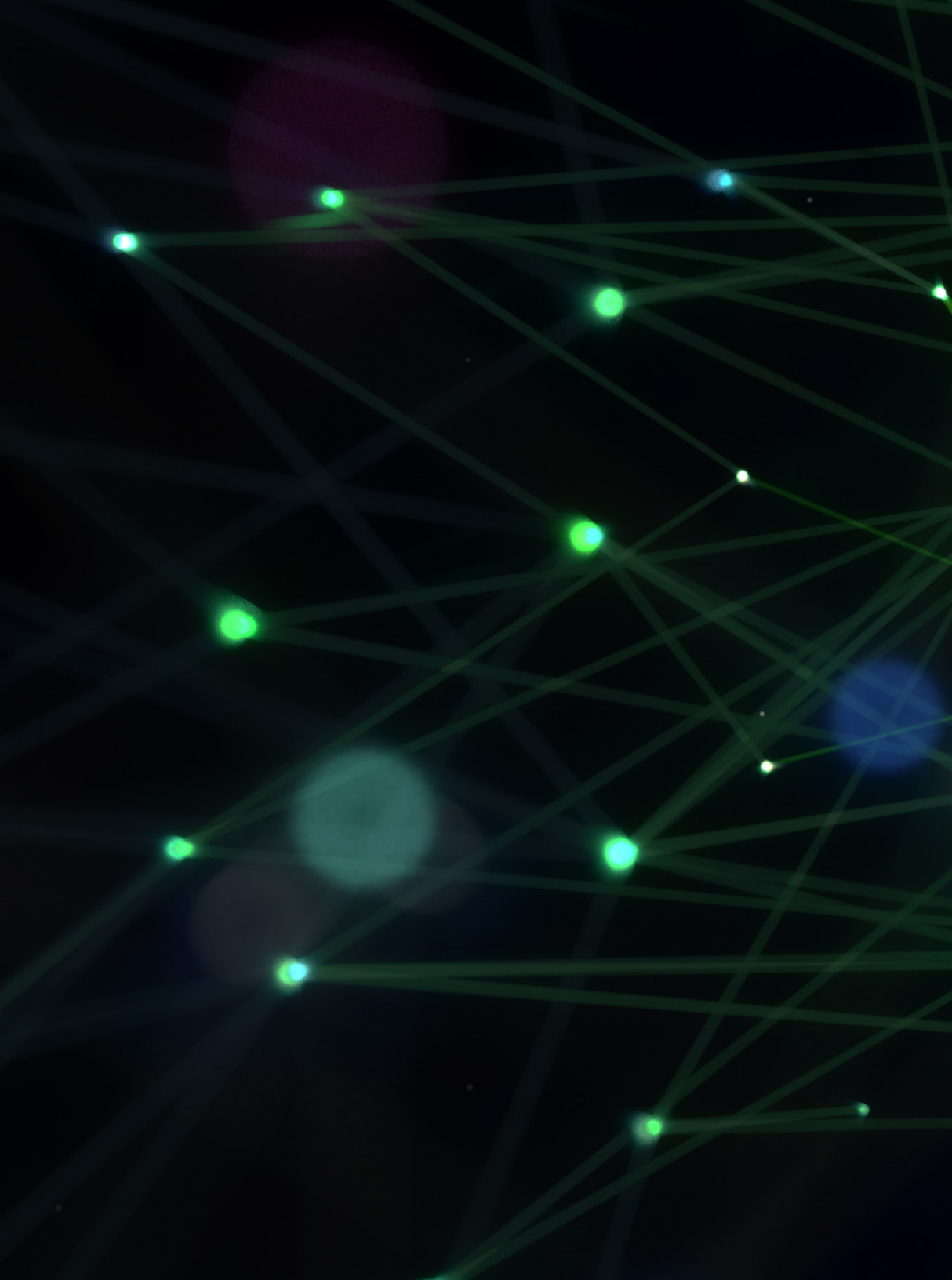# GETTING STARTED WITH AI SOFTWARE

NVIDIA

DRIVE BUSINESS INNOVATIONS WITH AI

Business leaders have long understood that the vast amounts of data generated in the technology-driven era had the potential to become the source of their success. However, without technology capable of sifting through, prioritizing, and refining massive amounts of data, organizations lack the resources necessary to extract meaningful conclusions. And the rate of data generation is growing exponentially, with billions of sensors in retail stores, hospital equipment, smart cameras, manufacturing machinery, and more.

Whether it's speech recognition in audio recordings, or information related to shopping behaviors, movies watched, or boxes moved, the data collected can be used to inform real-time actions and serve as use cases in conversational artificial intelligence (AI), natural language processing (NLP), computer vision, recommender systems, and more.

To derive insights from all this data, organizations need substantial compute power, coupled with developer tools and optimized algorithms. NVIDIA® GPU-powered laptops, workstations, and servers solve the compute requirements.. And now, with pre-trained AI models, application frameworks, and tools like NVIDIA Transfer Learning Toolkit and NVIDIA TensorRT™, your team can transform volumes of data into competitive insights.

PRE-TRAINED MODELS AND TRANSFER LEARNING

Instead of developing and training new models for specific skill sets, imagine the amount of time your team could save if they leveraged AI models previously built and trained.

Pre-trained models are state-of-the-art AI models that provide a specific skill, from recognizing boxes in a video feed in a warehouse to detecting words in audio recordings for NLP, so your teams can quickly benefit from AI. While these models are trained with generally available data, they need to be adapted to your specific use case to improve performance in your environment. Transfer learning enables your teams to fine-tune these pre-trained models to your specific needs. NVIDIA provides a wide selection of pre-trained models in **NGC**, our hub for GPU-optimized software, so you can kick-start your next AI project.

END TO END NVIDIA DEEP LEARNING WORKFLOW

TRANSFER LEARNING TOOLKIT

PRETRAINED MODEL | DATA | PRUNING | SCENE ADAPTATION | NEW CLASSES | OUTPUT MODEL
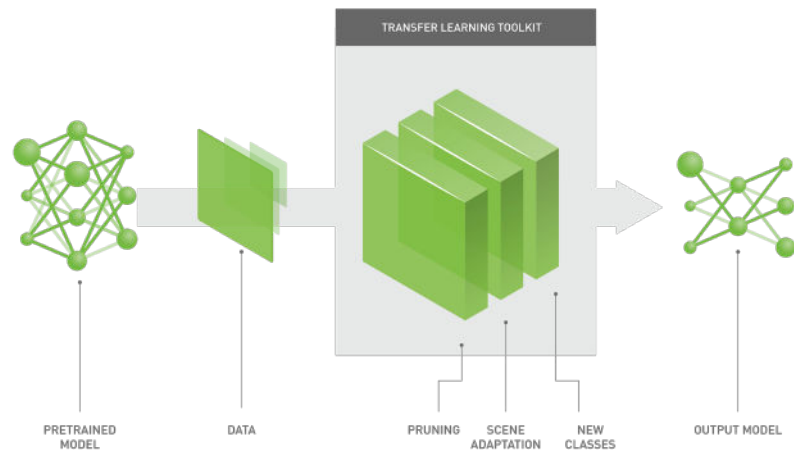
APPLICATION FRAMEWORKS

If your team is looking to build applications, application frameworks can help them get started more quickly. Application frameworks are reference applications that provide an example or template that demonstrates how several AI pre-trained or customized models can be used in your domain or industry. They can serve as a starting point for your specific application.

NVIDIA offers frameworks for several industries and use cases, including telecommunications, medical imaging, and genomics to intelligent video application (IVA) for retail, manufacturing, and cities. With pre-trained models and application frameworks, your team can rapidly build their own AI applications without the need to start from scratch.
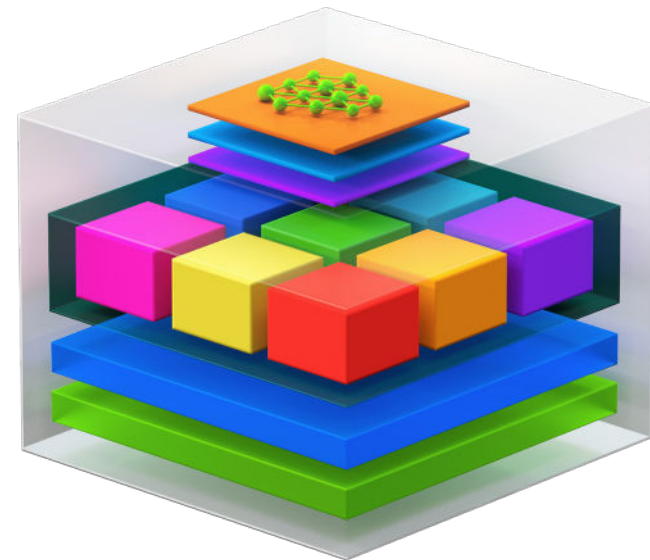
INDUSTRY SDKs

Industry SDKs provide the software and tools required by developers to build, test, and deploy AI solutions and integrate with applications unique to their industry. For example, NVIDIA Clara™ is used specifically in healthcare and NVIDIA Metropolis enables smart cities to leverage AI to boost efficient workflows from pre-trained models.

In healthcare, medical institutions can start with models pre-trained to classify MRI scans, use transfer learning to help models identify new diseases, or increase accuracy on existing diseases by fine tuning on new data as it becomes available.

Streaming analytics applications for smart cities can leverage pre-trained models for object detection and classification from camera video images. These pre-trained models can also be fine-tuned with company data for customized intelligent video analytics applications.
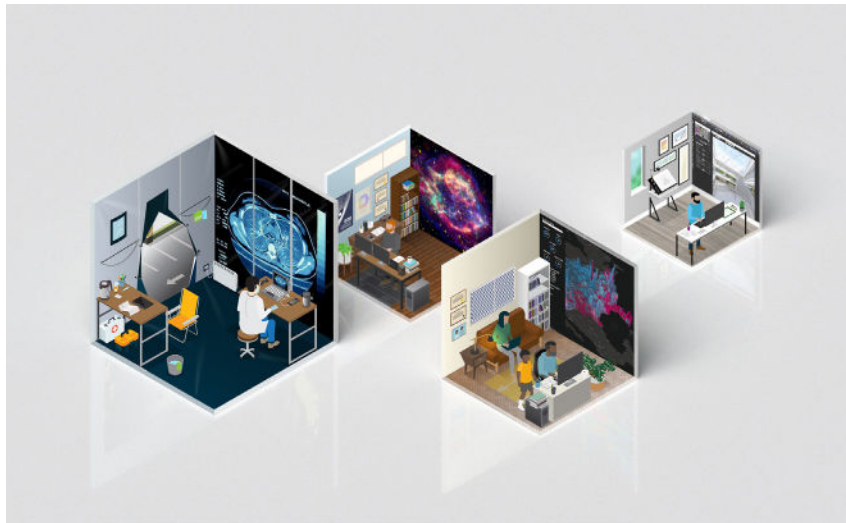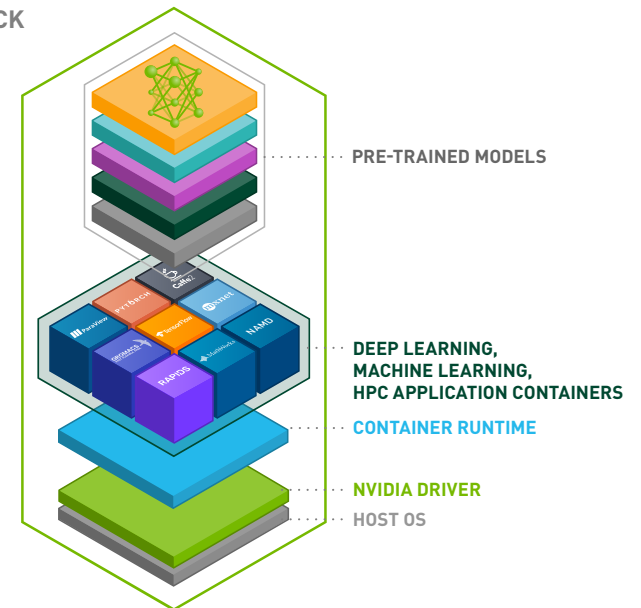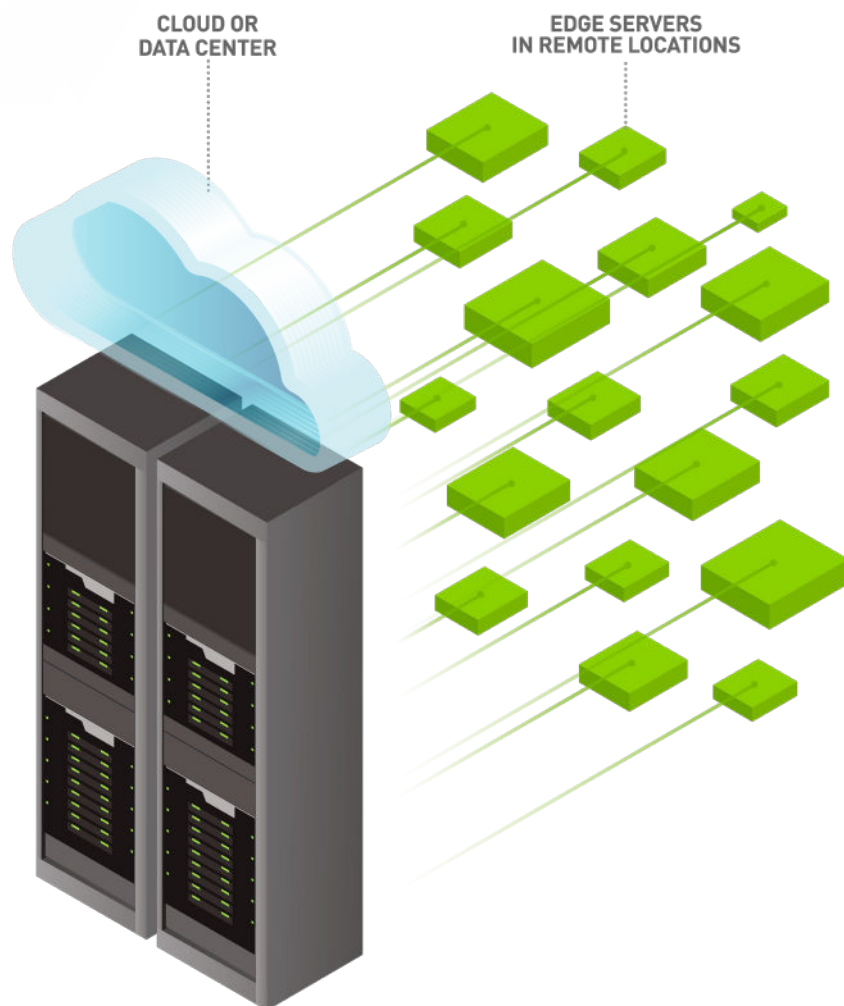

CONTAINERIZED AI SOFTWARE FOR EVERY INDUSTRY

AI is transforming every business and is proving critical in terms of getting products to market faster, reducing operational costs, and maximizing productivity. And, DevOps teams are moving to containerized applications as a deployment method on private, hybrid, or public clouds. To aid this effort, NGC hosts hundreds of AI software solutions, from frameworks to pre-trained models, to SDKs built by NVIDIA and other leading industry solution vendors.

These NGC offerings are performance-optimized and pre-configured to provide powerful and easy-to-deploy software, delivering faster results and allowing developers to build solutions from a tested framework with complete control.

NGC SOFTWARE STACK

PRE-TRAINED MODELS

DEEP LEARNING,
MACHINE LEARNING,
HPC APPLICATION CONTAINERS

CONTAINER RUNTIME

NVIDIA DRIVER

HOST OS

AUTOMATE DEPLOYMENTS WITH HELM CHARTS

CLOUD OR DATA CENTER

EDGE SERVERS IN REMOTE LOCATIONS

Helm charts automate software deployment on Kubernetes clusters, allowing users to focus on using—rather than installing—their software. NGC hosts Kubernetes-ready Helm charts that make it easy to deploy powerful third-party software. NGC also allows DevOps teams to push and share their Helm charts, so they can take advantage of consistent, secure, and reliable environments that speed up development-to-production cycles and encourage inter-team collaboration and asset reuse.

To simplify the installation and lifecycle of all necessary components when deploying applications on Kubernetes, NGC offers the NVIDIA GPU Operator Helm chart, which provides a suite of NVIDIA drivers, container runtime, device plug-in, and management software that IT teams can install on Kubernetes clusters, allowing them to remotely manage edge servers and run AI workloads.

GET STARTED

Quickly deploy AI frameworks with containers, get a head start with pre-trained models or model training scripts, and use domain-specific SDKs and Helm charts for the fastest AI implementations.

- **Learn More About NGC**
- **Explore NGC Now**

ADDITIONAL RESOURCES

- Provide hands-on and virtual training to your team through the **NVIDIA Deep Learning Institute (DLI)**. The DLI provides a variety of formats, from self-paced, online training for individuals to in-person workshops for teams, as well as downloadable course materials for university educators.
- Explore how to implement and scale AI workloads in the data center with the "**Introduction to AI in the Data Center**" course.
- **Download this e-book**: Find The Right Deep Learning Solution for Your Business
- **Download this e-book**: AI: 5 Steps to Get Started
- **Download this e-book**: Implementing AI Solutions for Every Industry

For additional information on models, frameworks, SDKs, applications, and more, visit **developer.nvidia.com**