

中兴通讯技术

简讯

ZTE TECHNOLOGIES | 第29卷 第6期 · 2025年6月

视点

04 通信大模型，引领网络智能化的创新力量

08 AI应用范式变革：Agentic AI



专题：星云通信大模型

13 星云通信大模型：聚合数据与知识，驱动通信网络迈向高阶自智





1996年创办

第29卷 总第441期

2025年6月 第6期 (月度出版)

中兴通讯技术 (简讯)

ZHONGXING TONGXUN JISHU (JIANXUN)

《中兴通讯技术 (简讯)》顾问委员会

主任: 刘健

副主任: 方晖 孙方平 俞义方 张万春

顾问: 柏钢 董伟杰 胡俊劼 华新海

阚杰 李伟正 刘明明 陆平

唐雪 王全 郑鹏

《中兴通讯技术 (简讯)》编辑委员会

主任: 林晓东

副主任: 卢丹

编委: 邓志峰 代岩斌 黄新明 姜永湖

孔建华 卢丹 梁大鹏 刘爽

林晓东 马小松 施军 余方宏

杨兆江

《中兴通讯技术 (简讯)》编辑部

总编: 林晓东

常务副总编: 卢丹

编辑部主任: 刘杨

执行主编: 方丽

发行: 王萍萍

主管: 中兴通讯股份有限公司

主办: 中兴通讯技术杂志社

出版: 《中兴通讯技术 (简讯)》编辑部

编辑部地址: 深圳市科技南路55号中兴通讯研发大楼

发行部地址: 合肥市金寨路329号国轩凯旋大厦12楼

发行部电话: 0551-65533356

<https://www.zte.com.cn/china/about/magazine>

发行范围: 国内业务相关单位

印数: 5000本

设计: 深圳市奥尔美广告有限公司

印刷: 深圳市旺盈彩盒纸品有限公司

印刷日期: 2025年6月15日



张万春

中兴通讯高级副总裁

星云通信大模型开启高阶自智网络新篇章

新一轮科技革命与产业变革正以指数级速度重构全球创新版图,以大模型为代表的人工智能技术浪潮,正以前所未有的广度与深度推动社会迈向AI平权时代。在通信领域,大模型也成为推动运营商网络进入高阶自智网络时代的核心技术驱动力。

中兴通讯持续推动AI技术在运营商运维场景的落地,实现AI范式的变革。通过在大模型、智能体和数字孪生等方向上的技术创新,中兴通讯帮助运营商打通流程堵点断点,提升整体网络运维自动化和智能化水平,推动网络运维向“Agentic Operation”演进。

依托自研的星云通信大模型,中兴通讯一方面运用通信领域的专有知识进行模型预训练和强化学习,并引入知识图谱、结构化数据模型和数字孪生等技术,有效减少大模型幻觉。同时,借助知识飞轮实现在线迭代和场景泛化,从而弥补了通用大模型在通信领域应用的鸿沟,使其逐步成为运营商的新质生产力。

今年初,中兴通讯全新升级了AIR Net高阶自智网络解决方案,基于星云通信大模型、大数据和数字孪生三大引擎,对基础设施、单域和跨域进行智能化改造,通过开放解耦,助力运营商运维完成从碎片化AI向端到端智能化升级。在此基础上,我们围绕运营商的高价值场景,通过构建价值评估体系与案例,达成提质创收、降本增效的商业成效,最终实现商业闭环。

未来,中兴通讯将立足通信大模型,持续深化与全球产业伙伴的合作,共同推动大模型等技术创新,探索新价值,建设开放共赢的新生态,为推动运营商网络数字化和智能化转型注入新动能。

目次

中兴通讯技术（简讯）2025年第6期



视点

04 通信大模型，引领网络智能化的创新力量
康红辉，刘昆麟

08 AI应用范式变革：Agentic AI
邵艳琴，杜永生

专题：星云通信大模型

13 星云通信大模型：聚合数据与知识，驱动通信网络迈向高阶自智
郑鹏，王成春，范雪峰

18 智能体工厂实践：工业级Agent的敏捷生产之路
巫江涛，付林舟

21 智能体和多智能体协同，推动网络实现L4高阶自智新高度
张文栓，冯媛

23 结构化信令大模型：让故障定位效率提升80%的协议级智能革命
陈燕，陈湘宁

25 看网智能体，实现AI内生的多维网络洞察和解决方案生成
赵欣，尹建华，严海波

27 AI大模型网络变更闭环智能体：运维模式革新
欧雪刚

30 通信大模型运维智能体，引领网络保障新趋势
沈远，沈毅

32 核心网运维智能体助力L4高阶自智持续提效
何伟

34 家庭知识库及智能体应用探讨
王猛

成功故事

36 无线大模型多智能体保障方案，为运营商春节保障“护航”
王洪新，王康，曹莉笠

38 山东移动携手中兴通讯：多智能体协同，使能网络综合故障监控处理
周波，宋培双

02 新闻资讯

中兴通讯发布全球首个50G-PON三代时分共存方案

5月16日，中兴通讯在“下一代光网络技术策源与发展论坛”上发布了全球首个50G-PON三代时分共存方案。在工业和信息化部全面推进万兆光网试点的战略背景下，此创新方案不仅填补了我国在超高速光通信领域的技术空白，更标志着我国在全球万兆光网核心技术领域实现了从跟随到引领的跨越，为全球光纤网络的演进提供了极具创新性的“中国方案”。

中兴通讯50G-PON三代时分共存方案，攻克了ITU 50G-PON与IEEE EPON制式兼容的行业性难题，采用世界首创技术，实现对称速率50G-PON、10G-EPON、EPON三代系统在同一光分配网络（ODN）中协同工作。

中兴通讯携手湖北移动、武汉大学打造全国首个50G-PON+Wi-Fi 7万兆全光校园示范网

5月，中兴通讯联合湖北移动、武汉大学，采用50G-PON+Wi-Fi 7创新网络技术，试点建设全国首个万兆全光校园示范网，推动万兆光网与高校场景深度融合，以超高速率、超低时延、全场景覆盖的技术特性全面赋能学校教学、科研与生活。

根据武汉大学统一规划，试点工作主要集中在水资源工程与调度全国

SuperCLUE 5月榜单发布：中兴星云大模型荣获推理榜总分第一

近日，中文大模型权威测评基准SuperCLUE发布《中文大模型基准测评2025年5月报告》。报告显示，中兴通讯自主研发的星云大模型NebulaCoder-V6在推理专项榜单中斩获榜单金牌，总分并列第一。同时在综合总榜中斩获银牌（并列第二），彰显了中兴通讯在AI核心赛道的强劲创新实力。

SuperCLUE的“推理”专项是检验大模型逻辑思维、分析解决问题能力的“硬核”考场，涵盖数学推理、科学推理、代码生成等多个维度。NebulaCoder-V6在SuperCLUE推理榜单中以总分67.4分获得第一，细分表现同样出色：

- 科学推理得分53.19分，全榜单排名第一，展示了其在复杂科学概念理解与推演方面的强劲实力。
 - 数学推理得分62.39分，排名第三，展现了在复杂科学概念理解与推演中的竞争力。
 - 代码生成得分86.59分，稳居第一梯队，工程实践能力扎实可靠。
- 星云大模型不仅“智商超群”，更通过了严格的多项安全“大考”，相继荣获中国软件评测中心与中国计算机行业协会数据安全专业委员会联合颁发的“大模型产品安全性检验证书”和CCRC认证企业实施测评的“信息安全风险评估认证证书”等。

中兴通讯联合绍兴电信率先完成基于AI识别的5G-A低空多元感知融合试点

5月，中兴通讯联合绍兴电信、中电无人科技、华航信、安邦护卫、君立华域等产业伙伴开展以5G-A为核心，结合TDOA、AOA、RID、视觉光电、AI等多种技术的低空多元感知融合试点，打造高性能、高可靠、一体化的低空智联网。本次试点基于AI识别技术，集成5G-A、雷达、摄像头、频谱侦测等多种数据，实现对低空环境的高精度、全天候、多维度感知与分析。

重点实验室、遥感信息工程学院、超算中心、工学部第一教学楼、工学部1食堂和宿舍楼等区域，中兴通讯联合湖北移动部署了三代全速率50G-PON+Wi-Fi 7万兆全光解决方案，构建了覆盖科研、教学、食堂、宿舍等场景的万兆全光校园网络体系，实现了“50G超万兆到学部、万兆到房间，千兆到终端”的网络能力提升，为武汉大学科研大模型极速入算、AR/VR教学、远程教育、校园生活提供充足的带宽保障，让“数智武大”信息化建设迈向新高度。



中兴通讯第九届“创兴日”：向智而行，共绘数智未来新蓝图

5月27日，为期3天的中兴通讯第九届“创兴日”在深圳召开。本次活动以“智行，向未来”为主题，汇聚来自政府、运营商、行业伙伴、高校机构及联盟协会的多位知名专家学者，共同探讨前沿趋势与创新机遇，围绕低空经济、智算、AIGC、ESG可持续发展、产业数字化等热点主题展开深度交流，共绘数智时代未来新蓝图。

中兴通讯高级副总裁、首席战略官王翔在开幕式致辞中表示，当前我们正身处高速变化的数智时代，中兴通讯以“连接+算力”为支点，以智算为战略主航道，向“成为网络连接和智能算力的领导者”的战略目标坚定迈进。

中兴通讯与印尼XLSMART签署战略合作协议，携手打造面向未来的数字基础设施

5月26日，中兴通讯与印度尼西亚电信公司XLSMART在雅加达举行战略合作协议签约仪式（XLSMART由印尼XL Axiata和Smartfren Telecom合并而成）。此次双方战略携手旨在加速印尼数字连接能力升级，联合构建面向未来的高效、智能化网络基础设施。

Sinar Mas Telecommunications and Technology董事长兼首席执行官Franky Oesman Widjaja、XLSMART首席网络整合官I Gede Darmayusa、XLSMART首席采购官Sanjay Sachdeva、中兴通讯董事长方榕、高级副总裁肖明、高级副总裁梅中华等人出席签约仪式。

珠海移动联合中兴通讯率先落地基于AI通感模型的低空多模态融合创新成果

5月，中国移动通信集团广东有限公司珠海分公司（珠海移动）联合中兴通讯，依托5G-A通感一体多源感知融合方案，借助AI大数据学习、训练与处理能力，实现对飞鸟、无人机多类型目标的精确识别，虚警率达业界最低，大幅提升低空经济安全监管效率与响应速度，筑牢低空经济发展安全根基。

中兴通讯携手31家企业启动低空产业生态平台 共建低空经济联合体

5月27日，中兴通讯联合深圳市低空经济产业协会共同举办低空经济产业发展论坛，并正式启动“中兴通讯低空产业生态平台”。中兴通讯高级副总裁、首席战略官王翔和中国民航局原副局长李健分别致辞，亿航智能副总裁张宏、中国信通院标准所无线信息化部副主任杜加懂、北京交通大学教授纪寿文、深圳市低空经济产业协会执行秘书长成涛、中兴通讯无

线通感产品总经理赵志勇先后做了专题分享。

本次论坛正式启动中兴通讯低空产业生态平台，该平台致力于打造行业智库和项目联合体，通过整合产业链上中下游资源，构建“技术共创、场景共探、市场共赢”的协同生态体系。平台依托中兴通讯开放实验室环境、5G-A低空示范网络（覆盖全国100+示范区）以及其他产业伙伴的业务环境等生态资源，规划分阶段推进低空典型业务的联合验证、低空经济示范区建设、低空规模化市场项目拓展等。



康红辉

中兴通讯无线网络智能化首席架构师



刘昆麟

中兴通讯星云通信大模型研发总工

通信大模型， 引领网络智能化的创新力量



人工智能（AI）与新一代信息技术深度融合，已成为自智网络变革的核心驱动力。随着大模型技术在网络领域的爆发式应用，通信大模型正深刻变革网络运维、运营及业务创新能力，推动网络智能化重塑。

方兴未艾，大模型时代通信模型和应用蓬勃发展

自2023年起，以ChatGPT为代表的大语言模型（LLM）引发全球技术竞赛，2025年初，DeepSeek发布开源模型DeepSeek-R1，凭借强推理能力与低成本优势，迅速渗透至金融、医疗、智能制造等垂直领域。为应对通信产业中低空经济、物联网、车联网、沉浸式XR等新兴业务的涌现，以及通信运营商的数智化转型需求，

生成式人工智能（GenAI）技术在通信产业快速引入和发展，基于通信大模型的自智网络和智慧应用迎来爆发式发展。

通信大模型的快速发展一方面得益于运营商和设备商的积极投入。三大运营商依托自身数据优势，加速行业大模型布局：中国移动自主打造“九天”人工智能基座，沉淀450项AI能力；中国电信星辰大模型完成全模态算法备案；中国联通元景大模型2.0已服务35个行业。据IDC报告显示，2024年中国通信行业大模型市场规模同比增长67%，多模态与科学计算模型成为新增长点。设备商方面，中兴通讯通过“全栈自研+生态协同”双轨并行态势，其星云通信大模型覆盖7B/14B至100B参数版本，通过“星云大模型+Agent工厂+系列化应用”架构，实现网络运维、反诈监测、信令分析等场景落地。另一方面，业界开源生态会推动技术普惠也推动通信模

型的进步，DeepSeek的开源策略大幅降低行业准入门槛。截至2025年Q1，国内超60%的通信企业基于开源模型开发定制化解决方案，这种“开源共享+垂直深耕”的模式，正在重塑通信网络智能化技术创新路径。

全面赋能，通信大模型重塑自智网络演进之路

通信大模型在自智网络智能化技术底座、系统架构和面向Agentic AI的运维新范式几方面，全面赋能自智网络的高阶演进。

大模型重构自智网络智能引擎

自智网络通过“三层四闭环”架构实现网络自动化与智能化（见图1）。面向L4、L5级别的自智网络演进，TM Forum对高阶自智使能技术进行了展望，网络AI大模型、网络可信技术、网络数字孪生等技术将支持3GPP网络内生的AI需求，推动自智网络向更高层次的智能化发展。在自智网络向高阶自智演进中，通信大模型起到“智能引擎”作用，通过和数据引擎、数字孪生

引擎相互协同，构建自智网络的数智引擎技术底座。

大模型重构自智网络技术架构

通信大模型重构自智网络演进的技术架构，形成“全场景AI范式+数智引擎+智能体”的三层体系架构，分层引入系列化模型，支持网络AI内生和全场景AI化。如在单域闭环到跨域协同，基于大模型的智能分析能力，运营商可在单域维优营场景（如无线网络优化）中率先实现闭环，并逐步向跨域端到端场景扩展，实现网络“历史可溯、现实可视、未来可观”的全局管控。基于通信大模型形成的“感知-决策-执行”一体化能力，在意图理解和自主学习、长流程闭环优化、准确率提升上都有诸多成功实践：如规划领域，在5G-A网络建设中，大模型辅助的数字孪生工具可模拟超10万种信道组合，缩短规划周期；运维领域，中国移动基于DeepSeek的智能运维系统将故障定位效率大幅提升，向“分钟级故障响应”演进；运营领域，中兴通讯反诈大模型通过分析通信行为模式，识别新型电信诈骗，大幅降低误报率。

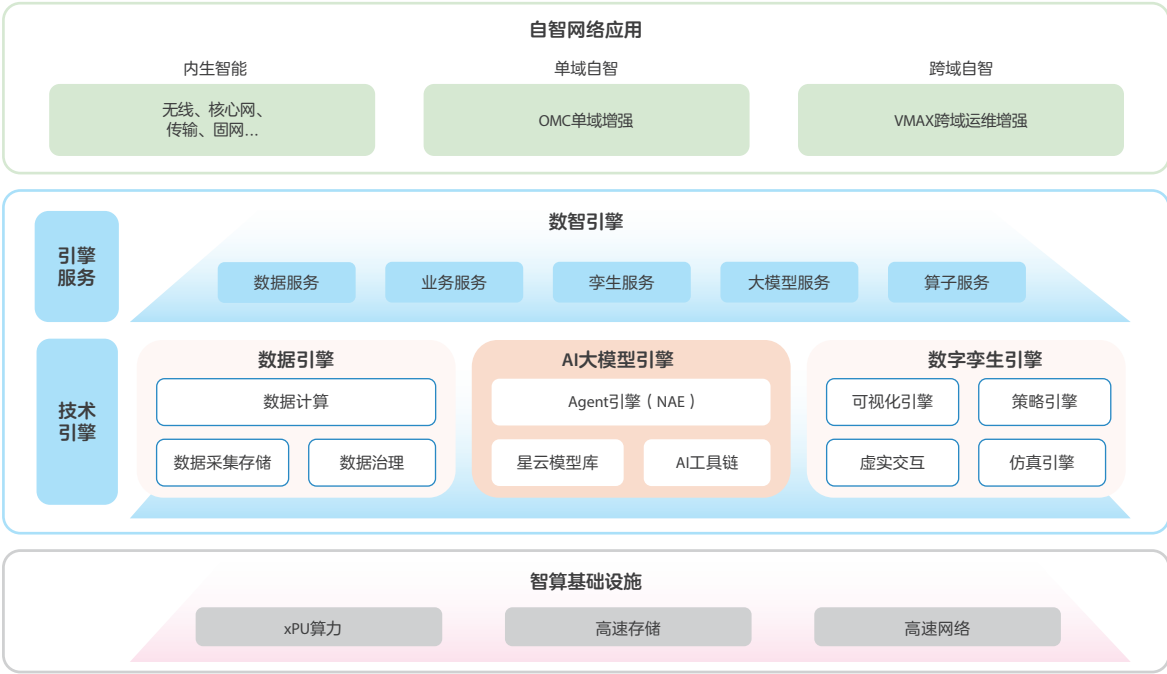


图1 自智网络技术架构



通信大模型催生新型网络智能体（network agent），传统自智网络运维方式向基于大模型的智能体Agent范式演进，凭借自主学习、任务执行及多任务协作能力，智能体为电信运营商的智能化转型提供了全新路径。

大模型重构自智网络Agent新范式

通信大模型催生新型网络智能体（network agent），传统自智网络运维方式向基于大模型的智能体Agent范式演进，凭借自主学习、任务执行及多任务协作能力，智能体为电信运营商的智能化转型提供了全新路径。通过部署智能体，运营商能够实现网络管理和优化的高度自动化，增强网络的自感知与自修复能力，显著降低运维成本并提升网络可靠性。

OpenAI O1和DeepSeek R1模型在复杂推理任务中表现卓越，开启了“慢思考”的大模型新范式，为通信网络的运维与运营智能体提供了全新解决方案。例如模型及其推理增强技术能够在复杂业务场景中实现任务编排、问题分析、灵活决策和优化执行，同时智能体应用在交互方式、接口形式、产品架构、业务能力、开发交付等方面，形成全场景AI新范式，智能体Agent范式可有效缩短实现自智网络L4的路径。

智领未来，通信大模型向网络世界模型发展

通信大模型的未来演进，一方面将向更强性能、更少资源占用、更多交互方式的模型发展，另一方面也会向模型和网络业务密切融合的网络世界模型方向发展，包括和数字孪生网络技术

的协同、网络智能中枢构建等技术发展方向。

通信大模型技术创新方向

通信大模型技术创新方向包括模型架构创新、多模态融合、领域增强强化学习等技术，也有模型和智能体的进一步融合演进。

● 模型架构创新

针对通信模型偏本地化部署场景，需要对模型结构做技术创新，达到更高性能、更小尺寸的通信大模型。在模型架构上，可以考虑在传统混合专家模型（MoE）架构基础上，引入时空分离式路由机制，把时间序列模型和空间数据结合形成网络时空大模型，然后时空模型和语言模型进一步融合。协同机制上，采用“中心大模型+边缘小模型”双向知识蒸馏机制，中心模型提炼全局特征（如全国网络流量模式），边缘模型捕获本地特性（如区域基站部署差异），通过联邦机制动态权重交换实现模型协同演进。

● 多模态融合

未来大模型将整合文本、语音、图像和网络信令、告警、日志等多模态数据，实现更精准的网络状态感知，为客户服务创造更具沉浸感的交互方式和个性化解决方案，提升用户体验。

● 轻量化与边缘部署

通过模型量化和蒸馏将大参数模型压缩，支持低功耗使大模型能力下沉至边缘节点，支持边



▲ 图2 网络世界模型

缘端低时延推理，满足工业互联网实时需求。

● 领域增强学习

结合通信领域知识图谱（如协议栈规则、故障案例库），提升大模型在垂直场景的专业性。

数字孪生与大模型的协同赋能演进

大模型和数字孪生技术协同演进是双向赋能。数字孪生为通信大模型提供虚实映射的试验场和合成数据，并支持孪生体驱动大模型训练；大模型则增强数字孪生的推演仿真能力，并赋予数字孪生动态优化能力。数字孪生体实时映射物理网络状态，大模型通过强化学习生成最优资源调度策略，可实现网络功能的“数字孪生-模型训练-网络部署”的闭环，推动网络从“被动监控”转向“主动设计”。

迈向统一的网络世界模型

6G时代，通信大模型将演化为“网络世界模型”，将物理网络、业务需求、用户行为统一表征与预测，形成网络世界模型，成为未来网络的智能中枢（见图2）。

其核心特征包括：

- **多域统一建模：**整合无线网、核心网、传输网及应用层数据，结合时空孪生模型形成全局建模。
- **因果推理能力：**从关联分析转向因果推断。
- **自主进化机制：**通过持续学习与反馈循环，实现模型能力的自我迭代，适应新型业务的动态需求。
- **孪生闭环：**未来的智能体网络将具备端到端场景孪生闭环的能力。

通信大模型正从技术工具进化为网络智能的核心引擎，重塑自智网络技术底座，推动其从单域优化走向全域协同，从人工干预迈向自主决策。随着6G网络智能化架构的演进，融合数字孪生、边缘推理和因果学习的网络世界模型，将推动自智网络实现从“功能自治”到“认知智能”的质变。这场由通信大模型驱动的智能革命，不仅重塑着网络基础设施的技术形态，更将重新定义自智网络的价值边界。

ZTE中兴



邵艳琴

中兴通讯通信领域大模型及RAN
智能体产品研发总工



杜永生

中兴通讯通信领域大模型及RAN
智能体产品技术总工

AI应用范式变革：Agentic AI

我们正处在AI能力飞速膨胀的时代。AI的发展本质上是技术范式不断迭代的过程，而每一次技术跃迁都会推动应用场景的革新。

当前，以大模型为驱动的AI Agent技术，带来了更深远的范式变革——从“工具化AI”向“Agentic AI”（具备自主性的智能体）的演进。“Agentic AI”具备“规划-感知-决策-执行”的全流程闭环，也将引发自智网络从“人工驱动”到“智能体主导”的范式转换。

AI范式跃迁：从“工具执行”到“自主执行”

AI Agent应用从对工具的编排执行逐步走向自主性的AI Agent系统。这类AI Agent应用能代表用户自主采取行动，承担复杂的任务、作出决

策、执行任务并适应不断变化的环境。

大模型发展重构Agent认知底座

大模型通过海量数据预训练，构建了通用化的“世界知识库”，能够理解复杂语境，处理跨领域任务，甚至展现初步的推理与创造能力。

- 通用智能底座：LLM（large language model）作为“大脑”，赋予Agent理解、推理和生成能力；
- 环境感知与行动：通过API、传感器、工具等接口，Agent可主动与环境交互；
- 目标驱动与进化：基于预设目标，Agent可规划路径、动态调整策略，并通过反馈持续优化。

这一范式下，AI不再是“工具”，而是能够自主决策、长期运行、动态进化的智能体。其运行逻辑是在一个任务目标循环中调用模型，根据

模型决策执行工具，并将结果反馈给模型，直至模型决定停止。

基于大模型的Agent能够自主拆解任务目标、调用工具链、动态调整策略，形成闭环决策系统。这种“目标驱动、自主规划”的模式，标志着AI应用从“功能模块化”向“智能体中心化”的范式转移。根据Gartner预测，到2026年，超过80%的企业将部署AI Agent以重构业务流程，效率提升幅度可达40%~60%。

AI Agent应用新范式催生AI新生产转变

当前的大模型厂商，如OpenAI、Anthropic、阿里、字节等，正从单一API输出转向构建Agent生态平台。微软推出的Copilot Studio允许企业定制专属Agent，集成内部数据与业务流程；字节推出的扣子及扣子空间，支持互联网智能体开发平台和企业应用。

在这一趋势下，随着AI Agent应用的火爆，AI应用的快速开发平台也成为了AI价值创造的枢纽。AI Agent的交互设计，端到端快速开发和发布、部署运行，以及AI生态的整合将是AI新生产的核心要素（见图1）。

Agentic AI进化新阶段：从“被动响应”到“主动进化”

AI正迈向Agentic新阶段，从被动执行指令转

向主动学习、适应和优化。通过自我迭代与环境交互，系统能动态调整策略，实现目标驱动的持续进化，标志着智能体从工具型向代理型转变的关键跃迁。

自主性跃升：从“被动响应”到“主动进化”

我们再来回顾一下Anthropic公司对智能体的定义：大语言模型（LLM）智能体能动态地自主指挥自己的行动和工具使用，并始终掌控完成任务的具体方式。

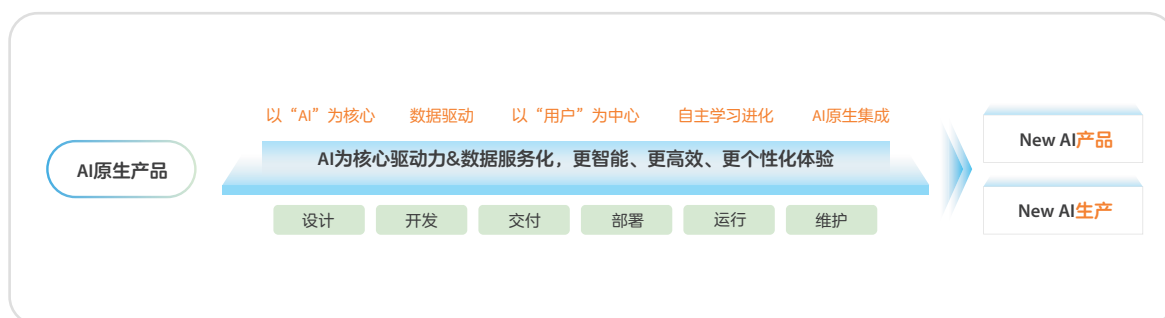
虽然当前Agentic AI仍处于过渡期，但2025年3月6日发布的Manus、3月13日夸克发布的“AI超级框”、字节推出的TARS Agent，以及Genspark产品的爆火，都标志智能体正在从“工具执行者”向“目标协作，主动进化”的超级智能体跃迁。

Agentic AI工作模式：多智能体目标协作式工作

要理解Agentic AI工作模式带来的革命性变化，我们将其与Workflow进行对比分析。

以搜索场景为例，这种变革体现得尤为显著——它不仅重构了技术实现路径，更颠覆了人机协作的底层逻辑。一个典型的Agentic AI的搜索过程可能是这样的：

- 用户提出问题，智能体分析并拆解问题，推测用户的真实意图；
- 如果问题模糊，智能体会主动向用户提问，



▲ 图1 AI原生产品生产定义

进一步确认；

- 智能选择搜索，可能选择进行一般性搜索，也可能根据情况选择特定的专业数据源；
- 智能整合搜索结果，并按照用户意图进行搜索输出。

针对这样的工作模式，模型会对搜索过程和搜索历史不断学习和优化，智能体能够自主判断搜索方向和有效路径。整个搜索过程中，智能体的每一步决策和推理都会留下清晰的内部记录，从而实现一定程度的可解释性。

从这个案例可以看出，相比传统基于规则或工作流串联的解决方案，Agentic AI实现了“静态Workflow”到“动态决策”的进化。如图2所示，整个工作由多Agent相互协同，完成“规划-决策-执行-评估”的全栈闭环。Agentic AI主要实现了3个维度的跃迁：

- 目标驱动的动态规划：摆脱预设流程的束缚，根据实时环境自主拆解任务，调整策略。

- 闭环自学习优化的能力：通过持续反馈学习改进决策路径，而非依赖人工调参。
- 资源整合的智能性：主动调用异构系统（如API/数据库/工具链），而非被动等待指令。

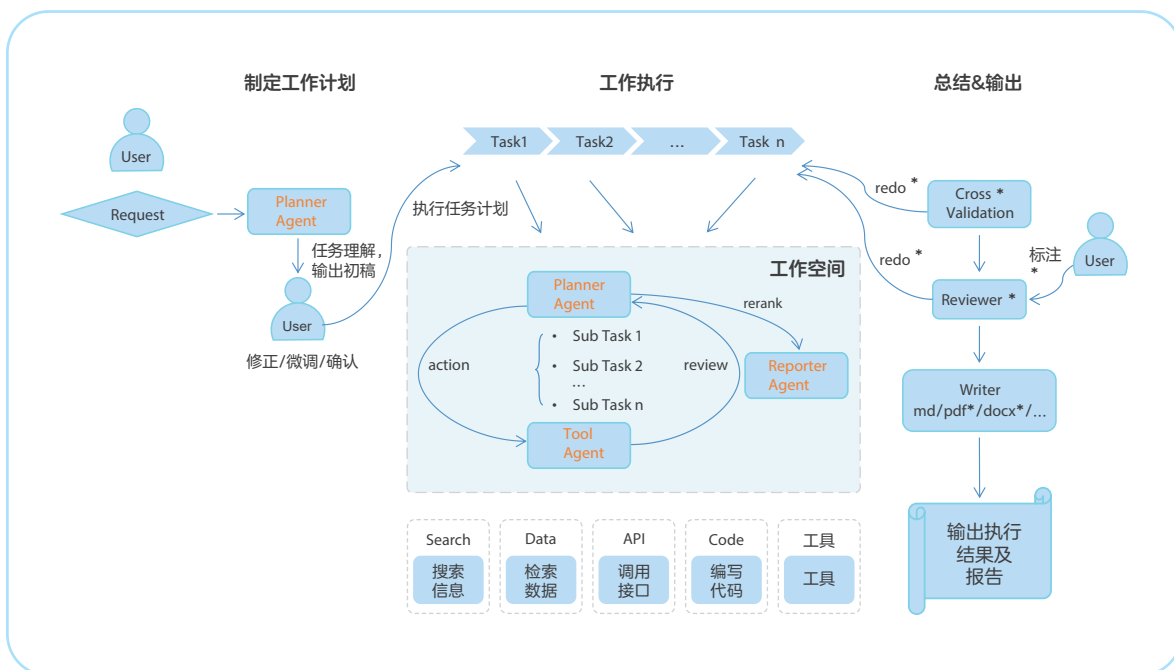
Agentic AI应用场景

我们将Agentic范式延伸至自智网络，变革同样显著。在无线接入网场景中，针对故障处理及网络优化这样两个实例，我们来说明通过数据共享、策略联动和任务接力，Agentic AI范式下的“感知-决策-执行-进化”闭环运维体系。

故障处理协同：工单自动闭环

场景为基站突发业务中断，系统自动处理流程如下：

- 故障感知：故障智能体实时监测到某5G基站的RRC连接成功率从99%骤降至72%，同时



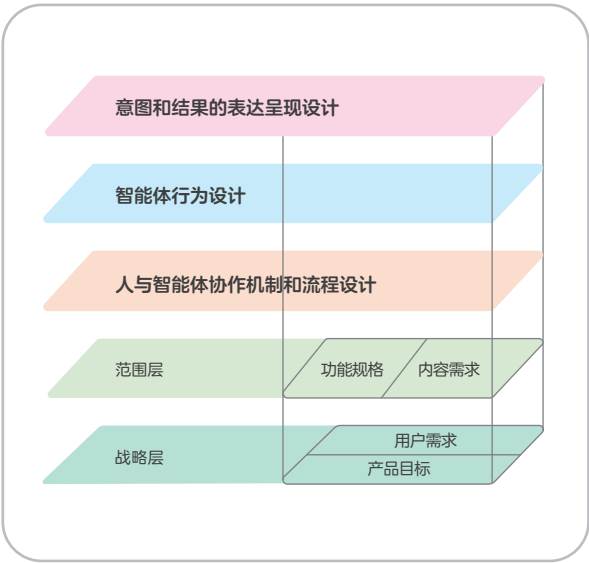
▲ 图2 Agentic AI通用工作模式定义

- SINR（信号干扰噪声比）均值下降6dB，触发3级告警；
- 根因诊断：调用知识图谱，匹配历史案例库，自动调用数据查看、根因诊断等API/命令/工具等，排除硬件故障（设备健康度指标正常），Agent主要提供了资源整合的智能性；
 - 模糊分析：根据基站-传输-动环-核心网关拓扑，定位到因邻区天线倾角偏差引发同频干扰，Agent主要提供了目标驱动的动态规划；
 - 自动修复：给出故障的修复方案及自动修复动作，针对对业务有影响的高危操作提醒用户确认执行方案及执行动作；
 - 执行动态参数调整：执行将故障基站的天线下倾角从8°修正为12°的命令，收缩覆盖范围；
 - 结果反馈：查看参数调整执行结果，并启动指标观测；
 - 验证指标：连接成功率回升至98%，SINR恢复至18dB；
 - 自动生成工单报告：推送至运维人员移动端，并标记“无需人工介入”；
 - 本次修复结果被计入到Agent的共享记忆中，Agent基于强化学习的方式从中提炼场景、规则、重大诊断和定位动作，用于下一次类似场景修复；Agent提供了闭环自学习优化的能力。

优化协同：预测性网络优化

场景为开展大型活动，网络受到流量冲击。系统处理流程如下：

- 联合预测（事件前24小时）：优化智能体预测场馆周边将出现10倍流量峰值，同时故障智能体预检该区域基站负载余量不足。
- 动态预配置（事件前2小时）：此处Agent主要提供了资源整合的智能性。优化智能体激活休眠态小区，扩容网络密度，同时调整QoS策略，优先保障视频带宽；故障智能体



▲ 图3 Agent设计要素

- 同步启动系统健康检查，对扩容基站进行压力测试，排除潜在硬件隐患。
- 网络优化（活动进行中）：当瞬时用户数突破预测值15%时，优化和故障智能体协同响应。优化智能体启动频谱分流流量；故障智能体监控基站CPU温度，动态限制非紧急业务以保证稳定性。此处Agent主要提供了目标驱动的动态规划能力。
 - 优化结果通知（活动结束后）：自动生成网络质量报告，通过短信/APP推送报告给用户：“峰值时段用户平均速率52Mbps，达标率97%”。

Agentic AI设计范式及关键技术

- Agentic AI产品设计逐步向用户自适应体验演进，设计逐步聚焦如图3所示的3个关键层次。
- 意图和结果的表达呈现设计：通过界面元素动态映射用户意图与系统响应，实现操作路径与反馈结果的可视化闭环表达。
 - 智能体行为设计：通过拟人化反馈与预期管理，体现AI智能体的可解释性、可控性及人

格化行为模式。

- 人与智能体协作机制和流程设计：聚焦人和AI双向适应性，通过动态角色分配、上下文感知、透明化决策路径和渐进式信任建立，设计可进化的协作框架，使智能体能主动预测需求、解释行为并接受干预，最终达成自然流畅的共同完成任务体验。

基于LLM的Agent核心关键技术

一个Agentic系统，越是由LLM决定其行为方式，就越具有“自主性”。围绕规划、记忆、工具与感知四大核心模块，我们来看看基于LLM的Agentic AI进化关键技术。

规划模块的关键技术包括：

- 分层任务推理：基于LLM进行全局规划，结合思维图谱和强化学习（RL），实现目标的多粒度拆解与动态路径优化。
- 因果推理增强：结合结构因果关系和经验数据，预判动作的长期影响，避免短视决策。
- 多Agent实时协作：采用多Agent实时协作、共享记忆、多Agent对抗等技术，实现多Agent任务分配与冲突消解。

记忆模块的关键技术包括：

- 长期记忆压缩：存储关键历史信息，结合自注意力机制实现高效检索与关联。
- 短期记忆管理：采用缓存保留对话状态与环境上下文，支持多轮交互连贯性。
- 知识蒸馏与更新：通过持续学习或检索增强（RAG），动态扩展外部知识库，避免模型幻觉。
- 记忆动态增强：构建动态记忆库存储交互轨迹，利用注意力机制提取关键经验，实现记忆动态增强。

工具模块的关键技术包括：

- 工具自动查找及使用：基于工具描述文本的语义嵌入匹配，自动构建工具调用方式。现实世界是开放、动态的，新的工具、API、

数据源层出不穷。对于一个像人类一样使用工具的智能系统，必须具备泛化能力，能够理解新工具的功能描述，并将其纳入自己的能力版图。

- 安全沙箱验证：在受限环境（如Docker容器）中预执行高风险操作（如网络请求），验证结果可靠性后再反馈至主流程。

Agents要基于环境、上下文、执行反馈等进行感知、判断和决策。感知模块的关键技术包括：

- 多模态信息融合：利用对齐跨模态表达空间，统一处理文本、图像、语音输入，构建环境状态表征。
- 环境动态建模：通过世界模型（world model）预测环境变化，辅助Agent预判动作影响。
- 主动感知与注意力控制：基于强化学习优化感知焦点，优先处理高价值信息（如对话中的偏好、习惯等）。

Agentic AI进化范式展望

AI Agent与模型正在相互促进、加速进化，能力边界不断突破，在特定领域甚至已接近或超过人类专家水平。面对智能技术“人间一日，智能一年”的快速进化模式，我们既要拿出大胆创新的勇气迎接变革，也要保持耐心让技术在实践中逐步验证。

技术层面，AI将从感知智能向自主决策跨越，多模态融合将推动智能体实现物理世界交互。应用层面，金融、医疗等领域率先落地，其他垂直领域也会加速向AI为中心转化。社会协作层面，人机协同将重构工作流，倒逼伦理规范与价值体系升级。设计模式将从“Reason+Tool”模式向“Learning+Reason”模式转变。

然而，真正的Agentic AI进化将伴随着人类认知方式、协作体系，乃至价值标准的全面升级，并将开启人类发展的新篇章。ZTE中兴

星云通信大模型： 聚合数据与知识， 驱动通信网络迈向高阶自智

中兴通讯 郑鹏，王成春，范雪峰





郑鹏
中兴通讯数据智能及服务
产品总经理



王成春
中兴通讯通信大模型规划
总工



范雪峰
中兴通讯大数据产品总工



前，通信网络正面临两大核心挑战：复杂度指数级攀升与智能化需求刚性增长。5G-A、6G等技术的加速商用让连接更泛在，驱动XR、工业互联网等新兴产业蓬勃发展，加速商业价值兑现；AI大模型等新技术与通信网络的深度融合，推动网络智能化向“智网”新范式演进。中兴通讯自智网络AIR Net高阶演进方案，以价值成效为牵引，通过开放解耦，构建面向多业务、多领域、全生命周期的全场景闭环自动化能力，通过运用大模型、智能体及数字孪生等技术，以商业成效、运维效率、生态价值为引领，加速网络迈向高阶自智。

- 商业成效变现：通信网络引入AI大模型等新生质生产力，拉通自智网络高价值场景端到端闭环，实现省人、省钱、省时的确定性收益，实现机器决策代替人工决策。
- 运维效率升级：运维模式从“人工经验驱动”转向“数据智能驱动”，故障处理效率提升数倍。
- 生态价值释放：多样化业务需求使得设备商、运营商在垂直行业深度协同，构建“网络即服务”新生态。

炼好大模型：构建灵活、动态、严谨的智能底座

在设计中兴通讯AIR Net自智网络技术架构时，核心目标是在网络自主运行能力提升与业务响应闭环效率优化两方面实现突破，消除传统人工干预导致的流程断点与效率瓶颈。当前运营商网络虽已实现基础自动化功能，如告警过滤、简单工单派发等，但其底层逻辑仍以预设规则为主、局部AI模型为辅，且面对跨域故障定界定位、投诉或质差分析等复杂场景时存在明显局限。要真正实现网络自智能力跃升，亟需构建具备深度网络认知能力的“智能中枢”——即基于大模型的智能决策引擎。这种将通用智能与领域知识深度结合的技术路径，已成为实现网络自动

化向智能化转型的关键范式。中兴通讯提出的星云通信大模型方案（见图1），其突破性在于：通过“基座灵活性、任务自主性、知识动态性、推理严谨性”能力升级，驱动网络从“人工规则预设”向“自主认知进化”跃迁。

全栈大模型体系：零感换擎实现“基座灵活性”

运营商网络环境的复杂性体现在：硬件异构、平台多元、场景长尾。星云通信大模型通过分层解耦架构，实现“零感换擎”，即基座模型的灵活替换无需改造业务逻辑。

依托自研星云通信大模型，中兴通讯通过高质量的专业通信语料训练，建设多智能体协同的体系，精准解决网络运维中的复杂问题。中兴通讯的大模型整体解决方案具备强大的开放解耦能力，能够灵活适配多种模型。在模型底座方面，不仅支持中兴星云通信大模型，还兼容业界优秀的开源模型，包括备受瞩目的DeepSeek模型。这一方案的核心优势在于“零感换擎”，用户可以在不同场景下无缝切换模型引擎，而无需担心性能或兼容性问题。

这种开放架构不仅规避了技术路线锁定的风险，更形成了“基座模型竞争择优”的良性生态。运营商可根据业务需求灵活选择基座模型，既可利用通用大模型的泛化能力，也可发挥垂直领域模型的精度优势，真正实现“最佳模型服务最佳场景”的技术愿景。

多智能体协同：目标导向的“任务自主性”

智能体是将大模型能力充分发挥出来的重要手段，通过不同智能体之间的相互协同可以自主快速解决通信网络的复杂问题。未来的网络运维运营面向大模型进行智能化改造，运维范式将从“人+机器”的模式演变为“机器+人”的模式。智能体之间通过LPI（language programming interface）的方式进行交互，改变了传统的API调用解决固定场景问题的方式，提升了泛化

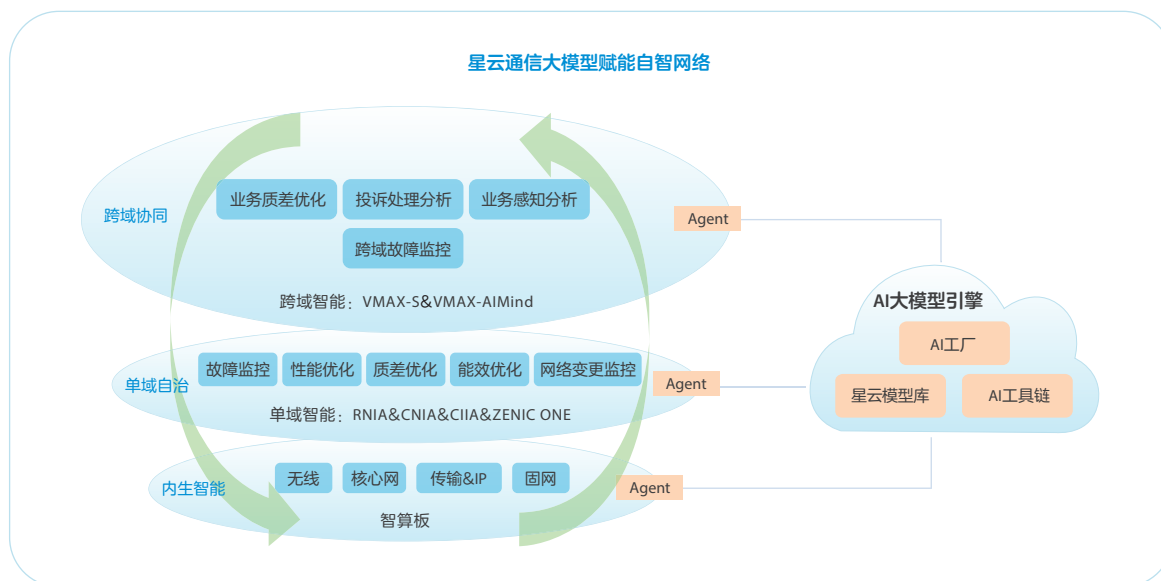


图1 星云通信大模型赋能自智网络全景图

能力。

当前的智能体主要在任务流固化的前提下，通过对话和少量的工具调用来完成任务。提升智能体自主完成任务的能力，需要让其具备更多的长流程自主规划和更全面的专业领域的工具调用能力。为此，中兴通讯不仅在基础模型、基础Agent和基础工具方面来打造Co-Sight智能体，也在通信网络的领域工具方面持续建设，使其更好地面对复杂的通信任务。

中兴通讯积极建设以智能体为核心的自主迭代的运维体系，根据场景按需编排跨域、单域的N个原子智能体，形成全域智能体间的协同。在跨域层，建立跨域多智能体协同中心，灵活编排调用跨域、单域的监控智能体。为了保障智能体能在需要的时候调用到相应的能力，采用可组装式模式开发原子能力，各专业智能体根据场景需要串接原子能力，通过跨域单域协同完成场景任务。

知识图谱：动态融合实现“知识动态性”

随着5G网络的快速发展，通信网络的规模与复杂度呈指数级增长。网络故障的监控与定位面临严峻挑战：多源异构数据（如日志、告警、拓扑、设备配置等）分散在独立系统中，存在数

据语义不一致、格式差异大、更新延迟等问题，信息孤岛现象严重。传统运维方法依赖人工经验与静态规则库，难以高效处理海量非结构化数据（如工单记录、维测日志），且在动态网络环境下缺乏实时根因推理能力。是否能将多源异构数据进行整合，并采用一种更有效的表示方式，从而实现更高效的应用，成为当前需要优先解决的课题。

知识图谱提供了存储和表达这些数据和知识之间关系的手段，将运维经验、指导手册和工单等分散的信息整合成知识图谱，应用图搜索及大模型推理增强能力提升故障定界定位效率和准确率。借助知识图谱，能够将通信领域的结构化数据以及经验语料数据以图形化方式存储，进而助力大模型实现高效使用与精准推理。

与此同时，知识图谱作为通信大模型的“记忆中枢”，其构建需突破数据时效性（网络配置动态变化）与知识完备性（跨厂商设备差异）两大瓶颈。现有知识图谱技术虽然能部分实现知识结构化，但其构建过程通常需要依赖人工标注和预定义本体，且存在自动化程度低、泛化性差、动态更新滞后等局限，无法满足复杂网络环境下故障传播路径的实时分析与决策需求。

由此，中兴通讯提出融合多源异构数据与大

规模语言模型的智能运维框架，结合“规则约束+LLM推理”混合模式，构建动态知识图谱。该技术突破了传统方法对结构化数据的依赖，利用LLM从非结构化日志中抽取隐含关系，并通过流式处理引擎实现知识图谱的实时更新与故障传播树的动态生成，为无线网络故障的精准定位与根因分析提供理论支持和技术保障，对提升网络可靠性和运维效率具有重要应用价值。

深度推理：知识增强保障“推理严谨性”

通信网络的复杂性体现为长因果链（用户感知→无线→传输→核心网）与多目标约束（时延、能耗、成本）。通用的语言大模型容易产生“幻觉”，需通过知识图谱注入领域逻辑约束。在使用大模型推理过程中，可以采用分段和反思来提升推理的准确性。

分阶段推理包括了假设生成、知识验证和迭代修正3个子环节。在假设生成环节，基于输入的多模态观测数据，如告警日志、性能指标、拓扑状态等，大模型通过预训练的领域知识经验与模式识别能力，生成初始假设集合；在知识验证环节，将假设队列输入动态知识图谱系统，进行领域逻辑约束验证；在迭代修正环节，基于验证

结果构建反馈强化机制。

中兴星云通信大模型基于创新性的模型自适应难度分级蒸馏技术，生成长思维链（CoT）语料并人工校对，对基础模型进行冷启动，该阶段使得模型具备完备的思考过程输出能力。然后使用高质量的问题-验证器语料，基于推理强化学习算法，提升模型在特定复杂领域的表现。

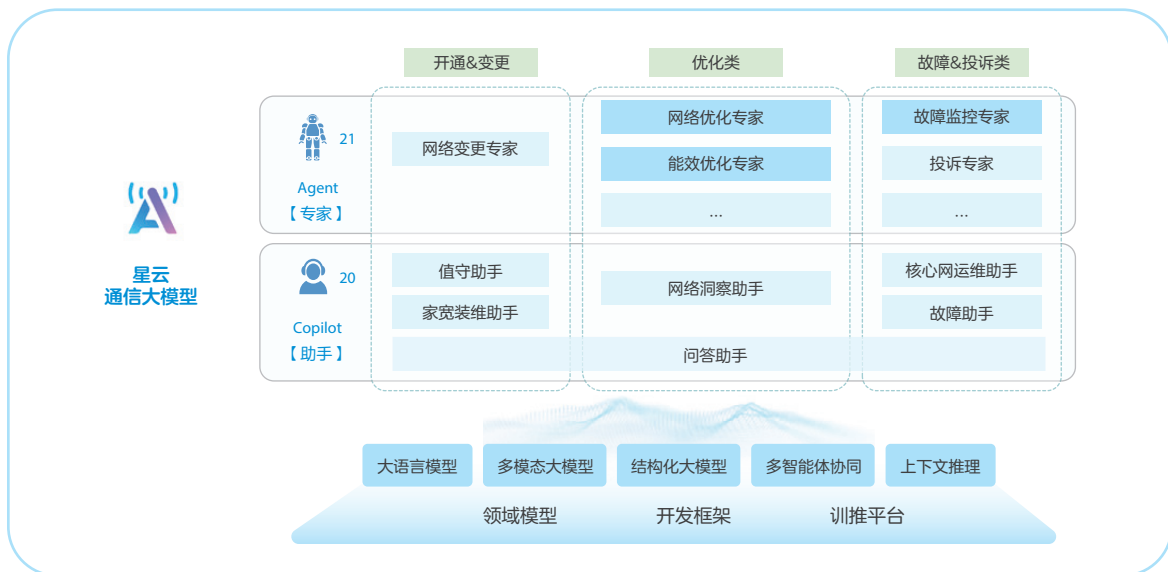
用好大模型：自主准确地解决复杂通信场景难题

在迈向高阶自智的实践中，如何解决长流程、多维度、结构化数据的应用问题一直是技术攻关的重点。我们通过上述四项基础能力的建设和持续迭代，建设面向各场景的Copilot和Agent来攻克这些难题（见图2）。

● 长流程场景的自动化闭环

自智网络的核心是引入大模型、智能体及数字孪生等技术，通过系统决策和执行代替人工判断与操作，实现端到端场景的自动化闭环。更高的自智等级对自动化闭环提出了更高要求，即不再是固定规则+人工兜底的“尽力而为”模式，而是要将长尾场景全部打开，实现全面的自动化

图2 自智网络Copilot与Agent规划



通信大模型智能体对复杂问题分析、推理以及调用工具的能力仍在持续提升，在准确性能够满足生产需要的状态下，“机器决策代替人工决策”的构想逐步变为现实。

闭环。这就要求运维体系具备泛化能力、更强自主流程规划和问题解决能力。

以跨域故障处理场景为例，完整流程包含告警发现、定界定位、方案执行、效果验证四大环节。传统方式需多系统间人工传递数据并协同，存在响应延迟风险。智能化改造需要考虑如何打通断点堵点，用多智能体协同的方式来解放人工操作。星云通信大模型通过灵活切换基座与多智能体协同进行持续迭代，在具备类Manus的长程规划能力的同时，可以随着业界模型能力提升深度思考能力。

● 多维度决策的智能化求解

多维度决策是通信网络重要的智能化能力，需在动态环境中同时权衡带宽、时延、能耗、安全、成本等多重目标参数，通过跨层协同与实时计算实现全局最优解。传统基于规则或单目标优化的方法已难以应对5G/6G的网络切片、边缘计算等复杂场景下的非线性耦合问题。而大模型智能体的引入，通过构建“感知-推演-决策-验证”的闭环认知体系，正在重塑通信网络优化的智能化范式。

以网优场景为例，需综合考虑用户分布、流量特征、设备状态等多维度数据，建立多目标优化模型，在时延、能耗、成本等约束条件下生成最优方案。而星云通信大模型通过知识图谱、结构化数据、深度推理的结合来寻求最优解。

● 结构化数据的深度价值挖掘

面对海量的网络运行数据，需突破传统阈值告警的局限，通过构建告警关联模型识别隐性故

障传播链。同时，对历史工单数据进行精准分析，从中提炼高频问题处置方案，预测潜在网络风险。

在生成式AI技术的加持下，垂直领域语料正从辅助训练数据升维为核心知识载体，其作用已超越传统特征工程的范畴，演变为塑造领域认知能力的基因图谱。星云通信大模型突破多源异构数据融合、专业知识蒸馏、轻量化推理、图搜索等挑战，最终将以系统性的方法论实现通信领域知识的深度内化与快速检索。

通信大模型智能体对复杂问题分析、推理以及调用工具的能力仍在持续提升，在准确性能够满足生产需要的状态下，“机器决策代替人工决策”的构想逐步变为现实。这一目标是建立在大模型理性且有完整逻辑推理、自主拉通流程堵点断点的基础之上，因此驯化大模型的核心任务则转化为通信思维链推理能力的增强、多智能体准确协同以及工具运用，并在具体生产中不断突破大模型对解决复杂问题的能力上限。中兴星云通信大模型在零感换擎、动态知识、深度推理方面持续提升自身能力，不断覆盖更广泛的复杂场景，推动网络迈向高阶自智。

5G-A已正式商用，6G行至将近。未来，6G网络将成为AI原生的网络，AI应用将推动6G网络在智能感知、智能建模、智能化运维及资源调度等方面实现质的飞跃，同时数据平面与控制平面等网络架构演进也将进一步促进AI应用的全面部署，实现AI与6G的双向奔赴。ZTE中兴

智能体工厂实践： 工业级Agent的敏捷生产之路



巫江涛
中兴通讯无线及算力产品
大模型规划经理



付林舟
中兴通讯应用软件开发
无线专家级工程师



Manus作为首个通用智能体引发全球关注时，智能体技术正经历着从实验室探索迈向工业化生产的关键跨域阶段。为了应对传统智能体（Agent）面临的开发周期长、质量标准不一、资产难复用三大痛点，智能体工厂应运而生。中兴通讯构建了业界首个实现“生产-测评-优化-闭环”的工业级Agent平台，以标准化、模块化生产推动智能体开发进入工业时代。

Agent工厂核心架构与生产模式

工业级Agent工厂基于云原生架构，底层支持异构计算如国产GPU和推理卡，为上层大模型应用提供Agent的设计开发和运行的全流程支撑（见图1）。

RAG工作台：知识装配车间

RAG工作台帮助用户快速构建基于检索增强生成（RAG）技术的应用。用户可通过工作台轻松完成知识库构建，包括语料上传、智能切分、Embedding模型选择等关键步骤，并结合测试集进行评测优化。同时，知识检索功能支持自定义 workflows，灵活调整检索策略，提升RAG系统的响应质量。通过RAG工作台，用户可一站式完成知

识管理及效果评估，快速打造高准确率、高性能的RAG应用。

WorkBridge工作台：自然语言与API的智能连接枢纽

WorkBridge工作台是桥接原子能力、扩展大模型能力边界的开发工具。它通过自然语言编程接口（LPI）将自然语言精准映射到现有API，通过自然语言驱动API执行，实现“语言即命令”的智能化调用。其核心功能包括：

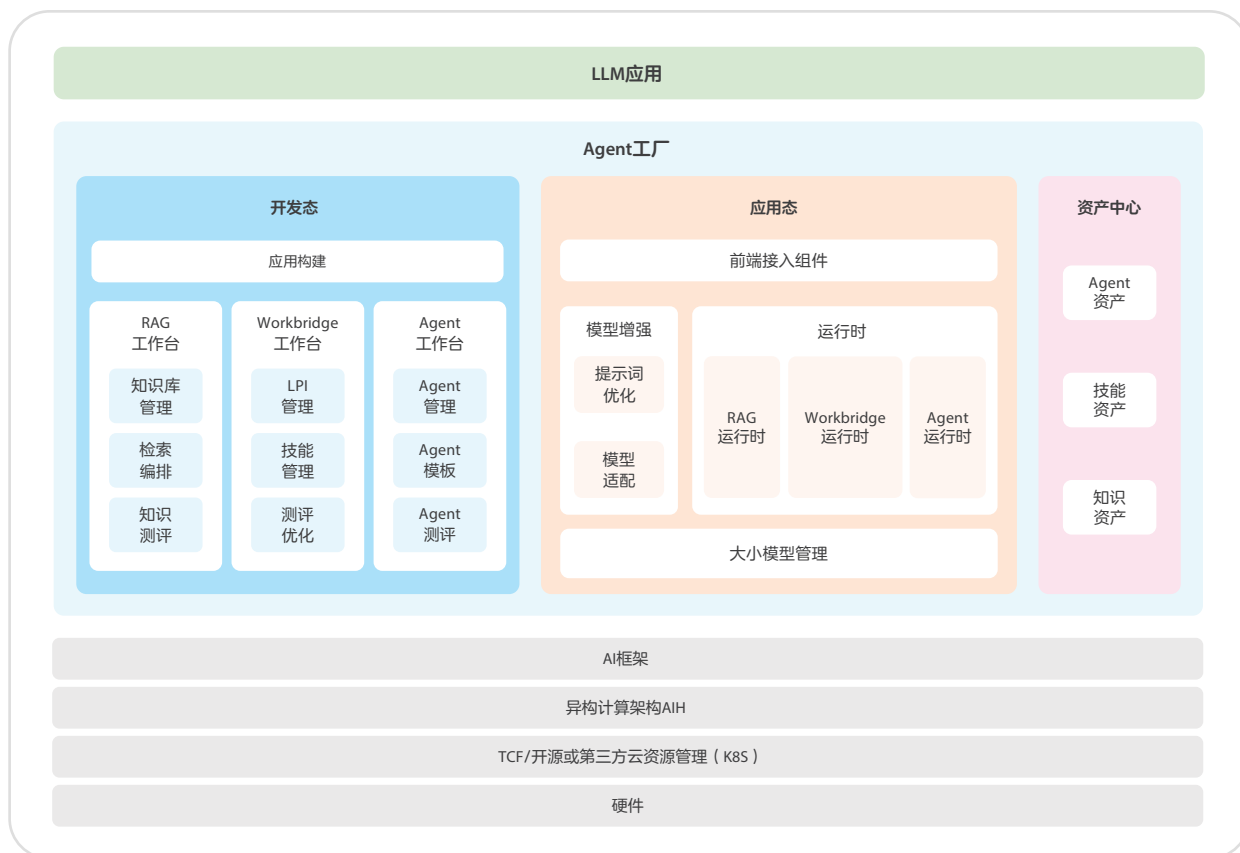
- LPI管理：支持LPI的创建、优化、评测及发布，确保自然语言与API的映射准确可靠。
- 技能管理：将多个LPI组合成可复用的业务技能，提供技能的全生命周期管理。

WorkBridge工作台有效降低了开发技术门槛，加速智能应用开发，成为自然语言到接口调用（NL2API）技术的工程实践典范。

Agent工作台：智能体的终极装配站

Agent工作台提供3种方式满足不同用户的Agent装配需求：基于预置模板的快速开发、AI辅助的智能构建以及自定义的专业开发。

在开发过程中，工作台提供可视化的配置界面和智能对话辅助功能，支持开发者快速完成智能体角色定义、知识库关联以及工具集成等核心



▲ 图1 Agent工厂核心技术架构

配置。开发者可通过手动测试或自动生成测试集的方式验证智能体的表现。

此外，工作台还配备了强大的评估优化功能，提供多维度指标分析和详尽评测报告，支持开发者持续调优。同时，系统还提供智能优化建议功能，通过AI技术辅助分析帮助开发者快速提升智能体性能。

双态协同：工业级Agent生产流水线

双态模式赋能智能体工业化生产：“开发态”通过三大工作台（即RAG、WorkBridge、Agent）协同完成知识生产、技能开发及智能体组装，经严格测试后进入资产中心；“应用态”则依托资产中心实现一键部署，形成从原材料到成品的标准化流水线，在兼顾质量与效率的同时，为AI应用规模化落地提供了可靠支撑。

深度实践：通信网络故障监控Agent炼成之路

传统监控系统存在故障发现滞后、根因定位准确率低、跨系统调度效率低下等问题，亟需通过AI技术构建智能化的故障监控系统，实现从被动响应到主动预防的质变。

中兴通讯采用“知识-技能-Agent”递进架构，打造通信网络故障监控专家。

知识工程建设：构建智能基础

在通信网络故障监控场景中，为使系统具备高效、精准的故障交互问答能力，需构建一套涵盖多领域故障定义和场景化处理指引的知识体系。具体实施包含2个重点：

- 知识库构建：支持文本/Word/PDF导入。

未来，Agent工厂将加速向自动化与专业化深度发展。在自动化方面，通过AI技术实现从需求分析到性能优化的端到端全流程闭环，降低开发门槛；在专业化方面，以通信领域为蓝本构建预定义Agent模板库，实现开箱即用，加速智能体应用部署。

重点建设故障知识库（包含故障定义标准、故障处理指南、告警处置手册等）与设备信息库（设备全称与简称映射关系），并通过RAG工作台完成文档导入和智能分片处理。

- 检索机制设计：采用向量+关键字混合检索模式，利用RAG工作台可视化流程编排功能，实现检索结果的智能重排。

能力组装：打通系统经脉

监控中心工程师经常需要在多个系统间频繁切换以处理故障，为解决此问题，中兴通讯将能力开放平台（获取告警、性能指标、日志等数据）、工单平台和单域工作台（拓扑查询）的API进行自然语言转换封装。使用Workbridge将API转为LPI，组合多个LPI为完整技能。目前已封装15个外部系统API，支撑故障监控全流程闭环处理。

Agent生产：打造数字专家团队

中兴通讯为故障监控场景打造了故障处理“特战队”，包括：故障识别Agent、故障分析Agent、故障调度Agent、报告生成Agent。

以故障分析Agent为例，其核心功能是通过多维数据分析实现故障定界定位。在实际网络运

维管理中，不同领域故障分析思路也有所不同。例如：处理数据网故障需综合机房设备、传输链路等维度，而网络云故障则需逐层分析硬件、虚层至网元。我们将定界定位方法设计为思维链存入知识库，然后为Agent配置思维链知识库，并配置告警分析、日志分析等技能。同时，Agent装配后需经过测试迭代优化，达标后发布至资产中心，再投入生产环境使用。

落地成效

在某省运维中心，中兴通讯故障监控专家“数字员工”实现了7×24小时全天候值守，可达到1分钟精准识别故障，复杂故障定界准确率达91%。中兴通讯故障监控专家目前正纵向拓展专业领域，横向扩大应用区域，持续提升智能运维覆盖面。

未来，Agent工厂将加速向自动化与专业化深度发展。在自动化方面，通过AI技术实现从需求分析到性能优化的端到端全流程闭环，降低开发门槛；在专业化方面，以通信领域为蓝本构建预定义Agent模板库，实现开箱即用，加速智能体应用部署。随着智能化技术的演进，Agent工厂有望成为企业智能化转型的基础设施，释放AI普惠价值。ZTE中兴

智能体和多智能体协同， 推动网络实现L4高阶自智新高度

“ 人工智能+ ” 行动已经连续两年写入政府工作报告，2025年政府工作报告中明确提出支持大模型的广泛应用。推动人工智能技术迈向商业成功，关键在于认知范式的迭代升级和技术的应用落地。智能体是人工智能技术迈向商业应用的桥梁。2025年，智能体和多智能体协同技术呈现科研和应用双轮驱动的高速发展态势。

在网络运维运营领域，智能体能够统一调度大小模型和各类网管系统能力，重构网络运维运营新范式。TM Forum在2024年11月发布的《自智网络L4级产业蓝图-高价值场景》中，引入全栈AI，更新体系架构，在资源运营、业务运营、商业运营各层引入智能体。单智能体工具调用及多智能体跨层跨域协同将成为分阶段推动自智网

络在2030年实现L4高阶自智的核心技术。

智能体和多智能体协同赋能L4高价值场景

在TM Forum的自智网络体系基础上，中兴通讯结合运营价值和技术可行性进行了扩充，确定了L4高阶自智网络一阶段的20多个高价值场景，如个人业务网络投诉处理、5G专网/物联网业务质量优化、网络故障监控处理、网络性能优化等。针对这些高价值场景，以星云通信大语言模型、星云通信专业模型（信令大模型、时空大模型等）为基础，中兴通讯重点打造六类助手、七类专家智能体应用（见图1）。

通过智能体、多智能体协作，中兴通讯开展



张文栓
中兴通讯SDI大模型技术规划总工



冯媛
中兴通讯IDA项目大模型研发经理



Copilot助手（6类20个）

应用名称	使用角色
故障助手	监控工程师
家宽装维助手	装维工程师
值守助手	运维工程师
看网助手	运维工程师
问答助手	运维工程师
核心网运维助手	运维工程师



Agent专家（7类21个）

应用名称	使用场景
网络变更专家	网络变更
网络优化专家	网络优化
故障监控专家	故障监控
投诉分析专家	投诉处理
能效优化专家	能效优化
保障专家	重大保障
看网专家	网络规划

图1 中兴通讯自智网络中的智能体应用

了多项高阶自智网络实践，如网络云故障智能化处理、移动网络业务投诉智能化处理、跨域故障智能化处理等。

网络云故障智能化处理

基于中兴通讯网络云运维产品CIIA，网络云故障处理应用多智能体技术，能够完成超过70次以上的任务动态分解，准确率达90%以上，解决了领域复杂任务并完成落地验证。在实际应用中，某运营商部署CIIA产品，应用日志大模型的故障监控诊断能力，提升了现场故障处理自动化率，降低对专家的要求，大大节省工作量，将交换机故障识别和诊断过程从140分钟以上，降低到20分钟以内。

移网业务投诉智能化处理

中兴通讯VMAX移动业务投诉解决方案引入投诉分析智能体，在投诉受理环节、投诉预处理、投诉处理环节和质检归档4个环节支撑运营商投诉工作的提质增效。相比传统的投诉处理方式，自动化的分析过程有效减少了投诉工单，极大缩短了投诉处理的响应时间，整体移动业务投诉处理时长降低50%，网络问题投诉拦截率提升20%，整体定界准确率84%，一线代维派工工时减少10%。

跨域故障智能化处理

中兴通讯故障监控专家，通过跨域分析智能体、单域分析智能体、方案生成智能体等智能体的协同运作，拉通从故障识别到定界定位再到调度执行的全流程无缝衔接，并实现业务层与网络层在跨域及单域层面的高度协同分析。在实际应用中，某省移动运营商针对IP网故障监控环节实施改进，效果显著：故障识别时长从5分钟降低到1分钟准实时级别，平均修复时间（MTTR）降低8%，而大模型在故障定界方面的准确率显著提升了20%，达到91%。

持续推进，应对高阶自智挑战

L4级别的高阶自智，需通过跨层跨域的多智能体协同实现复杂场景的端到端闭环，最终实现分布式决策、高阶灵活性、高阶扩展性。在持续推进过程中，仍面临多智能体协同技术、应用嵌入生产成效、生态共建协议平台等方面的挑战。

技术方面，多智能体协作框架尚处于发展阶段，幻觉易于在智能体间放大、去中心化的通信机制尚未在应用中充分验证等问题依然存在。中兴通讯打造的星云智能体引擎全面支持合作型、竞争型、混合型协作类型以及规则驱动、角色驱动、模型驱动的协作策略，务实分步应对。

应用方面，2024年，中兴通讯以节能、跨域故障分析、网络变更等高价值场景为标杆的落地验证成效显著，并实现了规模复制。2025年，中兴通讯将围绕自智网络无线网优和无线故障等重点高价值场景继续坐实价值成效的验证，嵌入生产流程，与合作伙伴协同推广复制。

生态方面，智能体相关的协议和平台不断涌现，但仍未形成统一标准，存在互联互通的挑战。中兴通讯持续拥抱开源生态，并为开源社区做贡献，推动开源标准的成熟。2025年4月，在权威GAIA基准测试中，中兴通讯开源了Co-Sight超级智能体，并以72.72的平均分荣登开源框架榜首。同时，星云智能体引擎已支持MCP（model context protocol）等开源协议。

展望未来，中兴通讯将携手业界合作伙伴，持续致力于大模型和智能体生态的全方位构建，联合攻关打破技术壁垒，繁荣以智能体为核心的端到端自智网络高价值场景的应用，推动网络迈向自智新高度。ZTE中兴

结构化信令大模型：

让故障定位效率提升80%的协议级智能革命

5 G-A (5G-Advanced) 正加速推动云游戏、工业元宇宙、远程医疗、车联网等高价行业的智能化升级，对网络质量的要求从“可用”提升至“极致可靠”。然而，信令交互复杂度呈指数级增长，单用户单日交互量突破万条。传统XDR（跨域数据记录）分析存在信元丢失（>30%）和场景覆盖不足（仅60%）的问题，导致运维人员仍需人工解析原始信令，平均耗时超8小时，且复杂故障（如时序冲突、协议兼容性问题）难以准确定位，影响投诉处理效率和网络优化能力。

中兴通讯创新推出结构化信令大模型，突破协议理解智能化壁垒，实现从“人工解码信令”到“模型认知协议”的范式变革，通过采用机器认知+端到端推理技术，彻底重构信令分析范式，显著提升运营商智能运维能力。相比传统方法，该方案可提升信令分析效率80%，故障定位准确率突破95%，并让运营商在微调阶段深度参与，实现业务场景定制化优化，助力5G-A时代智能化用户体验保障体系的构建。

基于结构化大模型的信令分析创新方案

方案基于数据层、模型层和业务层的协同架构，构建端到端的智能信令分析体系，实现从原始信令解析到智能推理分析的技术跃迁（见图1）。

- 数据层：支持对原始信令进行高效编码，覆盖5GC、EPC、IMS等100多种网元接口及相关协议，确保信令数据的完整性和一致性。
- 模型层：融合结构化信令大模型和LLM（large language model），提供预训练基础模型、思维链微调模型及智能推理分析能力，精准理解复杂信令逻辑，提升故障诊断和异常分析能力。

- 业务层：结合实际业务场景，提供信令可视化、异常检测、自动化分析报告及智能问答，赋能运营商实现智能化网络运维。

结构化信令大模型：从“规则驱动”到“智能认知”

方案采用结构化大模型学习原始信令交互模式，结合专家经验的思维链机制，赋予模型自主信令解析能力。

- LLM的适用性：信令本质上是网络设备间的“语言”，具有固定格式、较低信息离散度，与自然语言相比更适用于LLM方法。
- 智能信令分析突破：通过深度学习网络协议特征，模型可实现端到端信令解析、协议冲



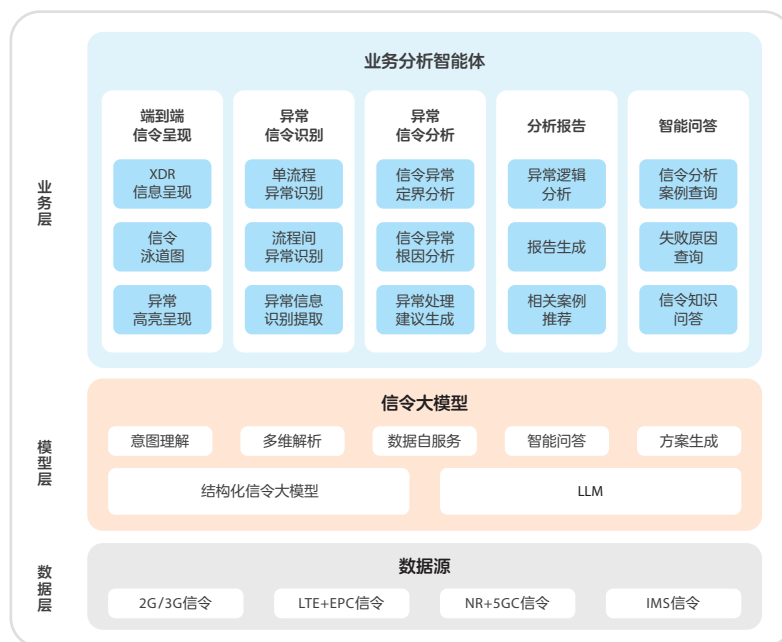
陈燕

中兴通讯数据智能及服务产品团队规划专家



陈湘宁

中兴通讯SDI产品开发经理



▲ 图1 结构化信令大模型三层架构

突检测和异常模式识别。

三大核心模型协同：端到端智能信令分析

方案通过编码模型、投影模型、领域解码模型的协同优化，实现完整的智能信令解析架构，同时为运营商在微调阶段提供深度参与机会，确保模型与真实网络环境精准适配。

- 信令编码模型：基于改进版Transformer和多层次注意力机制，深度解析信令协议的字节内容和层级结构，提供精准的信令语义表示。
- 投影模型：构建信令语义空间与业务规则空间的映射，统一不同协议格式的数据，自动实现特征对齐，降低适配复杂度。
- 领域解码模型：嵌入3GPP协议标准，结合信令流程、消息与业务场景，增强信令推理能力，支持多种智能运维应用。

RAG（检索增强生成）：精准信令知识推理

方案创新性地结合检索增强生成技术（retrieval-augmented generation, RAG），显著提升信令解析的准确性和效率：

- 检索优化：引入检索前路由、查询重写、索引优化、Rerank重排序技术，大幅提高信令数据检索的效率和相关性。
- 减少大模型幻觉：增强RAG结合大模型底座和外部知识源，优化信息生成过程，确保推理结果的准确性与可解释性。

应用赋能：投诉分析智能体

随着5G向5G-A及未来6G演进，网络业务日益多元化，用户体验保障成为核心诉求。面对新业务、新场景的复杂性，传统用户投诉处理模式已难以满足快速响应和精准定位需求。

中兴通讯基于结构化信令大模型，打造投诉分析智能体，引入数字员工“用户投诉处理专

家”，实现智能化、自动化的信令端到端分析，助力运营商实现提效降本。

- 智能投诉知识问答：降低信令分析门槛，提升响应效率

投诉信令数据庞杂、协议流程复杂，人工分析依赖专家经验，知识碎片化导致处理效率低下。基于结构化信令大模型，构建智能问答引擎，通过人机协同交互和历史知识沉淀，为信令分析人员提供即时精准的知识支持，降低处理门槛，提升分析效率。

- 自动化投诉信令分析：精准定位异常，快速给出解决方案

用户可点击智能体入口或直接输入问题意图，触发自动化投诉分析流程。智能体可实现异常信令自动识别，智能筛查；信令流程解析与可视化，提供直观根因分析；标准化信令解释，降低技术门槛；推荐处理建议与失败案例，助力精准决策。

- 智能投诉分析报告生成：释放运维资源，提高服务质量

融合RAG（检索增强生成）技术，整合多领域、多类型数据，生成全面、精准的投诉分析报告。报告生成时间大幅缩短，减少人工投入，提升投诉处理流程的标准化和自动化水平。

2025年，中兴通讯在江苏完成基于结构化信令大模型的投诉分析智能体关键技术验证，标志着智能化投诉处理迈入新阶段。未来，中兴通讯将深耕5G-A网络，推出面向ToC（个人用户）和ToB（企业用户）的智能投诉处理解决方案，助力运营商构建更高效、精准、智能的投诉管理体系。

作为全球领先的电信设备供应商，中兴通讯将继续坚持创新，以卓越的产品和服务，推动数字化、智能化通信网络的发展，为全球客户创造更优质的网络体验。 ZTE中兴

看网智能体，实现AI内生的 多维网络洞察和解决方案生成

作 为无线接入网络的大模型核心应用之一，看网智能体提供无线接入网流量激发解决方案，支持对网络结构、覆盖、容量、设备健康等维度的深入洞察。方案利用生成式AI技术，生成查询策略、总结归纳、解决建议和多维图表，显示了中兴通讯星云通信大模型的强大实力。看网智能体支持用户通过自然语言交互，更高效地理解用户在不同网络阶段、不同应用场景以及不同目标下的网络洞察和解决方案诉求，提升了网络规划和网络优化场景的看网讲网效率。

AI内生应用，智能高效实现网络数据分析和方案生成

传统的网络分析和方案输出，需要运维人员查询大量的网络数据，经过数据处理后提供网络规划和优化建议，高度依赖人工分析和专家经验。引入大模型之后，通过意图理解、RAG（检索增强生成）、NL2API、NL2SQL、NL2Code、Workflow、长短期记忆等关键算法和技术，使得数据分析和方案生成的自智能能力得以提升。

看网智能体带来以下用户价值：

- 意图理解：基于大模型的语言理解和推理能力，以自然语言对话的形式完成看网讲网的流程贯通。
- 报表生成：根据用户的诉求或偏好，灵活生

成文字总结、图表呈现和解决建议，以形成交流、汇报素材。

- 适应用户：根据用户对结果的评价，调整输出以更贴合用户诉求。
 - 专业提升：通过知识问答、引导式提问等功能，边用边学，使得用户对概念更清晰，对问题的描述更专业，提升用户解决问题的能力。
- 看网智能体的六大功能特性如下：
- 多智能体协同：多维网络洞察智能体，组织结构专家、覆盖专家、容量专家、健康专家等专家智能体，面向流量激发场景，协作进行多维洞察，提供洞察结论和解决建议。
 - 数据按需查询：实现无线接入网络8维洞察，含结构、覆盖、流量、负荷、健康等维度（未来将支持资产、性能、能效等3维），用户通过自然语言交互，按需查询统计数据，实现网络体检。
 - 解决方案生成：生成式AI，基于专家知识库和洞察数据，生成洞察结论和解决方案。
 - 创建个人智能体：按用户个人倾向创建智能体的组合，实现场景编排，覆盖多场景。
 - 查询策略生成、图表生成：基于大模型的语义理解、生成和逻辑推理能力，用于查询策略、洞察结论和解决建议的自动生成。
 - 输入联想：用户输入“5G”“覆盖”等简单词汇，看网智能体将自动联想“5G高负荷小区有多少”“弱覆盖小区分布”等问题，简



赵欣
中兴通讯无线网络AIOps
产品经理



尹建华
中兴通讯无线网络AIOps
技术预研总工



严海波
中兴通讯无线网络产品
规划总工

化用户输入的同时，也向用户展示了产品支持的能力范围。

多智能体协同，看网智能体在网络规划场景深入应用

智能体（Agent）可以为大模型核心提供强大的行动能力，是真正释放大模型潜能的关键。在明确的目标下，智能体能够针对目标独立思考并做出行动，它会根据给定任务详细拆解出每一步的计划，依靠来自外界的反馈和自主思考，为自己创建提示词以实现目标。

下面以一个网络规划的应用场景为例，说明Agent的工作过程。如用户输入“大唐不夜城周边的网络情况结果如何”，Agent会把该任务分解为4个主要步骤：分析大唐不夜城所在区域的站点有哪些、判断区域内站点是否稳定运行、判断区域内覆盖是否满足要求、判断区域内容量是否满足要求。

多维网络洞察智能体，首先召集专家智能体，如网络结构专家、网络覆盖专家、网络容量专家、设备健康专家，对该区域内的网络情况展开分析。

- 第一步：结构专家智能体，在工参配置表中筛选站点；
- 第二步：健康专家智能体，调用相关接口，查询退服、告警、隐患网元等信息；
- 第三步、第四步：覆盖专家智能体、容量专家智能体，分别选择覆盖洞察、容量洞察中

相应的工具进行执行；

- 最终，多维网络洞察智能体汇总各专家智能体的信息后，作为最终答复返回给用户。

为了支撑以上用户可感知的业务交互，看网智能体按图1所示的分层结构来设计Agent。

- 任务层：Agent对给定任务进行拆解和整体规划，进行任务编排。
- 执行层：根据任务描述查找合适的工具，并提供工具所需输入信息，完成工具执行动作。
- 适配层：包含会话管理、提示管理、外部知识管理等；提供短期和长期记忆的能力。
- 驱动层：包括对向量数据库、API库等实现访问的管理模块，以及对大模型进行访问的LLM API模块。

目前在网络规划场景中，就利用了多维网络洞察智能体以及多个单维度专家智能体的协同能力。多维网络洞察智能体作为核心智能体，能够规划任务、组织专家分析、汇总专家建议、提出改进举措。多个单维度专家智能体，查询单维度网络数据并生成解决方案。

通过“拟人化”专业分工的智能体，调用各维度的API、知识库以及在线数据，能最大程度的减少大模型的幻觉，提高生成解决方案的可靠性；而多智能体间的协作，可以实现多目标复杂场景的问题解决。

目前，看网智能体随中兴网管的智能化应用，已经在国内三大运营商的24个局点（省级网管）进行商用部署。在2025年3月初的MWC巴塞罗那期间，看网智能体代表中兴通讯无线侧的大模型应用进行了现场互动演示。

看网智能体等系列化智能体的推出，标志着中兴通讯的网络运维进入了智能化的新时代。通过深度学习和大模型技术，它不仅解决了传统运维模式中低效、信息不对称等问题，还在多维度数据分析和专家智能体的协作上取得了突破。未来，看网智能体将不断进化，向更精确的查询、更细粒度的网络分析以及更丰富的应用场景扩展。 ZTE中兴



▲ 图1 Agent分层概念设计

AI大模型网络变更闭环智能体： 运维模式革新

数

字化时代，网络已成为企业运营与社会发展的神经中枢。随着业务的不断拓展和技术的持续升级，网络变更成为保障网络高效运行、满足业务需求的必要手段。传统的网络变更运维模式在面对日益复杂的网络环境和业务需求时，逐渐力不从心。AI大模型技术的蓬勃发展，为网络变更运维带来了全新的思路和解决方案，催生出网络变更闭环智能体，有望彻底革新传统运维模式。

传统网络变更运维痛点剖析

传统网络变更运维体系存在诸多痛点，严重制约企业网络管理效率与稳定性，如运维技能门槛极高、自动化程度滞后、业务中断风险高等。

运维技能门槛高

在传统网络变更运维体系中，方案的设计与评审是极为关键的环节，高度依赖人工完成，要求运维人员必须具备深厚且全面的专业知识。从网络拓扑结构的深入理解，到各种网络设备的配置与管理，再到对不同业务场景下网络需求的精准把握，都需要运维人员有丰富的经验积累。一方面使得符合条件的运维人才数量相对有限，企业在人才招聘和储备上面临较大压力；另一方面，也增加了企业的人力成本。

自动化程度滞后

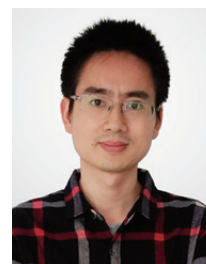
传统网络变更运维过程中，自动化程度严重不足。变更前检查、配置生成等核心环节，几乎完全依赖人工操作。人工操作不仅效率低下，而且容易受到人为因素的影响，出现操作失误或遗漏。此外，由于缺乏自动化工具的支持，难以对变更可能给网络及业务带来的影响进行全面、准确的评估，无法提前发现和规避潜在风险，增加了网络变更的不确定性和风险。

业务中断风险高悬

网络变更过程中，测试验证环节至关重要。传统运维模式下，测试验证环节多且复杂，人工对比效率极低。当在测试过程中发现问题需要执行回退操作时，由于人工操作的及时性，很容易超出预先设定的操作窗口限制，可能引发大面积的业务中断，给企业带来经济损失和声誉损害。同时，人工对比操作还存在主观性和局限性，难以发现一些细微但可能影响网络长期稳定运行的问题。

AI大模型网络变更闭环智能体方案

AI大模型网络变更闭环智能体方案，基于大模型智能体架构完成网络变更方案生成，基于数字孪生完成可信决策，通过自然语言串接完成变



欧雪刚
中兴通讯ITN产品规划经理

变更前检查、变更仿真、变更执行、变更验证、变更后值守完整的闭环流程。

自动/半自动变更方案生成

AI大模型具备强大的数据分析和处理能力，通过对网络相关API进行智能编排，能够生成专门用于变更前检查和变更后验证的工具。这些工具可以自动对网络状态进行全面检测，收集各种关键数据，并与预设的标准进行比对，快速准确地发现潜在问题，为网络变更提供可靠的前期评估和后期验证支持。

方案支持自然语言交互，并能自动串联起整体流程。通过自然语言交互技术，运维人员可以方便地与智能体沟通，输入网络变更需求和指令。智能体能够理解自然语言，并将其转化为具体的操作步骤，无缝衔接网络变更的各个环节，实现整个流程的自动化运行，大大降低了运维人员的操作难度和沟通成本。

可信决策机制

数字孪生技术作为一种新兴的数字化手段，能够在虚拟环境中构建与现实网络高度相似的数

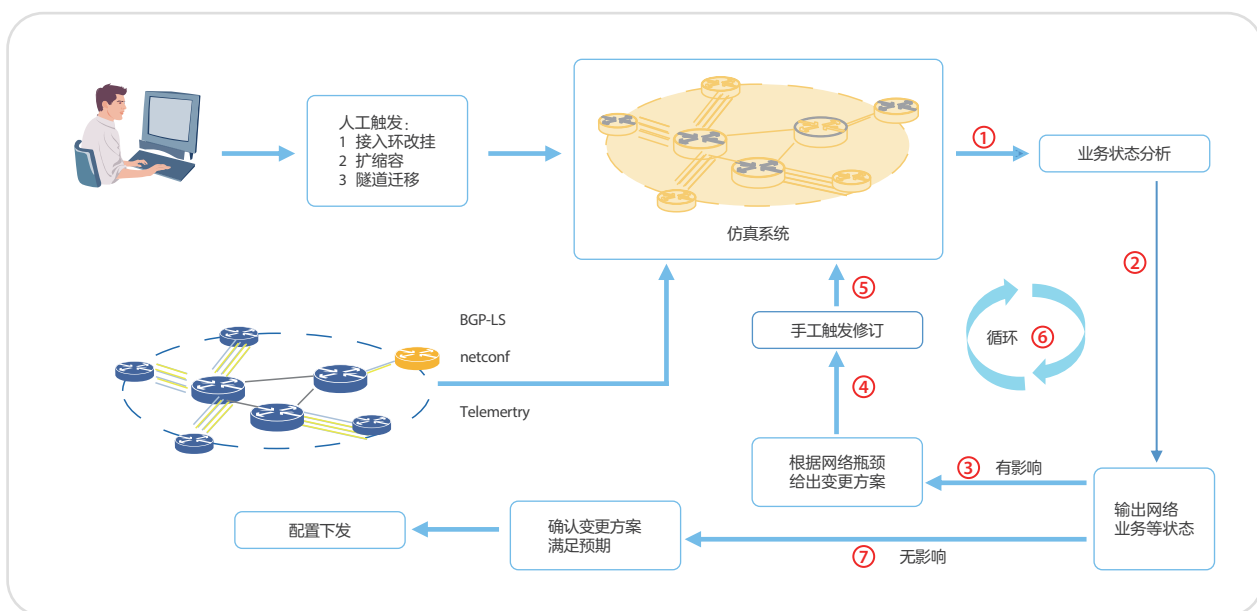
字模型。网络变更闭环智能体方案，以数字孪生底座为依托，对变更执行进行全方位的仿真模拟（见图1）。通过模拟变更在实际网络环境中的执行过程，智能体可以提前发现可能出现的问题，如网络拥塞、设备兼容性冲突等。基于仿真结果，为决策是否下发变更提供科学、可靠的依据，避免因变更而导致的网络故障和业务中断，从源头上保障了变更操作的安全性和可靠性。

可靠执行保障

AI大模型构建执行思维链，将变更前检查、配置变更、变更后验证等关键步骤进行有机整合和逻辑梳理。按照这一思维链，智能体可以自动、有序地执行各项操作，确保每个步骤都按照预设的标准和流程进行，避免了人工操作可能出现的混乱和错误。

原子化操作机制

方案将整个变更执行过程封装为一个原子操作。在执行过程中，要么所有操作完整执行，确保变更顺利实施；要么在出现问题时，自动回滚所有配置变更，恢复到变更前的状态，避免因部



▲ 图1 网络变更仿真

AI大模型网络变更闭环智能体在故障防控效果和效率提升方面取得了令人瞩目的成就。通过智能体的自动化和智能化操作，智能体成功实现人为重大故障为零的目标。

分执行导致的网络故障和数据混乱，极大地提高了变更执行的稳定性和可靠性。

AI大模型网络变更闭环智能体亮点

AI大模型驱动的网络变更闭环智能体，通过自动化与智能化手段，在故障防控和效率提升上实现重大突破。该智能体将传统网络变更运维耗时从14.5小时大幅压缩至4小时，效率提升超70%，同时达成人为重大故障零发生的目标。

故障防控效果与效率双提升

AI大模型网络变更闭环智能体在故障防控效果和效率提升方面取得了令人瞩目的成就。通过智能体的自动化和智能化操作，智能体成功实现人为重大故障为零的目标。传统网络变更运维全流程往往需要耗费14.5小时甚至更长时间，而在智能体的助力下，这一时间大幅缩短至4小时，效率提升超过70%。这一显著的效率提升，不仅意味着企业能够更快地完成网络变更，满足业务快速发展的需求，而且大大缩短了网络变更可能导致的业务中断时间窗口，降低了业务中断风险，提升了企业网络服务的质量和用户体验，为企业在激烈的市场竞争中赢得先机。

可信决策与稳健执行

基于数字孪生底座提供的可信决策能力，是

智能体的又一重要亮点。通过网络变更执行仿真，智能体能够在虚拟环境中对变更的实际效果进行预演，确保变更在实际执行过程中的准确性和可靠性。仿真模拟不仅可以发现明显的网络故障隐患，还能对一些细微的性能影响进行评估，为决策提供全面、深入的参考。

原子化操作机制则进一步保障了变更执行的稳健性。在变更执行过程中，严格按照原子化操作单元进行，避免了因操作失误或部分执行导致的重大故障。即使在出现意外情况时，也能迅速回滚配置，将网络恢复到稳定状态，为网络的稳定运行筑牢防线，降低网络变更带来的风险和损失。

AI大模型驱动的网络变更闭环智能体，无疑是网络运维领域的一次重大变革。它深刻洞察并有效解决了传统网络变更运维模式长期存在的痛点问题，以智能化、自动化的创新手段，实现了网络变更运维效率的大幅提升、风险的有效降低以及管理精细化水平的显著提高。随着AI技术的不断进步和发展，该智能体有望在更多领域和场景中得到拓展应用，与物联网、云计算等新技术实现更深度的融合，为网络运维带来更多的创新和突破，助力企业构建更加稳健、高效、智能的网络基础设施，在数字化时代的浪潮中稳步前行。 ZTE中兴

通信大模型运维智能体， 引领网络保障新趋势



沈远
中兴通讯无线产品规划
总工



沈毅
中兴通讯RNIA项目经理

在 国家“信号升格行动”的推动下，通信行业正加速向智能化转型，以应对网络规模扩大、业务场景复杂化带来的挑战。传统网络运维模式依赖人工经验，存在效率低、响应慢、操作繁琐等痛点，难以满足高价值区域（大型商业体、重大活动场馆等）的保障需求。中兴通讯推出的“星云通信大模型”及其衍生的运维智能体技术，通过人工智能与大模型的深度融合，开创了网络保障的新范式，成为行业智能化升级的核心引擎。

以高价值区域保障为例，中兴通讯联合合作伙伴，基于大模型技术构建的运维智能体方案，在全国多个项目的商用试点中，成功将保障人力投入降低83%，效率提升5倍，实现了网络保障的“智能化跃迁”。

传统保障模式的局限性

传统网络保障模式面临四大核心问题：

- 运维效率低下：网络资源配置与业务需求匹配周期长，难以快速响应突发流量或复杂场景需求。
- 专业门槛高：保障策略依赖专家经验，一线人员需掌握多系统操作技能，培训成本高昂。
- 策略与目标脱节：网络配置多关注故障处理，缺乏与商业目标（如用户体验优化）的深度关联。
- 动态响应不足：重要场合需频繁调整策略，但传统系统自动化水平低，难以实现实时动态优化。

态优化。

大模型运维智能体的创新实践

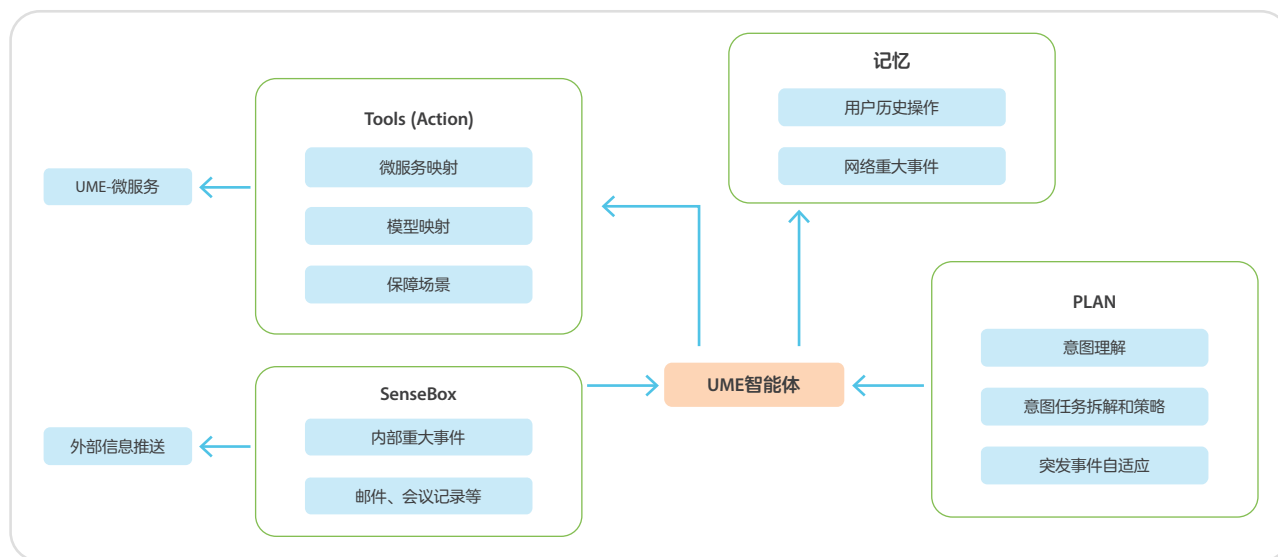
UME无线运维智能体（见图1），是中兴通讯基于大模型技术面向无线网络保障场景的智能体应用，适用于无线网络运维保障场景，如演唱会/体育赛事等重大活动、突发事件应急保障、日常运维中的潮汐保障等。

中兴通讯运维智能体基于“星云通信大模型”构建，采用模块化设计，核心包括：

- 感知模块：实时采集网络性能、用户行为及外部事件数据，通过特征提取与清洗为决策提供输入。
- 大模型模块：作为“智能大脑”，结合自然语言处理（NLP）与知识图谱技术，解析用户意图并生成保障策略。例如，通过日历或邮件触发保障流程，实现场景自动识别。
- 规划与执行模块：基于强化学习优化策略，调用网络原子能力（如资源调度、参数调整）实现闭环操作。
- 反馈学习机制：通过执行数据持续训练模型，提升策略精准度与自适应能力。

技术框架上，智能体融合了生成式AI、多智能体协同（Multi-Agent）及RAG（检索增强生成）技术，有效解决大模型“幻觉”问题，确保决策可靠。

UME无线运维智能体实现端到端闭环保障和动态优化与自适应。



▲ 图1 UME智能体技术架构

- 场景全覆盖：支持重大活动（如演唱会、体育赛事）、突发事件应急保障及日常潮汐流量管理。
- 智能生成式交互：通过自然语言指令（如“保障XX商业体晚高峰网络”）触发流程，自动生成策略并执行，操作效率提升5倍。
- 动态优化与自适应：实时监控网络指标（如用户数、PRB利用率、干扰水平），动态调整策略以应对流量波动。在杭州奥体中心演唱会保障项目中，网络流量预测准确率提升20%，保障效率显著提高。

中兴通讯运维智能体方案，不仅是技术突破，更是运维范式的重构：

- 技术融合创新：大模型与通信知识深度结合，解决了传统AI在结构化数据处理中的局限性，例如通过信令RAG知识库提升故障诊断精度。
- 生态协同共创：联合运营商、行业伙伴共建智能体技术标准与场景库，加速技术规模化落地。
- 商业价值延伸：从降本增效向体验经营延伸，为运营商提供差异化服务能力（如直播、云游戏专属保障），开辟新收入增长点。

应用试点与成效

选择湖北省武汉市全球最大商业体武商时代广场作为验证环境，覆盖31个物理小区、60个逻辑小区，日均用户数超1800人。该区域节假日人流峰值达日常3倍，传统保障需6人天完成，智能体方案仅需1人天。智能体将资源查询时间从2分钟缩短至15秒，故障定位耗时从15分钟降至1分钟，保障人力投入减少83%。

- 网络指标平稳：保障期间用户感知速率提升15%，干扰指标下降30%，PRB利用率稳定在健康阈值内。
- 效率革命性突破：通过自然语言指令下发保障任务，策略生成与执行全自动化，人力成本降低83%。
- 经济效益显著：节省的运维成本可转化为网络投资，助力运营商探索体验经营等新商业模式。

未来，随着6G与通算一体化的演进，运维智能体将进一步向“全域自智”升级，实现从网络保障到商业创新的全方位赋能，为全球通信行业的智能化转型树立标杆。ZTE中兴

核心网运维智能体

助力L4高阶自智持续提效



何伟
中兴通讯智能运维产品
规划总工

随着5G网络的快速部署和业务场景的多样化，核心网的运维复杂度呈指数级增长。传统的运维模式高度依赖人工经验，难以应对大规模网络动态调整、故障预测与自愈等需求。在此背景下，基于大模型（LLM）和智能体（Agent）技术的自智网络成为推动5G核心网智能化运维的关键解决方案。

智能体在智能运维中的应用

智能体技术在通信领域智能运维中得到了广泛应用，在核心网领域，典型的应用场景包括故障智能体和投诉智能体。

故障智能体，智能体实时感知外部告警信息，或接受外部流水线派发故障工单，针对信息中携带的告警完成故障异常检测、根因定位，得出故障根因，反馈故障处理结果并给出相应处理建议，故障工单可自动回复。同时，基于历史故障经验训练的故障处理模型，结合实时数据的监控和分析，智能体亦可预测潜在发生的故障，提前发出预警，使运维人员有足够的时间采取预防措施，降低故障发生的概率。

随着用户极致体验要求不断提升，通信网络用户投诉种类繁多，且波及从终端到无线、承载、核心网端到端的潜在问题，故投诉处理时长一直是投诉处理的最大痛点。投诉智能体针对投诉工单信息中的内容，可智能判断其中用户签约信息、网络配置信息、信令交互码流信息等，通过大模型技术智能分析其中问题，找出问题根

因，给出投诉问题的处理建议，并完成自动回单，投诉工单的处理过程实现自动无人化，极大缩短故障工单处理时长，提升投诉处理效率。

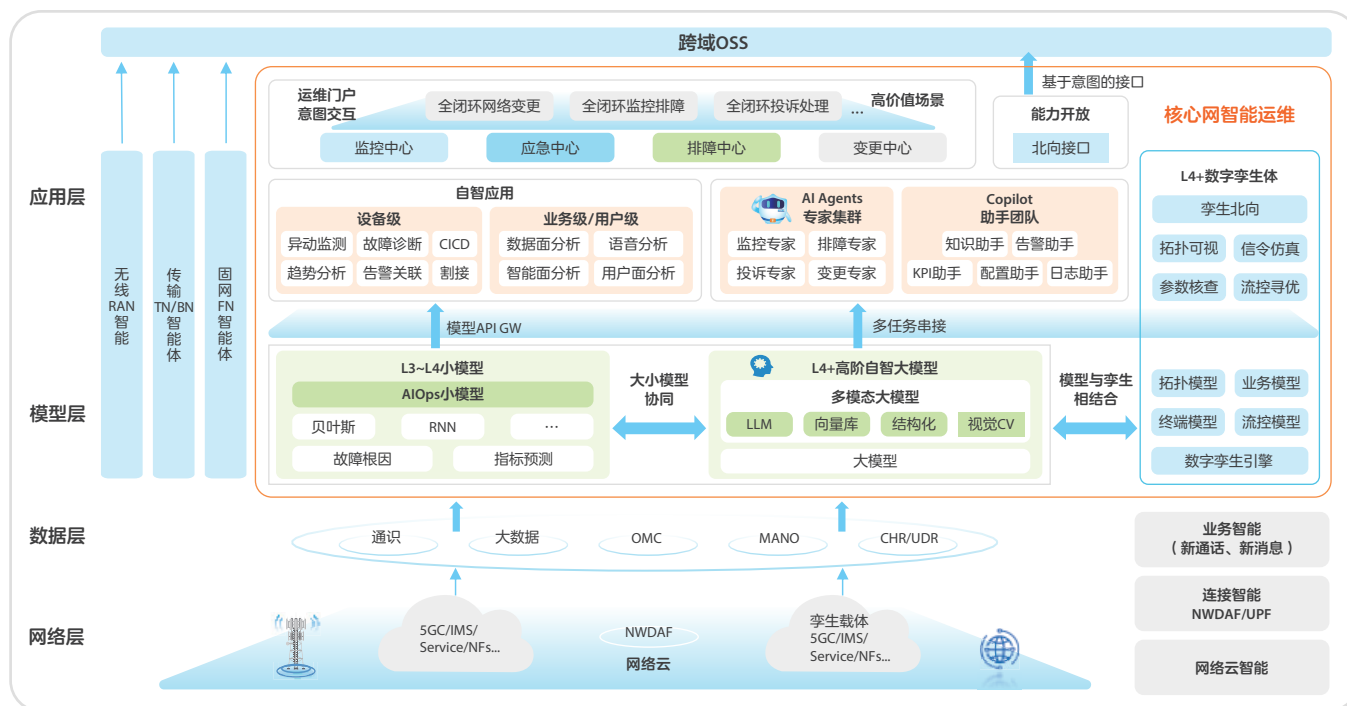
中兴通讯核心网运维智能体体系架构及关键技术

中兴通讯核心网运维智能体系统架构涵盖网络层、数据层、模型层、应用层和数字孪生体（见图1）。网络层作为网络运行根基，包含核心网现有原子网元；数据层提供高质量语料数据，为上层决策提供支撑；模型层以大/小模型为智能引擎，打造可组装编排、自主迭代的数智底座；应用层基于此数智底座，智能编排多种应用能力，满足多场景运维需求；数字孪生体构建“业务模型 + 孪生应用”架构，助力业务创新与网络优化。

为满足以上运维能力，核心网运维智能体分别应用了当前热门的关键技术，包括Graph-RAG（基于图搜索的检索增强生成）、AI Agent、模型上下文协议（model context protocol，MCP）、数字孪生等。

● Graph-RAG

Graph-RAG是一种结合知识图谱与检索增强生成（RAG）的技术，旨在通过结构化知识增强大语言模型（LLM）的推理能力，解决传统RAG在复杂查询和多跳推理中的局限性。其核心目标是通过结构化的知识图谱表示，捕捉数据中实体、关系及全局语义，从而提升LLM对私有未训



▲ 图1 中兴通讯核心网运维智能体产品系统架构图

练数据的理解与生成能力。

● AI Agent

多智能体协同是指多个智能体在共享环境中通过相互通信和协作，实现协同行动以达成共同目标的过程。每个智能体都具备一定的自主性和智能性，能够根据环境信息进行感知、决策和执行。多智能体协同通过相互之间的交互与合作，使整个系统能够从各个智能体的优势中受益，实现更高效、更智能的决策和行动。基于多智能体协同架构，我们可以创建知识专家、故障专家、值守专家、投诉专家等独立智能体个体，通过相互协同，共同构建网络的智能运维体系架构。

● MCP

MCP是一种用于规范大语言模型（LLM）或生成式AI系统如何理解、存储和利用上下文信息的机制或协议，目的是给AI模型和外部数据源或工具建立起统一的通信接口，通常包含上下文窗口管理、多轮对话维护、动态上下文更新等能力。可以将MCP类比为AI的USB接口，无论是AI模型还是外部工具，只要符合MCP标准，就可以实现快速“即插即用”的连接，不需要为每个工具单独编写接口程序，也没有编程语言的限制。

MCP相较于大模型初期的Function Calling，在交互模式、能力定位、通信协议标准化、生态开放程度等方面均有较大的提升。

● 数字孪生

该技术利用离散事件仿真算法等技术，构建核心网数字孪生体，提供核心网系统、设备和组件的数字孪生模型。此模型为运维提供可视化、仿真模拟、预测分析及策略反馈等多维度支持，以低成本实现定性与定量分析双重能力，推动决策过程从人工主导向机器辅助乃至机器自主决策转变，加速高阶自主智能运维演进，最终实现运维管理的全闭环自动化，有效削减运维成本。

智能体的发展正推动运维模式从“自动化”向“自治化”跃迁，但技术成熟仍需攻克可靠性、安全性、协同性等关键挑战。随着智能体、数字孪生等技术的突破，智能运维有望在未来5年内实现L4级广泛落地，为6G时代的“高度自智”奠定基础。我们需要不断地探索和创新，结合具体场景和需求进行应用和实践。同时，也需要进一步加强相关技术的研究和开发，推动智能运维技术的进步和发展。ZTE中兴

家庭知识库及智能体应用探讨



王猛
中兴通讯智慧家庭产品
规划经理

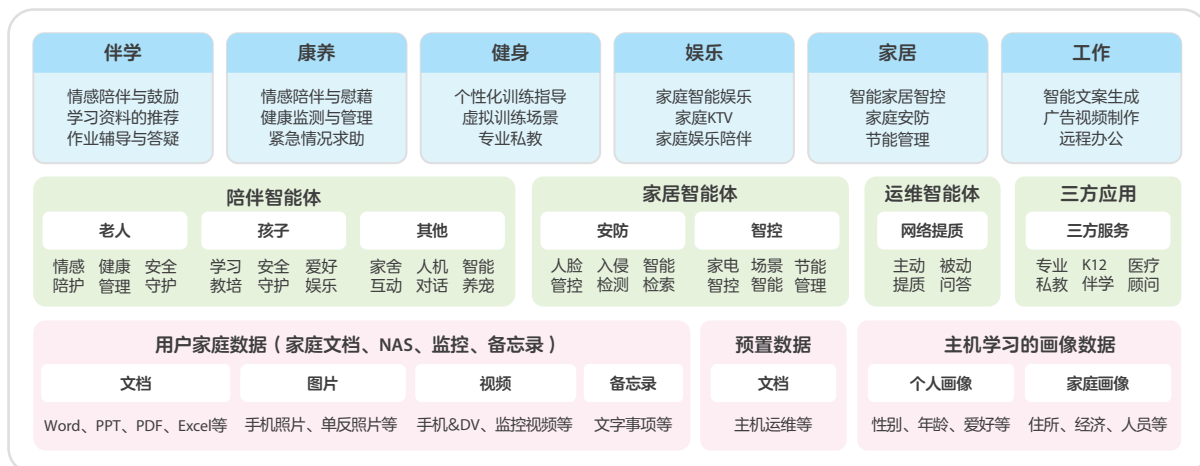
人工智能技术发展迅猛，进入2025年后，在用户体验等关键维度实现了从研究机构到终端用户的实质性跨越，达到了终端用户可接受并能有效应用的水平。在此背景下，家庭场景中的人工智能商业化应用正逐步兴起，为人们的生活带来前所未有的变革。

为精准满足家庭康养、伴学、娱乐、工作、健身以及智能家居等典型场景的多样化需求，中兴通讯深入展开了对家庭AI系统的规划研究。家庭数据作为智能家庭的核心资产，涵盖文档、图片、视频等多种格式文件，以及家庭画像、用户画像等重要数据类型。基于这些数据的家庭知识库规划成为家庭AI规划的核心环节。基于智能体的技术方案，能够同时实现家庭AI服务的主动与被动模式。家庭知识库与家庭智能体共同构成了家庭AI服务的基础支撑，二者协同运作，推动康养、伴学、娱乐、工作、健身、智能家居等多种场景智能化应用的实现（见图1）。

家庭知识库是实现家庭AI服务的核心

家庭知识库是一个整合育儿、伴学、康养、工作、娱乐、健身等多领域知识的综合性知识体系，旨在为家庭成员提供科学指导与个性化服务。其数据构成如下：

- 家庭文档数据：包括用户上传的各类文档、图片、视频数据，同时涵盖系统主动收集的用户备忘录数据。
- 用户和家庭画像数据：借助设备传感器与交互记录，收集用户行为、偏好、健康数据以及家庭环境状况等信息，用于构建用户与家庭画像。
- 家庭知识图谱：通过知识抽取和结构化表示手段，构建反映家庭成员、设备、环境之间关系的图谱。
- 家庭设备与环境数据：包括传感器采集的温度、湿度等数据，设备日志、环境变化记录，以及智能设备的基本信息、使用记录、



▲ 图1 基于家庭知识库与家庭智能体实现多种场景智能化应用

维护历史等。

- 家庭其他知识库：包含用户上传的通用知识库（如百科知识图谱）以及人工整理的数据等。

家庭智能体是实现家庭AI服务的途径

家庭智能体依托家庭知识库中的数据，运用机器学习、自然语言处理以及多模态大模型技术，使其能够精准理解家庭成员的需求，进而提供个性化、主动与被动相结合的服务。具体功能如下：

- 用户画像构建：从家庭知识库提取用户行为数据，并结合家庭成员基本信息与偏好设置完成画像构建；为每位家庭成员构建详尽的用户画像，涵盖生活习惯、偏好、行为模式等维度。
- 场景感知与动态调整：智能体具备感知家庭环境变化的能力，依据当前场景主动提供适配服务；通过整合家庭知识库中的环境数据与设备状态，运用相关技术识别当前场景，并动态调整服务内容。
- 多模态交互与自然语言理解：智能体支持通过语音、文字、手势等多种方式与家庭成员交互，能够理解复杂指令，并结合上下文及历史交互记录，提供更为精准的服务。
- 主动学习与反馈机制：智能体通过持续学习与优化，不断提升服务质量以契合家庭成员需求；收集家庭成员对服务的反馈，动态调整自身行为策略。
- 个性化推荐与主动服务：智能体依据家庭成员需求与习惯，主动提供个性化服务；通过结合用户画像与场景感知，预测家庭成员潜在需求。
- 隐私保护与安全：确保家庭知识库和智能体的使用严格遵循隐私和安全要求。数据存储于本地设备，避免数据泄露风险，采用加密技术保护敏感信息，同时配备用户权限管理机制。

基于场景化落地家庭AI服务

我们对家庭AI服务进行场景化分析，满足不同场景的需求。

● 康养场景

康养场景主要聚焦于健康监测与安全防护。未来，康养场景将更加侧重于主动干预与个性化服务。例如，系统能够识别老年人异常行为并自动触发急救响应。

● 伴学场景

AI在伴学领域的应用主要包括教育机器人和智能学习等应用。未来，伴学场景将深度融合多模态交互技术，如实现情感化互动与自适应教学。

● 娱乐场景

当前娱乐场景以语音控制和多设备联动为核心特征，例如，智能音箱可实现对灯光、音乐和电视的统一控制。未来，全息化与社交化将是发展方向，支持多用户虚拟互动。

● 工作场景

现阶段，智能家居在工作场景中主要提供环境优化和效率提升功能。未来，工作场景将更加注重智能协作等方面的拓展。

● 健身场景

目前，健身场景主要通过智能穿戴设备与家庭健身器材的联动实现。未来，健身场景将融合生物识别与虚拟教练技术，例如，制定个性化训练计划。

● 智能家居场景

当前，智能家居已实现设备互联与自动化控制。未来，智能家居将向“全屋智能”与能源自给方向发展，例如，整合家庭能源，实现能源的优化分配。

借助家庭知识库与家庭智能体，AI技术将实现从单一功能向全场景服务的跨越，家庭服务也从“被动响应”模式转变为“主动关怀”模式。家庭知识库与家庭智能体的结合，将成为提升生活质量、推动社会可持续发展的重要力量。ZTE中兴



无线大模型多智能体保障方案， 为运营商春节保障“护航”

中兴通讯 王洪新，王康，曹莉笠

在全球通信技术飞速发展的浪潮中，智能化运维已成为行业升级的核心赛道。中兴通讯作为通信领域的领军企业，始终以创新技术推动行业变革。2023年第10届乌镇世界互联网峰会上，中兴通讯推出无线大模型，开启通信网络运维智能化的新篇章。历经一年多的技术迭代与实践打磨，中兴通讯无线大模型实现了从单智能体架构到多智能体群智中心架构的跨越升级。这一变革不仅是技术架构的优化，更标志着通信网络运维向更高效、更智能的协同化运作迈进。多智能体群智中心架构赋予系统按需调用多个智能体的能力，如同为通信

网络配备了一支分工明确、协同作战的智能军团，极大提升了网络运维的效率与精准度。

截至目前，中兴通讯无线大模型应用已在全国23个局点成功部署，业务遍及浙江、安徽、云南、广东、吉林、新疆、湖北、重庆、海南、福建等多个省份或直辖市。在这些区域，中兴通讯与当地运营商深度战略合作，将无线大模型技术落地到实际网络运维场景中，推动商用局点的部署与应用，实现网络运维的智能化与自动化转型。

2024年12月，大模型国产化智算一体机首局在云南成功落地。国产化智算一体机的应用，

“大模型群智中心的多智能体组成的决策中枢，构建了应对复杂场景的核心力量。”

是对自主创新技术体系的成功验证，意味着在核心算力设备领域，中兴通讯已具备领先的国产化解决方案能力，为通信网络的自主可控发展提供了坚实支撑。无线大模型通过智能分析、自动化操作等功能，大幅提升了网络管理效率，降低了运维成本，成为区域通信网络智能化升级的标杆案例。

大模型群智中心的多智能体组成的决策中枢，构建了应对复杂场景的核心力量。在2025年春节保障这一重大实践中，中兴通讯三大智能体首次在运营商实现应用，成功经受住了春节保障范围广和持续时间长的考验，展现了多智能体协同作业的强大优势。

春节保障期间，网络面临高流量、复杂场景的双重挑战。大模型多智能体覆盖保障前、中、后全流程，实现高效应对。“看网专家”如同网络的“全景扫描仪”，通过采集网络数据，运用大数据分析 with AI 算法，提前识别潜在风险，如网络弱覆盖、网络低流量，将风险化解在萌芽阶段。“保障专家”聚焦重点事件，以“线”性全流程自动保障模式，实现从保障方案制定到执行的全程自动化。针对热门景区、交通枢纽等重点区域，可自动规划网络资源调度策略，实时监控用户体验指标，确保网络服务质量稳定。“故障助手”作为网络保障的“攻坚先锋”，对重大故

障进行单点突破，迅速启动智能诊断流程，结合历史故障库与实时数据，快速定位故障原因，提供精准解决方案，最大限度减少故障对网络服务的影响。

春节期间中兴通讯协同运营商对黄鹤楼春晚分会场、光谷春晚分会场、武汉站、SKP大型商超、东湖梨园景区等场景的238个5G小区进行保障，通过多智能体协同运作，保障了高清视频直播、现场互动等业务的顺畅进行。

大模型在运营商春节保障的成功应用，不仅为客户提供了全方位网络保障，更通过实际数据验证了多智能体保障方案的有效性，实现用户感知与客户满意度的双提升。

展望未来，中兴通讯将秉持创新驱动的发展理念，持续深化大模型智能体技术优化。在技术层面，将进一步探索大模型与通信网络的深度融合，优化智能体协同算法，提升多智能体在复杂场景下的协作效率；加强大模型对5G行业应用、6G前瞻技术等新型业务的适配能力。在服务层面，为三大运营商定制更高效的运维解决方案，从网络规划、建设到运维全周期提供智能化支持，与行业伙伴并肩前行，共同书写通信事业繁荣发展的新篇章，让智能运维的红利惠及社会各界，为数字中国建设贡献力量。 ZTE中兴



王洪新
中兴通讯网络技术专家级工程师



王康
中兴通讯网络技术系统研发总工



曹莉笠
中兴通讯网管软件开发无线资深专家



山东移动携手中兴通讯： 多智能体协同， 使能网络综合故障监控处理

中兴通讯 周波，宋培双

中国移动拥有超过10亿的移动用户，是全球最大的移动通信运营商，管理运营着全球规模最大的5G网络。作为移动信息现代产业链链长，中国移动积极承担国家科技创新工作，推动产业链的协同发展和技术创新。自智网络作为移动信息现代产业链中的关键技术方向，是产业首批设置的十大子链之一，其目标旨在通过AI智能化技术提升网络运维

效率，实现网络的自动化和智能化，加快向高阶自智网络演进。2024年，中国移动面向各省份分公司和行业合作伙伴发出自智网络子链联合攻关命题，号召业界聚焦AI前沿技术，凝聚合力解决自智网络瓶颈问题。

随着通信技术的飞速发展，网络规模和复杂度不断增加，维护工具和手段的离散化以及经验知识的分散化，使得运维资源整合与协同效率受

“智算一体机”全栈赋能，降低开发门槛，提供从数据准备、模型训练到部署与推理的全流程工具链，极大降低了企业开发AI应用的技术门槛。

到限制，如何加速故障识别、精准定界定位、高效调度故障处置流程、促进故障闭环管理，同时减轻运维人员的工作负荷，成为亟待解决的重大挑战。

中国移动山东公司携手中兴通讯积极参与自智网络子链，组建项目团队开展AI网络运维大模型赋能网络故障监控创新研究。

智算软硬一体化，筑牢大模型应用基石

缺少GPU资源成为大模型应用落地的关键风险，为此项目组多方协调，组织智算专家深度调研组网、工勘情况，深入讨论，最终形成全栈智算解决方案。从发起需求、方案规划到落地仅用45天，为大模型创新应用筑牢智算基石。

中兴通讯提供从算力、网络、能力、智力到应用的全栈智算解决方案，满足不同场景下的性能、成本和服务的差异化需求。智算一体机通过集成高算力硬件、易用的训练推理平台及主流大模型，解决大规模AI模型商业化落地的“最后一

公里”难题。

通过软硬协同优化，智算一体机能够最大限度地释放GPU算力，确保AI应用的卓越性能和高效性。智算一体机全栈赋能，降低开发门槛，提供从数据准备、模型训练到部署与推理的全流程工具链，极大降低了企业开发AI应用的技术门槛。

智算一体机适配中兴通讯星云通信大模型、业界主流开源大模型并支撑RAG服务本地化一键部署，保障运营商运维数据安全与隐私，实现全链路数据加密，结合身份认证机制，有效防止未授权访问，面向应用提供安全、可靠、稳定的推理服务接口。

探索“AI+”故障监控技术创新之路

中兴通讯围绕星云通信大模型引擎提出大小模型协同、多智能体协同，赋能“AI+”故障监控场景，实现更精准的故障识别与定界定位分析、更高效的串联流程加速故障闭环、更智能的意图交互运维体验，释放运维人工负荷。

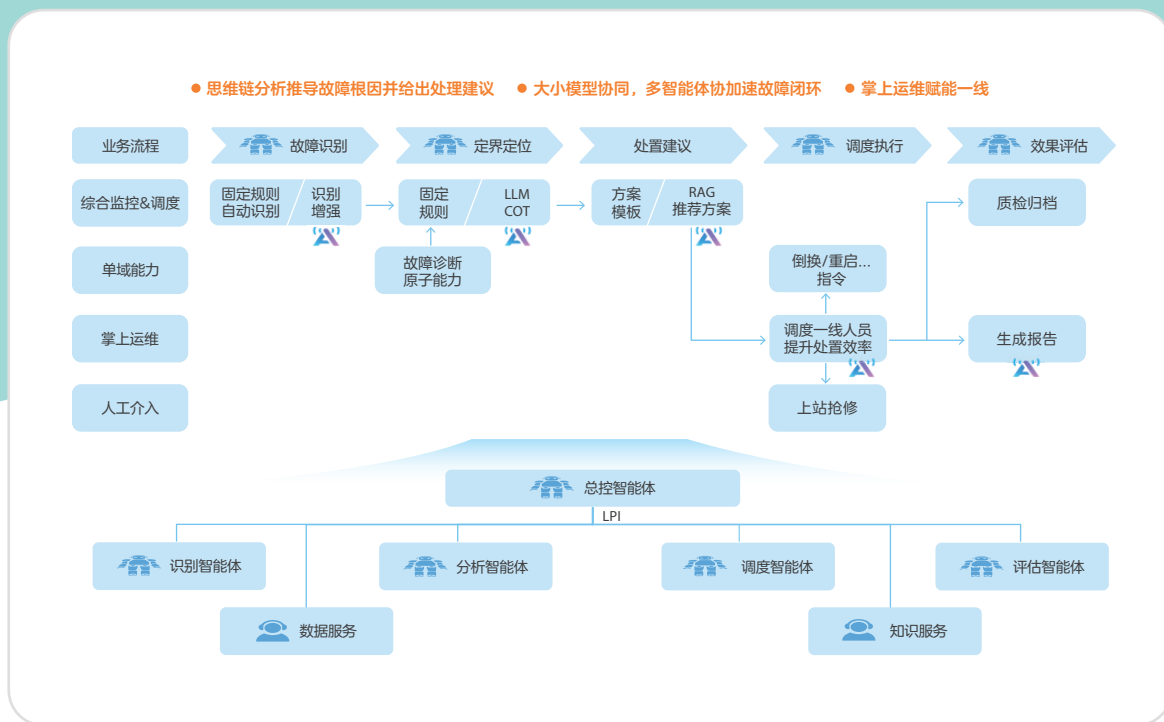


周波
中兴通讯服务产品
高级系统工程师



宋培双
中兴通讯技术支持
专家级工程师

图1 多智能体协同
赋能故障监控



如图1所示，方案创新采用二级意图路由策略，通过总控智能体进行意图识别和路由、流程控制，将任务分发到识别、分析、调度、评估等各个业务智能体，相应的业务智能体来完成具体的任务并将结果返回给总控智能体；总控智能体判断是否进入流程下一环节，协同配合来驱动故障流程闭环。同时，为了降低智能体协作间的误差累积，采用从API向LPI（language programming interface）演进范式，智能体间通过LPI进行协同交互，提升多智能体之间协作的准确性。

在事件识别环节，构建识别智能体，创新采用大小模型协同技术，小模型负责数据动态聚合，大模型负责事件摘要生成，给出言简意赅的事件总结，实现1分钟智能生成事件。

在定界定位环节，创新采用故障知识+大模型因果推理思维链融合推理技术，综合故障案例库与告警数据进行综合推理，将多重原因共同作用导致跨域故障分析准确率提升至91%以上。

基于大模型打造调度处理智能体和智能问答助手，将故障知识、后台数据和原子API能力前

移到掌上端，提升现场处置人员自助解决问题能力，提升前后台交互效率。

通过能力开放嵌入现网故障管理系统、调度系统、掌上运维，在IP网、传输网、动环跨专业故障场景试点验证中，可实现故障1分钟智能识别和生成，故障定界定位分析准确率达到91%，通过智能体协同减少人工参与工作量，驱动故障更高效闭环。

本项目实践入选TMF GenAI IG1345、CCSA中国通信标准化协会《网络运营管理大模型应用场景研究》、ICT中国2024优秀案例、2024年“兴智杯”金砖国家工业创新大模型赛道优秀项目奖、通信产业报网络运维创新力方案，为行业提供通信领域大模型应用实践参考。

未来，中兴通讯与山东移动将继续深入合作，全面拓展故障监控价值场景，加速故障监控机器人深度融入运维生产流程，降低监控和维护一线人员的工作负荷，提升运维效率，实现价值成效闭环。ZTE中兴

nubia
Be yourself

nubia Flip 2

国民小折叠 真AI随声动

全语音AI萌宠 | 全尺寸内嵌DeepSeek | 全方位轻薄抗摔 | 全视角悬停摄影



*图片仅供参考, 请以实物为准。

*画面仅供参考, 手机UI以实际体验为准。

ZTE中兴

成为网络连接和智能算力的领导者
让沟通与信任无处不在