SIGMA
A PIVOT COMPANY

SigmaUptime

volume 16 number 3

A Prescription for Success

Artificial intelligence helping healthcare organizations convert data into insight.

UPTIME

PRSRT STD
U.S. POSTAGE
PAID
Tulsa, OK
Permit No. 2146

Technology services to help you streamline operations, reduce costs and improve business processes.

Companies today must align IT strategy with their corporate objectives, strategy and business model. Pivot has created a portfolio of operating companies and partners with a focus on helping you enhance and extend the capabilities of your technology assets.

Pivot provides technology services ranging from initial needs assessment and design, through procurement and implementation, to on-going support. As an adjunct to your IT team, we provide the resources that allow your team to offload some of the day-to-day operational challenges and focus on innovations that will drive business value and competitive advantage.

Contact us to learn more.

PIVOT
TECHNOLOGY SERVICES

www.pivotts.com 888-895-0495

Contents

Sigma Uptime

Editorial Correspondence:

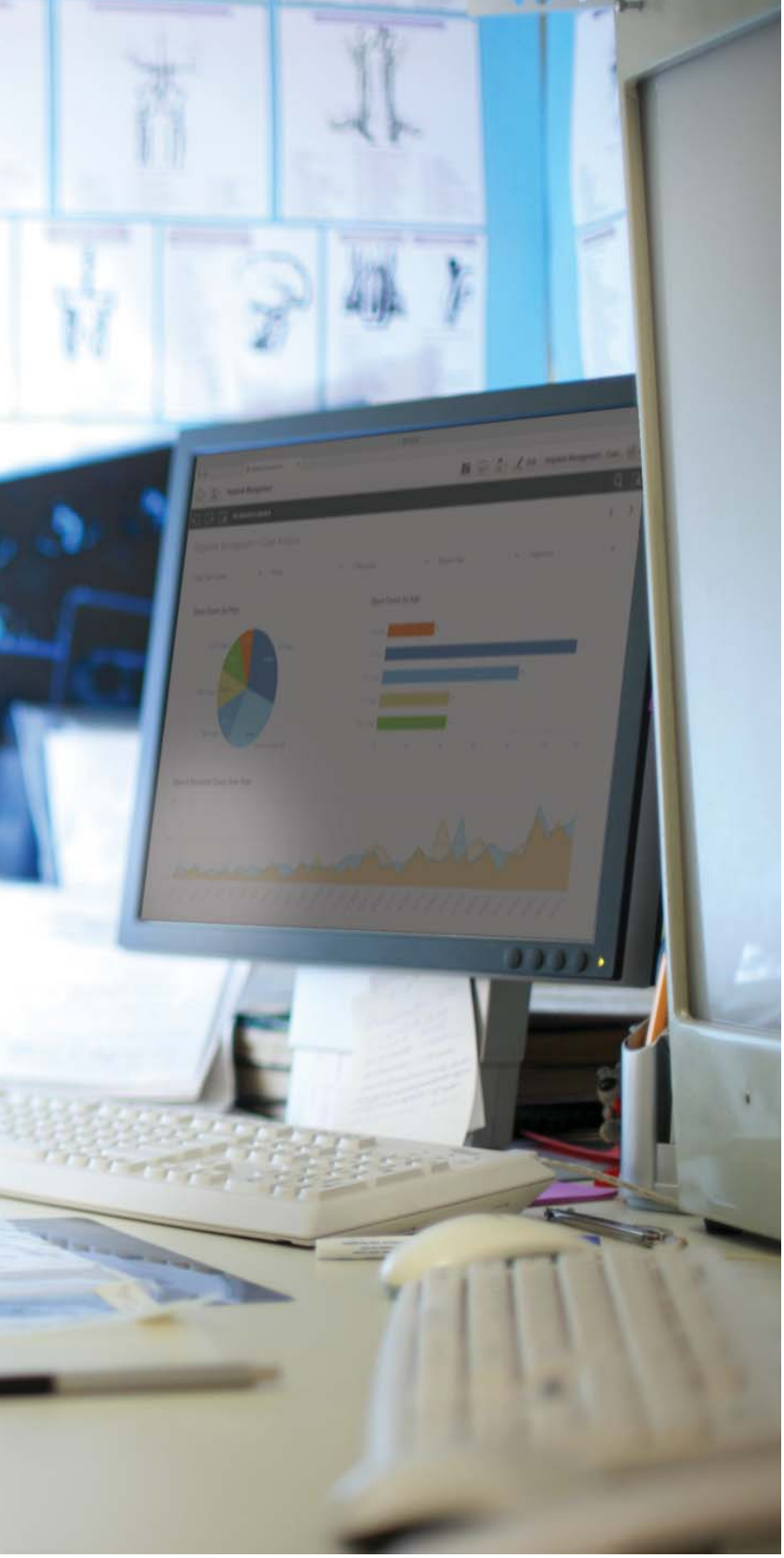
10221 East 61st Street, Tulsa, OK 74133

Phone (800) 726-7667 • Fax (918) 270-7134

Change of Address: Send corrected address label to the above address.

Sigma UPTIME is published bimonthly by CMS Special Interest Publications. Printed in the U.S.A. Product names may be trademarks of their respective companies.

A PRESCRIPTION FO



R SUCCESS

Artificial intelligence helping healthcare organizations convert data into insight.

With the transition to electronic health records (EHR) over the past decade, healthcare organizations have been able to collect and share data far more efficiently than ever before. The rise of artificial intelligence (AI) is finally enabling them to actually use that data to improve patient care and control costs.

There has been an exponential growth of healthcare data since the 2009 federal mandate for the adoption of EHR systems. Various studies find that data stores are doubling roughly every two years. But more data doesn't automatically mean more insight, according to Konstantin Cvetanov, Principal Architect with ProSys (which, like Sigma, is part of the Pivot family of companies).

"It's really been a 'first-things-first' situation with healthcare data," said Cvetanov. "To date, organizations have naturally concentrated on building the infrastructure that will allow them to acquire, store, manage, share and protect data.

"Now, they are turning their focus to data analysis that enables data-driven decisions. The ongoing evolution of AI technologies facilitates this by making analysis faster, more automated and more accurate than ever before."

Finding Patterns

AI is actually an umbrella term for a number of technologies such as deep learning, machine learning, predictive analytics, computer vision and natural language processing. All are aimed at embedding machines with the ability to analyze massive data sets and make autonomous decisions — eliminating the need for programmers to write code for every function.

Predictive analytics tools, in particular, are being used to evaluate data from sources all across the healthcare spectrum, including clinics, hospitals, private practices, insurance companies, labs and research facilities. Such analysis is uncovering trends and pat-

continued on page 6

continued from page 5

terns that suggest ways to improve disease management, enhance administration, streamline supply chains and more.

A 2016 study published in the Journal of the American Medical Association finds that predictive analytics is enabling "precision delivery" of healthcare services. The study from a team of research physicians at Brigham and Women's Hospital in Boston finds that analysis of EHR data enables doctors to predict risk and personalize care for the dual purpose of improving patient outcomes and controlling costs.

Reducing Risk

As an example, the study cites efforts by Parkland Health and Hospital System in Dallas to reduce hospital readmissions, a traditionally significant source of healthcare spending. The study notes that readmissions within 30 days of discharge cost the U.S. healthcare system $41 billion annually.

Using an algorithm based on 29 separate clinical, social and operational factors, the hospital is able to identify cardiac patients most at risk of requiring readmission. These patients receive special interventions, including enhanced education by a multidisciplinary team, post-discharge telephone contacts to ensure they are taking their medication, an outpatient follow-up appointment within a week, and an appointment with the primary-care physician. Early results found that this approach reduced readmissions by 26 percent.

In another example, Kaiser Permanente has used predictive analytics to dramatically reduce the administration of antibiotics to newborns. Because infants are at a higher risk of potentially fatal bacterial infections such as sepsis and pneumonia, pediatricians often prescribe antibiotics even if they aren't positive there is an infection. The downside of this practice is that early exposure to antibiotics can lead to increased risk of asthma and autoimmune disorders later in childhood, according to some studies.

Researchers at Kaiser Permanente, the University of Pennsylvania and the University of California-San Francisco developed a risk-prediction model based on data from more than 600,000 births. With this model, Kaiser Permanente has safely reduced antibiotic use by nearly 50 percent in newborns.

No healthcare organization has more experience with predictive analytics than the Veteran's Health Administration, which was digitizing patient records for a couple of decades before the federal mandate. The VHA analyzes this data to create targeted care plans delivered by multidisciplinary teams — a practice that generated more than $3 billion in savings over a 10-year period, according to one study.

"These studies and many others like them clearly demonstrate how AI tools can create direct improvements in patient care," said Cvetanov. "Additionally, they are powerful tools for bringing spiraling costs under control."

Controlling Costs

A new report from Grand View Research cites cost control as a chief driver of adoption of predictive analytics tools by healthcare providers. The firm says the market for such tools is expected to reach $19.5 billion by 2025. The report notes that financial applications, including fraud and waste detection, currently make up about 30 percent of the market for predictive analytics in healthcare.

"It's fairly obvious that current healthcare spending patterns are not sustainable," said Cvetanov. "Studies find that Americans are spending, on average, about $9,500 per person every year for healthcare services. That's more than twice the global average.

"The solutions for improving care and controlling costs are often locked inside healthcare's huge data stores. Artificial intelligence is giving us the key to unlock that data and open the door to new treatments, new medicines and new processes that improve quality of life."

NVIDIA Tesla V100

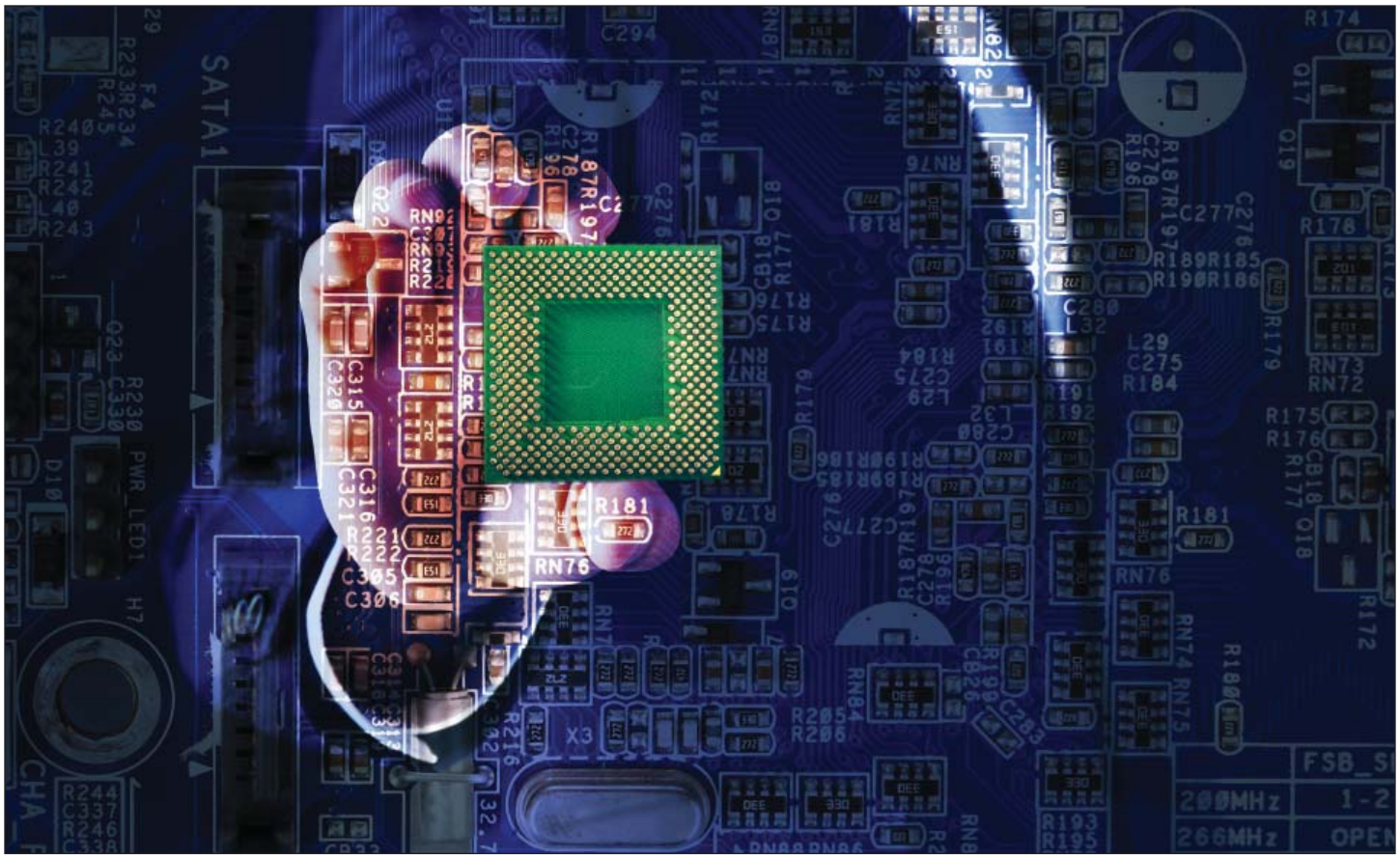
DESIGNED TO BRING AI TO EVERY INDUSTRY

NVIDIA

NVIDIA's new Tesla® V100 graphics processing unit is designed to accelerate adoption of artificial intelligence. It is powered by NVIDIA Volta™, the most powerful GPU architecture the world has ever seen with more than 21 billion transistors. Tesla V100 will fuel breakthroughs in every industry — including supercomputing, healthcare, financial services, big data analytics, and gaming. Contact Sigma to learn how GPU acceleration can boost your business operations.

SIGMA
A PIVOT COMPANY

www.sigmasol.com | 888.895.0495



A Whole New Game

Computer gaming pioneer Nvidia provides the boost that powers artificial intelligence.

Rapid advances in artificial intelligence (AI) are being enabled by more powerful hardware, sophisticated algorithms and big data analytics. Much of this has evolved from the graphics processing unit (GPU) Nvidia originally designed for the video game market.

Nvidia revolutionized computer gaming with its GPU. These specialized circuits perform multiple mathematic calculations simultaneously to manipulate and alter memory in order to produce cleaner, faster and smoother motion in video games.

In 2007, Nvidia pioneered the concept of "GPU-accelerated computing," which combines GPUs with traditional computer processing units (CPUs) in massively parallel processing environments that make compute-intensive programs run faster. That development provided the processing "oomph" required to enable essential AI functions such as deep learning.

Deep learning is a computing model designed to loosely mimic the way the human brain works with neurons and synapses. Nvidia's GPUs are used to create so-called "artificial neural networks" that use a large number of highly interconnected nodes working in unison to analyze large datasets. This gives a machine the ability to discover patterns or trends — and learn from those discoveries. This is the essence of artificial intelligence.

Need for Speed

Key architectural differences between a CPU and Nvidia's GPU made the difference. A CPU has a few cores with lots of cache memory that can handle a few software threads at a time, but a GPU has hundreds

of cores that can handle thousands of threads at the same time. Plus, CPUs are optimized for sequential processing — the execution of processes in the order they are received — as opposed to the GPU's ability to execute multiple processes at the same time.

"GPU-accelerated computing can run some software 100 times faster than with a CPU alone," said Konstantin Cvetanov, Principal Architect, ProSys. "Plus, it conserves power and is more cost-efficient. That makes it perfect for the deep learning type of algorithms that are powering a range of AI applications."

Nvidia GPUs are being used to accelerate more than 400 applications for uses such as quantum chemistry, fluid dynamics, video editing, medical imaging and geosciences. Over the past two years, the number of companies collaborating with Nvidia on deep learning has jumped to nearly 20,000.

The AI Server

To fully exploit the capabilities of its GPUs, Nvidia recently introduced the DGX-1 server. This so-called "AI supercomputer in a box" delivers 170 teraflops of processing power in a single system and is purpose-built for deep learning and AI-accelerated analytics. It comes fully integrated with hardware, deep learning software and development tools, and runs popular accelerated analytics applications.

The DGX-1 software stack includes DIGITS deep-learning training module, the CUDA programming model and a library of neural network designs. It also includes optimized versions of several widely used deep learning frameworks such as Caffe, Theano and Torch.

Because deep learning involves analysis of large datasets, AI platforms need a cloud element for accessing cloud storage. DGX-1 provides access to cloud management tools, software updates and a repository for containerized applications.

The cloud, in fact, represents a vital intersection for AI. It is inevitable that organizations will look to utilize deep learning and AI applications without implementing an AI framework onsite. This is why Nvidia recently partnered with Microsoft to allow users to run GPU-accelerated workloads in Microsoft's Azure cloud platform. Customers will be able to use Azure N-Series virtual machines powered by Nvidia Tesla K80 GPUs to run deep learning training jobs, high-performance computing simulations, data rendering, real-time analytics, DNA sequencing and other accelerated tasks.

Fastest-Ever GPUs Drive AI Advances

Artificial intelligence is being applied to a range of complex challenges, from financial analysis and medical diagnostics to speech recognition and self-driving cars. Such tasks obviously require significant amounts of computational power, and Nvidia is delivering with the recent launch of its new graphics-processing unit (GPU) architecture.

Volta, Nvidia's seventh-generation GPU architecture, features more than 21 billion transistors and delivers the equivalent performance of 100 central processing units (CPUs). It provides a 5x improvement in peak teraflops over Nvidia's current-generation Paschal architecture and 15x over the Maxwell architecture launched two years ago.

The company's first Volta-based processor, the Tesla V100 data center GPU, is built to drive the next wave of advancement in AI. The Tesla V100 features 672 tensor cores, a new type of core explicitly designed to accelerate AI workloads. These cores can deliver up to 120 teraflops of processing power, making it the world's first GPU to break the 100TFLOPs barrier.

Analysts say this new GPU can bring a convergence of AI and traditional high-performance computing (HPC) models. HPC is the practice of using parallel processing to aggregate the computing power of multiple machines. These large systems can create simulations based on analysis of large data sets. AI moves beyond simulations, using the HPC datasets to create so-called "artificial neural networks" that gives a machine the ability to learn from patterns or trends discovered.

"Artificial intelligence is driving the greatest technology advances in human history," said Jensen Huang, founder and chief executive officer of Nvidia. "It will automate intelligence and spur a wave of social progress unmatched since the industrial revolution.

"Deep learning, a groundbreaking AI approach that creates computer software that learns, has insatiable demand for processing power. Thousands of NVIDIA engineers spent over three years crafting Volta to help meet this need, enabling the industry to realize AI's life-changing potential."

Powerful Lessons

High-performance computing delivers the data resources required to 'teach' AI systems.

"The more that you read, the more things you will know." It turns out that quote from Dr. Seuss is as true for computers as it is for humans.

Artificial intelligence (AI) intends to endow computers with the ability to interpret data, learn from it and form its own conclusions without the need for human intervention. Just like human learning, however, machine learning requires lots of input. Special algorithms help computers to "teach" themselves to perform tasks by finding patterns within massive datasets. Studies find that error rates fall as more and more data is processed.

Though conceptually solid, AI faced practical limits for decades because of data processing constraints. Machine learning algorithms may need millions or billions of examples before they start working well. Standalone computers simply can't meet the data analysis and calculation requirements for effective AI modeling.

High-performance computing (HPC) systems have removed those limitations. By aggregating the computing power of hundreds or thousands of processors, HPC systems deliver exponential increases in processing power that are pushing AI into the mainstream.

Driving Growth

AI could double annual economic growth rates by 2035 by changing the nature of work and spawning a new relationship between man and machine, according to research

from Accenture. The global consulting firm says the impact of AI technologies on business is projected to boost labor productivity by up to 40 percent by fundamentally changing the way work is done and reinforcing the role of people to drive growth in business.

Not surprisingly, AI has spurred growth in the HPC server market. Hyperion Research reports the overall HPC server market set a new record at $11.2 billion in 2016, up from $10.7 billion in 2015 and surpassing 2012's high-water mark of $11.1 billion.

Standalone computers simply can't meet the data analysis and calculation requirements for effective AI modeling.

"An important factor driving growth is the market for big data needing HPC, which we call high-performance data analysis, or HPDA," said Steve Conway, senior vice president, Hyperion Research. "HPDA challenges have moved HPC to the forefront of R&D for machine learning, deep learning, artificial intelligence and the Internet of Things."

High-performance computing is more accessible than one might think. Although people often think of HPC in terms of monster IBM and Cray supercomputers crunching data in search of cancer cures or the nature of black holes, almost any size organization can achieve HPC performance using off-the-shelf components and the latest system management tools.

Business Uses

Affordable HPC solutions use x86 chips in highly scalable architectures based upon the Linux operating system to support a wide range of AI workloads for smaller enterprises, including financial modeling, research and development, and big data analysis.

Analysts say AI technologies have use cases and applications in almost every industry, and promise to significantly change existing business models while simultaneously creating new ones. Tractica, a research firm focused on the AI market, has identified nearly 200 real-world AI use cases across 27 industries.

Financial services organizations are using HPC-driven AI solutions to improve fraud detection through facial recognition. Retail companies are using AI to automate call centers with computers that can better understand speech and answer questions. AI automation is transforming business accounting practices by automating repeatable data-entry tasks.

"Leading commercial companies in a variety of market segments are turning to HPC technologies for challenging big data analytics workloads that enterprise IT technology alone cannot tackle effectively. HPC systems can handle more complex queries, more variables and faster turnaround requirements," said Conway. "We estimate that the move to HPC … is saving tens of millions of dollars per year for commercial companies, on top of the benefits reported by established HPC users in government, academia and industry."

Additionally, HPC solutions deliver significantly improved scalability, availability and reliability when compared to typical low-end servers. When the processing capacity of low-end servers is reached, more servers must be added, increasing management headaches. By reducing server sprawl, an HPC system reduces demand for power, cooling and network interconnects, further reducing costs and improving efficiency.

Accessing AI

Even if server replacement is out of the budget, smaller organizations can access a variety of resources for explore the possibilities of HPC-driven analytics and machine-learning applications. Stanford University and Columbia University offer AI-focused online courses, as does the online university Udacity. Tech incubators such as Techcode and Singularity University offer consulting services for organizations looking to incorporate AI tools.

Even the biggest players in the tech world understand the importance of making AI accessible to organizations of all sizes. IBM is working to bring the deep-learning and data analysis capabilities of its Watson AI supercomputing platform to users through the cloud. Microsoft, meanwhile, is enabling users to run deep-learning training jobs, data rendering, real-time analytics and other accelerated tasks in its Azure cloud platform.

"We're working hard to empower every organization with AI, so that they can make smarter products and solve some of the world's most pressing problems," said Harry Shum, executive vice president of the Artificial Intelligence and Research Group at Microsoft. "AI is now within reach of any business."

Long considered a science-fiction plot device, artificial intelligence has moved way beyond the realm of fantasy with a range of mainstream business applications. To a great degree, this is due to the intersection of AI and HPC concepts. The aggregated computing capacity of HPC platforms provide AI systems the data they need to learn and improve.

Focus on your business. We'll take care of the rest.

With a predictable, monthly cost structure, Sigma's comprehensive managed services reduce IT costs and risks. Contact us today to learn more.

SIGMA
A PIVOT COMPANY

www.sigmasol.com | 888.895.0495