

为什么通信服务提供商必须战略性地大规模实施人工智能的“垂直化”？

电信行业正处于一个转折点。过去二十年，从 3G/4G 到 5G/固定移动融合，再到早期的云原生应用，每一次技术演进都增加了网络各个层面的复杂性。

虽然这些进步释放了新的功能，但也带来了许多通信服务提供商 (CSP) 难以管理的复杂性和运营开销。

现在，随着世界在生成人工智能时代的快速发展，通信服务提供商面临着千载难逢的机会来改变网络的建设、管理和运营方式，使网络在每一层都具有高度的自主性、代理性和“对话性”。

为了实现这一目标，我们转向 Gen AI“垂直化”的概念。换句话说，通过一个自上而下、整体性强且融合了特定领域知识的 Gen AI 平台 (LLM ops、agentic AI ops) 来推动有影响力的业务成果。

我们已经看到针对零售、金融科技、生物科技和大型企业等行业的 Gen AI 垂直化加速应用。

对于通信服务提供商而言，网络垂直行业环境的特点是高速、实时/非实时数据、深度领域知识，以及对本地和自主云支持的需求、严格的监管治理和成本约束。这些独特的属性使得高度垂直化的 Gen AI 解决方案成为不二之选。

我们迫切需要采取行动，而且时机已经成熟。

自动化势在必行：简化复杂。

当今的通信服务提供商面临着双重使命：削减成本，同时构建面向未来、云原生且人工智能原生的网络，以应对新时代的服务需求。传统的运营效率提升方法——流程再造、外包和标准自动化——正在遭遇收益递减。通信服务提供商现在需要的是范式转变，而基于人工智能的自动化恰好能够满足这一需求。

随着每一代技术的升级——涵盖无线电、传输和路由器等多个领域，不断涌现的 5G 独立核心网络和相关切片运营，以及新的盈利平台——我们都添加了抽象层，例如控制平面、编排层、虚拟化/容器化网络功能、混合云……而且这个列表还在不断增长。这不仅使网络管理变得困难，而且维护成本也越来越高。

Gen AI 提供了一个三层平台，可将结构化/非结构化数据流转换为基于角色的代理式 AI 业务成果。这三层分别是：1) 具有模型服务的加速计算基础设施；2) AI 运维/LLM 运维堆栈；以及 3) 代理式 AI 应用。该平台的每一层都专为电信领域构建，具有强大的潜力，可以成为高效的简化基础工具。

Gen AI 能够理解并生成类似人类的响应、分析海量数据，甚至实现决策流程自动化，在推动所有网络层级的智能自动化方面拥有独特的优势。它可以简化服务交付、优化故障排除、简化网络规划并增强客户互动，为构建更精简、更智能、更敏捷的网络铺平道路。

随着通信服务提供商越来越认识到这种变革潜力，重点已从试点项目转向广泛采用和扩展，这反映了战略和投资重点的重大变化。

例如，麦肯锡公司在 2023 年的调查中，34% 的 CSP 专注于试点和快速见效。然而，到 2025 年，64% 的 CSP 则专注于以多智能体系统的形式在整个组织内推广 AI/Gen AI 的应用。

从财务角度来看，Appledore 预测，到 2028 年，通信服务提供商 (CSP) 在自动化领域的投资总额将达到 148 亿美元，其中 AIOps 将占 26%。此外，麦肯锡全球研究院估计，Gen AI 有望释放出比高级分析和“传统”AI 更高 35% 至 70% 的价值。

CSP 数据和成本敏感性：主权和控制问题。

当然，数据控制、主权和成本仍然是通信服务提供商最关心的问题，过去三年随着混合云的加速普及，我们亲身体会到了这一点。由于云服务提供商处理敏感的用户数据，并受到严格的数据保护和国家安全法律的监管，许多大规模的生产力用例必须保留在本地。

这使得自主云和自主 AI 战略至关重要，尤其对于网络和运营层而言。通信服务提供商需要能够拥有并端到端控制其数据管道的基础设施，无论托管在私有云、边缘、区域还是中央数据中心。他们必须采用混合模式，即在本地部署 Gen AI 以处理关键任务用例，并在云端部署以处理更具可扩展性的任务。

同样重要的是避免供应商/模型锁定。封闭的新一代人工智能堆栈会增加成本并限制创新，而以灵活的架构为基础并由模块化技术堆栈支持的开源优先方法，使通信服务提供商能够按照自己的方式，逐步、可持续地构建能力。

工业开源 Gen AI 技术栈的兴起。

Gen AI 在企业、金融科技、生物技术和其他行业的垂直化正在推动对工业级开源 Gen AI 堆栈的需求。

与早期由少数“封闭”商业平台主导的通用人工智能 (Gen AI) 不同，当前的通用人工智能 (Gen AI) 格局正因开放式创新而蓬勃发展。Llama、Mistral 和 Falcon 等模型正在迅速成熟，在特定任务中的表现往往优于专有模型，同时兼具成本效益和高度可定制性。共享训练集数据和开放权重的开放模型正迅速成为常态；更不用说 2025 年初 DeepSeek 3 模型带来的“深思熟虑”和高相关性响应的推理模

型的出现。

为这些努力提供支持的是支持垂直化的丰富的开源 Gen AI 工具生态系统：

- 检索增强生成 (RAG) 框架融合了特定领域的知识，以增强互动，从而推动有意义的业务成果。RAG 整合了两种模型的理解能力和相关的通信服务提供商知识，从而实现特定领域的对话体验。采用先进的 RAG 优化数据科学可以进一步提升响应质量。

- 利用参数高效的微调和 LoRA（低秩自适应）技术，利用深度通信服务提供商专用数据定制基础模型，对于实现电信垂直化至关重要。具体而言，这些模型经过高度调整（“自适应”）和优化，能够执行通信服务提供商的任务，旨在最大限度地减少人工时间投入。

- 知识向量化和索引工具，用于处理各种电信知识库。这包括执行高效分块、高吞吐量多模态数据处理、元数据形成、数据索引和灵活语义检索的能力。掌握这些垂直行业特定的知识管理技术以推动自动化速度，既是一门艺术，也是一门科学。

- 代理式人工智能 (Agentic AI) workflow 和管道，可将通用人工智能 (Gen AI) 无缝集成到现有业务、运营和 IT 生态系统中。这包括在多层、多代理环境中智能地使用函数调用和工具。这种先进的设计模式必须通过严格的测试驱动实践来适应，这种开发机制利用快速工程、可解释性、幻觉管理和稳健性护栏。传统的软件开发和测试体系对于通用人工智能 (Gen AI) 系统来说根本不够用。

通信服务提供商必须越来越多地利用这些类型的进步来实现 Gen AI 堆栈实施，以增强特定业务的成果并提升客户体验。

鉴于电信网络中数据速度高（实时和非实时）、多样性和复杂性，Gen AI 垂直化对通信服务提供商的潜在影响可能更具变革性。

领域知识：缺失的环节。

但这里有一个关键的成功因素：电信领域的知识，这与我们诺基亚息息相关。通用的 Gen AI 模型无论多么强大，都缺乏解决电信行业细微问题所需的情境理解能力。通信服务提供商的运营环境具有独特的术语、监管限制、运营指标和服务模式。

为了有效应用 Gen AI，必须将领域知识深度融入到堆栈的每一层。Omdia 报告显示，69% 的通信服务提供商 (CSP) 已开始将 Gen AI 纳入其关键的网络安全运营，而这需要对 5G 网络及其管理和运营流程有深入的了解。

幸运的是，通信服务提供商在这方面拥有两大优势：

1. 数据（知识库）：海量网络、运营、客户和性能结构化和非结构化数据，跨越数年。

通信服务提供商的独特之处在于其网络、管理和运营以及产品/IT 三个层面；在数据治理方面，每个层面都面临各自的挑战。

2. 电信领域专家：经验丰富的专业人士，在网络规划、运营、网络和服务以及运营方面拥有数十年的经验。

这些资源与通用人工智能能力结合在一起，可以形成良性循环。

例如，基于历史网络故障训练的模型不仅可以自动执行 L1 到 L4 级别的故障排除，还可以预测和预防未来的网络中断。同样，搭载领域训练的 Gen AI 的客户支持系统可以更快地解决问题，并实现更高的个性化。所有这些都开辟了一条战略性盈利途径：通信服务提供商可以将其数据集和领域训练模型产品化，用于智慧城市、汽车和工业物联网等相关领域。

为通信服务提供商构建垂直 Gen AI 平台。

垂直化并非即插即用。要想取得成功，通信服务提供商必须围绕通用人工智能实践构建深厚的内部软件能力。找到一个既能提供领域专业知识，又能带来独特通用人工智能技能组合的关键战略合作伙伴至关重要。

这意味着不仅仅要聘请人工智能科学家——还需要组建融合以下多学科的团队：

- LLM 运维专家，由熟悉生命周期管理、通用人工智能最佳实践以及最新工具包和实践的机器学习系统科学家和数据科学家组成。开源和第三方工具包涵盖多个层面，从数据管理到模型场管理，最终到代理人工智能层。每个层面都需要有各自的侧重点。
- 电信领域专家能够弥合 Gen AI 技术栈成果与有效业务运营之间的差距。这些专家负责配置、规划、部署、服务和维护运营，并对通信服务提供商面临的挑战拥有全球视野。
- 拥有可扩展分布式平台系统经验以及管理/维护通用人工智能平台生命周期的人工智能工程师和数据科学家。关键在于拥有精通微调、数据预处理和性能优化的各级数据科学工程师，尤其是在处理复杂的结构化和非结构化数据时。

在诺基亚，我们已经围绕开源软件平台建立了卓越中心，并通过核心举措开发高弹性、云原生的电信工作负载。如今，我们必须将同样规范且可复制的流程调整为一个“垂直化平台”，以适应代理式人工智能应用的采用。

组织还需要一个清晰的蓝图来实现可重用性，首先要建立一个模型农场框架，以确定何时使用哪些 LLM 和/或 SLM (开源或商业)。必须使用从整个组织收集的联合且干净的数据集来优化模型训练。

有了合适的团队，通信服务提供商就可以掌控自己的人工智能之旅，减少对第三方供应商的依赖，

并根据自己的确切需求定制解决方案。

中期回报：12 至 18 个月内实现明确、量化的收益

这并非一项长期的“登月计划”。通过有针对性的投资和战略协调，通信服务提供商可以在 12 到 18 个月内开始看到真正的生产力提升。

虽然我们专门为 Gen AI 挖掘了 50 多个基于通信服务提供商角色的用例，但一些具有高影响力的代理 AI 用例已经触手可及。想想您组织中的每个角色——从规划工程师、运营和维护工程师、网络策略师、安全专业人员到产品负责人——每个角色都拥有代理 AI 应用，可以帮助将效率提高 5 倍到 10 倍。

一些现实生活中的、简单的用例是：

1. 端到端故障排除自动化 (L1 到 L4)：Gen AI 代理可以通过自动化根本原因分析、解决指导和票证管理来减少平均修复时间 (MTTR)。
2. 设计自动化：高级和低级设计流程通常需要手动且耗时，但可以通过 Gen AI 实现半自动化，从而将规划周期从几周缩短到几天。
3. 服务编排、客户体验和关怀整合：Gen AI 可以统一供应、计费和支持方面的分散系统，创造无缝的客户旅程。

此外，许多此类案例与最终客户体验重叠，从而使它们能够跨 B2C、B2B 和 B2B2X 模式扩展。

结论：通信服务提供商 AI 垂直化并非可有可无，而是必然趋势。

三种力量的融合——对成本效益简化的需求、开源 Gen AI 的成熟度以及特定领域数据和专业知识的可用性——为电信领域的 Gen AI 垂直化创造了完美风暴。

这不仅仅是一场技术博弈，更是推动业务成果的战略要务。立即行动的通信服务提供商将降低成本、提高效率，在这个亟待重塑的行业中释放新的增长动力。工具、人才和技术都已具备——需要的是远见和速度。

人工智能垂直化已不再是未来的考虑。对于通信服务提供商而言，现在正是行动的时刻。