



加速推进 AI 在全球公共部门的应用

了解红帽如何通过专为实现以下目标而构建的平台，帮助政府机构从试点项目迈向规模化的生产级 AI 应用：

- 利用预测性、生成式和代理式 AI 功能，实现公民服务现代化。
- 在混合、边缘和隔离环境中，全面保障数据主权。
- 通过最大限度地提高硬件利用率，显著降低 AI 运维成本。

破解公共部门 AI 的应用悖论

全球各地的公共部门机构都面临着一项重要任务，即以更少的资源完成更多的工作，同时满足公民对数字原生体验日益增长的期待。各国政府正加速采用人工智能（AI），以提高工作效率、减少人为错误并为决策流程提供支持。然而，要规模化应用这些技术，在结构和运维方面仍面临诸多重大挑战。

要有效实现现代化，机构必须顺利完成 AI 从试点项目到规模化生产级应用的过渡。这需要清晰了解不断演变的 AI 格局，并制定一项以数据主权、成本控制和透明运作为核心的策略。

规划政府领域 AI 发展格局

随着各机构推进现代化进程，唯有了解不同 AI 技术的具体功能，才能使工具与使命目标相契合。例如，预测性 AI 技术能够通过分析历史数据来识别规律并预测未来事件。对于政府机构而言，预测性 AI 在风险缓解方面具有关键作用。常见用例包括检测税务欺诈、预测疾病爆发、预判关键基础架构的维护需求以及评估网络安全风险。生成式 AI（gen AI）突破分析范畴，可通过从海量数据中学习来生成、翻译或转换原创内容。它可通过知识检索、语义搜索以及自动执行总结文件、起草信函、重构现有代码等日常任务，显著提升公共部门的工作效率。最后，代理式 AI 代表着自动化领域的下一个发展阶段。代理式 AI 由能够在预定义参数范围内推理、决策和执行多步骤任务的自主系统组成。与等待提示的聊天机器人不同，AI 代理可以主动采取行动来实现目标，例如跨多个平台解决客户服务问题或自动执行 IT 修复。这使得机构能够从简单的任务自动化，迈向能够适应不断变化条件的自主运维。

突破创新进程中的结构层级障碍

尽管 AI 前景广阔，但公共部门领导者在采用 AI 的过程中仍面临诸多显著障碍。

- ▶ 基础架构老化与数据孤岛：IT 预算的很大一部分通常都用于维护传统系统，致使创新投入空间所剩无几。此外，孤立存储、互不关联的数据使机构无法利用全面的实时信息对 AI 进行训练，从而阻碍了大规模实现自动化。在缺乏统一数据视图的情况下，机构难以部署现代化服务交付所需的自动化工作流。
- ▶ 成本与可扩展性：生成式 AI 对算力的巨大需求可能会使云预算远超计划支出。随着各机构推进流程自动化，例如部署面向民众的聊天机器人，与之相关的推理（即生成响应的过程）成本会迅速攀升。这使机构陷入一个悖论：原本旨在减少人工工作量的工具，却产生了高昂的基础架构费用，迫使他们不得不在其他重要项目上做出权衡。



红帽官方微博



红帽官方微信



- ▶ 数据主权与合规性：《欧盟人工智能法案》等监管框架将许多公共部门用例归为高风险类别，要求提供严格的技术文档、进行系统性偏差测试，并保留防篡改的审计日志。各机构必须克服复杂的数据主权障碍，以确保敏感信息始终处于特定司法管辖区或组织边界之内。这一点对于医疗、执法与司法领域的数据尤为关键，因为这些领域的隐私保护不容妥协。
- ▶ 技能差距：随着 AI 逐渐融入公共服务，专业人才的需求缺口持续扩大。受薪资限制，公共部门组织在招揽高素质 AI 与数据科学人才方面，往往难以与私营企业竞争。各机构亟需配备能够降低技术门槛的工具，让现有员工无需掌握深厚的数据科学专业知识，也为 AI 计划做出贡献。

红帽的解决方案：打造开放、可扩展且注重安全的 AI

红帽提供了一个灵活、开放的平台，旨在帮助政府机构加速推进 AI 创新，同时控制成本并降低风险。依托开源技术，各机构可以避免对供应商的依赖，在开放混合云环境中保持灵活性，并受益于社区驱动的创新成果。

应对公共部门挑战的技术解决方案

- ▶ 混合云的灵活性和主权保障：红帽® OpenShift® AI 可帮助机构跨各种基础架构（本地、云端或网络边缘）构建、训练和部署模型。这种灵活性使数据能够保留在注重安全的环境中，以满足严格的数据隐私和主权要求，例如在隔离或断开连接的环境中。
- ▶ 经济高效的推理：为解决成本不断攀升的问题，红帽 AI 推理服务器可助力实现模型性能优化。该服务器使用虚拟大语言模型（vLLM）等技术来实现高级内存共享和动态批处理，同时结合模型压缩技术，以最大限度地提高硬件利用率。与传统的服务方式相比，这种方法可以显著降低每个词元的成本。
- ▶ 保护隐私的协作模式：红帽支持联合学习，使模型能够跨不同部门或区域进行训练，而无需转移敏感数据。例如，机构可以使用隐私保护技术在本地进行风险预测，避免暴露敏感记录。此外，对机密计算（安全飞地）的支持可确保即使在共享云环境中处理数据，信息也能始终受到保护。
- ▶ 自动化治理和安全防护：红帽可信应用管道将安全控制直接融入开发生命周期。它可以自动生成软件物料清单（SBOM），并进行签名认证和漏洞扫描，从而确保每个模型组件在部署之前都经过验证且符合合规要求。
- ▶ 弥合技能差距：红帽 Ansible® Lightspeed 等工具可使用生成式 AI，将简单语言命令转换为生产就绪型自动化代码。这有助于不同技能水平的多个团队保持一致且高效的工作状态，减轻手动编写脚本的负担，使员工能够专注于更具价值的任务。

红帽 AI 产品组合

- ▶ [红帽企业 Linux® AI](#): 一个基础模型平台，支持政府机构使用开源的 Granite 系列大语言模型 (LLM) 和旨在实现特定领域的一致性的 InstructLab 工具来开发、测试及部署生成式 AI 模型。
- ▶ [红帽 OpenShift AI](#): 一个灵活的机器学习运维 (MLOps) 平台，支持数据科学家和开发人员在开放混合云环境中大规模构建、训练和部署预测性和生成式模型。
- ▶ [红帽 AI 推理服务器](#): 一个可扩展的引擎，可帮助政府机构在不同硬件加速器上运行各种生成式 AI 模型，从而在更短的时间内实现一致且经济高效的推理。
- ▶ [红帽 Ansible 自动化平台](#): 一款智能自动化解决方案，通过 Ansible Lightspeed 集成生成式 AI，以简化基础架构管理并加速 IT 现代化。

红帽 AI 助您达成使命

依托专为全球公共部门特定需求而打造的平台，实现 AI 从试点项目到规模化生产级应用的过渡。红帽提供了丰富的工具，助您战略性地实现自动化、明智管控成本并高效服务民众。

如需进一步了解红帽如何帮助您打造 AI 赋能的未来，请访问[红帽 AI 网站](#)或[联系](#)我们的专家。



关于红帽

红帽通过[一流](#)的支持、培训和咨询服务，帮助客户跨环境实现标准化、开发云原生应用，并实现复杂环境的集成、自动化、安全防护和管理。



红帽官方微博



红帽官方微信

销售及技术支持

800 810 2100
400 890 2100

红帽北京办公地址

北京市朝阳区东大桥路 9 号侨福芳草地大厦 A 座 8 层 邮编: 100020
8610 6533 9300