

Solution Brief

Add-on Software for
Red Hat® OpenShift® Data Science



Optimizing Data Science Workflows with oneAPI AI Analytics and OpenVINO™ Toolkits

To create a comprehensive end-to-end environment for data scientists to develop container-based AI solutions, Red Hat built Red Hat OpenShift for Data Science. To strengthen the offering, Intel delivers integrated toolkits that amplify capabilities.

“With Red Hat OpenShift Data Science, data scientists can focus on developing models and gaining insights from data without having to worry about managing infrastructure and hardware.”¹

– Sophie Watson
Principal Data Scientist
Red Hat

Data Science Operations Enhanced with Powerful Toolkits from Intel

Data science tasks encompass a broad range of operations—spanning data acquisition through building models to continuous integration and development/deployment processes. As container technology has matured, data scientists increasingly rely on deploying and managing AI solutions using containers and Kubernetes to simplify portability, extend scalability, and manage operations and updates across private cloud, hybrid cloud, and multicloud networks. This can present a layer of unnecessary complexity for data science development teams.

To alleviate this complexity, Red Hat OpenShift Data Science delivers managed cloud services, giving organizations a common set of features across multiple platforms for supporting customers. Data scientists do not have to contend with the setup and maintenance of Kubernetes infrastructures and can focus on integrating models in a more secure environment with the latest open-source tooling and advanced lifecycle management.

Red Hat OpenShift Data Science gives data scientists access to common AI toolkits and libraries, providing an experimental sandbox for the development of data science and machine-learning (ML) projects. Updates and support for core open-source tooling encompasses JupyterHub, TensorFlow, PyTorch, and more. The source-to-image framework used across the Red Hat OpenShift cloud service is also fully supported. Red Hat OpenShift Data Science provides a core platform to streamline data science tasks, including data acquisition and preparation, experimentation with data sets, training of ML models, and other tasks.

The Red Hat OpenShift Data Science environment is offered as an add-on service on Red Hat OpenShift Dedicated and Red Hat Service on AWS. Within this environment, data scientists and machine-learning engineers can take advantage of many advanced AI/ML technologies that are made available from independent software vendors and Red Hat partners. In collaboration with Red Hat, Intel furnishes proven, optimized AI libraries and toolkits, tuned to Intel® architecture-based hardware, including Intel oneAPI AI Analytics Toolkit (AI Kit) and OpenVINO™ Pro for Enterprise to equip data scientists with the latest tools and technologies.

The unified, self-service sandbox environment is fortified with integrated tooling plus extended availability to a multitude of open-source data science projects and proprietary software. Red Hat OpenShift Data Science delivers a more secure, well-provisioned environment in which to accelerate the development and training of AI/ML models. The ML models built within the environment are highly portable for production to a diverse range of platforms, ready for container-based deployment to the hybrid cloud.



Key Benefits of Red Hat OpenShift Data Science

Key benefits available to data scientists and development teams include:

- **Eliminating complex Kubernetes setup tasks –**
Red Hat OpenShift Data Science includes integral support for a full-featured Kubernetes environment, ready for rapid development, training, and testing.
- **Managing software lifecycles efficiently –**
Through the managed cloud service, Red Hat updates the core Kubernetes environment, as well as the Jupyter, PyTorch, and TensorFlow libraries. Operators validate security provisions and automate management of components in the container stack, helping avoid downtime and minimizing manual maintenance tasks.
- **Providing specialized components and ISV support within Jupyter Notebooks –**
Data scientists can work with familiar tools or tap into a dynamic technology partner ecosystem for deeper artificial intelligence and machine-learning expertise, including the AI Analytics ToolKit and OpenVINO toolkit.
- **Streamlining development of network solutions –**
Create ML models and refine them—from initial pilots to containerized network deployments—on a shared,

“Another place that Intel and Red Hat are working together is in artificial intelligence and machine learning. Cloud, 5G, Edge, AI, and machine learning are all interrelated and open. An open, shared platform is really critical here.”²

– Paul Cormier
President and Chief Executive Officer
Red Hat

consistent platform. Data scientists can work efficiently with their choice of tools and access to a self-service infrastructure.

- **Publishing of models as end points –**
Using the Source-to-image (S2I) tool, models are container-ready, which makes it easier to integrate them into an intelligent app. Models can be rebuilt and redeployed as part of a continuous integration/continuous development process based on changes to the source notebook.

Fundamental Stages of a Machine Learning Workflow

The progression of a typical machine learning project follows the stages shown in Figure 1. During different stages, Intel AI components included in the Intel oneAPI AI Analytics Toolkit and OpenVINO Pro for Enterprise can be integrated into the environment to help optimize and refine the model development and deployment.

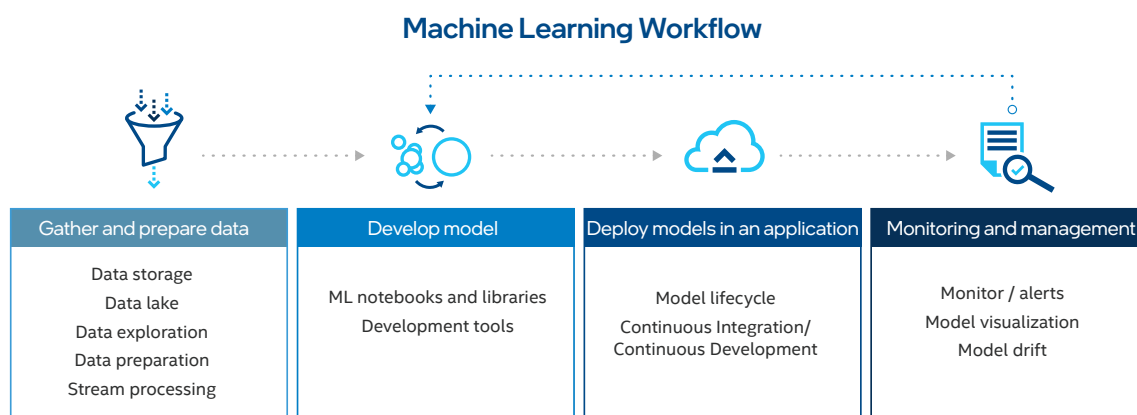


Figure 1. Machine learning workflow includes feedback loops for continuous integration and continuous delivery and deployment.

To provide a common collaboration space within which data scientists can share code, visualizations, multimedia elements, and equations, Red Hat OpenShift Data Science includes JupyterHub fully integrated with the environment. This open-source web application has gained popularity and been widely adopted in the developer community, enabling project teams to work together more efficiently and provide a shared knowledge base for advancing AI model design. As shown in Figure 2, the workflow across the pipeline lets developers prepare data, freely explore capabilities, experiment with different approaches, build models, and deploy those models for running on a container-based platform. The AI Analytics Toolkit covers the essential tasks for data scientists at each stage of the machine learning workflow, whereas the OpenVINO toolkit is slanted toward achieving inferencing efficiency on Intel architecture-based platforms.

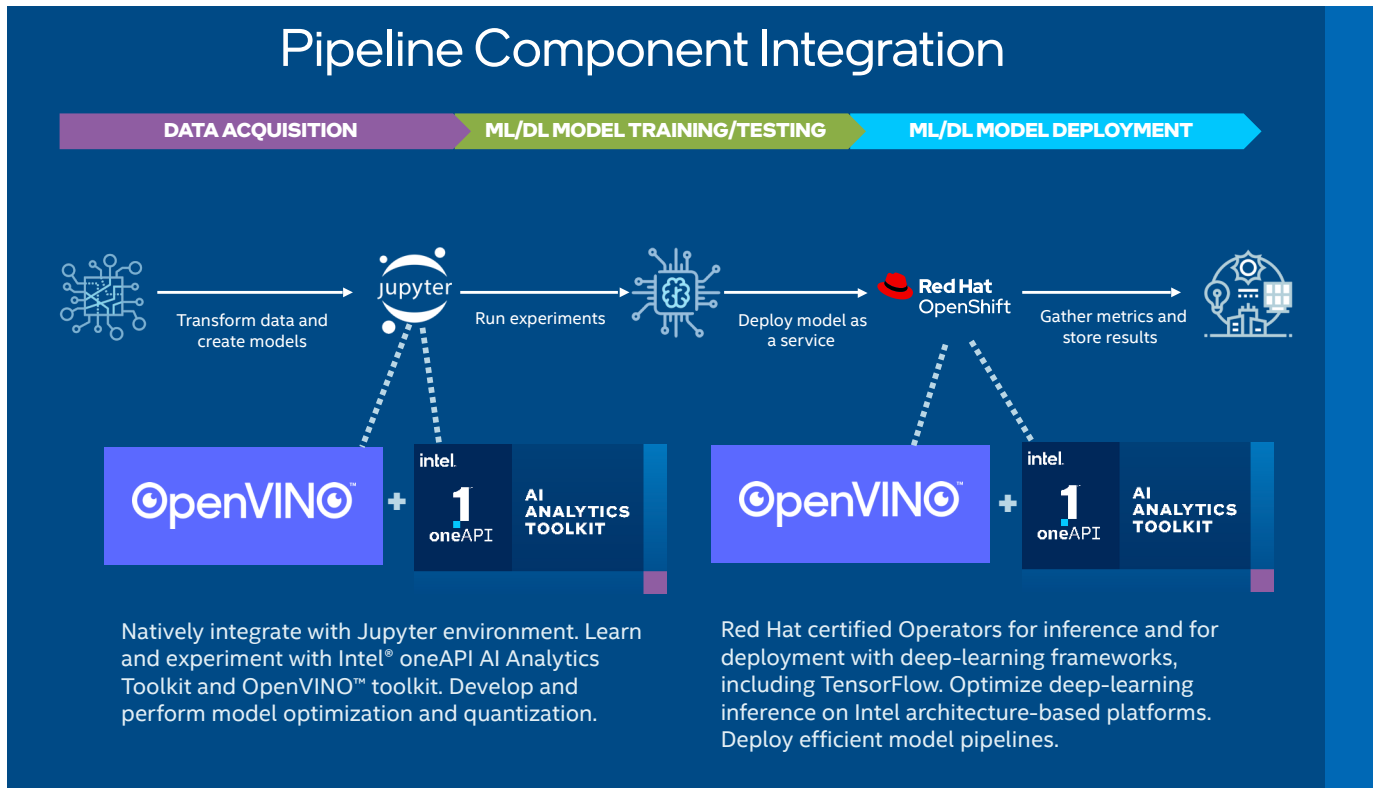


Figure 2. Intel oneAPI AI Analytics Toolkit and OpenVINO Toolkit integrated with Red Hat OpenShift Data Science.

Intel oneAPI AI Analytics Toolkit Overview

Optimized for maximum performance on Intel architecture-based CPUs and XPUs, the Intel oneAPI AI Analytics Toolkit brings a slate of tools and frameworks to streamline and accelerate the work of data scientists and AI researchers. Accessible directly from within a Jupyter Notebook in the Red Hat OpenShift Data Science environment, the toolkit components are constructed with oneAPI libraries, tuned to deliver exceptional performance for preprocessing tasks, model development, training, and inference.

As shown in Figure 3, for deep learning, the Intel oneAPI AI Analytics Toolkit includes:

- Intel Optimization for TensorFlow
- Intel Optimization for PyTorch
- Model Zoo for Intel Architecture
- Intel Low Precision Optimization Tool

Toolkit contents include accelerations for data preprocessing and machine learning frameworks, such as:

- Optimizations for NumPy and SciPy
- Drop-in optimizations with Intel Extensions for scikit-learn and XGBoost
- Drop-in optimizations with Intel Distribution of Modin (for scaling Pandas workflows)

Pretrained models give data scientists a valuable starting point for AI projects, improving interoperability across diverse architectures. Acceleration for data preprocessing and machine learning operations are gained through optimization of compute-intensive packages. The toolkit delivers easy access to powerful Intel analytics and AI optimizations from within Red Hat OpenShift Data Science.

Intel oneAPI AI Analytics Toolkit Operator

The Intel oneAPI AI Analytics Toolkit is available as a Red Hat OpenShift Operator. The functionality can be accessed through Red Hat OpenShift Data Science by having an Administrator enable the AI Kit Operator from OpenShift Dedicated. An install option in the OpenShift Data Science OperatorHub enables the Operator in OpenShift clusters, and AI Kit features are then available when OpenShift Data Science is started.

“The hardware requirements to manage data across the end-to-end AI workflow are getting more diverse, with unique accelerators coming to market to manage specific use cases. A recent Evans Data Corporation report notes that 40 percent of developers target heterogeneous systems that use more than one type of processor, processor core, or coprocessor.”³

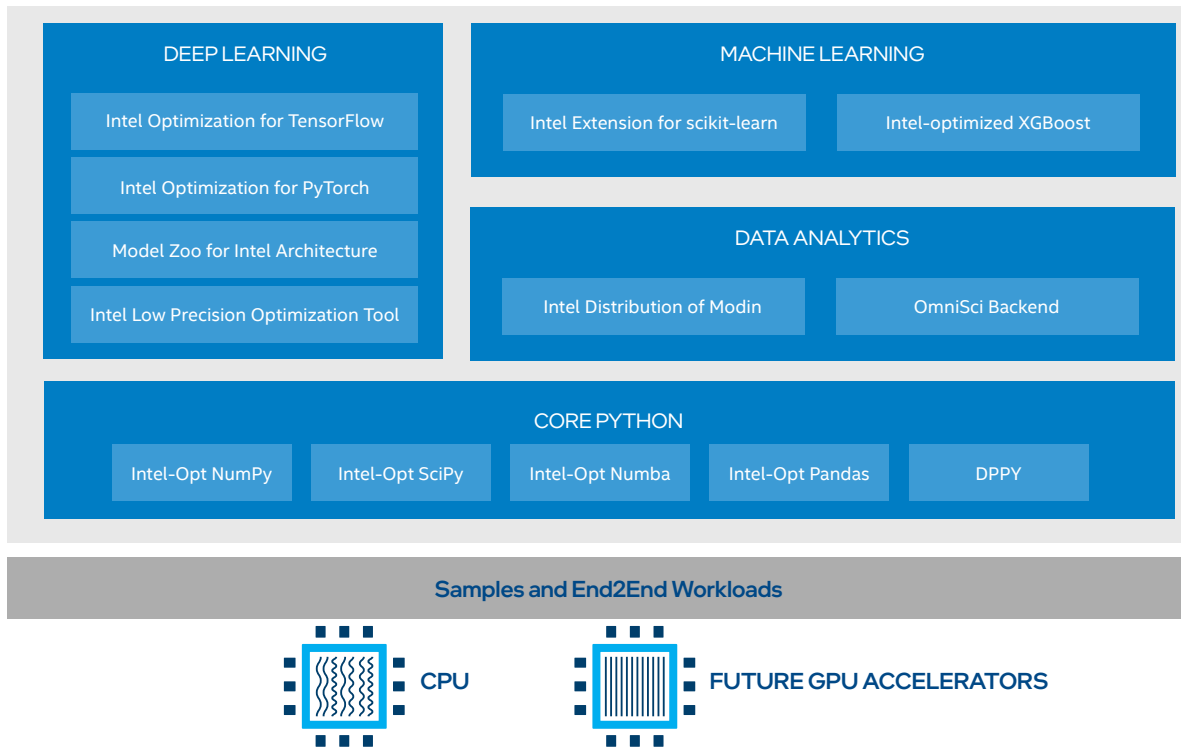


Figure 3. Overview of the Intel oneAPI AI Analytics Toolkit.

AI Analytics Toolkit Contents

Intel-optimized Python-based frameworks, tools, and instruction sets deliver exceptional performance running machine learning models and performing training and inference across a span of supported processors and accelerators. The toolkit contents include the following components.

Intel Optimization for TensorFlow

Intel and Google collaboratively optimized TensorFlow for maximum performance based on primitives from the Intel oneAPI Deep Neural Network Library (oneDNN). Intel Optimization for TensorFlow as part of the AI Kit also provides additional GPU support exclusive to AI Kit that is not upstreamed with Google.

Intel Optimization for PyTorch

Intel and Facebook collaboratively combined many Intel optimizations with the PyTorch deep learning framework. This performance-tuned package includes the latest PyTorch release for CPUs in binary form. It also adds Intel extensions and bindings with Intel oneAPI Collective Communications Library (oneCCL) to facilitate distributed training. Intel Extensions for PyTorch also provides additional optimizations.

Model Zoo for Intel Architecture

Pretrained models, sample scripts, coding guidelines, and tutorials in the Model Zoo give data scientists a rich set of resources for launching AI projects. Optimized open-source ML models are performance tuned for Intel Xeon Scalable processors.

Intel Low Precision Optimization Tool

This tool—featuring an open-source Python library—offers a low-precision inference interface spanning multiple Intel-optimized deep learning frameworks.

Intel Extension for Scikit-learn

This performance-tuned extension package patches scikit-learn estimators to use Intel oneAPI Data Analytics Library (oneDAL) as the underlying solver. Stock scikit-learn is included to provide a full-featured, complete Python environment.

Intel optimized XGBoost

Intel in concert with the XGBoost community has been upstreaming CPU optimizations for maximizing performance on Intel processors. This ML package features drop-in acceleration for accelerating model training and enhancing predictive operations.

Besides the upstream optimizations, there are further optimizations that users can utilize for XGBoost prediction with the daalpy library alongside GPU support. Packages in the AI Kit also utilize oneDAL solvers for fast gradient boosting inference with little to no accuracy loss.

Intel Distribution of Modin

This intelligent, distributed DataFrame library supports scalable data processing across multiple nodes. Optimizing the Pandas API, the library integrates with OmniSci in the backend to accelerate analytics operations, which is automatically installed on Red Hat OpenShift Data Science.

Intel Distribution for Python

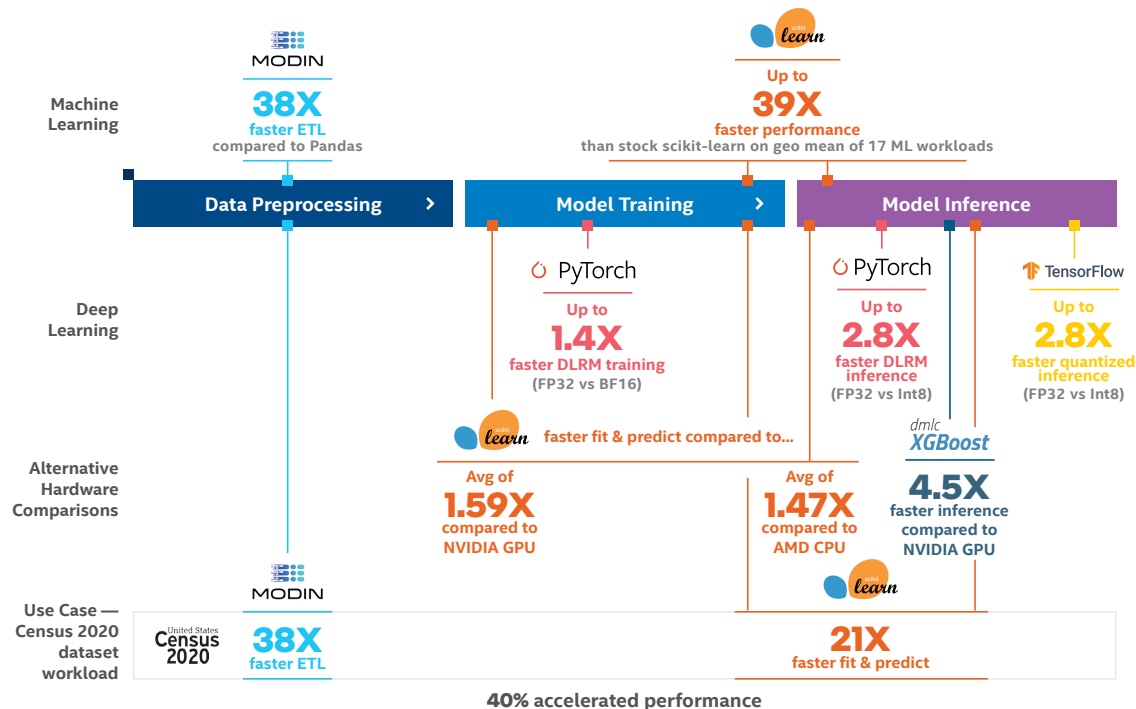
Core Python numerical and scientific packages deliver accelerated performance through the Intel Performance Libraries. Included in the package, Numba Compiler lets Python code use single instruction multiple data (SIMD) functionality and supports multicore execution on the latest Intel processors. Intel optimizations for NumPy and SciPy are also included, powered by oneAPI Math Kernel Library (oneMKL).

AI Analytics Toolkit Performance Benchmarks

AI Kit delivers value through the built-in acceleration that can be achieved by accessing optimized libraries equipped for top performance on Intel architectures. The cross-architecture capabilities of the AI Kit and oneAPI libraries minimize the need for code changes to support different platforms when performing operations from preprocessing to training of machine-learning and deep-learning models to inferencing within AI applications.

Examples of the productivity and performance gained from different Intel-optimized Python libraries and frameworks, part of the AI Analytics Toolkit, are illustrated in Figure 4. These benchmarks showcase acceleration achieved for popular AI workloads across each stage of the end-to-end data science pipeline.

Visit the [Benchmarks](#) section of this page to get the latest performance improvements available through the AI Analytics Toolkit.



See all benchmark configurations - <https://techdecoded.intel.io/resources/deliver-blazing-fast-python-data-science-and-ai-performance-on-cpus-with-minimal-code-changes>.

Each performance claim and configuration data is available in the body of the article listed under sections 1, 2, 3, 4, and 5.

Figure 4. Performance boosts achieved with the Intel oneAPI AI Analytics Toolkit.

OpenVINO Pro for Enterprise Overview

The Red Hat Marketplace now features OpenVINO Pro for Enterprise, maximizing the performance of computer workloads across Intel architecture-based hardware and supporting computer vision, as well as non-computer vision AI applications. Using the OpenVINO toolkit, deep learning models can be converted into optimized inference engines dynamically scalable to accommodate high-performance operations on thousands of Red Hat OpenShift nodes.

The toolkit includes a collection of pretrained models that perform a variety of inference tasks, such as emulation of human vision, automatic speech recognition, natural language processing, image retrieval, object detection and recognition, and more.

OpenVINO toolkit incorporates concepts and features from advanced neural networks, including convolutional neural networks (CNNs), as well as recurrent and attention-based networks. Optimization for Intel processors (CPUs) and graphics processing units (GPUs) gives data scientists a means to accelerate AI and deep learning inference from the edge network to the cloud.

Components of OpenVINO toolkit are powered by oneAPI libraries in the backend. Accelerated performance is achieved with OpenVINO Toolkit by utilizing oneAPI Deep Neural Network Library (oneDNN). This open-source, cross-platform performance library supplies the fundamental building blocks for deep-learning applications.

The capabilities of this professional toolkit are strengthened by proactive enterprise services, business-class support, and exclusive code releases.

"This [OpenVINO toolkit] is not a toolkit for faster training of your Deep Learning task, but for faster inference of your already trained Deep Neural model."⁴

Abhishek Dubey
Data Science Manager, Pune India Technology Center,
Schlumberger

OpenVINO Model Server

The model server automatically identifies new or updated versions of models stored in remote or local model repositories and monitors configuration changes. As models are served, any updates are hot swapped during runtime.

The OpenVINO Model Server handles trained models in a wide range of frameworks, including TensorFlow, Keras, and PyTorch/ONNX.

OpenVINO Toolkit Operator

Data scientists building container-based AI applications can take advantage of OpenShift Data Science component integration capabilities (see Figure 5). The Red Hat Operator for OpenVINO toolkit includes OpenVINO notebooks to simplify development, and provides access to the OpenVINO Model Server for deployment. With OpenVINO notebooks, data scientists can install toolkits within a container image. After optimizing the model for efficient performance on Intel architecture-based platforms, the model server can then be used to deploy inference services, fully scalable across complex network deployments consisting of thousands of nodes.

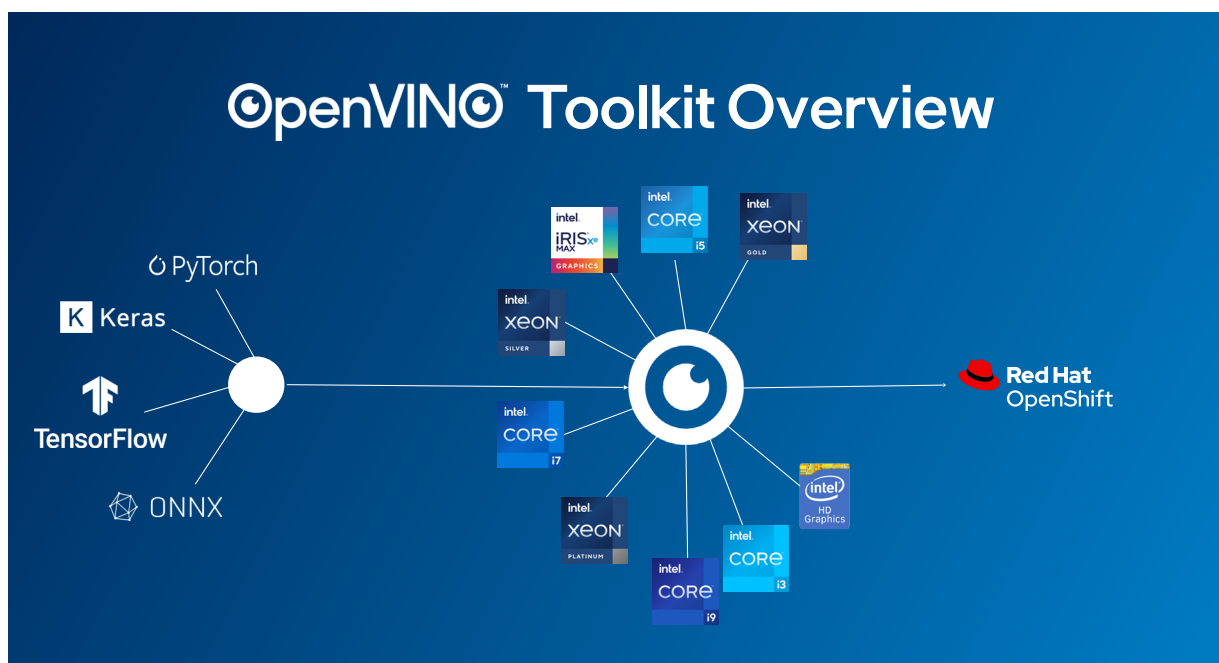


Figure 5. Overview of the OpenVINO Pro for Enterprise Toolkit.

Figure 5 shows the frameworks supported through OpenVINO, the Intel hardware components through which optimization is available, and the supported operating systems for which AI/ML applications can be built.

Enterprise Lifecycle Services and Software Updates

The value of OpenVINO Pro for Enterprise is extended through a wide range of available services that cover the full lifecycle of an AI application. This can include design, development, predeployment planning, application launch, and long-term AI application maintenance and improvement.

Subscribers to Red Hat OpenShift Science receive services from Red Hat, Intel, and partners, including 24x7 multichannel technical support.

OpenVINO Toolkit Performance Benchmarks

Performance benchmarks of AI and deep learning applications optimized by the Intel Distribution of OpenVINO toolkit demonstrate the substantial acceleration gains achieved across a diverse range of workloads. Benchmarks evaluate inference performance based on these factors: throughput, value, efficiency, and latency. Model server benchmarks use a measurement methodology based on two Ethernet-linked hardware platforms featuring multiple clients accessing a single server.

Visit these pages for detailed benchmarking results:

- [Intel Distribution of OpenVINO toolkit](#)
- [OpenVINO Model Server](#)

The data results provided help determine the Intel architecture-based hardware best suited for particular AI applications and solutions.

OpenVINO Toolkit Contents

The contents of the toolkit include:

Deep learning model optimizer

This cross-platform tool imports models and optimizes them for high-performance inferencing. Support includes models that were trained in familiar frameworks, including TensorFlow, PyTorch, and ONNX.

Deep learning inference engine

The unified API targets optimized performance across a broad slate of processors and accelerators, including Intel Xeon processors, Intel Core processors, Intel Iris® XE and Intel Iris Xe Plus graphics.

Inference engine samples

This collection of basic console applications demonstrate the best practices for adding inference engine operations in AI applications.

Post training optimization tool

This tool is designed to calibrate a model and execute it using INT8 precision.

Other tools

Additional tools for enhancing workflow with models include:

- Benchmark tool for estimating inference performance on a particular model
- Accuracy Checker tool that calculates the accuracy of models before and after optimizations are applied

Open Model Zoo

This collection of demos based on open-source templates provides shortcuts for implementing different types of deep learning projects. A repository for the Open Model Zoo includes documentation for working with the pretrained models.

Additional tools available through the Open Model Zoo include:

- Model Downloader provides a command-line tool to download pre-trained models from Open Model Zoo
- Model Converter converts Open Model Zoo models from frameworks to the OpenVINO IR format

[Visit the OpenVINO Toolkit Overview](#) for more information about the kit and its specific components.

Intel oneAPI Ecosystem

Designed to be a cross-industry, open, standards-based, unified programming model, oneAPI delivers a common developer experience across accelerator architectures—for fast application performance, enhanced productivity, and exceptional innovation. The worldwide Intel oneAPI ecosystem brings together more than 50 leading companies, research organizations, and universities that have endorsed oneAPI, many of which have taken advantage of Intel's implementation of oneAPI, using the oneAPI libraries and toolkits for experimentation, research, and commercial deployments.

Toolkits, including AI Analytics Toolkit and OpenVINO Toolkit, are part of the oneAPI ecosystem, in the middleware and frameworks section of the overall software stack. They utilize oneAPI libraries under the hood to maximize performance across Intel CPU and XPU architectures.

Figure 6 shows the layers encompassed by the full oneAPI product. The organizations involved in the oneAPI initiative are shown on the [oneAPI Ecosystem Support page](#).

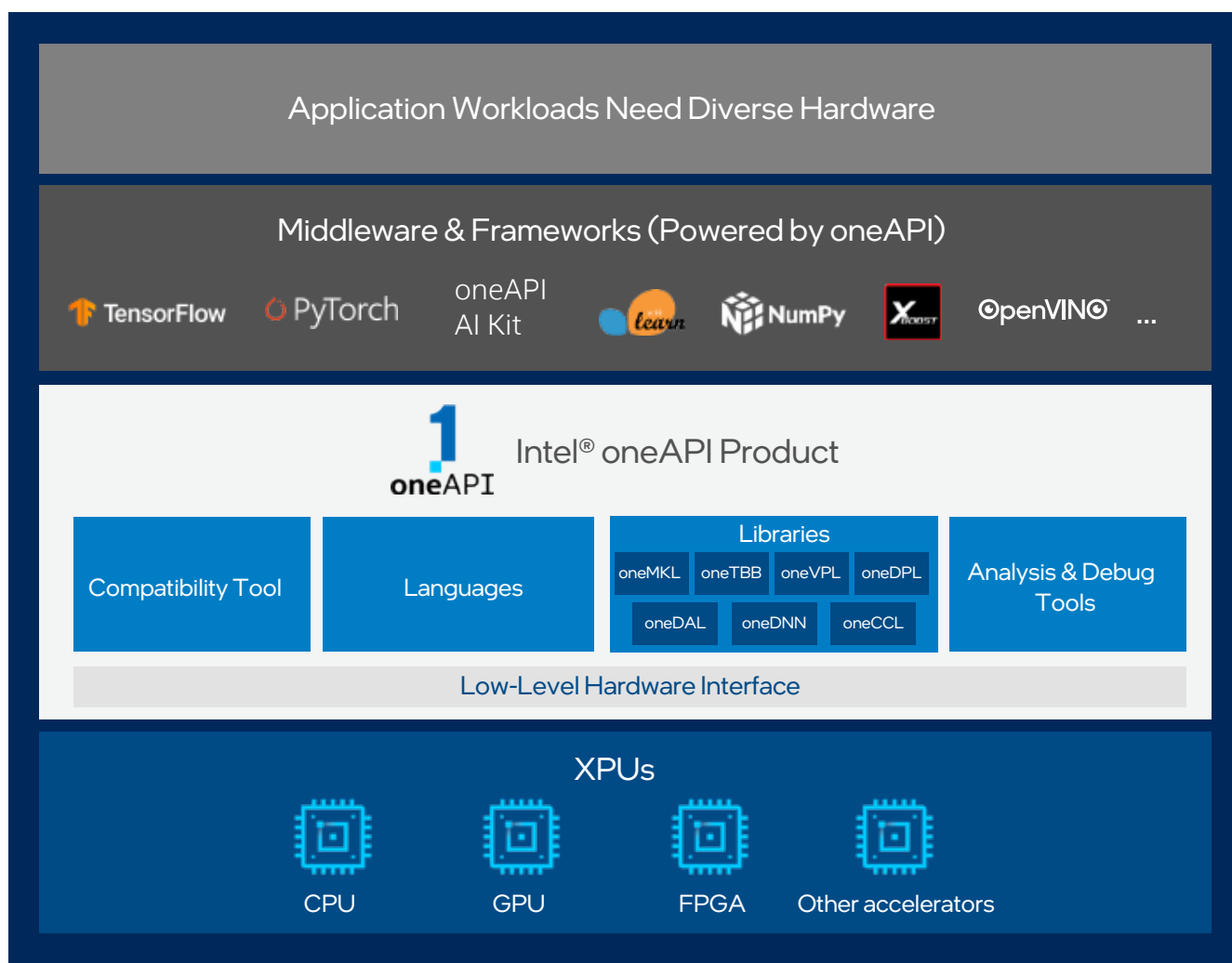


Figure 6. Intel oneAPI ecosystem includes compilers, libraries, and porting, analysis, and debugger tools.

Extending the XPU Vision to Data Science Solutions

The Intel vision of supporting an extensive mix of architecture within a unified software experience has been realized with the Intel oneAPI toolkits, giving data scientists, developers, and researchers a simplified cross-architecture programming model aimed at performance and interoperability. Through the oneAPI model, data scientists can select an architecture that is well-suited for a particular AI solution and then reuse the code, without rewriting it, for another architecture and platform. In this manner, solutions can be developed that exploit the performance of the underlying hardware while minimizing software development and maintenance costs.

Conclusion

Red Hat OpenShift Data Science with the addition of the Intel oneAPI AI Analytics Toolkit and Intel OpenVINO Pro for Enterprise Toolkit offers a winning combination for data scientists, researchers, and programmers working on AI and ML applications. With easy access to sample code, proven models, deployment tools, and a rich assortment of components from the Red Hat ecosystem, data scientists and intelligent application developers gain streamlined capabilities for building, training, and deploying ML models on a managed cloud service platform.

The platform architecture encourages AI/ML experimentation and innovation, as well as simplifying integration into intelligent applications. Open source software and ISV-certified software available through the

Red Hat Marketplace fills in any gaps in required pipeline tools or analytical components that might be needed to perform specialized work or address unique application requirements.

The collaborative efforts of Red Hat and Intel engineering teams have produced a platform ideally suited for data scientists to advance the work within a supportive and robust environment with access to the latest technologies.

Figure 7 shows the components that are included in the current release of Red Hat OpenShift Data Science.

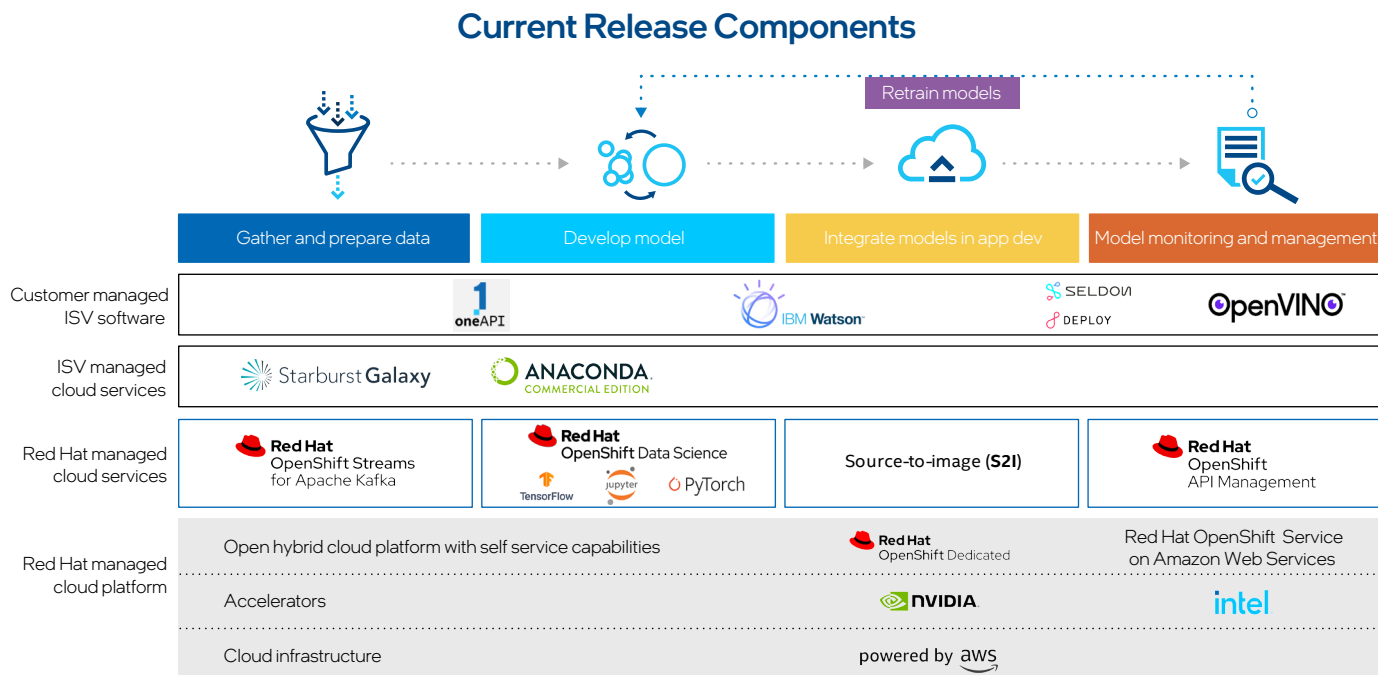


Figure 7. Components that are included in the current release of Red Hat OpenShift Data Science.

Learn More

Intel oneAPI AI Analytics Toolkit

For more information about building fast, efficient AI/ML applications for data science, visit [the AI Kit page](#).

[Learn more ›](#)

Intel oneAPI AI Analytics Toolkit Operator

Available through the Red Hat Ecosystem Catalog, this toolkit provides interoperable AI software tools across the Red Hat OpenShift pipeline.

[Learn more ›](#)

Intel Distribution of OpenVINO Toolkit

Discover techniques to enhance the development and deployment of data science projects using this toolkit.

[Learn more ›](#)

OpenVINO Pro for Enterprise by Intel

Unlock support, added services, and exclusive access to future code releases through the Red Hat Marketplace.

[Learn more ›](#)

Red Hat OpenShift Data Science

Learn more about the ways that Red Hat OpenShift Data Science provides a streamlined pipeline for launching AI/ML applications on a container platform.

[Learn more ›](#)

About Red Hat

Red Hat is the world's leading provider of enterprise open-source solutions, including high-performing Linux, cloud, container, and Kubernetes technologies. Visit redhat.com for more information.

“With the gold release of our oneAPI toolkits, we have extended the developer experience from familiar CPU programming libraries and tools to include our vector-matrix-spatial architectures.”⁵

– Raja Koduri,
Intel Senior Vice President,
Chief Architect, and General Manager of
Architecture, Graphics, and Software



1. Watson, Sophie. *Introducing Red Hat OpenShift Data Science*. Red Hat Blog. April 2021.
<https://www.redhat.com/en/blog/introducing-red-hat-openshift-data-science>

2. Red Hat Summit conversation. June 2021.
<https://events.summit.redhat.com/widget/redhat/sum21/sessioncatalog/session/1620405644376001RkhP>

3. *Deliver Blazing-fast Python Data Science and AI Performance on CPUs—with Minimal Code Changes*. Tech.Decoded.
<https://techdecoded.intel.io/resources/deliver-blazing-fast-python-data-science-and-ai-performance-on-cpus-with-minimal-code-changes/>

4. Dubey, Abhishek. *A quick intro to Intel's OpenVINO toolkit for faster deep learning inference*. Towards Data Science.
<https://towardsdatascience.com/a-quick-intro-to-intels-openvino-toolkit-for-faster-deep-learning-inference-d695c022c1ce>

5. *Intel Executing toward XPU Vision with oneAPI and Intel Server GPU*. Intel Newsroom. November 2020.
<https://www.intel.com/content/www/us/en/newsroom/news/xpu-vision-oneapi-server-gpu.html>

Performance varies by use, configuration and other factors. Learn more at [www.Intel.com/PerformanceIndex](https://www.intel.com/PerformanceIndex). Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Intel is committed to respecting human rights and avoiding complicity in human rights abuses. See Intel's [Global Human Rights Principles](#). Intel® products and software are intended only to be used in applications that do not cause or contribute to a violation of an internationally recognized human right.

Intel does not control or audit third-party data. You should review this content, consult other sources, and confirm whether referenced data is accurate.

Intel technologies may require enabled hardware, software or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.

Red Hat, the Red Hat logo, and other Red Hat marks are trademarks of Red Hat, Inc. or its affiliates.

© Intel Corporation. Intel, the Intel logo, OpenVINO, the OpenVINO logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

0921/RM/MESH/PDF