

Taking the Enterprise AI Journey

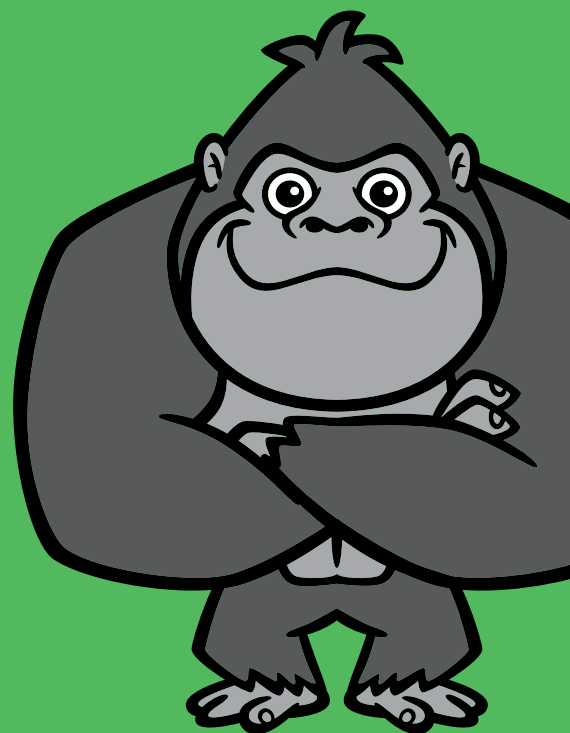
Ed Tittel

In This Paper

Getting more value from your data means adopting advanced technologies like artificial intelligence and machine learning. But the very thought of deployment, implementation, and management of these things can be daunting. That's where HPE comes in. They can help you with the scary stuff.



Hewlett Packard
Enterprise

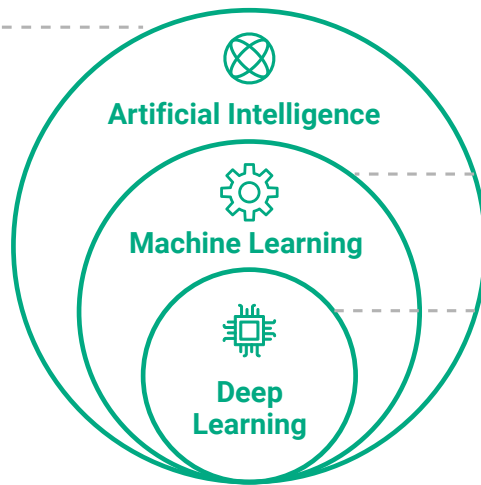


Contents

Key AI Terms: Machine Learning and Deep Learning.....	2
Confronting AI's Challenges.....	4
Take Your AI Journey with HPE.....	4
Moving AI Projects from Concept to Implementation.....	5

Artificial Intelligence (AI)

AI represents the field of technology focused on mimicking the reasoning and behaviors of humans using computer science. Modern AI narrows its focus to specialized areas of thought or action, such as virtual personal assistants (Siri or Alexa), autonomous vehicles, or automatic photo tagging on Facebook.



Machine Learning (ML)

Uses algorithms so computers can learn from examples (training data sets) without specific programming to drive selections and strategies.

Deep Learning (DL)

An area of ML that relies on neural networks—algorithms modeled after the biological structure of the brain—to perform complex reasoning tasks.

Figure 1: A visual comparison of AI technologies

Artificial intelligence (AI) has a unique ability to extract meaning from data. You may not know how to get to an answer, but if you can define what the answer should look like, AI can help. Further, AI can amplify human capabilities and understanding. This helps transform huge data collections (growing bigger all the time) into value, insight and action.

Organizations need training, the tools, and the people to help them get the most from their data, and to navigate the fast-moving and ever-changing AI arena. Here we explain how to get started in putting AI, machine learning, deep learning, and data analytics to work.

Key AI Terms: Machine Learning and Deep Learning

Two specific approaches to AI have emerged in the past decade that share a common focus on learning: one is called “machine learning,” and the other “deep learning.” Their relative position in the general field of AI is depicted in **Figure 1**.

UNDERSTANDING MACHINE LEARNING (ML)

ML starts with a model of how things work, then uses algorithms that can learn. This learning is cycled back into the model to improve its representational value and the understanding it conveys. ML doesn't rely on

detailed specific instructions on what to do, and how to do things—rather, its algorithms compare the results obtained by using various methods and making specific choices. Over time, by repeatedly choosing methods and options that produce the best results, the model evolves and improves.

ML usually begins with a training phase. This starts from a set of specially formulated data sets built around individual features and related data, used iteratively to create and then refine a basic set of understandings and selections.

AI/ML has proven to be of significant value and use across a broad range of industries.

ML uses specific categories for knowledge, such as *classification* and *regression*, to construct a model. Classification lets the model decide if data belongs to a specific class (“this is a sandwich” and “this is not a sandwich”). Regression depends on measurement and mathematical analysis to estimate values that depend on specific inputs (for example, a hypothetical investment portfolio's earning can be estimated based on inputs that include an initial balance, number of investments, returns on those investments, and other factors).

ML also employs a mathematical technique called clustering analysis. This identifies common patterns or structures observable within data sets. These clusters capture the understanding that certain data elements occur in concert or in connection with others in predictable ways.

THE MOVE TO DEEP LEARNING (DL)

ML uses algorithms and data sets to create and refine its own self-generated understandings and techniques. DL layers complex hierarchical models intended to reflect or represent human thought processes. For example, DL can use enormous data sets to offer personalized recommendations for consumption based on (a) a target buyer's prior choices and selections, (b) the buyer's expressed interests, and (c) the buying patterns of a large population, filtered to focus in on other buyers who share choices, selections, and interests with the target buyer. DL also works for applications related to computer vision, natural language processing, automated translation of human languages, speech recognition, text-to-speech conversion, and fraud detection (especially for anomalous transactions based on suspect locations, vendors, or purchase activity).

These collections of models are known as neural networks. They provide recommendations to target buyers based on their prior behavior and general purchasing activity among the total population of buyers. Newegg, Netflix, and other companies make daily, effective use of DL tools in their businesses. They are already using DL to boost their customers' buying activities and thus their bottom lines.

A typical DL dataflow is depicted in **Figure 2**. It shows how DL works in training mode (as it builds the associations and patterns it needs to make effective recommendations; path shown in light blue) and in production or inference mode, as it takes in inputs, transforms them into its various data sets, and uses that information to make recommendations by inference.

BENEFITS OF AI/ML

As we have seen, AI/ML has proven to be of significant value and use across a broad range of industries. It's especially insightful in industries with lots of customer interactions (for example, purchasing, order management, service and support delivery, account access and management, customer-supplied ratings and rankings data, unstructured [text] reviews and feedback, and more).

Too often, companies and organizations must manage a mix of systems and solutions, either poorly integrated or completely disjointed.

DL is particularly useful in industries such as banking, retail, financial services, credit card processing, and securities or commodities trading. All of these industries involve items of value that might be lost, hacked, or stolen, and must confront the possibilities of identity

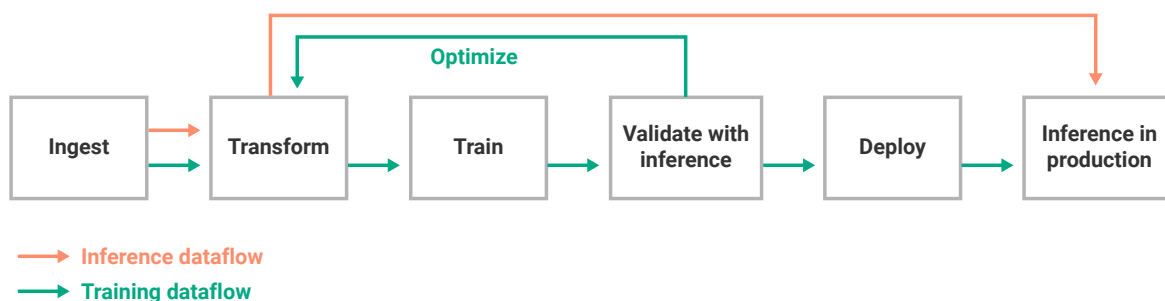


Figure 2: Deep learning dataflow

or data theft. AI/ML is also meaningful for healthcare, where it helps providers comply with privacy and confidentiality needs, and patients obtain improved treatment outcomes based on personalized medicine.

Confronting AI's Challenges

With unprecedented volumes of data to sift, companies and organizations now find themselves navigating vast, churning seas of information. ML and DL can help them put this data to good use, and to extract meaningful and actionable insights. But organizations often find themselves dealing with significant challenges to get there.

Organizations face numerous computing challenges when working with AI. These include the enormous volumes of data involved, which often require intelligent and faster networking capabilities (to move such volume) and added storage capacity (to store it somewhere accessible). But AI also poses internal organizational challenges, as well.

HPE can help with the skills, knowledge, and tools required to put AI to work in your organization, no matter what stage of the AI journey it might currently be on.

From a technology perspective, organizations may face a lack of tools to ingest, analyze, and operate on their data. This presents a complex, often daunting task—namely, how to deploy AI effectively and usefully within the organization. This can be exacerbated in many organizations because AI skills and knowledge are scarce and precious. There's a critical shortage of skills, talent, and expertise in adopting AI and achieving actionable insights. In fact, four in 10 companies mention [lack of analytical skills](#) as a key challenge hindering organizational growth and success.

Also, AI lives within a dynamic, fast-changing open source ecosystem. This can be daunting because of the volume and pace of change. Organizations must wade through a horde of niche companies and ISVs competing for AI market- and mindshare. Some of these players are HPE partners (for example H2O.ai and Dataiku), others commercialize open source frameworks. New open source frameworks and tools emerge at a frantic pace, with new versions popping up even faster. It's a case of too little time, and too many choices. And then, there are similar issues after adoption, when it comes to building, testing and deploying AI-based solutions.

Finally, some organizations drive digital transformation as an overriding imperative. Among other things, this means using AI as an innovative tool for competitive advantage. Keeping up is a major challenge: HPE can help. But first, organizations must acknowledge where they are in their AI journeys. Some might be early stage (still thinking about it, not sure where to begin). Others may have just started down the AI road with a use case identified and a Proof of Concept (PoC) underway or completed. The most advanced might already be running AI in production, but need to scale its use across their organization. They will want to maximize efficiency in their use of AI tools and make sure they've got the right infrastructure to support this mission.

Take Your AI Journey with HPE

HPE can help with the skills, knowledge, and tools required to put AI to work in your organization, no matter what stage of the AI journey it might currently be on. Look to HPE to provide expert insight, advice and resources to help all the way from exploration of options and alternatives, through experimentation with a test use case and PoC, into production use of AI in your organization.

Accelerated by NVIDIA solutions, HPE can take your organization all the way to extending AI across the entire infrastructure, wherever it makes sense to do so. In fact, AI is driving a fundamental shift in computing architectures from hyperconverged infrastructures (HCIs) to accelerated disaggregated infrastructures (which incorporate GPU acceleration and high degrees

of parallelism to greatly boost processing power while decreasing job run-lengths). NVIDIA and Mellanox are at the epicenter of this shift, with HPE bringing them into appropriate usage scenarios and deployments. The rewards will include increased innovation and productivity, improved insights into growth and new opportunities, and a better appreciation for the potential value that your data can deliver.



HPE TEAMS UP WITH NVIDIA ON AI PROBLEM SOLVING

HPE has teamed up with NVIDIA to drive DL and AI adoption across the enterprise, with an emphasis on GPU-accelerated compute and consultative expertise to aid enterprises in developing AI solutions. This partnership brings enormous strengths to interested enterprises, including proven AI expertise, the world's largest DL ecosystems, and a variety of powerful, ready-to-use AI software frameworks. Together, HPE and NVIDIA work with industries and customers to exploit AI and DL through online and instructor-led training and workshops, reference AI architectures, and benchmarking on NVIDIA GPU accelerated applications to enhance time to value, and improve AI project outcomes. The NVIDIA collaboration takes maximum advantage of NVIDIA GPU's superior parallel processing, which, when combined with HPE's computing and infrastructure innovations, streamline and hasten the process of obtaining real-time insights from an enterprise's DL initiatives. For more information, please consult [HPE and NVIDIA Deep Learning Collaboration](#).

Moving AI Projects from Concept to Implementation

For any organization to begin an AI journey, it must traverse a sequence of predictable steps. Before any technical work begins or any technology gets touched, it

starts with the involvement of, and ultimately support and sponsorship from, the senior management team.

A technical lead will have to pitch the need for investment and its potential to deliver value, insight, and possible benefits and advantages to senior staff. But the journey cannot really get underway until management buys in, and a sponsor emerges. The next step is to assemble (and fund) a project team to get the ball rolling.

The first real hurdle involves choosing one or more problem domains or focus areas around which to construct at least one AI project—select and analyze any number of possible use cases, then debate them within the project team, choosing a small number worth pursuing further. Next, the project team must establish the kinds of insights and decisions that the AI projects should deliver, and identify those projects' desired outcomes.



DOING AI USE CASES PROPERLY

At HPE's content portal, [enterprise.next](#), you can find a terrific article on the ins and outs of selecting and constructing use cases for AI projects. The piece comes from Dana Gardner, and is entitled "[AI Use Cases—and How To Do Them Right](#)." It takes readers through a conversation with two members of the HPE Pointnext Services team: Andy Longworth, senior solution architect, and Iveta Lohovska, data scientist. They discuss how to choose tasks and technology areas worthy of consideration for AI use cases, and then how to use the right tools and techniques to formulate and implement them. Along the way they discuss a wide range of possibilities, and how best to use AI to carry them forward as workable and meaningful use cases.

After that, they can assess the characteristics of the data those projects will need to ingest, transform, and train from, so that resulting inferences, associations,

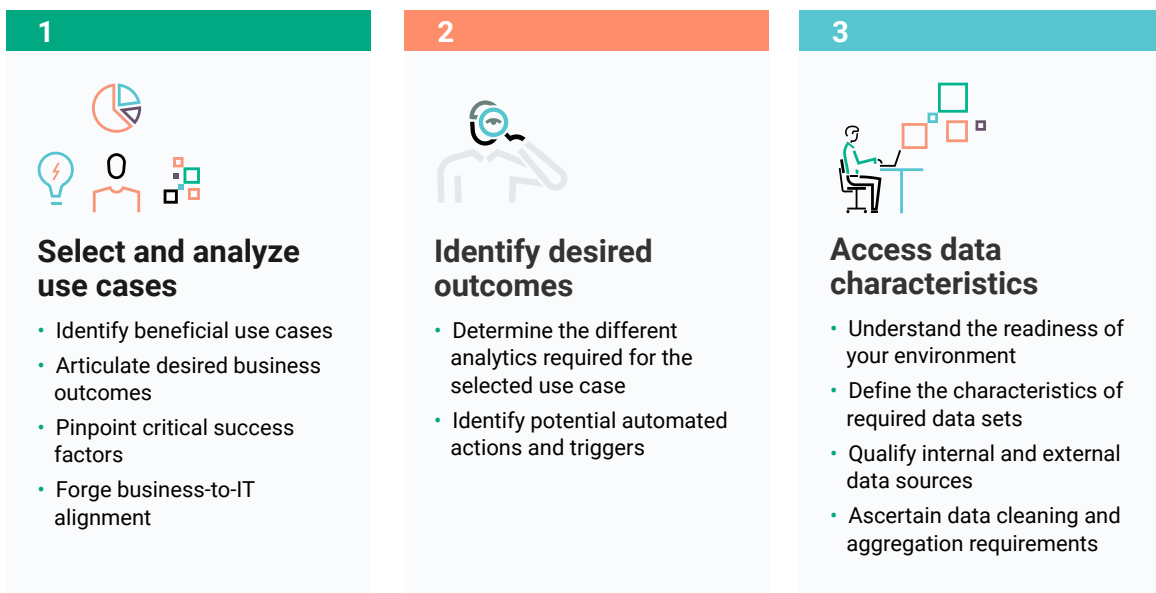


Figure 3: The three steps toward designing an AI project reflect best design practices

or recommendations can be checked against actual data sets and ultimately used to drive analysis and inference into production.

The team uses all this information to create a high-level roadmap of how they wish to proceed for designing and building PoCs and into possible pilots. **Figure 3** shows the three steps involved in working through this process.



GET DESIGN TRAINING STRAIGHT FROM THE PROS

The HPE Artificial Intelligence Transformation Workshop is a one-day workshop that teaches attendees how to build a data foundation, develop advanced analytics solutions, and start experimenting with AI for carefully chosen use cases. This is a great way for your AI project team to get underway with some expert input and guidance.

Though the odds may be small that any given project will actually go into production, teams need to plan out and understand what that might mean, and how

that would play out as a part of the overall upfront planning process.

Organizations pondering an AI journey should take careful note of the fact that the vast majority of AI projects do not go into production. To that end, various HPE offerings can help break PoC deadlocks, and help get AI projects underway (and, ultimately, into production). These include such things as AI configurations, training and analytical engagements, and examples that serve nicely as go-bys for those just getting started in AI.

The HPE and NVIDIA partnership makes it easy to stand up high-volume, processing-intensive AI projects in the cloud. This cloud structure makes it easy to pull the plug and walk away, should some particular project fail to pan out.

BUILDING PROOFS OF CONCEPT

Because of the learning curve, time, and effort involved in putting AI to work properly, most teams start with a single PoC as the next step in their AI journey. This involves choosing a framework, assembling and grooming data sets, then working to build the training activities that should ultimately produce the desired outcomes specified in your earlier design process.

HPE offers a pair of AI-ready configurations specifically tailored for building AI PoCs (or adding to their existing AI compute resources). These configurations include all the hardware and software necessary to build and run an AI PoC implementation. Both leverage NVIDIA GPU accelerators, and come ready to download and run NVIDIA GPU Cloud (NGC) software for analytics, DL, ML, and high-performance computing (HPC) projects.

The HPE and NVIDIA partnership makes it easy to stand up high-volume, processing-intensive AI projects in the cloud.

The two AI-ready configurations are designed to meet the demands of the training and inference dataflows mentioned earlier and support all NGC AI/ML libraries, including RAPIDS data preparation tools to speed how long it takes to start generating insights. In addition, NGC also makes containerized, optimized software available for all DL frameworks, along with pre-trained networks, Helm charts, and runs with inference TensorRT and Triton Inference Server. As an added bonus, the NVIDIA data center platform even accelerates over 700 different HPC applications. This lets your project team—including data scientists, ML engineers, developers, and researchers—focus on AI model development and deployment rather than the complex task of managing the open source software infrastructure.

ESTABLISHING AI PILOT PROGRAMS

Once a PoC implementation demonstrates enough value to be worth deploying, the AI team can begin making it available to its parent business unit for initial use. This means setting up a pilot program, during which the users will be brought more fully into the project, along with the team members who've taken the project to its current state of capability and readiness.



Please check these resources to help you make the most of your organization's AI journey, wherever it may be on that path:

- [HPE Makes AI Implementations Easier](#)
- HPE Centers of Excellence: HPE offers numerous research and development efforts focused around specific industries or use cases that include AI components. This HPE.com search for "[centers of excellence](#)" leads to all of them.
- [HPE Pointnext Services](#) for AI, data, and analytics is your one-stop source for training and expertise to accelerate your digital transformation, and guide you on your AI journey.

This may involve some scaling up, with additions to existing infrastructures to accommodate larger, more demanding workloads. It will also involve lots of back-and-forth between the users and the AI team to further refine, improve, and expand what the initial PoC brought to the table.

This is a stage during which the AI team and upper management should watch closely to see if the AI project delivers the value and insights in sufficient quantity and quality to justify its costs in terms of financial, staff, and IT resources going forward.

It will often be necessary to make some hard decisions based on results and user feedback at this stage of the AI journey. Many projects will be abandoned, while some projects will need rework and redesign (and may go back into PoC for another iteration). And indeed, some projects will demonstrate enough promise and value to justify rolling them into full-blown production. Putting AI into production is the next stage on the AI journey, and one well worth taking when PoCs or pilots show sufficient promise.

The [HPE Pointnext Services](#) organization can assist your organization with expertise and advice, plus design, implementation, and deployment assistance (including design reviews and recommendations). You don't have to travel alone—HPE is here to help at each step of your AI journey.

DEPLOYING PRODUCTION AI SOLUTIONS

Deploying AI solutions into full-blown production means making sure the infrastructure is prepared, with enough compute, storage, and networking resources to handle forecast workloads. Because AI is both compute- and data-intensive, the AI team and IT staff will have to adjust its design and growth metrics to accommodate AI's uniquely hungry consumption of those kinds of cloud/infrastructure resources.

Here's a sampling of HPE's various AI solutions designed to help organizations take PoCs into pilots and on to production:

- The [HPE Container Platform](#) is a cloud-based software-only platform that runs cloud-native and non-cloud native applications in containers.
- [HPE BlueData GPU-as-a-Service](#) delivers cloud computing resources specifically designed to support ML, DL, and advanced data analytics initiatives.
- [HPE Machine Learning Ops](#) (MLOps) is a container-based solution that brings DevOps-like speed and agility to ML workflows across the entire ML lifecycle (from sandbox experimentation with some chosen ML or DL framework, to model training on containerized distributed clusters, to deploying and tracking models in production).

Here again, HPE Pointnext Services can help your organization transition into full-blown production, and help you right-size your AI infrastructure to make sure production succeeds out of the gate, with ample planning and forecasting to handle increased demand, additional workloads, and ever-larger data sets and analytics tasks.

FINDING INNOVATION OPPORTUNITIES WITH AND FOR AI

As the old saying goes: “Nothing succeeds like success.” Once your line-of-business managers and upper management start to see results from AI projects, inevitably they will start asking for more. And as your AI team really starts to understand what AI can do with your organization's data, they will also start churning out new ideas and use cases that take advantage of those capabilities.

As your AI team really starts to understand what AI can do with your organization's data, they will also start churning out new ideas and use cases that take advantage of those capabilities.

The net result may be an explosion in demand, creativity, productivity, and value for your organization. Even for mature AI consumers, HPE stands ready to help them maximize its value, and make sure it remains lean and efficient in its consumption of resources of all kinds. The HPE Pointnext Services organization remains your partner in pushing AI to its limits, and squeezing maximum returns from your AI investment.

Learn more at [HPE AI Solutions](#). There you'll find a trove of information about HPE's AI-related products, services, and support, plus news and insights.