

Maximize the value of Red Hat Enterprise Linux AI on Lenovo

Artificial intelligence (AI) and generative AI (gen AI) workloads often require a combination of high-performance servers and software that can support intensive computing and streamline AI model training and fine-tuning. Red Hat and Lenovo have partnered to deliver these required components in a single solution, offering the high-performance computing capabilities of Lenovo ThinkSystem SR675 V3 servers factory loaded with Red Hat® Enterprise Linux® AI.

This checklist outlines how you can access Red Hat Enterprise Linux AI on your Lenovo server, what benefits you can expect, and how Red Hat can help you maximize the value of your investment.

1 Choose your subscription and launch the platform




As part of the purchasing process, you will need to choose the length of your subscription, with options for 1-year, 3-year, and 5-year subscriptions. Once you receive your physical Lenovo server, you can set up Red Hat Enterprise Linux AI in 4 steps:

1. Set up your server hardware.
2. Create a security password.
3. Wait for the server to unpack the factory loaded software for Red Hat Enterprise Linux AI.
4. Refer to documentation on Red Hat's website to complete installation and choose your AI models.

2 Access trusted AI models

Choosing your AI model is critical. The size of a model can affect the resources required to train it, the ownership and accuracy of the training data can create potential risks for your organization, and the licensing of the model can limit how you are able to use it.

Red Hat Enterprise Linux AI and the included Granite family large language models (LLMs) provide smaller, more specialized models that streamline fine-tuning and include complete training dataset transparency and full intellectual property (IP) model indemnification from Red Hat. They come with open source licensing that gives you freedom of use and allows you to retain complete ownership of the data you use to train the models and any enterprise applications you build with them.

 facebook.com/redhatinc
 twitter.com/RedHat
 linkedin.com/company/red-hat

3 Fine-tune models to your needs with InstructLab tools

With Red Hat Enterprise Linux AI running on a certified and supported ThinkSystem server, your organization can benefit from the exceptional performance of this combination to significantly increase the efficiency of LLM development and fine-tuning, which is further enhanced by the included models and tools.

The smaller, more specialized Granite LLMs included with Red Hat Enterprise Linux AI minimize the compute resources and data science expertise required to fine-tune a model to your specific use cases and domain expertise.

Red Hat Enterprise Linux AI includes a supported, lifecycle-based distribution of InstructLab that provides a scalable, cost-effective solution for enhancing LLM capabilities and making knowledge and skills contributions accessible to a much wider range of users.

By downloading the InstructLab command line interface (CLI) to your laptop or desktop environment, you can learn and experiment with quantized Granite models and InstructLab model alignment tools. This allows you to evaluate models against your specific use cases, while also being able to incorporate synthetic data generation and training to align those models with your organization's proprietary knowledge and existing internal skills.

Your organization can also benefit from engaging with the InstructLab open source community project, which can help you realize the benefits of using InstructLab model alignment tools in combination with Granite models.

4 Bring performance, scalability, and focus on security to gen AI innovation

Red Hat Enterprise Linux AI on ThinkSystem SR675 V3 servers allows you to develop, test, and run gen AI models on a trusted foundation model platform built on an AI-optimized operating system. This helps you bring performance, scalability, and a focus on security to your AI and gen AI initiatives.

With the combination of Red Hat's software and the built-in computing and GPU processing power provided by the AMD EPYC™ processors and NVIDIA GPUs in Lenovo's hardware, this joint solution allows you to:

- ▶ Run advanced gen AI models, such as the open source Granite LLMs, at optimized speeds, with the data processing and AI model training times you need for AI innovation.
- ▶ Implement your AI deployments with consistency and reliability across a hybrid or multicloud environment, with the performance needed for even the most resource-intensive AI models.
- ▶ Deploy AI models with the focus on security needed for critical AI applications, with robust security features such as [Image Mode deployment for Red Hat Enterprise Linux AI](#).

Get started

Contact [Lenovo](#) to discuss your options for purchasing this joint solution.

5 Kickstart your AI journey with consulting and training

Getting started with any new technology comes with a learning curve, but with the right product expertise, skills training, and technical support, you can expedite this process and learn how to successfully navigate your AI journey.

Lenovo has partnered with Red Hat to meet this need with a range of service offerings designed to help organizations achieve their business goals with AI, including numerous consulting and technical support services.

Red Hat also offers a suite of training courses to help you learn the AI skills you need to succeed, including [a course on how to get started with Red Hat Enterprise Linux AI](#), among a number of other companion courses.

Access these courses, and any others you need to maximize the value of your technology investments, with a free trial to [Red Hat Learning Subscription](#).

You can also access further knowledge, guidance, and support with your Red Hat subscription through the [Red Hat Customer Portal](#).

Learn more

[Read more](#) about the capabilities of Red Hat Enterprise Linux AI on Lenovo ThinkSystem SR675 V3 servers.



About Red Hat

Red Hat helps customers standardize across environments, develop cloud-native applications, and integrate, automate, secure, and manage complex environments with [award-winning](#) support, training, and consulting services.

North America

1 888 REDHAT1
www.redhat.com

Europe, Middle East, and Africa

00800 7334 2835
europe@redhat.com

Asia Pacific

+65 6490 4200
apac@redhat.com

Latin America

+54 11 4329 7300
info-latam@redhat.com

[f](#) facebook.com/redhatinc
[X](#) twitter.com/RedHat
[in](#) linkedin.com/company/red-hat

Copyright © 2024 Red Hat, Inc. Red Hat and the Red Hat logo are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries. Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries. All other trademarks are the property of their respective owners.