



借助红帽推进开源 AI/ML 计划

迁移到红帽 OpenShift AI，在整个企业内获得生产级的一致性、灵活性和支持

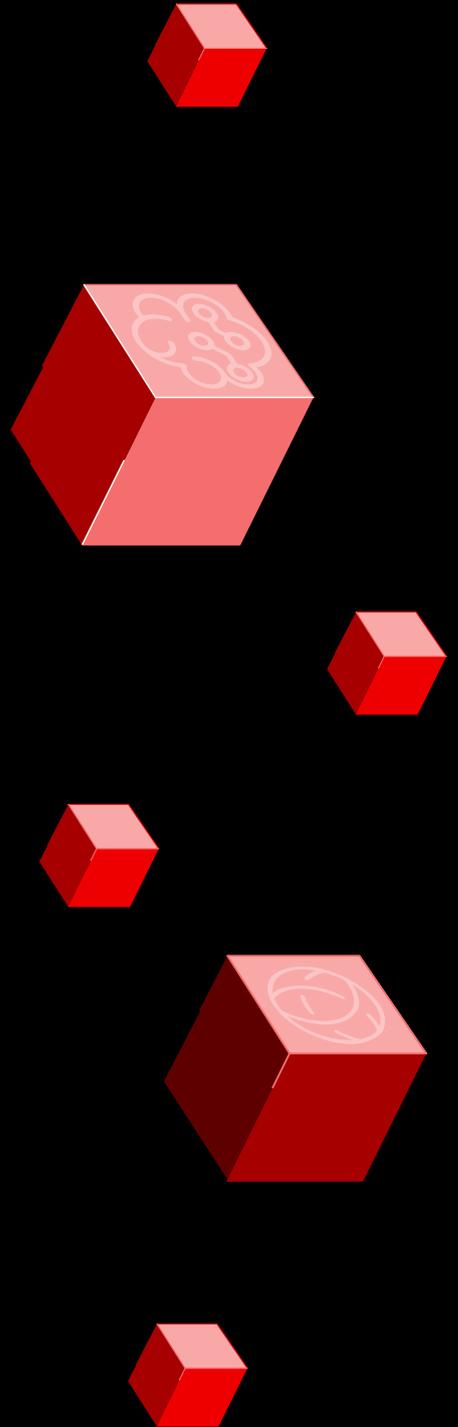
目录

1 借助红帽 OpenShift AI
推进 AI/ML 计划

2 规划迁移

3 红帽 OpenShift AI
适合贵企业吗？

4 准备好开始迁移了吗？



借助红帽 OpenShift AI 推进 AI/ML 计划

人工智能（AI）技术为创新应用提供了动力，这些应用可以加快流程、打造个性化体验，并从大量数据集中解锁见解。AI 是应用的科学与工程引擎，可以执行通常需要人类智能的任务，如解决问题、学习、感知和推理。机器学习（ML）是 AI 的一个子集，使用经过大量数据集训练的算法和统计模型，在不进行明确编程的情况下做出预测或决定。

对于想要实现业务优势的企业技术领导者来说，AI 和 ML 是关键工具。许多企业开始使用 Jupyter、PyTorch 和 Kubeflow 等开源项目创建自己的 AI 平台。虽然这种方法可以让团队在 [Open Data Hub](#) 这样的社区中进行创新，但也需要更大的团队投入更多的精力来测试、修改并整合这些项目。

红帽通过基础技术、可靠的专业知识以及战略合作伙伴关系来帮助您实现 AI 和 ML 目标。

打造 Open Data Hub 的计划经过自然演变后，成为了**红帽® OpenShift® AI** 这个 AI 平台，可为**红帽 OpenShift** 上的 AI/ML 实验和模型提供训练、调优、服务、监控和管理工具。OpenShift AI 为数据科学专家和开发人员提供了一个强大的技术平台，以便他们收集深度见解和构建智能应用。团队可以在一个协作一致的环境中，快速从实验阶段转移到生产阶段。

OpenShift AI 具有**多个版本**，包括一组核心开发和部署功能，如 AI/ML 库和框架、图形处理单元（GPU）加速器支持、数据科学管道和分布式工作负载功能，并与**可信赖的 AI 工具生态系统**集成。数据科学专家可以从他们选择的工具开始，创建自助开发环境并进行实时协作，而开发人员可以更轻松地将容器就绪模型集成到支持 AI 的应用中。同时，这两个团队都可以在一个安全至上的统一平台上部署容器化模型和应用，并快速扩展工作负载，以便在现场、云中或边缘处理数据量、训练时长、模型大小和所需加速等需求。

体验跨团队优势

红帽 OpenShift AI 可让您的整个企业受益匪浅。下面列出了将开源 AI/ML 技术、工具和应用迁移到 OpenShift AI 的诸多原因中的一部分。



生产级支持

OpenShift AI 订阅提供在线和电话支持，让您可以与专家合作，快速解决问题。从一开始就与产品专家交谈，或通过**红帽客户门户**在线查找产品文档、生命周期和支持信息、故障排除工具、红帽知识库文章、帐户和订阅管理工具以及安全更新。



经过简化的升级

红帽为 OpenShift AI 提供了一个经过简化的升级路径，让您无需跟踪开源项目的最新版本或在上游社区中进行错误修复。利用我们在创建企业级开源解决方案方面的专业知识，节省时间并将重点转移回您的业务优先事项。托管云服务用户可以从正常运行时间高达 99.95% 的服务级别协议（SLA）的自动升级中受益，而自我管理的用户则可以通过**透明的生命周期策略**更好地规划升级。



经过认证的集成

使用您首选的开源或付费工具自定义 OpenShift AI 部署。OpenShift AI 直接集成了一系列开源工具。您还可以整合经过认证的合作伙伴产品，用于数据工程、模型开发、硬件和软件加速以及模型管理。这些可选的产品和技术可直接从 OpenShift AI 仪表板访问，其中许多还提供了学习模块和教程。或者替代您首选的上游技术。



软件供应链安全性

红帽通过安全至上的文档化**软件供应链**提供企业就绪型解决方案，包括 OpenShift AI。我们有一个专门的安全工程师团队，负责分析威胁和漏洞、提供前瞻性建议，并为红帽产品开发安全补丁。

规划迁移

在迁移至红帽 OpenShift AI 之前，使用以下问题来评估您现有的开源 AI 环境和需要以及未来的需求。

硬件要求

- ▶ 您知道如何正确调整硬件（包括图形处理单元（GPU）和其他加速器）规模并进行配置吗？

管理策略

- ▶ 用户能够自助置备项目吗？
- ▶ 哪些现有用例需要一个项目？
- ▶ 如何管理对项目的访问权限？
- ▶ 将为团队预先部署哪些资源和工具？

流程和管道

- ▶ 您是否具备或需要一个用于模型训练和数据集成的自定义管道？
- ▶ 您的数据科学专家需要具有独特软件包的自定义 Notebook 吗？
- ▶ 是否需要使用自定义运行时配置您的 ModelMesh？

集成的组件和工具

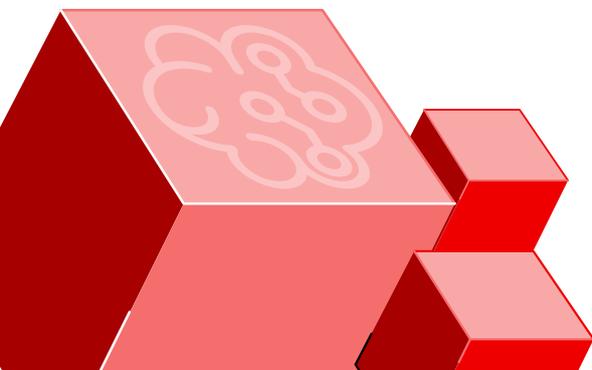
- ▶ 如何将其他数据科学工具集成到您的新环境中？
- ▶ 能否将您的数据存储系统与 OpenShift AI 集成？
- ▶ 您有需要集成的现有红帽 OpenShift 功能吗？
- ▶ 现有系统中是否存在需要额外工具的关键集成点？

解决方案部署

- ▶ 您将如何在多个开发、测试和生产集群上部署和管理 OpenShift AI？
- ▶ 您将如何在环境中部署和配置其他第三方工具？
- ▶ 您将如何跨环境和集群管理模型推广？

用户教育

- ▶ 团队将如何采用和使用 OpenShift AI？
- ▶ 谁来培训数据科学专家并提供自定义服务以改善用户体验？



向红帽专家寻求帮助

红帽咨询提供了以下几项服务来帮助您规划迁移：

- ▶ **红帽咨询：红帽 OpenShift AI Pilot** 是一项服务，可帮助您识别 AI/ML 平台中所需的工具、集成和自定义，并将其部署到您的环境中。
- ▶ **红帽咨询：MLOps 基础** 是一项服务，可帮助您推进 ML 策略、为生产就绪的推理服务创建可重用模式，并使用云原生工具和架构自动化整个 ML 模型生命周期。

[详细了解](#)红帽咨询。

常见迁移步骤

尽管迁移到红帽 OpenShift AI 的具体过程各不相同，本部分还是概要介绍了迁移过程，以及在此期间您可能会遇到的一些挑战。



注意：安装红帽 OpenShift AI 最简单、最安全的方法是在一个全新且未使用的集群上安装。如果需要重用现有集群，请务必先完全删除之前安装的任何组件，然后再安装红帽 OpenShift AI。

Open Data Hub 用户提示

Open Data Hub 是一个社区项目，提供了将常见开源 AI/ML 工具集成到 OpenShift 环境中的蓝图。因为 Open Data Hub 架构是基于 OpenShift 的，所以对于那些想要在迁移到生产级平台（如红帽 OpenShift AI）之前尝试 AI/ML 的企业来说，是常见的选择。

Open Data Hub 有多个版本，在本部分中，我们为准备将其 Open Data Hub 环境迁移到红帽 OpenShift AI 的企业提供了额外的特定于版本的任务。

步骤 1

调查您当前的 AI 平台安装情况。

- ▶ 收集所有清单和配置文件。
- ▶ 收集所有相关设置，包括 Kubernetes Secret 等机密数据和 ConfigMap 等非机密数据。
- ▶ 记录所有已安装组件的版本，如 Grafana 和 Strimzi。
- ▶ 找到所有 Notebook 镜像版本，包括您的自定义 Notebook。
- ▶ 查找您的 AI 平台或其他应用/用户使用的持久性卷声明（PVC）。
- ▶ 识别与您的 AI 平台安装在同一命名空间中的任何其他应用或工具。

建议用于 1.4 之前的 Open Data Hub 版本

- ▶ 针对任何已弃用的第 2 层组件，计划迁移到社区支持的替代选项。
- ▶ 准备将旧的 JupyterHub Notebook 迁移到 KubeFlow Notebook 控制器。
- ▶ 创建计划，将多实例部署整合到一个集中管理的实例。

建议用于 Open Data Hub 1.x 版本

- ▶ 评估部署在 KfDef 中的所有组件，以便迁移到红帽 OpenShift Dev Spaces。

建议用于 Open Data Hub 2.x 版本

- ▶ 确定哪些组件可以保留，哪些组件在迁移后必须删除并重新创建。

步骤 2

备份所有数据和配置。

在进行任何更改之前，执行完整的备份，包括原始数据、用户数据、配置和机密。

实际的备份过程将根据数据的类型和数量、允许的停机时间以及当前的备份和恢复能力而有所不同。

步骤 3 准备好当前的 AI 平台。

将当前的 AI 平台升级到最新的稳定版本，以确保兼容性。

建议用于 Open Data Hub 1.4 及更早版本

将单个用户和 Notebook 使用的 Kubernetes 资源（包括 Secret 和 ConfigMap）迁移到单个项目中。

步骤 4 卸载现有 AI 平台。

- ▶ 使用自动流程来删除尽可能多的组件。
- ▶ 手动删除任何剩余的工件和孤立对象。
- ▶ 根据需要清理旧组件。

步骤 5 安装红帽 OpenShift AI。

按照红帽 OpenShift AI 附带的[说明](#)完成安装。

步骤 6 恢复所有数据和配置。

执行先前保存的原始数据、用户数据、配置和机密的完整还原。

红帽 OpenShift AI

适合贵企业吗？

红帽 OpenShift AI 可能是您的项目或企业的正确选择。下面是成功将 OpenShift AI 用于 AI/ML 工作负载的企业分享的一些标准。

- ▶ 我们想要一个在红帽 OpenShift 上运行的数据科学平台。
- ▶ 我们想要一个能够提供基础构建组件的开放式平台。
- ▶ 我们希望将其他商业和开源产品简单轻松地集成到我们的环境中。
- ▶ 我们有合规性、数据主权和隐私顾虑，需要在现场部署某些工作负载。
- ▶ 我们希望将工作负载部署在混合、多云或断开连接的环境中。
- ▶ 我们希望集中管理多个团队的用户、数据访问权限和资源。
- ▶ 我们支持开源技术。
- ▶ 我们希望简化工作负载部署流程。
- ▶ 我们希望在多个集群资源上运行大语言模型（LLM）这样的 AI 工作负载，并且能够自动扩展。
- ▶ 我们缺乏时间或专业知识来集成、维护和交付一个基于多个上游开源项目的可靠平台。
- ▶ 我们想要一个专门的专家团队，为我们的 AI 平台提供支持。
- ▶ 我们希望我们的数据科学团队使用与其他开发团队相同的实践和工具。
- ▶ 我们已经拥有了红帽 OpenShift 技能、流程和工具，希望将其用于 AI 计划和工作负载。

如果以上标准适用于贵企业，红帽 OpenShift AI 将是您的 AI 现代化计划的有益补充。

准备好开始迁移了吗？

红帽 OpenShift AI 可帮助贵企业快速构建并推进 AI 驱动型应用，从开发阶段一直到生产阶段。

迁移到新的 AI 平台可能具有挑战性。红帽服务提供各种咨询、培训和认证服务，可帮助您构建经过优化的环境，以加快 AI/ML 模型和智能应用的部署。事实上，与红帽合作部署红帽 OpenShift 的企业提高了基础架构利用率，并减少了 50% 的虚拟机占用空间。¹此外，红帽的服务和支持还帮助这些企业实现了 703% 的投资回报率。¹

Red Hat Consulting

红帽咨询可提供实训指导，以帮助您培养技能、简化流程、协调团队并确保系统协同工作。无论您是刚刚开始，还是已经做好准备，想要制定将模型部署到生产环境的策略，红帽咨询都可以提供帮助。

Red Hat Training and Certification

红帽培训和认证可为红帽 OpenShift 管理员和开发人员提供**有针对性的技能路径**。实训培训和认证的主题包括 OpenShift AI 基础和 ML 模型创建。

Red Hat Technical Account Management

红帽大客户技术经理 (TAM) 是技术顾问，可帮助规划和部署红帽软件，为您的 AI/ML 计划提供支持。事实上，在关键项目中加入 TAM 后，3 年后的停机时间减少了 67%。²

¹ 由 Forrester Consulting 代表红帽进行的委托研究，是根据受访客户代表的¹数据综合整理而成。

“红帽 OpenShift 服务和支持的总体经济影响™”，2022 年 5 月。该结果

² 由 Forrester Consulting 代表红帽进行的委托研究，是根据受访客户代表的²数据综合整理而成。

“红帽大客户技术经理的总体经济影响™”，2023 年 3 月。该结果是根据