构建了包含约2万亿token的中英双语预训练数据集，DeepSeek-V3通过提高数学和编程样本的比例进一步优化预训练语料库，构建包含14.8万亿token的多语言数据集。

三是人工智能中文语料生态建设持续完善。上海人工智能实验室牵头成立全国首个大模型语料数据联盟。深圳、上海数据交易所支持开设语料交易专区，促进人工智能语料共享和交易。上海启动开源创新生态建设和语料普惠计划，为中小企业提供低成本语料服务，促进跨领域合作与知识技术的共享。此外，《语料库建设导则》《具身智能语料库建设导则》等团体标准已正式实施。

三大难题

尽管国内诸多行业企业、研究机构均已开始布局人工智能中文语料库建设并发布高质量数据集，但相较于Common Crawl等国际领先水平的语料库，人工智能中文语料库在规模、质量和领域覆盖度上仍落后于英文语料库。

人工智能中文语料库收集获取难度高。一是中文语料历史积累薄弱。互联网早期以英文为主导，中文互联网起步较晚，百科知识、专业语料、学术资源等优质中文内容沉淀不足。二是中文语料的数字化程度偏低。中文使用人口庞大，文化典籍丰富，但系统性、机器可读的高质量数字化建设难度高，许多中文语料资源仍以纸质或非结构化形式存在，无法直接用于人工智能模型训练。国家图书馆3700万册藏书中完成深度数字化处理并发布在中华古籍资源库中的古籍数字资源仅有13.9万部。三是中文语料分布分散。中文互联网内容分散于不同社交媒体、新闻网站和论坛平台，高质量中文语料则往往集中在政府机构、学术团体、出版社和头部企业中，出于版权保护、隐私安全等因素考虑，这些中文语料往往被封闭管理，形成数据壁垒和数据孤岛。

人工智能中文语料库建设规范性不足。一是中文语料来源广泛。网络文本中充斥着大量非正式、低质量甚至错误的内容，导致训练数据真实性、准确性和逻辑性难以保障。二是缺乏统一的标注标准和清洗规范。不同机构在构建语料库时采用的分词方式、句法结构、语义标注体系各不相同，导致不同语料库之间存在结构差异，难以兼容互通。三是专业领域语料稀缺且标注成本高昂。法律、医学、科技等领域通常依赖人工标注且

人工智能中文语料库 加快建设

□ 谢慧慧 赵燕

对标注者专业性要求高。语料标注也易受标注者主观判断影响，而不同标注者的判断标准差异可能影响数据标注的一致性和准确性。此外，随着生成式人工智能广泛应用，其生成内容反向污染训练语料库的现象日益凸显，“幻觉”信息也将进一步降低语料的可信度。

人工智能中文语料库建设机制尚不完善。一是缺乏统一的语料共享与协同机制。语料通常被企业视为核心数据资产和竞争壁垒。由于缺少合理的利益分配机制和版权归属界定，企业出于担心数据安全法、个人信息保护法等相关法规，但在语料方面的具体实施细则仍不明确，法律边界尚不清晰，导致企业在获取和使用语料时存在顾虑。三是尚未形成成熟的语料交易与授权模式。缺乏以市场为导向的语料流通机制和公共服务平台，难以发挥语料资源价值，进一步降低了企业积极性。

破解路径

面对困局，必须打通中文语料“收集—处理—共享”的全链条堵点。

收集层面。在国家层面统筹协调，牵头设立国家级语料库联盟，联合国家图书馆、档案馆、科研机构、重点高校、核心出版社等单位，系统性推进中华典籍、学术文献、政府公开信息、专利数据等重要资源的权威数字化与结构化处理。基于统一的数据标准和规范体系，实现语料的分级分类管理，构建覆盖广泛、结构清晰、质量上乘的国家基础语料资源池，为人工智能发展提供坚实支撑。

处理层面。加大对中文自然语言处理工具的研发投入。研发更精准的中文分词算法，提高分词准确率，利用深度学习技术实现自动去噪去重和标准化处理，降低人工成本。建议由行业协会牵头，联合各方专家制定涵盖语料收集、清洗、标注、存储等环节的统一标准，确保不同来源的语料在处理上具备结构一致性。建立数据质量评估体系，定期对入库语料进行质量检测和修改校验。

共享层面。制定明确的公共数据开放目录与实施细则，加快推动高质量公共数据的开放共享，在保障国家安全和个人隐私的前提下，优先释放科研论文、统计报告、政策法规、历史文化资源等非敏感内容。鼓励科研机构、高校、企业在合规前提下，开放共享高质量中文预训练语料、微调数据集或基准测试集。探索建立安全可控的跨境数据流动“白名单”，充分利用“一带一路”等平台，推动多语言语料资源的流通合作，提升中文在全球数字空间的话语权。

（作者单位：中国电子信息产业发展研究院）

