Unmatched Global Market Coverage

aiworld

Artificial Intelligence UPDATE

Machine Learning - Deep Learning - Software Research
Products - Technology - Applications - Industry News

AI World is a Global Publication of e2mos

Nov-Dec 2019

## Headlines:

- Dell New Servers for Modern Data Centers with 2nd Gen AMD EPYC
- Microsoft, Baker Hughes & C3.ai announce alliance for energy industry
- Marvell: NVIDIA CUDA-X AI & HPC Software Stack on Marvell ThunderX
- U.S. Defense: Without Effective AI, Military Risks Losing Next War

# Nvidia and ARM Strategic Partnership Paving the way for AI-enabled supercomputing

# In this Edition:

## Cover Story:

**Arm News: Paving the way for AI-enabled supercomputing** Arm continues to be the computing foundation for the convergence of AI, 5G, and IoT Arm and Nvidia in Strategic Partnership

## Headlines:

• **Microsoft alliance for energy industry** Baker Hughes, C3.ai, and Microsoft announce alliance to accelerate digital transformation of the energy industry

• **Dell Technologies Introduces New Solutions to Advance High Performance Computing and AI Innovation** At Supercomputing 2019, Dell Technologies is introducing several new solutions, reference architectures and portfolio advancements all designed to simplify and accelerate customers' high performance computing (HPC) and artificial intelligence (AI) efforts.

• **Marvell: NVIDIA CUDA-X AI and HPC Software Stack Now Available on Marvell ThunderX Platforms** Drives Industry-Leading Application Results for Energy-Efficient High Performance Computing

• **U.S. DEFENSE: Without Effective AI, Military Risks Losing Next War, General Says**
A next war against a near-peer competitor will be fast, chaotic and shockingly bloody, the director of the Joint Artificial Intelligence Center said. Air Force Lt. Gen. Jack Shanahan spoke at the National Security Commission on Artificial Intelligence in Washington.

**Daniel Dierickx CEO & co-Founder and Acting Chief Editor at e2mos**

40 Years Chips & Embedded Computing Business & Market Expertise
Together with my Partners we have built a PREMIER Global Customer Database

**Dear Reader,**

Here is your free copy of **A.I. World** one of our six High Focus e-magazines

**Your benefits:**

- Selected subjects HIGH FOCUS
- Easy reading
- TIME SAVER: includes many direct links
- UNIQUE: Dedicated Website for each e-magazine, remain visible for years
- FREE Distribution Worldwide
- FREE Subscription

**Discover our 6 e-magazines**
Click on the logos below

Semi Update World | Embedded Systems World
aiworld | Telecom COTS World (Broadband Broadcast IoT AI Convergence)
IoT World | ATCA World

**Editor/Publisher: e2mos**
WEB: www.e2mos.com
Contact: mgt@e2mos.com


# Paving the way for AI-enabled supercomputing

*Arm continues to be the computing foundation for the convergence of AI, 5G, and IoT*
*Arm and Nvidia in Strategic Partnership* ***

November 18, 2019 | Source: ARM

By Chris Bergey, SVP & GM, Infrastructure Line of Business, Arm

It's been a huge year for Arm and our ecosystem of infrastructure innovators as we work together to deliver the technologies and solutions that will ultimately enable the trillions of devices coming online by 2035. Arm continues to be the computing foundation for the convergence of AI, 5G, and IoT, fueling a massive change to today's data consumption models from the cloud to the edge. Just a few years ago, an Arm-based HPC was difficult to fathom for many. Today, Arm is becoming established as the compelling platform for AI and HPC innovation to build upon, scaling from the largest clusters to the emerging edge.

Earlier this year our strategic partner NVIDIA announced their intent to more broadly support Arm CPUs and make its full stack of AI and HPC software available to Arm partner platforms. This was a significant announcement for both companies. NVIDIA GPUs power some of the fastest supercomputers in the world, including the fastest, Summit at Oak Ridge National Labs. These systems will continue to play an important role in future exascale deployments where supercomputers will run at 5x – 10x the speed of today's best, but achieving this level of performance goes beyond the hardware.

Both Arm and NVIDIA know ecosystems and are great at building the software stacks needed for the success of our partners and end-users.  Enabling NVIDIA GPU's to run seamlessly in conjunction with Arm processors provides access to new markets for our HPC partners and opens further market opportunities for NVIDIA's market leading ML/AI technology. This important collaboration highlights our similar vision for enabling efficient compute from the datacenter to the edge and draws our companies closer in terms of collaboration on ecosystem goals.

Building on this momentum, today at SC'19, Jensen Huang, CEO of NVIDIA, introduced a technology blueprint for companies to quickly and easily build GPU-accelerated Arm-based servers. The reference platform, which consists of hardware and software building blocks, aims to enable more high-performance computing in a fast-growing range of applications across science and research. Furthermore, NVIDIA, Arm, and our collective ecosystem are already demonstrating  new proof points of solutions with Arm-based CPUs and NVIDIA GPUs.

This next step in delivering high-performance capabilities is significant because the combination of NVIDIA's CUDA-accelerated computing and Arm's energy-efficient CPU architecture will give the HPC community the advanced tools and technologies they need to enable compute from the edge all the way through to the largest HPC deployments. We are working diligently to deploy Arm technology all of the HPC space, addressing some of the world's most complex research challenges. We are very excited about the progress that has been made this year on joint enablement of the combined Arm and NVIDIA ecosystem. But, this is just the beginning.

Congratulations to the Arm and NVIDIA teams for a very exciting milestone achievement!

## About Arm

Arm technology is at the heart of a computing and connectivity revolution that is transforming the way people live and businesses operate. Our advanced, energy-efficient processor designs have enabled intelligent computing in more than 160 billion chips and our technologies now securely power products from the sensor to the smartphone and the supercomputer. In combination with our IoT device, connectivity and data management platform, we are also enabling customers with powerful and actionable business insights that are generating new value from their connected devices and data. Together with 1,000+ technology partners we are at the forefront of designing, securing and managing all areas of compute from the chip to the cloud.

*** **See News Nvidia** « NVIDIA Brings CUDA to Arm, Enabling New Path to Exascale Supercomputing »
in AI World Sep-Oct 2019 Click Here

# Baker Hughes, C3.ai, and Microsoft announce alliance to accelerate digital transformation of the energy industry

*Joint offerings will help make energy operations safer, cleaner, and more efficient*

HOUSTON, TX, REDWOOD CITY, CA, and REDMOND, WA. November 19, 2019 – Baker Hughes (NYSE:BKR), C3.ai, and Microsoft Corp. (NASDAQ:MSFT) today announced an alliance to bring enterprise artificial intelligence (AI) solutions to the energy industry on Microsoft Azure, an industry-leading cloud computing platform.

This alliance will enable customers to streamline the adoption of scalable AI solutions for the energy industry that help promote safety, reliability, and sustainability. It leverages the significant energy technology expertise of Baker Hughes, C3.ai's proven AI platform and applications, and the Microsoft Azure cloud computing platform. As a result, energy businesses will have a secure and reliable suite of enterprise-scale AI applications optimized to run on Azure. These solutions are tailored to address challenges across the entire value chain, from inventory optimization and energy management to predictive maintenance and process and equipment reliability.

"Shell supports the aim of this strategic alliance to improve efficiencies, increase safety, and reduce environmental impact through digital transformation, aligning seamlessly with our goals and ambitions," said Jay Crotts, Shell Group CIO. "Baker Hughes is one of our long-standing and valued partners in oilfield services and software development, and we use the C3.ai platform on Microsoft Azure to accelerate digital transformation across our business, helping to improve overall operations. The new technologies being developed will be critical as we all need to work together to reduce the net carbon footprint of the products and solutions that we put into society."

"The industry is adopting technologies that help manage the challenges and opportunities associated with the energy transition. The AI solutions offered through Baker Hughes and C3.ai deliver insights that can reduce risk and improve performance for operators as they navigate this transition," said Lorenzo Simonelli, Chairman and CEO, Baker Hughes. "With a singular offering that can accelerate digital transformation across the sector, energy businesses can now draw on the power of Microsoft's cloud, C3.ai's leading AI capabilities, and Baker Hughes's expertise in the energy industry."

"We are witnessing a massive market shift as oil and gas businesses undergo enterprise-level digital transformation to improve efficiencies and increase safety, while simultaneously reducing environmental impact," said Thomas M. Siebel, CEO, C3.ai. "With Microsoft's global reach and horizontal cloud platform, Baker Hughes's technology domain expertise, and C3.ai's industrial AI capabilities, organizations can rapidly improve core business operations and better serve customers with AI-enabled products and services. This strategic alliance is a complete game-changer for the industry."

The solutions will simplify the process of adopting AI capabilities for energy companies, starting with the shift of data management, storage, and compute onto Azure, through the development and enterprise-wide deployment of domain-specific AI applications built on the BHC3 AI Suite.

"For the energy industry, this is a time of significant transformation, and forward-thinking companies are exploring how to leverage technology to make their operations cleaner, safer and more efficient," said Judson Althoff, EVP, Worldwide Commercial Business, Microsoft. "By bringing together the domain expertise of Baker Hughes and the AI strengths of C3.ai to run on Microsoft's Azure cloud platform, customers can achieve new levels of digital transformation while advancing their sustainability commitments."

## About Baker Hughes

Baker Hughes (NYSE: BKR) is an energy technology company that provides solutions for energy and industrial customers worldwide. Built on a century of experience and with operations in over 120 countries, our innovative technologies and services are taking energy forward – making it safer, cleaner and more efficient for people and the planet.

## About C3.ai

C3.ai is a leading AI software provider for accelerating digital transformation. C3.ai delivers the C3 AI Suite for developing, deploying, and operating large-scale AI, predictive analytics, and IoT applications in addition to an increasingly broad portfolio of turn-key AI applications. The core of the C3.ai offering is a revolutionary, model-driven AI architecture that dramatically enhances data science and application development.

## About Microsoft

Microsoft (Nasdaq "MSFT" @microsoft) enables digital transformation for the era of an intelligent cloud and an intelligent edge. Its mission is to empower every person and every organization on the planet to achieve more.

# Dell Technologies Introduces New Solutions to Advance High Performance Computing and AI Innovation

- New improved Dell storage solutions address storage demands of high performance computing storage
- AI deployments simplified with new Dell EMC Ready Solutions for AI and reference architectures
- New 400GbE open networking switch introduced to support for compute and storage intensive cloud
- GPU & accelerator options added across the Dell EMC PowerEdge server portfolio to speed application

DENVER – Supercomputing 2019 – November 19 — At Supercomputing 2019, Dell Technologies (NYSE: DELL) is introducing several new solutions, reference architectures and portfolio advancements all designed to simplify and accelerate customers' high performance computing (HPC) and artificial intelligence (AI) efforts.

Continued adoption of AI to solve real-world problems has spurred growth across the HPC industry. According to a recent report from Hyperion Research, the global HPC industry is expected to grow by 7.1% to more than $39.2 billion by 20231 and HPC-server based AI is expected to rise by more than 29% from 2018 to 2023, reaching $2.7 billion in 2023.2

"There's a lot of value in the data that organizations collect, and HPC and AI are helping organizations get the most out of this data," said Thierry Pellegrino, vice president of HPC at Dell Technologies. "We're committed to building solutions that simplify the use and deployment of these technologies for organizations of all sizes and at all stages of deployment."

**Dell Technologies advances storage solutions for HPC**

Dell Technologies is expanding its portfolio of Dell EMC Ready Solutions for HPC Storage with new, turnkey solutions for ThinkParQ's BeeGFS and ArcaStream's PixStor file systems. Offering a combination of technology partners' software with Dell EMC hardware, networking and support, based on engineered and tested designs, Dell EMC Ready Solutions for HPC Storage simplify and speed deployment and solutions management.

Dell EMC Ready Solutions for HPC BeeGFS Storage, with ThinkParQ's software-defined parallel file system, speeds-up input/output (I/O)-intensive workloads with the ability to scale from small clusters to enterprise-class systems on premises or in the cloud.

Additionally, Dell EMC Ready Solutions for HPC PixStor Storage offers a high-performance parallel file system, enabling data management at scale with the ability to perform archive and analytics in place. The solution also includes a validated, scalable design with object, tape and cloud tiering capabilities using PixStor Ngenea along with PowerEdge servers, PowerSwitch and Mellanox® networking, PowerVault storage, supported by Dell Technologies deployment and support services.

With these ready solutions, customers have seen improvements in both performance and scale. For example:

- Using the Dell EMC Ready Solutions for HPC BeeGFS Storage for its two petabytes of all NVMe storage, The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia's national science agency, stands as one of the largest NVMe storage systems in the world, currently ranked 8th on the IO-500 list, a benchmark that showcases the world's fastest storage systems for HPC.3
- At Imperial College London, the Dell EMC Ready Solutions for HPC PixStor Storage can simultaneously serve its existing 2,500 node high-performance computing system and delivers over 20GB/s of throughput with no loss in interactive usage performance.4

Dell Technologies also is introducing expanded capacity for Dell EMC PowerVault ME4, offering 16TB HDDs that allow customers to scale to 4PB in 15U rack space – a 25% improvement in density, allowing for more HPC storage capacity in a smaller space. PowerVault brings scale, bandwidth and built-in data protection to HPC configurations with management simplicity.

**Dell Technologies simplifies path to AI with new Dell EMC Ready Solutions and reference architectures**

Dell EMC is expanding its Ready Solutions for AI portfolio with an all-new validated design for the Domino Data Science Platform. Developed in collaboration with Domino Data Lab, the Dell EMC validated design enables data scientists to develop and deliver models faster while providing IT with a centralized, extensible platform spanning the entire data science lifecycle –accelerating ideation and deployment.

To further simplify AI deployments, Dell Technologies also is introducing five new reference architectures for optimizing Dell EMC technologies with leading AI partners, such as those from DataRobot, Grid Dynamics, H20.ai, Iguazio and Kubeflow on Red Hat OpenShift. With these architectures, organizations can speed up deployment of AI solutions to modernize, automate and transform their data center using industry-leading Dell EMC converged infrastructure, servers, storage and data protection technologies.

... to next page

# Dell Technologies Introduces New Solutions to Advance High Performance Computing and AI Innovation

... from previous page

These architectures are optimized for Intel Xeon Scalable processors and enable organizations to speed up the deployment of AI solutions for training and inference to modernize, automate and transform their data centers using Dell EMC converged infrastructure, servers, storage and data protection technologies.

**New 400GbE networking switch for compute and storage intensive workloads**

Joining the Dell EMC PowerSwitch Z-Series portfolio, Dell is unveiling the Dell EMC PowerSwitch Z9332F-ON, a 400GbE open networking switch designed for high performance workloads. As an open networking pioneer, Dell Technologies delivers on the promise of software-defined networking, making network operations more flexible, programmable and easier to manage.

The Dell EMC PowerSwitch Z9332F-ON is purpose-built for cloud service provider data center networks with intensive compute and storage traffic, such as HPC, AI and streaming video. The new switch also delivers four times the throughput, double the price performance and near double the power efficiency of existing 100GbE platforms.5

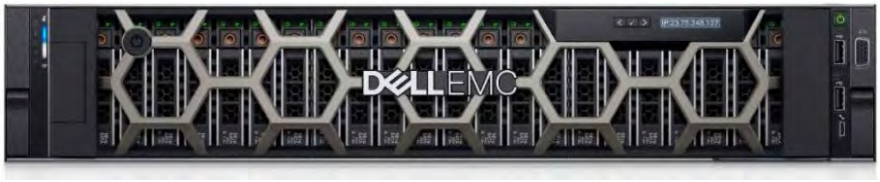
**Dell Technologies introduces new NVIDIA GPU and Intel FPGA options across its server portfolio**

Dell Technologies is also unveiling NVIDIA® T4 Tensor Core GPUs as a new accelerator option for the Dell EMC DSS 8440 server. With up to 16 accelerators, this offers high capacity, high performance machine learning inference with exceptional energy efficiency (70 watts per GPU). This is designed for multi-tenancy environments that need to share machine learning resources among users or departments.

Dell Technologies also is introducing new GPU and FPGA support for its PowerEdge servers, including:
● **NVIDIA Tesla V100S GPU** with up to 25% more bandwidth to communicate directly with Mellanox InfiniBand interconnect and PowerEdge Express Flash NVMe Performance PCIe SSDs for faster data transfers.
● **NVIDIA RTX GPUs,** designed to boost performance at a fraction of the cost, space and power requirements of a traditional render farm to produce high-quality content faster than ever before.
● **Intel FPGA Programmable Acceleration Card D5005** in Dell EMC PowerEdge R740xd and R940xa servers to boost inferencing, streaming analytics, video transcoding, and financial and genomic app• lications.

**Availability**

The Dell EMC PowerVault ME4 with 16TB HDDs will be available in early 2020.
Ready Solutions for AI - Design for Domino Data Science Platform will be globally available in December 2019.
AI Reference Architectures for Data Robot, Grid Dynamics, H20.ai and Iguazio have planned availability this quarter.
NVIDIA V100S GPUs will be available on Dell EMC PowerEdge servers in early 2020.
The Intel FPGA programmable acceleration card (PAC) D5005 will be available in early 2020.
All other announced offerings and updates are globally available now.




**Supporting quotes**

Harvey Newman, professor of Physics, Caltech

"We've been using Dell PowerSwitch Z Series products in both our network R&D and our HPC facility as part of the global Large Hadron Collider grid for years. We value the cost, performance, and reliability of the switches in high performance computing environments. Its Software Defined Network (SDN) support has enabled us to develop new distributed system paradigms and intelligent networks to better serve the high energy physics and other data intensive science disciplines. The new Z9332F-ON is a highly capable 400GbE open networking switch that can deliver the fat pipes necessary to allow us to extend our SDN and system developments to a new scale, in order to meet the needs of our next data taking run at the LHC in 2021-3, and beyond."

Nick Elprin, co-founder and CEO, Domino Data Lab

"It's been deeply validating to see Dell Technologies extend a validated Domino data science platform design to clients who are on a journey to become model-driven. We really value the first-hand perspective and experience Dell Technologies has embedded into this solution and look forward to helping global enterprises deliver central, reproducible, and measurable data science results with the Dell EMC Ready Solutions for AI."

**About Dell Technologies**

Dell Technologies (NYSE:DELL) is a unique family of businesses that helps organizations and individuals build their digital future and transform how they work, live, and play. The company provides customers with the industry's broadest and most innovative technology and services portfolio spanning from edge to core to cloud.

# NVIDIA CUDA-X AI and HPC Software Stack Now Available on Marvell ThunderX Platforms

## Drives Industry-Leading Application Results for Energy-Efficient High Performance Computing


Santa Clara | Nov. 18, 2019 – Marvell (NASDAQ: MRVL) today announced the availability of NVIDIA GPU support on its ThunderX® family of Arm®-based server processors. Following NVIDIA's June announcement to bring CUDA to the Arm architecture, Marvell has collaborated with NVIDIA to port its CUDA-X AI™ & HPC libraries, GPU-accelerated AI frameworks and software development tools to the ThunderX platform. The computational performance and memory bandwidth of ThunderX2®, Marvell's latest 64-bit Armv8-A based server processor, combined with the parallel processing capabilities of NVIDIA GPUs provide a compelling path to energy-efficient exascale computing.

Artificial intelligence (AI) and machine learning (ML) continue to become essential technology components to data center server requirements at the cloud and network edge. To address these evolving AI and ML workloads, as well as the most challenging and complex problems in science and research, supercomputers need processors that are optimized to provide cutting-edge throughput, application latency and power.

With an initial focus on computational science applications including GROMACS, NAMD, MILC and LAMMPS, the ThunderX2 configurations are demonstrating compelling performance with an enhanced ability to drive higher and more efficient combined application results in a GPU-enabled system.

"NVIDIA GPU support for our ThunderX2 server processor brings clear, differentiated value to meet the distinctive performance and power requirements of the exascale computing era," said Gopal Hegde, vice president and general manager, Server Processor Business Unit at Marvell Semiconductor, Inc. "The availability of NVIDIA GPU acceleration on the Arm architecture will further expand the ThunderX2 ecosystem for HPC, cloud computing and edge markets, spurring innovation across low level firmware through system software to commercial ISV applications."

"The availability of CUDA acceleration for ThunderX2 processors marks a significant milestone in bringing the power efficiency and high performance of the Arm architecture to the infrastructure market," said Chris Bergey, SVP & GM, Infrastructure Line of Business at Arm. "The breadth and depth of innovation across the ecosystem enables an easy migration path and robust support for existing and future GPU workloads from the edge to the cloud."

"NVIDIA GPU-accelerated computing on Arm provides customers worldwide with greater choice in building next-gen AI-enabled supercomputers," said Ian Buck, general manager and vice president of Accelerated Computing at NVIDIA. "Combining NVIDIA's unmatched platform for AI and HPC with Marvell's powerful ThunderX2 Arm-based server processors is already delivering impressive application performance."

### Ecosystem Support

ThunderX2 is the most widely supported **Armv8-A server processor with an ecosystem of over 100 partners** across commercial, open source and industry standards engagements. NVIDIA's full software suite support is enabling the acceleration of more than 600 HPC applications and AI frameworks on ThunderX2 systems.

"Our collaboration with Marvell enables us to support servers with the industry-leading performance of ThunderX2 with our SC8000 compute acceleration expansion platform, bringing data center AI capabilities to a host of edge applications," said Steve Cooper, CEO at One Stop Systems. "The SC8000 is the industry's first solution that incorporates NVIDIA Tesla GPUs with NVLink and Arm servers. The addition of Arm-based architecture into our solutions extends the value of use cases for AI on the Fly edge appliances for our customers."

"Red Hat and Marvell have a long history of collaborating in the Arm server ecosystem, helping to bring open, industry-wide standards to enterprise Arm architecture," said Chris Wright, senior vice president and chief technology officer at Red Hat. "Enabling NVIDIA GPUs on ThunderX-based systems paired with the CUDA-X SDK and libraries supports customer choice in terms of architecture for running HPC, AI and ML applications on top of Red Hat platforms."

"SUSE and Marvell have a successful track record of strong HPC collaboration as seen most recently in the Catalyst UK deployment," said Daniel Nelson, vice president, Products & Solutions Management at SUSE. "We are excited to build upon this collaboration with the addition of NVIDIA's GPUs and suite of HPC, AI and ML applications support on Marvell's ThunderX2. SUSE looks forward to driving additional optimizations and software infrastructure support in the Arm ecosystem and will be demonstrating GPU support at the SC19 industry conference in Denver."

### About Marvell

Marvell first revolutionized the digital storage industry by moving information at speeds never thought possible. Today, that same breakthrough innovation remains at the heart of the company's storage, processing, networking, security and connectivity solutions. With leading intellectual property and deep system-level knowledge, Marvell's semiconductor solutions continue to transform the enterprise, cloud, automotive, industrial, and consumer markets. To learn more, visit: https://www.marvell.com/

# Without Effective AI, Military Risks Losing Next War, General Says

Nov. 5, 2019 | BY David Vergun

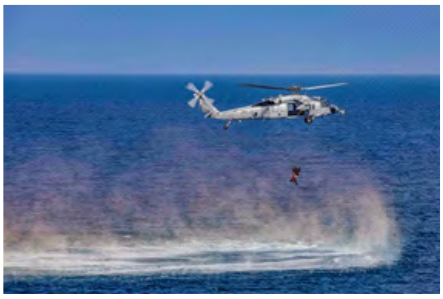
A next war against a near-peer competitor will be fast, chaotic and shockingly bloody, the director of the Joint Artificial Intelligence Center said.

Air Force Lt. Gen. Jack Shanahan spoke at the National Security Commission on Artificial Intelligence today in Washington.

The side with the best AI algorithms will put the other side at an extreme disadvantage, he said, particularly regarding the speed of battlefield decision making.

Shanahan described such a future battle as "algorithms vs. algorithms," with the best algorithm victorious.

In the future battle scenario, Shanahan said events will move so quickly that a traditional chain of command won't work. Junior, frontline troops will need to be empowered to make the decisions and to adjust AI algorithms on the fly. This decentralization of command entails higher risks and consequences, he said, but without it, "we risk losing the fight."



To get to the best AI, the department must rely on industry and academia, which are much further along in this endeavor than the DOD, he said.

Shanahan said there are lessons learned from Google's unwillingness to continue working with the DOD on Project Maven last year. The project had to do with AI's use in intelligence, surveillance and reconnaissance operations.

There needs to be a shared sense of responsibility and vision, along with trust and transparency from the DOD and industry, he said. "National security depends on it."

Another step to take in adopting the best of industry's AI is for service members to work directly in industry and academia and to bring AI experts from industry and academia into the DOD. "That's already happening but we need to scale that up," he said. "Peer-to-peer discussions and personal relationships matter."

Lastly, Shanahan said an important step that was taken last week was the Defense Innovation Board coming up with a set of AI ethical principles, which he said are excellent.

The DOD will act on the Defense Innovation Board's recommendations and then there will be deliberations on implementing them if the recommendations are accepted. "Implementation is not an overnight task," he said. The ethical use of AI by the military in training, research, product development and operations should inspire confidence in the industry that the department is making ethical use of this new technology.

"China and Russia didn't hold public hearings on the ethical use of their AI and I never expect them to do this," he added.