SUPERCHARGED COMPUTING FOR THE DA VINCIS AND EINSTEINS OF OUR TIME

We pioneered a supercharged form of computing loved by the most demanding computer users in the world — scientists, designers, artists, and gamers.

NVIDIA GPU computing has become the essential tool of the da Vincis and Einsteins of our time. For them, we've built the equivalent of a time machine.

Fueled by the insatiable demand for better 3D graphics and the massive scale of the gaming market, NVIDIA has evolved the GPU into a computer brain at the exciting intersection of virtual reality, high performance computing, and artificial intelligence.

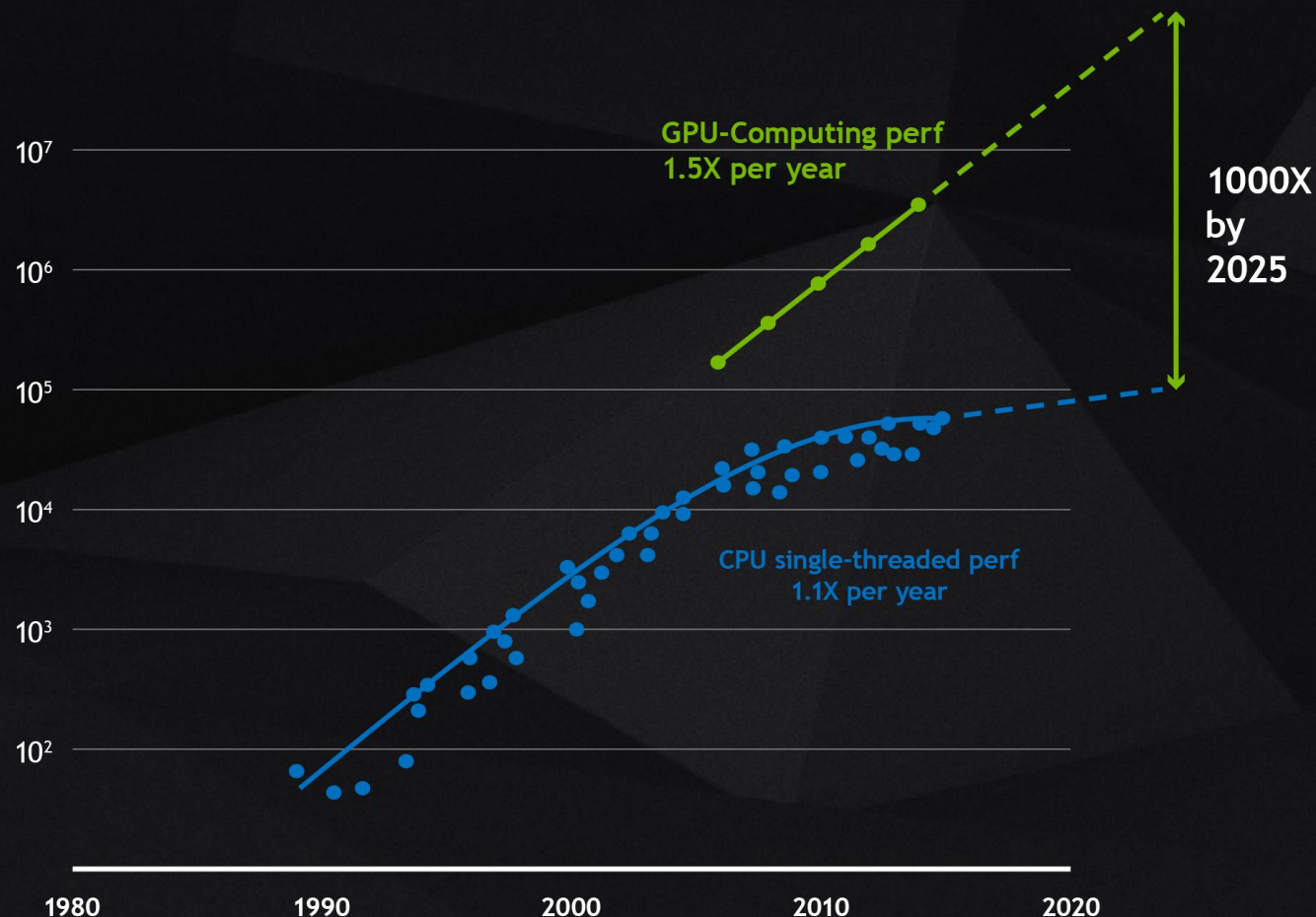*Image: Scientists at ETH Zurich use GPU-powered AI to sharpen our view of distant galaxies.*

THE TIME FOR GPU COMPUTING HAS COME

For 30 years, the dynamics of Moore's law held true. Microprocessor performance advanced at a rate of 50 percent per year as more and more transistors were fit onto a single chip. But that approach is hitting the limits of semiconductor physics, and, today, CPU performance only grows by 10 percent per year.

NVIDIA GPU computing has given the industry a path forward — and will provide a 1,000X speed-up by 2025. NVIDIA's CUDA® programming model complements the CPU with a specialized processor suited for parallel processing. And we innovate across the entire stack, from processor to systems to algorithms to applications.

GPU-Computing perf
1.5X per year

1000X
by
2025

CPU single-threaded perf
1.1X per year

10^7
10^6
10^5
10^4
10^3
10^2

1980   1990   2000   2010   2020

40 Years of Microprocessor Trend Data

GPU TECHNOLOGY CONFERENCE

THE EPICENTER OF GPU COMPUTING

GTC was started in 2009 to foster a new approach to high performance computing. Today, it's one of the most important AI events of the year. GTC 2017 hosted more than 7,000 attendees, 260 press and analysts, and 570 technical sessions. The world's top 15 tech companies and top 10 automakers were in attendance, along with more than 100 startups focused on AI and VR.
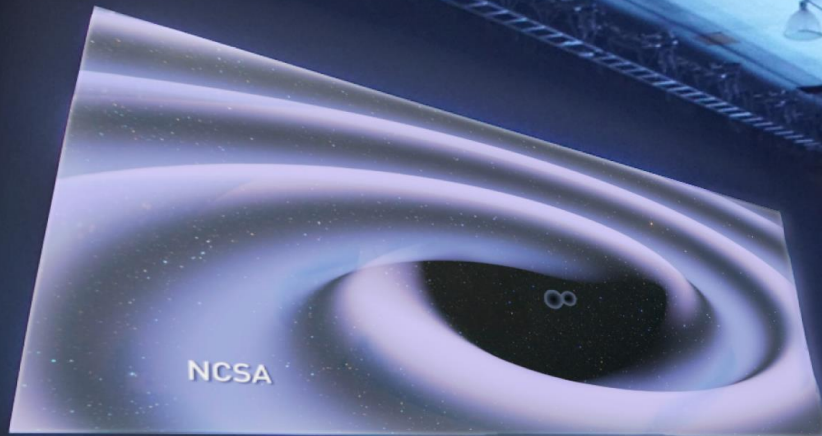
NVIDIA DEFINES MODERN COMPUTER GRAPHICS

Our invention of the GPU in 1999 made possible real-time programmable shading, which gives artists an infinite palette for expression. We've led the field of visual computing since.

THE DREAMS OF SCIENCE FICTION

NVIDIA's Project Holodeck is a photorealistic, collaborative VR environment that incorporates the feeling of real-world presence through sight, sound, and haptics. It allows creators to import high-fidelity, full-resolution models into VR to collaborate and share with colleagues or friends. A future where we create nearly everything in a virtual world is on the near horizon.

CREATING AMAZING WORLDS

Today's gaming industry is fueled by a steady stream of blockbuster titles with Hollywood-level production values. NVIDIA GPUs are the engines that make these games possible. And our GameWorks™ software allows developers to make games photorealistic and immersive.

GEFORCE — THE WORLD'S LARGEST GAMING PLATFORM

At $100 billion, computer gaming is the world's largest entertainment industry. And with 200 million gamers, NVIDIA GeForce is its largest platform. GeForce® GTX GPUs and the GeForce Experience™ application transform everyday PCs into powerful gaming machines. And we continuously innovate across the platform. One recent example, Max-Q design, sets a new standard for the thinnest, fastest, and quietest gaming laptops in the market.

SHIELD — THE MOST ADVANCED STREAMER BRINGS AI TO THE HOME

NVIDIA SHIELD™ boasts 1,000 games and offers the largest, most open catalog of media in stunning 4K. It's also a smart home hub with support for Google Assistant hands-free as well as SmartThings, which connects to hundreds of devices. With SPOT™, an AI mic accessory that plugs into an outlet, intelligent control can be extended throughout the house.

NVIDIA IS REVOLUTIONIZING COMPUTING

In 2006, the creation of our CUDA programming model and Tesla® GPU platform opened up the parallel-processing capabilities of the GPU to general-purpose computing.

A powerful new approach to computing was born.

POWERING THE WORLD'S FASTEST SUPERCOMPUTERS

Moore's law is coming to an end, but exponential advances in compute speed and capacity are just beginning. Today, GPU computing is the most pervasive, accessible, energy-efficient path forward for HPC and data centers, and powers the fastest supercomputers in the U.S. and Europe.
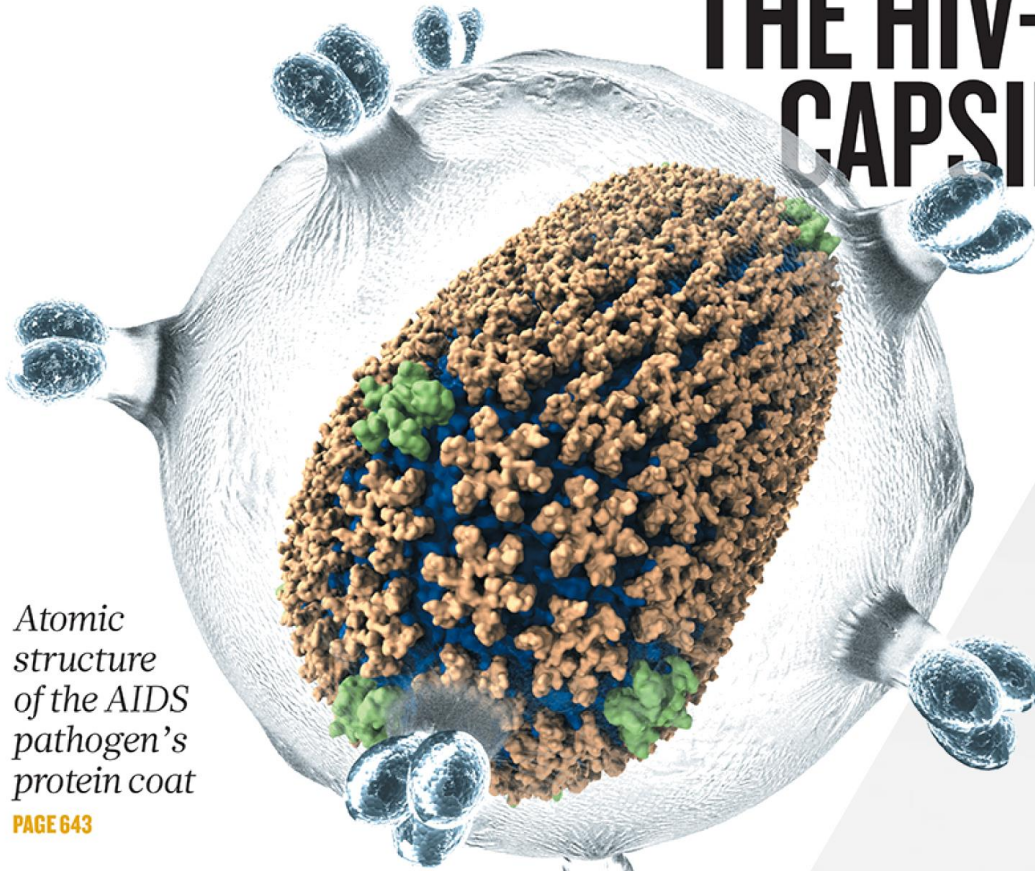
nature

THE HIV-1 CAPSID

Atomic structure of the AIDS pathogen's protein coat

ACCELERATING LIFE-CHANGING SCIENTIFIC DISCOVERIES

The GPU computing platform is essential to those who take on the challenges that matter most. Using a GPU-powered supercomputer, University of Illinois scientists achieved a breakthrough in HIV research by performing the first all-atom simulation of the capsid. GPUs also drive the GE Revolution CT scanner, which can produce high-quality imagery while reducing patient radiation exposure by up to 82%.
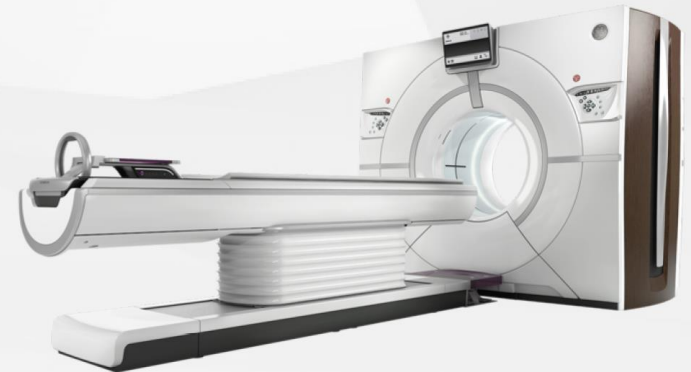
NVIDIA IGNITES
THE AI BIG BANG

Artificial intelligence is the use of computers to simulate human intelligence.

AI amplifies our cognitive abilities — letting us solve problems where the complexity is too great, the information is incomplete, or the details are too subtle and require expert training.

Learning from data — a computer's version of life experience — is how AI evolves. GPU computing powers the computation required for deep neural networks to learn to recognize patterns from massive amounts of data.
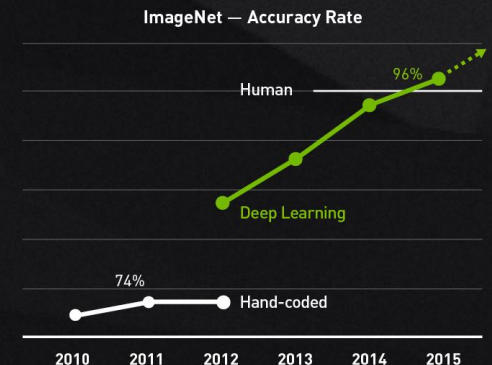
This new computing model sparked the AI era.

AI ACHIEVES "SUPERHUMAN" RESULTS

The big bang of modern AI set off a string of "superhuman" achievements. In 2015, Google and Microsoft both beat the best human score in the ImageNet challenge. DeepMind's AlphaGo recorded its historic win over Go champion Lee Sedol in 2016 and, more recently, beat the best player in the world, Ke Jie. Breakthroughs in AI happen almost every day.

ImageNet — Accuracy Rate

96%

Human

Deep Learning

74%

Hand-coded

2010 2011 2012 2013 2014 2015

THE ERA OF AI

With deep learning, we can teach AI to do almost anything. New internet services, like Google Assistant, have learned speech from sound and provide a more natural way to access information. Self-driving cars use deep learning to recognize the space the car inhabits, the lanes in which it drives, and the objects it must avoid. In healthcare, neural networks trained with millions of medical images can find clues in MRIs that until now could only be found through invasive biopsies. These are just a few examples. AI will spur a wave of social progress unmatched since the industrial revolution.
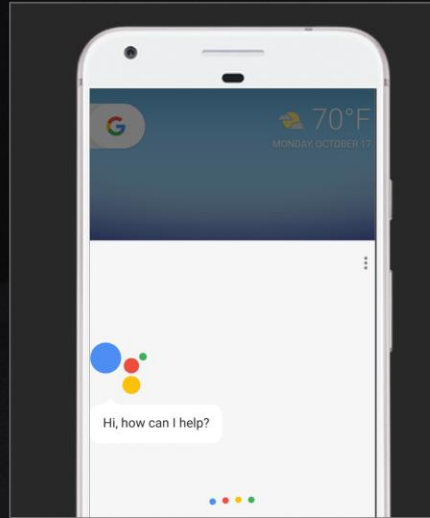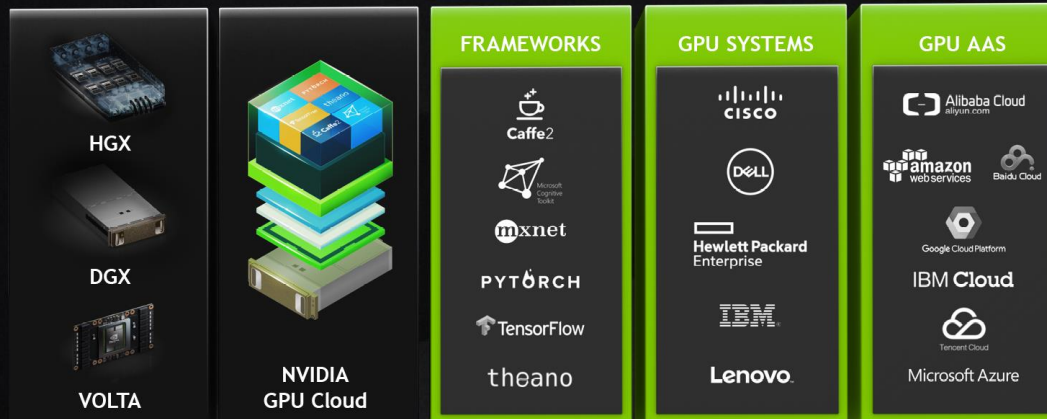

"Ok Google"
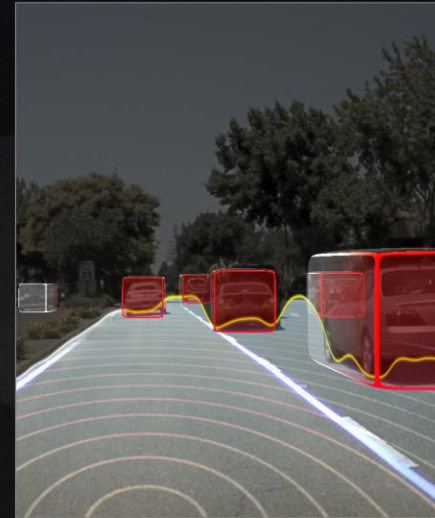
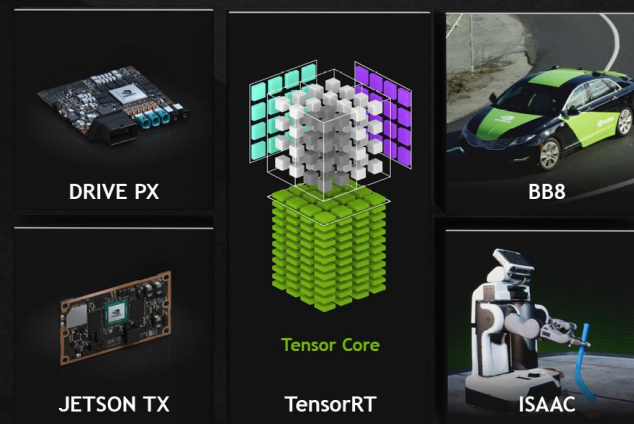Self-Driving Cars
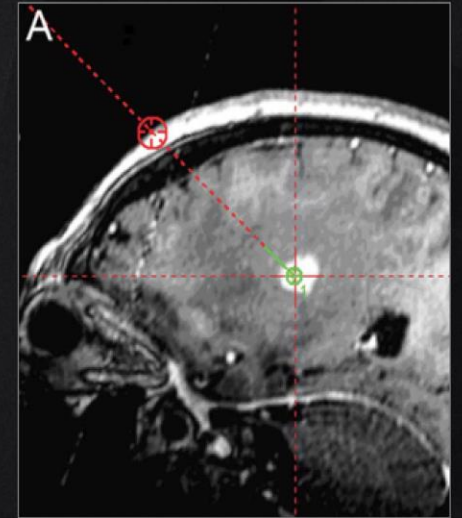
Medical Breakthroughs

POWERING
THE AI REVOLUTION

GPU computing is the most productive and pervasive platform for deep learning and AI. It begins with the most advanced GPUs and the systems and software we build on top of them. We integrate and optimize every deep learning framework. We work with the major systems companies and every major cloud service provider to make GPUs available in data centers and in the cloud. And we create computers and software to bring AI to edge devices, such as self-driving cars and autonomous robots.

AI in the Data Center and Cloud

		FRAMEWORKS	GPU SYSTEMS	GPU AAS
HGX		Caffe2	CISCO	Alibaba Cloud aliyun.com
DGX		Microsoft Cognitive Toolkit	DELL	amazon web services / Baidu Cloud
VOLTA	NVIDIA GPU Cloud	mxnet	Hewlett Packard Enterprise	Google Cloud Platform
		PYTORCH	IBM	IBM Cloud
		TensorFlow	Lenovo	Tencent Cloud
		theano		Microsoft Azure

AI at the Edge

DRIVE PX		BB8
JETSON TX	Tensor Core / TensorRT	ISAAC

VOLTA — A GIANT LEAP FOR AI AND HPC

Volta, the world's most powerful GPU computing architecture, was built to drive the next wave of AI and HPC. Tesla V100, based on Volta, is the world's first 120 TeraFLOPS processor and brings extraordinary speed and scalability for AI inferencing and training, as well as for accelerating HPC and graphics workloads. Volta is the biggest chip on the planet — boasting 21 billion transistors and able to deliver the equivalent performance of 250 CPUs for deep learning.
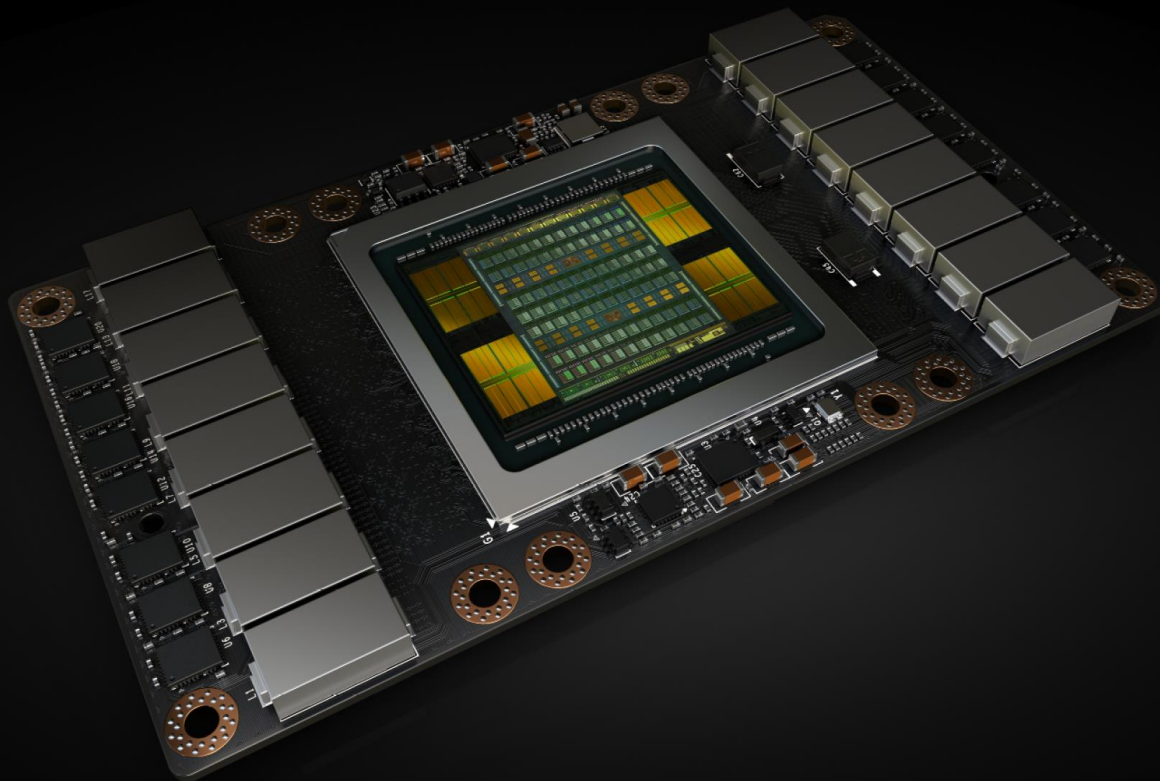
SUPERCOMPUTERS
PURPOSE-BUILT FOR AI

While deep learning holds enormous promise, it requires a massive amount of computing power. To arm data scientists in every organization, we created a lineup of AI supercomputers. NVIDIA DGX-1 with Volta, the world's first 1 PetaFLOPS computer, delivers the performance of up to 400 servers in a single box; while DGX Station packs 480 TeraFLOPS of computing power into a whisper-quiet workstation.

NVIDIA GPU CLOUD —
ONE PLATFORM, RUN
EVERYWHERE

The NVIDIA GPU Cloud gives AI developers
access to our comprehensive deep learning
software stack wherever they want it — on
PCs, in the data center, or via the cloud.
The containerized stack includes the latest
deep learning frameworks, libraries, OS, and
drivers. NVIDIA GPU Cloud makes it easier
for developers to do deep learning training,
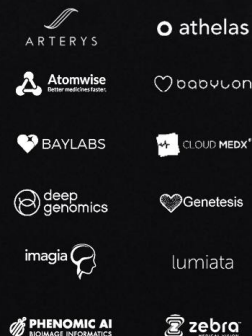experimentation, and deployment.

AI IS REVOLUTIONIZING EVERY INDUSTRY

In addition to our AI technologies, we advance fundamental research, foster universities and startups, and bring our full capabilities to industries where we can have the greatest impact.

NVIDIA INCEPTION — 1,300 DEEP LEARNING STARTUPS

The NVIDIA Inception program nurtures more than 1,300 startups that are revolutionizing industries with advances in AI and data science. The program helps startups during critical stages of product development, prototyping, and deployment. At GTC, our annual developer conference, we awarded six AI startups with a total of $1.5 million to accelerate their work.

HEALTHCARE

ARTERYS · athelas · Atomwise · babylon · BAYLABS · CLOUD MEDx · deep genomics · Genetesis · imagia · lumiata · PHENOMIC AI BIOIMAGE INFORMATICS · zebra MEDICAL VISION

RETAIL, ETAIL

focal Systems · GOFIND · GROKSTYLE · huew · MARIANA · mashgin · netra · ProductAI · shoppr

IOT AND MANFACTURING

3DSignals · ABEJA · CLOUDBRAIN · IFM · KONUX · Preferred Networks · PREDII · RS Reliability Solutions. The Art of Prediction · VERDIGRIS · VIMOC Technologies · IXG

AUTONOMOUS MACHINES

AIMOTIVE · ARGO AI · BLUE RIVER TECHNOLOGY · dispatch · drive.ai · marble · nuTonomy · pilot.ai · OPTIMUS RIDE · tuSimple · ZENUITY · ZOOX

FINANCIAL

Alpaca · CAPE ANALYTICS · Deep Learni.ng · Orbital Insight · Quantenstein

SECURITY, IVA

airXsys · ENTROPIX · herta · KUNA · UMBO CV

CYBER

CYLANCE · deepinstinct · GreatHorn · UNIFYID

AEC

CUBICASA · HOVER · PRENAV · SMARTVID.IO

PLATFORMS AND APIs

7INVENSUN · clarifai · deep gram · Gridspace · KITT.AI · VisionLabs visual recognition company

DATA MANAGEMENT

CrowdAI · DATALOGUE · understand.ai

DEVELOPMENT PLATFORMS

bonsai · FLOYD · H₂O.ai · skymind · SIGOPT

BUSINESS INTELLIGENCE AND VISUALIZATION

Blazing · graphistry · kinetica · MAPD · SQREAM TECHNOLOGIES

THE BRAIN OF AI CARS

Autonomous vehicles will modernize the $10 trillion transportation industry — making our roads safer and our cities more efficient. NVIDIA DRIVE™ PX is a scalable AI car platform that spans the entire range of autonomous driving.

Toyota recently joined some 225 companies around the world that have adopted the NVIDIA DRIVE PX platform for autonomous vehicles. They range from car companies and suppliers, to startups and research organizations.
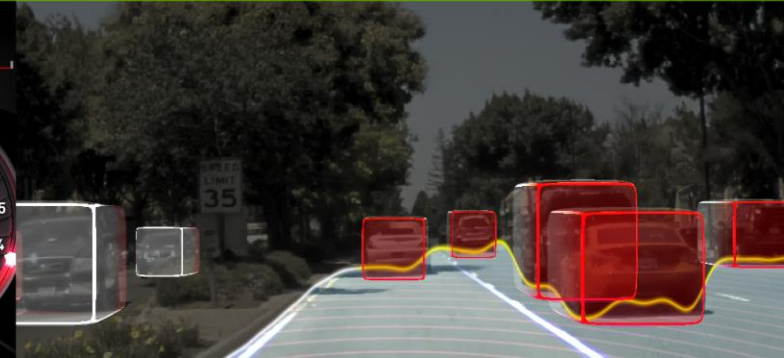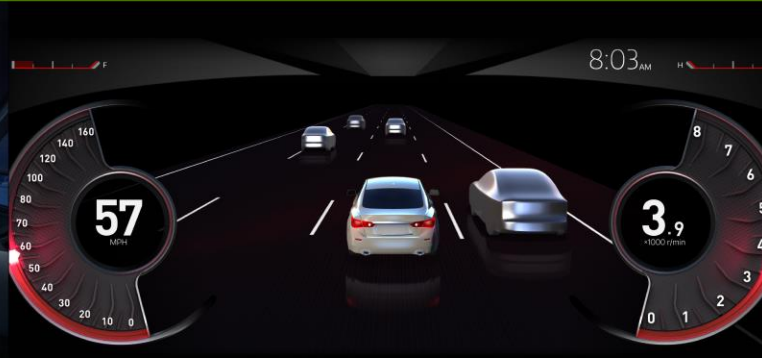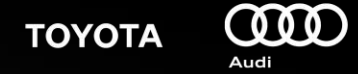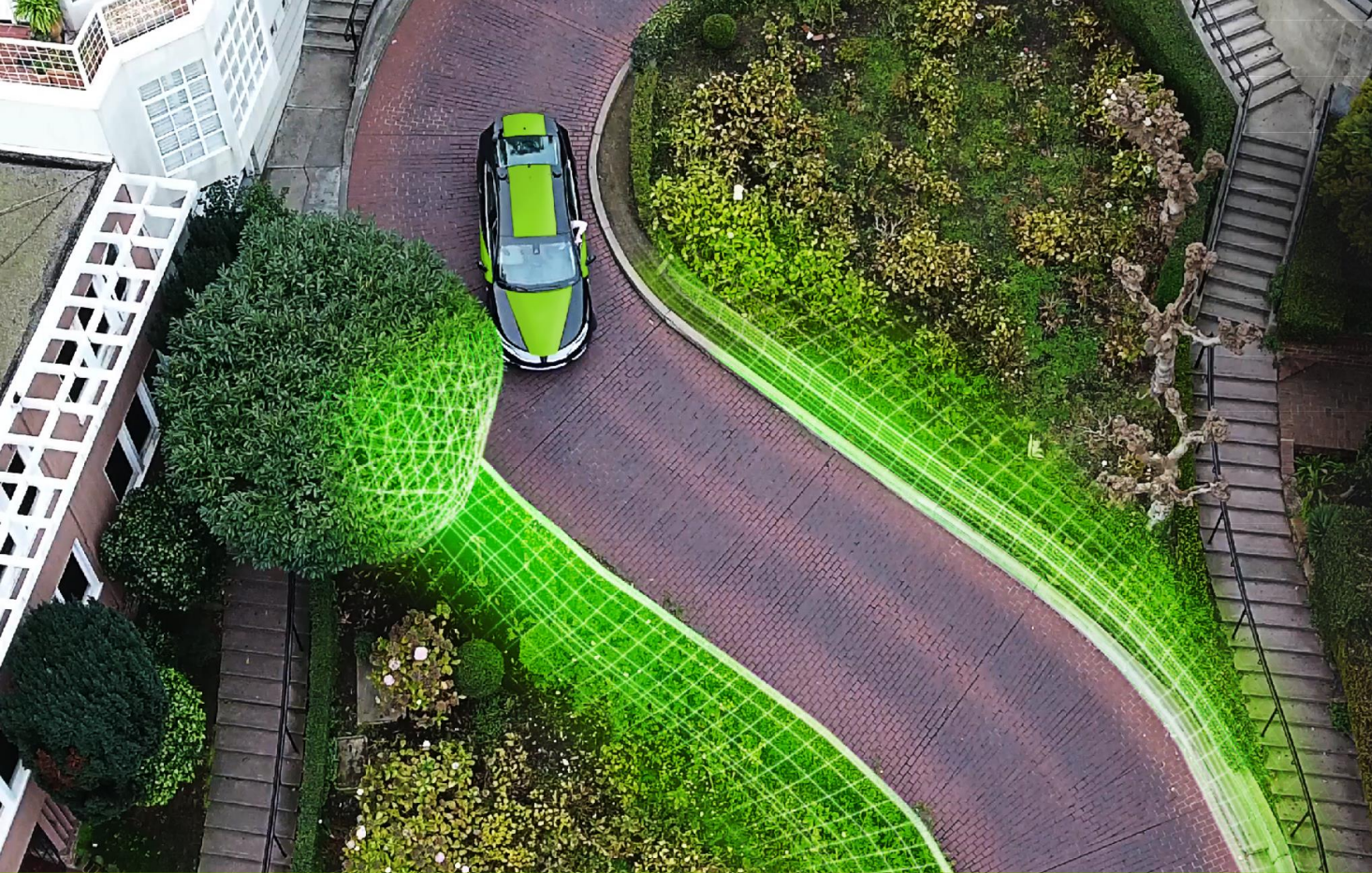
NVIDIA BB8 AI CAR — LEARNING BY EXAMPLE

Driving is a learned behavior that people do as second nature. Yet one that is impossible to program a computer to perform. Using all of the AI capabilities of NVIDIA DRIVE PX 2, our research AI car, BB8, watches humans drive, and has learned to drive in all kinds of conditions — on highways and dirt roads, through obstacle courses, at night, and in the rain. Processing data from multiple cameras, BB8 can even look both ways before safely crossing a busy road on its own.

Left Camera

Center Camera

Right Camera

THE BRAIN OF INTELLIGENT MACHINES & IoT

Deep learning and affordable sensors have created the conditions for a Cambrian explosion of autonomous machines — IoT with AI. NVIDIA Jetson™ TX2, an embedded AI supercomputer, delivers 1 TeraFLOPS of performance in a credit card-sized module. Such power will enable a new wave of automation in manufacturing, drones that can inspect hazardous places, and robots that can deliver the millions of packages shipped every day.
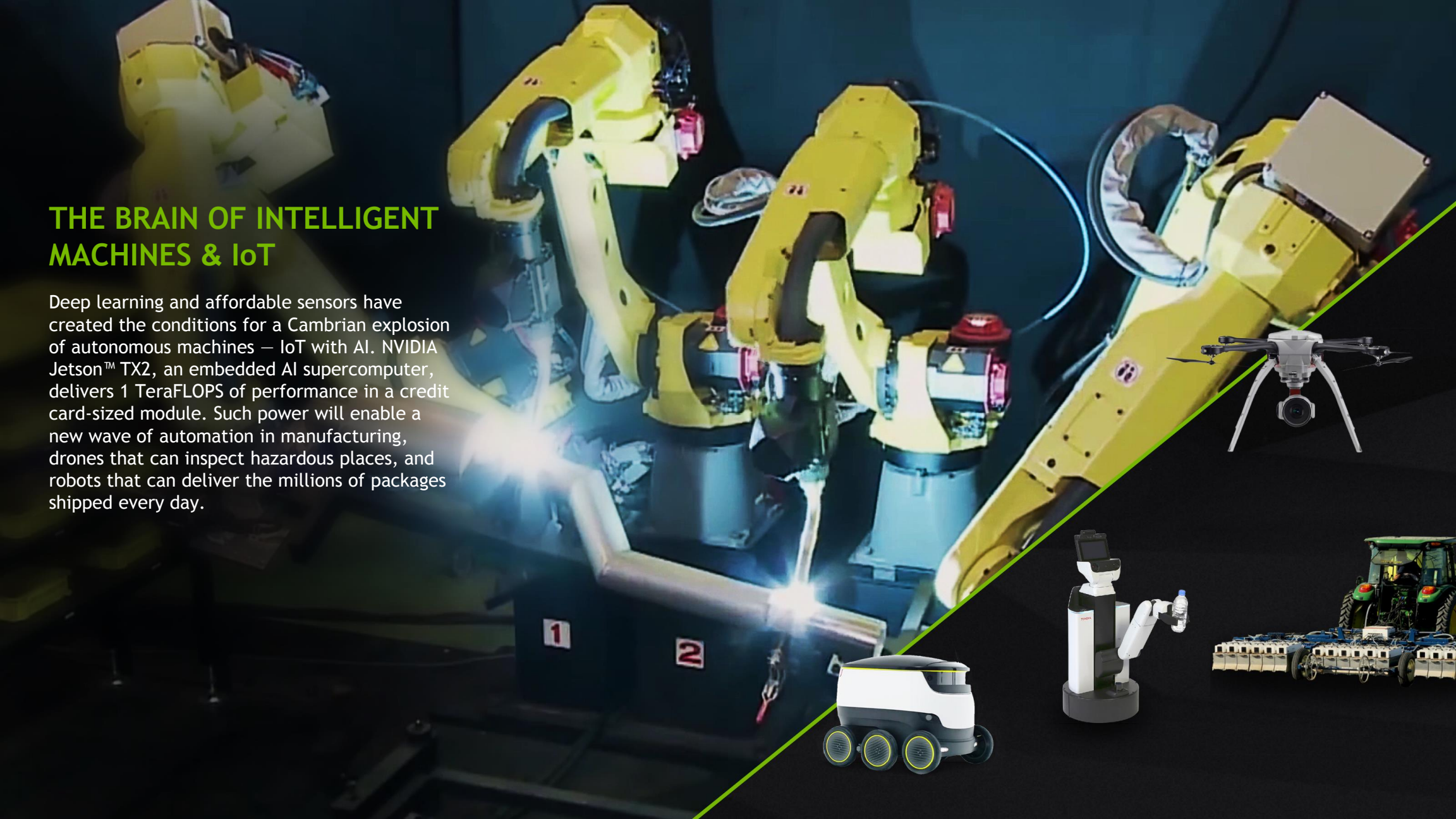
ISAAC — ACCELERATED LEARNING FOR A WORLD OF INTELLIGENT MACHINES

The Isaac robot simulator, an AI-based software platform, lets developers train robots in highly realistic, physics-based virtual environments and then transfer that knowledge to real-world units. Developers can set up extensive test scenarios using deep learning training, and then simulate them in minutes instead of months.

THE BRAIN OF THE AI CITY

There will be 1 billion cameras in the world by 2020. AI will power intelligent video analytics that can turn this massive amount of data into safer, more efficient cities. With Jetson TX2 at the edge and Tesla GPUs in the cloud, NVIDIA Metropolis is an end-to-end platform that provides a foundation for the AI City. More than two dozen partners around the world have adopted Metropolis for a variety of applications, from public safety to traffic management to city services.
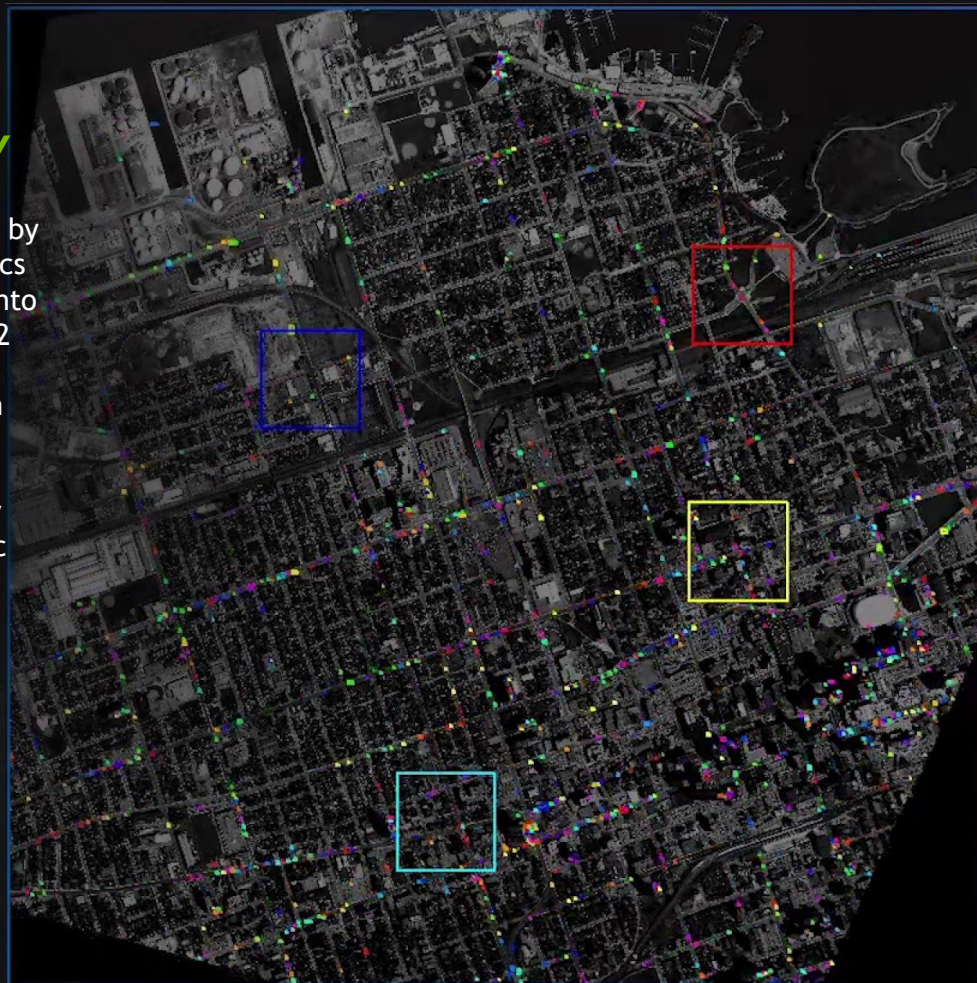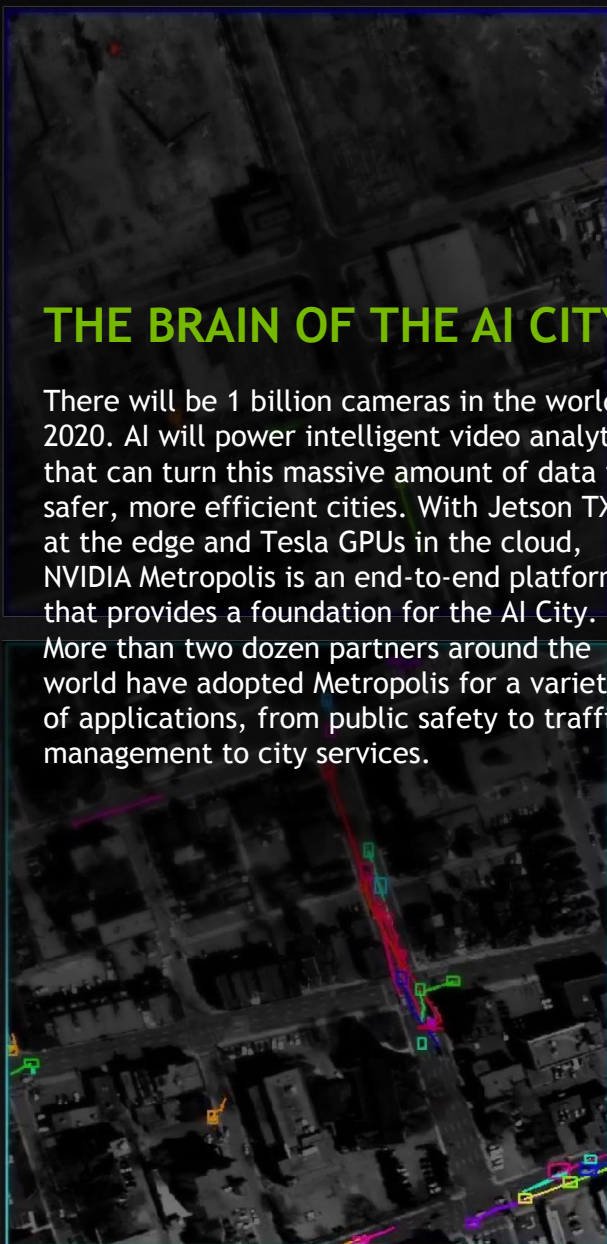
THE BRAIN OF AI HEALTHCARE

AI is transforming the spectrum of healthcare, from detection to diagnosis to treatment. GE Healthcare has reinvented the echocardiogram machine by embedding GPU-powered AI in its Vivid E95 system. Mayo Clinic used GPU-powered deep learning to discover that genomic data can be found in MRIs, hidden from traditional analysis methods.

NVIDIA is teaming up with the National Cancer Institute, the U.S. Department of Energy, and several national labs on the "Cancer Moonshot" to deliver a decade of advances in cancer prevention, diagnosis, and treatment in just five years.
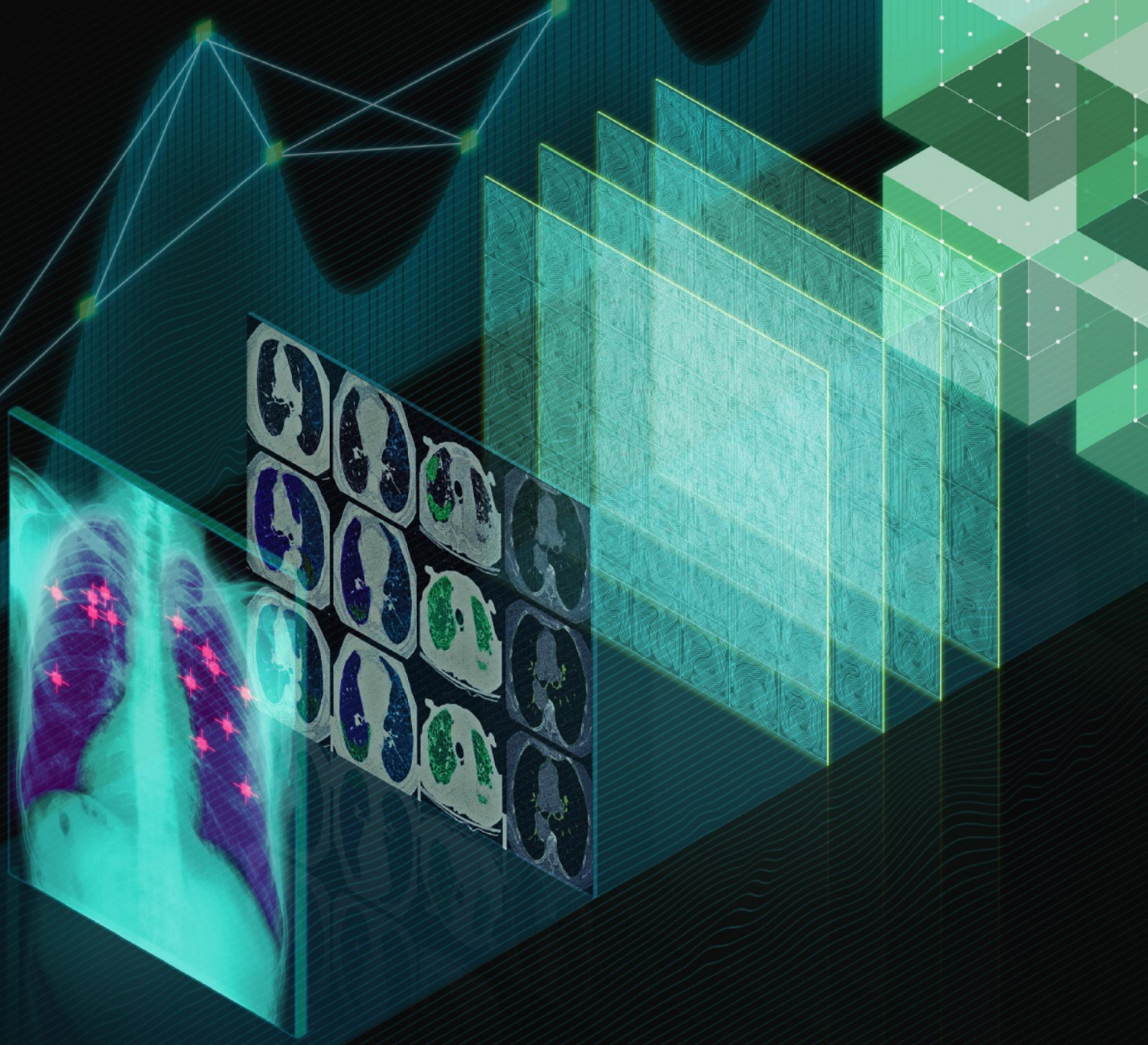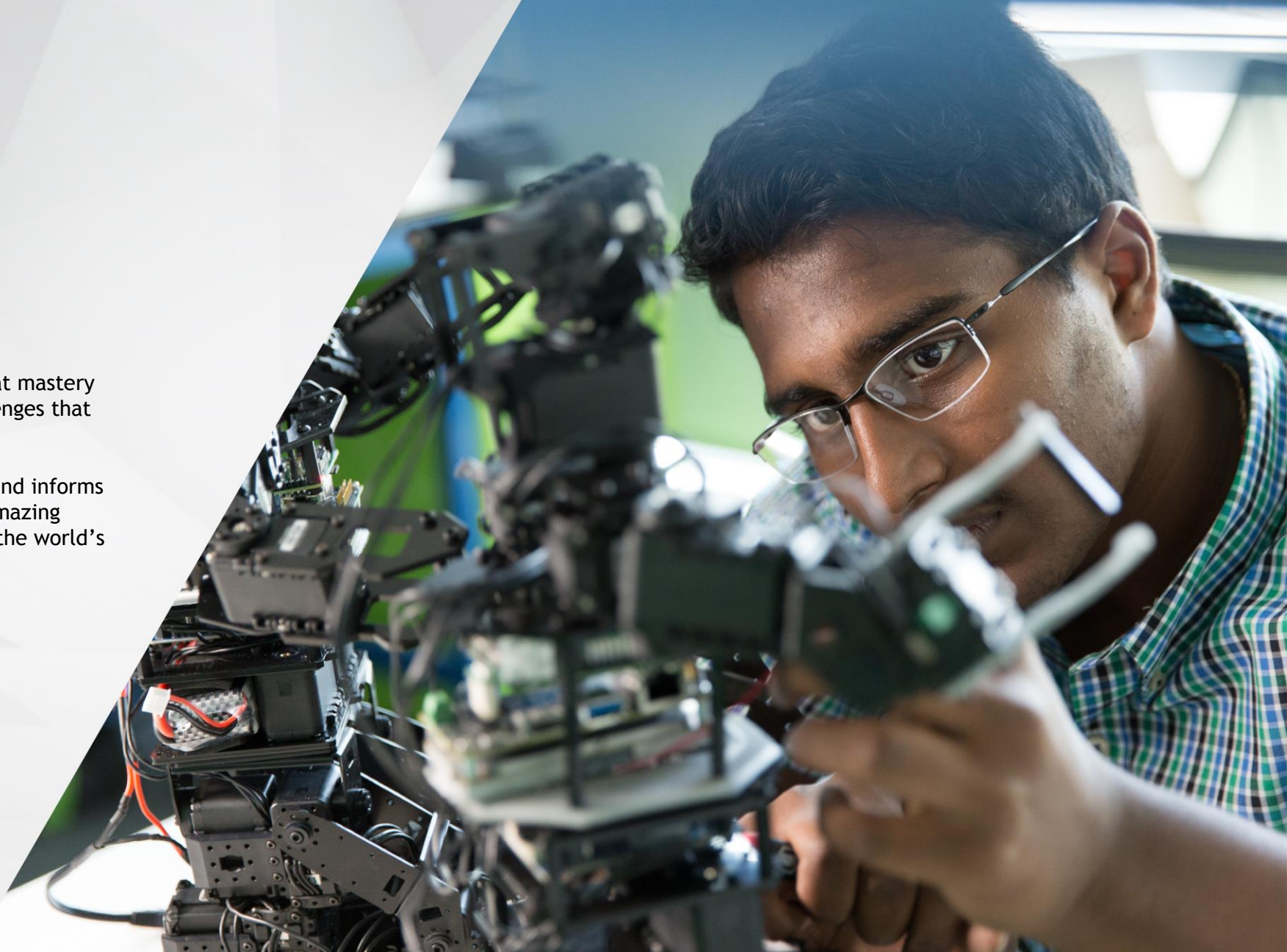
OUR CULTURE
A PASSION FOR CRAFTMANSHIP

NVIDIA is united by a core belief that mastery of our craft lets us tackle the challenges that matter to the world.

This attitude pervades our culture and informs everything we do, from designing amazing products to striving to build one of the world's great companies.

OUR CULTURE

INSPIRED TO GIVE TO OUR COMMUNITIES

NVIDIA's people share a strong sense of corporate responsibility. Our philanthropic giving this year exceeded $5 million.

Project Inspire, which brings our people together every year to transform local communities, continues to gain momentum. This year alone, employees volunteered more than 17,000 hours and supported education programs that benefited nearly 90,000 children.

OUR CULTURE
A LEARNING MACHINE

NVIDIA has continuously reinvented itself over two decades.

Our 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 computing ignited the era of AI.
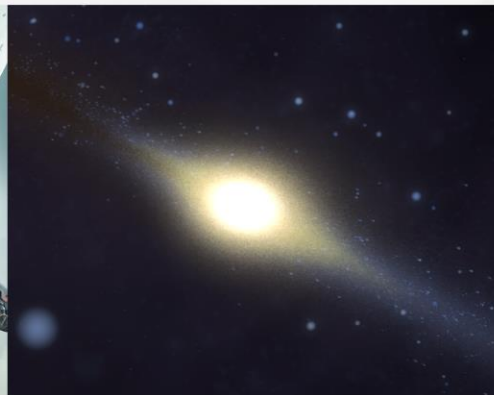
NVIDIA is a "learning machine" that constantly evolves by adapting to new opportunities that are hard to solve, that only we can tackle, and that matter to the world.
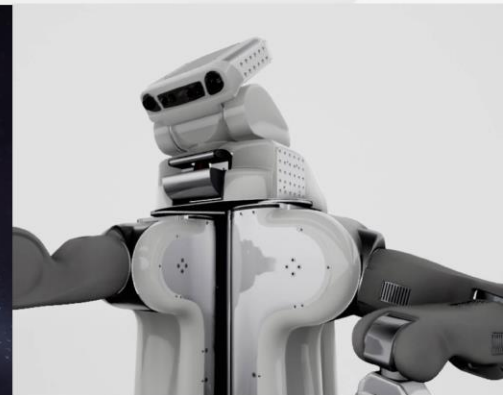
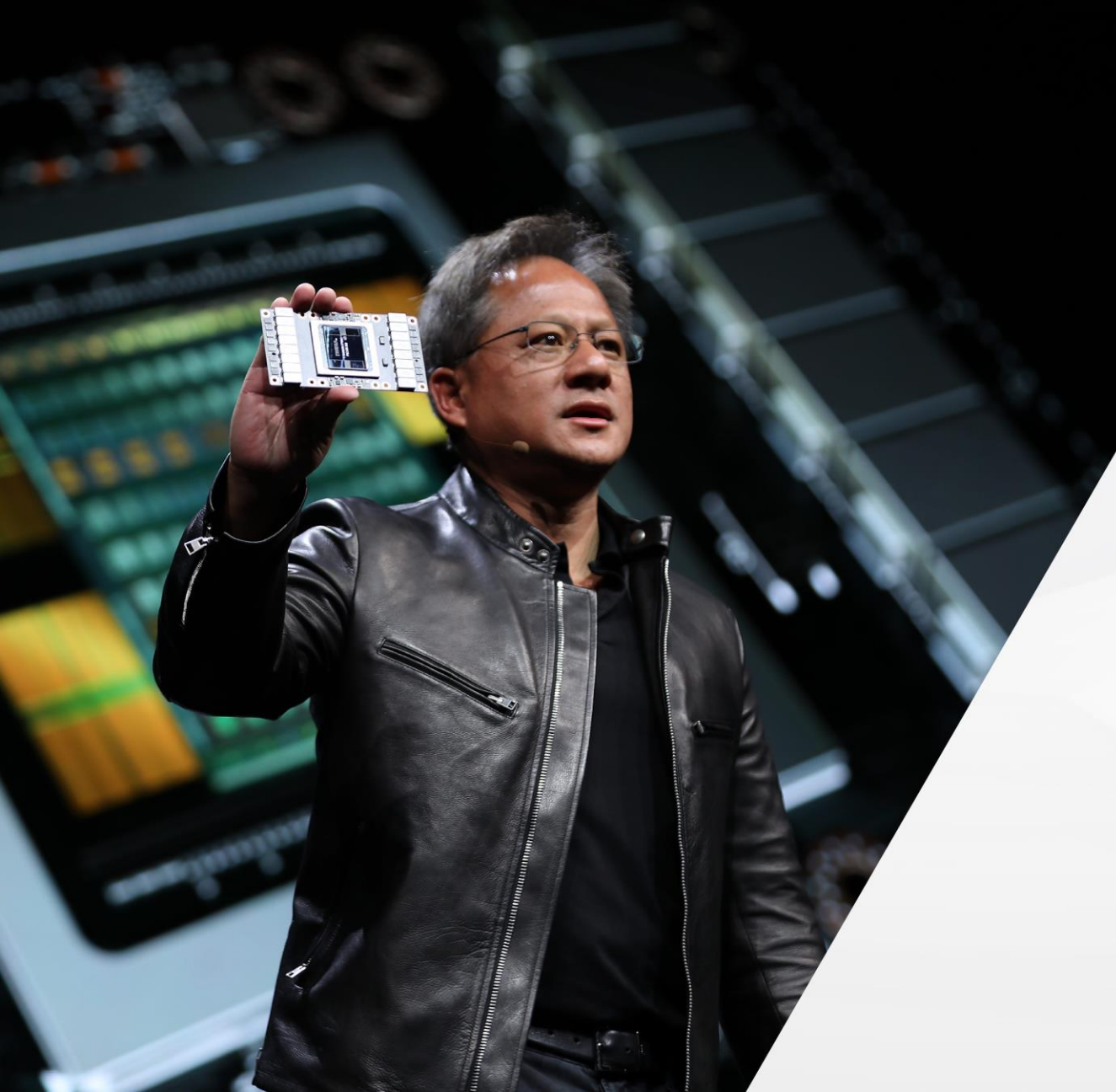**PC GRAPHICS**　　　　**GPU COMPUTING**　　　　**AI**

1996　　　　　　**2006**　　　　　　**2016**

NVIDIA

> Founded in 1993

> Jensen Huang, Founder & CEO

> 11,000 employees

> $6.9B in FY17

"World's Best Performing CEOs"
— Harvard Business Review

"World's Most Admired Companies"
— Fortune

"World's Best CEOs"
— Barron's

"Most Innovative Companies"
— Fast Company

"Employees' Choice: Highest Rated CEOs"
— Glassdoor

"50 Smartest Companies"
— MIT Tech Review