

## 电信行业需要将大规模实施生成式 AI 垂直化作为核心战略

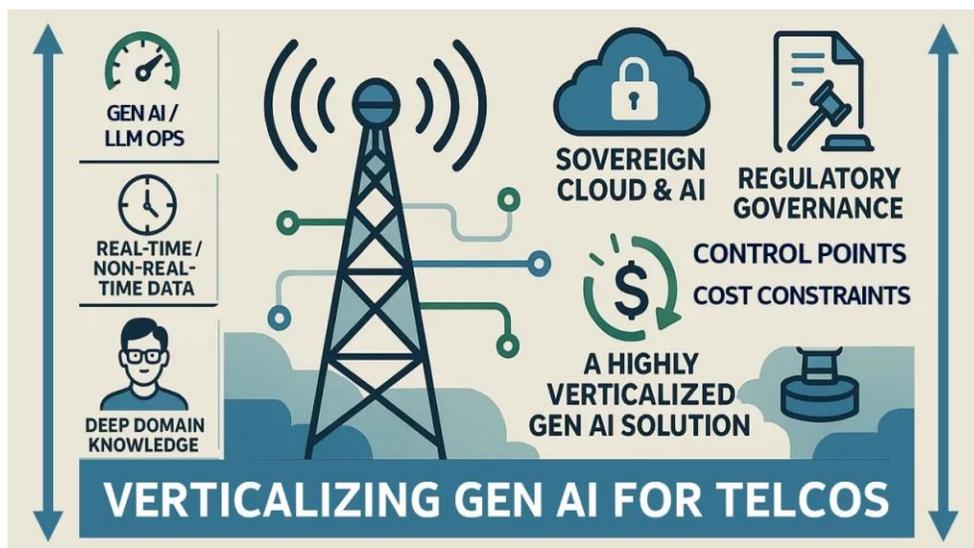
电信行业正处在一个转折点。

在过去的二十年中，每一次技术演变——从 3G/4G 到 5G/固定移动融合以及早期云原生技术的采用——都在网络的各个层面增加了复杂性。虽然这些进步解锁了新的能力，但它们也引入了复杂性和运营开销，许多通信服务提供商（CSPs）正在努力管理这些。

现在，随着世界在生成式人工智能时代迅速前进，电信运营商面临着一个千载难逢的机会，可以改变网络的构建、管理和运营方式——使它们在每个层面上都高度自治、使能和“会话化”。为了实现这一点，我们转向生成式人工智能“垂直化”的概念。换句话说，通过一个从上到下整体且充满特定领域知识的生成式 AI 平台（LLM 操作，使能 AI 操作），推动有影响力的业务成果。截止目前已经看到针对零售、金融科技、生物技术和大型企业等行业的生成式 AI 垂直化加速应用。

对于电信运营商来说，网络“垂直化（verticalization）”是一个以高速、实时/非实时数据和深度领域知识为特征的环境，以及对本地和主权云支持、严格的监管治理和成本限制的需求。这些独特的属性使得高度垂直化的生成式 AI 解决方案成为不二之选。

需求紧迫，时机已到。



## 自动化的必要性：简化复杂性

今天的电信运营商面临着双重任务：削减成本，同时为新时代服务构建面向未来的、基于云和人工智能原生的网络。实现运营效率的传统方法——流程再造、外包和标准自动化——正在达到收益递减的极限。电信运营商现在需要的是一个范式转变，基于生成式人工智能（Gen AI）的自动化正好提供了这一点。

随着每一代的升级——跨越无线电、传输和路由器的多个领域，不断上升的 5G 独立核心网络及其相关的切片操作，以及新的货币化平台——我们增加了诸如控制平面、编排层、虚拟化/容器化网络功能、混合云等抽象层.....这个列表还在不断增长。这不仅使网络难以管理，而且维护成本也越来越昂贵。

生成式 AI 提供了一个三层平台，将结构化/非结构化数据流转化为基于角色的使能 AI 业务成果。这些层包括 1) 加速计算基础设施与模型服务，2) AI 运营/LLM 运营堆栈，以及 3) 使能 AI 应用。当这些层专门为电信领域构建时，这个平台有潜力作为一个高效的基础工具来简化网络。

凭借理解和生成类似人类响应的能力，分析大量数据，甚至自动化决策过程，生成式 AI 独特地定位于推动所有网络层的智能自动化。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——为更精简、更智能和更敏捷的网络铺平道路。

随着电信运营商越来越认识到这种变革潜力，重点已经从试点项目转移到广泛采用和扩展，反映了战略和投资优先级的重大变化。例如，2023 年，麦肯锡公司调查的 34% 的通信服务提供商（CSPs）专注于试点和快速胜利。然而，到 2025 年，64% 的公司专注于在组织中以多代理系统的形式扩大他们的 AI/生成式 AI 采用。

从财务角度来看，Appledore 预测到 2028 年，通信服务提供商将在自动化方面投资 148 亿美元，其中 AI 运营将占 26%。此外，麦肯锡全球研究所估计，生成式 AI 有潜力在先进分析和“传统”AI 之前提供的价值之外，解锁 35% 到 70% 的更多价值。

## 通信服务提供商（CSP）的数据和成本敏感性：主权和控制至关重要

当然，数据控制、主权和成本仍然是电信运营商最关心的问题，我们在过去三年中亲眼见证了随着混合云加速采用而带来的这些挑战。由于通信服务提供商处理受严格的数据保护和国家安全法律监管的敏感用户数据，许多大规模的生产力用例必须保留在本地。

这使得主权云和主权人工智能战略变得至关重要，尤其是对于网络和运营层。电信运营商需要能够让他们拥有和控制端到端数据管道的基础设施，无论这些基础设施是托管在私有云、边缘、区域还是中央数据中心。他们必须接受混合模式，在这种模式下，生成式人工智能被部署在本地以应对关键任务用例，而在云端则用于更可扩展的任务。

同样重要的是避免供应商/模型锁定。封闭的生成式人工智能堆栈会抬高成本并限制创新，而以开源为先的方法，基于灵活的架构并由模块化技术堆栈支持，使电信运营商能够按照自己的条件逐步和可持续地构建能力。

### 工业级开源生成式人工智能技术栈的崛起

生成式人工智能（Gen AI）在企业、金融科技、生物技术和其他行业的垂直化推动了对工业级开源生成式人工智能技术栈的需求。

与生成式人工智能早期由少数“封闭”商业平台主导的情况不同，当前的生成式人工智能领域正以开放创新蓬勃发展。像 Llama、Mistral 和 Falcon 这样的模型正在迅速成熟，经常在特定任务中超越它们的专有同行，同时具有成本效益和高度可定制性。开放模型，共享训练集数据和开放权重，正迅速成为常态；更不用说由 DeepSeek 3 模型在 2025 年初引入的“深思熟虑”和高相关性响应的推理模型的出现。

支持垂直化的丰富开源生成式人工智能工具生态系统正在助力这些努力：

- 检索增强生成（RAG）框架注入特定领域的知识，以增强推动有意义业务成果的互动。RAG 整合了模型的理解能力和相关的电信知识，以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高响应质量。
- 针对定制基础模型参数的高效微调和 LoRA（low-rank adaptation）对于实现电信垂直化至关重要。具体来说，这些模型经过高度调整（“适应”）和优化以执行电信任务，目的是最小化涉及的人力时间。
- 知识向量化和索引工具处理广泛的电信知识库。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。掌握这些特定于垂直领域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。
- AI 智能体工作流无缝集成生成式人工智能到现有的业务、运营和 IT 生态系统中。这包括在多轮、多代理环境中智能使用函数调用和工具使用。这种高级设计模式必须由一种有纪律的实践来适应，该实践利用提示工程、可解释性、幻觉管理和防护措施以实现鲁棒性。在生成式人工智能系统中，传统的软件开发和测试制度是不够的。

电信运营商将不得不越来越多地利用这些类型的生成式人工智能技术栈实施进步，以增强特定于业务的成果并提升客户体验。鉴于电信网络中高数据速度（实时和非实时）、多样性和复杂性，生成式人工智能在电信运营商中的垂直化潜在影响可能更具变革性。

领域知识：缺失的关键环节

这里有一个关键的成功因素：电信领域的知识，这是诺基亚非常擅长的领域。无论通用生成式人工智能模型多么强大，它们都缺乏解决微妙电信问题所需的情境理解。电信运营商在一个具有独特术语、监管限制、运营指标和服务模式的环境中运营。

为了有效应用生成式人工智能，必须将领域知识深度整合到技术栈的每一层。Omdia 报告称，69%的通信服务提供商（CSPs）已经在将生成式人工智能整合到他们的关键网络安全运营中，这需要对 5G 网络及其管理和运营流程有深入了解。

幸运的是，电信运营商在这方面有两个主要资产：

数据（知识库）：大量的网络、运营、客户和性能的结构化和非结构化数据跨越多年。对于电信运营商来说，网络、管理和运营以及产品/IT 的三层结构在数据治理方面各有其挑战。电信领域专家：在网络规划、运营、网络 and 客户服务以及运营方面拥有数十年经验的资深专业人士。这些资源与生成式人工智能能力结合在一起，可以创造一个良性循环。

例如，基于历史网络故障训练的模型不仅可以自动化 L1 到 L4 的故障排除，还可以预测和防止未来的中断。同样，通过领域训练的生成式人工智能丰富的客户支持系统可以更快地解决更个性化的问题。所有这些都开辟了一条战略货币化途径：电信运营商可以将其数据集和领域训练模型产品化，用于相邻领域，如智能城市、汽车和工业物联网。

为电信运营商构建垂直生成式人工智能平台

垂直化并不是即插即用的简单任务。为了成功，电信运营商必须围绕生成式人工智能实践建立深厚的内部软件能力。找到一个注入领域专业知识并带来独特生成式人工智能技能的关键战略合作伙伴是关键。

这意味着不仅仅是雇佣人工智能科学家——它需要组建多学科团队，融合：

- 大型语言模型（LLM）运营专家，包括熟悉生命周期管理、生成式人工智能最佳实践以及最先进的工具包和实践的机器学习系统和数据科学家。从数据管理到模型农场管理，再到最终的使能人工智能层，有多个开源和第三方工具包层。每一层都需要自己的特别关注。
- 电信领域专家，他们能够弥合生成式人工智能技术栈输出与有效的业务运营成功之间的差距。这些是从事供应、规划、部署、服务和关怀运营的专家，并且对电信运营商面临的挑战有全球视野。
- 具有可扩展、分布式平台系统经验和管理/维护生成式人工智能平台生命周期的人工智能工程师和数据科学家。关键是拥有各个级别的数据科学工程师，他们擅长微调、数据预处理和性能优化，特别是在处理复杂的结构化和非结构化数据时。

这些都不是未知领域

在诺基亚，我们已经通过中心计划建立了基于开源软件平台的卓越中心，以开发高度弹性的、云原生的电信工作负载。现在必须将同样有纪律且可复制的流程适应于“垂直化平台”，以采用使能人工智能应用程序。

组织还需要一个清晰的蓝图，以实现可重用性，从模型农场框架开始，以确定使用哪些大型语言模型（LLM）和/或状态空间模型（SLM）（开源或商业）以及何时使用。必须使用从整个组织中收集的联合和清洁数据集来最佳地训练模型。

有了合适的团队，电信运营商可以掌握他们的人工智能旅程，减少对第三方供应商的依赖，并根据他们的确切需求定制解决方案。

中期回报：12 到 18 个月内清晰、可量化的收益

这并不是一个长期的难以达到的目标。通过有针对性的投资和战略性对齐，电信运营商可以在 12 到 18 个月内开始看到真正的生产力提升。

虽然我们针对生成式人工智能挖掘了 50 多个基于电信角色的用例，但一些使能人工智能的高影响力用例已经触手可及。想象一下你们组织中的每个角色——从规划工程师、运营与关怀工程师、网络策略师、安全专业人员到产品负责人——每个角色都有使能人工智能应用程序帮助提高 5 到 10 倍的效率。

一些真实、容易实现的用例包括：

- 端到端故障排除自动化（L1 到 L4）：生成式人工智能代理可以通过自动化根本原因分析、解决方案指导和工单管理来减少平均修复时间（MTTR）。
- 设计自动化：通常手动且耗时的高级和低级设计流程可以通过生成式人工智能半自动化，将规划周期从几周缩短到几天。
- 服务编排、客户体验和关怀集成：生成式人工智能可以统一分散在供应、计费和支持系统，创造无缝的客户旅程。

此外，这些用例中的许多都与最终客户体验重叠，允许它们在 B2C、B2B 和 B2B2X 模型中扩展。

结论：电信行业的生成式人工智能垂直化不是可选项——而是必然趋势

三种力量的汇聚——对成本效益简化的需求、开源生成式人工智能的成熟度以及特定领域数据和专业知识的可用性——为电信行业的生成式人工智能垂直化创造了完美的契机。



这不仅仅是一场技术游戏；它是一个推动业务成果的战略必要性。现在采取行动的电信运营商将降低成本并提高效率，在这个行业正待重新发明的时候，开启新的增长向量。工具、人才和技术都已具备——需要的是远见和速度。

对于电信运营商来说，生成式人工智能垂直化不再是未来的思考，现在是采取行动的时候了。