FREE Download www.aiworld.be

aiworld

Artificial Intelligence UPDATE

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

AI World is a Global Publication of e2mos

Q3-2017

■ Avitas Systems, a GE Venture, Partners with NVIDIA to Enhance AI for Robotic Inspection and Automated Defect Recognition

■ IBM and MIT to Pursue Joint Research in Artificial Intelligence, Establish New MIT–IBM Watson AI Lab

■ NVIDIA Partners with World's Top Server Manufacturers to Advance AI Cloud Computing

Foxconn, Inventec, Quanta, Wistron Using NVIDIA HGX Reference Architecture to Build AI Systems for Hyperscale Data Centers

■ The Race For AI: Google, Baidu, Intel, Apple In A Rush To Grab Artificial Intelligence Startups

■ Amazon Strategy Teardown: Building New Business Pillars In AI, Next-Gen Logistics, And Enterprise Cloud Apps

In this Edition

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

Dear Reader,

Here is your free copy of **AI World**, one of our five magazines published by e2mos.

Our aim is to provide you with relevant information.

Those five magazines are part of the e2mos « **Go-to-Market Platform** »

This GLOBAL Platform is a UNIQUE Set of Services for Telecom ICT, Video Broadcast, Embedded Computing, IoT and AI Vendors from Multicore Chips to Application-ready Systems & Rack Space Servers.

Our WORLDWIDE Services:
- Business Discovery
- Customer Meeting Setup
- Telemarketing
- Call Campaigns
- e-mailings Worldwide
- and our 5 e-magazines, each magazines has its **own Website** (see below).

Global Expertise & Added-Value
- Over 30 Years Global Expertise in those Markets
- we use a **UNIQUE Database** started in 1980 - **Daily UPDATES**

FREE and Direct
No Login, No Password
Just Click on the LOGO

aiworld

IoT World

TelecomCOTSWorld
Broadband Broadcast IoT Convergence

Embedded Systems World

ATCA World

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

Avitas Systems, a GE Venture, Partners with NVIDIA to Enhance AI for Robotic Inspection and Automated Defect Recognition

Avitas Systems uses NVIDIA DGX systems and AI expertise to innovate the inspection services industry while making oil and gas, transportation, and energy industries safer.

BOSTON, MA – SEPTEMBER 7, 2017 – Avitas Systems, a GE Venture, is partnering with NVIDIA to use some of the latest advances in artificial intelligence computing to optimize the use of robotics for inspection and better detect defects on industrial assets with advanced data analytics.

Avitas Systems can target specific points of inspection and develop paths to collect data in the form of images and video for a variety of robotics, including drones, robotic crawlers, and autonomous underwater vehicles (AUV). These paths, driven by 3D models, can be repeated from the same angles and locations. The paths' repeatability means a wide variety of images captured over time can be inputted into the Avitas Systems cloud-based platform, so advanced image analytics can detect changes and measure exact defects on an industrial asset, such as cracks and corrosion. The platform can also rate the severity of defects, oftentimes not visible to the human eye, allowing customers to determine when equipment needs to be replaced and enabling earlier resolution of potential issues.

The company is using NVIDIA DGX-1 and DGX Station systems for AI training involved in automated defect recognition. Avitas Systems data scientists build convolutional neural networks for image classification and generative adversarial neural networks to minimize the amount of work involved in labeling captured images. NVIDIA allows Avitas Systems to train software to process many different images and determine when it is ready to identify defects, following a variety of models.

Avitas Systems stores deep learning models in an AI Workbench, an innovative solution that can process inspection data in real-time and retrain the models to adapt to new use cases.

"Working with NVIDIA allows us to fully commercialize our cutting-edge, self-service AI Workbench, and we look forward to expanding its capabilities using the new NVIDIA DGX Stations with Volta," said Alex Tepper, Founder and Head of Corporate and Business Development at Avitas Systems. "With our workbench, our engineers can easily create and access new deep learning models that train the software deployed to recognize defects automatically at inspection sites.

Avitas Systems uses global expertise to push the boundaries of AI and inspection services.

"Avitas Systems is breaking new ground by bringing NVIDIA DGX Station beyond the deskside and into the field for the first time," said Jim McHugh, General Manager of DGX Systems for NVIDIA. "Using our latest DGX systems to help train robots and better predict industrial defects increases worker safety, protects the environment, and leads to substantial cost savings for companies."

For more developments, visit http://www.avitas-systems.com/, or follow on Twitter (@Avitas_Systems) and LinkedIn.

About Avitas Systems, a GE Venture

Avitas Systems is a GE Venture advancing the inspection services industry across oil and gas, transportation, and energy sectors through predictive data analytics, robotics, and artificial intelligence. Its solutions increase safety and decrease inspection costs by providing state-of-the-art robotic-based autonomous and semi-autonomous inspection management, smart scheduling, and a cloud-based platform that analyzes and stores comprehensive inspection data. Avitas Systems delivers advanced insights based on anticipated risk, boosting facility productivity. For more information, visit http://www.avitas-systems.com/, or follow on Twitter (@Avitas_Systems) and LinkedIn.

About NVIDIA

NVIDIA's (NASDAQ: NVDA) invention of the GPU in 1999 sparked the growth of the PC gaming market, redefined modern computer graphics and revolutionized parallel computing. More recently, GPU deep learning ignited modern AI -- the next era of computing -- with the GPU acting as the brain of computers, robots and self-driving cars that can perceive and understand the world. For more information, visit www.nvidia.com/ai.

IBM and MIT to Pursue Joint Research in Artificial Intelligence, Establish New MIT–IBM Watson AI Lab

IBM plans to make a 10-Year, $240 Million Investment in new lab with MIT to advance AI hardware and software and algorithms

Cambridge, MA - 07 Sep 2017: IBM (NYSE: IBM) and MIT today announced that IBM plans to make a 10-year, $240 million investment to create the MIT–IBM Watson AI Lab in partnership with MIT. The lab will carry out fundamental artificial intelligence (AI) research and seek to propel scientific breakthroughs that unlock the potential of AI. The collaboration aims to advance AI hardware, software and algorithms related to deep learning and other areas, increase AI's impact on industries, such as health care and cybersecurity, and explore the economic and ethical implications of AI on society. IBM's $240 million investment in the lab will support research by IBM and MIT scientists.

The new lab will be one of the largest long-term university-industry AI collaborations to date, mobilizing the talent of more than 100 AI scientists, professors, and students to pursue joint research at IBM's Research Lab in Cambridge—co-located with the IBM Watson Health and IBM Security headquarters in Kendall Square, in Cambridge, Massachusetts—and on the neighboring MIT campus.

The lab will be co-chaired by IBM Research VP of AI and IBM Q, Dario Gil, and Anantha P. Chandrakasan, dean of MIT's School of Engineering. IBM and MIT plan to issue a call for proposals to MIT researchers and IBM scientists to submit their ideas for joint research to push the boundaries in AI science and technology in several areas, including:

· • AI algorithms: Developing advanced algorithms to expand capabilities in machine learning and reasoning. Researchers will create AI systems that move beyond specialized tasks to tackle more complex problems, and benefit from robust, continuous learning. Researchers will invent new algorithms that can not only leverage big data when available, but also learn from limited data to augment human intelligence.

· • Physics of AI: Investigating new AI hardware materials, devices, and architectures that will support future analog computational approaches to AI model training and deployment, as well as the intersection of quantum computing and machine learning. The latter involves using AI to help characterize and improve quantum devices, and also researching the use of quantum computing to optimize and speed up machine-learning algorithms and other AI applications.

· • Application of AI to industries: Given its location in IBM Watson Health and IBM Security headquarters and Kendall Square, a global hub of biomedical innovation, the lab will develop new applications of AI for professional use, including fields such as health care and cybersecurity. The collaboration will explore the use of AI in areas such as the security and privacy of medical data, personalization of healthcare, image analysis, and the optimum treatment paths for specific patients.

· • Advancing shared prosperity through AI: The MIT-IBM Watson AI Lab will explore how AI can deliver economic and societal benefits to a broader range of people, nations, and enterprises. The lab will study the economic implications of AI and investigate how AI can improve prosperity and help individuals achieve more in their lives.

In addition to IBM's plan to produce innovations that advance the frontiers of AI, a distinct objective of the new lab is to encourage MIT faculty and students to launch companies that will focus on commercializing AI inventions and technologies that are developed at the lab. The lab's scientists also will publish their work, contribute to the release of open source material, and foster an adherence to the ethical application of AI.

"The field of artificial intelligence has experienced incredible growth and progress over the past decade. Yet today's AI systems, as remarkable as they are, will require new innovations to tackle increasingly difficult real-world problems to improve our work and lives," said Dr. John Kelly III, IBM senior vice president, Cognitive Solutions and Research. "The extremely broad and deep technical capabilities and talent at MIT and IBM are unmatched, and will lead the field of AI for at least the next decade."

"I am delighted by this new collaboration," says MIT President L. Rafael Reif. "True breakthroughs are often the result of fresh thinking inspired by new kinds of research teams. The combined MIT and IBM talent dedicated to this new effort will bring formidable power to a field with staggering potential to advance knowledge and help solve important challenges."

... to next page

IBM and MIT to Pursue Joint Research in Artificial Intelligence, Establish New MIT–IBM Watson AI Lab

... from previous page

Both MIT and IBM have been pioneers in artificial intelligence research, and the new AI lab builds on a decades-long research relationship between the two. In 2016, IBM Research announced a multi-year collaboration with MIT's Department of Brain and Cognitive Sciences to advance the scientific field of machine vision, a core aspect of artificial intelligence. The collaboration has brought together leading brain, cognitive, and computer scientists to conduct research in the field of unsupervised machine understanding of audio-visual streams of data, using insights from next-generation models of the brain to inform advances in machine vision. In addition, IBM and the Broad Institute of MIT and Harvard have established a five-year, $50 million research collaboration on AI and Genomics.

MIT researchers were among those who helped coin and popularize the very phrase "Artificial Intelligence" in the 1950s. MIT pushed several major advances in the coming decades, from neural networks to data encryption to quantum computing to crowdsourcing. Marvin Minsky, a founder of the discipline, collaborated on building the first artificial neural network and he, along with Seymour Papert, advanced learning algorithms. Currently, the Computer Science and Artificial Intelligence Laboratory, the Media Lab, the Department of Brain and Cognitive Sciences, and the MIT Institute for Data, Systems, and Society serve as connected hubs for AI and related research at MIT.

For more than 20 years, IBM has explored the application of AI across many areas and industries. IBM researchers invented and built Watson, which is a cloud-based AI platform being used by businesses, developers and universities to fight cancer, improve classroom learning, minimize pollution, enhance agriculture and oil and gas exploration, better manage financial investments and much more. Today, IBM scientists across the globe are working on fundamental advances in AI algorithms, science and technology that will pave the way for the next generation of artificially intelligent systems. For more information, visit MITIBMWatsonAILab.mit.edu.

Full article with slide show and Links CLICK HERE

Globe Telecom taps Cloudera to harness machine learning

Staff writer telecomasia.net | September 13, 2017 | B/OSS Asia

Globe Telecom in the Philippines has deployed the Cloudera platform to enhance customer experience and deliver real-time targeted marketing campaigns and offers to its 60 million customers.

Globe Telecom is using machine learning with Cloudera to enrich customer experiences across channels and deliver targeted and optimized products and services, while maintaining compliance with the latest industry data regulations.

The operator's mobile data traffic grew 85% from 151 petabytes (PB) in 2016 to 280 PB this year. Mobile data is a growing and significant contributor to total mobile revenues for the first half of 2017 versus the similar period a year ago.

"To sustain our growth, we are always looking for ways to improve customer experiences across our channels and touchpoints," said Gil Genio, chief technology and information officer at Globe Telecom. "Our ability to strategically manage and monetize information about our customers will enable us to deliver value-added products and further differentiate ourselves in today's competitive business landscape."

With Cloudera Enterprise now at the core of Globe Telecom's data management architecture, the increasing volumes of data are ingested from different sources and channels into a centralized data hub and made available to all employees across the organization with full fidelity and security.

Mark Micallef, regional VP for Asia Pacific and Japan at Cloudera, said Globe Telecom can now use data to gain valuable insights, make accurate business decisions faster and deliver targeted marketing campaigns and offers to enhance their customer's experience.

About Global Telecom
Global Telecom Holding is a leading international telecommunications company operating mobile networks in high growth markets in Pakistan, Algeria and Bangladesh, having a total population under license of approximately 400 million as of December 31, 2016.
Headquarters are located at Gustav Mahlerlaan314, 1082 ME Amsterdam, the Netherlands
The representative office is located at 2005C, Nile City Towers- North Tower, CornicheEl Nile- RamletBeaulac11221, Cairo, Egypt.
Global Telecom operates networks in Algeria ("Djezzy"), Pakistan ("Jazz") and Bangladesh ("banglalink").
Total customers exceeded 98.3 million as of December 31, 2016. MORE: CLICK HERE

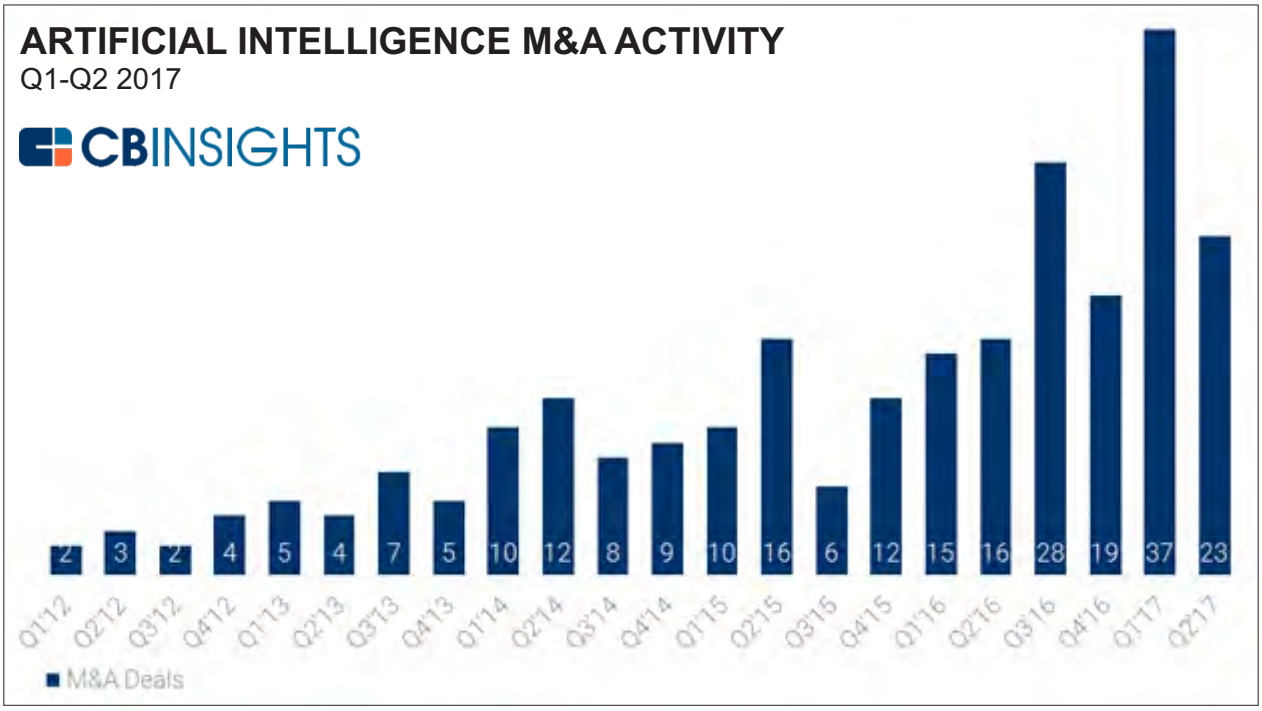
About Cloudera CLICK HERE

The Race For AI: Google, Baidu, Intel, Apple In A Rush To Grab Artificial Intelligence Startups

July 21, 2017 -- SPECIAL REPORT from CB INSIGHTS

Around 47% of the AI companies acquired since 2012 have had VC backing.

Corporate giants like Google, IBM, Yahoo, Intel, Apple, and Salesforce are competing in the race to acquire private AI companies, with Ford, Samsung, GE, and Uber emerging as new entrants. Over 250 private companies using AI algorithms across different verticals have been acquired since 2012, with 37 acquisitions taking place in Q1'17 alone. That quarter also saw one of the largest M&A deals: Ford's acquisition of Argo AI for $1B. Baidu has been particularly aggressive in its AI acquisitions in 2017, with 3 M&A deals so far this year, including its acquisition of Amazon Alexa Fund-backed Kitt.ai this quarter.



Google is the most active acquirer of AI startups, with 12 acquisitions under its belt since 2012. In 2013, Google picked up deep learning and neural network startup DNNresearch from the computer science department at the University of Toronto. This acquisition reportedly helped Google make major upgrades to its image search feature. In 2014 Google acquired British company DeepMind Technologies for some $600M (Google's DeepMind program recently beat a human world champion in the board game "Go"). Last year, it acquired visual search startup Moodstock, and bot platform Api.ai. More recently, in Q3'17, Google acquired India-based Halli Labs.

Apple has been ramping up its M&A activity, and ranked second with a total of 8 acquisitions. It recently acquired California-based Lattice Data for $200M and Tel Aviv-based RealFace, valued at $2M. (The table below excludes computer vision-based AR/VR startups like SensoMotoric Instruments, which Apple acquired in Q2'17.)

Intel, Microsoft, and Facebook are tied for third place. Intel acquired 3 startups in 2016 alone: Itseez, Nervana Systems, and Movidius, while Facebook acquired Belarus-based Masquerade Technologies and Switzerland-based Zurich Eye recently. Microsoft recently acquired Genee and conversational AI startup Maluuba.

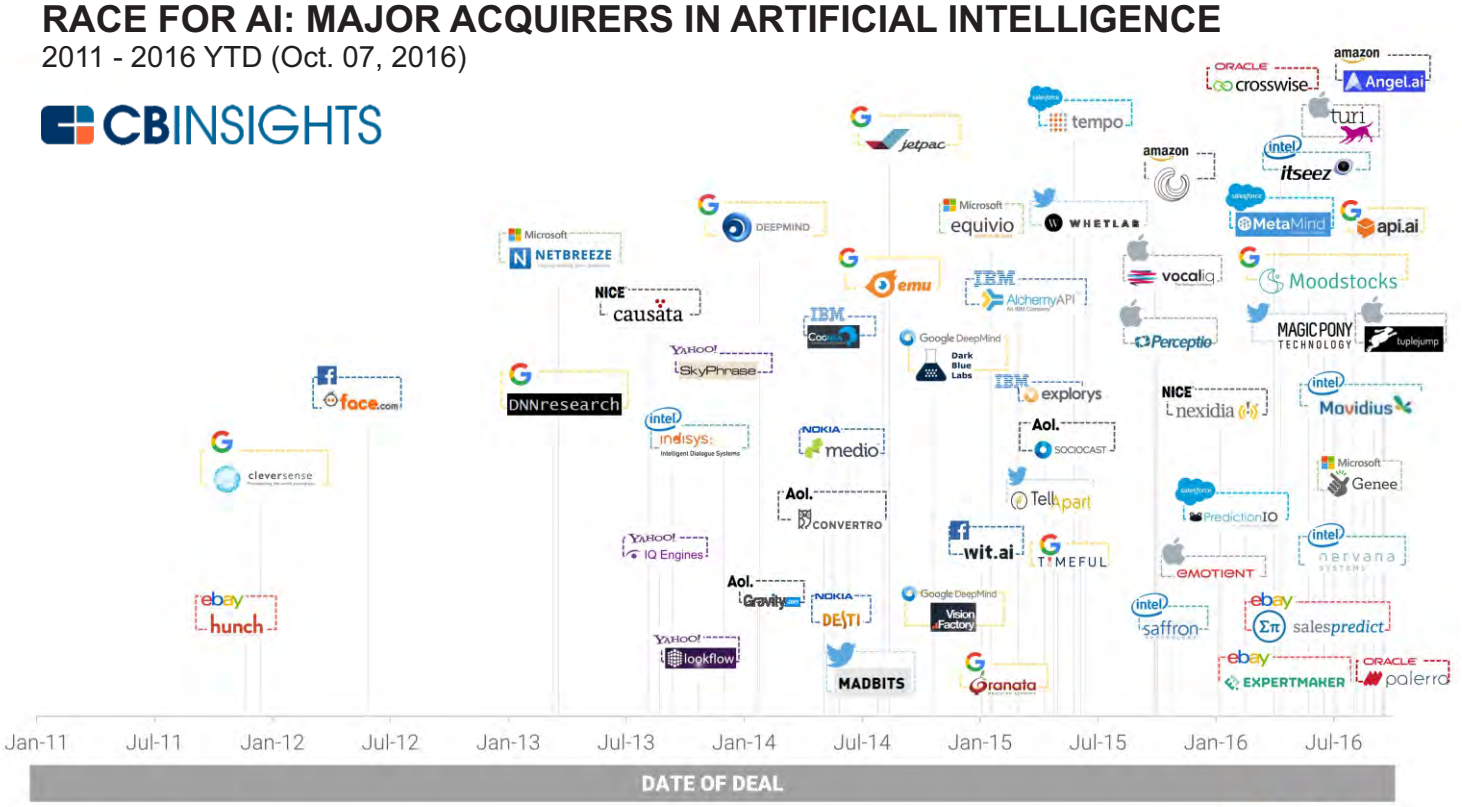
Salesforce and Twitter have acquired 4 AI startups each. Twitter's most recent M&A deal was image-processing startup Magic Pony. Salesforce, which joined the race in 2015 with the acquisition of Tempo AI, made two major acquisitions last year: Khosla Ventures-backed MetaMind and open-source machine-learning server PredictionIO. GE made 2 acquisitions in November 2016: AI-IoT startup Bit Stew Systems, and CRM-focused Wise.io.

The timeline below shows the M&A activity of corporations that have made 2 or more acquisitions since 2012. (Note: The exact dates for Apple's Novauris and Amazon's Orbeus acquisitions are not known. They are marked with a star to indicate approximate date of acquisition.)

... from previous page

The Race For AI: Google, Baidu, Intel, Apple In A Rush To Grab Artificial Intelligence Startups *... from previous page*

The timeline below shows the M&A activity of corporations that have made 2 or more acquisitions since 2012. (Note: The exact dates for Apple's Novauris and Amazon's Orbeus acquisitions are not known. They are marked with a star to indicate approximate date of acquisition.)



Why AI-Driven Predictive Logistics is Exactly What the Industry Needs

Stuart Reichenbach, Vice President of Supply Chain Development, ClearMetal -- 20-Sep-2017

Artificial intelligence and machine learning are dominating news headlines and events across a wide range of industries. As business segments begin to adopt these technologies in meaningful ways, the $4 trillion logistics industry still runs in a largely analog fashion. This is surprising given the heavily data-driven nature of logistics.

The industry is ready for digital transformation. There's a clear need for increased efficiency. A typical supply chain transaction involves five to 25 different parties and handoffs. A staggering amount of data and potential outcomes are introduced in this process – far too many for humans to optimize without advanced technology. With the overwhelming complexity, supply chain leaders are increasingly interested in how data intelligence can simplify and supplement their predictions and business decisions.

The current approach & tools can't keep up

According to the IHS Global Insight World Trade Services, many players in today's global supply chain are using technology created decades ago – or, in some cases, collecting data completely offline. Because these tools aren't sophisticated and haven't been updated for today's supply chain, data ends up inaccessible and inaccurate which leads to being reactive and likely wrong. We all know what that means - more inventory, lower margins and less than happy customers. Big isn't better. It is time to get smarter and deploy data science. Here's how predictive logistics can help.

Most supply chain decisions are already based on predictions (defined as the best information we have, along with years of experience and tribal knowledge). With the advancement of machine learning and artificial intelligence, it is in the industry's best interest to embrace this technology. The way predictive logistics works is by first pooling and then cleaning data that logistics decision makers have at their disposal. AI and machine learning are then applied to interpret and understand the vast quantities of data in a way the human mind cannot. Ultimately, this process produces an accurate prediction of what logistics leaders can expect to see happen in the global supply chain weeks in advance. This technology is useful for carriers, forwarders, terminals, shippers and 3/4PLs, because it enables them to plan further in advance and far more accurately.

Moving forward

Supply chain leaders need to embrace this new technology and start by simply reexamining their data through the lens of AI and machine learning. Then, they should think about the corresponding change to your operating model and decision making process. Being proactive and more accurate with the help of intelligent, predictive data will drive less inventory, lower cost, and greater revenue.

IBM opens first 'Machine Learning Hub' in Bengaluru

Software major IBM on Thursday opened its first "Machine Learning (ML) Hub" in Bengaluru which would provide a physical space to organisations for hands-on training on machine learning.

IANS | August 04, 2017, 08:49 IST -- Bengaluru: Software major IBM on Thursday opened its first "Machine Learning (ML) Hub" in Bengaluru which would provide a physical space to organisations for hands-on training on machine learning.

Through the 'ML Hubs', data professionals, business analysts and engineers could work with IBM's data science experts to understand and learn the technology to visualise, analyse and interpret data.

"'Machine Learning' termed by an IBMer decades ago has evolved significantly. Today, it is the entry point to the cognitive era, enabling enterprises to drive critical insights. With India's focus on digitisation, it's an apt time for organisations to make this transition," said Gaurav Sharma, Vice President, IBM India Software Labs and Vice President, Growth IBM India & South Asia.

IBM 'ML Hub' also provides a platform for like-minded enterprises to collaborate and transform their data science processes.

The company has similar 'ML Hubs' in Toronto, San Jose, California, at IBM's Silicon Valley Lab, Beijing, and Boblingen, Germany.

Artificial intelligence, machine learning to impact workplace practices in India: Adobe

According to a global report by software major Adobe, over 50 per cent respondents did not feel concerned by artificial intelligence (AI) or machine learning.

IANS | August 04, 2017, 08:50 IST -- NEW DELHI: Over 60 per cent of marketers in India believe new-age technologies are going to impact their workplace practices and consider it the next big disruptor in the industry, a new report said on Thursday.

According to a global report by software major Adobe that involved more than 5,000 creative and marketing professionals across the Asia Pacific (APAC) region, over 50 per cent respondents did not feel concerned by artificial intelligence (AI) or machine learning.

However, 27 per cent in India said they were extremely concerned about the impact of these new technologies.

Creatives in India are concerned that new technologies will take over their jobs. But they suggested that as they embrace AI and machine learning, creatives will be able to increase their value through design thinking.

"While AI and machine learning provide an opportunity to automate processes and save creative professionals from day-to-day production, it is not a replacement to the role of creativity," said Kulmeet Bawa, Managing Director, Adobe South Asia.

"It provides more levy for creatives to spend their time focusing on what they do best -- being creative, scaling their ideas and allowing them time to focus on ideation and creativity," Bawa added.

A whopping 59 per cent find it imperative to update their skills every six months to keep up with the industry developments.
The study also found that merging online and offline experiences was the biggest driver of change for the creative community, followed by the adoption of data and analytics, and the need for new skills.

It was revealed that customer experience is the number one investment by businesses across APAC.

Forty-two per cent of creatives and marketers in India have recently implemented a customer experience programme, while 34 per cent plan to develop one in the one year.

The study noted that social media and content were the key investment areas by APAC organisations, and had augmented the demand for content. However, they also presented challenges.

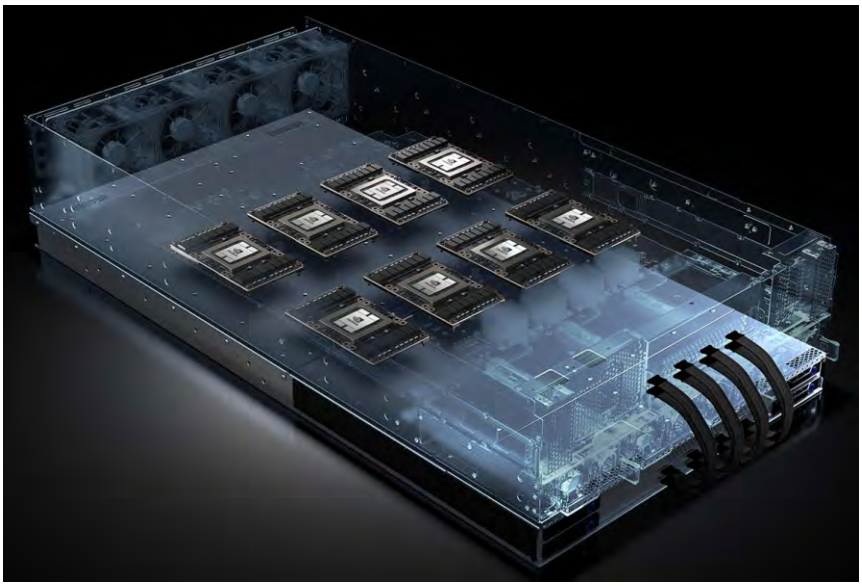
"Budgets were identified as the biggest challenge, followed by conflicting views and internal processes. Data and analytics become their primary tool to ensure that what they are creating is relevant, and delivering an amazing experience for customers," Bawa said.

NVIDIA Partners with World's Top Server Manufacturers to Advance AI Cloud Computing

Foxconn, Inventec, Quanta, Wistron Using NVIDIA HGX Reference Architecture to Build AI Systems for Hyperscale Data Centers

TAIPEI, TAIWAN--(Marketwired - May 30, 2017) - Computex -- NVIDIA (NASDAQ: NVDA) today launched a partner program with the world's leading original design manufacturers (ODM) -- Foxconn, Inventec, Quanta and Wistron -- to more rapidly meet the demands for AI cloud computing.

Through the NVIDIA HGX Partner Program, NVIDIA is providing each ODM with early access to the NVIDIA HGX reference architecture, NVIDIA GPU computing technologies and design guidelines. HGX is the same data center design used in Microsoft's Project Olympus initiative, Facebook's Big Basin systems and NVIDIA DGX-1™ AI supercomputers.



Using HGX as a starter "recipe," ODM partners can work with NVIDIA to more quickly design and bring to market a wide range of qualified GPU-accelerated systems for hyperscale data centers. Through the program, NVIDIA engineers will work closely with ODMs to help minimize the amount of time from design win to production deployments.

As the overall demand for AI computing resources has risen sharply over the past year, so has the market adoption and performance of NVIDIA's GPU computing platform. Today, 10 of the world's top 10 hyperscale businesses are using NVIDIA GPU accelerators in their data centers.

With new NVIDIA® Volta architecture-based GPUs offering three times the performance of its predecessor, ODMs can feed the market demand with new products based on the latest NVIDIA technology available.

"Accelerated computing is evolving rapidly -- in just one year we tripled the deep learning performance in our Tesla GPUs -- and this is having a significant impact on the way systems are designed," said Ian Buck, general manager of Accelerated Computing at NVIDIA. "Through our HGX partner program, device makers can ensure they're offering the latest AI technologies to the growing community of cloud computing providers."

Flexible, Upgradable Design
NVIDIA built the HGX reference design to meet the high-performance, efficiency and massive scaling requirements unique to hyperscale cloud environments. Highly configurable based on workload needs, HGX can easily combine GPUs and CPUs in a number of ways for high performance computing, deep learning training and deep learning inferencing.

The standard HGX design architecture includes eight NVIDIA Tesla® GPU accelerators in the SXM2 form factor and connected in a cube mesh using NVIDIA NVLink™ high-speed interconnects and optimized PCIe topologies. With a modular design, HGX enclosures are suited for deployment in existing data center racks across the globe, using hyperscale CPU nodes as needed.

Both NVIDIA Tesla P100 and V100 GPU accelerators are compatible with HGX. This allows for immediate upgrades of all HGX-based products once V100 GPUs become available later this year.

HGX is an ideal reference architecture for cloud providers seeking to host the new NVIDIA GPU Cloud platform. The NVIDIA GPU Cloud platform manages a catalog of fully integrated and optimized deep learning framework containers, including Caffe2, Cognitive Toolkit, MXNet and TensorFlow.

... to next page

NVIDIA Partners with World's Top Server Manufacturers to Advance AI Cloud Computing

... from previous page

"Through this new partner program with NVIDIA, we will be able to more quickly serve the growing demands of our customers, many of whom manage some of the largest data centers in the world," said Taiyu Chou, general manager of Foxconn/Hon Hai Precision Ind Co., Ltd., and president of Ingrasys Technology Inc. "Early access to NVIDIA GPU technologies and design guidelines will help us more rapidly introduce innovative products for our customers' growing AI computing needs."

"Working more closely with NVIDIA will help us infuse a new level of innovation into data center infrastructure worldwide," said Evan Chien, head of IEC China operations at Inventec Corporation. "Through our close collaboration, we will be able to more effectively address the compute-intensive AI needs of companies managing hyperscale cloud environments."

"Tapping into NVIDIA's AI computing expertise will allow us to immediately bring to market game-changing solutions to meet the new computing requirements of the AI era," said Mike Yang, senior vice president at Quanta Computer Inc. and president at QCT.

"As a long-time collaborator with NVIDIA, we look forward to deepening our relationship so that we can meet the increasing computing needs of our hyperscale data center customers," said Donald Hwang, chief technology officer and president of the Enterprise Business Group at Wistron. "Our customers are hungry for more GPU computing power to handle a variety of AI workloads, and through this new partnership we will be able to deliver new solutions faster."
"We've collaborated with Ingrasys and NVIDIA to pioneer a new industry standard design to meet the growing demands of the new AI era," said Kushagra Vaid, general manager and distinguished engineer, Azure Hardware Infrastructure, Microsoft Corp. "The HGX-1 AI accelerator has been developed as a component of Microsoft's Project Olympus to achieve extreme performance scalability through the option for high-bandwidth interconnectivity for up to 32 GPUs."

About NVIDIA
NVIDIA's (NASDAQ: NVDA) invention of the GPU in 1999 sparked the growth of the PC gaming market, redefined modern computer graphics and revolutionized parallel computing. More recently, GPU deep learning ignited modern AI -- the next era of computing -- with the GPU acting as the brain of computers, robots and self-driving cars that can perceive and understand the world. More information at http://nvidianews.nvidia.com/.

Amazon Strategy Teardown: Building New Business Pillars In AI, Next-Gen Logistics, And Enterprise Cloud Apps

Seattle-based Amazon is doubling down on AWS and its AI assistant, Alexa. It's seeking to become the central provider for AI-as-a-service.

Amazon is the exception to nearly every rule in business. Rising from humble beginnings as a Seattle-based internet bookstore, Amazon has grown into a propulsive force in at least five different giant industries: retail, logistics, consumer technology, cloud computing, and most recently, media and entertainment. The company has had its share of missteps — the expensive Fire phone flop comes to mind — but is also rightly known for strokes of strategic genius that have put it ahead of competitors in promising new industries.

This was the case with the launch of cloud business AWS in the mid-2000s, and more recently the surprising consumer hit in the Echo device and its Alexa AI assistant. Today's Amazon is far more than just an "everything store," it's a leader in consumer-facing AI and enterprise cloud services. And its insatiable appetite for new markets mean competitors must always be on guard against its next moves.

As the biggest online retailer in America, the company accounts for 5% of all retail spending in America, and the company has been publicly traded for two decades. While its market capitalization has swelled recently, so too have expectations. Wall Street banks like Morgan Stanley expect Amazon to continue growing at a rate that no company its size has ever done before: 16% average compound growth in sales through 2025. If Amazon were able to satisfy the lofty goals, it would be "the most aggressive expansion of a giant company in the history of modern business."

Understanding the many-headed beast that is Amazon is no easy feat, especially because Amazon is far less transparent than its peers. As the Times has written, "It isn't just secretive, the way Apple is, but in a deeper sense, Jeff Bezos' e-commerce and cloud-storage giant is opaque. Amazon rarely explains either its near-term tactical aims or its long-term strategic vision. It values surprise."

From: CB Insights -- April 2017 MORE: CLICK HERE

HPE Accelerates Real-Time Insights for Deep Learning

New purpose-built compute portfolio, partner ecosystem collaboration and comprehensive services optimize the Deep Learning experience

May 10, 2017 17:00 ET | Source: Hewlett Packard Enterprise

PALO ALTO, Calif., May 10, 2017 (GLOBE NEWSWIRE) -- Hewlett Packard Enterprise (NYSE:HPE) today announced a comprehensive set of computing innovations to accelerate Deep Learning analytics and insights across all organizations with innovations spanning systems design, partner ecosystem collaboration, and expertise including flexible consumption models from HPE Pointnext Services.

Advanced artificial intelligence (AI) techniques, such as Deep Learning, are growing in popularity across various sectors including financial services, life sciences, manufacturing, energy, government and retail. HPE has a strong track record of delivering comprehensive, workload optimized compute solutions for all AI and Deep Learning with its purpose-built HPE Apollo portfolio that maximizes performance, scale and efficiency. With the latest innovations specifically targeted to Deep Learning, leveraging capabilities from the recent SGI acquisition, HPE now offers greater choice for larger scale, dense GPU environments and addresses key gaps in technology integration and expertise with integrated solutions and services offerings.

"Customers pursuing Deep Learning projects face a variety of challenges including a lack of mature IT infrastructure and technology capabilities leading to poor performance, efficiency and time to value," said Bill Mannel, Vice President and General Manager, High Performance Computing and Artificial Intelligence at Hewlett Packard Enterprise. "To address these challenges, HPE is introducing new optimized GPU compute platforms, an enhanced collaboration with NVIDIA and HPE Pointnext Services from the Core Datacenter to the Intelligent Edge."

The new portfolio of capabilities includes:

- **New HPE SGI 8600:** Based on the SGI ICE XA architecture, High Performance Computing platform with support for optimal combination of **liquid-cooled** GPU performance with NVIDIA® Tesla® GPU accelerators with NVLinkTM interconnect technology to provide scale and efficiency for the most complex, largest environments - up to thousands of nodes with leading power efficiency.
- **Interactive Rendering from the** Datacenter with the HPE Apollo 6500 and NVIDIA Tesla GPUs certified with NVIDIA VCA software
- **Support for NVIDIA's next generation Tesla GPUs based on its Volta® architecture** when available in production quantities in the Apollo 2000, Apollo 6500 and Proliant DL380 servers

Apollo 6500

Hewlett Packard Enterprise

HPE and NVIDIA Enhanced Collaboration for Deep Learning

Through their collaboration HPE and NVIDIA will jointly address GPU technology integration and Deep Learning expertise challenges to accelerate the adoption of technologies that provide real-time insights from massive data volumes.

"As the artificial intelligence era takes hold, enterprises are increasingly adopting NVIDIA's GPU computing platform to generate insights from decades of untapped data," said Ian Buck, General Manager of Accelerated Computing at NVIDIA. "Expanding our collaboration with HPE around deep learning will help enterprises deploy, manage and optimize their GPU computing infrastructure and realize the benefits of AI and deep learning in their business."

Building on the recent HPE Supercomputer win at Tokyo Institute of Technology, which is one of the largest NVIDIA Tesla P100 GPU based clusters, this collaboration will deliver:
Enhanced Centers of Excellence for benchmarking, code modernization and proof of concept initiatives. The locations include Korea, Sydney, Grenoble, Bangalore and Houston
Early access program for Volta-based NVIDIA Tesla SXM2 GPU systems powered with eight GPUs for selected customers in 4Q 2017

... to next page

HPE Accelerates Real-Time Insights for Deep Learning

... from previous pgeNew

"Through our partnership with SGI, and now HPE, the Tokyo Institute of Technology has worked successfully to deliver a converged world-leading HPC and Deep Learning platform that can address our requirements and those of our nation," said Satoshi Matsuoka, Professor and TSUBAME Leader, Tokyo Institute of Technology. "The NVIDIA Tesla P100 SXM2 node solution enables GPU based Deep Learning capability to be scalable to the entire size of our TSUBAME 3.0 system. We look forward to continuing our partnership with HPE to work together on future projects in HPC and Deep Learning."

Partner Ecosystem Collaboration for Deep Learning based Fraud Detection

As part of HPE's partner ecosystem collaboration, HPE is working with Kinetica, a leading software application provider leveraging Deep Learning frameworks to develop a solution to automate, real-time fraud detection with GPU acceleration. Designed specifically for consumer credit card transaction processing, the new performance optimized, cost effective solution, will be demoed at the HPE booth during GTC.

"We look forward to advancing the Kinetica GPU database with HPE and jointly offering a best-of-breed GPU-accelerated analytics solution that converges Artificial Intelligence and Business Intelligence workloads for financial services as well as for retail, healthcare and other industries," said Chris Prendergast, Vice President of Business Development and Alliances from Kinetica.

Services to Enable and Support AI and Deep Learning Environments

As customers begin the journey to adopt these powerful and scalable IT solutions, HPE Pointnext provides the knowledge and expertise through its Advisory, Professional and Operational Services to help achieve desired business outcomes, including faster time to value. With AI and Deep Learning requiring scalable infrastructure, Pointnext offers HPE Flexible Capacity, a service that provides on-demand capacity, combining the agility and economics of public cloud with the security and performance of on-premises IT.

"With the need to embed more intelligence and automation into data analytics to address scientific and business challenges, artificial intelligence-based techniques are growing in importance. HPE's systems and solutions innovations announced today are designed to address key performance and expertise constraints affecting deep learning," said Steve Conway, Senior Vice President for Research at Hyperion Research. "HPE's enhanced collaboration with NVIDIA for Deep Learning and comprehensive infrastructure capabilities, from the Core Datacenter to the Intelligent Edge, aims to use automated intelligence to enable real-time insights for customers."

HPE will showcase the new Deep Learning portfolio along with our fraud detection demo at GTC in San Jose at booth # 811. Additionally, Dr. Goh, and Professor Matsuoka will discuss Tokyo Institute of Technology's large scale system, the latest TSUBAME3.0 supercomputer during a breakout session on the Scalable Learning Platform.

Groq TPU-killer aimed at AI and machine learning

Groq could have the key to machine learning and AI – Silicon Valley's top-priority foci.

21-Apr-2017 -- By David Manners, Electronics Weekly

Groq is a **start-up in stealth mode** thought to be working on a TPU. It has raised $10.3 million.

The company is formed from some of the Google people working on Google's TPU – Tensor Processing Unit – which was invented by Jonathan Ross. Ross is one of the founders of Groq.

TPUs are designed for machine learning. According to Google they run AI workloads 15-20x faster than current processors with 30-80x better efficiency.

VC Chamath Palihapitiya, who backs Groq, says "we think what they're building could become a fundamental building block for the next generation of computing."